RELATIVISTIC QUANTUM MECHANICS

Lectures given at M.I.T. in 1953/4

by

Victor F. WEISSKOFF

(Notes prepared by A. Bincer)
Propriété littéraire et scientifique réservée pour tous les pays du monde. Ce document ne peut être reproduit ou traduit en tout ou en partie sans l'autorisation écrite du Directeur général du CERN, titulaire du droit d'auteur. Dans les cas appropriés, et s'il s'agit d'utiliser le document à des fins non commerciales, cette autorisation sera volontiers accordée.

Le CERN ne revendique pas la propriété des inventions brevetables et dessins ou modèles susceptibles de dépôt qui pourraient être décrits dans le présent document; ceux-ci peuvent être librement utilisés par les instituts de recherche, les industriels et autres intéressés. Cependant, le CERN se réserve le droit de s'opposer à toute revendication qu'un usager pourrait faire de la propriété scientifique ou industrielle de toute invention et tout dessin ou modèle décrits dans le présent document.

Literary and scientific copyrights reserved in all countries of the world. This report, or any part of it, may not be reprinted or translated without written permission of the copyright holder, the Director-General of CERN. However, permission will be freely granted for appropriate non-commercial use.

If any patentable invention or registrable design is described in the report, CERN makes no claim to property rights in it but offers it for the free use of research institutions, manufacturers and others. CERN, however, may oppose any attempt by a user to claim any proprietary or patent rights in such inventions or designs as may be described in the present document.
PREFACE

These notes are the first part of a course in advanced quantum mechanics held at M.I.T. in 1953-54. They contain an introduction into some fundamental concepts, all connected with the Dirac equation of the spinning electron. Special emphasis is laid upon the concepts of angular momentum, spinors and upon that famous two-spinor wave function which is the Dirac electron.

The presentation is far from complete. In particular, the treatment of charge-conjugation and the treatment of the scattering problem are rudimentary to the extreme. These shortcomings are the reason why the author has abstained from the publication of these notes. He relented, not because he thinks that the treatment is not so bad, but because he realized that he will never have the time to do it better.

Many thanks are due to Dr. H. Rollnik for correcting mistakes and improving many details of the manuscript.

Victor F. Weisskopf
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>ANGULAR MOMENTUM</td>
<td></td>
</tr>
<tr>
<td>Commutation Relations</td>
<td>1</td>
</tr>
<tr>
<td>Spherical Harmonics</td>
<td>15</td>
</tr>
<tr>
<td>Addition of Angular Momenta</td>
<td>18</td>
</tr>
<tr>
<td>Vector Fields</td>
<td>26</td>
</tr>
<tr>
<td>Spinors</td>
<td>32</td>
</tr>
<tr>
<td>II</td>
<td></td>
</tr>
<tr>
<td>DIRAC EQUATION FOR THE ELECTRON</td>
<td></td>
</tr>
<tr>
<td>Analogy with Maxwell Equations</td>
<td>43</td>
</tr>
<tr>
<td>Maxwell's Equations in Dirac's Formalism</td>
<td>43</td>
</tr>
<tr>
<td>The Physics in Dirac's Equations</td>
<td>51</td>
</tr>
<tr>
<td>Non-relativistic Approximation</td>
<td>54</td>
</tr>
<tr>
<td>Solutions of the Dirac Equation</td>
<td>62</td>
</tr>
<tr>
<td>Velocity of the Dirac Electron</td>
<td>69</td>
</tr>
<tr>
<td>III</td>
<td></td>
</tr>
<tr>
<td>THE DIRAC EQUATION UNDER THE LORENTZ TRANSFORMATION</td>
<td></td>
</tr>
<tr>
<td>Comparison of the Dirac Spinor and the Maxwell Vectors under the Lorentz Transformation</td>
<td>75</td>
</tr>
<tr>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>DIRAC'S EQUATION IN THE PRESENCE OF AN ELECTROMAGNETIC FIELD</td>
<td></td>
</tr>
<tr>
<td>The Charge Conjugate Solution</td>
<td>85</td>
</tr>
<tr>
<td>V</td>
<td></td>
</tr>
<tr>
<td>SOLUTION OF DIRAC'S EQUATION IN AN EXTERNAL FIELD</td>
<td></td>
</tr>
<tr>
<td>First Dirac Corrections to the Schrödinger Equation</td>
<td>87</td>
</tr>
<tr>
<td>Presence of Magnetic Field (Zeeman effect)</td>
<td>89</td>
</tr>
<tr>
<td>Solution of Dirac's Equation in a Central Field</td>
<td>95</td>
</tr>
<tr>
<td>Ground State of the Hydrogen Atom</td>
<td>101</td>
</tr>
<tr>
<td>VI</td>
<td></td>
</tr>
<tr>
<td>THE SCATTERING PROBLEM. BORN APPROXIMATION</td>
<td></td>
</tr>
<tr>
<td>Born Approximation in Schrödinger's formalism</td>
<td>103</td>
</tr>
<tr>
<td>Born Approximation in Dirac's Formalism</td>
<td>117</td>
</tr>
<tr>
<td>Polarization in the Scattered Beam</td>
<td>123</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>Detailed Table of Contents</td>
<td>146</td>
</tr>
</tbody>
</table>
Chapter I: Angular Momentum

We start with the definitions of several symbols and concepts that we shall be dealing with. We shall denote state functions by symbols such as \( \psi \); we shall be able to build up a quite extensive formalism without specifying what \( \psi \) is a function of.

If \( \psi = \psi \), where \( \psi \) is a state function and \( \psi \) is another state function, we call \( \psi \) an operator. In addition, if \( \psi(a\psi_1 + b\psi_2) = a\psi_1 + b\psi_2 \), where \( \psi_1 = \) state functions, \( a, b = \) scalars, then \( \psi \) is a linear operator.

The inner (or scalar) product of two state functions \( \psi \) and \( \psi \) will be denoted by \( (\psi, \psi) \). Some of the properties of the inner product:

\[
(\psi, \psi) = \text{number} \\
(\psi, a\psi) = a(\psi, \psi) \\
(a\psi, \psi) = a^*(\psi, \psi) \\
(\psi, \psi) = \text{square of magnitude of } \psi
\]

If \( (\psi, \psi) = 1 \), \( \psi \) is said to be normalized. If \( \psi \) is normalized and \( (\psi, 0\psi) = 1 \), \( \psi \) is said to be a unitary operator. If \( (\psi, P\psi) = (P\psi, \psi) \) then \( P \) is said to be an hermitian operator. If \( (\psi, a\psi) = (a^*\psi, \psi) \), then \( Q^* \) is called the hermitian conjugate or complex adjoint of \( Q \). If \( Q^* = -Q \), then \( Q \) is said to be anti-hermitian. (For example, if \( P \) is hermitian, then \( iP \) is anti-hermitian.)
If \( \mathcal{O} \mathcal{Q} = \sum_k O_{ik} \mathcal{Q}_k \) and the \( \mathcal{Q}_k \) 's form a complete orthonormal set, then \( O_{ik} \) 's are the matrix elements of the operator \( \mathcal{O} \).

\[ [A, B] = AB - BA \] is called the commutator of \( A \) and \( B \). Next we introduce the concept of momentum. In classical physics for a given coordinate \( q \) we can find its canonically conjugate momentum \( P_q \) given by \( P_q = -\frac{2H}{\partial q} \) where \( H \) is the Hamiltonian of the system. In quantum theory we say that the effect of \( P_q \) operating on a state function is the same as that obtained by multiplying by \( \frac{\hbar}{i} \) the change in the state function when the latter is shifted by an infinitesimal displacement in the direction of \(-q\), i.e.

\[
P_q = \frac{\hbar}{i} \frac{\partial}{\partial q} \mathcal{Q}.
\]

For example,

\[
P_x(x) = \frac{\hbar}{i} \frac{\partial}{\partial x} \mathcal{Q}(x) = \frac{\hbar}{i} \frac{\mathcal{Q}(x+dx)-\mathcal{Q}(x)}{dx} = \frac{\hbar}{i} \text{ state shifted by } ( -dx ) - \text{ original state } \frac{dx}{dx}
\]

Before we proceed to the discussion of angular momentum, it will be useful to define the rotation operator \( R \). (Note: We shall always consider the physical system as being rotated and not the coordinate frame.) A rotation through an angle \( \Theta \) around an axis \( \hat{n} \) will be denoted by \( R(\Theta, \hat{n}) \), and the rotated state function \( \mathcal{Q}' \) is then given by

\[
\mathcal{Q}' = R(\Theta, \hat{n}) \mathcal{Q}.
\]
Rotations through finite angles do not form a commutative group under addition as one can easily be convinced by the example below:

If, however, the rotation is through an infinitesimal angle $d\theta$ and we neglect terms involving $(d\theta)^2$, we find that such rotations commute. Under such an infinitesimal rotation, the state function is the same as before except for an infinitesimal change which is proportional to $d\theta$, i.e.:

$$R(\theta, \hat{n}) = 1 + r(\hat{n}) d\theta.$$ 

An infinitesimal rotation can be split into rotations around the $x$, $y$, $z$ axes:

$$R_x = 1 + r_x d\theta$$
$$R_y = 1 + r_y d\theta$$
$$R_z = 1 + r_z d\theta$$
with the obvious relation $\mathbf{r}(\hat{A}) = \hat{n}_x r_x + \hat{n}_y r_y + \hat{n}_z r_z$. The $r$'s satisfy very important commutation relations:

\[
\begin{align*}
[r_x r_y - r_y r_x] &= r_z \\
[r_y r_z - r_z r_y] &= r_x \\
[r_z r_x - r_x r_z] &= r_y
\end{align*}
\]  
(1)

These relations are based upon the fundamental properties of rotations and can be seen as follows: Consider the unit sphere around the center of a coordinate system and watch the point $P$: $x=1, y=z=0$, on this sphere. We now rotate the sphere around the $y$-axis by $d\varphi$ and then around the $x$-axis (Note: The sphere only is rotated, the coordinate axes remain fixed!)

![Diagram of rotation](image)

The two displacements $r_x d\varphi$, $r_y d\varphi$ of $P$ are seen in Fig. la. Fig. lb shows the result of the operations $r_y d\varphi r_x d\varphi$, where we rotate the sphere first around $x$ (no displacement of the point $P$) and then around $y$. The difference is the little arrow in Fig. la, which represents a rotation around the $z$-axis by an angle $d\varphi^2$. Hence, $r_x d\varphi r_y d\varphi - r_y d\varphi r_x d\varphi = r_z d\varphi^2$ from which follows (1). The fact that $[r_i, r_k] \neq 0$ doesn't violate the statement made before that
infinitesimal rotations commute. The amount of rotation is not $r$ but $r\,d\theta$, hence $(r_x\,d\theta)(r_y\,d\theta)-(r_y\,d\theta)(r_x\,d\theta)=r_z(d\theta)^2$. Thus we see that if we perform the rotations in the two different orders, the difference is of second order in $d\theta$ and, therefore, can be neglected.

Since $R$ represents only a rotation, it must be unitary:

\[
(R\varphi, R\varphi) = ((1 + \hat{r}d\theta)\varphi, (1 + \hat{r}d\theta)\varphi) = (\varphi, \varphi) + (\hat{r}d\theta \varphi, \varphi) + (\varphi, \hat{r}d\theta \varphi) + (\hat{r}d\theta \varphi, \hat{r}d\theta \varphi) = 1
\]

Since $\varphi$ is normalized $(\varphi, \varphi) = 1$, and since $(\hat{r}d\theta \varphi, \hat{r}d\theta \varphi)$ is of second order in $d\theta$, we neglect it; hence we get

\[
(\hat{r}\varphi, \varphi) = -(\varphi, \hat{r}\varphi)
\]

and therefore $r$ is anti-hermitian.

We are now ready to define the angular momentum operator $\mathbf{J}$, which then according to the rule stated on page 2, must be

\[
\mathbf{J} = \frac{h}{i} \mathbf{\hat{r}}
\]

Since $\mathbf{\hat{r}}$ is anti-hermitian, it follows that $\mathbf{J}$ is hermitian. Also the commutation rules follow:

\[
\begin{align*}
J_x J_y - J_y J_x &= i\hbar J_z \\
J_y J_z - J_z J_y &= i\hbar J_x \\
J_z J_x - J_x J_z &= i\hbar J_y 
\end{align*}
\]
Using only above commutation relations and nothing else, the following properties of $\hat{J}$ can be proved:

a) $\left[ J^2, J_i \right] = 0$

b) $J_+ \psi_j = \hbar \ A_j \psi_{j+1}$ ; $J_- \psi_j = \hbar \ B_j \psi_{j-1}$

c) $\hat{J}^2 \psi_j = \hbar^2 (j+1) \psi_j$ ; $J_z \psi_j = \hbar \ m \ \psi_j$

with $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ and $m = -j, -j+1, \ldots, j-1, j$

where

$J_+ \equiv J_x + i J_y$ \hspace{1cm} $J_- \equiv J_x - i J_y$

and

$A_{jm} = \sqrt{j(j+1) - m(m+1)}$, $B_{jm} = \sqrt{j(j+1) - m(m-1)}$

Before proving (a), (b) and (c) we first state the following laws of composition for commutators which follow directly from their definition:

$[A, B] = -[B, A]$

$[A+B, C] = [A, C] + [B, C]$


$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0$ \hspace{1cm} (Jacobi's identity)

Proof of (a):

$\left[ \hat{J}^2, J_i \right] = \left[ J^2, J_i \right] + \left[ J_+^2, J_i \right] + \left[ J_-^2, J_i \right] = 0$

$= J_+ \left[ J_+, J_i \right] + \left[ J_+, J_+ \right] + J_+ \left[ J_+, J_i \right] + \left[ J_-^2, J_i \right] = 0$

$= -i \hbar J_k J_e - i \hbar J_e J_k + i \hbar J_k J_e + i \hbar J_e J_k = 0$

\[463/e/p]
Proof of (b): Since $J^2$ and one component of $J$, say $J_z$, can be simultaneously well defined because they commute, we can write

$$J_z \phi_{jm} = \hbar \phi_{jm}, \quad J^2 \phi_{jm} = \hbar^2 j (j+1) \phi_{jm}.$$ 

So far we have made no statements about what $j$ and $m$ are or about their relation to each other. All that the two equations, $J_z \phi_{jm} = \hbar \phi_{jm}$ and $J^2 \phi_{jm} = \hbar^2 j (j+1) \phi_{jm}$, tell us is that $J^2$ and $J_z$ are diagonal matrices. Using the fundamental commutation relations and the definitions

$$J_+ = J_x + i J_y$$
$$J_- = J_x - i J_y$$

we obtain the following further commutation relations:

$$[J_z, J_-] = -\hbar J_, \quad [J_+, J_-] = -\hbar J_-$$

and

$$[J_+, J_-] = 2 \hbar J_z, \quad [J^2, J_+] = [J^2, J_-] = 0$$

From the first one we get

$$J_z J_+ = J_+ (J_z + \hbar)$$
or

$$J_z J_+ \phi_{jm} = J_+ (J_z + \hbar) \phi_{jm} = (m+1) \hbar J_+ \phi_{jm}$$

and from the fourth one we get

$$J^2 J_+ = J_+ J^2$$
or

$$J^2 J_+ \phi_{jm} = J_+ J^2 \phi_{jm} = J_+ (j+1) \hbar^2 J_+ \phi_{jm}$$

5463/e/p
which shows that the function \((J_+ \phi_j^m)\) is an eigenfunction of \(J_m\) to the eigenvalue \((m+1)\hbar\) and also an eigenfunction of \(J^2\) to the eigenvalue \(j(j+1)\hbar^2\). Therefore, it must be possible to write

\[
J_+ \phi_j^m = \hbar A_{jm} \phi_j^{m+1}
\]

where \(A_{jm}\) is some constant dependent on \(j\) and \(m\). By an analogous argument one proves

\[
J_- \phi_j^m = \hbar B_{jm} \phi_j^{m-1}
\]

To find \(A_{jm}\) we make use of the relation

\[
J^2 = J_+ J_+ + J_2 \left( J_2 + \frac{\hbar}{2} \right)
\]

which when written out for the matrix elements gives:

\[
\begin{align*}
\langle j' \mid J_+^2 \mid j \rangle & = \langle j' \mid J_+ \mid j' \rangle \langle j' \mid J_+ \mid j' \rangle + \langle j' \mid J_2 \mid j' \rangle + \hbar \langle j' \mid J_2 \mid j' \rangle \\
\end{align*}
\]

by the usual matrix multiplication rules and using the fact that \(J^2\), \(J_2\) has only diagonal elements, \(J_-\) has only the element for which \(\Delta m = -1\), and \(J_+\) only the one for which \(\Delta m = +1\). Since in addition \(J_- = (J_+)^*\), we have

\[
\langle j' \mid J_+ \mid j \rangle \langle j_1 \mid J_+ \mid j_2 \rangle = \frac{1}{2} \langle j' \mid J_+ \mid j \rangle \langle j' \mid J_+ \mid j \rangle \]

463/4/1

1
\[ j(j+1) = |A_j|^2 + m(m+1) \]

or
\[ |A_j|^2 = j(j+1) - m(m+1) \]

and similarly we find for \( B_j \):
\[ |B_j|^2 = j(j+1) - m(m-1) \]

It is customary to set the phase angles for \( A_j \) and \( B_j \) equal to 0, then
\[ A_j = \sqrt{j(j+1) - m(m+1)} \]
\[ B_j = \sqrt{j(j+1) - m(m-1)} \]

Proof of (c): Before determining \( j \) in (c) we show first that for any \( \varphi_j \), we must have \( m^2 \leq j(j-1) \):
\[ J^2 \varphi_j = (J_x^2 + J_y^2 + J_z^2) \varphi_j \]

or
\[ j(j+1) \varphi_j = (J_x^2 + J_y^2) \varphi_j + m^2 \varphi_j \]

Since \( J_x, J_y \) are hermitian, \( J_x^2 + J_y^2 \) is also hermitian, and the expectation value of the square of an hermitian operator must be non-negative. Hence it follows that
\[ j(j+1) \geq m^2 \]
Let the largest value of \( m \) be denoted by \( k \). This then means that \( \varphi _{j,k+1} = 0 \) because otherwise we could have

\[
\sum_{\mathcal{Z}} \varphi _{j,k+1} = (k+1) \varphi _{j,k+1}
\]

and \( m = k+1 > k \) in contradiction of the statement that the largest value of \( m \) is \( m = k \).

But if \( \varphi _{j,k+1} = 0 \), then \( J \varphi _{j,k} + m \Lambda _{j,k} \varphi _{j,k+1} = 0 \) (since \( \Lambda _{j,k} \) is finite). Using \( J \varphi _{j,k} = (J^2 - J_{\alpha \alpha} - m^2 \mathcal{J}) \varphi _{j,k} \), we get

\[
0 = \int (J^2 - J_{\alpha \alpha} - m^2 \mathcal{J}) \varphi _{j,k} \varphi _{j,k}
\]

or

\[
J^2 (j+1) = k (k+1)
\]

(\( J \) is an integer, hence \( j = -k, -k+1, \ldots, k \)).

Similarly one shows that if \( t = \) smallest value of \( m \), then \( t = -j \), and thus we prove that \( m \) varied from \( -j \) to \( +j \), and we have shown before that neighbouring values of \( m \) differ by unity. Hence it follows that \( 2j = \) integer, hence

\[
j = 0, \frac{1}{2}, 1, \frac{3}{2}, \text{ etc.}
\]

and this concludes the proof of (c). The preceding argument shows that the eigenfunctions \( \varphi _{j,m} \) can be grouped into subgroups \( \varphi _{j,k} \) where in each subgroup we have the same value of \( j \) but different values of \( m \). Such a subgroup \( \varphi _{j,j}, \varphi _{j,j-1}, \ldots, \varphi _{j,-j+1}, \varphi _{j,-j} \) can be thought
of as a state vector with \(2j+1\) components, each component being a state vector itself.

We can immediately write the matrix elements for \(J^2\), \(J_z\), \(J_+\), and \(J_-\) in such a subspace:

\[
J^2 = \hbar^2 (j+1) \begin{pmatrix}
1 & & & \\
& 1 & & \\
& & \\n& & & 1
\end{pmatrix} \quad \text{a (2j+1)x(2j+1) matrix}
\]

\[
J_z = \hbar \begin{pmatrix}
j & & & \\
& j-1 & & \\
& & \\
& & & -j+1
\end{pmatrix} \quad \text{also (2j+1)x(2j+1) matrix}
\]

\[
(J_+)_i^k = \begin{cases}
0 & \text{if } k \neq i+1 \\
\hbar A_{ji} & \text{if } k = i+1
\end{cases} \quad \text{also (2j+1)x(2j+1) matrix}
\]

\[
(J_-)_i^k = \begin{cases}
0 & \text{if } k \neq i-1 \\
\hbar B_{ji} & \text{if } k = i-1
\end{cases} \quad \text{also (2j+1)x(2j+1) matrix}
\]

In particular, if \(j=0\) all the (2j+1)x(2j+1) matrices reduce to a single element and that element is 0. This, of course, was to be expected since \(j=0\) means a spherically symmetric system, and we
would be rather surprised were we to find that a rotation changes something in such a system. When \( j=1 \), we obtain the following 3x3 matrices:

\[
\begin{align*}
J_2 &= 2 \hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
J_x &= \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & i \end{pmatrix} \\
J_y &= \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \\
J_z &= \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{align*}
\]

The form of these 3x3 matrices suggests that a system with \( j=1 \) can be thought of as a vector in our usual three-dimensional space, since the components of such a vector would transform under a rotation exactly by the same rules that the above written matrices specify. Let us check this statement: We want the transformation properties of a vector

\[
\vec{\nu}' = \xi' \vec{u}_x + \eta' \vec{u}_y + \xi' \vec{u}_z
\]

under a rotation \( R = (1 + \vec{r} \cdot d\phi) \). Let the rotation be around the z-axis; \( R_z = (1 + r_z d\phi) \). We can specify the rotation by either finding the new components \( \xi', \eta', \xi' \) of \( \vec{v} \) on the basis vectors \( u_x, u_y, u_z \) or by finding a new basis \( u'_x, u'_y, u'_z \) in terms of which the components of \( \vec{v} \) remain \( \xi, \eta, \xi' \). Doing
it the second way we find
\[
\hat{R}_z = (1 + i \frac{\alpha}{\hbar} d\varphi) = \begin{pmatrix} 1 & i \frac{\alpha}{\hbar} \varphi \\ -i \frac{\alpha}{\hbar} \varphi & 1 \end{pmatrix} = 1 + \alpha \varphi \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]
hence
\[
\hat{J}_z = i \hat{K}_z = \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]
This obviously is not the same expression as we have written before, and the reason is that deriving \( J_z \) this time we didn't specify that we want it diagonal. However, two matrices \( A \) and \( B \) are equivalent if there exists a matrix \( C \) such that
\[
A = CBC^+.
\]
Hence, if we can find a transformation from the basis vectors \( u_x, u_y, u_z \) to some other basis, such that the matrix of that transformation would make it possible to change
\[
\begin{pmatrix} 0 & i \\ -i & 0 \\ 0 & 0 \end{pmatrix}
\]
into an equivalent diagonal matrix, and if this equivalent diagonal matrix should turn out to be
\[
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}
\]
we'll have completed the proof. There are standard methods for diagonalizing a matrix which we shall not go through here. Be it sufficient to say that instead of \( u_x, u_y, u_z \), we use as basis the vectors
\[
\begin{align*}
u_+ &= -(u_x + i u_y) \frac{1}{\sqrt{2}} \\
u_0 &= u_z \\
u_- &= (u_x - i u_y) \frac{1}{\sqrt{2}}
\end{align*}
\]
5463/e/p
The matrix for this transformation is obviously

\[ C = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & -i & 0 \end{pmatrix} \]

Then

\[ C J_2 C^+ = \kappa \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \]

We note that with the basis \( u_+ \), \( u_0 \), \( u_- \), the vector \( v \) becomes

\[ \vec{v} = \xi_+ \vec{u}_+ + \xi_0 \vec{u}_0 + \xi_- \vec{u}_- \]

where

\[ \xi_+ = -\frac{\xi - i\eta}{\sqrt{2}} \]

\[ \xi_0 = \xi \]

\[ \xi_- = \frac{\xi + i\eta}{\sqrt{2}} \]

The reader can convince himself that using for the basis vectors \( u_+ \), \( u_0 \), \( u_- \), and considering rotations of a vector around the \( x \) (or \( y \)) axes, the matrix for \( J_x \) (or \( J_y \)) comes out identical to the \( J_x \) (or \( J_y \)) matrix calculated on page 11.

We can now state: If a state function when rotated in space transforms as a vector, it must have for its angular momentum \( j=1 \); if it transforms as a scalar, \( j \) must be \( 0 \).
Spherical Harmonics

We now look at state functions which can be written as scalar functions of the three-space variables: \( \varphi = \varphi(x, y, z) \).

The following expansion is possible:

\[
\varphi(x, y, z) = \varphi(r, \theta, \phi) = \sum_{\ell_m} \ell_m (r) Y_{\ell m} (\theta, \phi)
\]

where the \( Y_{\ell m} \)'s are called spherical harmonics and are eigenfunctions of the angular momentum operator. We use \( \ell \) and \( m \) instead of \( j \) and \( m \) for the subscripts on the spherical harmonics because we want to reserve \( j \) to stand for a general angular momentum, regardless of the functional dependence of \( \varphi \).

The functional form of the spherical harmonics corresponding to different values of \( \ell \) and \( m \) can be obtained in several ways. We shall try to deduce them in a special way, useful in other realizations of state functions.

\( \ell = 0 \). We know that a state function with angular momentum 0 is rotation invariant, hence \( Y_{00} \) must be a constant, and normalization requirements give

\[
Y_{00} = \frac{1}{(4\pi)^{1/2}}
\]

\( \ell = 1 \). We know that a state function with angular momentum 1 must have three components that transform like components of a vector. Hence by analogy to our previous considerations, we put

\[
Y_{1,1} = x_+
Y_{1,0} = x_0
Y_{1,-1} = x_-
\]
where by the symbol $\doteq$ we mean equal except for constants. As far as the spherical harmonics are concerned, any function of $r$ is also considered a constant. The $x_+, x_0, x_-$ are defined as

$$
x_+ = -(x + iy) \frac{i}{\sqrt{2}}
$$

$$
x_0 = z
$$

$$
x_- = (x - iy) \frac{i}{\sqrt{2}}
$$

in analogy with the definitions of $\bar{u}_+, \bar{u}_0, \bar{u}_-$. However, the quantities $x_+, x_0, x_-$ should not be confused with $\bar{u}_+, \bar{u}_0, \bar{u}_-$: the former are scalars and the latter are vectors. The only thing that they have in common is their transformation properties when operated upon by the angular momentum operator.

Before obtaining the spherical harmonics for $\ell = 2, 3, \ldots$, we prove the following theorem:

**Theorem:** Given two eigenfunctions $\Phi_\ell$ and $\Phi_\ell'$, then

$$
\Phi_\ell \Phi_\ell' \doteq \Phi_\ell \Phi_\ell', \ell + \ell'
$$

**Proof:** First we recall that if $O$ is a linear first-order differential operator and $\varphi, \varphi'$ are two state functions, then

$$
O(\varphi \cdot \varphi') = \varphi(O \varphi') + \varphi'(O \varphi)
$$

Now let

$$
\psi = \Phi_\ell \cdot \Phi_\ell'
$$

then

$$
J_+ \psi = \Phi_\ell \cdot J_+ \Phi_\ell' + \Phi_\ell' \cdot J_+ \Phi_\ell = 0
$$

Also

$$
\frac{J_z}{\ell} \psi = \Phi_\ell \cdot \frac{J_z}{\ell} \Phi_\ell' + \Phi_\ell' \cdot \frac{J_z}{\ell} \Phi_\ell =
$$

$$
= e' \ell \cdot \Phi_\ell \cdot \Phi_\ell' + e \ell \cdot \Phi_\ell \cdot \Phi_\ell' =
$$

$$
= (e' + e) \ell \cdot \psi
$$
and using

\[ J^2 \psi = J^- J^+ \psi + J_2 J_{2+} \psi \]

we get

\[ J^2 \psi = 0 + \hbar^2 (e' + e) (e'+ e + 1) \psi \]
\[ = \hbar^2 L (L+1) \psi \]

where \( L = e + e' \)

and this proves that

\[ \varphi_{e,e',e'} = \varphi_{e+e',e+e'} \]

Now we are ready to build up all spherical harmonics: We know \( \varphi_e \) for \( e = 1 \), hence we can find

\[ \varphi_{2,2} = \varphi_{1+1,1+1} = \varphi_{1,1} \varphi_{1,1} \]

From \( \varphi_{2,2} \) and \( \varphi_{1,1} \), we can get \( \varphi_{3,3} \) and so on. Then by applying \( J^- \) to \( \varphi_{e,e} \) we can get \( \varphi_{e-1,e} \) and so on down the line.

In particular, we obtain for \( e = 2 \):

\[ \varphi_{2,2} = \varphi_{1,1} \varphi_{1,1} = x_+^2 \]

hence

\[ \varphi_{2,2} = x_+^2 \]

\[ \varphi_{2,1} = x_0 x_+ \]

\[ \varphi_{2,0} = x_+^2 + x_0^2 \]

\[ \varphi_{2,-1} = x_0 x_- \]

\[ \varphi_{2,-2} = x_-^2 \]

\[ \varphi_{2,1} = x_0 x_- \]

\[ \varphi_{2,0} = x_+^2 + x_0^2 \]

\[ \varphi_{2,-1} = x_0 x_- \]

\[ \varphi_{2,-2} = x_-^2 \]
where we have made use of the relations
\[ J_x = \frac{\hbar}{i} \sqrt{\Sigma} x_o, \quad J_y = \frac{\hbar}{i} \sqrt{\Sigma} x_0, \quad J_z = 0 \]

By iterating the above procedure we can find \[ Y_{\ell m} \] for any integer value of \( \ell \) and the corresponding values of \( m \).

**Addition of Angular Momenta**

We now consider state functions that are scalar functions of two sets of spatial coordinates: \[ \varphi(x, y, z; x_1 y_1 z_1 x_2 y_2 z_2) \] as they occur in a problem involving two particles. We would then be working in a "six"-dimensional space \( x_1 y_1 z_1 x_2 y_2 z_2 \), where the subscripts refer to the two particles. In such a space, by a rotation \( R \) we mean \( R^{(1)} + R^{(2)} \). The operator \( J \) of the total angular momentum is then
\[ \overrightarrow{J} = \overrightarrow{J}_1 + \overrightarrow{J}_2, \quad J_k = \frac{\hbar}{i} (\overrightarrow{r}_k), \quad k = 1, 2, \]

where \( \overrightarrow{r}_k \) is the infinitesimal rotation in the space 1 or 2.

We will be looking into the question of adding angular momenta, which is posed as follows: Given the eigenfunctions corresponding to two angular momenta \( \ell_1 \) and \( \ell_2 \) of each particle, what is the resultant angular momentum and eigenfunction of the system formed by combining the two?

We write for the state function in terms of polar coordinates for each particle:
\[ \varphi(r, \theta, \phi, r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2) = \sum R_{\ell_1 m_1} R_{\ell_2 m_2} Y_{\ell_1 m_1} (\theta_1, \phi_1) Y_{\ell_2 m_2} (\theta_2, \phi_2) \]
and introduce the definition
\[ \psi_{\ell_1, m_1; \ell_2, m_2} = \psi_{\ell_1, m_1} \begin{pmatrix} \ell_1 \ 0 \\ 0 \ \ell_2 \end{pmatrix} \psi_{\ell_2, m_2}. \]

We then get
\[ \frac{J}{2} \psi_{\ell_1, m_1; \ell_2, m_2} = \hbar \left( m_1 + m_2 \right) \psi_{\ell_1, m_1; \ell_2, m_2} \]

but
\[ \frac{J}{2} \psi_{\ell_1, m_1; \ell_2, m_2} = \text{ad} \psi_{\ell_1, m_1; \ell_2, m_2} \]

is not true in general.

In fact, only if \( \ell_1 = \ell_1 \) and \( m_2 = \ell_2 \) does the above relation hold as we have proved on p.16. In that case
\[ \frac{J}{2} \psi_{\ell_1, m_1; \ell_2, m_2} = \hbar^2 L \left( L + 1 \right) \psi_{\ell_1, m_1; \ell_2, m_2} \]

where \( L = \ell_1 + \ell_2 \).

Hence \( \psi_{\ell_1, m_1; \ell_2, m_2} \) is the eigenfunction of the operator \( J^2 \)
with the eigenvalue \( L(L + 1) \) and of \( J_z \) with eigenvalue \( M = \ell_1 + \ell_2 \).

From this we can find the state functions with lower \( M \) by applying the operator \( J_+ \) a sufficient number of times. It then turns out that the eigenfunction for the total system \( \psi_{L,M} \) can be written as a linear combination of the products of the eigenfunctions of each of the two particles:
\[ \psi_{L,M} = \sum_{m_1, m_2} C_{\ell_1, \ell_2} \left( L M; m_1, m_2 \right) \psi_{\ell_1, m_1} \psi_{\ell_2, m_2} \]

(where we leave out the \( r \)-dependence of \( \psi_{L,M} \)). The double summation actually collapses to a single one because it turns out that
\[ C_{\ell_1, \ell_2} \left( L M; m_1, m_2 \right) \neq 0 \text{ only if } m_1 + m_2 = M. \]

In addition \( L \) can be only equal to \( \left( \ell_1 + \ell_2 \right), \left( \ell_1 + \ell_2 - 1 \right), \ldots, \left( \ell_1 = \ell_2 \right) \)
(if we assume \( \ell_1 > \ell_2 \)).
The coefficients \( \ell_1, \ell_2 \) are called the Clebsch-Gordan or Wigner coefficients.

Starting with \( \Psi_L \ell_1 \ell_2 = \Psi_{LM} \), we get the so-called "stretched" case defined by \( L = \ell_1 + \ell_2 \) (i.e. \( \ell_1 \) and \( \ell_2 \) are lined up): the eigenfunction must then be

\[
\Psi_{LM} = \Psi_L \ell_1 \ell_2
\]

because then

\[
\begin{align*}
J_+ \Psi_{LM} &= 0 \\
J_- \Psi_{LM} &= \ell_1 \ell_2 \Psi_{LM} \\
\hat{J}^2 \Psi_{LM} &= \ell_1^2 + \ell_2^2 (L+1) \Psi_{LM}
\end{align*}
\]

and we find all other \( \Psi_{LM} \) by simply applying the \( J_- \) operator.

Thus, for example:

\[
\Psi_{L, L-1} = J_- \Psi_{L, L} = B_{\ell_1 \ell_2} \Psi_L \ell_1 \ell_2 - 1 \Psi_L \ell_1 \ell_2 + B_{\ell_1 \ell_2} \Psi_L \ell_1 \ell_2
\]

and there are, of course, \( 2L+1 \) such \( \Psi_{LM} \). It is interesting to note that for this, i.e. the "stretched," case setting \( \theta_2 = \theta_1, \sigma_2 = \sigma_1 \) reduces \( \Psi_{LM} \) to \( Y_{LM}(\theta_1 \sigma_1) \). This, of course, is to be expected for the eigenfunctions of the stretched system, since \( \Psi_{LL} \) goes over into \( Y_{LL} \), and the other \( Y_{LM} \)s follow from \( Y_{LL} \) uniquely by applying the operator \( J_- \). The state in which \( L = \ell_1 + \ell_2 \) is only one of many possible states. In general, \( L = \ell_1 + \ell_2 - k \), where \( k \) can vary from 0 to \( 2 \ell_2 \) if \( \ell_1 \geq \ell_2 \) or from 0 to \( 2 \ell_1 \) if \( \ell_1 < \ell_2 \).
to insure that \( L \) takes only one of the following values:

\[
\ell_1 + \ell_2, \quad \ell_1 + \ell_2 - 1, \quad \ldots, \quad |\ell_1 - \ell_2| + 1, \quad |\ell_1 - \ell_2|
\]

as we will see presently.

As usual we first look for that eigenfunction for which

\( M = L \). This is \( \psi_{L,L} \):

\[
\psi_{L,L} = \sum_{n=0}^{K} \psi_{L,j} (L \ell_1, \ell_2-n, \ell_2-k+n) \psi_{L,j}^{(1)} \psi_{L,j}^{(2)} \psi_{L,j-1} \psi_{L,j}^{(1)} \psi_{L,j}^{(2)}
\]

\[
= \psi_{L,j} (L \ell_1, \ell_2-k) \psi_{L,j}^{(1)} \psi_{L,j}^{(2)} + \psi_{L,j} (L \ell_1, \ell_2-k+1) \psi_{L,j-1} \psi_{L,j}^{(1)} \psi_{L,j}^{(2)}
\]

\[
+ \ldots + \psi_{L,j} (L \ell_1, \ell_2-k, \ell_2) \psi_{L,j}^{(1)} \psi_{L,j}^{(2)}
\]

i.e. \( \psi_{L,L} \) is a linear combination of all possible products of \( \psi_{L,j}^{(1)} \) and \( \psi_{L,j}^{(2)} \) such that \( m_1 + m_2 = M \), and in this case \( M = \ell_1 + \ell_2 - k \), hence \( m_1 + m_2 = \ell_1 + \ell_2 - k \). To find all the Clebsch-Gordan coefficients we make use of the requirement that

\[
J_+ \psi_{L,L} = 0
\]

If we use the abbreviations

\[
C_{\ell_1 \ell_2} (L \ell_1, m_1, m_2) \equiv C_{(m_1, m_2)}
\]

and

\[
\psi_{L,j}^{(1)} \psi_{L,j}^{(2)} \equiv \left[ m_1, m_2 \right]
\]

then the effect of \( J_+ \) on \( \psi_{L,L} \) can be written as follows showing how each term (except the first and last) is split into two.

---

5463/e/p
\[ \psi_{LL} = C(\ell_j, \ell_{2-k})[\ell_j, \ell_{2-k}] + C(\ell_{j-1}, \ell_{2-k+1})[\ell_{j-1}, \ell_{2-k+1}] + \ldots \]

\[ \frac{J_+}{k} \psi_{LL} = A_{\ell_2, \ell_{2-k}} C(\ell_j, \ell_{2-k})[\ell_j, \ell_{2-k}] + \]

\[ + A_{\ell_2, \ell_{2-k+1}} C(\ell_{j-1}, \ell_{2-k+1})[\ell_{j-1}, \ell_{2-k+1}] + \]

\[ + A_{\ell_2, \ell_{2-k+2}} C(\ell_{j-2}, \ell_{2-k+2})[\ell_{j-2}, \ell_{2-k+2}] + \]

\[ + \ldots \]

We now pair together (as indicated) terms involving products of the same spherical harmonics and require that the coefficient of each product vanishes and thus obtain equations of the type:

\[ A_{\ell_2, \ell_{2-k}} C(\ell_j, \ell_{2-k}) + A_{\ell_2, \ell_{2-k+1}} C(\ell_{j-1}, \ell_{2-k+1}) = 0 \]

We have started in \( \psi_{LL} \) with (k+1) terms, hence we have (k+1) Clebsch-Gordan coefficients to evaluate. Applying \( J_+ \) to \( \psi_{LL} \) gave us \( 2(k-1) + 2 = 2k \) terms, and by pairing them up we obtain
k equations for the \((k+1)\) unknowns. This, together with normalization requirements, is enough to determine all the \(0(m_1m_2)'s\) uniquely. The \(A_{\ell_m}'s\) are not unknowns—they are given by the formula on pg. 6 with \(j\) replaced by \(\ell\). If \(k\) should exceed its maximum value of \(2\ell_2\) (assuming \(\ell_2 < \ell_1\)), our expression for \(\psi_{LL}\) will automatically give \(\psi_{LL} = 0\). (Note: Our formulae are all written on the assumption that \(\ell_2 < \ell_1\), if this is not the case, simply exchange everywhere \(\ell_1\) and \(\ell_2\).

The eigenfunction corresponding to such a general \(L\) is not unique, since obviously if \(L\) is, for example, 5, it can be cooked up by taking \(\ell_1 = 6\), \(\ell_2 = 1\), \(k = 2\) or by taking \(\ell_1 = 4\), \(\ell_2 = 4\), \(k = 3\) and many more combinations. However, for a given \(\ell_1\) and given \(\ell_2\) there is only one eigenfunction corresponding to any permissible value of \(L\).

We would now like to find explicitly the eigenfunctions for the case when \(\ell_1 = \ell_2 = 1\). Although the procedure described above is quite straightforward, it is somewhat clumsy, and for the simple case \(\ell_1 = \ell_2 = 1\), we shall proceed in a different fashion. Since \(\ell_1 = \ell_2 = 1\), it follows that \(L\) can be 2, 1, or 0.

**Case (a): \(L = 0\)**

We have seen that the \(\psi_{L,M}\) is bilinear in spherical harmonics, i.e. linear in spherical harmonics for each particle. Hence, we want \(\psi_{00}\) also to be bilinear—that is linear in the coordinates of particle 1 and linear in the coordinates of particle 2. For \(L = 0\), the system must be spherically symmetric (rotation invariant). The only form which is bilinear in both
coordinates and which is rotation invariant is
\[ \mathcal{U}_{00} = x^{(1)} x^{(2)} + y^{(1)} y^{(2)} + z^{(1)} z^{(2)} \]
\[ = -x_+^{(1)} x_-^{(2)} - x_-^{(1)} x_+^{(2)} + x_0^{(1)} x_0^{(2)} \]

hence
\[ C_{ll}(001-1) = -1 \quad C_{ll}(0001-1) = \frac{-1}{\sqrt{3}} \]
\[ C_{ll}(00-1+1) = -1 \text{ or normalized: } C_{ll}(0001+1) = \frac{-1}{\sqrt{3}} \]
\[ C_{ll}(0000) = 1 \quad C_{ll}(0000) = \frac{1}{\sqrt{3}} \]

all other \[ C_{ll}(00 \ldots ) = 0 \]

Case (b): \( L = 1 \)

For \( L = 1 \), \( M \) can be \( +1, 0, -1 \), hence we can have
\[ \psi_{1,+1}, \quad \psi_{1,0}, \quad \psi_{1,-1} \]

or in other words our state function has three component state functions, and they must transform by the rules of transformation for \( L = 1 \), i.e. they must transform as a vector. We therefore want a vector linear in the coordinates of particle 1 and linear in the coordinates of particle 2--or in other words a bilinear vector. The cross product of two vectors satisfies these requirements, hence we shall take for the state function
\[ \frac{\chi(1)}{\chi(2)} \]

Then we find \( \psi_{1,+1}, \psi_{1,0}, \psi_{1,-1} \) by taking the appropriate components of \( \frac{\chi(1)}{\chi(2)} \);
\[
\psi_{l+1} = - \frac{(\tilde{\phi}^{(1)} x \tilde{\phi}^{(2)} x + i \tilde{\phi}^{(1)} x \tilde{\phi}^{(2)} y)}{\sqrt{2}} = \frac{i}{2} \left( x_+^{(2)} y_0 - x_0^{(2)} y_+ \right)
\]
\[
\psi_{l,0} = \frac{(\tilde{\phi}^{(1)} x \tilde{\phi}^{(2)} x)}{\sqrt{2}} = \frac{i}{2} \left( x_+^{(1)} x_+^{(2)} - x_-^{(1)} x_-^{(2)} \right)
\]
\[
\psi_{l,-1} = \frac{(\tilde{\phi}^{(1)} x \tilde{\phi}^{(2)} x - i \tilde{\phi}^{(1)} x \tilde{\phi}^{(2)} y)}{\sqrt{2}} = \frac{i}{2} \left( x_0^{(2)} x_- - x_-^{(2)} x_0 \right)
\]

and therefore
\[
C_{ll} (1110) = \frac{i}{2} \quad C_{ll} (101-1) = \frac{i}{2} \quad C_{ll} (1100) = \frac{i}{2}
\]
\[
C_{ll} (1101) = i \quad C_{ll} (1011) = 2i \quad C_{ll} (1101) = 2i
\]
other $C_{ll} (1\ldots) = 0$; other $C (10\ldots) = 0$; other $C_{ll} (1\ldots1) = 0$

or if we normalize them all we get:
\[
C_{ll} (1110) = -\frac{i}{\sqrt{2}} \quad C (101-1) = -\frac{i}{\sqrt{2}} \quad C_{ll} (1100) = \frac{i}{2}
\]
\[
C_{ll} (1101) = \frac{i}{\sqrt{2}} \quad C_{ll} (1011) = \frac{i}{\sqrt{2}} \quad C_{ll} (1101) = \frac{i}{2}
\]

Case (c) : $L = 2$

For $L = 2$, $M$ can be $2,1,0,-1,-2$. The five component state functions can be deduced directly from the expressions derived on page 17. Since those are bilinear as they stand, we simply rewrite them by replacing, for example, $x_0^{(2)}$ by $x_0^{(1)} x_0^{(2)}$. We must make sure, however, that all the expressions are symmetric in the
coordinates of particles 1 and 2—thus, for example, we replace
\( x^+ x^- \) not by \( x^{(1)}_+ x^{(2)}_- \) but by \( \frac{x^{(1)}_+ x^{(2)}_+ + x^{(1)}_- x^{(2)}_-}{2} \). In this fashion
we get

\[
\begin{align*}
\Psi_{2,2} &= x^{(1)}_+ x^{(2)}_- \\
\Psi_{2,1} &= x^{(1)}_+ x^{(2)}_- + x^{(1)}_- x^{(2)}_+ \\
\Psi_{2,0} &= x^{(1)}_+ x^{(2)}_- + x^{(1)}_- x^{(2)}_+ \\
\Psi_{2,-1} &= x^{(1)}_+ x^{(2)}_- + x^{(1)}_- x^{(2)}_+ \\
\Psi_{2,-2} &= x^{(1)}_+ x^{(2)}_-
\end{align*}
\]

and, therefore,

\[
C_{11} (2211) = 1; \quad \text{all other } C_{11} (22..) = 0
\]

\[
C_{11} (2101) = \frac{1}{\sqrt{2}}; \quad C_{11} (2110) = \frac{1}{\sqrt{2}}; \quad C_{11} (21..) = 0
\]

\[
C_{11} (2011) = \frac{1}{\sqrt{2}}; \quad C_{11} (2010) = \frac{1}{\sqrt{2}}; \quad C_{11} (2000) = 1; \quad C_{11} (20..) = 0
\]

\[
C_{11} (2100) = \frac{1}{\sqrt{2}}; \quad C_{11} (2110) = \frac{1}{\sqrt{2}}; \quad C_{11} (21..) = 0
\]

\[
C_{11} (2211) = 1; \quad C_{11} (2210) = \frac{1}{\sqrt{2}}; \quad C_{11} (22..) = 0
\]

and we can normalize all of them by dividing by the square root
of the sum of the squares in each row.

**Vector Fields**

Before we start on this subject, we would like to mention
that just as it was possible to have two realizations of systems
of angular momentum 1, namely the spherical harmonics (or scalar) representation and the vector representation, it is also possible to have similar two representations for higher integer values of angular momentum. For $\ell = 1$, we had:

\[
\begin{align*}
    x_+ &= - (x + i \gamma) \frac{1}{\sqrt{2}} \\
    x_0 &= \gamma \\
    x_- &= (x - i \gamma) \frac{1}{\sqrt{2}}
\end{align*}
\]

spherical harmonics realization

\[
\begin{align*}
    \hat{u}_+ &= - (\hat{u}_x + i \hat{u}_y) \frac{1}{\sqrt{2}} \\
    \hat{u}_0 &= \hat{u}_z \\
    \hat{u}_- &= (\hat{u}_x + i \hat{u}_y) \frac{1}{\sqrt{2}}
\end{align*}
\]

basis vectors in vector realization

For $\ell = 2$, we can write:

\[
\begin{align*}
    x_+^2 & \\
    x_+ x_0 & \\
    x_+ x_- + x_0^2 & \\
    x_- x_0 & \\
    x_-^2 & \\
    \hat{u}_+ \hat{u}_+ & \\
    \hat{u}_+ \hat{u}_0 & \\
    \hat{u}_+ \hat{u}_- + \hat{u}_0 \hat{u}_0 & \\
    \hat{u}_- \hat{u}_0 & \\
    \hat{u}_- \hat{u}_-
\end{align*}
\]

spherical harmonics realization

basis "vectors" in vector realization
The products of the two vectors in the vector realization are the so-called direct products. Given any two vectors we can find their direct product in terms of the basis vectors as follows:

\[
\mathbf{v}_i = \sum_k a_{ik} \mathbf{u}_k \quad \text{and} \quad \mathbf{v}_j = \sum_e a_{je} \mathbf{u}_e
\]

then

\[
\mathbf{v}_i \cdot \mathbf{v}_j = \left( \sum_k a_{ik} \mathbf{u}_k \right) \left( \sum_e a_{je} \mathbf{u}_e \right) = \sum_{k,e} a_{ik} a_{je} \mathbf{u}_k \cdot \mathbf{u}_e
\]

This summation over two indices suggests that we are actually dealing with a tensor of second rank, and the five direct products that we have called basis "vectors" can be looked upon as basic elements of such a tensor. We deduce therefore that an irreducible tensor of second rank must transform like a system of angular momentum two and vice versa. Similar arguments for higher integer values of angular momentum, and so we can conclude that a system characterized by an angular momentum \( J \) (integer) has the same transformation properties as a tensor of rank \( J \). In particular, this proves that spherical harmonics of order \( n \) are isomorphic with irreducible tensors of rank \( n \) as far as transformation properties under rotation are concerned.

We now consider a different realization of angular momentum state vectors, namely vector fields. A vector field \( \mathbf{\nabla}(x,y,z) \) is a vector state function defined at every point in space by a vector:

\[
\mathbf{\nabla}(x,y,z) = \xi (xy) \mathbf{u}_x + \eta (xy) \mathbf{u}_y + \zeta (xyz) \mathbf{u}_z
\]

\[
= \xi x \mathbf{u}_x + \eta (x \mathbf{u}_x + \mathbf{u}_y) + \zeta (x \mathbf{u}_x + \mathbf{u}_y) \mathbf{u}_z
\]

\[
= \sum_{l=0}^{\infty} \xi_l \mathbf{u}_l
\]
where $\hat{u}_+, \hat{u}_-, \hat{u}_0$ is the same vector basis we have been using before and \(\xi_+, \xi_0, \xi_-\) are related to \(\xi, \gamma, \zeta\) in such a way as to satisfy:

\[ \xi_+ \hat{u}_+ + \xi_0 \hat{u}_0 + \xi_- \hat{u}_- = \xi \hat{u}_x + \gamma \hat{u}_y + \zeta \hat{u}_z \]

**Angular Momentum Operator in Vector Fields**

In the case of the six-dimensional scalar field, the operator \(J\) is specified by requiring \(J = J^{(1)} + J^{(2)}\) or \(J = L^{(1)} + L^{(2)}\). For the vector field, we replace all these scalar operators by their corresponding vector equivalents \(\vec{J} = \vec{L}^{(1)} + \vec{L}^{(2)}\) and further change the notation by letting \(\vec{L}^{(2)} = \vec{\nu}\) and dropping the superscripts:

\[ \vec{J} = \vec{L} + \vec{\nu} \]

\(\vec{\nu}\), of course, is nothing else but the angular momentum operator corresponding to \(j = 1\) (or \(l = 1\)); when operating on a state function, it rotates the vectors at each point. \(\vec{L}\), on the other hand, rotates the spatial structure of the state function (i.e. its \(x, y, z\) dependence). The components of \(\vec{\nu}\) are (see p. 12):

\[
\begin{align*}
\nu_x &= \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \nu_y &= \frac{\hbar i}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \nu_z &= \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{align*}
\]

We can think of \(L\) as the orbital angular momentum related to the structure in space of the system and of \(\vec{\nu}\) as a fixed intrinsic angular momentum of value 1.

Hence, it is useful to expand the components \(\xi_0(x, y, z)\) of the vector field in spherical harmonics (upon which \(\vec{L}\) operates):

*) This notion (cp. p. 27) must not be confused with the transformation on properties of \(\vec{J}\), which behaves always like a vector!
\[ \xi_i \left( x_+, x_0, x_- \right) = \xi_i \left( x, \Omega, \delta \right) = \sum_{\ell m} R_{\ell m}^{(\ell)} \Psi_{\ell m}(\Omega, \delta) \]

we can then think of \( \Psi(x, y, z) = \sum_i \xi_i \hat{u}_i \) as a product of two state functions, one for a general \( \ell \) (the \( \xi \) part) and one for \( \ell = 1 \) (the \( \hat{u}_i \) part). Therefore, we can write down immediately an expression for \( \Psi(x, y, z) \), which is a state function of given \( J \) and \( M \) by simple analogy with the state function obtained for the six-dimensional (two-particle) scalar field.

In the two-particle case we had (see p.19)

\[ \Psi_{JM, \ell_1, \ell_2} = \sum_{m_1, m_2} \xi_{\ell_1, \ell_2}^{(JM; m_1, m_2)} \chi_{\ell_1, m_1} \chi_{\ell_2, m_2} \]

If in this expression we set \( \ell_1 = \ell \), \( \ell_2 = 1 \) and replace \( \chi_{\ell_1, m_1} \) by \( \hat{u}_{m_2} \), to which it is equivalent, we obtain:

\[ \Psi_{JM, \ell} = \xi_{\ell}^{(JM)} \sum_{m_1, m_2} C_{\ell_1, \ell_2}^{(JM; m_1, m_2)} \chi_{\ell, m_1} \hat{u}_{m_2} \]

where we have now included the \( r \) dependence of \( \Psi \). The double summation again collapses to a single one since we must have \( m_1 + m_2 = M \). Any given value of \( J \) when \( J \geq 1 \) can be obtained by taking \( \ell = \begin{cases} J+1 \medskip \text{ when } J = 0, \ell \text{ must be } 1, \\ J-1 \end{cases} \]

Let us now look at some examples of vector fields:

\[ J = 0 \]

When \( J = 0, \ell = 1 \)--we therefore want the eigenfunction made up of a product of eigenfunctions of two states each with \( \ell = 1 \),
and total $J = 0$. This is exactly the case worked out on page 24.

and the answer there was

$$
\psi_0 = x^{(1)} x^{(2)} + y^{(1)} y^{(2)} + z^{(1)} z^{(2)}
$$

$$
= -x_+ x_- - x_- x_+ + x_0 x_0
$$

Therefore for the vector field by analogy:

$$
\vec{\psi} = (x \vec{u}_x + y \vec{u}_y + z \vec{u}_z) f(r)
$$

$$
= (-x_+ \vec{u}_x + x_- \vec{u}_- + x_0 \vec{u}_0) f(r)
$$

The vectors $\vec{\psi}$ always point away from or towards the origin.

Thus, we have a spherically symmetric field (as it must be since $J = 0$). Perhaps the best-known example of such a field is the Coulomb field of a point charge. This kind of spherically symmetric vector field is known as "porcupine" field.

$J = 1$

When $J = 1$, $\ell$ could be 0, 1, or 2; for $\ell = 0$, all the $\ell_1$'s must be constants (since a spherical harmonic of order 0 is just a constant) and we must obviously have

$$
\vec{P}_{\ell = 0} = f(r) \vec{u}_0 = \frac{\partial}{\partial \theta} - \frac{\partial}{\partial \phi}
$$

For $M = 0$

$$
\vec{P}_{00} = f(r) \vec{u}_0 = f(r) \vec{u}_z
$$

a constant vector parallel to z-axis

For $M = +1$

$$
\vec{P}_{10} = f(r) \vec{u}_+ = f(r)(\vec{u}_x + i \vec{u}_y)
$$

rotating vectors in the x-y plane always perpendicular to z-axis

For $M = -1$

$$
\vec{P}_{-0} = f(r) \vec{u}_- = f(r)(\vec{u}_x - i \vec{u}_y)
$$

5463/e/p
$J = 1, \ell = 1$

Here again we are adding two $\ell$'s, each equal to 1 except instead of getting $J = 0$ as for the case described on pages 30, 31 we now want $J = 1$. By looking back to the two-particle case, we find that we need a bilinear structure that transforms like a vector—a cross product of two vectors will do. In the two-particle case we had

for $M = 0$

$$\mathcal{U}_{10} = \frac{i}{2} \left( \chi^{(1)}_+ \chi^{(2)}_- - \chi^{(1)}_- \chi^{(2)}_+ \right)$$

hence

$$\mathcal{U}_{101} = \frac{i}{2} f(r) \left( \hat{u}_+ \hat{u}_- - \hat{u}_- \hat{u}_+ \right) =$$

$$= (\chi \hat{u}_y - y \hat{u}_x) f(r) = f(r) (\hat{r} \times \hat{u}_z) \text{ component}$$

This is a field which has at each point a vector perpendicular to the z-axis in the x-y plane—a so-called "circular" vector. By continuing this analogy with the two-particle case, we find for $M = 1$

$$\mathcal{U}_{11} = f(r) \left( \chi_+ \hat{u}_y - \chi_0 \hat{u}_+ \right)$$

for $M = -1$

$$\mathcal{U}_{1-1} = f(r) \left( \chi_0 \hat{u}_y - \chi_- \hat{u}_- \right)$$

The case $J = 1, \ell = 2$ is left as an exercise for the reader.

**Spinors**

All the realizations for the angular momentum that we have discussed so far are valid only for integer $j$'s. But we know that $j$ can take on half integral values as well. By referring back (see page 11) to our initial discussion of $J$ when we did not specify the functional form of the state function on which $J$ operates,
we can immediately write down the matrices for the components of \( J \) for, for example, \( j = 1/2 \):

\[
J_x = S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_y = S_y = \frac{\hbar}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad J_z = S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\]

For the special case of \( j = 1/2 \), these components are designated by \( S_x, S_y, S_z \) instead of \( J_x, J_y, J_z \), and they are said to be the three components of a spin operator \( \mathbf{S} \). Since \((2j+1)\) in this case is 2, there are only two basis "vectors"—these are called spinors and written \( \widetilde{u}_+ \) and \( \widetilde{u}_- \), where the symbol \( \sim \) indicates that these are spinors. Any spinor can be written as \( \varphi = c_+ \widetilde{u}_+ + c_- \widetilde{u}_- \), where \( c_+ \) and \( c_- \) are the spinor components. It is usual to write this as a 1 column, 2 row matrix:

\[
c_+ \begin{pmatrix} \widetilde{u}_+ \\ \widetilde{u}_- \end{pmatrix} + c_- \begin{pmatrix} \widetilde{u}_- \\ -\widetilde{u}_+ \end{pmatrix} = \begin{pmatrix} c_+ \\ c_- \end{pmatrix}
\]

The coefficients \( c \)'s must be complex; otherwise, we would have only two numbers to describe a rotation, and this is not enough.

It may appear that having four numbers is more than necessary. After all, a vector needs only three numbers to describe it. Actually this is due to the fact that a vector has no "thickness"; if we should rotate a vector around its own axis, we would be unable to detect any effect on it. A spinor, on the other hand, is sensible to such a rotation.

The inner product in spinor space is defined as follows:

If \( \varphi = c_+ \widetilde{u}_+ + c_- \widetilde{u}_- \) and \( \psi = \alpha_+ \widetilde{u}_+ + \alpha_- \widetilde{u}_- \), then

\[
\left\langle \varphi, \psi \right\rangle = c_+^* \alpha_+ + c_-^* \alpha_-
\]

5463/e/p
and the square of magnitude: \( |\psi| = \sqrt{\sum \psi_i \psi_i^*} = |c_+|^2 + |c_-|^2 \)

To ask "what is a spinor?" is rather meaningless. After all, we cannot really answer the question, "what is a vector?" Since we can draw arrows on the blackboard, we feel that we know what a vector is. All that we really know are a vector's transformation properties, and when we know that, we know all that there is to know.

Let us, therefore, investigate the transformation properties of a spinor. For this we need to know some of the mathematical properties of \( \mathbf{\sigma} \). If we write \( \vec{\sigma} = \frac{\pi}{2} \vec{\sigma} \)

then we get for the components of \( \mathbf{\sigma} \) the so-called Pauli matrices:

\[
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\]

By direct multiplication we find

\[
\begin{align*}
\sigma_i^2 &= 1 \\
\sigma_i \sigma_k - \sigma_k \sigma_i &= 2 \delta_{ik} \\
\sigma_i \sigma_k + \sigma_k \sigma_i &= 2 \delta_{ik} \\
& \text{for } n \text{ odd}
\end{align*}
\]

\[
\sigma_i \sigma_j = \begin{cases} \sigma_i & \text{for } n \text{ odd} \\ 1 & \text{for } n \text{ even} \end{cases}
\]

and \( i = x, y, z \) any cyclic permutation of \( x, y, z \).

If we rotate the spinor \( \varphi = c_+ u_+ + c_- u_- \) through an angle \( \theta \) around an axis \( \mathbf{\hat{n}} \), we obtain a new spinor \( \varphi' \):

\[
\varphi' = R_{\theta} (\mathbf{\hat{n}}) \varphi
\]
and $\varphi'$ can be written either as $\varphi' = c'_+ u'_+ + c'_- u'_-$ or

$$\varphi' = c'_+ u'_+ + c'_- u'_-$$

We ask for the form of the operator $R_\theta (\hat{n})$ when the latter form is obtained, i.e., the components of $\varphi$ remain the same, and we ask for the transformation properties of the "spinor basis."

Let the amount of rotation be infinitesimal and the axis of rotation be the z-axis. Then

$$R_{d\theta} (\hat{z}) = 1 + \frac{d\theta}{2} J_2 \neq d\theta = 1 - \frac{\imath \sigma_2}{2} d\theta$$

If we repeat this process $n$ times, then the spinor will be rotated through a finite angle $\theta = n d\theta$ and

$$R_\theta (\hat{z}) = \left( R_{d\theta} (\hat{z}) \right)^n = \left( 1 - \frac{\imath \sigma_2}{2} d\theta \right)^n$$

Since $d\theta$ is an infinitesimal, we can make use of the relation

$$(A + B)^n = e^{nB}$$

for small $B$, (here we have $B = -\frac{\imath \sigma_2}{2} d\theta$)

hence

$$R_\theta (\hat{z}) = e^{-\frac{\imath \sigma_2}{2} n d\theta} = e^{-\frac{\imath \sigma_2}{2} \theta}$$

If we now make use of the expansion

$$e^x = 1 + x + \frac{x^2}{2} + \ldots$$

we get :-
\[ R_\theta (z) = \begin{pmatrix} 1 & i \sigma_2 \frac{\gamma}{2} - \bar{c} \frac{\gamma^3}{3!} + i \sigma_2 \frac{\gamma}{2} \left( \frac{\gamma^3}{3!} \right) + \ldots \\
\end{pmatrix} \]
\[ = 1 - i \sigma_2 \frac{\gamma}{2} - \frac{\gamma^2}{2!} + i \sigma_2 \frac{\gamma}{2} \left( \frac{\gamma^3}{3!} \right) + \ldots \\
\[ = \left( 1 - \frac{\gamma^2}{2!} + \ldots \right) - i \sigma_2 \left( \frac{\gamma}{2} - \frac{\gamma^3}{3!} \right) + \ldots \\
\[ = \cos \frac{\gamma}{2} - i \sigma_2 \sin \frac{\gamma}{2} = \cos \frac{\gamma}{2} \begin{pmatrix} 1 & 0 \\
0 & 1 \\
\end{pmatrix} - i \sin \frac{\gamma}{2} \begin{pmatrix} 0 & 1 \\
-1 & 0 \\
\end{pmatrix} \\
\[ = \begin{pmatrix} \cos \frac{\gamma}{2} - i \sin \frac{\gamma}{2} & 0 \\
0 & \cos \frac{\gamma}{2} + i \sin \frac{\gamma}{2} \\
\end{pmatrix} = \begin{pmatrix} e^{-i \frac{\gamma}{2}} & 0 \\
0 & e^{i \frac{\gamma}{2}} \\
\end{pmatrix} \]

and similarly we find for rotations around the other axes the expressions:

\[ R_\theta (x) = \cos \frac{\varphi}{2} \begin{pmatrix} 1 & 0 \\
0 & 1 \\
\end{pmatrix} - i \sin \frac{\varphi}{2} \begin{pmatrix} 0 & 1 \\
-1 & 0 \\
\end{pmatrix} = \begin{pmatrix} \cos \frac{\varphi}{2} & -i \sin \frac{\varphi}{2} \\
-i \sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \\
\end{pmatrix} \]

\[ R_\theta (y) = \cos \frac{\varphi}{2} \begin{pmatrix} 1 & 0 \\
0 & 1 \\
\end{pmatrix} - i \sin \frac{\varphi}{2} \begin{pmatrix} 0 & -i \\
i & 0 \\
\end{pmatrix} = \begin{pmatrix} \cos \frac{\varphi}{2} & -\sin \frac{\varphi}{2} \\
\sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \\
\end{pmatrix} \]

If we want the matrix for a rotation around an arbitrary axis, we can obtain that by making use of Euler theorem. This theorem states that there exist three angles, \( \varphi, \theta, \) and \( \gamma, \) the so-called Euler angles, such that a rotation around an arbitrary axis by an angle can be represented as a rotation by \( \varphi \) around \( z \)-axis, then by \( \theta \) around the \( y \)-axis, and then by \( \gamma \) around the \( z \)-axis.
Hence

\[
R = \begin{pmatrix}
e^{-i \frac{\theta}{2}} & 0 \\
0 & e^{i \frac{\theta}{2}}
\end{pmatrix}
\begin{pmatrix}
\cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\
\sin \frac{\theta}{2} & \cos \frac{\theta}{2}
\end{pmatrix}
\begin{pmatrix}
e^{-i \frac{\theta}{2}} & 0 \\
0 & e^{i \frac{\theta}{2}}
\end{pmatrix}
\]

We now know all that there is to know about spinor, but to get some feeling for them let us consider a few examples:

1. Rotate \( \varphi = C_{\mathbf{u}_+} \begin{pmatrix} c \\ 0 \end{pmatrix} \) through an angle \( \pi \) around the x axis:

\[
\varphi' = R_{\pi}^{(\hat{x})} \varphi = \begin{pmatrix} \cos \frac{\pi}{2} & -i \sin \frac{\pi}{2} \\
-i \sin \frac{\pi}{2} & \cos \frac{\pi}{2}
\end{pmatrix} \begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -ic \end{pmatrix} = -ic C_{\mathbf{u}_-}
\]

2. Rotate \( \varphi = C_{\mathbf{u}_+} \begin{pmatrix} c \\ 0 \end{pmatrix} \) through an angle \( \pi \) around the y axis:

\[
\varphi'' = R_{\pi}^{(\hat{y})} \varphi = \begin{pmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} \\
\sin \frac{\pi}{2} & \cos \frac{\pi}{2}
\end{pmatrix} \begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ c \end{pmatrix} = C_{\mathbf{u}_-}
\]

Thus we obtain the surprising result that a rotation through around x of a spinor pointing in the z direction differs from a rotation through \( \pi \) around y. For a vector we would expect in both cases the result to be \( c \mathbf{u}_- \). This just indicates what we have mentioned before, namely that a vector has no thickness. The two rotations performed above differ by a rotation through \( \pi \) around z:

\[
R_{\pi}^{(\hat{z})} \varphi' = \begin{pmatrix} \ i & 0 \\
0 & -i \end{pmatrix} \begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -ic \end{pmatrix} = \varphi''
\]

and

\[
R_{\pi}^{(\hat{z})} \varphi'' = \begin{pmatrix} \ i & 0 \\
0 & -i \end{pmatrix} \begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -ic \end{pmatrix} = \varphi'
\]
3. Rotate $\varphi$ around the $z$-axis by angle $\theta$:

$$
\varphi' = R_\theta(z)\varphi = \left( \begin{array}{c} e^{-i\frac{\theta}{2}} \begin{pmatrix} 0 \\ e^{i\frac{\theta}{2}} \end{pmatrix} \\ 0 \end{array} \right) = \left( \begin{array}{c} e^{-i\frac{\theta}{2}} c \\ 0 \end{array} \right) = e^{-i\frac{\theta}{2}} c \varphi
$$

4. Rotate $\varphi$ around the $y$-axis by $\frac{\pi}{2}$:

$$
\varphi' = R_{\frac{\pi}{2}}(y)\varphi = \left( \begin{array}{cc} i^{-\frac{\pi}{4}} & -\frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & i^{-\frac{\pi}{4}} \end{array} \right) \begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{c}{\sqrt{2}} \nu_++\frac{c}{\sqrt{2}} \nu_- \\ \frac{c}{\sqrt{2}} \nu_+ \end{pmatrix}
$$

5. Rotate around any axis by $2\frac{\pi}{2}$—then we find that

$$
\varphi' = -\varphi
$$

Example 5 shows that spinors are a two-valued representation. Does this violate any laws of nature? No. The only quantities that we can measure or observe physically are bilinear expressions of the form $\left\{ \varphi, \psi \right\}$ and these are always single-valued, even for spinors.

We can use our basis spinors, in the way we used basis vectors, to construct systems characterized by angular momentum made up out of sums or differences of $j = 1/2$. Thus we can form the direct product of the basis spinors:

$$
\begin{align*}
\frac{1}{\sqrt{2}} (\nu_+ \nu_- + \nu_- \nu_+) & \rightarrow V_+ \\
\frac{1}{\sqrt{2}} (\nu_+ \nu_- - \nu_- \nu_+) & \rightarrow V_0 \\
\nu_- \nu_- & \rightarrow V_-
\end{align*}
$$

The three $V$'s are the three components of a system of total angular momentum $= \frac{1}{2} + \frac{1}{2} = 1$, hence this is a system that has transformation properties of a vector. One could say that a vector in ordinary space is a tensor of second rank in spinor space.
If instead of taking the three possible symmetric direct products we would take the only possible asymmetric one \( u_+ u_+ u_- u_- \), we obtain a system with just one element necessary to describe it and total angular momentum = 0, i.e. obviously a scalar. All the above direct products are bilinear in the spinors and hence single-valued, as vectors and scalars ought to be.

**Direction of a Spinor**

On page 37 we said that a certain spinor was pointing in the z direction. That statement must now be explained since we do not know yet what is meant by "direction" when applied to a spinor. As was pointed out before, a spinor is described by four quantities and since what we ordinarily mean by "direction" needs only three numbers to be completely specified, we shall ascribe to a spinor a direction and a phase.

By the "direction" of a spinor \( \sim c_+ u_+ c_- u_- \), we mean the direction of the z-axis of a co-ordinate system in which the spinor is expressed by \( \sim' = cu_- \). As is suggested by the diagram, we must rotate the spinor through an angle \( \varphi \) around the z-axis and through an angle \( \Theta \) around the y-axis in order to have \( \sim' = cu_- \). The angle \( \varphi \) is the phase angle and the spinor is considered to be lined up with the z-axis only when both \( \Theta \) and \( \varphi \) are reduced to 0.
Thus in order to know the direction and phase of a spinor, we must find $\theta$ and $\varphi$. These we can find from $c_+$ and $c_-$ as follows:

Consider the spinor to be lined up with the $z$-axis so that

$$\tilde{k}' = c_+ u_+ + c_- u_- = \begin{pmatrix} c \\ \varphi \end{pmatrix}$$

We can bring it into the position where

$$\tilde{k}'' = R_{\varphi}(\tilde{z}) \tilde{k}' = \begin{pmatrix} \cos \frac{\varphi}{2} & -\sin \frac{\varphi}{2} \\ \sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \end{pmatrix} \begin{pmatrix} c \\ \varphi \end{pmatrix} = \begin{pmatrix} c \cos \frac{\varphi}{2} \\ c \sin \frac{\varphi}{2} \end{pmatrix}$$

$$\tilde{k} = R_{\theta}(\tilde{r}) \tilde{k}'' = \begin{pmatrix} e^{-i \frac{\theta}{2}} & 0 \\ 0 & e^{i \frac{\theta}{2}} \end{pmatrix} \begin{pmatrix} c \cos \frac{\varphi}{2} \\ c \sin \frac{\varphi}{2} \end{pmatrix} = \begin{pmatrix} c_+ \\ c_- \end{pmatrix}$$

hence $\cot \frac{\varphi}{2} = \frac{|c_+|}{|c_-|}$ and $\varphi = (\text{phase of } c_-) - (\text{phase of } c_+)$.  

**Spinor Field**

By a spinor field we mean a field at every point of which a spinor is prescribed:

$$\Psi(x, y, z) = \psi_+ (x, y, z) u_+ + \psi_- (x, y, z) u_-$$

The angular momentum operator for spinor fields consists of two parts (analogously to the vector field case):

$$\vec{J} = \vec{L} + \vec{S}$$

and the eigenfunctions are:

$$\Psi_{JLM} = C_{\frac{L}{2}} \left( \frac{JM \cdot M - L}{\ell + \frac{1}{2}} \right) \chi_{M+L, \ell} \chi_{M+\frac{1}{2}} + C_{\frac{L}{2}} \left( JM \cdot M + \frac{1}{2} \right) \chi_{M+L, \ell} \chi_{M+\ell - \frac{1}{2}}$$

where $J$ can be only $(\ell + 1/2)$ or $(\ell - 1/2)$. If $J = \ell + 1/2$, we have
the stretched case:

\[ \mathcal{V}_{J, \ell, M=J} = \mathcal{V}_{J+1, \ell-1, M=J} = \mathcal{V}_{J-1, \ell+1, M=J} = 0 \]

If \( J = \ell - \frac{1}{2} \):

\[ \mathcal{V}_{J, \ell, J} = a \mathcal{Y}_{\ell, \ell-1} \mathcal{V}_{J, \ell+1} + b \mathcal{Y}_{\ell, \ell+1} \mathcal{V}_{J, \ell-1} \]

where the coefficients \( a, b \) are found from

All this follows by direct comparison with the case of a vector field, and if something in this section is not clear, the reader is referred back to pages 26 and following.

To close this chapter we include two tables of the Clebsch-Gordan coefficients

**TABLE I**

Clebsch-Gordan Coefficients \( C_{j, j'} (J, M; \ell, m') \) for \( j' = 1/2 \)

<table>
<thead>
<tr>
<th>( J = \ell + \frac{1}{2} ) (stretched case)</th>
<th>( m' = \frac{1}{2} )</th>
<th>( m' = -\frac{1}{2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( j + M + \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} ) ( \frac{1}{2} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( j = j - \frac{1}{2} \)

\( \frac{1}{2} \) \( \frac{1}{2} \) \( \frac{1}{2} \) \( \frac{1}{2} \) \( \frac{1}{2} \) \( \frac{1}{2} \)
TABLE II

\[ C_{jj'} (J M; m m') \text{ for } j'+1 \]

<table>
<thead>
<tr>
<th></th>
<th>( m' = 1 )</th>
<th>( m' = 0 )</th>
<th>( m' = -1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J = j + 1 )</td>
<td>[ \frac{(j+1)(j+M+1)}{(2j+1)(2j+2)} \right)^{1/2} ]</td>
<td>[ \frac{(j-M)(j+M+1)}{2j(j+1)} \right)^{1/2} ]</td>
<td>[ \frac{(j+M)(j-M+1)}{2j(j+1)} \right)^{1/2} ]</td>
</tr>
<tr>
<td>( J = j )</td>
<td>[ -\frac{(j-M)(j-M+1)}{2j(j+1)} \right)^{1/2} ]</td>
<td>[ M \left[ \frac{1}{j(j+1)} \right]^{1/2} ]</td>
<td>[ \frac{(j-M)(j+M+1)}{2j(j+1)} \right)^{1/2} ]</td>
</tr>
<tr>
<td>( J = j - 1 )</td>
<td>[ \frac{(j-M)(j-M+1)}{2j(2j+1)} \right)^{1/2} ]</td>
<td>[ -\frac{(j-M)(j+M+1)}{2j(2j+1)} \right)^{1/2} ]</td>
<td>[ \frac{(j+M+1)(j+M)}{2j(2j+1)} \right)^{1/2} ]</td>
</tr>
</tbody>
</table>
There is experimental evidence (e.g. the Stern-Gerlach experiment) that the intrinsic angular momentum of the electron is half integer. Therefore, we assume, after Pauli, that the electron wave function is a spinor field and we would like to obtain the equation for this spinor field. We will approach this problem by trying to build a formalism for spinor fields in as close analogy as possible to the formalism for the vector fields $\vec{E}$ and $\vec{B}$ of the electromagnetic field.

We therefore, in the first place, assume the existence of two, and not one, spinors $\psi$ and $\chi$ in analogy to the two vectors $\vec{E}$ and $\vec{B}$. This should not be very surprising, since for any field, with the exception of scalar fields, some such extension is necessary to have the equations of the field invariant under the Lorentz transformation. (For example, one way of making the equations of motion in classical mechanics relativistically invariant is to treat time as a fourth co-ordinate.)

Next, we want the wave equation satisfied by the electron. We want this to be relativistically correct. We know that the relation between mass, momentum and energy of a particle in free space is

$$p^2 c^2 + m^2 c^4 = E^2$$

If we recall that the operators for momentum and energy are

$$\vec{p} \rightarrow i \hbar \vec{\nabla}, \quad E \rightarrow i \hbar \frac{\partial}{\partial t}$$

the preceding equation in operator form becomes

$$-\frac{\hbar^2}{2m} \nabla^2 \psi + m c^2 \gamma = -\frac{\hbar^2}{2} \frac{\partial^2 \psi}{\partial t^2}$$
or
\[ (\nabla^2 - \kappa^2) \psi = \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} \]  
(2.1)

where \( k = \frac{mc}{\hbar} \) = reciprocal Compton wavelength.

This equation is known as the Klein–Gordon equation.

However, just as in the case of the electromagnetic field, the wave equation tells us nothing about the relation between the two spinors. We know that each of the two spinors, or better yet, each of the two components of the two spinors, must satisfy Klein–Gordon's equation. To find out the relation between them we need the equivalent of Maxwell equations which give the relation between \( \vec{E} \) and \( \vec{B} \).

As we know Maxwell's equations in free space are
\[
\text{curl } \vec{E} = \frac{1}{c} \frac{\partial \vec{B}}{\partial t} \quad \text{div } \vec{E} = 0
\]
\[
\text{curl } \vec{B} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t} \quad \text{div } \vec{B} = 0
\]  
(2.2)

From these relations we obtain the wave equation for the electromagnetic field, making use of the identity
\[
\text{curl } \text{curl } \vec{A} = -\nabla^2 \vec{A} + \nabla \text{rad div } \vec{A} = -\nabla^2 \vec{A} \quad \text{since } \text{div } \vec{A} = 0
\]

We see that the fundamental characteristics of Maxwell's equations are: they are all linear, first order differential equations, and, when iterated, they result in the wave equation. In addition, the operators involved when operating on a vector produce a vector.

Let us therefore introduce an operator which we will call "spin curl", and write \( \vec{\omega} \), such that when it operates on a spinor it produces another spinor. In addition, we want this spin \( \vec{\omega} \) to be a first order differential operator. The following expression has
the required properties:

\[
\text{curl } \sim = \left( \begin{array}{cc}
\frac{\partial}{\partial z} & -i \frac{\partial}{\partial y} \\
\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} & -\frac{1}{\partial x}
\end{array} \right)
\text{ (a 2 x 2 matrix)}
\]

\( ( \vec{\sigma} \cdot \vec{\nabla} ) \) is obviously a first order differential operator and to convince ourselves that when operating on a spinor it produces a spinor, as far as proper rotations are concerned, we merely note that \( \vec{\nabla} \) transforms as a vector, \( \vec{\sigma} \) transforms as a vector, in proper rotations, hence \(( \vec{\sigma} \cdot \vec{\nabla} )\) transforms as a scalar, therefore if \( \psi \) transforms as a spinor, \(( \vec{\sigma} \cdot \vec{\nabla} ) \psi = \psi \) must transform as a spinor, and we can safely put the spinor symbol \( \sim \) under \( \psi \).

We now prove the relation

\[
\text{curl } \sim = \vec{\nabla} \sim \quad \text{(2.3)}
\]

by putting

\[
\text{curl } \sim = \sum_{i,k} \left( \sigma^i \, \nabla_i \sigma_k \right) = \sum_i \sigma^i \nabla_i \sim
\]

\[+ \frac{1}{2} \sum_{i,k} \left( \sigma^i \sigma^k + \sigma^k \sigma^i \right) \, \nabla_i \nabla_k \nabla^i \nabla_k \]

Since \( \sigma^i \nabla_i \) and \( \sigma^i \sigma^k + \sigma^k \sigma^i \sim \) equation (2.3) follows as an identity. This is similar to the identity that holds for vectors except for the sign of \( \vec{\nabla} \). Furthermore, (2.3) is somewhat more general since it holds for all spinors, while the corresponding vector identity is restricted to vectors \( \vec{F} \) such that \( \text{div} \vec{F} = 0 \).
We are now ready to write Dirac's equations in analogy to (2.2) as follows

\[ \nabla \psi = a_1 \varphi + a_2 \varphi' \]  
\[ \nabla \varphi = b_1 \psi + b_2 \psi' \]  

and then by taking either of the two, applying to it the spin curl operator and making use of the other equation to obtain an equation containing just one of the spinors, we find that in order that the equation so obtained be identical with the Klein-Gordon equation, we must set

\[ a_1 = i k \quad a_2 = \frac{i}{c} \]
\[ b_1 = -i k \quad b_2 = \frac{i}{c} \]

This choice of the a's and b's is not unique unless we require that \( a_2 = b_2 \). The requirement \( a_2 = b_2 \) involves no loss of generality, we have four coefficients to find and only three conditions on them, hence we can impose the fourth condition in any arbitrary way, provided it gives us unique solutions for the a's and b's.

Hence we finally obtain the Dirac equations in what we call the **Maxwell form**:

\[ \nabla \psi - i k \varphi = -\frac{i}{c} \varphi' \]
\[ \nabla \varphi + i k \psi = -\frac{i}{c} \psi' \]  

(2.5)

Since \( \varphi \) has two components, \( \psi^+ \) and \( \psi^- \), and \( \varphi' \) has two components, \( \varphi^+ \) and \( \varphi^- \), we can talk about a four-spinor \( \Psi \) which has four components:

\[ \Psi = \left( \begin{array}{c} \psi^+ \\ \psi^- \end{array} \right) \]

*In fact, the Lorentz invariance implies \( a_2 = b_2 \).
and for this $\Psi$, called Dirac spinor, we can write the two Dirac equations obtained previously as just one equation.

To do this we note that since $\vec{V} = \frac{1}{\hbar} \vec{p}$, we can write $\text{curl} = \frac{i}{\hbar} (\vec{\sigma} \vec{p})$.

Also $E = i\hbar \frac{\partial}{\partial t}$.

Therefore equations (2.5) become

$$\frac{1}{i} (\vec{\sigma} \vec{p}) \Psi + \kappa \Psi = \frac{E}{\hbar c} \Psi$$

$$\frac{1}{i} (\vec{\sigma} \vec{p}) \Psi - \kappa \Psi = \frac{E}{\hbar c} \Psi \quad (2.6)$$

The first equation actually represents two simultaneous equations for the components $\Psi_+$ and $\Psi_-$ and the second equation similarly represents two simultaneous equations for the components $\Psi_+$ and $\Psi_-$. These four simultaneous equations can therefore be written as one equation for a four spinor with the four components $\Psi_+, \Psi_-, \Psi_+,$ and $\Psi_-$. i.e. the Dirac spinor $\Psi$:

$$\left[ \frac{-i}{\hbar} (\vec{\sigma} \vec{p}) + i \kappa \right] \Psi = \frac{E}{\hbar c} \Psi \quad (2.7)$$

where $\beta$ is the 4x4 matrix

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

and $\sigma$ has as its components the 4x4 matrices

5463/e/p
\[ \chi_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \]

i.e.

\[ \chi_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \chi_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \end{pmatrix}, \quad \chi_z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \]

By direct calculation we check that the \( \chi \)'s anticommute with each other and with \( \beta \), and in addition

\[ \chi_x^2 = \chi_y^2 = \chi_z^2 = \beta^2 = 1 \]

The equation, (2.7), is the Dirac equation for the electron in Dirac form—it means nothing else but four simultaneous differential equations, the same equation that we have written before as Dirac equations in "Maxwell form". Actually there are different ways in which to write the Dirac equation in Dirac form. The one chosen here is distinguished by the fact that \( \beta \) is diagonal.

**Maxwell's Equations in Dirac's Formalism**

In order to get better acquainted with Dirac's formalism, we now write Maxwell's equations for the electromagnetic field in Dirac's form.

The two Maxwell equations

\[ \nabla \times \mathbf{E} = \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \]

\[ - \nabla \times \mathbf{B} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} \quad (2.2) \]
actually stand for six simultaneous equations for each of the three components of the vectors $\vec{E}$ and $\vec{B}$. We therefore introduce a six-vector $F$ (just like we had to introduce the four spinor $\Psi$) with components consisting of the three components of $\vec{E}$ and three components of $\vec{H}$:

$$
F = \begin{pmatrix}
E_x \\
E_y \\
E_z \\
H_x \\
H_y \\
H_z \\
\end{pmatrix}
$$

Then all the six equations for the components of $F$ can be written as a single equation

$$
(\vec{\alpha} \cdot \vec{\nabla}) F = \frac{i}{c} \dot{F}
$$

(2.8)

To see what $\vec{\alpha}$ is in above equation, we write out the six Maxwell equations:

\begin{align*}
\text{curl} \vec{\Psi} & = \frac{i}{c} \dot{\vec{E}} \\
-\text{curl} \vec{\Psi} & = \frac{i}{c} \dot{\vec{B}}
\end{align*}

$$
\begin{align*}
\text{curl} \vec{\Psi} & = \frac{i}{c} \dot{\vec{E}} \\
-\text{curl} \vec{\Psi} & = \frac{i}{c} \dot{\vec{B}}
\end{align*}
$$

\begin{align*}
\begin{array}{cccccc}
\frac{\partial}{\partial z} E_y & - \frac{\partial}{\partial x} E_z & \frac{\partial}{\partial z} H_x & - \frac{\partial}{\partial x} H_y & \frac{\partial}{\partial z} H_z & = \frac{i}{c} \dot{E}_x \\
\frac{\partial}{\partial z} E_z & - \frac{\partial}{\partial y} E_y & - \frac{\partial}{\partial z} H_y & \frac{\partial}{\partial y} H_z & \frac{\partial}{\partial z} H_z & = \frac{i}{c} \dot{E}_y \\
\frac{\partial}{\partial z} E_x & - \frac{\partial}{\partial y} E_z & \frac{\partial}{\partial z} H_x & - \frac{\partial}{\partial y} H_y & \frac{\partial}{\partial z} H_z & = \frac{i}{c} \dot{E}_z \\
\frac{\partial}{\partial z} H_y & - \frac{\partial}{\partial x} E_x & \frac{\partial}{\partial z} H_x & - \frac{\partial}{\partial x} H_y & \frac{\partial}{\partial z} H_z & = \frac{i}{c} \dot{H}_x \\
\frac{\partial}{\partial z} H_z & - \frac{\partial}{\partial y} E_y & \frac{\partial}{\partial z} H_y & - \frac{\partial}{\partial y} H_y & \frac{\partial}{\partial z} H_z & = \frac{i}{c} \dot{H}_y \\
\frac{\partial}{\partial z} H_x & - \frac{\partial}{\partial x} E_x & \frac{\partial}{\partial z} H_x & - \frac{\partial}{\partial x} H_y & \frac{\partial}{\partial z} H_z & = \frac{i}{c} \dot{H}_z \\
\end{array}
\end{align*}

63/e/p
Now since \( \tilde{\alpha} \cdot \tilde{D} = \alpha_x \frac{\partial}{\partial x} + \alpha_y \frac{\partial}{\partial y} + \alpha_z \frac{\partial}{\partial z} \)
we see immediately that \( \alpha_x \) must be a 6x6 matrix with 0 everywhere except for 1 or -1 at the intersection of those rows and columns where \( \frac{\partial}{\partial x} \) or \( \frac{\partial}{\partial x} \) appears in the array formed by the six Maxwell equations. The requirements on \( \alpha_y \) and \( \alpha_y' \) are similar. Thus we get:

\[
\alpha_x F = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & +1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & +1 & 0 & 0 & 0 & 0 \\
0 & 0 & +1 & 0 & 0 & 0
\end{pmatrix}
\]

\[
\alpha_y F' = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & +1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

\[
\alpha_z F = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & +1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

Instead of above, we can write

\[
\alpha_x F = \begin{pmatrix}
0 & K_x \\
-K_x & 0
\end{pmatrix}, \quad \alpha_y F = \begin{pmatrix}
0 & K_y \\
-K_y & 0
\end{pmatrix}, \quad \alpha_z F = \begin{pmatrix}
0 & K_z \\
-K_z & 0
\end{pmatrix}
\]

which brings out more clearly the similarity in structure of these \( \alpha 's \) with the ones that are contained in the Dirac equation for the four-spinor. Therefore, it is not surprising that the \( k 's \) turn out to be nothing else but the \( r_x, r_y, r_z \) matrices for infinitesimal rotations. As we know,\(^{*)} \) the equivalence transformation that makes

\(^{*)} \) compare p. 13.
r_z diagonal changes r_x, r_y, r_z into v_x, v_y, v_z (except for the $\frac{\hbar}{i}$ factor), i.e. into the matrix elements of the angular momentum operator to the eigenvalue j - 1. Thus we can write (forgetting the $\frac{\hbar}{i}$)

$$\alpha_i^F = \begin{pmatrix} 0 & \nu_i \\ -\nu_i & 0 \end{pmatrix}$$

a 6x6 matrix which clearly shows the similarity with the $\alpha$ matrices for the four-spinor.

The Physics in Dirac's Equations

To obtain physical information from Dirac's equation, we make use of the relation defining the average value $\bar{O}$ of an operator $O$:

$$\bar{O} = \int \{ \Psi, O \Psi \} \, d\tau$$

(2.9)

where $d\tau$ represents a volume element, $\Psi$ represents Dirac's four-spinor and $\{ \}$ stands for the inner product in four-spinor space. The latter is defined in analogy to our previous definition of inner products:

let $\Psi^{(1)}$ be a four-spinor with components $\psi_+^{(1)}, \psi_-^{(1)}, \varphi_+^{(1)}, \varphi_-^{(1)}$

and $\Psi^{(2)}$ be a four-spinor with components $\psi_+^{(2)}, \psi_-^{(2)}, \varphi_+^{(2)}, \varphi_-^{(2)}$

then

$$\{ \Psi^{(1)}, \Psi^{(2)} \} = \{ \psi^{(1)}, \psi^{(2)} \} + \{ \varphi^{(1)}, \varphi^{(2)} \}$$

$$= \sum_{i=1}^{4} \Psi_i^{(1)} \Psi_i^{(2)}$$
where in the last expression \( \hat{\Psi} \) stands for any one of the four components of the four-spinor.

If Dirac's equation is to give us a proper physical picture of the electron, then we must expect the inner product of \( \hat{\Psi} \) with itself to represent the probability (as a function of space coordinates) of finding the electron at some point in space and therefore, that inner product multiplied by the electronic charge should give us the charge density \( \rho(x, y, z) \). If

\[
\rho = e \left\{ \hat{\Psi}, \hat{\Psi} \right\} = e \sum_{i=1}^{4} \left| \hat{\Psi}_i \right|^2
\]

(2.10)

then what expression do we get for \( \mathbf{J} \), the current density? To find \( \mathbf{J} \), we make use of the continuity equation which states

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0
\]

(2.11)

Hence, if our expression for \( \rho \) is correct, taking the time derivative of it should lead to the divergence of some vector:

\[
\frac{\partial \rho}{\partial t} = \left\{ \hat{\Psi}, \hat{\Psi} \right\} + \left\{ \hat{\Psi}, \hat{\Psi} \right\}
\]

\[
= \left\{ \left( \frac{\mathbf{c}}{\mathbf{i} \hbar} \hat{\alpha} \hat{\rho} + \frac{\mathbf{c}}{\mathbf{i} \beta \mathbf{k}} \right) \hat{\Psi}, \hat{\Psi} \right\} + \left\{ \hat{\Psi}, \left( \frac{\mathbf{c}}{\mathbf{i} \hbar} \hat{\alpha} \hat{\rho} + \frac{\mathbf{c}}{\mathbf{i} \beta \mathbf{k}} \right) \hat{\Psi} \right\}
\]

where we have substituted for \( \hat{\Psi} \) its expression in terms of \( \hat{\Psi} \) from Dirac's equation:

\[
\hat{\Psi} = \left( \frac{\mathbf{c}}{\mathbf{i} \hbar} \hat{\alpha} \hat{\rho} + \frac{\mathbf{c}}{\mathbf{i} \beta \mathbf{k}} \right) \hat{\Psi}
\]

To reduce the expression for \( \rho \) to something more decent looking, we recall that \( \hat{\alpha} \), \( \hat{\rho} \) and \( \hat{\beta} \) are all hermitian operators,
therefore, $\frac{\alpha}{i}$ and $\frac{\beta}{i}$ are antihermitian, and the expression for $\dot{\hat{p}}$ would reduce identically to 0 if it was not for the presence of the operator $\vec{\nabla} \cdot \vec{\omega}$.

Hence, omitting the terms involving $\frac{\beta}{i}$ which cancel each other, we have

$$\dot{\hat{p}} = \left\{ \frac{c}{i \hbar} \vec{\omega} \cdot \hat{p} \psi, \psi \right\} + \left\{ \psi, \frac{c}{i \hbar} \vec{\omega} \cdot \hat{p} \psi \right\}$$

$$= -c \left[ \left\{ \vec{\omega} \cdot \vec{\nabla} \psi, \psi \right\} + \left\{ \psi, \vec{\omega} \cdot \vec{\nabla} \psi \right\} \right]$$

Consider now taking the divergence of $\left\{ \psi, \vec{\omega} \psi \right\}$

$$\text{div} \left\{ \psi, \vec{\omega} \psi \right\} = \vec{\nabla} \cdot \left\{ \psi, \vec{\omega} \psi \right\}$$

$$= \left\{ (\vec{\nabla} \cdot \psi, \vec{\omega} \psi) + (\psi, \vec{\omega} \cdot \vec{\nabla} \psi) \right\} =$$

$$= \left\{ \vec{\nabla} \cdot \psi, \vec{\omega} \psi \right\} + \left\{ \psi, \vec{\omega} \cdot \vec{\nabla} \psi \right\}$$

and since $\vec{\omega}$ and $\vec{\nabla}$ commute with each other, we conclude

$$\dot{\hat{p}} = -e c \text{div} \left\{ \psi, \vec{\omega} \psi \right\}$$

and therefore,

$$\vec{j} = e c \left\{ \psi, \vec{\omega} \psi \right\}$$  \hspace{1cm} (2.12)

In the case of the electromagnetic field, the inner product of $\vec{F}$ with itself should represent energy density $u$:

$$u = \frac{1}{8\pi} \left\{ \vec{F}, \vec{F} \right\}$$

The continuity equation for the electromagnetic field is

$$\dot{u} = -\text{div} S$$

where $S$ must be the Poynting vector. We therefore investigate the

5463/e/p
time derivative of \( \{F,F\} \):

\[
\dot{\mathbf{u}} = \frac{1}{8\pi} \left( \{ \dot{F}, F \} + \{ F, \dot{F} \} \right) = \frac{1}{8\pi} \left( \{ c(\nabla \cdot \mathbf{F}), F \} + \{ F, c(\nabla \cdot \mathbf{F}) \} \right)
\]

\[
= \frac{c}{8\pi} \nabla \cdot \{ F, \alpha F \}
\]

and so we conclude that \( \dot{S} = \frac{c}{8\pi} \{ F, \alpha F \} \). To check this, let us compute the x-component of \( \dot{S} \):

\[
\dot{S}_x = \frac{c}{8\pi} \{ F, \alpha_x F \} = -\frac{c}{8\pi} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \\ H_x \\ H_y \\ H_z \end{pmatrix}
\]

\[
= -\frac{c}{8\pi} \begin{pmatrix} E_x & E_y & E_z & H_x & H_y & H_z \end{pmatrix} \begin{pmatrix} 0 \\ -H_x \\ H_y \\ 0 \\ -E_y \\ 0 \end{pmatrix}
\]

\[
= -\frac{c}{8\pi} \begin{pmatrix} 2H_y & E_z & -2H_z & E_y \end{pmatrix}
\]

Non-relativistic Approximation

The statement that \( E \) can be represented by the operator \( \frac{\partial}{\partial t} \), actually means that

\[
\frac{i\hbar}{\partial t} \psi = E\psi
\]

5463/e/p
or in other words, this is to say that the wave function must contain as its time dependence the factor $e^{-\frac{1}{\hbar} \frac{\mathbf{p} \cdot \mathbf{E}}{m}}$. For a relativistically correct theory, this $E$ must be the total energy of the particle

$$E = mc^2 + \epsilon$$  \hspace{1cm} (2.13)

where $m$ = rest mass, and $\epsilon$ = kinetic energy. We can, therefore, write the two spinors $\psi_0$ and $\psi$ as

$$\psi = \psi_0 e^{-\frac{i}{\hbar} \frac{mc^2}{E} t} \quad \text{and} \quad \psi = \psi_0 e^{-\frac{i}{\hbar} \frac{mc^2}{E} t}$$

where $\psi_0$ and $\psi_0$ are time dependent but contain only the non-relativistic part of the energy. Introducing this into (2.5) gives, after cancelling the common exponential factor:

$$\nabla \times \psi_0 - \frac{i}{\hbar} \frac{mc}{E} \psi_0 = -\frac{1}{c} \psi_0 + i \frac{mc}{E} \psi_0$$

$$\nabla \times \psi_0 + \frac{i}{\hbar} \frac{mc}{E} \psi_0 = -\frac{1}{c} \psi_0 + i \frac{mc}{E} \psi_0$$  \hspace{1cm} (2.14)

At non-relativistic energies $\epsilon \ll mc^2$ and, therefore, we can drop $-\frac{1}{c} \psi_0$ in comparison with $i \frac{mc}{E} \psi_0$ (Note that because of the cancelling of terms, we cannot do the same in the second equation). We can, therefore, write the first equation as

$$\nabla \times \psi_0 \approx 2i \frac{mc}{E} \psi_0$$

or

$$\psi_0 \approx \frac{1}{2i} \frac{\hbar}{mc} \nabla \times \psi_0$$
Recalling that \( \text{curl} = (\mathbf{\sigma} \cdot \nabla) \), we see that

\[
\hat{\psi} = \frac{1}{2i} \frac{\mathbf{r}}{m} \text{curl} \ \psi_0
\]

(2.15)

If \( \psi_0 \) is an eigenfunction of \( \hat{\rho} \) then when operating on \( \psi_0 \) \( \hat{\rho} \) produces a product of its eigenvalue with \( \psi_0 \); if, however, \( \psi_0 \) is not an eigenfunction, we can write it as a linear combination of such eigenfunctions, and we then obtain \( \hat{\rho} \psi_0 = \sum_i \rho_i \psi_i \), where \( \rho_i \) are the eigenvalues and \( \psi_i \) the eigenfunctions.

Since we are interested here in orders of magnitude, we see that we can write \( (\mathbf{\sigma} \cdot \hat{\rho}) \psi_0 = \overline{\hat{\rho}} \psi_0 \) where \( \overline{\hat{\rho}} \) is the average momentum. Hence we conclude that

\[
\hat{\psi} = \frac{\overline{\hat{\rho}}}{mc} \hat{\psi} = \frac{\overline{\rho}}{c} \psi_0
\]

(2.16)

at non-relativistic energies. This is why \( \hat{\psi} \) is sometimes called the "small" spinor.

We now look at the form the Klein-Gordon equation assumes in the non-relativistic approximation. We have as the two equations for the spinors:

\[
\hat{\psi} = \frac{i}{2i mc} \text{curl} \ \psi_0
\]

\[
\text{curl} \ \hat{\psi} = \frac{1}{c} \psi_0
\]
Taking the curl of the first equation, and substituting for curl \( \mathcal{L}_0 \) from the second one, we obtain

\[
-\frac{\hbar^2}{2m} \nabla^2 \Psi_0 = i \frac{e}{\hbar} \dot{\Psi}_0
\]

(2.17)

which is nothing else but Schrödinger's equation for the function \( \Psi_0 \) (actually, we have here Pauli's equation, since we are dealing with spinors).

Thus we see that in the non-relativistic limit, the Klein-Gordon equation, which states that \( E^2 = m^2 c^4 + p^2 c^2 \), changes into Schrödinger's (Pauli's) equation, which states that \( \frac{p^2}{2m} = \varepsilon \) as we might have expected.

We now investigate the expression for \( \mathbf{j} \) in the non-relativistic approximation. We recall that

\[
\mathbf{j} = e \mathbf{c} \left[ \Psi, \dot{\Psi} \right] = e \mathbf{c} \left[ \left\{ \Psi, \dot{\Psi} \right\} + \left\{ \Psi, \mathbf{\sigma} \cdot \nabla \Psi \right\} \right]
\]

To simplify matters, we consider one, say the x, component of \( \mathbf{j} \) and we make the non-relativistic approximation by taking

\[
\mathcal{L}_0 = \frac{\hbar}{2i mc} \text{curl} \ \dot{\Psi}_0
\]

\[
\mathbf{j}_x = e \mathbf{c} \left[ \left\{ \Psi_0, \sigma_x \dot{\Psi}_0 \right\} + \left\{ \Psi_0, \sigma_x \Psi_0 \right\} \right]
\]

\[
= \frac{e \hbar}{2im} \left[ \left\{ \Psi_0, \sigma_x \left( \mathbf{\sigma} \cdot \nabla \right) \dot{\Psi}_0 \right\} - \left\{ \sigma_x \mathbf{\sigma} \cdot \nabla \Psi_0, \sigma_x \dot{\Psi}_0 \right\} \right]
\]
\[
\frac{e\hbar}{2im} \left[ \{ \psi_0, \sigma_\xi \frac{\partial}{\partial \xi} \} + \sigma_\xi \frac{\partial}{\partial \xi} + \sigma_\xi \frac{\partial}{\partial \zeta} \right] \psi_0 = \left[ \{ \psi_0, \sigma_\xi \frac{\partial}{\partial \xi} \} - \{ \sigma_\xi \frac{\partial}{\partial \xi} \}, \psi_0 \right] = \\
= \frac{e\hbar}{2im} \left[ \{ \psi_0, \frac{\partial}{\partial \xi} \} - \{ \frac{\partial}{\partial \xi} \}, \psi_0 \right] + \frac{e\hbar}{2m} \left[ \psi_0, \sigma_\xi \psi_0 \right] - \frac{\partial}{\partial \zeta} \left\{ \psi_0, \sigma_\xi \psi_0 \right\} = \\
= \left( \nabla_{\text{Schrödinger}} \right)_\xi + \frac{e\hbar}{2m} \left( \text{curl} \left\{ \psi_0, \sigma_\xi \psi_0 \right\} \right)_\xi.
\]

Where \( \nabla \) Schrödinger stands for the first two terms, since they are of the same form as the expression obtained for \( \nabla \) in Schrödinger theory (except for the fact that the \( \psi_0 \)'s are spinors, therefore, we should perhaps call those terms \( \nabla \) Pauli).

Letting \( \nabla \vec{M} = \frac{e\hbar}{2mc} \left\{ \psi_0, \sigma_\xi \psi_0 \right\} \) we can write

\[ \vec{j} = \vec{j}_{\text{Schr.}} + \epsilon \left( \text{curl} \ nabla \vec{M} \right) \quad (2.18) \]

As we know, a current density which is the curl of some vector \( \vec{M} \), produces the same field as a magnetic dipole with magnetic dipole density per unit volume equal to \( \vec{M} \). Hence, we conclude that the electron must have just such magnetic dipole density, and therefore, its total magnetic dipole must be

\[
\int \vec{M} \, d\tau = \frac{e}{mc} \int \left\{ \psi_0, \frac{\hbar}{2} \sigma_\xi \psi_0 \right\} \, d\tau = \frac{e}{mc} \int \left\{ \psi_0, \frac{\sigma_\xi}{2} \psi_0 \right\} \, d\tau
\]

5463/a/p
But $\int \vec{y}_0 \cdot \vec{\gamma} dt$ is nothing else but the intrinsic angular momentum of the electron and so we obtain the remarkable result that the gyromagnetic ratio of the electron is $\frac{e}{2mc}$ or $g \frac{e}{2mc}$ with $g=2$, and not $\frac{e}{2mc}$ ($g=1$).

What is remarkable about getting the correct $g$-factor is that it follows directly from Dirac's equation and is not put in artificially. One could argue that it really is put in artificially, by the assumption that the electron wave function is a spinor field. Of course this is so. However, should we take simply Schrödinger's equation and replace the vector functions in it by spinors, we would not be able to say anything about the electron's magnetic moment.

What we would obtain is simply Pauli's theory, which, as we know, requires the additional statement that, due to experimental evidence, we must ascribe to the electron a magnetic moment and furthermore, that the $g$-factor should be 2.

We now proceed to compute the current density $j$ exactly. If we take equation (2.7), and premultiply it by $\beta$ we obtain

$$\left( \frac{k}{i} \beta (\alpha \cdot \vec{p}) + \kappa - \beta \frac{e}{kc} \right) \vec{\psi} = 0$$

hence

$$\vec{\psi} = \frac{i}{k} \left[ \frac{\beta E}{kc} - \frac{k}{i} \beta \alpha \cdot \vec{p} \right] \vec{\psi} = \frac{2i}{kc} \left[ \beta \frac{\delta \vec{F}}{\delta \vec{P}} + \beta c \alpha \cdot \vec{F} \right] \vec{\psi} \quad (2.19)$$

The expression for the $x$-component of $j$ is given by eq. (2.12):

$$j_x = \frac{2e}{\hbar} \left\{ \frac{\delta \vec{F}}{\delta \vec{P}} \alpha_x \vec{\psi} \right\}$$

By replacing first the first $\vec{\psi}$ by (2.19), and then the second $\vec{\psi}$ and taking half of the sum of the two expressions, we obtain:
\[ J_x = \frac{ie}{2\hbar} \left[ \left\{ \Psi, \alpha_x (\beta c \overrightarrow{\nabla} + \beta \frac{\partial}{\partial x}) \Psi \right\} - \{ c \overrightarrow{\nabla} + \beta \frac{\partial}{\partial x} \}_{\Psi, \alpha_x \Psi} \right] \]

The terms 1 and 3 together with 5 and 6 give

\[ \frac{\partial}{\partial t} \left\{ \Psi, \alpha_x \beta \Psi \right\} \]

(note that \( \alpha_x \) and \( \beta \) anticommutate)

The terms 1 and 2 together with 4 and 6 give two types of terms:

when we take from \( \alpha \overrightarrow{\nabla} \) the part \( \alpha_x \Psi \) we get the terms

\[ \left\{ \frac{\partial}{\partial x} \Psi, \beta \Psi \right\} - \left\{ \Psi, \beta \frac{\partial}{\partial x} \Psi \right\} \]

and when we consider the terms involving \( \alpha \overrightarrow{\nabla} + \alpha_x \overrightarrow{\nabla} \) we get the terms that correspond to

\[ -ic \left[ \text{curl} \left\{ \Psi, \overrightarrow{\sigma} \beta \Psi \right\} \right] \]

Combining all these terms with what we get when we compute \( j_y \) and \( j_z \), we finally obtain

\[ \overrightarrow{j} = \frac{e\hbar}{2im} \left[ \left\{ \frac{\Psi}{\beta \overrightarrow{\nabla}} \Psi \right\} - \left\{ \overrightarrow{\nabla} \frac{\Psi}{\beta} \Psi \right\} + ic \text{curl} \left\{ \Psi, \beta \overrightarrow{\sigma} \Psi \right\} \right] + \frac{1}{c} \frac{\partial}{\partial t} \left\{ \Psi, \beta \overrightarrow{\sigma} \Psi \right\} \]

which we can write as

\[ \overrightarrow{j} = \overrightarrow{j}_{\text{Schr.}} + c \text{curl} \overrightarrow{\nabla} + \frac{\partial}{\partial t} \overrightarrow{p} \]

where

\[ \overrightarrow{j}_{\text{Schr.}} = \frac{e\hbar}{2im} \left[ \left\{ \frac{\Psi}{\beta \overrightarrow{\nabla}} \Psi \right\} - \left\{ \overrightarrow{\nabla} \frac{\Psi}{\beta} \Psi \right\} \right] \]

\[ \overrightarrow{\nabla} = \left\{ \Psi, m \frac{\overrightarrow{e}}{c} \beta \overrightarrow{\sigma} \Psi \right\} \]

\[ \overrightarrow{p} = \frac{e\hbar}{2mc} \left\{ \Psi, \beta \overrightarrow{\sigma} \Psi \right\} \]

463/e/p
We know already why the curl \( \mathbf{A} \) term appears in addition to \( \mathbf{j} \) Schr., but what is the meaning of \( \frac{\partial \mathbf{P}}{\partial t} \)? As we know, a current that can be written as the time rate of change of some vector \( \mathbf{P} \) implies the presence of electric polarization with density \( \mathbf{P} \).

There is nothing unusual in the appearance of this \( \mathbf{P} \). We know that in the theory of relativity, electric and magnetic phenomena are tied up together quite intimately, and therefore, if the electron possesses a magnetic moment, we should expect it to have an electric moment too if it is in motion. This follows from the fact that magnetic fields have electric parts if seem from a moving system. Thus if \( \mathbf{P} \) had not appeared in our equation for the current density, we would need some explanation for it.

To finish our discussion of \( \mathbf{P} \) we would like to point out that \( \mathbf{P} \) is not imaginary as would appear from the presence of \( i \) in the expression for it. To find the actual value of \( \mathbf{P} \), we need the eigenvalues of the operator \( \beta \tilde{\alpha} \). \( \beta \) is hermitian and \( \tilde{\alpha} \) is hermitian, therefore

\[
\left\{ \Psi, \beta \tilde{\alpha} \Psi \right\} = \left\{ \beta \Psi, \tilde{\alpha} \Psi \right\} = \left\{ \mathbf{2}/\beta \Psi, \Psi \right\} = -\left\{ \beta \mathbf{2} \Psi, \Psi \right\}
\]

where the minus sign follows from \( \alpha \beta = -\beta \alpha \). Hence we conclude that the operator \( \beta \tilde{\alpha} \) is antihermitian. The eigenvalues of an anti-hermitian operator are purely imaginary.

\( \frac{\beta \tilde{\alpha}}{\mathbf{i}} \), therefore, is hermitian and its eigenvalues are real.

Beside the addition of the term \( \frac{\partial \mathbf{P}}{\partial t} \), the expression (2.20) for \( \mathbf{j} \) differs from the corresponding nonrelativistic equation (2.18).
by the presence of the factor \( \beta \) in both \( \hat{J} \) Schr. and \( \hat{M} \). We shall show later that the average value of \( \beta \) is \( \sqrt{1 - \frac{v^2}{c^2}} \) and so it reduces to 1 in the nonrelativistic case. It expresses the fact, however, that the mass \( m \) which appears in all expressions changes in the familiar way if the particle is in motion.

Solutions of the Dirac Equation

In obtaining the solutions for Dirac equation, we again start with the analogy of the Maxwell equations for the electro-magnetic field. In those, as we know, the space-time dependence of the solutions must be such as to satisfy the space-time dependence of the wave equation

\[
\nabla^2 \psi + \frac{k^2}{c^2} \psi = 0
\]

From that we obtain

\[
\psi = \psi_0 \exp \left( \frac{i}{\hbar} \vec{k} \cdot \vec{r} - i \nu t \right)
\]

where \( |\vec{k}| = \frac{\nu}{c} \) and a similar expression for \( \psi \). From Maxwell equations themselves, we obtain \( \psi \perp \vec{E} \) and both \( \vec{E} \) and \( \vec{B} \) perpendicular to the direction of propagation.

Therefore, we will expect the solutions of Dirac equation to have a space-time dependence such as to satisfy

\[
\nabla^2 \psi - k^2 = \frac{\hbar^2}{c^2} \frac{\partial^2 \psi}{\partial t^2}
\]

This leads to:

\[
\psi = \psi_0 \exp \left( \frac{i}{\hbar} \vec{k} \cdot \vec{r} - i \nu t \right)
\]

\[
\psi = \psi_1 \exp \left( \frac{i}{\hbar} \vec{k} \cdot \vec{r} - i \nu t \right)
\]

\[
(2.21)
\]

where \( k \) must satisfy

\[
\vec{k}^2 + \nu^2 = \frac{\hbar^2}{c^2}
\]

(2.22)

or putting in (2.22) and (2.21) \( \vec{p} = \hbar \vec{k} \), \( \vec{K} = \frac{mc}{\hbar} \), \( E = \hbar \nu \) we get

5463/e/p
\[ \rho^2 + m^2 \frac{c^2}{\epsilon^2} = \frac{E^2}{c^2} \quad \text{or} \quad E = \pm c \sqrt{\rho^2 + m^2 \frac{c^2}{\epsilon^2}} \pm c \sqrt{W} \] (2.22b)

and

\[ \psi = \chi \left( \xi \left( \rho^2 - \frac{c^2}{\epsilon^2} - E \right) \right) \]

\[ \chi = \left( \xi \left( \rho^2 - \frac{c^2}{\epsilon^2} - E \right) \right) \]

(2.21b) gives us the desired space-time dependence of \( \psi \) and \( \chi \).

We observe in (2.22b) the most important fact that the energy can be positive and negative. Since we are considering here free electrons, the energy value \( E = \pm c \sqrt{W} \) would mean that the electron has a negative kinetic energy. Such states do not make sense from the physical point of view. The energy spectrum of a free electron, therefore, has the form:

\[ \begin{array}{c}
\pm mc^2 \\
- \infty \quad \cdots \quad + \infty
\end{array} \]

with energy states from \( +mc^2 \) to \( +\infty \) and from \( -mc^2 \) to \( -\infty \). It should be emphasized that the negative values of \( E = \sqrt{\rho^2 + m^2 \frac{c^2}{\epsilon^2}} \) also appear in a purely classical theory of the electron. In a classical theory, however, the gap between \( +mc^2 \) and \( -mc^2 \) could prevent any transition from the positive region into the negative one since transitions must be continuous. Hence we may assume in classical theory that the states with negative energy do not exist at all.

In quantum theory such non-continuous transitions can be induced by perturbations; for example, a radiative transition with the emission of two quanta could take place. Hence it is not possible to exclude the solutions with negative energy. Later on we will see that they correspond to states of the antiparticle: the positron.
The amplitudes $u$ and $v$ are as yet unspecified except that we know that they have no space or time dependence. To obtain a relation between them, we must consider Dirac's equations themselves. We expect then, to find a relation between $u$ and $v$, just as in the electromagnetic case we also find $\xi_0$ and $\eta_0$ in terms of each other. For positive energy states we have $v$ large and $u$ small, hence we will take $u$ arbitrarily and find $v$ in terms of $u$. For negative energy states, the situation is reversed, and we will accordingly take $v$ arbitrarily and find $u$ in terms of $v$.

Case a: $E = +cW$

Inserting (2.21b) into one of Dirac's equations we obtain:

$$\frac{i}{\hbar} \left( \xi \cdot p \right) v = \left( \frac{iE}{\hbar c} + i \kappa \right) v$$

In the above, $p$ and $E$ are a vector and a number—not operators.

Hence

$$v = \frac{\xi \cdot p}{E/c + \sqrt{W}} \ n = \frac{\xi \cdot p}{W + \mu} v$$

where $\mu = m c (2.23)$

Since $W = \sqrt{p^2 + m^2 c^2}$, at low energies $W + u \approx 2mc$ which shows clearly that $v$ is small compared to $u$.

Case b: $E = -cW$

Using the other of the two Dirac equations (2.5) we get

$$\frac{i}{\hbar} \left( \xi \cdot p \right) u = \left( \frac{iE}{\hbar c} - i \kappa \right) u$$

hence

$$u = \frac{-\xi \cdot p}{W + \mu} v$$

(2.24)
and we see—as we have anticipated before—that for the negative energy states \( \nu \) is "big" and \( \nu \) is "small".

We would like now to write down the general solution of Dirac's equation. Since the quantity involved in Dirac's equation is a four-spinor with four components, we will expect Dirac's equation to have four linearly independent solutions—a general solution can then be obtained by a linear combination of the four independent solutions. We obtain two independent solutions for \( E = cW \) and two for \( E = -cW \). In the first case we have taken \( \nu \) to be arbitrary and the simplest two forms for \( \nu \) are

\[
\nu_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \nu_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}
\]

and \( \nu_3 \) and \( \nu_4 \) follow from (2.23). No more independent solutions can be obtained since any other \( \nu \) can be written as \( \nu = a_1 \nu_1 + a_2 \nu_2 \).

For the case \( E = -cW \), we have agreed to take \( \nu \) arbitrary, hence again the simplest forms are

\[
\nu_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \nu_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}
\]

and \( \nu_5 \) and \( \nu_6 \) follow from (2.24). We can summarize our results in the following table, where in the first row we put the \( \nu_+ \) component, in the second \( \nu_- \), in the third \( \nu_+ \), and in the fourth \( \nu_- \) of the four-spinor. The first two columns correspond to \( E = +cW \), the second two columns to \( E = -cW \).
\[
\Psi = (\frac{W + \mu}{2W})^{1/2} \begin{pmatrix}
\gamma^0 & \gamma^1 & \gamma^2 & \gamma^3 \\
1 & 0 & -\frac{p_3}{W+\mu} & -\frac{p_1}{W+\mu} \\
0 & 1 & -\frac{p_1}{W+\mu} & \frac{p_3}{W+\mu} \\
\frac{p_2}{W+\mu} & \frac{p_1}{W+\mu} & 1 & 0 \\
\end{pmatrix}
\] (2.25)

where \( p_\pm = p_x \pm i p_y \). In order that the \( \gamma^\alpha \) be normalized, we have divided them by the square root of the sum of the squares of their components. This turns out to be \( \left(\frac{2m}{W + \mu}\right)^{1/2} \) for every \( \gamma^\alpha \).

The four fundamental solutions of the Dirac equations can then be written in the form:

\[
\begin{pmatrix}
\gamma^{(1)} \\
\gamma^{(0)} \\
\end{pmatrix} = \gamma^\alpha \frac{\frac{2}{\sqrt{V}} (P^\alpha - E)}{\sqrt{V}}
\]

\[E = +cW \quad \text{for } i = 1, 2\]

\[E = -cW \quad \text{for } i = 3, 4 \] (2.26)

Here the factor \( V^{-1/2} \) is added in order to normalize the function.

\( V \) is a large volume in which we include all our problems and over which we have to integrate the normalization integrals:

\[
\left( \Psi^{(i)}, \overline{\Psi}^{(i)} \right) = \left\{ \gamma^\alpha \right\} \int \frac{e^{i \frac{\hat{p} \cdot \hat{x} - \hat{E} t}}}{\sqrt{V}} \left| d\sigma \right| d\gamma d\hat{t} = 1
\]

The general solution for a given energy \( E \) and momentum \( \hat{p} \) is

\[
\Psi = \sum_i A_i \overline{\Psi}^i
\]

where the sum is extended over \( i = 1, 2 \) for positive \( E \), over \( i = 3, 4 \) for negative \( E \).
Now that we know the form of the general solution, we would like to be able to make statements about the spinors $\psi$ and $\varphi$, similar to the statement that $\vec{E}$, $\vec{B}$, and the direction of propagation of electromagnetic waves are all orthogonal. Let us take for example $\varphi_0$ and see what are the angles between $\vec{y}$, $\vec{p}$, and $\vec{r}$. For $\varphi_0$, $\varphi$ is along the $z$-axis and the angle $\theta$ of the orientation of $\varphi$ from our formula is

$$\tan \frac{\theta}{2} = \frac{|c|}{|c+1|} = \frac{|p_+|}{p_z} = \frac{\sqrt{p_x^2 + p_y^2}}{p_z} = \tan \delta$$

where $\delta$ is the angle $\vec{p}$ is making with the $z$-axis. The angle $\varphi$ of the spinor orientation is equal to the phase of $p_+$ minus the phase of $p_z$ which is the same as the phase of $\vec{p}$. Thus we see that $\vec{p}$, $\vec{y}$, $\vec{p}$ are all in the same plane with the angle between $\vec{y}$ and $\vec{p}$, being bisected by $\vec{p}$:

Hence, $\vec{y}$, $\vec{p}$, and $\vec{p}$ are all lined up when $\delta = 0$, i.e. $p = p_z$ $\vec{y}$ and $\vec{p}$ are antiparallel and perpendicular to $\vec{p}$ when $\delta = \frac{\pi}{2}$, i.e. $p_z = 0$. Thus the question: "what is the direction of the spin?" cannot really be answered in general. What we mean, therefore, whenever we talk of the direction of the spin is its direction in that frame of reference in which the electron is at rest. In that case $\vec{p} = 0$, $\vec{y} = 0$, and the direction of the spin is given by $\vec{y}$. $\vec{y}$ is, in that case, well defined, while for an electron in motion, we need the eigenfunctions of the total angular momentum, neither $\psi$
nor $\xi$. We can write our general solution (2.26) in a somewhat different fashion as

$$\nabla = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{pmatrix} + \frac{i}{\sqrt{\lambda}} \begin{pmatrix} \frac{\hat{p}\cdot \hat{r}}{\lambda} - \hat{r} \cdot \hat{E} \end{pmatrix}$$

(2.27)

and we introduce the following set of fundamental four-spinors:

$$\eta^{(i)}(o) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \eta^{(a)}(o) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \eta^{(3)}(o) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \eta^{(4)}(o) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

We write the $\eta^{(i)}$ with an argument equal to $0$ since they are the same as the $\eta^{(i)}$ of (2.25) if we write the latter as functions of $\hat{p}$, $\eta^{(i)}(\hat{p})$ and let $\hat{p} = 0$. We then can connect $\eta^{(i)}(\hat{p})$ with $\eta^{(i)}(0)$ by some matrix $D(\hat{p})$ which is a function of $\hat{p}$:

$$\eta^{(i)}(\hat{p}) = D \eta^{(i)}(0)$$

(2.28)

where (1) specifies the "basis" four-spinor and $k$, $e$, specify the component of that spinor. It follows from its definition by (2.28) that $D(\hat{p})$ is the matrix obtained by writing the table (2.25) as a matrix:

$$D = \left( \frac{W + M}{2W} \right)^{1/2} \begin{pmatrix} 1 & 0 & -\frac{P_3}{W + M} & -\frac{P_0}{W + M} \\ 0 & 1 & -\frac{P_0}{W + M} & \frac{P_2}{W + M} \\ \frac{P_2}{W + M} & \frac{P_0}{W + M} & 1 & 0 \\ \frac{P_0}{W + M} & -\frac{P_2}{W + M} & 0 & 1 \end{pmatrix}$$
or

\[ D(\vec{\omega}) = \left( \frac{W+\mu}{2W} \right)^{1/2} \left( 1 - \frac{\beta(\vec{\omega}, \vec{p})}{W+\mu} \right) \]  

(2.29)

We write \( D \) as in (2.29) in terms of \( \vec{\omega} \) and \( \beta \), by simply realizing that this is how the table (2.25) was obtained in the first place.

At this point we again remind the reader that \( \vec{p} \) in \( D(\vec{p}) \) is not an operator. The matrix \( D(\vec{p}) \) will be very useful in interpreting the solutions of Dirac equation as we see below.

**Velocity of the Dirac Electron**

To find the velocity operator, we need the time derivative of the position vector operator. The time derivative of any operator \( O \) is found as follows:

\[
\frac{d}{dt} \{ \psi, O \psi \} = \{ \frac{\partial}{\partial t} \psi, \psi \} + \{ \psi, \frac{\partial O}{\partial t} \psi \} + \{ \psi, O \frac{\partial \psi}{\partial t} \}
\]

\[
= \frac{i}{\hbar} \left( \{ \psi, O \mathcal{E} \psi \} - \{ \mathcal{E} \psi, O \psi \} \right) + \{ \psi, \frac{\partial O}{\partial t} \psi \}
\]

\[
= \frac{i}{\hbar} \left( \{ \psi, O \mathcal{E} \psi \} - \{ \psi, O \mathcal{E} \psi \} \right) + \{ \psi, \frac{\partial O}{\partial t} \psi \}
\]
Hence if the operator $O$ does not involve time explicitly (i.e. $\frac{d}{dt} = 0$) we obtain
\[
\frac{d}{dt} \{ \varphi, O \varphi \} = \frac{1}{i\hbar} \{ \varphi, [O, E] \varphi \}
\]
or
\[
\frac{d}{dt} O = \frac{2}{\hbar} [E, O] = \frac{2}{\hbar} [H, O]
\]  \hspace{1cm} (2.30)

if the Hamiltonian $H$ is given by the total energy of the system $E$.

From eq. (2.7) we have for the Hamiltonian

\[
H = c \alpha \cdot \vec{p} + \beta m c^2
\]

hence we find for the velocity operator:
\[
\frac{d}{dt} (\vec{v}) = \frac{2}{\hbar} \left[ c \alpha \cdot \vec{p} + \beta m c^2, \vec{r} \right] = \frac{2}{\hbar} \left[ c \alpha \cdot \vec{p}, \vec{r} \right] = \frac{c^2}{\hbar} \sum_k \left[ \alpha_k \rho_k, x_k \vec{u}_k \right] = \frac{c^2}{\hbar} \sum_k \frac{\hbar \alpha_k}{\hbar \alpha_k} \vec{u}_k = c \vec{\alpha}
\]  \hspace{1cm} (2.31)

where we have made use of the fact that $\beta$ and $\alpha$ commute with $\vec{r}$

and that $[\rho_k, x_k] = \frac{\hbar}{\alpha_k}$. This velocity operator is not a constant of motion because $[H, c \alpha] \neq 0$. The square of $c \alpha$, however, is a constant of motion since

\[
(c \alpha)^2 = c^2 (\alpha_x^2 + \alpha_y^2 + \alpha_z^2) = 3c^2
\]

Thus we find that the instantaneous velocity of the electron is such that its square is $3c^2$. Since the average velocity is, of course, different (we will compute it presently), we must conclude that the electron moves in some curly fashion such that the instantaneous value of each component of its velocity is $\pm c$. Only one of
the three components can be made diagonal then the other two are not sharply defined. The eigenvalues of the one that we choose to be diagonal are then \(-c\) and \(-c\). It can be shown that this motion, called "Zitterbewegung" gives rise to the magnetic moment of the electron. (See K. Huang, "On the Zitterbewegung of the Dirac Electron", Amer. J. Phys., 20, 479-84 (Nov. 1952))

Let us now compute the average, or expectation value, of the velocity. For the simplicity we will assume that the solution \(\Psi\) has only one component \(\psi^{(i)}\) -- the generalization to four components is quite straightforward. Then

\[
\bar{v} = \left\{ \psi, c \frac{\alpha}{\sqrt{V}} \psi \right\} \frac{d\tau}{}\frac{\sqrt{V}}{V}
\]

\[
= \int \left[ \frac{1}{\sqrt{V}} e^{\frac{i}{\hbar} (\hat{p} \cdot \vec{x} - \vec{E} \cdot \vec{t})} \psi^{(i)}(\vec{p}) \right]^* \left[ c \frac{\alpha}{\sqrt{V}} e^{\frac{i}{\hbar} (\hat{p} \cdot \vec{x} - \vec{E} \cdot \vec{t})} \psi^{(i)}(\vec{p}) \right] d\tau
\]

\[
= \psi^{(i)}(\vec{p})^* c \frac{\alpha}{\sqrt{V}} \psi^{(i)}(\vec{p}) = \left( D(\vec{p}) \psi^{(i)}(\vec{p}) \right)^* c \frac{\alpha}{\sqrt{V}} \left( D(\vec{p}) \psi^{(i)}(\vec{p}) \right)
\]

\[
= (\psi^{(i)}(\vec{p}))^* D(\vec{p})^* c \frac{\alpha}{\sqrt{V}} D(\vec{p}) \psi^{(i)}(\vec{p}) = \left( D(\vec{p})^* c \frac{\alpha}{\sqrt{V}} D(\vec{p}) \right)_{\vec{p}}
\]

Since \(D(\vec{p}) = \left( \frac{w+\mu}{2w} \right)^{1/2} \left( 1 - \frac{\vec{p} \cdot \vec{\alpha}}{w+\mu} \right)\) it follows that

\[
D^* = \left( \frac{w+\mu}{2w} \right)^{1/2} \left( 1 - \frac{\vec{p} \cdot \vec{\alpha}}{w+\mu} \right) = \left( \frac{w+\mu}{2w} \right)^{1/2} \left( 1 - \frac{\vec{p} \cdot \vec{\alpha}}{w+\mu} \right)
\]

\[
= \left( \frac{w+\mu}{2w} \right)^{1/2} \left( 1 + \frac{\vec{p} \cdot \vec{\alpha}}{w+\mu} \right) = \left( \frac{w+\mu}{2w} \right)^{1/2} \left( 1 + \frac{\vec{p} \cdot \vec{\alpha}}{w+\mu} \right)
\]

5463/e/p
Since $\beta, \check{\alpha}, \check{p}$ are all selfadjoint and $\check{\alpha} \beta' = -\beta \check{\alpha}$, hence

$$ (D^+ \check{\alpha}, \check{\alpha})_{ii} = \frac{W+\mu}{2W} \left( (1 + \frac{\beta \check{\alpha} \cdot \check{p}}{W+\mu}) \check{\alpha} \check{\alpha} \right)_{ii} \left( 1 - \frac{\beta \check{\alpha} \cdot \check{p}}{W+\mu} \right)_{ii} \check{\alpha} \check{\alpha} $$

Since $\check{\alpha}_K = \begin{pmatrix} 0 & 0 \\ 0 & \check{\alpha}_K \end{pmatrix}$ has no diagonal elements, it follows that any product of the $\check{\alpha}$'s involving an odd number of factors will have no diagonal elements. We are therefore left with:

$$ \frac{W+\mu}{2W} \left( \frac{\beta \check{\alpha} \cdot \check{p}}{W+\mu} \check{\alpha} \check{\alpha} - \frac{\check{\alpha} \beta \check{\alpha} \cdot \check{p}}{W+\mu} \right)_{ii} \check{\alpha} \check{\alpha} = \frac{\check{p}}{W} \beta_{ii} $$

$$ = \begin{cases} \frac{\check{p}}{W}, & i' = 1, 2 \\ -\frac{\check{p}}{W}, & i' = 3, 4 \end{cases} $$

that is to say for the positive energy states

$$ \check{p}_+ = \frac{\check{p}}{W} \quad (2.32a) $$

and for the negative energy states

$$ \check{p}_- = -\frac{\check{p}}{W} \quad (2.32b) $$

which will perhaps look more familiar if we write it for $\check{p}$ in terms of $\check{\nu}$:

Positive States: $\check{p} = \frac{m \check{\nu}}{\sqrt{1 - \frac{\check{\nu}^2}{c^2}}}$, $E = \frac{mc^2}{\sqrt{1 - \frac{\check{\nu}^2}{c^2}}}$

Negative States: $\check{p} = -\frac{m \check{\nu}}{\sqrt{1 - \frac{\check{\nu}^2}{c^2}}}$, $E = -\frac{mc^2}{\sqrt{1 - \frac{\check{\nu}^2}{c^2}}}$
Thus for the positive energy states we have recovered the normal relations between velocity, momentum, mass and energy. In this relationship, however, we have to take the average velocity and not the eigenvalues of the velocity operator $\alpha^2$. For negative energy states we find that the average velocity and momentum are in opposing directions and an increase in velocity causes a decrease in the energy—a peculiar behaviour which would seem to indicate that we are dealing with a particle of negative mass.

We next show, as we have promised we would, that the average value of $\beta$ is $\sqrt{-\frac{v^2}{c^2}}$

$$\bar{\beta} = \left\{ \beta^i (\vec{p}) \right\} = \left( \beta^i \right)_{ii} = \frac{W+\mu}{2W} \left( \beta^i \left( \beta^i \right) \right)_{ii}$$

$$= \frac{W+\mu}{2W} \left( \beta^i - \frac{\beta^i \beta^j \beta^i}{(W+\mu)^2} \right)_{ii} = \frac{W+\mu}{2W} \left( 1 - \frac{\beta^2}{(W+\mu)^2} \right)_{ii}$$

$$= \frac{(W+\mu)^2}{2W(W+\mu)} \beta^i = \frac{\mu}{W} \beta^i = \int \frac{\mu}{W} \beta^i = \begin{cases} \frac{\mu}{W} & i = 1,2 \\ -\frac{\mu}{W} & i = 3,4 \end{cases}$$

but $\frac{\mu}{W} = \frac{1}{\sqrt{1+\frac{\beta^2}{c^2}}} = \left( 1 + \frac{\frac{\beta^2}{c^2}}{1-\frac{\beta^2}{c^2}} \right)^{-\frac{1}{2}} = \sqrt{1-\frac{v^2}{c^2}}$ hence

$$\bar{\beta} = \begin{cases} \sqrt{1-\frac{v^2}{c^2}} & i = 1,2 \\ -\sqrt{1-\frac{v^2}{c^2}} & i = 3,4 \end{cases}$$ (2.33)
To finish this chapter we note that since $\overline{\sigma} = \frac{c\overline{\rho}}{W} \beta; \iota$

we can rewrite the matrix $D(\overline{\rho})$ in terms of $\overline{\sigma}$:

$$D = \left(\frac{W+\mu}{2W}\right)^{1/2} \left(1 - \frac{\beta \overline{\sigma} \overline{\rho}}{W+\mu}\right) = \left(\frac{W+\mu}{2W}\right)^{1/2} \left(1 + \frac{\beta \overline{\sigma} \overline{\rho}}{W+\mu}\right)$$

$$= \left(\frac{W+\mu}{2W}\right)^{1/2} \left(1 + \frac{\overline{\sigma} \cdot \overline{\rho} \cdot W}{c(W+\mu)}\right)$$

(2.34)

As stated above this matrix is unitary, but is not hermitian.

The last property is due to the fact that the mean velocity in (2.34) is not a simple number but a (diagonal) matrix given by (2.32a) and (2.32b).

For the discussion of the Lorentz transformations in the next chapter it is important to reinterpret equation (2.34):

We regard $\overline{\sigma}$ as simple number given by $\frac{c \overline{\rho}}{W}$. The reason is that for the Lorentz transformations the velocity is the more fundamental quantity. In that way (2.34) gives a matrix $D(\overline{\sigma})$ which is now hermitian, but no longer unitary. A better justification for this redefinition can be found in the fact that the negative energy solutions $\gamma^{(i)}(\overline{\rho})$, with $i = 3,4$, describe a positron of positive energy and momentum $-\overline{\rho}$. Therefore $D(\overline{\sigma})$ generates as well electron as positron states of momentum $\overline{p}$ out of the spinors $\gamma^{(i)}(0)$ ($i = 1,2,3,4$).
Chapter III: The Dirac Equation Under the Lorentz Transformation

We have derived Dirac's equation by requiring that, when iterated, it should give Klein-Gordon's equation, since the latter expresses the relativistically correct relation between the energy and momentum of a free particle. However, we have not demonstrated as yet the relativistic invariance of Dirac's equation itself. To do this we introduce the $\gamma$-matrices defined by

$$\gamma_k = -i \sigma_k \quad k=1,2,3$$

$$\gamma_4 = \beta$$

A direct calculation will show that these $\gamma$'s anticommute and that $\gamma_\nu^2 = 1 \ (\nu = 1,2,3,4)$. Furthermore, we consider time as a fourth coordinate setting $x_4 = i ct$. Writing Dirac's equation as

$$\left[ \frac{i}{\hbar} \left( \partial \cdot \vec{A} \right) + \beta \mu - \frac{i \hbar}{c} \frac{\partial}{\partial t} \right] \Psi = 0$$

we obtain on multiplying by $\beta / \hbar$ from the left

$$\left[ \beta \frac{\partial}{\partial x^\nu} + \kappa + \frac{\beta}{4c} \beta \frac{\partial}{\partial x^4} \right] \Psi = 0$$

which becomes when we introduce the $\gamma$'s and $x_4 = i ct$

$$\left( \sum_{\nu=1}^{4} \gamma_\nu \frac{\partial}{\partial x^\nu} + \kappa \right) \Psi = 0 \quad (3.1)$$

We shall use the usual convention of greek letters for subscripts which run from 1 to 4, and we shall from now on omit the summation sign in front of expressions in which the greek subscript appears
twice. If (3.1) is relativistically invariant, then in a moving frame of reference $S'$ with coordinates $x'_1$:

$$
\left( \gamma'_\nu \frac{\partial}{\partial x'_\nu} + \kappa \right) \Psi' = 0 \quad (3.2)
$$

We do not put a prime on $\phi'_\nu$ since we claim that the Dirac equation should be the same in the moving system: hence, the moving observer will use the same matrices $\phi'_1 \ldots \phi'_4$ for his equations. The wave function $\Psi'$, however, is different in the moving system, since it describes the same electron in the system $S'$, where it has a different velocity. [It is formally possible to define the equation in the new system such that $\Psi$ is invariant and the $\phi'_\nu$'s are transformed. We prefer the other way of definition.] Note that the transformation properties of $\frac{\partial}{\partial x'_\nu}$ are known, since $X'_\nu$ must transform as the components of a four-vector. The transformation of the $\Psi$ can be expressed by

$$
\Psi' = S \Psi \quad (3.3)
$$

where $S$ is the Lorentz transformation operator. The structure of $S$ must be such as to make equations (3.1) and (3.2) compatible.

We recall that a solution of Dirac's equation was

$$
\Psi = \gamma \frac{i}{\sqrt{\nu}} e^{\frac{i}{\hbar} (\vec{p} \cdot \vec{r} - \nu t)}
$$

hence in the primed coordinate system

$$
\Psi' = \gamma' \frac{i}{\sqrt{\nu'}} e^{\frac{i}{\hbar} (\vec{p}' \cdot \vec{r}' - \nu' t')}
$$
Since \( \mathbf{p} \cdot \mathbf{r} \)-\( E t \) is simply the scalar product of the momentum four-vector with the position four-vector, it is relativistically invariant, i.e.,

\[
\mathbf{p} \cdot \mathbf{r} = \mathbf{p}' \cdot \mathbf{r}' - E' t' = p v x_\nu
\]

Furthermore, the relation between \( V \) and \( V' \) is that given by the Lorentz contraction

\[
V' = \frac{V}{\sqrt{1 - \frac{v^2}{c^2}}}
\]

where \( v \) represents the relative velocity of the two coordinate systems. Hence from (3.3) we obtain for \( S' \):

\[
\gamma' = \frac{\gamma}{\sqrt{1 - \frac{v^2}{c^2}}} S \gamma
\]

Now comes the main point: Let us assume that in the unprimed system the electron is at rest. Then \( \gamma(1) = \gamma(1) (0) \). Let the primed system move with velocity \( \mathbf{v} \) with respect to the unprimed system. The average velocity of the electron in that system will then be \( -\mathbf{v} \) and therefore

\[
\gamma'(i) = \mathcal{D}(-\mathbf{v}) \gamma(i) (0)
\]

and therefore

\[
\mathcal{S} (\mathbf{v}) = \left( 1 - \frac{v^2}{c^2} \right)^{-\frac{1}{2}} \mathcal{D} (-\mathbf{v})
\]

At this point we introduce a few symbols that will permit us to write (3.4) more elegantly. We define \( \mathcal{S} \) as

\[
\mathcal{S} = \left( 1 - \frac{v^2}{c^2} \right)^{-\frac{1}{2}}
\]

from which the useful relation \( \frac{\sqrt{v^2 - 1}}{c} = \frac{1}{\mathcal{S}} \) follows.
Recalling that $W = \frac{|E|}{c}$ and $\mu = mc$, we have in terms of $\xi$

$$W = \mu \xi \quad \frac{W + \mu}{2W} = \frac{1 + \xi}{2\xi}$$

In addition, we introduce the symbol $\alpha_\nu$ for the component of $\mathbf{a}$

in the direction of $\mathbf{v}$, i.e. $\mathbf{a} \cdot \mathbf{v} = a_\nu |\mathbf{v}|$. With these changes and

with $D(\mathbf{v})$ given by Eq. (2.34), Eq. (3.4) becomes:

$$S(\mathbf{v}) = \left(1 - \frac{\mathbf{v}^2}{c^2}\right)^{-\frac{1}{2}} \left(\frac{W + \mu}{2W}\right)^{\frac{1}{2}} \left(1 - \alpha_\nu \frac{\mathbf{v} \cdot \mathbf{v}}{c^2} \frac{W}{W + \mu}\right)$$

$$= \sqrt{\frac{\xi + 1}{2}} - \alpha_\nu \sqrt{\frac{\xi - 1}{2}}$$

or, introducing the angle $\chi$ by $\cosh \chi = \frac{\xi}{2}$

$$S(\mathbf{v}) = \cosh \frac{\chi}{2} - \alpha_\nu \sin \frac{\chi}{2} = e^{-\alpha_\nu \frac{\chi}{2}} \quad (3.7)$$

The form of (3.7) seems to imply that the Lorentz transformation

operator $S(v)$ corresponds to a rotation of a spinor through an

imaginary angle $\chi = \tanh^{-1} \frac{\xi}{1}$. (Note: The operator $S(\mathbf{v})$ as

defined here corresponds to a pure Lorentz transformation excluding

and spytial rotations.)

We have still not shown that the operator $S$ actually leaves

Dirac's equation invariant. This can be done in two ways: by

showing that the operator $S$ has the same form or transforming

from any coordinate system to any other coordinate system or by

directly testing with it Dirac's equation.

a) Consider the three systems I, II, and III. Let the

electron be at rest in system I, let II move with velocity $\mathbf{v}_1$
relative to I, and let III move with velocity $\vec{v}_2$ relative to II.

For simplicity let us assume that $\vec{v}_1$ and $\vec{v}_2$ are in the same direction (then $\chi_v$ will be the same for $S(v_1)$ and $S(v_2)$). Let the wave function of the electron be $\Psi_I$ in system I, $\Psi_{\text{II}}$ in system II, and $\Psi_{\text{III}}$ in system III. We know that

$$\Psi_{\text{II}} = S(\vec{v}_1) \Psi_I \quad (3.8a)$$

and we claim that

$$\Psi_{\text{III}} = S(\vec{v}_3) \Psi_{\text{II}} \quad (3.8b)$$

We also know that

$$\Psi_{\text{III}} = S(\vec{v}_3) \Psi_I \quad (3.8c)$$

where $\vec{v}_3$ is the velocity of system III with respect to I. Therefore (3.8b) will be true if we can show that

$$S(\vec{v}_3) = S(\vec{v}_2) S(\vec{v}_1)$$

which in terms of (3.7) means

$$e^{-\alpha v_3 x_3} = e^{-\alpha v_2 x_2} e^{-\alpha v_1 x_1}$$

This is to say we must show that

$$\chi_3 = \chi_1 + \chi_2$$

We obtain the required result by making use of the recipe for adding velocities:

$$v_3 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}} \quad (3.9)$$
Since \( \tanh \chi_1 = \frac{\nu_1}{\nu_0} \) (3.9) in terms of \( \chi_1 \) becomes
\[
\tanh \chi_3 = \frac{\tanh \chi_1 + \tanh \chi_2}{1 + \tanh \chi_1 \cdot \tanh \chi_2} = \tanh (\chi_1 + \chi_2)
\]
and therefore \( \chi_1 + \chi_2 = \chi_3 \) as required.

b) We now prove that \( S \) leaves Dirac's equations invariant by a direct test.

In that frame of reference where the electron wave function is given by \( S\Psi \), Dirac's equation reads:
\[
\left( \gamma\nu \frac{\partial}{\partial \tau\nu} + \kappa \right) S\Psi = 0 \tag{3.10a}
\]
Operating on above with \( S^{-1} \) from the left, we obtain
\[
\left( S^{-1} \gamma\nu S \frac{\partial}{\partial \tau\nu} + \kappa \right) \Psi = 0 \tag{3.10b}
\]
where we have made use of the fact that \( S \) commutes with \( \frac{\partial}{\partial \tau\nu} \) since \( S \) commutes with \( \frac{\partial}{\partial \tau\nu} \). However, in a frame of reference where the electron's wave function is given by \( \Psi \), Dirac's equation is supposed to read
\[
\left( \gamma\nu \frac{\partial}{\partial \tau\nu} + \kappa \right) \Psi = 0 \tag{3.10c}
\]
In order that (3.10b) and (3.10c) be identical, we must have
\[
\left( S^{-1} \gamma\nu S \right) \frac{\partial}{\partial \tau\nu} = \gamma\nu \frac{\partial}{\partial \tau\nu} \tag{3.11}
\]
The right-hand side of (3.11) has the form of a scalar product of two four-vectors and therefore will be an invariant if the \( \gamma\nu \)'s are indeed components of a four-vector. If they are, then
\((S^{-1} y, S)\) must represent the transformed components of that four-vector

\[
S^{-1} y, S = \frac{y_1 + i y_4}{\sqrt{1 - \frac{v^2}{c^2}}} = y_1 \cosh \chi + i y_4 \sinh \chi
\]

\[
S^{-1} y_2 S = y_2
\]

\[
S^{-1} y_3 S = y_3
\]

\[
S^{-1} y_4 S = \frac{y_4 - i y_1}{\sqrt{1 - \frac{v^2}{c^2}}} = y_4 \cosh \chi - i y_1 \sinh \chi
\]

(3.12)

assuming the velocity \(v\) to be in the \(x\)-direction. But

\[
S^{-1} y, S = \left( \cosh \frac{x}{2} + \alpha, \sinh \frac{x}{2} \right) \beta \left( \cosh \frac{x}{2} - \alpha, \sinh \frac{x}{2} \right)
\]

\[
= \beta \left( \cosh^2 \frac{x}{2} + \sinh^2 \frac{x}{2} - 2 \alpha \sinh \frac{x}{2} \cosh \frac{x}{2} \right)
\]

\[
= \beta \left( \cosh \frac{x}{2} - \alpha, \sinh \frac{x}{2} \right) - \beta \alpha, 2 \sinh \frac{x}{2} \cosh \frac{x}{2}
\]

\[
= y_4 \cosh \chi - i y_1 \sinh \chi
\]

and for \(k = 2\) or \(3\):

\[
S^{-1} y_k S = \left( \cosh \frac{x}{2} + \alpha, \sinh \frac{x}{2} \right) \beta \alpha_k \left( \cosh \frac{x}{2} - \alpha, \sinh \frac{x}{2} \right)
\]

\[
= \beta \alpha_k \left( \cosh^2 \frac{x}{2} - \sinh^2 \frac{x}{2} \right) = y_k
\]
Hence the $\gamma$-matrices do transform as the components of a four-vector and (3.11) is automatically satisfied.

We now summarize briefly the transformation properties of the Dirac four-spinor $\Psi$:

**Space rotations** through an angle $\Theta$ around an axis $\hat{A}$: *)

$$R = e^{\frac{i}{2} \sigma_A \frac{\gamma}{2}} = \cos \frac{\gamma}{2} + i \sigma_A \sin \frac{\gamma}{2}$$

where $\sigma_A$ is the component of $\vec{\sigma}$ in the direction $A$. In the case of a two-component spinor the components of $\vec{\sigma}$ are the usual 2 x 2 Pauli matrices. In the case of the four-spinor, $\vec{\sigma}$ stands for the extended Pauli matrices defined by

$$\text{extended } \sigma_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \text{ a 4 x 4 matrix}$$

(In our notation we shall not distinguish between the usual and extended $\sigma$, it being always clear from the context whether we need a 2 x 2 or 4 x 4 matrix.)

**Pure Lorentz transformation** to a system moving with constant velocity $\vec{v}$:

$$S = e^{-\lambda_v \frac{\gamma}{2}} = \cosh \frac{\gamma}{2} - \lambda_v \sinh \frac{\gamma}{2}$$

where $\lambda_v$ is the component of $\vec{\lambda}$ in the direction of $\vec{v}$ and $\gamma$ is given by $\gamma = \tanh^{-1} \left( \frac{v}{c} \right)$.

We note that $R$ is unitary, as it must be since it represents a spatial rotation, but $R$ is not hermitian. On the other hand, $S$ is hermitian but not unitary. The non-unitarity of $S$ is the result of normalization of the wave function not being preserved.

*) Replace on page 35 $\mathcal{O}$ by $-\mathcal{O}$, because we now consider a transformation of the frame and not of the spinor.
since the volume of the box in which we are working is contracted in the moving system.

From the above properties we can deduce that given any two four-spinors \( \Psi \) and \( \Phi \) the expression

\[
\{ \Psi, \beta \Phi^* \}
\]

is an invariant.

**Proof:** Replace in (3.13) \( \Psi \) and \( \Phi \) by \( S \Psi \) and \( S \Phi \):

\[
\{ S \Psi, \beta S \Phi \} = \{ S \Psi, S \beta S \Phi \} = \{ \Psi, \beta S^{-1} S \Phi \} = \{ \Psi, \beta \Phi \}
\]

where we make use of \( S \) being hermitian and the obvious relation \( S \beta = \beta S^{-1} \). Now replace in (3.13) \( \Psi \) and \( \Phi \) by \( R \Psi \) and \( R \Phi \):

\[
\{ R \Psi, \beta R \Phi \} = \{ R \Psi, R^{-1} \beta R \Phi \} = \{ \Psi, \beta R^{-1} R \Phi \} = \{ \Psi, \beta \Phi \}
\]

where we make use of the relation \( \beta R = R \beta \), and the fact that the hermitian conjugate of \( R \) is \( R^{-1} \). In particular, if we define the adjoint wave function by \( \Psi^+_K = \beta \Psi_K^* \) we have

\[
\{ \Psi, \beta \Psi \} = \{ \beta \Psi^*, \Psi \} = \sum_K \beta \Psi_K^* \Psi_K = \sum_K \Psi_K^+ \Psi_K
\]

and in virtue of (3.13) \( \sum_K \Psi_K^+ \Psi_K \) is an invariant.

The expression \( \{ \Psi, \alpha_K \Psi \} \) is not an invariant. We can, however, show that it represents the space part of a four-vector given by

\[
j_K = \epsilon_{\alpha} \{ \Psi, \alpha_K \Psi \}
\]

\( k = 1, 2, 3 \)
whose fourth (time-) component is the charge density

\[ j^4 = i c e \{ \Psi, \bar{\Psi} \} = i c \mathcal{P} \]

**Proof:**

We assume the transformation to be \( S(\tilde{\alpha}) = S(|v| \tilde{u}_x) \). Then

\[ S \alpha_1 S = \alpha_1 S S = \alpha_1 \cosh \gamma - \sinh \gamma = \frac{\alpha_1 - \frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} \]

\[ S \alpha_2 S = \alpha_2 S S' = \alpha_2 \]

\[ S \alpha_3 S = \alpha_3 S S' = \alpha_3 \]

\[ SS' = \cosh \gamma - \alpha_1 \sinh \gamma = \frac{1 - \alpha_1 \frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} \]

With these preliminary results established we have

\[ j^1' = e c \{ S \Psi, \alpha_1 S \bar{\Psi} \} = e c \{ \Psi, S \alpha_1 S \bar{\Psi} \} \]

\[ = \frac{e c}{\sqrt{1 - \frac{v^2}{c^2}}} \left[ \{ \Psi, \alpha_1 \bar{\Psi} \} - \frac{v}{c} \{ \Psi, \bar{\Psi} \} \right] = \frac{j^1 - \mathcal{P} \mathcal{F}}{\sqrt{1 - \frac{v^2}{c^2}}} \]

\[ j^2' = e c \{ S \Psi, \alpha_2 S \bar{\Psi} \} = e c \{ \bar{\Psi}, S \alpha_2 S \bar{\Psi} \} = e c \{ \bar{\Psi}, \alpha_2 \bar{\Psi} \} = j^2 \]

\[ S' = e \{ S \bar{\Psi}, S \bar{\Psi} \} = e \{ \bar{\Psi}, S S \bar{\Psi} \} = \]

\[ = \frac{e}{\sqrt{1 - \frac{v^2}{c^2}}} \left[ \{ \bar{\Psi}, \bar{\Psi} \} - \frac{v}{c} \{ \bar{\Psi}, \alpha \bar{\Psi} \} \right] = \frac{S - \mathcal{P}}{\sqrt{1 - \frac{v^2}{c^2}}} \]
which proves that the components \( j_y \) transform as components of a four-vector should.

From this we deduce that

\[
\frac{\partial}{\partial x'} j_y = \text{invariant}
\]

In fact, \( \frac{\partial}{\partial x'} j_y = 0 \), which, when expanded gives us the continuity equation \( \frac{\partial}{\partial t} \mathcal{S} = \text{div} \mathbf{\jmath} \).

Comparison of the Dirac Spinor and the Maxwell Vectors under
the Lorentz Transformation.

Let

\[
\mathcal{S} = e^{-i \frac{x}{2}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cosh \frac{x}{2} & -\sinh \frac{x}{2} \\ \sinh \frac{x}{2} & \cosh \frac{x}{2} \end{pmatrix}
\]

With

\[
\Psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \quad \text{and} \quad \mathcal{S} \Psi = \Psi' = \begin{pmatrix} \psi'_+ \\ \psi'_- \end{pmatrix}
\]

get for the equations for each of the components of \( \Psi' \) the following:

\[
\psi'_+ = \psi_+ \cosh \frac{x}{2} - \psi_- \sinh \frac{x}{2}
\]

\[
\psi'_- = \psi_+ \cosh \frac{x}{2} + \psi_- \sinh \frac{x}{2}
\]

or since

\[
\cos \frac{x}{2} = \sqrt{\frac{1 + \frac{1}{c^2}}{2}}, \quad \sinh \frac{x}{2} = \sqrt{\frac{1 - \frac{1}{c^2}}{2}} \quad \text{and} \quad \mathcal{S} = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}}
\]

\[
\psi'_+ = \frac{(\sqrt{1 - \frac{v^2}{c^2}} + 1)^{\frac{1}{2}} \psi_+ - (1 - \sqrt{1 - \frac{v^2}{c^2}})^{\frac{1}{2}} \psi_-}{\sqrt{2} (1 - \frac{v^2}{c^2})^{\frac{1}{4}}}
\]  \( (3.14) \)

\[
\psi'_- = \frac{(\sqrt{1 - \frac{v^2}{c^2}} + 1)^{\frac{1}{2}} \psi_+ - (1 - \sqrt{1 - \frac{v^2}{c^2}})^{\frac{1}{2}} \psi_-}{\sqrt{2} (1 - \frac{v^2}{c^2})^{\frac{1}{4}}}
\]  \( (3.15) \)
In the limit of $\frac{v}{c} \ll 1$ the equations become

$$\psi'_\pm = \frac{\psi'_\pm - \frac{v^2}{c^2} \psi'_\pm}{(1 - \frac{v^2}{c^2})^{1/4}}$$

and

$$\varphi'_\pm = \frac{\varphi'_\pm - \frac{v^2}{c^2} \varphi'_\pm}{(1 - \frac{v^2}{c^2})^{1/4}}$$

On the other hand the Maxwell vectors under the corresponding transformation become

$$E'_x = E_x$$
$$H'_x = H_x$$

$$E'_y = \frac{E_y - \frac{v^2}{c^2} H^2}{\sqrt{1 - \frac{v^2}{c^2}}}$$
$$H'_y = \frac{H_y + \frac{v^2}{c^2} E^2}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$E'_z = \frac{E_z + \frac{v^2}{c^2} H^2}{\sqrt{1 - \frac{v^2}{c^2}}}$$
$$H'_z = \frac{H_z - \frac{v^2}{c^2} E^2}{\sqrt{1 - \frac{v^2}{c^2}}}$$

(3.16)

Both (3.15) and (3.16) exhibit the characteristic mixing of the components and both involve the factor $(1 - \frac{v^2}{c^2})$ characteristic of the Lorentz transformation. Note, however, that for spinors this factor is raised to the $(-\frac{1}{4})$ power as compared to the $(-\frac{1}{2})$ power for the vectors of the electromagnetic field. This is a result of the fact that a vector is a "tensor of second rank" in the spinor space; hence the vector transforms with a matrix that is in some sense the square of the matrix $S$. 

\[\text{i463/e/p}\]
Chapter IV: Dirac's Equation in the Presence of an Electromagnetic Field

So far we have discussed only the motion of a particle in free space. When we enter the subject of a particle in an external field we find that relativistic quantum mechanics can deal only with a restricted class of external fields. Since the total Hamiltonian must contain the energy of interaction of the particle with the field we shall be able to solve only those problems which involve fields with known transformation properties under the Lorentz transformation. This limits us, almost exclusively, to the problem of an electron in an electromagnetic field.

The electromagnetic field will be specified if $\mathbf{E}$ and $\mathbf{B}$ are given. Alternately the field is specified by the electromagnetic potentials $\mathbf{A}$ and $\phi$ from which $\mathbf{E}$ and $\mathbf{B}$ follow uniquely as

$$\mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} - \text{grad} \phi$$
$$\mathbf{B} = \text{curl} \mathbf{A}$$

The classical interaction energy $H'$ of a point particle with this field is given by

$$H' = e \phi(\mathbf{r}) - e \frac{\mathbf{A}(\mathbf{r})}{c} \cdot \mathbf{A}(\mathbf{r})$$  (4.1)

where $\mathbf{A}$ and $\phi$ are evaluated at the position of the particle and $e$ stands for the particle's charge (e.g. for an electron $e$ is a negative number.). The electrodynamic potentials as we use them at this time are simply parameters specifying the external field.
This field is not quantized, and the potentials are functions of the position coordinates only, hence in particular they commute with all our $\alpha$'s, $\beta$'s, $\gamma$'s, etc. On the other hand the velocity appearing in (4.1) represents the velocity of the electron and therefore must be replaced by the velocity operator $\hat{\mathbf{v}}_c$. Combining (4.1) with the field free Hamiltonian we obtain

$$H = c(\hat{\mathbf{a}} \cdot \hat{\mathbf{p}}) + \beta \mu c + e \phi - e \hat{\mathbf{A}} \cdot \hat{\mathbf{x}}$$  \hspace{1cm} (4.2)$$

Dirac's equation can then be written as

$$c \left[ \hat{\mathbf{a}} \cdot (\hat{\mathbf{p}} - \frac{e}{c} \hat{\mathbf{A}}) + \beta \mu \right] \Psi = (E - e \phi) \Psi$$  \hspace{1cm} (4.3)$$

To see that (4.3) is still relativistically invariant we write it as

$$\left[ \gamma_\nu \left( \frac{\partial}{\partial x_\nu} - \frac{\gamma_5}{c} A_\nu \right) + \kappa \right] \Psi = 0$$  \hspace{1cm} (4.4)$$

where $A_\nu$ is the four-vector whose first three components are the components of $\hat{\mathbf{A}}$ and whose fourth operator is $A_4 = i \phi$.

With introduction of the operator

$$\pi_\nu = \rho_\nu - \frac{\alpha}{c} A_\nu$$

we can also write (4.4) in the form

$$\left( \gamma_\nu \pi_\nu - i \mu \right) \Psi = 0$$  \hspace{1cm} (4.4a)$$
This form is identical with the Dirac equation for a free particle with the exception that $P_\nu$ is replaced by the components of another four-vector $\mathbf{v}_\nu$. The proof of Lorentz invariance and of the validity of the transformation $\mathbf{v}' = S \mathbf{v}$ can be applied to (4.4a) in the same way as was done in the previous chapter, since the only fact used was that the magnitudes multiplied with the $\mathbf{v}_\nu$ are the components of a four-vector.

Eq. (4.3) is then the proper Dirac equation for an electron moving in an external field $A_\nu$. But we know that an electron in motion produces its own field $A_\nu^e$ since with an electron in motion there is associated a current density four-vector with components \[ ec \{\mathbf{v}, \mathbf{v}'\} \quad \text{and} \quad e \{\mathbf{v}, \mathbf{v}'\}. \] Hence the actual problem that we are faced with is the motion in a field $A_\nu + A_\nu^e$. We assume therefore that $A_\nu \gg A_\nu^e$ and proceed to solve (4.3). We would like to point out, however, that the assumption $A_\nu \gg A_\nu^e$ is not a very good one since sufficiently close to the position of the electron $A_\nu^e > A_\nu$, no matter how large $A_\nu$ is.

**The Charge Conjugate Solution**

So far no statement has been made as to the sign of $e$, and we should therefore expect that (4.3) should give us valid solutions for both particles with charge $e$ and $-e$. Suppose we have solved (4.3) for a particle with charge $e$ and have obtained the eigenfunction $\mathbf{v}_E'$ corresponding to the energy eigenvalue $E$. Suppose we then solve (4.3) for a particle with charge $-e$ and obtain the eigenfunction $\mathbf{v}_E$ corresponding to the same eigenvalue $E$. We can then write:
\[ c \left[ \hat{\alpha} \cdot (\mathbf{p} + \frac{e}{c} \mathbf{A}) + \beta \mu \right] \Psi_E = (E - e\phi) \Psi_E \quad (4.5a) \]
\[ c \left[ \hat{\alpha} \cdot (\mathbf{p} - \frac{e}{c} \mathbf{A}) + \beta \mu \right] \Phi_E = (E + e\phi) \Phi_E \quad (4.5b) \]

Note that in above equations \( E \) no longer represents an operator but simply stands for that eigenvalue to which \( \Psi_E \) and \( \Phi_E \) correspond. \( \hat{p} \), however, is still the momentum operator.

Let us take the complex conjugate of (4.5a) and multiply it by \(-1\):
\[ c \left[ \hat{\alpha}^* \cdot (\mathbf{p} + \frac{e}{c} \mathbf{A}) - \beta^* \mu \right] \Psi^*_E = (-E + e\phi) \Psi^*_E \quad (4.5c) \]
and let us replace \( E \) by \(-E\) in (4.5b):
\[ c \left[ \hat{\alpha} \cdot (\mathbf{p} + \frac{e}{c} \mathbf{A}) + \beta \mu \right] \Phi^{-}_E = (-E + e\phi) \Phi^{-}_E \quad (4.5d) \]

(4.5c) and (4.5d) are still not quite the same, but suppose that a matrix \( C \) exists such that
\[ C^{-1} \hat{\alpha} C = \hat{\alpha}^* \quad \text{and} \quad C^{-1} \beta C = -\beta^* \quad (4.6) \]

Then (4.5c) becomes
\[ c \left[ C^{-1} \hat{\alpha} C \cdot (\mathbf{p} + \frac{e}{c} \mathbf{A}) + C^{-1} \beta C \mu \right] C \Psi^*_E = (-E + e\phi) C \Psi^*_E \]
or, multiplying by \( C \) from the left:
\[ c \left[ \hat{\alpha} \cdot (\mathbf{p} + \frac{e}{c} \mathbf{A}) + \beta \mu \right] C \Psi^*_E = (-E + e\phi) C \Psi^*_E \quad (4.5c) \]

\( C \) has been defined as an operation on \( \hat{\alpha} \) and \( \beta \), and it therefore commutes with \( \hat{p} \).) Comparing (4.5c) and (4.5d) we see that
\[ \Phi^{-}_E = C \Psi^*_E \quad (4.7) \]
and
\[ \Phi^{+}_E = C \Psi^*_E \]
The matrix $C$, besides satisfying (4.6), must be unitary to preserve the normalization of the wave function. If we take

$$C = \begin{pmatrix} \frac{\beta \alpha_4^*}{r} & \alpha_2^* \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

we can satisfy both requirements since

$$C^{-1} = \left( \frac{\beta \alpha_4^*}{r} \right)^{-1} = i \alpha_y^* \beta^* = i \alpha_y^* \beta$$

$$C^{-1} \alpha_k = (i \alpha_y \beta) \alpha_k \left( \frac{1}{r} \beta \alpha_4 \right) = \alpha_k = \alpha_k^* \quad \text{for } k = x, y$$

$$C^{-1} \alpha_y = (i \alpha_y \beta) \alpha_y \left( \frac{1}{r} \beta \alpha_4 \right) = \alpha_y = \alpha_y^*$$

$$C^{-1} \beta = (i \alpha_y \beta) \beta \left( \frac{1}{r} \beta \alpha_4 \right) = \beta = -\beta^*$$

$$C^T = \left( \frac{1}{r} \beta \alpha_4 \right)^* = i \alpha_y \beta^* = i \alpha_y \beta = C^{-1}$$

With these properties of $C$ established, we see that from

$$\Phi_{-E} = C \Psi_E^*$$

we get

$$C^{-1} \Phi_{-E} = \Psi_E^*$$

and taking complex conjugate on both sides:

$$\Psi_E = C^{-1*} \Phi_{-E}^*$$

but

$$C^{-1*} = (i \alpha_y \beta)^* = -i \alpha_y^* \beta^* = i \alpha_y \beta = C$$
hence
\[ \Psi_E = C \Phi^*_E \]
and
\[ \Psi_{-E} = C \Phi^*_E \] (4.9)

Eqs. (4.7) and (4.9) show the complete symmetry of the relation between solutions \( \Psi \) and \( \Phi \). In words: One gets a solution of energy \( E \) for a given charge by taking the complex conjugate solution of energy \( -E \) of the opposite charge, and by operating \( C \) on it. Either of the solutions is called the charge conjugate solution of the other. We have obtained these results by purely formal manipulations of Dirac's equation, and yet we have arrived at conclusions of significant physical interest.

We see that the mysterious solutions to negative energy eigenvalues are in a 1-1 correspondence with solutions to the positive energy eigenvalues, which, however, describe a particle with a charge of opposite sign.

Let us now investigate what happens when we find the charge conjugate solution of a solution of the free particle problem. In that case the charge \( e \) doesn't enter our equations and we therefore expect that the new solution so obtained should still describe the same particle, and, in fact, be simply a linear combination of the regular solutions of that energy set. Let us take a solution corresponding to a negative energy, say \( \eta^{(3)} \).
\[\Psi_{-E} = \left(\frac{W+\mu}{2w} \right)^{1/2} \gamma^{(3)}(\hat{p}) \frac{i}{\sqrt{v}} e^{-\frac{i}{\hbar} \left(\hat{p} \cdot \hat{r} - E t\right)}\]

\[= \left(\frac{W+\mu}{2w} \right)^{1/2} \left(\begin{array}{c}
\frac{-p_z}{w+\mu} \\
\frac{p_y}{w+\mu} \\
1 \\
0
\end{array}\right) e^{-\frac{i}{\hbar} \left(\hat{p} \cdot \hat{r} - (-E) t\right)}\]

then

\[\Phi_E = C \Psi^*_E = \left(\frac{W+\mu}{2w} \right)^{1/2} C \left(\begin{array}{c}
\frac{-p_z^*}{w+\mu} \\
\frac{p_y^*}{w+\mu} \\
1 \\
0
\end{array}\right) e^{-\frac{i}{\hbar} \left(\hat{p} \cdot \hat{r} - E t\right)}\]

\[= \left(\frac{W+\mu}{2w} \right)^{1/2} \left(\begin{array}{c}
0 \\
1 \\
\frac{-p_z}{w+\mu} \\
\frac{p_y}{w+\mu}
\end{array}\right) e^{-\frac{i}{\hbar} \left(-\hat{p} \cdot \hat{r} + E t\right)} = \left(\frac{W+\mu}{2w} \right)^{1/2} \gamma^{(2)}(\hat{p}) e^{-\frac{i}{\hbar} \left(-\hat{p} \cdot \hat{r} + E t\right)}\]

Thus \(\Phi_E\) turns out to be a solution to a positive energy eigenvalue for a particle with opposite momentum \(-p\) and the spin upside down. In other words the operation \(C \Psi^*_E\) leads to a solution in which time runs backwards. This is an indication of the fact that the motion of a particle with a given charge is a backwards motion.
in time of a particle with opposite charge in a negative energy state.

That we must reverse the sign of the charge when we reverse the time is obvious if we consider a current \( \mathbf{j} = \mathbf{e} \frac{d \mathbf{s}}{dt} \) since

\[
e \frac{d \mathbf{s}}{dt} = (-e) \frac{d \mathbf{s}}{d(-t)}
\]

and that we must reverse the sign of the energy is again obvious since

\[
e^{\frac{i}{\hbar} E t} = e^{\frac{i}{\hbar} (-E)(-t)}
\]

To end this section we should remark that in the literature a slightly different charge conjugation matrix \( C \) is frequently used. The definition of this matrix which we shall denote by \( C_s \) can be gained replacing (4.2) by

\[
\Phi_E^\dagger \psi = C_s \left( \psi^\dagger \beta \right)
\]

so that

\[
C_s = C \beta^* = \gamma_2 \beta^* = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
\]
Chapter V: Solution of Dirac's Equation in an External Field

First Dirac Corrections to the Schroedinger Equation

We have shown (see p. 55-57) that, in the non-relativistic limit, Dirac's equation for the four-spinor $\Psi$ reduces to the Schroedinger (Pauli) equation for the "large" spinor $\Psi$. Essentially what was done was to replace $W + \mu$ by $2\mu$ on the assumption that $\mu \gg \rho$ (i.e. $\frac{\rho}{c} \ll 1$). We now proceed to obtain the equation for the "large" spinor $\Psi$ which is accurate up to second order terms in $\frac{\rho}{c}$.

With application to the hydrogen atom in mind, we consider Dirac's equation in an external electrostatic field, i.e.

$$\hat{A} = 0, \quad \phi \neq 0$$

Dirac's equation in this case is

$$\left( c \alpha \gamma + \beta \mu c \right) \Psi = \left( E - e\phi \right) \Psi$$

or

$$\left( \alpha \gamma + \beta \mu \right) \Psi = \left( W - U \right) \Psi$$

(5.1)

where $U = \frac{\rho}{c}$, $\rho = e\phi$

In terms of the two spinors $\Psi$ and $\varphi$, (5.1) becomes

$$\left( \alpha \gamma \right) \varphi = \left( W - \mu - U \right) \Psi$$

(5.2a)

$$\left( \alpha \gamma \right) \varphi = \left( W + \mu - U \right) \varphi$$

(5.2b)

From (5.2b)

$$\varphi = \frac{1}{W + \mu - U} \left( \alpha \gamma \right) \Psi$$
It is at this stage that the non-relativistic approximation is made by replacing \( W + \mu - U \) by just \( 2\mu \) (the potential energy \( V \) is also assumed to be small compared to \( mc^2 \)), and changing \( \psi \) to \( \psi_0 \), \( \psi \) to \( \psi_0 \), where the time dependence of \( \psi_0 \), \( \psi_0 \) consists of only the kinetic energy. Instead of this approximation, we expand \( \frac{1}{W + \mu - U} \):

\[
\frac{1}{W + \mu - U} = \frac{1}{2\mu} \left( 1 - \frac{W - U - \mu}{2\mu} \right) = \frac{1}{2\mu} \left( 1 - \frac{\varepsilon - V}{\mu c^2} \right)
\]

where \( \varepsilon = E - mc^2 = c(W - \mu) \)

Hence

\[
\psi = \frac{1}{2\mu} \left( 1 - \frac{\varepsilon - V}{2\mu c^2} \right) \mathbf{\sigma} \cdot \mathbf{p} \psi 
\]

and therefore 5.2a becomes

\[
\frac{1}{2\mu} \mathbf{\sigma} \cdot \mathbf{p} \left( 1 - \frac{\varepsilon - V}{2\mu c^2} \right) \mathbf{\sigma} \cdot \mathbf{p} \psi = \frac{\varepsilon - V}{c} \psi
\]

or

\[
\frac{\mathbf{p}^2}{2m} \psi - \frac{\mathbf{\sigma} \cdot \mathbf{p} (\varepsilon - V) \mathbf{\sigma} \cdot \mathbf{p}}{4\mu^2} \psi = \frac{\varepsilon - V}{c} \psi
\]

Eq. (5.3) would be Schroedinger's equation except for the term

\[-\frac{1}{4\mu^2} \mathbf{\sigma} \cdot \mathbf{p} (\varepsilon - V) \mathbf{\sigma} \cdot \mathbf{p} \] . This correction term is obviously proportional to \( \frac{\mathbf{p}^2}{\mu^2} \) and therefore does not appear when we confine ourselves to terms accurate up to \( \frac{\mathbf{p}}{\mu} \) (or \( \frac{\mathbf{v}}{c} \)). We rewrite this correction term as follows:

\[
-\frac{\mathbf{\sigma} \cdot \mathbf{p} (\varepsilon - V) \mathbf{\sigma} \cdot \mathbf{p}}{4\mu^2} = \frac{(\varepsilon - V) \mathbf{p}^2}{4\mu^2} + \frac{[\mathbf{\sigma} \cdot (\mathbf{p} V)] (\mathbf{\sigma} \cdot \mathbf{p})}{4\mu^2}
\]
We now make use of the relation
\[
(\hat{\sigma} \cdot \hat{A})(\hat{\sigma} \cdot \hat{B}) = \hat{A} \cdot \hat{B} + i \hat{\sigma} \cdot (\hat{A} \times \hat{B})
\]
which holds for any two vectors \(\hat{A}\) and \(\hat{B}\) that commute with \(\hat{\sigma}\).

Taking \(\hat{A} = \hat{p} V\) and \(\hat{B} = \hat{p}\) the correction term becomes
\[
- \frac{\varepsilon - V}{4 \mu^2} \hat{p}^2 + \frac{(\hat{p} \cdot V) \hat{p}}{4 \mu^2} + i \hat{\sigma} \cdot \frac{(\hat{p} \cdot V) \times \hat{p}}{4 \mu^2}
\]
\[
= - \frac{\varepsilon - V}{4 \mu^2} + \frac{(\hat{p} \cdot V) \hat{p}}{4 \mu^2} + \frac{i}{4 \mu^2} \hat{\sigma} \cdot \nabla \times \hat{p}
\]
and equation (5.3a) becomes
\[
\left[ \frac{\hat{p}^2}{2 \mu} - \frac{\varepsilon - V}{4 \mu^2} \hat{p}^2 + \frac{(\hat{p} \cdot V) \hat{p}}{4 \mu^2} + \frac{i}{4 \mu^2} \hat{\sigma} \cdot \nabla \times \hat{p} \right] \psi = (\varepsilon - V) \psi \quad (5.3b)
\]

We note that \(\hat{p}\) in \((\hat{p}V)\) operates on \(V\) only. We now must insert a further correction. Eq. (5.3b) is a two-component equation and therefore the probability of finding the particle at a given place is
\[
\mathcal{P} = \{\psi, \psi\} = \sum_{i=1}^{2} \left| \psi_i \right|^2
\]
and \(\int \mathcal{P} d^4x\) should be unity. This, however, is not the case since actually
\[
\mathcal{P} = \{\Psi, \Psi\} = \sum_{i=1}^{2} \left( \left| \psi_i \right|^2 + \left| \varphi_i \right|^2 \right)
\]
Hence we must correct for this. By using the first term in Eq. (5.2c) we can write
\[ S = \{ \psi, \bar{\psi} \} + \{ \bar{\psi}, \psi \} = \{ \psi, \bar{\psi} \} + \{ \frac{1}{2 \mu} \bar{\psi} \gamma^\mu \bar{\psi}, \frac{1}{2 \mu} \bar{\psi} \gamma^\mu \bar{\psi} \} \]

\[ = \{ \psi, \bar{\psi} \} + \{ \psi, \frac{p^2}{\mu^2} \bar{\psi} \} = \{ (1 + \frac{p^2}{\mu^2}) \psi, (1 + \frac{p^2}{\mu^2}) \bar{\psi} \} \]

if magnitudes of order \( \left( \frac{p^2}{\mu^2} \right)^2 \) are neglected. Thus we get the correct normalization if we introduce \( \psi' \) instead of \( \psi \) where

\[ \psi' = (1 + \frac{p^2}{\mu^2}) \psi \quad \text{or} \quad \psi = (1 - \frac{p^2}{\mu^2}) \psi' \]

We now replace \( \psi \) by \( (1 - \frac{p^2}{\mu^2}) \psi' \) in (5.3b) but only in those terms which are not small in themselves. That is to say in the first term on the left and in the term on the right. In all other terms we set \( \psi = \psi' \). We then obtain

\[ \left[ \frac{p^2}{2m} - \frac{p^2}{2m} \frac{p^2}{\mu^2} - \frac{\mu^2}{4\mu^2} \frac{p^2}{\mu^2} + \frac{(\bar{\psi} V \psi) \cdot \bar{\psi} \psi}{4\mu^2} + \frac{\bar{\psi} \gamma^\mu \nabla \bar{\psi} \cdot \nabla \psi}{4\mu^2} \right] = \left[ (\epsilon - V - \frac{\mu^2}{\mu^2} \bar{\psi} \psi) \right] \psi' \]

(5.3c)

Combining the third term on the left with the last term on the right to obtain \( \frac{\mu^2}{\mu^2} \frac{p^2}{\mu^2} \) on the left, we then write this term as follows:

\[ - \frac{\mu^2}{\mu^2} \frac{p^2}{\mu^2} \psi' = \frac{1}{2 \mu^2} \left[ \bar{\psi} \psi \right] \psi' \]

where \( \bar{\psi} \) in \( (p^2 \psi) \) and \( (\bar{\psi} V) \) operates on \( V \) only. Then Eq. (5.3c) becomes

\[ \left[ \frac{p^2}{2m} - \frac{p^2}{2m} \frac{p^2}{\mu^2} - \frac{p^2}{\mu^2} (\epsilon - V) - \frac{1}{2 \mu^2} (p^2 \psi) + \frac{\bar{\psi} \gamma^\mu \nabla \bar{\psi} \cdot \nabla \psi}{4\mu^2} \right] \psi' = (\epsilon - V) \psi' \]
We now make use of the approximate relation
\[(\varepsilon - V)\psi' \approx \frac{p^2}{2m} \psi'\]

the accuracy of which is sufficient to replace the second order term \(\frac{p^2}{\mu^2} (\varepsilon - V)\psi'\) by \(\frac{p^2}{\mu^2} \frac{p^2}{2m} \psi'\). We then finally obtain
\[
\left[\frac{p^2}{2m} - \frac{p^4}{8m^3 c^2} + \frac{2}{8m^3 c^2} (\vec{\nabla}^2 V) + \frac{2}{4m^2 c^2} \vec{\nabla} \times \vec{\rho}\right] \psi' \approx (\varepsilon - V)\psi' \quad (5.4)
\]

The first term on the left and the term on the right represent the non-relativistic Schrödinger equation. The second term on the left represents second order relativistic mass correction to the kinetic energy:
\[
\varepsilon = E - mc^2 = (p^2 c^2 + m^2 c^4)^{1/2} - mc^2 = \frac{p^2}{2m} - \frac{p^4}{8m^3 c^2}
\]

The last term on the left represents the spin-orbit coupling energy. For a central force field (spherically symmetric potential) we have \(\vec{\nabla} V = \frac{\vec{r}}{r} \frac{\partial V}{\partial r}\), hence
\[
\frac{2}{4m^2 c^2} \vec{\nabla} \times \vec{p} = \frac{2}{4m^2 c^2} \left( \frac{1}{r} \frac{\partial V}{\partial r} \vec{r} \times \vec{p} \right) = \frac{1}{2m^2 c^2} \frac{1}{r} \frac{\partial V}{\partial r} \vec{s} \cdot \vec{L}
\]

which is, perhaps, a more familiar form for the spin-orbit coupling term. This term can be also written as follows: Since for an observer moving with velocity \(\vec{v}\) an electrostatic field \(\vec{E}\) appears to have a magnetic part \(\vec{\mathcal{E}} = \frac{(\vec{E} \times \vec{v})}{c}\), we obtain, letting \(\vec{p} \rightarrow m\vec{v}\) (and using \(\vec{\nabla} V = -e \vec{E}\)),
\[
\frac{2}{4m c^2} \hat{s} \cdot (-e \vec{\mathcal{E}}) \times m \hat{\sigma} = -\frac{e}{4mc} \hat{s} \cdot \vec{\mathcal{E}} = -\frac{1}{2} \vec{p} \cdot \vec{\mathcal{E}}
\]
where \( \hat{\mathbf{\mu}} = \frac{e \hat{\mathbf{n}}}{2mc} \). Classically we would expect from the Larmor precession an energy of interaction

\[
\Delta E = - \hat{\mathbf{\mu}} \cdot \hat{\mathbf{E}}
\]

The 1/2 factor is the Thomas factor and it is the Thomas precession that gives an energy of interaction of 1/2 \( \hat{\mathbf{\mu}} \cdot \hat{\mathbf{E}} \) thus making the combined contributions of the Larmor and Thomas precessions equal to -1/2 \( \hat{\mathbf{\mu}} \cdot \hat{\mathbf{E}} \). The Thomas precession is a rather involved kinematic relativistic effect, and it is another triumph for Dirac's equation that it gives directly the correct value for the spin orbit coupling energy.

The term \( \frac{\hbar^2}{8m^2c^2} (\nabla^2V) \) is an interaction taking place only where \( \nabla^2V \) is different from 0. This is at the sources of the potential \( \frac{\mathcal{E}}{e} \). If we call the charge density of these sources \( \varphi_s \) then \( \nabla^2V = 4\pi \varphi_s e \), and this term becomes \( -\frac{\hbar^2 e}{2m^2c^2} \varphi_s \mathcal{V}' \).

A physical explanation of this term can be given as follows. We know that the electron actually performs a curly motion which extends over a linear distance \( \Delta \mathbf{r} \) of the order \( |\Delta \mathbf{r}| \sim \frac{\hbar}{mc} = \) Compton wavelength. Hence we cannot talk of the value of the potential at the position of the electron but must use a series expansion of the form

\[
\sqrt{V(\mathbf{r} + \Delta \mathbf{r})} = V(\mathbf{r}) + \frac{\nabla V \cdot \Delta \mathbf{r}}{2} + \frac{1}{2} \nabla^2 V (\Delta \mathbf{r})^2 + \ldots
\]
The second term in the above expansion vanishes on the average since \( \Delta \hat{\mathbf{r}} = 0 \). However \( (\Delta \hat{\mathbf{r}})^2 \sim \frac{\hbar^2}{m^2 c^2} \), and so we obtain a term 
\[ \sim \frac{1}{2} \frac{\hbar^2}{m^2 c^2} \left( \nabla^2 v \right) \], whose form and order of magnitude is the same as 
that of \( \frac{\hbar^2}{8m^2 c^2} \nabla^2 v \).

Presence of Magnetic Field (Zeeman Effect)

In the preceding section we have assumed that the external field was specified by \( \mathbf{A} = 0, \phi \neq 0 \). We now specify that \( \mathbf{A} \) also be different from 0. The Dirac equation in that case is (see Eq. 4.3)

\[
\frac{\hbar}{2m} \left( \mathbf{\sigma} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A}) + \beta \mu \right) \Phi = (E - e \phi) \Phi
\]

or in Maxwell form

\[
\mathbf{\nabla} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A}) \Phi = (W - \mu - U) \Phi \tag{5.5a}
\]

\[
\mathbf{\nabla} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A}) \Phi = (W + \mu - U) \Phi \tag{5.5b}
\]

and from (5.5b) we get

\[
\Phi = \frac{\mathbf{\nabla} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A})}{W + \mu - U} \Phi \sim \frac{\mathbf{\nabla} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A})}{2\mu} \Phi \tag{5.5c}
\]

To obtain (5.5c) we have made the non-relativistic approximation

\( W + \mu = U \sim 2\mu \). We also replace \( W - \mu - U \) by \( \frac{e^2}{c^2} A^2 \), and furthermore we shall neglect terms proportional to \( \frac{e^2}{c^2} A^2 \).

Introducing (5.5c) into (5.5a) we obtain

\[
\mathbf{\nabla} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A}) \left( \frac{1}{2\mu} \mathbf{\nabla} \cdot (\mathbf{\hat{p}} - \frac{e}{c} \mathbf{A}) \right) \Psi = \frac{e - \gamma}{c} \Psi
\]

or

\[
\left\{ \frac{p^2}{2m} - \frac{e}{2\mu} \left[ (\mathbf{\sigma} \cdot \mathbf{\hat{p}})(\mathbf{\sigma} \cdot \mathbf{\hat{A}}) + (\mathbf{\hat{A}})(\mathbf{\nabla} \cdot \mathbf{\hat{p}}) \right] \right\} \Psi = (E - \gamma) \Psi
\]
The expression in the square brackets can be rewritten as

\[ \hat{p} \cdot \hat{A} + \hat{A} \cdot \hat{p} + i \hat{\sigma} \cdot \hat{A} \times \hat{p} + i \hat{\sigma} \cdot (\hat{p} \times \hat{A}) \]

Now \( \hat{p} \cdot \hat{A} = \hat{A} \cdot \hat{p} \) because \( [\hat{p}, \hat{A}] = \frac{\hbar}{i} \) \( \text{div} \hat{A} = 0 \). Also

\[ \hat{p} \times \hat{A} \psi = (\hat{p} \times \hat{A}) \psi + (\hat{p} \psi) \times \hat{A} \]

where by \( (\hat{p} \times \hat{A}) \psi \) we mean that \( \hat{p} \) operates on \( \hat{A} \) only, i.e.

\( (\hat{p} \times \hat{A}) \psi = \frac{i}{\hbar} \text{(curl} \hat{A}) \psi \). Hence we have

\[ \hat{p} \cdot \hat{A} + \hat{A} \cdot \hat{p} + i \hat{\sigma} \cdot (\hat{A} \times \hat{p} + \hat{p} \times \hat{A}) = i \hat{\sigma} \cdot \hat{B} + \hat{\sigma} \cdot \text{curl} \hat{A} \]

But \( \text{curl} \hat{A} = \hat{\mathbf{E}} \) = magnetic field intensity. Furthermore, if \( \hat{\mathbf{E}} \)

is constant, we can take \( \hat{A} = \frac{1}{2} (\hat{\mathbf{E}} \times \hat{r}) \), then

\[ i \hat{\sigma} \cdot \hat{B} = \hat{\mathbf{E}} \times \hat{r} \cdot \hat{p} = (\hat{\mathbf{E}} \times \hat{p}) \cdot \hat{r} = \hat{L} \cdot \hat{\mathbf{H}} \]

Thus the equation for \( \psi \) becomes

\[ \left( \frac{\hat{p}^2}{2m} - \frac{e}{2mc} \left( \hat{L} + 2 \hat{S} \right) \cdot \hat{r} \right) \psi = (\varepsilon - \mu) \psi \quad (5.5d) \]

The term \( - \frac{e}{2mc} (\hat{L} + 2 \hat{S}) \cdot \hat{r} \) represents the interaction energy with
the external magnetic field and gives rise to the anomalous Zeeman

effect. Eq. (5.5d) is by no means exact but represents only a
non-relativistic approximation. The exact relativistic equation
is considerably more complicated. As we have seen in the preceding
section, even the equation that is only one order of \( \hat{p} \) more
accurate than (5.5d) contains several additional terms. Even the
term causing the Zeeman effect has to be changed considerably,
since the relativistic mass correction is different for
and because the first one requires the proper correction for motion on an orbit whereas the second one requires the proper correction for the "curly" motion.

Solution of Dirac's Equation in a Central Field

The solution \( \Psi \) of Schroedinger's equation can be factored into a radial part, \( R(r) \), and a part involving the angles, \( Y^\ell_m(\theta, \varphi) \):

\[
\Psi(r, \theta, \varphi) = R(r) Y^\ell_m(\theta, \varphi)
\]

This is possible because in a central field orbital angular momentum \( \vec{L} \) is conserved, i.e. \( [\vec{L}, H] = 0 \Rightarrow \vec{L} = \vec{L} \times \vec{p} \), and therefore the \( Y^\ell_m \)'s can simultaneously be eigenfunctions of \( L^2, L^z \) and \( H \):

\[
\begin{align*}
H \Psi & = \varepsilon \Psi \\
L^2 \Psi & = \ell(\ell+1) \Psi \\
L^z \Psi & = m \Psi
\end{align*}
\]

The radial equation for \( R \) is, as we know,

\[
-\frac{\hbar^2}{2m} \left( \frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} \right) + \frac{\hbar^2}{2m} \frac{\ell(\ell+1)}{r^2} R = (\varepsilon - V) R
\]

To solve Dirac's equation in a central field we proceed in a similar fashion. The orbital angular momentum no longer commutes with the Hamiltonian, however the total angular momentum

\[
\vec{J} = \vec{L} + \vec{S}
\]

does, as it should since the Hamiltonian is invariant to a rotation in space.
We prove the commutability nevertheless in detail:

\[
[ J_x, H ] = [ L_x + S_x, c(\hat{\vec{\rho}} + i \alpha \vec{r} + V(r) ) ]
= [ L_x + S_x, c(\hat{\vec{\rho}}) ]
\]

since \( mc^2\beta + V(r) \) commutes with both \( L_x \) and \( S_x \) for a spherically symmetric potential. Now

\[
[ L_x, \hat{\vec{\rho}} ] = \hat{\alpha} \cdot \left( (y \rho_z - z \rho_y) \hat{\vec{p}} - \hat{\vec{p}} (y \rho_z - z \rho_y) \right)
= \hat{\alpha} \cdot \left( (y \rho_z \rho_y - y \rho_y \rho_z) \hat{\vec{n}}_y - (z \rho_z \rho_y - z \rho_y \rho_z) \hat{\vec{n}}_z \right)
= \hat{\alpha} \left( i \mathbf{\tau} \rho_z \hat{\vec{n}}_y - i \mathbf{\tau} \rho_y \hat{\vec{n}}_z \right) = i \mathbf{\tau} \left( \alpha_y \rho_z - \alpha_z \rho_y \right)
\]

\[
[ S_z, \hat{\vec{\rho}} ] = \frac{i}{2} \left[ \sigma_x, \alpha_x \rho_x + \alpha_y \rho_y + \alpha_z \rho_z \right]
= \frac{i}{2} \left[ \sigma_x, \alpha_y \rho_y + \alpha_z \rho_z \right]
= \frac{i}{2} \left( 2i \mathbf{\alpha} \rho_y - 2i \mathbf{\alpha} \rho_z \right) = i \mathbf{\tau} \left( \alpha_z \rho_y - \alpha_y \rho_z \right)
\]

where we make use of the relation

\[
\sigma_i \mathcal{L}_k - \sigma_k \mathcal{L}_i = 2i \mathbf{\alpha} \mathcal{L}
\]

\( i, k, \mathcal{L} = \) cyclic permutation of \( x, y, z \)

Hence we conclude \( [ J_x, H ] = 0 \)
One similarly shows that \([J_y, H] = 0, [J_z, H] = 0\). From this it follows that
\[
\left[ J^2, H \right] = [J^2_y, H] + [J^2_z, H] = 0
\]
because
\[
\left[ J^2_x, H \right] = J_x \left[ J^2_x, H \right] + [J^2_x, H] = 0 \text{ etc.}
\]
Furthermore, \(J^2\) and any component of \(J\) commute (see p. 6); therefore the eigenfunctions of \(J^2\) and, say, \(J_z\), can simultaneously be eigenfunctions of the Hamiltonian which allows us to factor the solution \(\Psi\) of Dirac's equation into a radial part and a part involving the angles. The latter part must, of course, be a spinor field corresponding to the eigenfunctions of \(J^2\) and \(J_z\).

These we have calculated before (see p. 40, 41) for \(j = \ell + \frac{1}{2}\):
\[
\Psi^+_{j m} = \frac{\ell + m + \frac{1}{2}}{2\ell + 1} \Psi^+_{\ell, \frac{\ell}{2}, m + \frac{1}{2}} + \frac{\ell - m + \frac{1}{2}}{2\ell + 1} \Psi^+_{\ell, \frac{\ell}{2}, m - \frac{1}{2}}
\]
for \(j = \ell - \frac{1}{2}\):
\[
\Psi^-_{j m} = \frac{\ell - m + \frac{1}{2}}{2\ell + 1} \Psi^-_{\ell, \frac{\ell}{2}, m + \frac{1}{2}} - \frac{\ell + m + \frac{1}{2}}{2\ell + 1} \Psi^-_{\ell, \frac{\ell}{2}, m - \frac{1}{2}}
\]
We can write above as a two-row column (the way we usually write spinors). To indicate that these are, in a sense, spherical harmonics, but spinor spherical harmonics, we write them as \(\Psi^j_{jm}\).

Furthermore, to distinguish between the stretched and non-stretched cases, we denote the former by \(\Psi^+\) and the latter by \(\Psi^-\):
\[
\Psi^\pm_{jm} = \begin{pmatrix}
\Psi^+_{j m} \\
\Psi^-_{j m}
\end{pmatrix}
\]

\((5.6a)\)
\[ \tilde{Y}_{jm} = \begin{pmatrix} \tilde{y}_{j+\frac{1}{2}, m-\frac{1}{2}} \\ \frac{(-1)^j}{2^j + 2} y_{j+\frac{1}{2}, m+\frac{1}{2}} \\ \frac{(-1)^{j+1}}{2^j + 2} y_{j+\frac{1}{2}, m-\frac{1}{2}} \end{pmatrix} \]

We note that for a given value of \( j \) we obtain two, not one, spinor spherical harmonics, and that \( \tilde{Y} \) and \( \tilde{\bar{Y}} \) have opposite parity. The property of parity depends on whether or not the wave function changes sign under an inversion in space, i.e. under replacement of the value of every space coordinate by the negative of that coordinate. If there is no change in sign under inversion, the wave function is said to have even parity; in the other case it is said to have odd parity. It is easy to show that spherical harmonics obey the following relation

\[ \tilde{Y}_{j', m'} (\theta, \phi) = (-1)^{j'} \tilde{Y}_{j, m} (\pi - \theta, \pi + \phi) \]

Therefore the parity of \( \tilde{Y} \) is given by \((-1)^{j+\frac{1}{2}}\), and the parity of \( \tilde{\bar{Y}} \) is given by \((-1)^{j+\frac{1}{2}}\), which proves that \( \tilde{Y} \) and \( \tilde{\bar{Y}} \) have opposite parity.

Since \( \tilde{Y} \) and \( \tilde{\bar{Y}} \) form a complete set, the actual solution of Dirac's equation must be some linear combination of them. The coefficients in this linear combination may be function of \( r \) (such functions behave as constants as far as operating with \( J^2 \) or \( J_z \) is concerned). This leads us to assume for a solution

\[ \tilde{\psi} = \begin{pmatrix} \tilde{\psi}_1(r) \tilde{Y}_{j m}^+ + \tilde{\psi}_2(r) \tilde{\bar{Y}}_{j m} \\ \frac{d}{dr} \tilde{\psi}_1(r) \tilde{Y}_{j m}^+ + \frac{d}{dr} \tilde{\psi}_2(r) \tilde{\bar{Y}}_{j m} \end{pmatrix} \]
We write (5.7) in that particular form for reasons that will become apparent later. The upper row in (5.7) gives us the solution for the "large" spinor $\Psi$ and the lower row gives the solution for the "small" spinor $\tilde{\Psi}$. Now consider either of Eq. (5.2a) or (5.2b), the two Dirac equations for $\Psi$ and $\tilde{\Psi}$.

To be specific we take Eq. (5.2a)

$$(\tilde{\sigma} \cdot \tilde{p}) \Psi = (\mathcal{W} - \mu - U) \Psi$$

Under an inversion in the space coordinates, $\tilde{p}$ will become $-\tilde{p}$, while $\tilde{\sigma}$, $\mathcal{W}$ and $U$ will remain the same (U doesn't change sign because we are dealing with a spherically symmetric potential). However, above equation must obviously be still valid, and therefore we conclude that under a space inversion either $\tilde{\Psi}$ or $\tilde{\Psi}$ must change sign, but not both. This is to say $\tilde{\Psi}$ and $\tilde{\Psi}$ must have opposite parity. Since $\tilde{\Psi}$ and $\tilde{\Psi}$ each must have definite parity, we see that if $\Psi = \tilde{\Psi}_1(r) \tilde{\Psi}_{jm}$, we must have $\Psi = \frac{\tilde{\Psi}_1(r)}{i} \tilde{\Psi}_{jm}$. In other words we actually have two solutions:

$$\Psi_{n_{jm}} = \left( \begin{array}{c} \tilde{\Psi}_1(r) \tilde{\Psi}_{jm} \\ \tilde{\Psi}_2(r) \tilde{\Psi}_{jm} \end{array} \right) \quad \text{or} \quad \left( \begin{array}{c} \tilde{\Psi}_1(r) \tilde{\Psi}_{jm} \\ \tilde{\Psi}_2(r) \tilde{\Psi}_{jm} \end{array} \right)$$

The statement that $\Psi$ and $\tilde{\Psi}$ each must have definite parity in non-degenerate states follows from the fact that the Hamiltonian is invariant under the following parity operation: If $P$ stands for
the usual parity operator (inversion in space coordinates),
then the operator $P$ leaves the Hamiltonian invariant. This
then implies that $\mathcal{F}$ must have definite parity under the
$P$ operation, hence $\mathcal{U}$ must have definite parity under the $+P$
operation, and $\mathcal{Q}$ must have definite parity under the $-P$ operation,
which is just another way of saying that $\mathcal{U}$ and $\mathcal{Q}$ each must have
definite parity.

We now proceed to find the equations satisfied by the func-
tions $f(r), g(r)$, i.e. we ask for the equivalent of the Schroedinger
radial equation. In order to answer this, we must be able to
evaluate expressions of the form

$$(\hat{\mathcal{S}} \cdot \vec{p}) \left[ R(r) \frac{\vec{r}}{r} Y_{jm} \right] \quad \text{and} \quad (\hat{\mathcal{S}} \cdot \vec{p}) \left[ R(r) \frac{\vec{r}}{r} Y_{jm} \right]$$

Making use of the relations

$$(\hat{\mathcal{S}} \cdot \vec{A}) (\hat{\mathcal{S}} \cdot \vec{B}) = \vec{A} \cdot \vec{B} + i \hat{\mathcal{S}} \cdot \vec{A} \times \vec{B}$$

we have:

$$\begin{align*}
\hat{\mathcal{S}} \cdot \vec{p} & = \frac{(\hat{\mathcal{S}} \cdot \vec{r})(\hat{\mathcal{S}} \cdot \vec{p})}{r^2} = \frac{\hat{\mathcal{S}} \cdot \vec{r}}{r^2} \left[ \hat{\mathcal{S}} \cdot \vec{p} + i \hat{\mathcal{S}} \cdot \vec{r} \times \vec{p} \right] \\
& = \frac{\hat{\mathcal{S}} \cdot \vec{r}}{r^2} \left[ \frac{\hat{\mathcal{S}}}{i} r + \frac{2}{r} + i \hat{\mathcal{S}} \cdot \vec{L} \right] \quad (5.8a)
\end{align*}$$

Taking the expression in the square brackets and operating with
it on $R(r) \frac{\vec{r}}{r} Y_{jm}$ we obtain
\[
\left( \frac{\alpha}{i} \frac{\partial}{\partial r} + i e \mathbf{E} \right) \mathbf{R} \mathbf{y}_{j}^{+}\mathbf{m} = \left[ \frac{\alpha}{i} \frac{\partial R}{\partial r} + R \mathbf{i} (\mathbf{e} \cdot \mathbf{L}) \right] \mathbf{y}_{j}^{+}\mathbf{m} \tag{5.8b}
\]

and

\[
(\mathbf{e} \cdot \mathbf{L}) = \mathbf{k} \begin{pmatrix}
L_2 & L_3 \\
L_3 & L_2
\end{pmatrix}
\begin{pmatrix}
\frac{j + m}{2j} & \mathbf{y}_{j - \frac{1}{2}, m - \frac{1}{2}} \\
\frac{j - m}{2j} & \mathbf{y}_{j - \frac{1}{2}, m + \frac{1}{2}}
\end{pmatrix}
\]

\[
= \mathbf{k} \begin{pmatrix}
\frac{j + m}{2j} L_2 \frac{\mathbf{y}_{j - \frac{1}{2}, m - \frac{1}{2}}}{z_j} + i \frac{j - m}{2j} L_3 \frac{\mathbf{y}_{j - \frac{1}{2}, m + \frac{1}{2}}}{z_j} \\
\frac{j + m}{2j} L_3 \frac{\mathbf{y}_{j - \frac{1}{2}, m - \frac{1}{2}}}{z_j} - i \frac{j - m}{2j} L_2 \frac{\mathbf{y}_{j - \frac{1}{2}, m + \frac{1}{2}}}{z_j}
\end{pmatrix}
\]

\[
= \mathbf{k} \begin{pmatrix}
\left( m - \frac{1}{2} \right) \frac{j + m}{2j} + v (j - m) (j + m) \frac{j - m}{2j} & \mathbf{y}_{j - \frac{1}{2}, m - \frac{1}{2}} \\
\sqrt{(j + m)(j - m)} \frac{j + m}{2j} - (m + \frac{1}{2}) \frac{j - m}{2j} & \mathbf{y}_{j - \frac{1}{2}, m + \frac{1}{2}}
\end{pmatrix}
\]

\[
= \mathbf{k} \begin{pmatrix}
\left( j - \frac{1}{2} \right) \frac{j + m}{2j} \mathbf{y}_{j - \frac{1}{2}, m - \frac{1}{2}} \\
\left( j - \frac{1}{2} \right) \frac{j - m}{2j} \mathbf{y}_{j + \frac{1}{2}, m + \frac{1}{2}}
\end{pmatrix}
= \mathbf{k} \left( j - \frac{1}{2} \right) \mathbf{y}_{j}^{+}\mathbf{m}
\]

and therefore (5.8b) becomes

\[
\frac{\alpha}{i} \left[ \frac{\partial R}{\partial r} - R \left( j - \frac{1}{2} \right) \right] \mathbf{y}_{j}^{+}\mathbf{m} \tag{5.8c}
\]
Introducing (5.8c) into (5.8a) we have:

\[
(\vec{\sigma} \cdot \vec{p}) R \overset{\sim}{Y}^{+}_{jm} = \frac{i}{\hbar} \left[ \frac{\partial R}{\partial r} - (j + \frac{1}{2}) \frac{R}{r} \right] \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} \quad (5.9a)
\]

where we have written \( \overset{\sim}{\sigma} \cdot \vec{r} = \sigma \) after the square bracket rather than before because \( \sigma \) commutes with \( R \), which is just a function of the space coordinate \( r \). By a similar procedure one obtains

\[
(\vec{\sigma} \cdot \vec{p}) R \overset{\sim}{Y}^{+}_{jm} = \frac{i}{\hbar} \left[ \frac{\partial R}{\partial r} + (j + \frac{3}{2}) \frac{R}{r} \right] \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} \quad (5.9b)
\]

We still have to evaluate \( \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} \) and \( \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} \).

Under a space inversion \( r \) becomes \( -r \), therefore \( \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} \) must have the opposite parity to \( \overset{\sim}{Y}^{+}_{jm} \). However, as far as behaviour under a rotation in space is concerned \( \overset{\sim}{\sigma} \cdot \vec{r} \) is just a scalar, therefore \( \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} \) must be characterized by the same \( j \) and \( m \) as \( \overset{\sim}{Y}^{+}_{jm} \). This forces us to conclude

\[
\overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} = A \overset{\sim}{Y}^{+}_{jm}
\]

and

\[
\overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} = A^* \overset{\sim}{Y}^{+}_{jm}
\]

Now

\[
\frac{\partial \overset{\sim}{Y}^{+}_{jm}}{\partial r} + (\overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm}) = A \overset{\sim}{\sigma} \cdot \vec{r} \overset{\sim}{Y}^{+}_{jm} = AA^* \overset{\sim}{Y}^{+}_{jm}
\]

hence

\[
AA^* = \frac{\partial \overset{\sim}{Y}^{+}_{jm}}{\partial r} \overset{\sim}{\sigma} \cdot \vec{r} = I
\]

and we must have \( A = e^{i\hat{S}} \). The question of what \( \hat{S} \) is can be decided only by an actual computation of the
expression \( \frac{\hat{\sigma} \cdot \hat{p}}{r} Y_{jm} \), which is what we are trying to avoid. It turns out that if \( \tilde{\psi} \) and \( \tilde{\varphi} \) are defined as in (5.6a,b) we have \( A = +1 \). We would like to point out here that our definition of \( \tilde{\varphi} \) in (5.6b) is actually the negative of the expression given by the Clebsch-Gordan coefficients in Chap. I, p.41. The choice of the phases of eigenfunctions is in general arbitrary; had we defined \( \varphi \) as given in Chap I, we would now have to take \( A = -1 \).

Our final result from all the preceding computations is

\[
\begin{align*}
(\hat{\sigma} \cdot \hat{p}) \tilde{\psi}(r) \tilde{Y}_{j}^{+}j_{m} &= \frac{\hbar}{i} \left[ \frac{\partial \hat{R}}{\partial r} - \left( j - \frac{1}{2} \right) \hat{R} \right] \tilde{Y}_{j}^{-}j_{m} \\
(\hat{\sigma} \cdot \hat{p}) \tilde{\varphi}(r) \tilde{Y}_{j}^{+}j_{m} &= \frac{\hbar}{i} \left[ \frac{\partial \hat{R}}{\partial r} + \left( j + \frac{1}{2} \right) \hat{R} \right] \tilde{Y}_{j}^{-}j_{m}
\end{align*}
\] (5.10)

We are now ready to obtain the radial equations.

a) The Stretched Case

As we have shown before \( \varphi \) and \( \psi \) have opposite parity; hence if one corresponds to the stretched case, the other must correspond to the non-stretched case. Therefore when we refer to the whole system as being in the stretched state of affairs, we actually mean that the "large" spinor is in the stretched situation

\[
\begin{align*}
\tilde{\varphi} &= f(r) \tilde{Y}_{j}^{+}j_{m} \\
\tilde{\psi} &= \frac{\hbar}{i} \tilde{Y}_{j}^{-}j_{m}
\end{align*}
\] (5.11)

Substituting (5.11) into

\[
\begin{align*}
(\hat{\sigma} \cdot \hat{p}) \tilde{\varphi} &= (W - \mu - U) \tilde{\varphi} \\
(\hat{\sigma} \cdot \hat{p}) \tilde{\psi} &= (W + \mu - U) \tilde{\psi}
\end{align*}
\] (5.2)
and making use of (5.10) we obtain:

\[
\frac{d\jmath}{d\tau} + \left( j + \frac{3}{2} \right) \frac{\jmath}{\tau} = \frac{-i}{\kappa} (W - \mu - U) f
\]

\[
\frac{df}{d\tau} - \left( j - \frac{1}{2} \right) \frac{f}{\tau} = \frac{i}{\kappa} (W - \mu - U) g
\]

(5.12)

b) The Non-stretched Case

In this case we have

\[
\Psi = f^{(\tau)} \frac{\Psi}{\jmath} \jmath m, \quad \phi = \frac{g^{(\tau)} + j}{\jmath} \frac{\phi}{\jmath} \jmath m
\]

(5.13)

which when substituted in (5.2) gives (making use of (5.10)):

\[
\frac{df}{d\tau} + \left( j + \frac{1}{2} \right) \frac{f}{\tau} = \frac{i}{\kappa} (W - \mu - U) f
\]

\[
\frac{df}{d\tau} + \left( j - \frac{1}{2} \right) \frac{f}{\tau} = \frac{i}{\kappa} (W + \mu - U) g
\]

(5.14)

Eqs. (5.12) and (5.14) can be combined into

\[
\frac{df}{d\tau} + \left( 1 + \kappa \right) \frac{f}{\tau} = \frac{i}{\kappa} (W + \mu - U) \phi
\]

(5.15)

\[
\frac{df}{d\tau} + \left( 1 - \kappa \right) \frac{f}{\tau} = \frac{-i}{\kappa} (W - \mu - U) f
\]

where

\[
\kappa = \begin{cases} \left( j + \frac{1}{2} \right) & \text{for the stretched case} \\ \left( j - \frac{1}{2} \right) & \text{for the non-stretched case} \end{cases}
\]

We rewrite the radial equations once more setting

\[
f = \frac{F}{r}, \quad \jmath = \frac{G}{r}:
\]

\[
F' + \kappa \frac{F}{r} = \frac{i}{\kappa} (W + \mu - U) G
\]

(5.16a)

\[
G' - \kappa \frac{G}{r} = \frac{-i}{\kappa} (W - \mu - U) F
\]

(5.16b)
Before solving (5.16) we show that in the non-relativistic limit it is equivalent to Schrödinger's radial equation. Assuming the kinetic and potential energies small compared with the rest energy, we can set \( W + \mu - \dot{U} \approx 2\mu \). Then from (5.16a) we have:

\[
G \approx \frac{\lambda}{2\mu} \left( F' + \kappa \frac{F}{\tau} \right) \tag{5.17a}
\]

hence

\[
G' \approx \frac{\lambda}{2\mu} \left( F'' + \kappa \frac{F'}{\tau} - \kappa \frac{F}{\tau^2} \right) \tag{5.17b}
\]

Substituting (5.17a,b) into (5.16b) we obtain:

\[
\frac{\lambda}{2\mu} \left[ (F'' + \kappa \frac{F'}{\tau} - \kappa \frac{F}{\tau^2}) - \left( \kappa \frac{F'}{\tau} + \kappa \frac{F}{\tau^2} \right) \right] = -\frac{1}{\kappa} \frac{\epsilon - V}{c} F
\]

which is the same as

\[
-\frac{\lambda^2}{2\mu} F'' + \frac{\lambda^2}{2m^2 \tau^2} \kappa(k+1) F = (\epsilon - V) F \tag{5.18}
\]

Eq. (5.18) is identical to Schrödinger's radial equation if we replace in the latter \( R \) by \( \frac{F}{\tau} \) and provided \( k(k+1) = \ell(\ell + 1) \).

From the relation defining \( k \) we have

a) for the stretched case \((j = \ell + 1/2)\)

\[
k(k+1) = (-j - \frac{1}{2})(-j + \frac{1}{2}) = (-\ell)(-\ell + 1) = \ell(\ell + 1)
\]

b) for the non-stretched case \((j = \ell - 1/2)\):

\[
k(k+1) = (-j - \frac{1}{2})(-j + \frac{1}{2}) = (-\ell)(-\ell - 1) = \ell(\ell + 1)
\]
and this completes the proof.

We now proceed to solve (5.16) to a better approximation

than the non-relativistic one. We solve (5.16a) for G and approx-

imate \((W + \mu - U)^{-1}\) by \(\frac{1}{2\mu} (1 - \frac{\epsilon - V}{2\mu c})\) (see p. 95). Thus we obtain

\[
G \approx \frac{k}{2\mu} \left(1 - \frac{\epsilon - V}{2\mu c}\right) \left(F' + \frac{k}{r} F\right)
\]

\[
G' \approx \frac{k}{2\mu} \left[(1 - \frac{\epsilon - V}{2\mu c}) \left(F'' + \frac{k F'}{r} - \frac{k F}{r^2}\right) + \frac{1}{2\mu c} \frac{\partial V}{\partial r} \left(F' + \frac{k}{r^2} \right)\right]
\]

which when substituted into (5.16b) gives:

\[
\left(1 - \frac{\epsilon - V}{2mc^2}\right) \left(-\frac{\kappa^2}{2m} F'' + \frac{\kappa^2}{2mr^2} k(k+1) F\right) - \frac{k^4}{4mc^2} \frac{\partial V}{\partial r} \left(F' + \frac{k F}{r^2}\right) = (\epsilon - V) F
\]

Eq. (5.19) is a two-component equation, and again we must insure

its normalization (see p. 96 and following). Since to the lowest approximation \(G \approx \frac{k}{2\mu} \left(F' + \frac{k F}{r^2}\right)\), we can write

\[
\left\{ G, G \right\} \equiv \left\{ \frac{k}{2\mu} \left(\frac{\partial}{\partial r} + \frac{k}{r}\right) F, \frac{k}{2\mu} \left(\frac{\partial}{\partial r} + \frac{k}{r}\right) F \right\}
\]

\[
= \left\{ F, \frac{\kappa^2}{4m^2} \left(\frac{\partial}{\partial r} + \frac{k}{r}\right) \left(\frac{\partial}{\partial r} + \frac{k}{r}\right) F \right\}
\]

\[
= \left\{ F, \frac{\kappa^2}{4m^2} \left(-\frac{\partial^2}{\partial r^2} + \frac{k(k+1)}{r^2}\right) F \right\}
\]

5463/e/p
where we have made use of the fact that \( \frac{\partial}{\partial r} \) is antihermitian and \( \frac{1}{r} \) is hermitian. Hence if we neglect terms of order \( \left( \frac{\hbar^2}{\mu^2} \right) \frac{1}{r^4} \) we can write

\[
\{ F, F \} + \{ G, G \} \cong \left\{ 1 + \frac{\hbar^2}{\mu^2} \left( \frac{k(k+1)}{r^2} - \frac{\partial^2}{\partial r^2} \right) F, \left[ 1 + \frac{\hbar^2}{\mu^2} \left( \frac{k(k+1)}{r^2} - \frac{\partial^2}{\partial r^2} \right) \right] F \right\}
\]

Hence if we take

\[
F_0 = \left[ 1 + \frac{\hbar^2}{\mu^2} \left( \frac{k(k+1)}{r^2} - \frac{\partial^2}{\partial r^2} \right) \right] F \quad \text{and} \quad F = \left[ 1 - \frac{\hbar^2}{\mu^2} \left( \frac{k(k+1)}{r^2} - \frac{\partial^2}{\partial r^2} \right) \right] F_0
\]

then

\[
\{ F_0, F_0 \} \cong \{ F, F \} + \{ G, G \}
\]

and the equation for \( F_0 \) will be properly normalized. We therefore introduce \( F_0 \) into (5.19) but only in those terms that are not small in themselves. This leads to

\[
\left( 1 - \frac{\epsilon - \nu}{2mc^2} \right) \left[ - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m} \frac{k(k+1)}{r^2} \right] F_0 - \frac{\hbar^2}{4mc^2} \frac{\partial V}{\partial r} \left( \frac{\partial}{\partial r} + \frac{k}{r} \right) F_0 + \frac{\hbar^2}{2m} \frac{\hbar^2}{\mu^2} \left( \frac{\partial^2}{\partial r^2} - \frac{k(k+1)}{r^2} \right) \left( \frac{k(k+1)}{r^2} - \frac{\partial^2}{\partial r^2} \right) F_0 + \frac{\epsilon - \nu}{2mc^2} \frac{\hbar^2}{\mu^2} \left( \frac{k(k+1)}{r^2} - \frac{\partial^2}{\partial r^2} \right) F_0 = (\epsilon - \nu) F_0
\]

The second order term singled out by the arrow can be written (by making use of the zeroth order approximation given by Eq. 5.18) as

\[
\frac{\hbar^2}{\mu^2} \left( \frac{\partial^2}{\partial r^2} - \frac{k(k+1)}{r^2} \right) (\epsilon - \nu) F_0
\]

Also the last term on the left can be combined with a similar term in the equation, and the whole becomes:

5463/e/p
\[-\frac{\hbar^2}{2m} \left[ \frac{\partial^2}{\partial t^2} - \frac{K(k_{i+1})}{T^2} \right] F_0 + \frac{(\epsilon - V)\hbar^2}{\delta \mu^2} \left[ \frac{\partial}{\partial r} - \frac{K(k_{i+1})}{r^2} \right] F_0 = (\epsilon - V) F_0 \]

\[-\frac{\hbar^2}{4m} \frac{\partial^2}{\partial r^2} \frac{F_0}{(\epsilon - V)} + \frac{\hbar^2}{\delta \mu^2} \left[ \frac{\partial}{\partial r} \frac{K(k_{i+1})}{r^2} \right] (\epsilon - V) F_0 = (\epsilon - V) F_0 \]

We now further transform the second term as follows:

\[
\left[ \frac{\partial^2}{\partial r^2} - \frac{K(k_{i+1})}{r^2} \right] = \frac{\hbar^2}{\delta \mu^2} \left[ \frac{\partial}{\partial r} - \frac{K(k_{i+1})}{r^2} \right] (\epsilon - V) + \frac{\hbar^2}{\delta \mu^2} \left[ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \frac{K(k_{i+1})}{r^2} \right] F_0
\]

and this makes our equation read:

\[-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial t^2} - \frac{K(k_{i+1})}{r^2} \right) \left( \frac{1}{1 - \frac{\epsilon - V}{2\hbar c^2}} \right) F_0 + \frac{\hbar^2}{\delta \mu^2} \left[ \frac{\partial^2}{\partial r^2} - \frac{2}{r} \frac{\partial}{\partial r} \frac{K(k_{i+1})}{r} \right] F_0 = (\epsilon - V) F_0 \quad (5.20)
\]

The terms singled out by the arrows represent our correction terms. The first of these is simply a second-order relativistic mass correction to the kinetic energy (see p. 98). The remaining two terms can be written in two groups as follows:

\[-\frac{\hbar^2}{4m\epsilon c^2} \left( \frac{K(k_{i+1})}{r} \right) \frac{\partial}{\partial r} F_0 \quad (5.21)
\]

\[-\frac{\hbar^2}{\delta \mu^2} \left[ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right] F_0 = \frac{\hbar^2}{\delta \mu^2} (\nabla^2 V) F_0 \quad (5.22)
\]

This splitting is motivated by the following relation:

\[
\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{i}{2} \left[ (\hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2) - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2 \right] = \frac{\hbar^2}{2} \left[ \hat{\mathbf{d}} (j + \frac{3}{2}) - \epsilon (\ell + 1) \right] + \frac{\hbar^2}{2} \left[ \frac{3}{2} (j + \frac{3}{2}) - \epsilon (\ell + 1) \right] = -\frac{\hbar^2}{2} (k + 1)
\]

5463\text{/e/p}
Hence (5.12) becomes 
\[
\frac{1}{2m^2c^2} \frac{1}{r} \frac{\partial}{\partial r} \mathbf{r} \mathbf{F}_e^r \quad (\text{L.S.})
\]
which is the spin-orbit coupling term. The other term, (5.22b), is the one whose
physical explanation was given on p. 99. Since \( \nabla^2 \mathbf{V} = -4\pi \rho_s \mathbf{e} \),
where \( \rho_s \) is the source of the external field, this term will
be different from 0 only at the position of the source. It
will therefore affect the s-states since only in those states
the electron spends any appreciable length of time near the
source. Therefore, the s-states will be shifted even though
the spin-orbit coupling term vanishes in that case.

**Ground State of the Hydrogen Atom**

For the hydrogen atom (or hydrogen-like atoms) the potential
\( V \) is given by

\[
V = -\frac{Ze^2}{r}
\]

and eqs. (5.16a,b) can be solved exactly. The procedure con-
sists in solving one of the equations (5.16a,b) for \( G \) (or \( F \)),
then differentiating it and substituting the values of \( G, G' \) (or
\( F, F' \)) into the other equation. One then obtains a second
order differential equation whose solution is the confluent
hypergeometric function. We must impose on the solution cer-
tain conditions in order that it be square-integrable. These
conditions lead to the Sommerfeld formula for the allowed energy
levels:

\[
E = mc^2 \left[ 1 + \left( \frac{\alpha}{(\kappa - \alpha)^{1/2} + n - \kappa} \right)^2 \right]^{-1/2}
\]

(5.23)

where

\[
\alpha = \frac{Ze^2}{kc} \quad \kappa = |j + \frac{1}{2}| \quad n = 1, 2, \ldots \quad n > |j + \frac{1}{2}|
\]
With the substitutions \( \gamma = \frac{W}{m_c} \), \( \lambda_c = \frac{a}{m_c} \), \( \omega = \frac{2c^2}{k_c} \)

the potential becomes \( V = -\frac{\omega}{r} \) and Eqs. (5.16a,b) can be written as:

\[
F' + \frac{k}{r} F = \left( \frac{\gamma^2}{\lambda_c} + \frac{\omega}{r} \right) G
\]
\[
G' - k \frac{G}{r} = -\left( \frac{\gamma^2}{\lambda_c} + \frac{\omega}{r} \right) F
\]

(5.24)

Instead of actually solving (5.24) in the way outlined before we assume a trial solution of the form

\[
F = A + d e^{-\gamma a}
\]
\[
G = B + d e^{-\gamma a}
\]

(5.25)

The form (5.25) for the solutions is suggested by the asymptotic behaviour of (5.24) as \( r \to \infty \). Setting \( \frac{B}{A} = s \) and substituting (5.25) into (5.24) leads to four equations for \( s, d, s \) and \( \gamma \).

Solving those we obtain for \( d \):

\[
d = \pm \sqrt{k^2 - \alpha^2}
\]

and we take \( d = + \sqrt{k^2 - \alpha^2} \) in order that the solutions be finite at the origin. We then obtain for \( s, \alpha \), and \( \gamma \):

\[
\gamma = \frac{1}{k} \sqrt{k^2 - \alpha^2}
\]
\[
s = \frac{\alpha}{\sqrt{k^2 - \alpha^2}}
\]
\[
\beta = \frac{2c^2}{k_c} \frac{k}{\alpha}
\]

We now recall that \( \lambda + 1 \) was shown to be the eigenvalue of \(-2 \hat{L} \cdot \hat{S} \). For the ground state of hydrogen \( L.S = 0 \), hence we must have \( \lambda = -1 \) which gives:
\[-119-\]

\[ S = -\frac{1 - \frac{1}{\alpha^2}}{\alpha} \approx -\frac{\alpha}{2} \quad \text{for} \quad \alpha \ll 1 \]

\[ \gamma \approx \frac{1}{\alpha^2} \quad \gamma \approx 1 - \frac{\alpha^2}{2} \quad \gamma \]

\[ \lambda = \frac{2}{\sqrt{1 - \alpha^2}} \quad \lambda \approx 1 - \frac{\alpha^2}{2} \quad \lambda \]

\[ \alpha = \frac{\alpha}{\alpha} \]

Recalling that by the ground state we mean the ground state of the large spinor \( \psi \) we have from (5.11) \( \psi = f(r) \frac{\gamma}{2}^{1/2} \frac{m}{3m} \) and \( \psi = -ig(r) \frac{1}{2} \frac{1}{2}, m \). In particular if we take \( m = +1/2 \), we have

\[
\begin{align*}
\gamma^{\dagger} \frac{1}{2} \frac{1}{2} \gamma &= \begin{pmatrix} \gamma_{0}^{0} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{4\pi}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\end{align*}
\]

\[
\begin{align*}
\gamma^{\dagger} \frac{1}{2} \frac{1}{2} \gamma &= \begin{pmatrix} \gamma_{0}^{1} \\ \gamma_{1}^{0} \end{pmatrix} = \frac{1}{\sqrt{4\pi}} \begin{pmatrix} \frac{\gamma_{0}}{\gamma_{1}} \\ \sin \psi \exp(\psi) \end{pmatrix} \\
\end{align*}
\]

hence finally for \( j = 1/2, m = 1/2 \).

\[
\begin{align*}
\psi &= N \frac{r^{d-1} \exp(\gamma \alpha)}{\sqrt{4\pi}} \begin{pmatrix} 1 \\ 0 \\ \frac{\gamma_{0}}{\gamma_{1}} \\ \sin \psi \exp(\psi) \end{pmatrix} \\
\end{align*}
\]

(N = normalization factor)

(5.26)

and one obtains a similar expression for \( j = 1/2, m = -1/2 \). From (5.26) we see that the ground state is not entirely an s-state as non-relativistic theory would have it, but contains a little of the P-state. However the probability density
$\{\Psi, \overline{\Psi}\}$ is still spherically symmetric.

Before leaving this subject we point out an interesting hierarchy of lengths

$$r_o = \alpha \frac{\hbar}{c}, \quad \lambda_c = \alpha c, \quad a = \alpha \frac{\hbar}{2c}$$

where $\alpha = \frac{\epsilon^2}{\hbar c} = \frac{1}{137}$ = fine structure constant

$r_o = \frac{\epsilon^2}{mc^2}$ = classical radius of the electron

$\lambda_c = \frac{\pi}{mc}$ = Compton wavelength

$a = \frac{\hbar^2}{ma^2}$ = Bohr radius

$\lambda_R = \frac{2 \hbar^3 c}{me^4}$ = Rydberg wavelength = $\frac{\hbar c}{E_R}$, where $E_R$ is the energy of ionization of the hydrogen atom.

We consider the problem of scattering of an electron beam by a potential $V(\vec{r})$. We require $V(\vec{r})$ to be a static potential and to be localized sufficiently well so that $V(r) \to 0$ as $r \to \infty$ at least as fast as $1/r$. Except for the above restrictions, the potential $V(\vec{r})$ can be anything at all. Before proceeding with the relativistic calculation, we review briefly what is meant by the Born approximation in the nonrelativistic case.

Born Approximation in Schroedinger Formalism

We have the Schroedinger equation

$$\nabla^2 \psi + \frac{\hbar^2}{2m} \psi = \frac{2m}{\hbar^2} V(\vec{r}) \psi \quad (6.1)$$

where $\hbar^2 = (2m/\kappa^2) \varepsilon$, $\varepsilon > 0$. Let $L(\vec{r}) = \frac{2m}{\hbar^2} V(\vec{r}) \psi$

Then (6.1) can be written in integral form (see, e.g. Bohm, p.543) as

$$\psi(\vec{r}) = -\frac{1}{4\pi} \int \frac{e^{i \frac{\hbar}{\kappa^2} |\vec{r} - \vec{r}'|}}{|\vec{r} - \vec{r}'|} L(\vec{r}') d\vec{r}' + \psi_0(\vec{r}) \quad (6.2)$$

where $\psi_0(\vec{r})$ is the solution of (6.1) when $V(\vec{r}) = 0$. It would seem that (6.2) gives us the required solution. However, (6.2) gives $\psi(\vec{r})$ in terms of $L(\vec{r})$ which in turn contains $\psi(\vec{r}')$. We can get around this difficulty by guessing what $\psi$ ought to be and putting that expression into $L(r')$. Then the value of $\psi$ obtained by solving (6.2) will result in a better estimate of the actual correct $\psi$. This new value of $\psi$ can again be put into $L(r')$, and the
procedure repeated. The original guess at the value of \( \psi \) need not be very good; it must, however, be such that the process described above be a convergent one. In the Born approximation one replaces \( \psi \) in \( \mathcal{L}(\vec{r}) \) by the solution \( \psi_0 \) valid in the absence of the potential. This amounts to assuming that the incident plane wave \( \psi_0(\vec{r}) = e^{i \vec{k}_0 \cdot \vec{r}} \) representing the incoming beam of electrons is not distorted by the presence of the potential -- an assumption which is reasonably good if the energy of the incoming beam is large compared to the scattering potential, or if the potential is sufficiently weak. The + and - signs in the exponential in (6.2) represent outgoing and incoming waves, respectively. Hence, if we are interested in solutions at large values of \( \vec{r} \), we take the positive exponential. Furthermore, the contributions to the integral in (6.2) come from limited values of \( \vec{r}' \) only. This is so because we have assumed \( V(\vec{r}') \to 0 \) as \( r' \to \infty \), hence \( L(\vec{r}') \to 0 \) as \( \vec{r}' \to \infty \), since \( \psi(\vec{r}') \) remains finite. Therefore, for large values of \( \vec{r} \), we have approximately

\[
|\vec{r} - \vec{r}'| \approx \gamma - \vec{\kappa} \cdot \vec{r}'
\]

For large \( \vec{r} \) we have \( \vec{\kappa} \cdot \vec{r}' \ll \gamma \) and we can replace \( |\vec{r} - \vec{r}'| \) by \( r \) in the denominator of (6.2). In the exponential we must keep \( |\vec{r} - \vec{r}'| \approx \gamma - \vec{\kappa} \cdot \vec{r}' \). Hence in the Born approximation and for large values of \( r \), (6.2) becomes

\[
\psi(\vec{r}) \xrightarrow{r \to \infty} e^{i \vec{k}_0 \cdot \vec{r}} - \frac{2m}{\hbar^2} \frac{e^{i \vec{\kappa} \cdot \vec{r}}}{4\pi r} \int e^{i (\vec{k}_0 - \vec{\kappa}) \cdot \vec{r}'} V(\vec{r}') d\vec{r}' (6.3)
\]
where $\hat{\mathbf{b}} = \mathbf{r} \hat{\mathbf{n}}$.

The scattering cross-section $\sigma(\theta, \varphi)$ per unit solid angle is given by the square of the amplitude of the outgoing wave, i.e. by the square of the coefficient of $\frac{e^{i\mathbf{k}_0 \cdot \mathbf{r}}}{r}$ in (6.3).

Hence

$$\sigma(\theta, \varphi) = \left[\frac{2m}{\hbar^2} \frac{1}{4\pi} V(\hat{\mathbf{k}}_0 - \hat{\mathbf{k}})\right]^2$$  \hspace{1cm} (5.4)

where

$$V(\hat{\mathbf{k}}_0 - \hat{\mathbf{k}}) \equiv \int e^{i(\hat{\mathbf{k}}_0 - \hat{\mathbf{k}}) \cdot \mathbf{r}} V(\mathbf{r}) d\mathbf{r}$$

In particular, for a Coulomb potential $V = \frac{Z e^2}{r}$ and

$$V(\hat{\mathbf{k}}_0 - \hat{\mathbf{k}}) = \frac{4\pi Z e^2}{|\hat{\mathbf{k}}_0 - \hat{\mathbf{k}}|^2} = \frac{-Z e^2}{\mathbf{k}_0 \cdot \mathbf{n} \sin^{\frac{3}{2}} \varphi}$$

which, when introduced in (6.4), gives the $\sin^{-\frac{3}{2}} \varphi$ dependence given by Rutherford law.

**Born Approximation in Dirac's Formalism**

We now proceed to find the scattering cross-section starting from Dirac's equation

$$(\mathbf{H}_0 - E) \Psi = V \Psi \quad \text{where} \quad \mathbf{H}_0 = c \alpha \cdot \mathbf{p} + \beta m c^2.$$

We can make the left hand side of above equation look like the left hand side of (6.1) by operating from the left with $\mathbf{H}_0 + E$. We then obtain

$$\left(\mathbf{H}_0^2 - E^2\right) \Psi = (\mathbf{H}_0 + E) V \Psi$$

or

$$\left(\nabla^2 + \mathbf{k}^2\right) \Psi = -\frac{(\mathbf{H}_0 + E)}{\hbar^2 c^2} V \Psi$$  \hspace{1cm} (6.5)
where \( \hat{p}^2 = \frac{E^2 - m^2 c^4}{\hbar^2 c^2} = \frac{p^2}{\hbar^2} \). Letting \( \Psi(\vec{r}) = \frac{(H_0 + E) V(\vec{r})}{\hbar^2 c^2} \Psi(\vec{r}) \),

we obtain by analogy with the nonrelativistic case

\[
\Psi \xrightarrow{\vec{r} \to \infty} \Psi_0 + \int e^{-i \vec{k} \cdot \vec{r}'} \, L(\vec{r}) d\vec{r}'
\]

(6.6)

where \( \Psi_0 \) is a solution of (6.5) with \( V(\vec{r}) = 0 \).

Omitting the time dependence, we have from (2.27):

\[
\Psi_0 = \gamma_0(\vec{p}_0) e^{i \frac{\vec{p}_0 \cdot \vec{r}}{\hbar}}
\]

where \( \gamma_0 \) is some linear combination of the four \( \gamma^{(i)} \) properly normalized. In particular, if the electrons in the incident beam are all in the positive energy states, \( \gamma_0 \) will be a linear combination of \( \gamma^{(1)} \) and \( \gamma^{(2)} \). Calling the coefficient of \( \frac{e^{i \vec{k} \cdot \vec{r}}}{\vec{r}} \) in eq. (6.6) \( f \), we have for large \( \vec{r} \) in Born approximation:

\[
f = \frac{i}{4\pi} \int e^{-i \vec{p}_0 \cdot \vec{r}'} \frac{H_0 + E}{\hbar^2 c^2} V(\vec{r}) \gamma_0(\vec{p}_0) e^{i \frac{\vec{p}_0 \cdot \vec{r}'}{\hbar}} d\vec{r}'
\]

(6.7)

The part of the Hamiltonian \( H_0 \) that does not commute with \( V(\vec{r}) \) is \( c \vec{d} \cdot \vec{p}' \), where \( \vec{p}' = \frac{\vec{k}}{\alpha} \vec{V}' \); hence

\[
H_0 V(\vec{r}) = V(\vec{r}) H_0 + c (\vec{d} \cdot \vec{p}') V(\vec{r})
\]

\[
= V(\vec{r}) H_0 + \frac{c\hbar}{\alpha} \vec{d} \cdot \vec{V} V(\vec{r})
\]

(6.8)

We then get

\[
f = \frac{i}{4\pi} \int e^{-i \vec{p}_0 \cdot \vec{r}'} \frac{V(\vec{r}) (H_0 + E)}{\hbar^2 c^2} \gamma_0(\vec{p}_0) e^{i \frac{\vec{p}_0 \cdot \vec{r}'}{\hbar}} d\vec{r}'
\]

\[
+ \frac{i}{4\pi} \int e^{-i \frac{\vec{p}_0 \cdot \vec{r}'}{\alpha} \vec{d} \cdot \vec{V} V(\vec{r}') \gamma_0(\vec{p}_0) e^{i \frac{\vec{p}_0 \cdot \vec{r}'}{\hbar}} d\vec{r}'
\]

(6.9)
We find, however,

\[ \int e^{i (\vec{k}_o - \vec{k} \cdot \hat{\tau} ) \cdot \hat{\tau} \cdot } \nabla (\vec{\tau}) \cdot \hat{\tau} \cdot \cdot d\tau \cdot = -i (\vec{k}_o - \vec{k} \cdot \hat{\tau} ) \int e^{i (\vec{k}_o - \vec{k} \cdot \hat{\tau} ) \cdot \hat{\tau} \cdot } \nabla (\vec{\tau}) \cdot \hat{\tau} \cdot \cdot d\tau \cdot \cdot \]  \quad (6.10)

Furthermore:

\[ H_o \ \gamma_o (\vec{p}_o) \ e^{i \vec{k}_o \cdot \hat{\tau} \cdot } = \ (c \ \vec{x} \cdot \vec{p}_o + \beta \ mc^2) \ \gamma_o (\vec{p}_o) \ e^{i \vec{k}_o \cdot \hat{\tau} \cdot } \]

\[ = H_o (\vec{p}_o) \ \gamma_o (\vec{p}_o) \ e^{i \vec{k}_o \cdot \hat{\tau} \cdot } \]  \quad (6.11)

Here \( H_o (\vec{p}_o) = c \ (\vec{x} \cdot \vec{p}_o) + \beta \ mc^2 \). This is a spinor operator only, and \( \vec{p}_o \) is not an operator but the actual momentum of the initial state. (In contrast to this \( H_o = \frac{c^2}{\hbar^2} (\vec{x} \cdot \vec{v}) + \beta \ mc^2 \).)

Introducing (6.10) and (6.11) into (6.9), we encounter the following combination:

\[ H_o (\vec{p}_o) - c \ \vec{k} \ \vec{x} \cdot (\vec{k}_o - \vec{k}) = c (\vec{x} \cdot \vec{p}_f ) + \beta \ mc^2 = H_o (\vec{p}_f ) \]

where \( \vec{p}_f = \vec{k} \cdot \vec{x} \cdot \vec{k} \) is the momentum after scattering.

We then get

\[ f = \frac{1}{4\pi} \int e^{i \vec{k}_o \cdot \hat{\tau} \cdot } \nabla (\vec{\tau}) \ \frac{H_o (\vec{p}_f ) + E}{\hbar^2 \ c^2} \ \gamma_o (\vec{p}_o) \ e^{i \vec{k}_o \cdot \hat{\tau} \cdot } \cdot d\tau \cdot \cdot \]  \quad (6.12)

\[ = \frac{1}{4\pi} \ \nabla (\vec{k}_o - \vec{k} \cdot \hat{\tau} \cdot ) \ \frac{H_o (\vec{p}_f ) + E}{\hbar^2 \ c^2} \ \gamma_o (\vec{p}_o) \]

It is of interest to express \( f \) in terms of \( \gamma (\vec{p}_f ) \) rather than \( \gamma_o (\vec{p}_o) \), where the \( \gamma (\vec{p}_f ) \) are the spinor amplitudes of the four solutions of a free electron with a momentum \( \vec{p}_f \).

We write \( \gamma_o (\vec{p}_o) \) in terms of \( \gamma (\vec{p}_f ) \) as follows:
We recall that  

\[ \gamma^{(i)}(\hat{p}) = \mathcal{D}(\hat{p}) \gamma^{(i)(0)} \]  

hence  

\[ \gamma^{(i)}(\hat{p}_o) = \mathcal{D}(\hat{p}_o) \gamma^{(i)(0)} = \mathcal{D}(\hat{p}_o) \mathcal{D}^{-1}(\hat{p}_f) \gamma^{(i)}(\hat{p}_f) \]

Since \( \mathcal{D}(\hat{p}_o) \mathcal{D}^{-1}(\hat{p}_f) \gamma^{(i)}(\hat{p}_f) \) is simply some linear combination of the four four-spinors \( \gamma^{(i)}(\hat{p}_f), i = 1, 2, 3, 4 \) we can write  

\[ \gamma_o(\hat{p}_o) = \sum_{k=1}^{4} a_{ki} \gamma^{(k)}(\hat{p}_f) \]  

(6.13)

where we have assumed for simplicity that the incident beam is described by just one basis four-spinor \( \gamma^{(i)}(\hat{p}_o) \) rather than a linear combination of several. Since we are considering elastic scattering, we have  

\[ E(\hat{p}_o) = E(\hat{p}_f) = E \]  

(6.14)

Making use of (6.13) and (6.14), we rewrite part of (6.12) as follows:

\[ \bigg[ H_0(\hat{p}_f) + E \bigg] \gamma_o(\hat{p}_o) = \bigg[ H_0(\hat{p}_f) + E \bigg] \sum_{k=1}^{4} a_{ki} \gamma^{(k)}(\hat{p}_f) \]

\[ = (E + E) \sum_{k=1}^{2} a_{ki} \gamma^{(k)}(\hat{p}_f) + (E - E) \sum_{k=3}^{4} a_{ki} \gamma^{(k)}(\hat{p}_f) \]

\[ = 2E \bigg[ a_{1i} \gamma^{(1)}(\hat{p}_f) + a_{2i} \gamma^{(2)}(\hat{p}_f) \bigg] \]  

(6.15)

Hence equation (6.12) now becomes  

\[ f = \frac{1}{4\pi} \frac{2E}{\hbar c^2} \sqrt{(\hat{p}_o - \hat{p}_0)} \bigg[ a_{1i} \gamma^{(1)}(\hat{p}_f) + a_{2i} \gamma^{(2)}(\hat{p}_f) \bigg] \]  

(6.16)
We thus have for the scattered wave at large r:
\[
\frac{e^{i \frac{2 \pi}{r} \vec{k} \cdot \vec{r}}}{4\pi r} \nabla \cdot \left( \frac{\vec{k}}{k^2} \cdot \vec{r} \right) \frac{2E}{k^2 c^2} \sum_{k=1}^{2} \alpha_{ki} \gamma^{(k)}(\vec{r})
\]
(6.17)

and the scattering cross-section becomes
\[
\sigma(\theta, \phi) = \sum_{k} \sigma_{ki}
\]
(6.18a)

where \( \sigma_{ki} \) represents the cross-section for the spin going from the state \( \gamma^{(i)}(\vec{p}_0) \) to \( \gamma^{(k)}(\vec{p}_f) \) and is given by
\[
\sigma_{ki} = \left| \frac{\nabla \cdot \left( \frac{\vec{k}}{k^2} \cdot \vec{r} \right) \frac{E}{k^2 c^2} \alpha_{ki} \right|^2
\]
(6.18b)

We now evaluate the coefficients \( \alpha_{ki} \). We first note that since the \( \gamma^{(i)}(\vec{p}) \) form an orthonormal set, we have
\[
\left\{ \gamma^{(i)}(\vec{p}_f), \gamma^{(k)}(\vec{p}_f) \right\} = \delta_{ik}
\]
We therefore form the inner product of \( \gamma^{(j)}(\vec{p}_f) \) with both sides of equation (6.13) and obtain:
\[
\left\{ \gamma^{(d)}(\vec{p}_f), \gamma^{(i)}(\vec{p}_0) \right\} = \alpha_{ji} \quad j = 1, 2, 3, 4
\]
(6.19)

or, replacing \( j \) by \( k \):
\[
\alpha_{ki} = \left\{ \mathcal{D}(\vec{p}_f) \gamma^{(k)}(\vec{0}), \mathcal{D}(\vec{p}_0) \gamma^{(i)}(\vec{0}) \right\} = \\
= \left\{ \gamma^{(k)}(\vec{0}), \mathcal{D}^+(\vec{p}_f) \mathcal{D}(\vec{p}_0) \gamma^{(i)}(\vec{0}) \right\} = \\
= (\mathcal{D}^+(\vec{p}_f) \mathcal{D}(\vec{p}_0))_{ki} \text{- matrix element}
\]
Since $W$ is the same for both $D^+(\hat{\rho}_f)$ and $D(\hat{\rho}_p)$, we have

$$D^+(\hat{\rho}_f)D(\hat{\rho}_p) = \frac{W+\mu}{2W} \left( 1 + \frac{\beta \hat{\omega} \cdot \hat{\rho}_f}{W+\mu} \right) \left( 1 - \frac{\beta \hat{\omega} \cdot \hat{\rho}_p}{W+\mu} \right)$$

$$= \frac{W+\mu}{2W} \left( 1 + \frac{\beta \hat{\omega} \cdot \hat{\rho}_f}{W+\mu} - \frac{\beta \hat{\omega} \cdot \hat{\rho}_p}{W+\mu} - \frac{\beta (\hat{\omega} \cdot \hat{\rho}_f) (\hat{\omega} \cdot \hat{\rho}_p)}{(W+\mu)^2} \right)$$

We are interested only in $\alpha_{\kappa i}$ for $\kappa = 1, 2$ and $i = 1, 2$ and for those values it turns out that

$$\left( \beta \hat{\omega} \right)_{\kappa i} = \left( \beta \hat{\omega} \right)_{\kappa i} = (\beta \hat{\omega})_{\kappa i} = 0$$

and since $\beta (\hat{\omega} \cdot \hat{\rho}_f) \beta (\hat{\omega} \cdot \hat{\rho}_p) = -\left( \hat{\omega} \cdot \hat{\rho}_f \right) \left( \hat{\omega} \cdot \hat{\rho}_p \right)$ we have

$$\alpha_{\kappa i} = \left( D^+(\hat{\rho}_f)D(\hat{\rho}_p) \right)_{\kappa i} = \frac{W+\mu}{2W} \left( 1 + \frac{\hat{\omega} \cdot \hat{\rho}_f}{(W+\mu)^2} \right) \kappa i$$

$$= \frac{W+\mu}{2W} \left( 1 + \frac{\hat{\omega} \cdot \hat{\rho}_f}{(W+\mu)^2} + i \sigma \cdot \hat{\rho}_f \times \hat{\rho}_p \right) \kappa i \quad (6.20)$$

for $\kappa$ and $i$ equal 1 and 2. In obtaining (6.20), we have made use of the identity

$$\left( \hat{\omega} \cdot \hat{A} \right) \left( \hat{\omega} \cdot \hat{B} \right) = \hat{A} \cdot \hat{B} + i \hat{\sigma} \cdot \hat{A} \times \hat{B}$$

where $A$ and $B$ commute with $\hat{\omega}$ but not necessarily with each other, and $\hat{\sigma}$ is the extended Pauli $\hat{\sigma}$. For convenience in notation, we now introduce a unit vector $\hat{n}$ and the symbol $P$ by:

$$\hat{n} = \frac{\hat{\rho}_f \times \hat{\rho}_p}{|\hat{\rho}_f \times \hat{\rho}_p|}, \quad P = \frac{W-\mu}{W}$$

5463/o/p
then
\[
\frac{W+\mu}{IZ} = 1 - \frac{W-\mu}{IZ} = 1 - \frac{\mathcal{P}}{2}
\]
\[
\frac{W+\mu}{IZ} \frac{\vec{p}_f \cdot \vec{p}_0}{(W+\mu)^2} = \frac{\mathcal{P}^2 \cos \theta}{IZ (W+\mu)} = \frac{\mathcal{P}}{2} \sigma_3 \theta
\]
\[
\frac{W+\mu}{IZ} i \frac{\vec{r}_f \times \vec{r}_0}{(W+\mu)^2} = \frac{\mathcal{P}}{2} i \sigma_3 \sin \theta
\]

We can now write the matrix \( \mathcal{A} \) whose elements are \( a_{ki} \) \( (k, i = 1 \text{ or } 2) \) as

\[
\mathcal{A} = 1 - \frac{\mathcal{P}}{2} \left[ 1 - \cos \theta - i \sigma_3 \sin \theta \right]
\]
\[
= 1 - \mathcal{P} + \frac{\mathcal{P}}{2} \left[ 1 + \cos \theta + i \sigma_3 \sin \theta \right]
\]
\[
= \frac{\mu}{W} + \mathcal{P} \cos \frac{\Theta}{2} \left( \cos \frac{\Theta}{2} + i \sigma_3 \sin \frac{\Theta}{2} \right)
\]
\[
= \frac{\mu}{W} + \mathcal{P} \cos \frac{\Theta}{2} e^{i \sigma_3 \frac{\Theta}{2}}
\]

(6.21)

where \( \sigma_3 \) stands for the component of \( \vec{\sigma} \) on the unit vector \( \vec{n} \) and is just the regular 2x2 Pauli matrix (since we have restricted ourselves to \( k, i = 1 \text{ or } 2 \)). The term \( \mathcal{P} \cos \frac{\Theta}{2} e^{i \sigma_3 \frac{\Theta}{2}} \) shows that the effect of \( \mathcal{A} \) on the spinors is that of rotating them by an angle \( \Theta \) around an axis perpendicular to the plane of the initial and final momentum.
So far we have not specified the orientation of our co-ordinate system. It would be convenient to choose it so that \( A \) turns out to be a diagonal matrix. Since \( \Omega \) is diagonal, we therefore choose the z-axis to be parallel to \( \hat{n} \). Since the direction of \( n \) depends on the direction of \( \hat{\Omega}_f \) we will have a different z-axis for every \( \hat{P}_f \) - this should not be disturbing since we can always express our final answers in terms of a single co-ordinate system by performing appropriate rotations. We then find for the elements of \( \hat{\Omega} \):

\[
\begin{align*}
a_{11} &= \frac{\mu}{W} + \frac{P}{2} (1 + \cos \varphi + i \sin \varphi) \\
a_{22} &= \frac{\mu}{W} + \frac{P}{2} (1 + \cos \varphi - i \sin \varphi) \\
a_{12} &= a_{21} = 0
\end{align*}
\]

(6.22)

From this we see immediately that in the nonrelativistic case \( a_{11} = 1, a_{22} = 1 \), because \( \mu \ll W, P \ll 0; \) hence we obtain

\[
\hat{\Omega} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\]

is the unit matrix. On the other hand, in the ultrarelativistic case, when \( \mu \) can be neglected in comparison to \( W \), we have \( \mu/W \ll 1 \), \( P = \frac{W - \mu}{W} \approx 1 \) and the matrix

reduces to

\[
\hat{\Omega} = \cos \frac{\varphi}{2} e^{i \frac{\varphi}{2} \hat{n} \cdot \hat{\xi}}
\]

\( i463/e/p \)
Thus, in the nonrelativistic case, the spin does not change direction on scattering, while in the ultrarelativistic case, the spin is turned around so as to form the same angle with the direction of motion before and after scattering. When we talk about the direction of the spin, we are always referring to its direction in the electron's rest frame, hence we are talking only about the "large" spinor $\psi$ (the components of $\psi$ are 0 in that frame of reference.)

So far we have been assuming that the electron before scattering is described by $\gamma_i(\vec{p}_0)$ or $\gamma_i(\vec{p}_o)$. It could, of course, be a linear combination

$$\gamma_o(\vec{p}_o) = \sum_{i=1}^{2} \beta_i \gamma_i(\vec{p}_o)$$

(6.23)

which we assume normalized so that $|\beta_1|^2 + |\beta_2|^2 = 1$. We then proceed in the same fashion as before for each $\gamma_i(\vec{p}_o)$, multiply the results by the corresponding $\beta_i$ and add. For each $\gamma_i(\vec{p}_o)$, the scattering cross section is given by (see eq.6.18a, 6.18b)

$$\sigma_i(S, \phi) = \left| \frac{\nu(\vec{k}_o - \vec{k})}{4\pi} \right|^2 \frac{2E}{\kappa^2 c^2} \left| \sum_{k=1}^{2} a_{k,i} \right|^2$$

hence the total scattering cross section is

$$\sigma(\vec{S}, \phi) = \left| \frac{\nu(\vec{k}_o - \vec{k})}{4\pi} \right|^2 \frac{2E}{\kappa^2 c^2} \left| \sum_{k=1}^{2} \sum_{i=1}^{2} a_{k,i} \beta_i \right|^2$$

(6.24)
The sum over $i$ collapses to a single term since $\alpha_k \neq 0$ only for $k = i$. Furthermore, $|a_{kl}|^2 = |a_{2l}|^2$ as is obvious from eq. (6.22). Hence

$$
\sum_{k=1}^{2} \left| \sum_{l=1}^{2} a_{ki} \beta_l \right|^2 = \sum_{k=1}^{2} |a_{kk} \beta_k|^2 = |a_{kk}|^2 \sum_{k=1}^{2} |\beta_k|^2 = |a_{kk}|^2
$$

which shows that our result is independent of the $\beta_l$. That means that the scattering is independent of the state of polarization of the incoming beam. This is not true in general, but is characteristic of the Born approximation. Noting that

$$
|a_{kk}|^2 = \left| 1 - \frac{\hbar^2}{\mathcal{K}^2} \sin^2 \frac{\theta}{2} \right|
$$

we can write eq. (6.24) as

$$
\sigma(\theta, \varphi) = \left| \frac{\mathcal{V}(\mathcal{R}_0 - \mathcal{K}) \frac{2E}{k^2 c^2}}{\frac{\hbar^2}{\mathcal{K}^2} \sin \frac{\theta}{2}} \right|^2 \left( 1 - \frac{\hbar^2}{\mathcal{L}^2} \sin^2 \frac{\theta}{2} \right) \quad (6.25)
$$

Comparing (6.25) with (6.4), we can summarize our results by writing:

$$
\frac{\sigma_{\text{relativistic}}}{\sigma_{\text{nonrel.}}} = \left( \frac{E}{m c^2} \right)^2 \left( 1 - \frac{c^2}{E^2} \sin^2 \frac{\theta}{2} \right) \quad (6.26)
$$

which obviously reduces to unity in the nonrelativistic limit when $E \approx mc^2$ and $\frac{\hbar^2}{\mathcal{K}^2} \ll 1$. Eq. (5.26) also shows that besides the overall difference between $\sigma_{\text{rel}}$ and $\sigma_{\text{nonrel}}$, due to $\left( \frac{E}{m c^2} \right)^2$, the two cross sections also differ in their dependence on $\theta$. Since $\sin^2 \frac{\theta}{2}$ increases as $\theta$ increases from 0 (forward scattering) to $\pi$ (backward scattering), we see that as we go backwards, the relativistic cross section falls off faster than the nonrelativistic one. This is the
known relativistic effect of "shoving" things forward.

Before closing this section on scattering, we would like to mention a somewhat different way of expressing our results. We have shown that at large distances from the scattering centre we have

$$\psi \rightarrow \eta^{(i)}(\hat{p}_o) e^{\frac{i\hat{p}_o \cdot \hat{r}}{r}} + \frac{e^{i\hat{q} \cdot \hat{r}}}{r} \sum_{\kappa} f_{k\kappa} \eta^{(\kappa)}(\hat{p}_p)$$  \hspace{1cm} (6.27)

where

$$f_{k\kappa} = \frac{1}{4\pi} \frac{2E}{k^2 c^2} \left\langle \hat{r} \left| V(\hat{p}_o - \hat{p}_f) \right| \hat{r} \right\rangle \alpha_{k\kappa}$$

$$= \frac{1}{4\pi} \frac{2E}{k^2 c^2} \int e^{i\hat{p}_o \cdot \hat{r}} \left\langle \hat{r} \left| V(\hat{p}_o - \hat{p}_f) \right| \hat{r} \right\rangle \left\langle \eta^{(\kappa)}(\hat{p}_f) \right| \eta^{(k)}(\hat{p}_o) \right\rangle d^3r$$

$$= \frac{1}{4\pi} \frac{2E}{k^2 c^2} \sum_{\nu \sigma} \int e^{i\hat{p}_o \cdot \hat{r}} \left\langle \nu \left| V(\hat{p}_o - \hat{p}_f) \right| \sigma \right\rangle \left\langle \nu \left| \eta^{(k)}(\hat{p}_o) \right| \sigma \right\rangle e^{i\hat{p}_f \cdot \hat{r}} d^3r$$  \hspace{1cm} (6.28)

$$= \frac{1}{4\pi} \frac{2E}{k^2 c^2} \left( \Psi_f , V(\hat{r}) \Psi_o \right)$$

But \( \left( \Psi_f , V(\hat{r}) \Psi_o \right) \) is simply the matrix element of \( V \) (which is treated as a perturbation) between the final and initial states, i.e. between the outgoing and incoming plane spinor waves. To emphasize this, the following notation is also used

$$\left( \Psi_f , V(\hat{r}) \Psi_o \right) = \left( \hat{p}_f , \kappa \right| V \left| \hat{p}_o , i \right)$$

where \( k \) and \( i \) indicate which of the basis spinors describes the final and initial momentum. In obtaining (6.28) we have assumed the initial state to be described by just \( \eta^{(i)} \) - the generalization to a linear combination of \( \eta^{(i)} \) and \( \eta^{(a)} \) is quite straightforward.
Polarization in the Scattered Beam

We have seen that in the Born approximation, the scattering cross-section is independent of the way in which the incident beam was made up from the two spinors \( \gamma_i(\hat{p}_o) \) (\( i = 1, 2 \)). The final result was independent of the \( \delta_i \) 's in (6.23). This is not true when one uses exact solutions. In what follows, we will derive some general rules as to the dependence of the solutions upon the polarization of the incident beam. These rules are exact.

We restrict ourselves to an electrostatic scattering field which is a central field. We will investigate the angular dependence of the solutions and we will consider only the large spinors \( \gamma' \) since the small ones are determined by the large ones. (Far from the scattering centre all waves are waves of free particles, for which the small components are well known functions of the large ones.)

Let us first consider the solution of our scattering problem when the incident beam has its spin in the positive \( z \)-direction. (The \( z \)-direction is the direction of incidence). We then write

\[
\begin{align*}
\gamma_+^{\pm} & \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{\pm i k z} + \frac{e^{i k r}}{r} \begin{pmatrix} f(\vartheta, \varphi) \\ g(\vartheta, \varphi) \end{pmatrix} \tag{6.29}
\end{align*}
\]

Here \( f \) and \( g \) are the two spinor components of the out-going wave. We can determine their \( \varphi \)-dependence from the following general
considerations: Suppose we should rotate our whole system by an angle $f^3$ around the z-axis. (We leave our co-ordinate system fixed). Then $f(\vartheta, \varphi)$ goes over into $f(\vartheta, \varphi - f^3)$ and $g(\vartheta, \varphi)$ goes over into $g(\vartheta, \varphi - f^3)$. Besides this, since we are dealing with a spinor field, we must multiply (6.29) by the matrix

$$R(f^3) = \begin{pmatrix} e^{-i f^3} & 0 \\ 0 & e^{i f^3} \end{pmatrix}$$

With all these changes, (6.29) becomes

$$\psi' \sim_{2^+} \rightarrow \begin{pmatrix} e^{-i f^3} \\ 0 \end{pmatrix} e^{ikz} + \frac{e^{ikx}}{r} \begin{pmatrix} e^{-i f^3} f(\vartheta, \varphi - f^3) \\ e^{i f^3} g(\vartheta, \varphi - f^3) \end{pmatrix}$$

(6.30)

Certainly $\psi'$ is also a solution of our scattering problem since the potential is invariant to this rotation. However, if $\psi'$ is a solution of our problem, then $\psi'' = e^{i f^3} \psi'$ also must be a solution:

$$\psi'' \sim_{2^+} \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{ikz} + \frac{1}{r} \begin{pmatrix} f(\vartheta, \varphi - f^3) \\ e^{i f^3} g(\vartheta, \varphi - f^3) \end{pmatrix}$$

(6.31)

Comparing (6.31) with (6.29), we note that the incident wave has the same form in both of them. Therefore, the scattered wave as given in $\psi''_{2^+}$ must be the same as that given in $\psi'_{2^+}$ since the incident waves are the same. This means that

$$f(\vartheta, \varphi - f^3) = f(\vartheta, \varphi)$$

$$g(\vartheta, \varphi - f^3) = g(\vartheta, \varphi)$$

or

$$g(\vartheta, \varphi - f^3) = e^{i f^3} g(\vartheta, \varphi)$$
We therefore conclude that \( f \) must be independent of \( \beta \) whereas \( g \) depends on \( \beta \) through the exponential. Setting \( \varphi = 0 \) and \( -\beta = \varphi \) to return to a more conventional notation we conclude that

\[
\begin{align*}
  f &= f(\gamma) \\
  g &= g(\gamma) e^{i\varphi}
\end{align*}
\]

(6.32)

where \( f(\gamma) \), \( g(\gamma) \), are functions of \( \gamma \) only. We now have the \( \varphi \)-dependence for the incident wave \((|0\rangle)\):

\[
\psi_{2+} \underset{\varphi \to \infty}{\longrightarrow} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{ikz} + \frac{e^{ikr}}{r} \begin{pmatrix} f(\gamma) \\ e^{i\varphi} g(\gamma) \end{pmatrix}
\]

(6.33)

We now ask for the \( \varphi \)-dependence when the incident spinor is \((|0\rangle)\).

We proceed as follows: we first apply the space inversion (reflection) operation to \( \psi_{2+} \). This operation (see p. 108) does not change the components of a spinor, but it changes

\[
\begin{align*}
  z &\longrightarrow -z \\
  \gamma &\longrightarrow -\gamma \\
  \varphi &\longrightarrow \pi + \varphi \\
  \varphi &\longrightarrow \varphi
\end{align*}
\]

\[
\begin{align*}
  (\psi_{2+})_{\text{inverted}} \underset{\varphi \to \infty}{\longrightarrow} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-ikz} + \frac{e^{ikr}}{r} \begin{pmatrix} f(\pi-\gamma) \\ e^{i(\pi+\varphi)} g(\gamma-\pi) \end{pmatrix}
\end{align*}
\]

463/0/0
Now let us rotate above wavefunction around the x-axis by $\pi$. The matrix for the effect of this rotation on a spinor is 

$$ R_x (\frac{\pi}{2}) = \begin{pmatrix} \cos \frac{\pi}{2} & -i \sin \frac{\pi}{2} \\ i \sin \frac{\pi}{2} & \cos \frac{\pi}{2} \end{pmatrix} = -i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} $$

Furthermore, under this rotation $\chi \rightarrow -\chi$; $\pi-\theta \rightarrow \theta$; $\pi + \varphi \rightarrow \pi - \varphi$. Disregarding the common $(-i)$ factor, we obtain a solution where the incident beam is polarized as $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. This solution we call $\psi_2$: 

$$ \psi_2 \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{ikr} + \frac{e^{ikr}}{r} \begin{pmatrix} -g(\theta)e^{i\varphi} \\ f(\varphi) \end{pmatrix} $$

(6.34)

To see clearer how the angles change under inversion and then rotation by $\pi$ around the x-axis, we draw symbolic diagrams for these operations.

Initial state $\psi$
$\tilde{\psi}' = \tilde{\psi}$ after inversion

$\tilde{\psi}'' = \tilde{\psi}'$ after the $R(\frac{\pi}{2})$ rotation

We see that the scattering of the incident beam $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is completely determined when the scattering of the beam $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is known.

Having found the angular-dependence of $\tilde{\psi}_{x+}$ and $\tilde{\psi}_{x-}$, we can find immediately the angular dependence for other incident beams.

Thus, suppose the spinor in the incident beam is in the $+y$ direction; the incident beam then is $\eta e^{i\frac{k}{\tau} x}$ with $\eta = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$, and therefore we find $\tilde{\psi}_{y+}$ by the recipe

$$\tilde{\psi}_{y+} = \frac{1}{\sqrt{2}} \left( \tilde{\psi}_{x+} + i \tilde{\psi}_{x-} \right)$$

$$\tilde{\psi}_{y+} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} e^{i\frac{k}{\tau} x} + \frac{e^{i\frac{k}{\tau} x}}{\tau \sqrt{2}} \begin{pmatrix} f(\varphi) - i q(\varphi) e^{-i \varphi} \\ q(\varphi) e^{i \varphi} + i f(\varphi) \end{pmatrix}$$

(6.35)
Similarly we find $\psi_2^-$:

$$
\psi_2^- = \frac{1}{\sqrt{2}} \left( \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{ikr} + \frac{e^{i\kappa r}}{\sqrt{2}} \left( f(\theta) + ig(\theta) e^{-i\varphi} \right) \right)
$$

(6.36)

We can now talk about the direction of the scattered spinor.

We first introduce the notation

$$
f(\theta) = |f| e^{i\alpha(\theta)}$$

$$
g(\theta) = |g| e^{i\beta(\theta)}$$

$$
\Delta(\theta) = \beta(\theta) - \alpha(\theta)
$$

(6.37)

Calling the two angles of the spinor orientation $\Theta$ and $\Phi$, we have

for $\psi_2^+$:

$$
tan \frac{\Theta}{2} = \frac{|g|}{|f|} \ ; \ \phi = \Delta(\Theta) + \varphi
$$

for $\psi_2^-$:

$$
tan \frac{\Theta}{2} = \frac{|f|}{|g|} \ ; \ \phi = -\Delta(\Theta) + \varphi - \pi
$$

Hence, when $\Delta(\Theta) = 0$ (f and g having the same phase) $\phi = \varphi$ or $\varphi + \pi$; this means that for $\Delta(\Theta) = 0$ the scattered spin is in the scattering plane defined by the incident and scattered beam.

If $\Delta(\Theta) \neq 0$, the scattered spin points out of this plane and points out equally but in opposite directions for $\psi_2^+$ and $\psi_2^-$. This, of course, has no effect on the scattering cross-section, which is the same for $\psi_2^+$ and $\psi_2^-$, namely $\sim |f|^2 + |g|^2$ and it is independent of $\varphi$ since any spin along the z-direction leaves the

5463/e/p
problem completely symmetric around the z-axis.

The situation is quite different when we consider either \( \gamma^+ \) or \( \gamma^- \) where we take the incident beam polarized in the \( ±y \)-direction. From eq. (6.35-36), we see that if we set \( \varphi = 0 \), i.e. consider scattering into the plane perpendicular to the direction of the spinor, we get:

\[
\gamma^+ = \frac{1}{\sqrt{2}} (1) e^{ikr} + \frac{e^{ikr}}{\sqrt{2}} (f - i \gamma) \gamma^+
\]

and

\[
\gamma^- = \frac{1}{\sqrt{2}} (-i) e^{ikr} + \frac{e^{ikr}}{\sqrt{2}} (f + i \gamma) \gamma^-
\]

So we conclude that in the plane perpendicular to the direction of the spinor there is no flipping of the spin. For arbitrary \( \varphi \), the angles \( \phi \) and \( \Theta \) are no longer the simple expressions we have found before. We shall not look for them but instead ask for the \( \varphi \)-dependence of the scattering cross section. We recall that \( \sigma \) is proportional to \( |C_+|^2 + |C_-|^2 \) where \( C_+ \) is the scattered spinor.
Consider the case $\psi_+^y$:

$$\sigma \sim \frac{1}{2} \left[ |f - ig e^{i\psi}|^2 + |g e^{i\psi} + if|^2 \right]$$

$$= \frac{1}{2} \left[ 2 |f|^2 + 2 |g|^2 + 2 \text{Re} \left( ifg^* e^{-i\psi} + if^*g e^{i\psi} \right) \right]$$

$$= |f|^2 + |g|^2 + |f||g| \text{Re} \left[ i \left( e^{-i(\Delta + \psi)} + e^{-i(\Delta - \psi)} \right) \right]$$

Hence

$$\sigma \sim |f|^2 + |g|^2 + 2 |f||g| \sin \Delta \cos \varphi \quad (6.38a)$$

For the case $\psi_-^y$ a similar calculation leads to

$$\sigma \sim |f|^2 + |g|^2 - 2 |f||g| \sin \Delta \cos \varphi \quad (6.38b)$$

Thus we manage to prove that the scattering cross-section will indeed depend on $\psi$, provided $\Delta \neq 0$. We cannot find $\Delta$ without having the exact solution. We expect, from physical considerations, that $\Delta$ will come out $\neq 0$ in an exact solution if the incident beam is polarized. It is not fair to ask that an argument of the type presented here give us more information than that contained in eq. (6.38a,b). All we were asking for was the functional form of the $\psi$-dependence of $\sigma$ and we now can say that this dependence is a $\cos \varphi$ law.
One could perhaps present an alternate argument, proving that $\sigma$ must depend on $\varphi$ through $\cos\varphi$.

Consider the spin-orbit coupling interaction energy. This, as we know, is proportional to $\mathbf{\mu} \cdot \mathbf{H}$ where $\mathbf{\mu}$ is the intrinsic magnetic moment (proportional to $\mathbf{s}$) and $\mathbf{H}$ is the magnetic field due to the orbital motion. We have assumed a spherically symmetric electrostatic potential for the scattering centre, hence its $\mathbf{E}$-field must be radial. Therefore, the $\mathbf{E}$-field due to motion through such an $\mathbf{E}$-field will be perpendicular to the orbit. Now in the case where the incident spinor is perpendicular to the direction of the beam, it follows that in the plane $\varphi = 0$ (plane perpendicular to direction of spinor) $\mathbf{\mu}$ and $\mathbf{H}$ will be parallel or anti-parallel. In either case the spin finds itself in an eigenstate with respect to the magnetic field, and does not want to flip. However, depending on whether $\mathbf{\mu}$ and $\mathbf{H}$ are parallel or anti-parallel, the interaction energy will be minimum or maximum. As we leave the $\varphi = 0$ plane, the interaction energy proportional to $\mu H \cos \varphi$ brings about the flipping of the spin in a way obviously proportional to $\cos \varphi$. Thus we see why the scattering cross-section will depend on $\varphi$ through $\cos \varphi$. On the other hand, when the incident spinor is parallel to the direction of the beam, $\mathbf{\mu}$ and $\mathbf{H}$ are always perpendicular to each other, $\mathbf{\mu} \cdot \mathbf{H} = 0$, and the scattering cross-section is independent of $\varphi$.

In our discussion of a polarized incident beam, we have been assuming that the incident spinor was perpendicular to the beam.
This need not be the case; however, a spinor at any other angle to the beam can be obtained from a proper combination of the two spinors \( \eta_{\pm} \). By a "proper" combination, we mean that not only the amplitudes of \( \eta_{\pm} \) but also their phases must be chosen right. For example, a spinor \( \eta \) made up out of 50% \( \eta_{\pm} \) and 50% \( \eta_{\mp} \) with arbitrary phases does not correspond to a spinor in some definite direction. In fact, above \( \eta \) could also be written as a combination of 50% \( \eta_{\pm} \) and 50% \( \eta_{\mp} \) with some other phases.

We now show that in the case of the Born approximation \( \Delta(\mathbb{S}) \) comes out 0, and so \( \mathbb{S} \) comes out independent of \( \varphi \). We recall that the functions we now call \( f,g \) (or some combination of \( f,g \)) are the elements of the matrix \( A \) given by (6.20):

\[
A = \frac{W+\mu}{2W} \left( 1 + \frac{\mathbf{p}_f \cdot \mathbf{p}_0}{(W+\mu)^2} + i \mathbf{\hat{e}} \cdot \frac{\mathbf{p}_f \times \mathbf{p}_0}{(W+\mu)^2} \right)
\]

We cannot use the elements of \( A \) as given by (6.22), since those were obtained by choosing the z-axis of our co-ordinate system to be parallel to \( \mathbf{p}_f \times \mathbf{p}_0 \), whereas now we require the z-axis to be parallel to \( \mathbf{p}_0 \). In this co-ordinate system \( \mathbf{p}_f \times \mathbf{p}_0 \) will have no z-component, hence \( \mathbf{\hat{e}} \cdot \mathbf{p}_f \times \mathbf{p}_0 \) will have no diagonal terms and we get for the diagonal elements

\[
A_{11} = A_{22} = \frac{W+\mu}{2W} \left( 1 + \frac{\rho^2}{(W+\mu)^2} \cos \varphi \right)
\] (6.39a)
For the off-diagonal we get

\[
a_{12} = \frac{i}{2W(w+\mu)} \left[ \left( \hat{p}_f \times \hat{p}_o \right)_x - i \left( \hat{p}_f \times \hat{p}_o \right)_y \right]
\]

\[
a_{21} = \frac{i}{2W(w+\mu)} \left[ \left( \hat{p}_f \times \hat{p}_o \right)_x + i \left( \hat{p}_f \times \hat{p}_o \right)_y \right]
\]

which can be written as

\[
a_{12} = \frac{\rho^2}{2W(w+\mu)} \sin \phi \ \text{e}^{i\phi} \quad (6.39b)
\]

We now recall that in terms of the matrix elements of \( \mathcal{A} \), the scattered wave is as given by eq. (6.17):

\[
\frac{e^{ik r}}{4\pi} \sqrt{\left( \frac{\tilde{\omega}}{\tilde{k}_0 - \tilde{k}} \right) \frac{2E}{\tilde{k}^2 c^2}} \sum_{k=1}^{2} a_{ki} \eta^{(k)}(\hat{p}_f)
\]

where the subscript \( i \) in \( a_{ki} \) specifies the incident spinor \( \eta^{(i)}(\hat{p}_o) \).

On the other hand, from equation (6.33) we have for the scattered wave

\[
\frac{e^{ik r}}{\nu} \begin{pmatrix} f(\theta) \\ e^{i\psi} q(\theta) \end{pmatrix}
\]

the incident wave being \( \eta^{(i)}(\hat{p}_o) \).
It is therefore clear that

\[ f(\vartheta) = A \ a_{11} \ e^{i \varphi} \ q(\vartheta) = A \ a_{21} \]

where \( A = \frac{\sqrt{\Delta \ a_{11}}}{4\pi} \frac{2E}{\kappa^2 c^2} \)

or

\[ f(\vartheta) = A \ \frac{w + \mu}{2w} \left( 1 + \frac{E^2}{(w + \mu)^2} \cos \vartheta \right) \]

\[ q(\vartheta) = A \ \frac{P^2}{2w(w + \mu)} \sin \vartheta \]

(6.40)

Recalling the definition of \( \Delta^{(3)} \) by eq. (6.37), we see from (6.40) that \( \Delta(\vartheta) = 0 \) and so if we use the Born approximation, we cannot expect to obtain any information about the polarization of the scattered wave.
# APPENDIX

## Chapter I

### Angular Momentum

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Properties of Rotations</td>
<td>2</td>
</tr>
<tr>
<td>Commutation Relations of Angular Momenta</td>
<td>5</td>
</tr>
<tr>
<td>Their Eigenvectors</td>
<td>6</td>
</tr>
<tr>
<td>Matrix elements of the infinitesimal Rotations</td>
<td>11</td>
</tr>
<tr>
<td>Special case: Spin 1</td>
<td>12</td>
</tr>
<tr>
<td>Spherical Harmonics</td>
<td>15</td>
</tr>
<tr>
<td>Addition of Angular Momenta</td>
<td>18</td>
</tr>
<tr>
<td>Special case: $\ell_1 = \ell_2 = 1$</td>
<td>23</td>
</tr>
<tr>
<td>Vector Representation of Angular Momentum and Vector Fields</td>
<td>26</td>
</tr>
<tr>
<td>Angular Momentum in Vector Fields</td>
<td>29</td>
</tr>
<tr>
<td>Spinors</td>
<td>32</td>
</tr>
<tr>
<td>Transformations properties of Spinors</td>
<td>34</td>
</tr>
<tr>
<td>Special Rotations</td>
<td>37</td>
</tr>
<tr>
<td>Direction of a Spinor</td>
<td>39</td>
</tr>
<tr>
<td>Spinor Field</td>
<td>40</td>
</tr>
<tr>
<td><strong>TABLE I</strong> Clebsch-Gordan Coefficients for $j' = \frac{1}{2}$</td>
<td>41</td>
</tr>
<tr>
<td><strong>TABLE II</strong> Clebsch-Gordan Coefficients for $j' = 1$</td>
<td>42</td>
</tr>
</tbody>
</table>

## Chapter II

### Dirac Equation for the Electron

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analogy with Maxwell Equations</td>
<td>43</td>
</tr>
<tr>
<td>&quot;Maxwell form&quot; of the Dirac equation</td>
<td>46</td>
</tr>
<tr>
<td>&quot;Dirac form&quot; of the Dirac equation</td>
<td>47</td>
</tr>
<tr>
<td>Maxwell's Equations in Diracs formalism</td>
<td>48</td>
</tr>
<tr>
<td>The Physics in Dirac's equation</td>
<td>51</td>
</tr>
<tr>
<td>Charge density</td>
<td>52</td>
</tr>
<tr>
<td>Current density</td>
<td>53</td>
</tr>
<tr>
<td>Electromagnetic analogy: energy density and the Poynting vector</td>
<td>54</td>
</tr>
<tr>
<td>&quot;Large&quot; and &quot;small&quot; spinors</td>
<td>56</td>
</tr>
<tr>
<td>Non-relativistic form of the current: the magnetic moment</td>
<td>58</td>
</tr>
<tr>
<td>Relativistic generalisation: the electric polarisation</td>
<td>59</td>
</tr>
<tr>
<td>Solution of the Dirac equation</td>
<td>62</td>
</tr>
<tr>
<td>The four fundamental solutions</td>
<td>66</td>
</tr>
<tr>
<td>The Spin direction</td>
<td>67</td>
</tr>
<tr>
<td>The unitary matrix $D$ ($^\dagger$)</td>
<td>68</td>
</tr>
<tr>
<td>Velocity of the Dirac Electron</td>
<td>69</td>
</tr>
<tr>
<td>Average value of the velocity</td>
<td>71</td>
</tr>
<tr>
<td>Average value of $\beta$</td>
<td>73</td>
</tr>
<tr>
<td>Redefinition of the matrix $D$: $D$ ($\bar{\psi}$)</td>
<td>74</td>
</tr>
<tr>
<td>Chapter III</td>
<td>The Dirac Equation under the Lorentz Transformation</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>The matrix: $S (\mathcal{V})$</td>
<td>76</td>
</tr>
<tr>
<td>The two proofs of Lorentz in variance</td>
<td>78</td>
</tr>
<tr>
<td>Proof a) Physical argument</td>
<td>78</td>
</tr>
<tr>
<td>Proof b) Formal argument</td>
<td>80</td>
</tr>
<tr>
<td>Space rotations</td>
<td>82</td>
</tr>
<tr>
<td>The Lorentz invariant ${\psi', \beta \psi}$</td>
<td>83</td>
</tr>
<tr>
<td>The transformation law of the current</td>
<td>84</td>
</tr>
<tr>
<td>Comparison of the Dirac Spinor and the Maxwell</td>
<td></td>
</tr>
<tr>
<td>vectors under the Lorentz transformation</td>
<td>85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter IV</th>
<th>Dirac's Equation in the Presence in an External Field</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The classical interaction energy</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>The Dirac equation with external fields</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>The Charge Conjugate Solution</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Properties of the matrix $C$</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>Charge conjugation for free particles</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Charge conjugation and time reversal</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Schwinger's charge conjugation matrix</td>
<td>94</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter V</th>
<th>Solution of Dirac's Equation in an External Field</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>Dirac Corrections to the Schrödinger equation</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Physical interpretation of the correction terms</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Zeeman effect</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>Solution of Dirac's equation in a central field</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Conservation of total angular momentum</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>Spinor spherical harmonics</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Parity invariance</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Evaluation of $\vec{\mathcal{V}} \cdot \vec{P}$</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>Radial equations</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Non-relativistic limit of these equations</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>First Dirac corrections to the radial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Schrödinger Equation</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>Ground state of the Hydrogen atom</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>Hierarchy of lengths</td>
<td>120</td>
</tr>
<tr>
<td>Chapter VI</td>
<td>The Scattering Problem, Born Approximation</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>Born approximation in Schrödinger formalism</td>
<td>121</td>
<td></td>
</tr>
<tr>
<td>Born approximation in Diracs formalism</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>Different forms of the scattering amplitude</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>The differential cross-section</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td>Evaluation of the spin dependence</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>The Born-cross-section is independent of the polarisation</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Another formulation of the Born approximation</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Polarisation in the scattered beam</td>
<td>134</td>
<td></td>
</tr>
<tr>
<td>General consequences of rotation and parity-invariance</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Direction of the scattered spinor</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>The azimuthal dependence of the scattering cross-section</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>Application to the Born approximation</td>
<td>143</td>
<td></td>
</tr>
</tbody>
</table>