RELATIVISTIC QUANTUM

MECHANICS AND

INTRODUCTION TO

QUANTUM FIELD THEORY

Anton Z. Capri

Theoretical Physics Institute,
Department of Physics,
University of Alberta

World Scientific
New Jersey • Singapore • London • Hong Kong
Preface

These notes arose from a graduate course that I taught for several years to students interested in particle physics and field theory. Most of the material in the first part is strictly quantum mechanical and fairly standard. In this part I develop in detail the important concepts of discrete and continuous symmetries of the Klein-Gordon and Dirac equations. I also deemed it important to include a section on central force problems for the Dirac equation since this is once again of interest in nuclear physics.

In the second part I wanted to introduce the students to some of the main ideas of field theory via the canonical formalism. Although I know that, to quote Schwinger, "Like the silicon chip of more recent years, the Feynman diagram was bringing computation to the masses", these days it is functional methods that are the computational tools of the masses, I nevertheless felt that students should be first exposed to the canonical formalism that has proved so robust for almost three-quarters of a century. It is a beautiful subject, worthy of study.

As in all such endeavours, I have been helped most by all the students who took this course from me over the past several years and forced me to refine my explanations. I thank them most sincerely and hope they got as much out of my course as I got from them.

A.Z. Capri
Edmonton, Alberta
January 2002
## Contents

1 Relativistic Quantum Mechanics  
1.1 The Poincaré Group  
1.2 The Casimir Operators  
1.3 The Irreducible Representations  
1.4 Problems

2 Spin 0: The Klein-Gordon Equation  
2.1 The Equation for a Free Particle  
2.2 Transformation Properties  
2.3 The Current  
2.4 Solutions of the K-G Equation  
2.5 Inner Product  
2.6 Normalization  
2.7 Charged Particles  
2.8 Charge Conjugation  
2.9 The Klein Paradox  
2.10 Coulomb Interaction

3 The Dirac Equation  
3.1 Introduction  
3.2 The Dirac Equation  
3.3 Free Particle Solutions  
3.4 Normalization, Orthogonality, Traces  
3.4.1 Normalization  
3.4.2 Orthogonality  
3.4.3 Traces  
3.4.4 Gamma Gymnastics  
3.5 Relativistic Invariance  
3.6 The Discrete Transformations  
3.6.1 Charge Conjugation  
3.6.2 Parity  
3.6.3 Time Reversal

vii
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6.4</td>
<td>Charge Conjugate Spinors</td>
<td>51</td>
</tr>
<tr>
<td>3.7</td>
<td>Problems</td>
<td>52</td>
</tr>
<tr>
<td>4</td>
<td>Structure of Dirac Particles</td>
<td>54</td>
</tr>
<tr>
<td>4.1</td>
<td>Electromagnetic Interaction</td>
<td>54</td>
</tr>
<tr>
<td>4.2</td>
<td>Constants of the Motion</td>
<td>59</td>
</tr>
<tr>
<td>4.3</td>
<td>Problems</td>
<td>61</td>
</tr>
<tr>
<td>5</td>
<td>Dirac Equation: Central Potentials</td>
<td>62</td>
</tr>
<tr>
<td>5.1</td>
<td>Separation of Variables</td>
<td>62</td>
</tr>
<tr>
<td>5.2</td>
<td>The Hydrogenic Atom</td>
<td>68</td>
</tr>
<tr>
<td>5.3</td>
<td>Problems</td>
<td>72</td>
</tr>
<tr>
<td>6</td>
<td>The Weyl or Neutrino Equation</td>
<td>75</td>
</tr>
<tr>
<td>6.1</td>
<td>Derivation</td>
<td>75</td>
</tr>
<tr>
<td>6.2</td>
<td>Solutions</td>
<td>76</td>
</tr>
<tr>
<td>6.3</td>
<td>Problems</td>
<td>77</td>
</tr>
</tbody>
</table>

II Relativistic Quantum Field Theory

| 7 | The Neutral Klein-Gordon Field | 80 |
| 7.1 | Introduction | 80 |
| 7.2 | Classical Real Scalar Field | 80 |
| 7.3 | Particle Interpretation | 82 |
| 7.4 | Quantization of the Field | 83 |
| 7.5 | Annihilation and Creation Operators | 84 |
| 7.6 | The Hamiltonian | 85 |
| 7.7 | Normal Ordering | 88 |
| 7.8 | The Energy-Momentum Tensor | 89 |
| 7.9 | The Field Operators | 92 |
| 7.9.1 | Initial Value Problem | 94 |
| 7.10 | Retarded and Advanced Green’s Functions | 95 |
| 7.11 | c-number External Scalar Source | 97 |
| 7.12 | The S-Matrix and the S-Operator | 100 |
| 7.13 | Single Particle Matrix Elements | 103 |
| 7.14 | Problems | 104 |

| 8 | The Charged Klein-Gordon Field | 105 |
| 8.1 | Gauge Invariance | 105 |
| 8.2 | Interaction with an External Electromagnetic Field | 109 |
| 8.3 | Perturbation Theory | 110 |
| 8.4 | Problems | 111 |
9 **The Dirac Field**  
9.1 The Free Dirac Field .................................. 112  
9.2 Quantization .......................................... 113  
9.3 Positive Energy ....................................... 116  
9.4 Green’s Functions for Dirac Equation ............... 117  
9.5 Perturbation of Electromagnetic Interaction ........ 119  
9.6 The S-Operator ....................................... 120  
9.7 Scattering of an Electron by an External EM Field .... 122  
9.8 Cross-Section for Coulomb Scattering ................ 122  
9.9 Pair Production in an External Field ................ 127  
9.10 Vacuum Polarization .................................. 131  
9.11 Charge Renormalization ................................ 136  
9.12 Dispersion Relations .................................. 138  
9.13 Problems ............................................. 141  

10 **Asymptotic Fields: LSZ Formulation** .......... 143  
10.1 The Scalar Field .................................... 143  
10.2 Reduction of the S-matrix: Scalar Field ............. 145  
10.3 Reduction of the S-matrix: Dirac Field ............... 148  
10.3.1 Proton-Meson Scattering ........................... 150  
10.4 Problems ............................................. 151  

11 **Perturbation Theory**  
11.1 The U-matrix ......................................... 153  
11.2 Perturbation of VEV of T-Ordered Products ......... 156  
11.3 Wick’s Theorem ....................................... 160  
11.4 Graphical Representation .............................. 162  
11.5 The Pion-Nucleon Interaction ......................... 165  
11.6 Self-Interacting Boson Field ........................ 166  
11.7 Feynman Rules: External Field Problems ............. 170  
11.8 Electron Scattered by External EM Field ............. 172  
11.9 Pair Production ....................................... 173  
11.10 Vacuum Polarization .................................. 174  
11.11 Second Order Scattering .............................. 174  
11.12 Conclusion .......................................... 175  
11.13 Problems ............................................. 175
Chapter 1

The Poincaré Group

1.1 Notation

Throughout this book we use the following notation. A four vector is specified by its four components as follows

\[ A \equiv A^\mu = (A^0, \vec{A}) \]  \hspace{1cm} (1.1.1)

The metric tensor is given by

\[ g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix} \]  \hspace{1cm} (1.1.2)

This metric tensor is used to raise and lower indices according to

\[ A_\mu = A^\nu g_{\nu\mu} \quad A^\mu = A_\nu g^{\nu\mu} \]  \hspace{1cm} (1.1.3)

Here, as throughout the rest of the book we have summed over repeated upper and lower indices.

Using this summation convention, we write the inner product as

\[ A.B = A^\mu g_{\mu\nu} B^\nu = A^\mu B_\mu = A^0 B^0 - \vec{A} \cdot \vec{B} \]  \hspace{1cm} (1.1.4)

the "length" of a vector is given by

\[ A^2 = A.A = (A^0)^2 - \vec{A} \cdot \vec{A} \]  \hspace{1cm} (1.1.5)

A Lorentz transformation \( \Lambda \) is a linear transformation mapping Minkowski spacetime onto itself and therefore preserves the inner product. In other words we always have for any two four vectors \( x \) and \( y \)

\[ (Ax) \cdot (Ay) = x \cdot y \]  \hspace{1cm} (1.1.6)

If we write out the Lorentz transformation explicitly so that

\[ (Ax)^\mu = \Lambda_{\mu\nu} x^\nu \]  \hspace{1cm} (1.1.7)
then we find that the real matrix $\Lambda_{\mu\nu}$ must satisfy

$$\Lambda_{\mu\nu} \Lambda_{\mu\lambda} = g_{\nu\lambda}$$  \hspace{1cm} (1.1.8)

Finally, for later use, we also define the Levi-Civita pseudo-tensor

$$\epsilon_{\mu\nu\lambda\rho} = \begin{cases} 
1 & \text{if } \mu\nu\lambda\rho \text{ is an even permutation of } 0, 1, 2, 3 \\
-1 & \text{if } \mu\nu\lambda\rho \text{ is an odd permutation of } 0, 1, 2, 3 \\
0 & \text{if any two indices of } \mu\nu\lambda\rho \text{ are equal}
\end{cases} \hspace{1cm} (1.1.9)$$

### 1.2 The Poincaré Group

The Poincaré group is the set of Lorentz transformations and space-time translations $(\Lambda, a)$ such that

$$x^\mu \rightarrow x'^\mu = \Lambda_{\mu\nu} x^\nu + a^\mu \hspace{1cm} (1.2.10)$$

The group multiplication law is given by

$$(\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda_2 \Lambda_1, \Lambda_2 a_1 + a_2) \hspace{1cm} (1.2.11)$$

The generators of these transformations are called $P^\lambda$, $M^{\mu\nu}$ and are such that the corresponding unitary operators that represent these transformations may be written

$$U(a,1) = \exp(i P \cdot a)$$
$$U(0,\Lambda) = \exp(i M^{\mu\nu} \Lambda_{\mu\nu}) \hspace{1cm} (1.2.12)$$

They satisfy the following commutation relations. (See problem 1.1)

$$[M_{\mu\nu}, M_{\rho\sigma}] = -i (g_{\mu\rho} M_{\nu\sigma} + g_{\nu\sigma} M_{\mu\rho} - g_{\nu\rho} M_{\mu\sigma} - g_{\mu\sigma} M_{\nu\rho})$$
$$[M_{\mu\nu}, P_\sigma] = i (g_{\nu\sigma} P_\mu - g_{\mu\sigma} P_\nu)$$
$$[P_\mu, P_\nu] = 0 \hspace{1cm} (1.2.13)$$

In 1939 Wigner [1] presented a complete classification of all the unitary irreducible representations of the Poincaré group. These are the possible elementary states in any theory which respects this symmetry. We now give his classification. However, we do not follow his method. Instead we find a set of operators whose eigenvalues label these irreducible representations. These so-called Casimir operators are constructed from the generators and commute with all the generators.

A familiar example, from ordinary quantum mechanics, is the Casimir operator $J^2$ for the rotation group. It commutes with the three generators $J^i$ of the rotation group and its eigenvalues $j(j+1)\hbar^2$ label the “simplest” states of particles with angular momentum.
1.3 The Casimir Operators

For the Poincaré group we have two Casimir operators. They are

$$P^2 = P_\mu P^\mu$$ \hspace{1cm} (1.3.14)

and

$$w^2 = w_\mu w^\mu.$$ \hspace{1cm} (1.3.15)

where the Pauli-Lubanski vector $w_\mu$ is defined by

$$w_\sigma = \frac{1}{2} \epsilon_{\sigma \mu \nu \lambda} M^{\mu \nu} P^\lambda.$$ \hspace{1cm} (1.3.16)

Writing this out we find

$$w^0 = \vec{P} \cdot \vec{J}, \quad \vec{w} = P_0 \vec{J} - \vec{P} \times \vec{N}$$ \hspace{1cm} (1.3.17)

where

$$\vec{J} = (M_{32}, M_{13}, M_{21})$$ \hspace{1cm} (1.3.18)

are the three components of angular momentum and (with units such that $\hbar = 1$) satisfy

$$[J_1, J_2] = iJ_3 \quad \text{and cyclic permutations}.$$ \hspace{1cm} (1.3.19)

and

$$\vec{N} = (M_{01}, M_{02}, M_{03})$$ \hspace{1cm} (1.3.20)

are the boosts in the three Cartesian directions.

1.4 The Irreducible Representations

The irreducible unitary representations of the Poincaré group are classified according to the eigenvalues of $P^2$ and $w^2$. They fall into several classes according to the eigenvalues of $P^2$ and $P_0$. These are (here we have units such that $\hbar = c = 1$)

1. a) $P^2 = m^2 > 0$ and $P_0 > 0$

1. b) $P^2 = m^2 > 0$ and $P_0 < 0$

2. a) $P^2 = 0$ and $P_0 > 0$

2. b) $P^2 = 0$ and $P_0 < 0$

3. $P^2 = 0$ and $P_0 = 0$

In this case $P_\mu = (0, 0, 0, 0)$
4. $P^2 = m^2 < 0$

This case corresponds to faster-than-light particles or "tachyons".

Case 3. corresponds to particles with continuous spin and is also unphysical. Thus, only cases 1. and 2. are of interest. For case 1. we have $m \neq 0$ and thus we can transform to a Lorentz frame in which the three-momentum $\vec{p} = 0$. In this rest frame the eigenvalues of $P^\mu$ are

$$p^\mu = (m, 0, 0, 0)$$

so that

$$p^2 = p_\mu p^\mu = m^2$$

and

$$-w^2 = -w_0^2 + \vec{w}^2 = \vec{p}^2 J^2 = m^2 s(s + 1) .$$

(1.4.21)

(1.4.22)

(1.4.23)

Now, the eigenvalues of $J^2$ in the rest frame of the particle are just the values of the total intrinsic angular momentum or spin, namely $s(s + 1)$. This means that massive particles may be classified according to their mass and spin. We next consider the simplest cases, namely spin 0 and spin 1/2.

1.5 Problems

1.1 Show that the generators of the Poincaré group satisfy the commutation relations given by 1.5.24

$$[M_{\mu\nu}, M_{\rho\sigma}] = -i (g_{\mu\rho} M_{\nu\sigma} + g_{\nu\sigma} M_{\mu\rho} - g_{\nu\rho} M_{\mu\sigma} - g_{\mu\sigma} M_{\nu\rho})$$

$$[M_{\mu\nu}, P_\sigma] = i (g_{\nu\sigma} P_\mu - g_{\mu\sigma} P_\nu)$$

$$[P_\mu, P_\nu] = 0 .$$

(1.5.24)

1.2 Verify that $P^2$ and $w^2$ are indeed Casimir operators, i.e. that they commute with $P_\mu$ and $M_{\mu\nu}$. 
Chapter 2

Spin 0: The Klein-Gordon Equation

2.1 The Equation for a Free Particle

We recall that the free Schrödinger equation was obtained by writing the relation

\[ E = \frac{\vec{p}^2}{2m} \]  \hspace{1cm} (2.1.1)

as an operator equation. Now in the relativistic case we have the energy-momentum relation for a free particle

\[ E^2 = c^2 \vec{p}^2 + m^2 c^4 . \]  \hspace{1cm} (2.1.2)

Proceeding as for the Schrödinger equation and replacing \( E \) and \( \vec{p} \) by the corresponding operators \( \hbar \partial / \partial t \) and \( -i\hbar \nabla \) we get the Klein-Gordon equation.

\[ -\hbar^2 \frac{\partial^2 \phi}{\partial t^2} = (\hbar^2 c^2 \nabla^2 + m^2 c^4) \phi . \]  \hspace{1cm} (2.1.3)

We now rewrite this equation in several different ways to introduce more of our notation

\[ \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi + \frac{m^2 c^2}{\hbar^2} \phi = 0 . \]  \hspace{1cm} (2.1.4)

Also, setting \( x^0 = ct \) or else setting \( c = 1 \) we get, if we further choose units such that \( \hbar = 1 \), that the equation becomes

\[ \frac{\partial^2 \phi}{\partial x^0^2} - \nabla^2 \phi + m^2 \phi = 0 . \]  \hspace{1cm} (2.1.5)

Next, we introduce the symbols \( \partial^\mu \) and \( \partial_\mu \)

\[ \partial^\mu = \frac{\partial}{\partial x_\mu} = (\partial^0, -\vec{\nabla}) , \hspace{0.5cm} \partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_0, \vec{\nabla}) \]  \hspace{1cm} (2.1.6)
as well as
\[ \Box = \partial_\mu \partial^\mu = \partial_0^2 - \nabla^2. \] (2.1.7)

So the Klein-Gordon equation reads
\[ (\Box + m^2) \phi = 0. \] (2.1.8)

### 2.2 Transformation Properties

Under a proper (no parity change), orthochronous (no time reversal) Lorentz transformation and a space-time translation
\[ x \to x' = \Lambda x + a \] (2.2.9)

the field \( \phi \) transforms as a scalar. This means that the transformed field \( \phi' \) at \( x' \) is related to the untransformed field \( \phi \) at \( x \) by
\[ \phi'(x') = \phi(x). \] (2.2.10)

So, the field \( \phi(x) \) is transformed into the field \( \phi'(x') \) under the coordinate transformation 2.2.9. Equation 2.2.10 gives what is usually called the passive interpretation. In practice, however, one usually uses the active interpretation.\(^1\) We use the active interpretation throughout. This is given by
\[ \phi'(x) = \phi(\Lambda^{-1}(x - a)). \] (2.2.11)

The meaning here is that if \( \phi(x) \) is a solution then so is \( \phi'(x) \).

If we also allow discrete transformations such as parity, namely
\[ x'^0 = x^0 \quad \bar{x}' = -\bar{x} \] (2.2.12)

then we can distinguish between scalar and pseudoscalar fields. Thus, \( \phi \) is a scalar field if
\[ \phi'(x^0, \bar{x}) = \phi(x^0, -\bar{x}) = \phi(x^0, \bar{x}) \] (2.2.13)

and \( \phi \) is a pseudoscalar field if
\[ \phi'(x^0, \bar{x}) = \phi(x^0, -\bar{x}) = -\phi(x^0, \bar{x}) \] (2.2.14)

Physical examples of pseudoscalar fields are furnished by the pions.

\(^1\)There are difficulties for interpretation of transformations such as parity and time-reversal if one stays with the passive interpretation.
2.3 The Current

The Klein-Gordon field has a conserved current. The density is given by

$$\rho = \frac{i\hbar}{2mc^2} (\phi^* \partial_t \phi - \partial_t \phi^* \phi) = \frac{i\hbar}{2mc} (\phi^* \partial_0 \phi - \partial_0 \phi^* \phi)$$  \hspace{1cm} (2.3.15)

and the current is the same as in the case of the Schrödinger equation

$$\vec{j} = \frac{\hbar}{2im} (\phi^* \nabla \phi - \nabla \phi^* \phi) .$$  \hspace{1cm} (2.3.16)

As a consequence of the Klein-Gordon equation we have

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0 .$$  \hspace{1cm} (2.3.17)

If we introduce the four-vector current

$$j^\mu = (c\rho, \vec{j})$$  \hspace{1cm} (2.3.18)

then the continuity equation for $j^\mu$ reads.

$$\frac{\partial}{\partial x^\mu} j^\mu = \partial_\mu j^\mu = 0$$  \hspace{1cm} (2.3.19)

This shows that $j^\mu$ is indeed a four-vector since 0 is an invariant and $\partial_\mu$ transforms as a four-vector.

Now, suppose we look at stationary states so that we have

$$i\hbar \partial_t \phi = E \phi$$
$$-i\hbar \partial_t \phi^* = E \phi^*$$  \hspace{1cm} (2.3.20)

Then we find that the "probability density" $\rho$ becomes

$$\rho = \frac{E}{mc^2} \phi^* \phi .$$  \hspace{1cm} (2.3.21)

If we look at the non-relativisitic limit $E \approx mc^2$ we recover the non-relativistic probability density

$$\rho = \phi^* \phi .$$  \hspace{1cm} (2.3.22)

However, the relativistic energy-momentum relation

$$E^2 = c^2 \vec{p}^2 + m^2 c^4$$  \hspace{1cm} (2.3.23)

has two roots

$$E = \pm \sqrt{c^2 \vec{p}^2 + m^2 c^4} .$$  \hspace{1cm} (2.3.24)

This means that $\rho$ may be both negative as well as positive and thus can no longer be consistently interpreted as a probability density. This difficulty will be circumvented, in a later section, by interpreting the Klein-Gordon equation as a field equation in the same spirit as, for example, Maxwell's equations. In that case $e J^\mu$ will be interpreted as an electric current four-vector.
2.4 Solutions of the K-G Equation

We first look for plane wave solutions. These are of the form (again we have set \( \hbar = c = 1 \))

\[
\phi = \frac{1}{\sqrt{2(2\pi)^3}} e^{-i(p^0 x^0 - \vec{p} \cdot \vec{x})} = \frac{1}{\sqrt{2(2\pi)^3}} e^{-i p \cdot x}.
\] (2.4.25)

The normalization is arbitrarily set to \( 1/\sqrt{2(2\pi)^3} \). For this plane wave to be a solution requires that

\[
p^2 \equiv p^0^2 - \vec{p}^2 = m^2
\] (2.4.26)

so that we have

\[
p^0 = \pm E = \pm \sqrt{p^2 + m^2}.
\] (2.4.27)

Now, as we saw, the existence of negative energy solutions poses a problem for a consistent interpretation of the solutions of the K-G equation to yield a probability amplitude. However, if there are no interactions, this is really not a problem since if at some time, say \( t = 0 \), the solution has \( E > 0 \) then the solution maintains this condition. This can be made explicit by replacing the K-G equation by the following equation

\[
 i\partial_t \phi = \sqrt{m^2 - \nabla^2} \phi.
\] (2.4.28)

Here the square root of the operator is defined in terms of the Fourier transform by

\[
\sqrt{m^2 - \nabla^2} \phi = \int d^3k e^{i\vec{k} \cdot \vec{x}} \sqrt{m^2 + \vec{k}^2} \chi(\vec{k}, x^0)
\] (2.4.29)

where

\[
\phi(\vec{x}, x^0) = \int d^3k e^{i\vec{k} \cdot \vec{x}} \chi(\vec{k}, x^0).
\] (2.4.30)

Clearly, \( \chi(\vec{k}, x^0) \) satisfies the equation

\[
 i\partial_0 \chi(\vec{k}, x^0) = \omega(\vec{k}) \chi(\vec{k}, x^0)
\] (2.4.31)

with

\[
\omega(\vec{k}) = \sqrt{m^2 + \vec{k}^2}.
\] (2.4.32)
2.5 Inner Product

The density \( \rho \) provides us with a means to define an invariant, time-independent inner product. We start with

\[
(\phi, \phi) = \int_{t=\text{constant}} \rho(t, \vec{x}) \, d^3x \tag{2.5.33}
\]

and extend this to define

\[
(\phi_1, \phi_2) = i \int_t d^3x \left[ \phi_1^*(x) \partial_0 \phi_2(x) - \partial_0 \phi_1^*(x) \phi_2(x) \right]
\equiv i \int_t d^3x \phi_1^*(x) \partial_0^* \phi_2(x) . \tag{2.5.34}
\]

In the last step we have defined the operator \( \partial_0^* \). To show that this expression is indeed invariant under Lorentz transformations we rewrite it in a somewhat neater form using Fourier transforms. To do this we begin by noticing that the condition

\[
k^0 = +\sqrt{k^2 + m^2} \tag{2.5.35}
\]

can be obtained from the expression

\[
\int dk^0 \delta(k^2 - m^2) \theta(k) \tag{2.5.36}
\]

where \( \theta(k) \equiv \theta(k^0) \) is the Lorentz-invariant \(^2\) step function

\[
\theta(k) = \begin{cases} 
1 & \text{for } k^0 > 0 \\
0 & \text{for } k^0 < 0
\end{cases} \tag{2.5.37}
\]

In this case any positive energy solution of the K-G equation may be written

\[
\phi(x) = \frac{\sqrt{2}}{\sqrt{(2\pi)^3}} \int d^4k \, e^{-ik \cdot x} \, \delta(k^2 - m^2) \, \theta(k) \, \Phi(k) \tag{2.5.38}
\]

Now, with \( \omega(\vec{k}) \) as defined by 2.4.32 we have

\[
\delta(k^2 - m^2) \theta(k) = \frac{1}{2\omega(\vec{k})} \, \delta(k^0 - \omega(\vec{k})) \tag{2.5.39}
\]

so we can write

\[
\phi(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int_{\nu} \frac{d^3k}{\omega(\vec{k})} \, e^{-ik \cdot x} \, \Phi(\vec{k}) \tag{2.5.40}
\]

\(^2\)The stepfunction is Lorentz invariant because it only distinguishes between the past and future and this is a Lorentz invariant concept, when \( k \) is restricted to the interior of the light cone.
where the subscript $V_+$ on the integral signifies that the integral is to be performed over the forward mass hyperboloid defined by
\[ k^0 = \omega(\vec{k}) = \sqrt{\vec{k}^2 + m^2} \tag{2.5.41} \]

Using this result we find
\[ i \int d^3 x \frac{\phi_1^*}{\partial_0} \phi_2 = \frac{i}{2(2\pi)^3} \int d^3 x \int_{V_+} \frac{d^3 k}{\omega(\vec{k})} \frac{d^3 q}{\omega(\vec{q})} \Phi_1^*(k) \Phi_2(q) \left[ -i\omega(\vec{q}) - i\omega(\vec{k}) \right] e^{-i(q-k) \cdot x} \tag{2.5.42} \]

We change the order of integration. Then, the integral over $d^3 x$ produces a delta function and we use this to integrate over $d^3 q$ to finally get
\[ (\phi_1, \phi_2) = \int_{V_+} \frac{d^3 k}{\omega(\vec{k})} \Phi_1^*(k) \Phi_2(k) \tag{2.5.43} \]

Since, this expression is equivalent to the expression
\[ (\phi_1, \phi_2) = \int d^4 k \delta(k^2 - m^2) \theta(k) \Phi_1^*(k) \Phi_2(k) \tag{2.5.44} \]

it is manifestly Lorentz invariant. Furthermore, it also makes it clear that
\[ \frac{d^3 k}{\omega(\vec{k})} \]

is a Lorentz invariant measure.

### 2.6 Normalization

We have normalized the plane wave solutions in the form
\[ \phi_p(x) = \frac{1}{\sqrt{2(2\pi)^3}} e^{-ip \cdot x}. \tag{2.6.45} \]

It now follows that, with the energies $p_0 = \omega(\vec{p})$ and $q_0 = \omega(\vec{q})$ fixed, we have
\[ (\phi_p, \phi_q) = p_0 \delta(\vec{p} - \vec{q}) \tag{2.6.46} \]

This is clearly a Lorentz invariant result since if we multiply by $d^3 p / \omega(\vec{p})$ and integrate we get 1. The completeness relation over the positive energy solutions
\[ \int_{V_+} \frac{d^3 p}{\omega(\vec{p})} \phi_p(x) \phi_p^*(x') = \frac{1}{2(2\pi)^3} \int_{V_+} \frac{d^3 p}{\omega(\vec{p})} e^{-ip \cdot (x - x')} \tag{2.6.47} \]
does not yield a δ-function since these solutions are only complete over the positive energy solutions. However, even if we include the negative energy solutions then the completeness relation

\[
\frac{1}{2(2\pi)^3} \int_{V^+} \frac{d^3p}{\omega(p)} e^{-ip.(x-x')} + \frac{1}{2(2\pi)^3} \int_{V^-} \frac{d^3p}{\omega(p)} e^{-ip.(x-x')}
\]

does not yield a δ-function even for \(x_0 = x'_0\) since the solutions of the Klein-Gordon equation are complete only over functions \(f(p)\) such that \(p_0^2 = \vec{p}^2 + m^2\).

Since we do not get δ-functions we therefore define the following functions which prove very useful later on.

\[
\Delta^+(x) = \frac{-i}{2(2\pi)^3} \int_{V^+} \frac{d^3p}{p^0} e^{-ip.x} \tag{2.6.48}
\]

\[
\Delta^-(x) = \frac{-i}{2(2\pi)^3} \int_{V^-} \frac{d^3p}{p^0} e^{-ip.x}
\]

\[
= \frac{i}{2(2\pi)^3} \int_{V^+} \frac{d^3p}{p^0} e^{ip.x}
\]

\[
= (\Delta^+(x))^* \tag{2.6.49}
\]

as well as

\[
\Delta(x) = \Delta^+(x) + \Delta^-(x) \tag{2.6.50}
\]

The function \(\Delta(x)\) has the interesting property that it vanishes if its argument is spacelike, that is if \(x^2 = (x^0)^2 - \vec{x}^2 < 0\). To prove this we simply write out \(\Delta(x)\).

\[
\Delta(x) = -\frac{1}{(2\pi)^3} \int_{V^+} \frac{d^3p}{\omega(p)} \sin(p^0x^0 - \vec{p} \cdot \vec{x}) \tag{2.6.51}
\]

Now we let \(x\) be a spacelike vector. In this case we can always perform a Lorentz transformation to a frame in which \(x^0 = 0\). In this special frame we then have

\[
\Delta(x)|_{\text{special frame}} = \frac{1}{(2\pi)^3} \int_{V^+} \frac{d^3p}{\omega(p)} \sin \vec{p} \cdot \vec{x} . \tag{2.6.52}
\]

Next, we let

\[
\vec{p} \cdot \vec{x} = \|\vec{p}\|\|\vec{x}\| \cos \theta = pru . \tag{2.6.53}
\]

Then,

\[
\Delta(x)|_{\text{special frame}} = \frac{2\pi}{(2\pi)^3} \int_0^\infty \frac{p^2 dp}{\sqrt{p^2 + m^2}} \int_{-1}^{1} du \sin pru . \tag{2.6.54}
\]

But,

\[
\int_{-1}^{1} du \sin pru = 0 . \tag{2.6.55}
\]

We can now Lorentz transform back to the original frame and conclude that for spacelike \(x\) we have \(\Delta(x) = 0\).
2.7 Charged Particles

The coupling to an electromagnetic field has to be done in a gauge invariant manner and we consider only minimal coupling. To this end we introduce the electromagnetic four-vector potential \( A_\mu \) where

\[
A_0 = \text{electric potential and } \vec{A} = \text{vector potential}
\]

so that

\[
\vec{B} = \nabla \times \vec{A} , \quad \vec{E} = -\nabla A_0 - \frac{1}{c} \frac{\partial \vec{A}}{\partial t} .
\]  

(2.7.57)

Later, when looking at the Dirac equation, we also consider direct coupling to the electromagnetic field tensor

\[
F_{\mu \nu} = \partial_\mu A_\nu - \partial_\nu A_\mu .
\]  

(2.7.58)

For minimal coupling we simply make the replacement

\[
p_\mu \to p_\mu - \frac{e}{c} A_\mu .
\]  

(2.7.59)

With \( \hbar \) and \( c \) explicitly included, the Klein-Gordon equation now becomes

\[
(i\hbar \partial_t - eA_0(x)) \phi(x) - (-i\hbar c \vec{\nabla} - e\vec{A}(x))^2 \phi(x) = m^2 c^2 \phi(x) .
\]  

(2.7.60)

Also, the conserved current becomes

\[
\rho = \frac{i\hbar}{2mc^2} (\phi^* \partial_t \phi - \partial_t \phi^* \phi) - \frac{e}{mc^2} A_0 \phi^* \phi
\]  

(2.7.61)

\[
\vec{j} = \frac{\hbar}{2im} (\phi^* \vec{\nabla} \phi - \vec{\nabla} \phi^* \phi) - \frac{e}{mc} \vec{A} \phi^* \phi .
\]  

(2.7.62)

In this case we find that the simple interpretation for \( \rho \) as a probability density, that was available in the case of no interaction, is no longer possible. To see this consider the case with \( \vec{A} = 0 \) for all time and \( A_0 = 0 \) except for a finite time interval \( 0 \leq t \leq T \). Suppose we have an incoming state of a particle with only positive energies so that

\[
\phi_{\text{incident}} = \int \text{positive energy solutions} .
\]  

(2.7.63)

Due to the interaction we then find that the state \( \phi_{\text{out}} \) for \( t > T \) always contains an admixture of negative energy solutions. This implies that the probability for a transition to occur from a state with positive energy to a state with negative energy is non-zero and therefore, we can longer have a consistently positive \( \rho \). This effect is not due to the time-dependence of \( A_0 \). It also occurs for static fields. To see this consider the Klein-Gordon equation for the static case where

\[
A_0(x) = A_0(\vec{x}) , \quad \vec{A} = 0 .
\]  

(2.7.64)
In this case we can look for stationary state solutions

$$\phi(t, \vec{x}) = u(\vec{x}) e^{-iEt/\hbar}.$$  \hfill (2.7.65)

The density $\rho$ now becomes

$$\rho(\vec{x}) = \frac{E - eA_0(\vec{x})}{mc^2} u^*(\vec{x})u(\vec{x}).$$  \hfill (2.7.66)

This shows that $\rho$ can become negative. In fact for the Coulomb potential

$$eA_0(\vec{x}) = \frac{Ze^2}{r}$$  \hfill (2.7.67)

we shortly solve the Klein-Gordon equation and find both positive and negative energy solutions. In this case if we start with a positive energy solution then it evolves into a sum of only positive energy solutions. Nevertheless, a consistent interpretation of $\rho$ as a probability density is not possible since as the expression 2.7.66 shows $\rho$ becomes negative whenever we get too close to the source of the Coulomb potential i.e. whenever

$$r < r_0 = \frac{Ze^2}{mc^2} = Z \times \text{the classical electron radius.}$$  \hfill (2.7.68)

Thus, we are forced to conclude that a consistent one-particle theory is not possible for the Klein-Gordon equation.

### 2.8 Charge Conjugation

We now try to find an interpretation for $\rho$. For this purpose we begin with positive energy solutions for a particle with charge $e$. These solutions can be characterized as satisfying the equation

$$i\hbar \partial_t \phi(e, +) = \sqrt{m^2c^4 + \hbar^2(-ic\vec{\nabla} - e\vec{A})^2} \phi(e, +) + eA_0 \phi(e, +).$$  \hfill (2.8.69)

The negative energy solutions satisfy

$$i\hbar \partial_t \phi(e, -) = -\sqrt{m^2c^4 + \hbar^2(-ic\vec{\nabla} - e\vec{A})^2} \phi(e, -) + eA_0 \phi(e, -).$$  \hfill (2.8.70)

If we complex conjugate this equation we find

$$i\hbar \partial_t \phi^*(e, -) = \sqrt{m^2c^4 + \hbar^2(ic\vec{\nabla} - e\vec{A})^2} \phi^*(e, -) - eA_0 \phi^*(e, -)$$

$$= \sqrt{m^2c^4 + \hbar^2(-ic\vec{\nabla} + e\vec{A})^2} \phi^*(e, -) - eA_0 \phi^*(e, -).$$  \hfill (2.8.71)

But this is just the equation for $\phi(-e, +)$, the positive energy solution for a particle with charge $-e$. Thus, we have

$$\phi(e, -)^* = \text{constant} \times \phi(-e, +).$$  \hfill (2.8.72)

The solution $\phi(e, -)^*$ is called the charge conjugate solution of $\phi(e, +)$. 
2.9 The Klein Paradox

We consider the scattering of a Klein-Gordon particle of energy \( E = \sqrt{k^2 + m^2} \) from a step function potential

\[
A_0 = V(z) \quad , \quad \vec{A} = 0 .
\]  
(2.9.73)

where

\[
V(z) = \begin{cases} 
0 & z < 0 \\
V > 0 & z > 0 
\end{cases} .
\]  
(2.9.74)

Then, the Klein-Gordon equation becomes

\[
(i\partial_t - V)^2 \phi + \nabla^2 \phi - m^2 \phi = 0 \quad z > 0 \\
-\partial_t^2 \phi + \nabla^2 \phi - m^2 \phi = 0 \quad z < 0 .
\]  
(2.9.75)

The incoming beam (for \( z < 0 \)) is \( \exp -i(Et - kz) \). Thus, we look for solutions of the form

\[
\phi_\prec = e^{-iEt} \left[ e^{ikz} + Re^{-ikz} \right] \quad z < 0 \\
\phi_\succ = Te^{-iEt} e^{ik'z} \quad z > 0 .
\]  
(2.9.76)

After substituting these expressions into the equation of motion we find

\[
E^2 - k^2 - m^2 = 0 \quad \text{or} \quad E = \sqrt{k^2 + m^2} \\
(E - V)^2 - k'^2 - m^2 = 0 \quad \text{or} \quad k' = \sqrt{(E - V)^2 - m^2} .
\]  
(2.9.77)

Here we have already imposed the condition that the incoming wave carries positive energy. Also notice that \( k' \) may be purely imaginary as well as real. Next, we impose the usual boundary conditions that both \( \phi \) and \( \partial_z \phi \) are continuous at \( z = 0 \). Thus, we get

\[
1 + R = T \\
\frac{ik(1 - R)}{\partial_z} = ik'T
\]  
(2.9.78)

Solving for \( R \) and \( T \) we find a result that seems to look similar to the nonrelativistic case

\[
R = \frac{k' - k}{k' + k} \\
T = \frac{2k}{k' + k}
\]  
(2.9.79)

The current for \( z < 0 \) is given by

\[
j_\prec = \frac{1}{2im} \left[ (e^{-ikz} + R^* e^{ikz}) ik \left( e^{ikz} - Re^{-ikz} \right) + ik \left( e^{-ikz} - R^* e^{ikz} \right) \left( e^{ikz} + Re^{-ikz} \right) \right] \\
= \frac{k}{m} \left( 1 - |R|^2 \right)
\]  
(2.9.80)
Similarly the current for \( z > 0 \) is given by

\[
\jmath_{>} = \frac{k' + k'}{2m} |T|^2 \tag{2.9.81}
\]

Since the incident current is just

\[
\jmath_{\text{incident}} = \frac{k}{m} \tag{2.9.82}
\]

we find that for the kinetic energy \( E - m > V \), both \( k \) and \( k' \) are real and the transmission and reflection coefficients \( T \) and \( R \) are given by

\[
T = \frac{\jmath_{>}}{\jmath_{\text{incident}}} = \frac{4kk'}{(k' + k)^2} \tag{2.9.83}
\]

\[
R = \frac{\jmath_{<} - \jmath_{\text{incident}}}{\jmath_{\text{incident}}} = \frac{(k' - k)^2}{(k' + k)^2} \tag{2.9.84}
\]

This result is just like the non-relativistic result that one gets with the Schrödinger equation.

Throughout we have conservation of probability in the sense that

\[
R + T = 1 \tag{2.9.85}
\]

as one would expect from our experience with non-relativistic scattering.

For the case that the incident kinetic energy is less than the height of the potential barrier and \( |V - E| < m \) we see that \( k' \) is purely imaginary and therefore \( \jmath_{>} = 0 \) so that we have

\[
T = 0 \quad , \quad R = 1. \tag{2.9.86}
\]

This is also in agreement with non-relativistic scattering.

However, for the case of an extremely strong potential \( V > E + m \) we see that \( k' \) is again real, but negative. In this case we find

\[
T = -\frac{4k|k'|}{(k - |k'|)^2} < 0 \tag{2.9.87}
\]

and

\[
R = \frac{(k + |k'|)^2}{(k - |k'|)^2} > 1 \tag{2.9.88}
\]

although 2.9.85 is still obeyed. However, a negative transmission probability is totally unphysical. This is the famous Klein paradox. The accepted interpretation is that, for such a strong potential, particle pairs are created and the whole single-particle theory that we have been trying to apply no longer works. These created particles add to the reflected beam and so we find \( R > 1 \).
2.10 Coulomb Interaction

Here we have
\[ A_0 = -\frac{Ze}{r}, \quad \vec{A} = 0. \] (2.10.89)

In this case (with \(\hbar = c = 1\)) the Klein-Gordon equation reads
\[ \left[ -i\partial_t - eA_0 \right]^2 - \nabla^2 + m^2 \Phi = 0. \] (2.10.90)

We are interested in stationary states and therefore let
\[ \Phi = e^{-iEt} \phi(t) . \] (2.10.91)

This gives us the equation
\[ \left( E + \frac{Ze^2}{r} \right)^2 \phi = \left( -\nabla^2 + m^2 \right) \phi . \] (2.10.92)

Next, we separate out the angular momentum and put
\[ \phi(t) = R(r) Y_{lm}(\theta, \varphi) . \] (2.10.93)

The radial equation now reads
\[ \frac{1}{r} \frac{d^2}{dr^2} (rR) - \frac{l(l+1) - Z^2 e^2}{r^2} R + \frac{2EZ e^2}{r} R - (m^2 - E^2) R = 0 . \] (2.10.94)

Since we are looking for bound states we have \(E^2 < m^2\) and therefore we set
\[ \beta = \sqrt{m^2 - E^2} , \quad \nu = 2EZ e^2 \] (2.10.95)
as well as
\[ rR(r) = u(r) . \] (2.10.96)

This clearly requires that \(u(0) = 0\) and so we look for the asymptotic behaviour near \(r = 0\) by putting \(u(r) \approx r^{\lambda+1}\) and let \(r \to 0\). This yields
\[ \lambda(\lambda + 1) - l(l+1) + Z^2 e^4 = 0 \] (2.10.97)
and therefore
\[ \lambda = \frac{-1}{2} \pm \frac{1}{2} \sqrt{1 + 4l(l+1) - 4Z^2 e^4} . \] (2.10.98)

Since we require \(\lambda > -1\) we must use the plus sign in this expression. For a bound state the wavefunction must vanish at infinity. The correct asymptotic behaviour is found to be
\[ u(r) \approx e^{-\sqrt{m^2 - E^2} r} . \] (2.10.99)

Therefore, we look for a solution of the form
\[ u(r) = r^{\lambda+1} e^{-\sqrt{m^2 - E^2} r} f(r) . \] (2.10.100)
The differential equation for \( f(r) \) now becomes

\[
rf'' + 2[\lambda + 1 - \beta r] f' + [\nu - 2\beta(\lambda + 1)] f = 0 .
\]  
(2.10.101)

Since we have already extracted the asymptotic behaviour we look for a series solution

\[
f = A[1 + a_1 r + a_2 r^2 + \cdots + a_N r^N] \quad a_N \neq 0
\]  
(2.10.102)

where \( A \) is a normalization constant. Then, we obtain the following equation

\[
2.1a_2r + 3.2a_3r^2 + 4.3a_4r^3 + \cdots + N(N - 1)a_N r^{N-1} + 2(\lambda + 1)[a_1 + 2a_2 r + \cdots + N a_N r^{N-1}] - 2\beta[a_1 r + 2a_2 r^2 + \cdots + N a_N r^{N}] + [\nu - 2\beta(\lambda + 1)][1 + a_1 r + a_2 r^2 + \cdots + a_N r^N] = 0 .
\]  
(2.10.103)

Equating to zero the coefficients of the different powers of \( r \) we get

\[
\nu - 2\beta(\lambda + 1) + 2(\lambda + 1)a_1 = 0 .
\]  
(2.10.104)

\[
2a_2[3 + 2\lambda] + a_1[\nu - 2\beta(2 + \lambda)] = 0 .
\]  
(2.10.105)

\[
\vdots
\]  

\[
N a_N [N + 1 + 2\lambda] + a_{N-1}[\nu - 2\beta(N + \lambda)] = 0 .
\]  
(2.10.106)

\[
a_N[\nu - 2\beta(N + \lambda + 1)] = 0 .
\]  
(2.10.107)

Since \( a_N \) is assumed to be different from zero we have

\[
\nu = 2\beta(N + \lambda + 1) .
\]  
(2.10.108)

Writing this out and squaring we get

\[
4(m^2 - E^2)(N + \lambda + 1)^2 = 4E^2 Z^2 e^4 .
\]  
(2.10.109)

Thus, we obtain the quantized energy levels

\[
E = \frac{N + \lambda + 1}{\sqrt{(N + \lambda + 1)^2 + Z^2 e^4}} m = \frac{N + \lambda + 1}{\sqrt{(N + \lambda + 1)^2 + \frac{Z^2 e^4}{\hbar^2 c^2}}} mc^2
\]  
(2.10.110)

where in the last step we have restored \( \hbar \) and \( c \).

The various coefficients may be written in terms of \( a_0 = 1 \). We then first find

\[
a_N = \frac{2\beta(N + \lambda) - \nu}{N(N + 2\lambda + 1)} a_{N-1}
\]  
(2.10.111)
These equations may be iterated to get all the coefficients

\[
a_N = \frac{2\beta(N + \lambda) - \nu}{N(N + 2\lambda + 1)} a_{N-1}
\]

\[
= \frac{2\beta(N + \lambda) - \nu}{N(N + 2\lambda + 1)} \frac{2\beta(N - 1 + \lambda) - \nu}{(N - 1)(N + 2\lambda)} a_{N-2}
\]

\[
= \frac{2\beta(N + \lambda) - \nu}{N(N + 2\lambda + 1)} \frac{2\beta(N - 1 + \lambda) - \nu}{(N - 1)(N + 2\lambda)} \ldots \frac{2\beta(1 + \lambda) - \nu}{1(2 + 2\lambda)}.
\]

(2.10.112)

Since,

\[
N + \lambda + 1 = N + \frac{1}{2} + \sqrt{(l + 1/2)^2 - Z^2e^4}
\]

\[
= N + l + 1 - [(l + 1/2) - \sqrt{(l + 1/2)^2 - Z^2e^4}]
\]

\[
= n - \epsilon_l
\]

(2.10.113)

where

\[
n = N + l + 1 = 1, 2, 3, \ldots
\]

\[
l = 0, 1, 2, \ldots, n - 1
\]

\[
\epsilon_l = (l + 1/2) - \sqrt{(l + 1/2)^2 - Z^2e^4}
\]

(2.10.114)

we can rewrite the energy as

\[
E_{n,l} = \frac{n - \epsilon_l}{\sqrt{(n - \epsilon_l)^2 + Z^2e^4}} m
\]

\[
= \left[ \frac{(n - \epsilon_l)^2 + Z^2e^4}{(n - \epsilon_l)^2} \right]^{-1/2} m
\]

\[
= \left[ 1 + \frac{Z^2e^4}{(n - \epsilon_l)^2} \right]^{-1/2} m
\]

(2.10.115)

or restoring \(h\) and \(c\) and introducing the finestructure constant

\[
\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}
\]

(2.10.116)

we find

\[
E_{n,l} = \left[ 1 + \frac{(Z^2e^4)/((\hbar \omega)^2)}{(n - \epsilon_l)^2} \right]^{-1/2} mc^2
\]

\[
\approx mc^2 - \frac{1}{2} mc^2 \alpha^2 Z^2 \frac{1}{(n - \epsilon_l)^2}
\]

(2.10.117)

where

\[
\epsilon_l = l + \frac{1}{2} - \sqrt{(l + 1/2)^2 - Z^2\alpha^2}
\]

(2.10.118)

This dependence of the spectrum on the quantum number \(l\) is sufficiently different from the hydrogen spectrum to have caused Schrödinger to reject the
Klein-Gordon equation as the correct description of the electron. He later returned to this problem with a non-relativistic formulation and discovered the Schrödinger equation. The Klein-Gordon equation was later rediscovered independently by [1] O. Klein, W. Gordon, V. Fock, J. Kudar, and The. de Donder and H. van Dungen. It is for this reason that Pauli dubbed it "the equation with many fathers".
3.1 Introduction

The difficulties of obtaining a consistent single particle theory from the Klein-Gordon equation led Dirac to look for an equation that 
a) had a positive-definite conserved density and 
b) was first order in time to give a unique evolution equation of the form

\[ i\hbar \frac{\partial}{\partial t} \Phi = H\Phi \]  \hspace{1cm} (3.1.1)

We now show that it is always possible to rewrite the Klein-Gordon equation as such a set of first order differential equations in time. To maintain a manifestly relativistic form of the equation requires that it should also be first order in space. Consider the Klein-Gordon equation (with \( \hbar = c = 1 \))

\[ (\partial^\mu \partial_\mu + m^2)\phi = 0 \]  \hspace{1cm} (3.1.2)

and define the four-vector

\[ \chi_\mu = \frac{i}{m} \partial_\mu \phi \]  \hspace{1cm} (3.1.3)

so that

\[ -i\partial_\mu \phi + m\chi_\mu = 0 \]  \hspace{1cm} (3.1.4)

Then, clearly

\[ -i\partial^\mu \chi_\mu = \frac{1}{m} \partial^2 \phi = -m\phi \]  \hspace{1cm} (3.1.5)

or

\[ -i\partial^\mu \chi_\mu + m\phi = 0 \]  \hspace{1cm} (3.1.6)
If we now define the five-component column matrix

\[ \Phi = \begin{pmatrix} \phi \\ \chi_0 \\ \chi_1 \\ \chi_2 \\ \chi_3 \end{pmatrix} \]  

(3.1.7)

Then, the pair of equations 3.1.4, 3.1.6 can be combined into the single matrix equation

\[ -i \beta \mu \partial^\mu \Phi + m \Phi = 0 \]  

(3.1.8)

where the four matrices \( \beta \mu \) can be read off from 3.1.4 and 3.1.6.

\[ \beta^0 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \beta^1 = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \]

\[ \beta^2 = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \beta^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \]  

(3.1.9)

The matrix equation 3.1.8 with the matrices 3.1.9 is known as the Duffin-Kemmer equation [1] and is equivalent to the Klein-Gordon equation. However, just like the Klein-Gordon equation it does not have a positive definite, conserved probability density.

### 3.2 The Dirac Equation

Although the Duffin-Kemmer equation was discovered after the Dirac equation, we presented the argument above to motivate the next step. As already stated, Dirac [2] was motivated by a desire to find an equation with a positive definite probability density. He also wanted an equation that was first order in time. For the equation to be manifestly relativistically covariant then implies that the equation should also be first order in space. It is now fairly clear that, to have an equation that is first order in time and space, a matrix equation is required. To this end we look for an equation of the form

\[ \frac{\hbar}{i c} \frac{\partial \psi_k}{\partial t} + \sum_n \bar{\alpha}_{kn} \frac{\hbar}{i} \nabla \psi_n + mc \sum_n \beta_{kn} \psi_n = 0 \]  

(3.2.10)

where we have introduced the four matrices \( \beta \) and \( \bar{\alpha} = (\alpha^1, \alpha^2, \alpha^3) \). The symbol \( \psi_n \) indicates the components of a column matrix \( \psi \). We also want a positive definite, conserved density

\[ \rho = \psi^\dagger \psi = \sum_n \psi_n^* \psi_n = \sum_n |\psi_n|^2. \]  

(3.2.11)
This means we need a conserved current $\tilde{j}$ such that
\[ \frac{\partial \rho}{\partial t} + \nabla \cdot \tilde{j} = 0 \]  
(3.2.12)

Now, from 3.2.10 we see that $\psi^\dagger$ satisfies
\[ \frac{1}{c} \frac{\partial \psi^\dagger}{\partial t} + \nabla \psi^\dagger \cdot \tilde{\alpha}^\dagger - \frac{i m c}{\hbar} \psi^\dagger \beta^\dagger = 0 \]  
(3.2.13)

If we multiply eqn 3.2.10 by $\psi^\dagger$ from the left and 3.2.13 by $\psi$ from the right and add the resultant equations we get
\[ \frac{1}{c} \frac{\partial}{\partial t} (\psi^\dagger \psi) + (\psi^\dagger \tilde{\alpha} \cdot \nabla \psi + \nabla \psi^\dagger \cdot \tilde{\alpha}^\dagger \psi) + \frac{i m c}{\hbar} (\psi^\dagger \beta \psi - \psi^\dagger \beta^\dagger \psi) = 0 \]  
(3.2.14)

For this to reduce to an equation of the form 3.2.12 requires that
\[ \beta = \beta^\dagger \]  
(3.2.15)

and
\[ \tilde{\alpha} = \tilde{\alpha}^\dagger \]  
(3.2.16)

In this case we find that we indeed have a positive definite probability density
\[ \rho = \psi^\dagger \psi \]  
(3.2.17)

and with
\[ \tilde{j} = c \psi^\dagger \tilde{\alpha} \psi \]  
(3.2.18)

the equation of continuity 3.2.12 is satisfied.

Of course we also want the new equation 3.2.10 to imply the Klein-Gordon equation so that we can satisfy the proper energy-momentum relation for a particle of mass $m$. To this end we “square” equation 3.2.10 by acting on it from the left with
\[ \frac{1}{c} \frac{\partial}{\partial t} - \tilde{\alpha} \cdot \nabla - \frac{i m c}{\hbar} \beta = \frac{1}{c} \frac{\partial}{\partial t} - \sum_m \alpha^m \frac{\partial}{\partial x_m} - \frac{i m c}{\hbar} \beta. \]  
(3.2.19)

Imposing the requirement that we obtain the Klein-Gordon equation then yields
\[ \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = \sum_{m,n} \frac{1}{2} [\alpha^m \alpha^n + \alpha^n \alpha^m] \frac{\partial^2}{\partial x_m \partial x_n} \psi \]

\[ - \frac{m^2 c^2}{\hbar^2} \beta^2 \psi + \frac{i m c}{\hbar} \sum_m [\alpha^m \beta + \beta \alpha^m] \frac{\partial}{\partial x_m} \psi \]

\[ = \left[ \nabla^2 - \frac{m^2 c^2}{\hbar^2} \right] \psi . \]  
(3.2.20)
Therefore, we get the additional conditions

\[
\frac{1}{2} [\alpha^m \alpha^n + \alpha^n \alpha^m] = \delta^{mn} \mathbf{1} \tag{3.2.21}
\]

\[
\alpha^m \beta + \beta \alpha^m = 0 \tag{3.2.22}
\]

\[
\beta^2 = 1 \tag{3.2.23}
\]

where \( \mathbf{1} \) represents an \( N \times N \) unit matrix. From now on we no longer write the unit matrix unless required for clarity. This should not create any confusion since matrices can only equal matrices.

Next we show that \( N \) is even. To do this we use the fact that

\[
\alpha^m \beta = -\mathbf{1} \beta \alpha^m \tag{3.2.24}
\]

Taking the determinant of this equation we get

\[
det (\alpha^m \beta) = (-1)^N \ det (\beta \alpha^m) \tag{3.2.25}
\]

Therefore,

\[
det(-\mathbf{1}) = (-1)^N = 1 \tag{3.2.26}
\]

so that \( N \) is even.

An alternate proof proceeds as follows. Since \( \beta \) is hermitian, it can be brought to diagonal form

\[
\beta = \begin{pmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_N
\end{pmatrix} \tag{3.2.27}
\]

Then, using the fact that

\[
\beta^2 = \mathbf{1} \tag{3.2.28}
\]

we see that each of the eigenvalues \( b_i = \pm 1 \). Furthermore the fact that

\[
(\alpha^k)^2 = \mathbf{1} \tag{3.2.29}
\]

guarantees that \( \alpha^k \) has an inverse. Then, we find

\[
\alpha^k \beta (\alpha^k)^{-1} = -\beta \tag{3.2.30}
\]

Taking the trace of this equation and using the cyclic property of the trace, namely that

\[
\text{Tr}(ABC) = \text{Tr}(CAB) \tag{3.2.31}
\]

we get that

\[
\text{Tr}(\beta) = -\text{Tr}(\beta) \tag{3.2.32}
\]
or
\[
\text{Tr}(\beta) = 0 . \quad (3.2.33)
\]
This means that the eigenvalues +1 and −1 occur in pairs and so \( N \) must again be even. By a similar argument we find that
\[
\text{Tr}(\alpha^k) = 0 . \quad (3.2.34)
\]
Actually, for irreducible representations of \( \alpha^k \) and \( \beta \) we require that \( N = 4 \). An explicit such representation is
\[
\alpha^k = \begin{pmatrix} 0 & \sigma^k \\ \sigma^k & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.2.35)
\]
Here the \( \sigma^k \) are the Pauli matrices and \( \mathbf{1} \) is a \( 2 \times 2 \) unit matrix.
Since \( \beta^2 = 1 \) we can make the Dirac equation look covariant by putting
\[
\gamma^0 = \beta , \quad \gamma^k = \beta \alpha^k . \quad (3.2.36)
\]
Then, after multiplying the Dirac equation by \( \beta \) we get
\[
(-i\hbar \beta \partial_0 - i\hbar \beta \alpha^k \partial_k) \psi + mc \psi = 0 \quad (3.2.37)
\]
or,
\[
(-i\gamma^\mu \partial_\mu + \frac{mc}{\hbar}) \psi = 0 . \quad (3.2.38)
\]
The algebra of the \( \gamma^\mu \) is obtained from that of the \( \alpha^k \) and \( \beta \). It reads
\[
\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu \nu} \mathbf{1} . \quad (3.2.39)
\]
To show that the Dirac equation is truly covariant, we must still show that the \( \gamma^\mu \)-matrices transform as the components of a four-vector. First some notation,
\[
\gamma \cdot p = \slashed{p} = \gamma^\mu p_\mu = \gamma^0 p^0 - \vec{\gamma} \cdot \vec{p} \quad (3.2.40)
\]
where
\[
\vec{\gamma} = (\gamma^1, \gamma^2, \gamma^3) , \quad \vec{p} = (p^1, p^2, p^3) . \quad (3.2.41)
\]
Also, we lower the indeces on the \( \gamma^\mu \) in the usual manner
\[
\gamma_\mu = g_{\mu \nu} \gamma^\nu \quad (3.2.42)
\]
Furthermore, we can combine the probability density \( \rho \) and the current density \( \vec{j} \) into what looks like a four-vector
\[
c \rho = j^0 = c \psi^\dagger \psi = c \psi^\dagger \gamma^0 \gamma^0 \psi , \quad j^k = c \psi^\dagger \alpha^k \psi = c \psi^\dagger \gamma^0 \gamma^k \psi \quad (3.2.43)
\]
Thus, if we define the \textit{Dirac adjoint}
\[
\overline{\psi} = \psi^\dagger \gamma^0 \quad (3.2.44)
\]
we can write the combined $j^0$ and $j^k$ simply as
\[ j^\mu = c\overline{\psi}\gamma^\mu \psi \]  
(3.2.45)

The continuity equation now reads
\[ \partial_\mu j^\mu = 0 \]  
(3.2.46)

Since it is known that $\partial_\mu$ transform as the covariant components of a four-vector and the right side of this equation is invariant, it follows that the four components $j^\mu$ transform as the components of a contravariant four-vector. This suffices to show that the $\gamma^\mu$-matrices transform as the components of a four-vector. We give a more explicit proof a little later.

The $\gamma^\mu$ are completely defined by their algebra
\[ \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \mathbf{1} \]  
(3.2.47)

Furthermore there are exactly 16 linearly independent elements that one can construct from this algebra, namely
\[
\begin{align*}
\mathbf{1} & = \Gamma_1 \\
\gamma^0, i\gamma^1, i\gamma^2, i\gamma^3 & = \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5 \\
i\gamma^2\gamma^3, i\gamma^3\gamma^1, i\gamma^1\gamma^2, i\gamma^0\gamma^1, i\gamma^0\gamma^2, i\gamma^0\gamma^3 & = \Gamma_6, \Gamma_7, \Gamma_8, \Gamma_9, \Gamma_{10}, \Gamma_{11} \\
i\gamma^0\gamma^2\gamma^3, i\gamma^0\gamma^3\gamma^1, i\gamma^0\gamma^1\gamma^2, \gamma^1\gamma^2\gamma^3 & = \Gamma_{12}, \Gamma_{13}, \Gamma_{14}, \Gamma_{15} \\
i\gamma^0\gamma^1\gamma^2\gamma^3 & = \Gamma_{16} 
\end{align*}
\]  
(3.2.48)

We have inserted appropriate factors of $i$ so that all of these $\Gamma_j$ satisfy
\[ (\Gamma_j)^2 = \mathbf{1} \]  
(3.2.49)

Furthermore, any permutations of 4 or fewer $\gamma^\mu$ can always be brought back to one of the 16 listed above, by using the algebra 3.2.47. Also if there is a product of more than 4 $\gamma^\mu$ we can use the same algebra to bring the two of the $\gamma^\mu$ that are the same next to each other and then using the fact that $(\gamma^\mu)^2 = \pm 1$ reduces the number of factors in the product by 2. This process can be continued until we are again left (up to a factor of $\pm 1$ or $\pm i$) with one of the 16 $\Gamma$'s listed above. For example, the product
\[ \gamma^0\gamma^1\gamma^2\gamma^3 \gamma^0 = -\gamma^0\gamma^0\gamma^1\gamma^2\gamma^3 = -\gamma^1\gamma^2\gamma^3 = -\Gamma_{15} \]  
(3.2.50)

It is furthermore worth noting that, up to a factor of $\pm 1$ or $\pm i$, we have always a $j$ such that
\[ \Gamma_k \Gamma_j = \Gamma_j \]  
(3.2.51)

and for every $\Gamma_k$, except $\Gamma_1$ we can always find a $\Gamma_j$ such that
\[ \Gamma_j \Gamma_k \Gamma_1 = -\Gamma_k \]  
(3.2.52)

The proof of this is simply to display the appropriate $\Gamma_j$ for each of the 15 $\Gamma_k$. Thus, for
\[ k = 2, 3, 4, 5 \text{ we use } \Gamma_{16} = i\gamma^0\gamma^1\gamma^2\gamma^3 \]

\[ k = 6, 7, 8, 9, 10, 11 \text{ we simply use one of } \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5. \]

Thus, for example for \( i\gamma^2\gamma^3 = \Gamma_6 \) we use \( i\gamma^2 = \Gamma_4 \)

\[ k = 12, 13, 14, 15 \text{ we use } \Gamma_{16} = i\gamma^0\gamma^1\gamma^2\gamma^3. \text{ Finally, for} \]

\[ k = 16 \text{ we use } \gamma^0 = \Gamma_2. \]

We now use these results to show that the 16 \( \Gamma \)'s are linearly independent. This means that we show that the equation

\[
\sum_{k=1}^{16} a_k \Gamma_k = 0
\]  

implies \( a_k = 0 \) for \( i = 1, \ldots, 16 \). We begin with the fact that

\[
\text{Tr}(\Gamma_k) = 0 \text{ for } k \neq 1, \quad \text{Tr}(\Gamma_1) = N.
\]

Taking the trace of the sum in (3.2.53) we get that

\[
a_1 N = 0
\]

Thus, the sum now runs from 2 to 16. Next we multiply the sum by \( \Gamma_i \) \( i \neq 1 \) and again take the trace to obtain

\[
a_i N = 0, \quad i = 2, 3, \ldots, 16.
\]

So we have shown that we have 16 linearly independent matrices and that means that the representation must consist of at least \( 4 \times 4 \) matrices, which is exactly what we have.

**Fundamental Theorem**

Let \( \gamma^\mu, \gamma'^\mu \) be two sets of irreducible matrices satisfying

\[
\gamma^\mu\gamma^\nu + \gamma'^\nu\gamma'^\mu = 2g^{\mu\nu} \mathbf{1}
\]

\[
\gamma'^\mu\gamma'^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \mathbf{1}
\]

then there exists a nonsingular matrix \( S \) such that

\[
\gamma'^\mu = S\gamma^\mu S^{-1}.
\]

This means that any irreducible representation of these matrices is unique up to a similarity transformation.
3.3 Free Particle Solutions

We look for plane wave solutions. To this end we set

\[ \psi_p(x) = e^{-ip \cdot x} u(p) \]  \hspace{1cm} (3.3.60)

where \( u(p) \) is a spinor, a column matrix. After substituting this into the Dirac equation

\[ \left(-i\gamma^\mu \partial_\mu + \frac{mc}{\hbar} \right) \psi = 0 \]  \hspace{1cm} (3.3.61)

with units such that \( \hbar = c = 1 \) we get

\[ (-\gamma \cdot p + m)u(p) = 0 \]  \hspace{1cm} (3.3.62)

Multiplying this equation from the left by \((\gamma \cdot p + m)\) yields

\[ (\gamma \cdot p + m)(-\gamma \cdot p + m)u(p) = 0 \]  \hspace{1cm} (3.3.63)

or

\[ (-\gamma^\mu \gamma^\nu p_\mu p_\nu + m^2)u(p) = 0 \]  \hspace{1cm} (3.3.64)

But,

\[ \gamma^\mu \gamma^\nu p_\mu p_\nu = \gamma^\nu \gamma^\mu p_\mu p_\nu = \frac{1}{2} (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) p_\mu p_\nu = g^{\mu\nu} p_\mu p_\nu = p^2 \]  \hspace{1cm} (3.3.65)

Thus, we find that

\[ (-p^2 + m^2)u(p) = 0 \]  \hspace{1cm} (3.3.66)

So, the momentum vector \( p_\mu \) must satisfy the Klein-Gordon condition

\[ p^2 = m^2 \]  \hspace{1cm} (3.3.67)

i.e.

\[ p^0 = \pm \sqrt{p^2 + m^2} \]  \hspace{1cm} (3.3.68)

This shows that \( p \) is a \textit{timelike vector} and therefore a rest frame exists. In this frame we have

\[ \vec{p} = 0 \]  \hspace{1cm} (3.3.69)

so that

\[ p^0 = \pm m \]  \hspace{1cm} (3.3.70)

and

\[ p = (\pm m, 0, 0, 0) \]  \hspace{1cm} (3.3.71)
Next, we use one of the explicit representations of the gamma matrices, namely
\[
\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}
\]
\[
\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\]
\[\text{(3.3.72)}\]

In this case the Hamiltonian in the rest frame becomes
\[H = \beta m + \vec{\alpha} \cdot \vec{p} = \beta m\,.
\]
\[\text{(3.3.73)}\]

Furthermore, if we define the three matrices
\[
\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}
\]
\[\text{(3.3.74)}\]

we see that all three of them commute with \(H\). So we can simultaneously diagonalize \(H\) and one of them, say \(\Sigma_3\). Then, we get four independent spinor solutions corresponding to the three-momentum \(\vec{p} = 0\).

\[
u_+ = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad u_- = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_+ = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_- = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
\]
\[\text{(3.3.75)}\]

As we see very soon, \((h/2) \vec{\Sigma}\) is the spin operator and the solutions \(u_\pm\) correspond to positive energy solutions of spin \(\pm h/2\), while the solutions \(v_\pm\) correspond to negative energy solutions of spin \(\pm h/2\).

We now repeat the explicit determination of the spinors for a general Lorentz frame. To begin notice that the Hamiltonian
\[H = \vec{\alpha} \cdot \vec{p} + \beta m\]
\[\text{(3.3.76)}\]
is hermitian if we use the obvious matrix inner product \(u^\dagger Hu\). Also, with the same explicit matrix representation as above, we have
\[
\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}.
\]
\[\text{(3.3.77)}\]

Furthermore, we already know that the eigenvalues of the Hamiltonian \(H\) are
\[
p^0 = \pm E(\vec{p}) = \pm \sqrt{\vec{p}^2 + m^2}.
\]
\[\text{(3.3.78)}\]

From now on \(E(\vec{p})\) always represents the positive square root \(\sqrt{\vec{p}^2 + m^2}\).

Writing out the eigenvalue equations explicitly in terms of two-component spinors \(u_1, u_2\) and \(v_1, v_2\) where
\[
u_\pm = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad v_\pm = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}
\]
\[\text{(3.3.79)}\]
we get
\[ \vec{\sigma} \cdot \vec{p} u_2 + m u_1 = E(\vec{p}) u_1 \]  
(3.3.80)
\[ \vec{\sigma} \cdot \vec{p} u_1 - m u_2 = E(\vec{p}) u_2 \]  
(3.3.81)

These two equation are linearly dependent since
\[ E(\vec{p})^2 = \vec{p}^2 + m^2. \]  
(3.3.82)

For the positive energy solution \( u_+ \) we have that
\[ E(\vec{p}) + m \geq 2m \neq 0. \]  
(3.3.83)

Therefore from 3.3.81 we find that
\[ u_2 = \frac{\vec{\sigma} \cdot \vec{p} u_1}{E(\vec{p}) + m}. \]  
(3.3.84)

If we substitute this into 3.3.80 we simply recover that \( E(\vec{p})^2 = \vec{p}^2 + m^2 \). Thus, up to normalization we have
\[ u_+ = \begin{pmatrix} \frac{u_1}{E(\vec{p}) + m} \\ \frac{d_1}{E(\vec{p}) + m} u_1 \end{pmatrix} \]  
(3.3.85)

where we still have
\[ u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \]  
(3.3.86)

corresponding to spin \(+1/2\) and \(-1/2\) respectively. After normalization so that
\[ u^\dagger u = \frac{E(\vec{p})}{mc^2} \]  
(3.3.87)
rather than
\[ u^\dagger u = 1 \]  
(3.3.88)

we get \( (c = 1) \)
\[ u_+ = \sqrt{\frac{E(\vec{p}) + m}{2m}} \begin{pmatrix} \frac{u_1}{E(\vec{p}) + m} \\ \frac{d_1}{E(\vec{p}) + m} u_1 \end{pmatrix}. \]  
(3.3.89)

### 3.4 Normalization, Orthogonality, Traces

#### 3.4.1 Normalization

Rather than the normalization \( u^\dagger u = 1 \) we have chosen \( u^\dagger u = E(\vec{p})/(mc^2) \). The reason for this is that this normalization is Lorentz covariant since now both sides of the equation transform as the 0-component of a Lorentz four-vector. Furthermore, this normalization makes the normalization in terms of the Dirac
adjoint $\bar{u} = u^\dagger \gamma^0$ very simple. To see this consider the Dirac equation for the spinors.

$$p^0 u = (c\bar{\alpha} \cdot \bar{p} + \beta mc^2) u$$  \hspace{1cm} (3.4.90)

$$p^0 u^\dagger = u^\dagger (c\alpha \cdot \bar{p} + \beta mc^2)$$  \hspace{1cm} (3.4.91)

Remember that $\beta$ and $\bar{\alpha}$ are hermitian and multiply 3.4.90 by $u^\dagger \beta$ from the left and 3.4.91 by $\beta u$ from the right and add the two equations to get

$$2p^0 u^\dagger \beta u = 2mc^2 u^\dagger u + u^\dagger (c\bar{\alpha} \cdot \bar{p} \beta + \beta c\alpha \cdot \bar{p}) u$$  \hspace{1cm} (3.4.92)

$$= 2mc^2 u^\dagger u$$  \hspace{1cm} (3.4.93)

So, we find that

$$\bar{u} u = \frac{mc^2}{p^0} u^\dagger u = \frac{mc^2}{p^0} \frac{|p^0|}{mc^2}$$  \hspace{1cm} (3.4.94)

Therefore,

$$\bar{u} u = \epsilon(p^0)$$  \hspace{1cm} (3.4.95)

where

$$\epsilon(p^0) = \begin{cases} +1 & \text{if } p^0 > 0 \\ -1 & \text{if } p^0 < 0 \end{cases}$$  \hspace{1cm} (3.4.96)

The technique of using the Dirac equation to obtain these matrix elements may also be used in other cases. For example, to evaluate $\bar{u}(\vec{p})\gamma^\mu u(\vec{p})$ we again start with

$$(\gamma \cdot p - mc) u(\vec{p}) = 0 \quad \text{and} \quad \bar{u}(\vec{p})(\gamma \cdot p - mc) = 0$$  \hspace{1cm} (3.4.97)

and multiply the first from the left by $\bar{u}(\vec{p})\gamma^\mu$ and the second from the right by $\gamma^\mu u(\vec{p})$ and add to get

$$2mc \bar{u}(\vec{p})\gamma^\mu u(\vec{p}) = \bar{u}(\gamma^\mu \gamma \cdot p + \gamma \cdot p \gamma^\mu) u(\vec{p})$$  \hspace{1cm} (3.4.98)

$$= p_\nu \bar{u}(\gamma^\nu \gamma^\mu + \gamma^\nu \gamma^\mu) u(\vec{p})$$  \hspace{1cm} (3.4.99)

$$= 2g^{\mu\nu} p_\nu \bar{u}(\vec{p}) u(\vec{p})$$  \hspace{1cm} (3.4.100)

$$= 2p^\mu \bar{u}(\vec{p}) u(\vec{p})$$

So,

$$\bar{u}(\vec{p})\gamma^\mu u(\vec{p}) = \frac{p^\mu}{mc} \bar{u}(\vec{p}) u(\vec{p}) = \frac{p^\mu}{mc} \epsilon(p^0)$$  \hspace{1cm} (3.4.101)

Similarly we easily see that

$$\bar{u}(\vec{p})\gamma^5 u(\vec{p}) = 0$$  \hspace{1cm} (3.4.102)

because by the same approach we find

$$2mc \bar{u}(\vec{p})\gamma^5 u(\vec{p}) = p_\nu \bar{u}(\vec{p}) (\gamma^5 \gamma^\nu + \gamma^\nu \gamma^5) u(\vec{p}) = 0$$  \hspace{1cm} (3.4.103)

since $\gamma^5$ anticommutes with all the other $\gamma$ matrices.
3.4.2 Orthogonality

In this section we introduce more notation that is conventional. Although it may seem as if we are introducing too many symbols, all of these occur in the literature and are therefore required. Let \( u^r_+ (\vec{p}) \), \( r = 1, 2 \) be two positive energy solutions corresponding to different eigenvalues of some convenient observable (related to spin) that commutes with the Hamiltonian. A convenient such observable is the helicity, defined as the component of spin parallel to the momentum. Then, \( u^1_+ \) and \( u^2_+ \) are orthogonal. This means

\[
\overline{u}^r_+ (\vec{p}) u^s_+ (\vec{p}) = \delta_{rs}, \quad r, s = 1, 2 .
\] (3.4.104)

Similarly, as we now show, the solutions \( u^r_+ (\vec{p}) \) corresponding to energy \(-E\), momentum \(-\vec{p}\), and definite helicity are orthogonal to \( u^r_+ (\vec{p}) \). They also satisfy

\[
(\gamma \cdot p + mc) u^- (\vec{p}) = 0
\] (3.4.105)

and

\[
\overline{u}^- (-\vec{p}) (\gamma \cdot p + mc) = 0 .
\] (3.4.106)

If we multiply (3.4.106) on the right by \( u^r_+ (\vec{p}) \) and the equation for \( u^r_+ (\vec{p}) \), namely

\[
(\gamma \cdot p - mc) u^r_+ (\vec{p}) = 0
\] (3.4.107)

on the left by \( \overline{u}^- (-\vec{p}) \) and subtract the resultant equations we find that

\[
2mc \overline{u}^- (-\vec{p}) u^r_+ (\vec{p}) = 0
\] (3.4.108)

as stated.

We define two linearly independent orthogonal solutions corresponding to energy \(-E\) and momentum \(-\vec{p}\) by

\[
v^s (\vec{p}) = u^s_- (-\vec{p}) \quad s = 1, 2 .
\] (3.4.109)

Here again, the index \( s \) labels helicity eigenstates. Then, we have

\[
\overline{v}^s (\vec{p}) v^r (\vec{p}) = -\delta_{rs} .
\] (3.4.110)

So, altogether we get

\[
\begin{align*}
\overline{u}^r_+ (\vec{p}) u^r_+ (\vec{p}) &= -\overline{v}^s (\vec{p}) v^r (\vec{p}) = \delta_{rs} , \\
\overline{u}^r_+ (\vec{p}) v^s (\vec{p}) &= \overline{v}^r (\vec{p}) u^r_+ (\vec{p}) = 0 .
\end{align*}
\] (3.4.111)

These form a complete set of solutions so that

\[
\sum_{r=1}^{2} \left\{ u^r_+ (\vec{p}) \overline{u}^r_+ (\vec{p}) - v^r_+ (\vec{p}) \overline{v}^r_+ (\vec{p}) \right\} = \delta_{\alpha \beta} 1 .
\] (3.4.112)

Here, \( \alpha, \beta \) are the spinor indeces and take the values 1, 2, 3, 4. What we have on the left side of this equation is nothing other than the exterior product

\[
u \otimes \overline{u} - v \otimes \overline{v} .
\] (3.4.113)
At this point we refine the notation a little further. Thus, we define

\[ w^r(\vec{p}) = u^r_+(\vec{p}) \quad r = 1, 2 \]  
(3.4.114)

and

\[ w^{r+2}(\vec{p}) = v^r(\vec{p}) = u^r_-(-\vec{p}) \quad r = 1, 2 . \]  
(3.4.115)

The orthogonality relations now read

\[ \bar{w}^m(\vec{p})w^n(\vec{p}) = \epsilon^m\delta_{mn} \quad m, n = 1, \ldots, 4 \]  
(3.4.116)

and

\[ \epsilon^m = \begin{cases} 
1 & m = 1, 2 \\
-1 & m = 3, 4 
\end{cases} \]  
(3.4.117)

The completeness relation can then be written

\[ \sum_{m=1}^{4} \epsilon^m w^m(\vec{p})\bar{w}^m(\vec{p}) = 1 \]  
(3.4.118)

Furthermore,

\[ \sum_{m=1}^{4} \epsilon^m \bar{w}^m(\vec{p})w^m(\vec{p}) = \sum_{m=1}^{4} (\epsilon^m)^2 = 4 \]  
(3.4.119)

### 3.4.3 Traces

In many computations, as we see a little later, one winds up having to compute quantities like

\[ \Omega = \sum_{r=1}^{2} (\bar{f}Qw^r)(\bar{w}^rPg) \]  
(3.4.120)

where \( Q \) and \( P \) are some combination of \( \Gamma \)'s and \( f, g \) are spinors. Written out in terms of all the spinor components and recalling that \( r = 1, 2 \) indicates positive energy we find

\[ \Omega = \sum_{r=1}^{2} \sum_{\alpha=1}^{4} (\bar{f}_\alpha Q_{\alpha\beta}w_\beta^r) \sum_{\rho=1}^{4} (\bar{w}_\rho^rP_{\rho\sigma}g_\sigma) \]  
(3.4.121)

If we can insert projection operators for positive energy then we can carry out the sum for \( r \) from \( r = 1 \) to \( r = 4 \) and use the completeness relation to simplify the computation. To achieve this we consider the equations

\[ (\not{p} - m)w^r(\vec{p}) = 0 \quad r = 1, 2 \]
\[ (\not{p} + m)w^r(\vec{p}) = 0 \quad r = 3, 4 . \]  
(3.4.122)
Next we consider

\[ \Lambda_{\pm} = \frac{m \pm \hat{p}}{2m} \]  \hspace{1cm} (3.4.123)

and find

\[ \Lambda_{+} w^r(\vec{p}) = \frac{m + \hat{p}}{2m} w^r(\vec{p}) = w^r(\vec{p}) \quad r = 1, 2 \]
\[ \Lambda_{-} w^r(\vec{p}) = \frac{m - \hat{p}}{2m} w^r(\vec{p}) = w^r(\vec{p}) \quad r = 3, 4 \]
\[ \Lambda_{+} w^r(\vec{p}) = 0 \quad r = 3, 4 \]
\[ \Lambda_{-} w^r(\vec{p}) = 0 \quad r = 1, 2 \]  \hspace{1cm} (3.4.124)

Also, when acting on free particle solutions of the Dirac equation so that \( p^2 = m^2 \)

\[ \Lambda_{+}^2 = \left( \frac{m + \hat{p}}{2m} \right)^2 = \frac{p^2 + 2m \hat{p} + m^2}{4m^2} = \frac{m + \hat{p}}{2m} = \Lambda_{+} \]
\[ \Lambda_{-}^2 = \left( \frac{m - \hat{p}}{2m} \right)^2 = \frac{p^2 - 2m \hat{p} + m^2}{4m^2} = \frac{m - \hat{p}}{2m} = \Lambda_{-} \]  \hspace{1cm} (3.4.125)

Considered as an operator on the four-dimensional vector space with the inner product \( \bar{u}u \) we find that \( \Lambda_{\pm} \) is hermitian.

\[ \Lambda_{\pm}^\dagger = \Lambda_{\pm} \]  \hspace{1cm} (3.4.126)

and finally

\[ \Lambda_{+} + \Lambda_{-} = 1 \]  \hspace{1cm} (3.4.127)

Thus, \( \Lambda_{\pm} \) are the desired projection operators. On the other hand, \( w^r(\vec{p}) \quad r = 1, 2 \) form a complete set of positive energy solutions and \( w^r(\vec{p}) \quad r = 3, 4 \) form a complete set of negative energy solutions. Therefore, we have

\[ \Lambda_{+} = \frac{m + \hat{p}}{2m} = \sum_{r=1}^{2} w^r(\vec{p}) \otimes \overline{w^r(\vec{p})} \]
\[ \Lambda_{-} = \frac{m - \hat{p}}{2m} = -\sum_{r=3}^{4} w^r(\vec{p}) \otimes \overline{w^r(\vec{p})} \]  \hspace{1cm} (3.4.128)

Also \( e^r \Lambda_{+} \) has the same properties as \( \Lambda_{+} \) since for \( r = 3, 4 \) the whole expression vanishes.

With these preliminaries out of the way we are ready to evaluate \( \Omega \).

\[ \Omega = \sum_{r=1}^{2} (\bar{f}Q w^r)(\overline{w^r P g}) \]
\[ = \sum_{r=1}^{4} (\bar{f}Q \Lambda_{+} w^r)(\overline{w^r P g}) \]
\[
\sum_{r=1}^{4} (\bar{f} Q A_+ e^r w^r) (\bar{w}^r P g)
\]
\[
= (\bar{f} Q A_+ \sum_{r=1}^{4} e^r w^r \otimes \bar{w}^r P g)
\]
\[
= (\bar{f} Q A_+ P g)
\] (3.4.129)

The same technique can be used if we have negative energy intermediate states.

A particular instance where this is frequently used is if we have to evaluate a matrix element of the form

\[
M = \bar{w}_f Q w_i
\] (3.4.130)

which describes a transition from the state labelled \(i\) to the state labelled \(f\). In this case the rate or probability to be calculated is proportional to \(|M|^2\). Also, in many cases the final spin is not of interest as, for example, in a scattering experiment in which the detector accepts all particles regardless of the spin. In such a case one sums over the final spin states. This leads to an expression of the form

\[
\sum_{\text{sum over final spin states of } |M|^2} \sum_{r=1}^{2} (w_i Q^\dagger w_f^r)(\bar{w}_f Q w_i)
\]
\[
= \bar{w}_i Q^\dagger A_+ (\bar{p}) Q w_i
\] (3.4.131)

where in the last step we have used the trick above. If the incoming beam is also unpolarized, we have to average over the initial spin states to obtain the result.

\[
\sum_{\text{Sum over final spins and average over initial spins of } |M|^2} \frac{1}{2} \sum_{r=1}^{2} \bar{w}_i^r Q^\dagger A_+ (\bar{p}) Q w_i^r
\]
\[
= \frac{1}{2} \sum_{r=1}^{2} \sum_{\alpha\beta=1}^{4} \bar{w}_i^r,\alpha (Q^\dagger A_+ Q)_{\alpha\beta} w_i^r,\beta
\]
\[
= \frac{1}{2} \sum_{r=1}^{2} \sum_{\alpha\beta=1}^{4} \bar{w}_i^r,\alpha (Q^\dagger A_+ Q A_+),_{\alpha\beta} e^r w_i^r,\beta
\]
\[
= \frac{1}{2} \sum_{\alpha\beta=1}^{4} (Q^\dagger A_+ Q A_+),_{\alpha\beta} \delta_{\alpha\beta}
\]
\[
= \frac{1}{2} \text{Tr} (Q^\dagger A_+ Q A_+)
\] (3.4.132)

So the whole computation has been reduced to the evaluation of a trace.
3.4.4 Gamma Gymnastics

In this section we develop some simple rules for evaluating traces of products of $\gamma$-matrices such as occur in the expression in 3.4.132. We begin by recalling that

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = (\gamma^5)^\dagger$$ \hspace{1cm} (3.4.133)

and that

$$(\gamma^5)^2 = 1$$ \hspace{1cm} (3.4.134)

as well as that $\gamma^5$ anticommutes with all the other $\gamma^\mu$. It then follows immediately that

1. The trace of an odd number of $\gamma$'s is 0.

Proof

Since,

$$\gamma^5\gamma^\mu = -\gamma^\mu\gamma^5$$ \hspace{1cm} (3.4.135)

it follows that

$$\gamma^5\gamma^{\mu_1}\ldots\gamma^{\mu_n}\gamma^5 = (-1)^n\gamma^{\mu_1}\ldots\gamma^{\mu_n}$$ \hspace{1cm} (3.4.136)

Then, using the cyclic property of the trace we immediately find that

$$\text{Tr}(\gamma^{\mu_1}\ldots\gamma^{\mu_n}) = (-1)^n\text{Tr}(\gamma^{\mu_1}\ldots\gamma^{\mu_n})$$ \hspace{1cm} (3.4.137)

Thus, the trace vanishes if $n$ is odd. The same argument also shows that

$$\text{Tr}(\gamma^5) = 0.$$ \hspace{1cm} (3.4.138)

The proof is again straightforward and relies on the cyclic property of the trace

$$\text{Tr}(\gamma^5) = \text{Tr}(\gamma^5\gamma^0(\gamma^0)^{-1}) = -\text{Tr}(\gamma^0\gamma^5(\gamma^0)^{-1}) = -\text{Tr}(\gamma^5)$$ \hspace{1cm} (3.4.139)

2. If $n$ is even one uses the algebra

$$\gamma^\nu\gamma^\mu + \gamma^\mu\gamma^\nu = 2g^\nu\mu$$

to reduce $n$ to $n-2$.

Thus, for example,

$$\text{Tr}(\gamma^{\mu}\gamma^{\nu}) = \text{Tr}(\gamma^{\nu}\gamma^{\mu})$$

$$= \frac{1}{2}\text{Tr}(\gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu})$$

$$= g^{\mu\nu}\text{Tr}(1) = 4g^{\mu\nu}. $$ \hspace{1cm} (3.4.140)

In a similar manner one can show (see problem 3.1) that

$$\text{Tr}(\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma}) = 4(g^{\mu\sigma}g^{\nu\rho} + g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma}).$$ \hspace{1cm} (3.4.141)

Another easy corollary to these results is that

$$\text{Tr}(A) = 0$$ \hspace{1cm} (3.4.142)

for any vector operator $A^\mu$ that does not contain $\gamma$ matrices. Furthermore, if we have two vector operators $A^\mu$, $B^\nu$ such that they commute it immediately follows that

$$\text{Tr}(AB) = 4A.B.$$ \hspace{1cm} (3.4.143)
3.5 Relativistic Invariance

We want to find the correct transformation such that the Dirac spinor equation should remain form invariant under the Lorentz transformation

\[ x'^\mu = \Lambda^\mu_\nu x^\nu \]  \hspace{1cm} (3.5.143)

or simply

\[ x' = \Lambda x \]  \hspace{1cm} (3.5.144)

This means that we want to find a representation of the Lorentz group, that is, a set of nonsingular matrices \( S(\Lambda) \) such that the transformation law for a Dirac spinor is given by

\[ \psi'(x') = S(\Lambda)\psi(x) \]  \hspace{1cm} (3.5.145)

Form invariance of the Dirac equation means that if \( (\hbar = c = 1) \)

\[ (-i\gamma^\mu \partial_\mu + m)\psi(x) = 0 \]  \hspace{1cm} (3.5.146)

then

\[ (-i\gamma^\mu \partial'_\mu + m)\psi'(x') = 0 \]  \hspace{1cm} (3.5.147)

Note that we have the same \( \gamma \)-matrices in both equations. Also,

\[ \partial'_\mu = \frac{\partial}{\partial x'^\mu} \]  \hspace{1cm} (3.5.148)

But,

\[ \frac{\partial}{\partial x^\mu} = \frac{\partial x'^\nu}{\partial x^\mu} \frac{\partial}{\partial x'^\nu} \]  \hspace{1cm} (3.5.149)

and using 3.5.143 we get,

\[ \frac{\partial}{\partial x^\mu} = \Lambda^\nu_\mu \frac{\partial}{\partial x'^\nu} \]  \hspace{1cm} (3.5.150)

Therefore, writing

\[ \psi(x) = S(\Lambda)^{-1}\psi'(x') \]  \hspace{1cm} (3.5.151)

we can rewrite equation 3.5.146 to read

\[ (-i\gamma^\mu \Lambda^\nu_\mu \frac{\partial}{\partial x'^\nu} + m)S^{-1}\psi'(x') = 0 \]  \hspace{1cm} (3.5.152)

After multiplying from the left by \( S \) we find

\[ -iS\gamma^\mu \Lambda^\nu_\mu S^{-1} \frac{\partial}{\partial x'^\nu} \psi'(x') + m\psi'(x') = 0 \]  \hspace{1cm} (3.5.153)

and this equation is exactly of the form that we want if

\[ S\gamma^\mu \Lambda^\nu_\mu S^{-1} = \gamma^\nu \]  \hspace{1cm} (3.5.154)
This means that
\[ S^{-1} \gamma^\nu S = \Lambda^\nu_{\, \mu} \gamma^\mu . \] (3.5.155)

Therefore, it follows that \( \Lambda^\nu_{\, \mu} \gamma^\mu \) must satisfy the algebra of the \( \gamma \)-matrices. This is easy to check.

\[
\begin{align*}
\gamma^\mu \Lambda^\nu_{\, \mu} \gamma^\rho \Lambda^\tau_{\, \rho} + \gamma^\rho \Lambda^\tau_{\, \rho} \gamma^\mu \Lambda^\nu_{\, \mu} &= \Lambda^\nu_{\, \mu} \Lambda^\tau_{\, \rho} (\gamma^\mu \gamma^\rho + \gamma^\rho \gamma^\mu) \\
&= 2 \Lambda^\nu_{\, \mu} \Lambda^\tau_{\, \rho} g^{\mu \rho} \\
&= 2 g^{\nu \tau} .
\end{align*}
\] (3.5.156)

This shows that \( S(\Lambda) \) exists. We construct \( S(\Lambda) \) explicitly a little later, but first we determine some further conditions on \( S \). From
\[
(\gamma^0)^2 = 1 , \quad \gamma^0 \dagger = \gamma^0 , \quad \gamma^k \dagger = -\gamma^k , \quad (\gamma^0 \gamma^\mu) \dagger = \gamma^0 \gamma^\mu
\] (3.5.157)
we can conclude that
\[
\gamma^\mu \dagger = \gamma^0 \gamma^\mu \gamma^0 .
\] (3.5.158)

Taking the hermitean adjoint of 3.5.155 we get
\[
(S^{-1} \gamma^\nu S)^\dagger = (\gamma^\mu \Lambda^\nu_{\, \mu})^\dagger .
\] (3.5.159)

After multiplying by \( \gamma^0 \) from the left and right we find
\[
\begin{align*}
\gamma^0 \gamma^\mu \dagger \gamma^0 \Lambda^\nu_{\, \mu} &= \gamma^\mu \Lambda^\nu_{\, \mu} = S^{-1} \gamma^\nu S \\
&= \gamma^0 (S^{-1} \gamma^\nu S)^\dagger \gamma^0 \\
&= \gamma^0 S^\dagger \gamma^0 \gamma^0 \gamma^\nu \dagger \gamma^0 S^\dagger -1 \gamma^0 \\
&= (\gamma^0 S^\dagger \gamma^0) \gamma^\nu (\gamma^0 S^\dagger \gamma^0)^{-1} .
\end{align*}
\] (3.5.160)

Therefore,
\[
S \gamma^0 S^\dagger \gamma^0 \gamma^\nu = \gamma^\nu S \gamma^0 S^\dagger \gamma^0 .
\] (3.5.161)

Thus,
\[
[S \gamma^0 S^\dagger \gamma^0 , \gamma^\nu] = 0 .
\] (3.5.162)

This means that
\[
S \gamma^0 S^\dagger \gamma^0 = b 1
\] (3.5.163)

where \( b \) is a constant. Furthermore, since \( \gamma^0 \) is hermitean we see that \( b \) must be real
\[
b = b^* .
\] (3.5.164)

We choose the normalization of \( S \) such that
\[
\text{det} S = 1 .
\] (3.5.165)
In that case we find that
\[ b^4 = 1 \quad , \quad b = \pm 1 \quad . \] (3.5.166)

Now consider
\[ S^\dagger S = S^\dagger \gamma^0 \gamma^0 S \]
\[ = b \gamma^0 S^{-1} \gamma^0 S \]
\[ = b \gamma^0 \gamma^\mu \Lambda^0_\mu \]
\[ = b(\Lambda^{00} - \Lambda^{0k} \alpha_k) \quad . \] (3.5.167)

Next, recalling that \( \text{Tr} \alpha_k = 0 \) and that \( S^\dagger S > 0 \), we take the trace of this equation and get that
\[ 4b \Lambda^{00} > 0 \quad . \] (3.5.168)

Thus,
\[ b = 1 \text{ if } \Lambda^{00} > 0 \quad , \quad b = -1 \text{ if } \Lambda^{00} < 0 \quad . \] (3.5.169)

Now,
\[ \overline{\psi} = \psi^\dagger \gamma^0 \quad . \] (3.5.170)

But,
\[ \psi' = S\psi \quad . \] (3.5.171)

Hence, we find
\[ (\psi')^\dagger = \psi^\dagger S^\dagger \] (3.5.172)
so that
\[ \overline{\psi'} = \psi'^\dagger \gamma^0 = \psi^\dagger S^\dagger \gamma^0 \] (3.5.173)
or
\[ \overline{\psi'} = b \overline{\psi} S^{-1} \quad . \] (3.5.174)

This again shows that \( \overline{\psi} \gamma^\mu \psi \) is indeed a four-vector for orthochronous Lorentz transformations \( (b = 1) \) since
\[ (\overline{\psi} \gamma^\mu \psi)^\prime(x) = \overline{\psi'}(x) \gamma^\mu \psi'(x) \]
\[ = b \overline{\psi}(x) S^{-1}(\Lambda) \gamma^\mu S(\Lambda) \psi(x) \]
\[ = b \Lambda^\mu_\nu(\overline{\psi} \gamma^\nu \psi)(x) \quad . \] (3.5.175)

Now, we are ready to construct the matrices \( S(\Lambda) \). But before doing so we again look at the infinitesimal Lorentz transformations.

A finite Lorentz transformation is written
\[ x'^\mu = \Lambda^\mu_\nu x^\nu \quad \text{or simply} \quad x' = \Lambda x \quad . \] (3.5.176)
Corresponding to this we consider $\Lambda_{\mu\nu}$ infinitesimal. In this case we write

$$\Lambda_{\mu\nu} = \delta_{\mu\nu} + \epsilon \omega_{\mu\nu}$$ \hspace{1cm} (3.5.177)

where

$$\omega^{\mu\nu} = -\omega^{\nu\mu}$$ \hspace{1cm} (3.5.178)

are 6 independent matrices corresponding to 3 rotations and 3 pure Lorentz transformations or "boosts". To construct the corresponding $S$ for this infinitesimal transformation we try

$$S = 1 - \frac{i}{4} \epsilon \sigma_{\mu\nu} \omega^{\mu\nu}$$ \hspace{1cm} (3.5.179)

and look for a matrix

$$\sigma_{\mu\nu} = -\sigma_{\nu\mu}$$ \hspace{1cm} (3.5.180)

such that 3.5.155

$$S^{-1} \gamma^\nu S = \gamma^\mu \Lambda_{\mu\nu}$$ \hspace{1cm} (3.5.181)

is satisfied. From 3.5.179 we get that to order $\epsilon$

$$S^{-1} = 1 + \frac{i}{4} \epsilon \sigma_{\mu\nu} \omega^{\mu\nu}.$$ \hspace{1cm} (3.5.182)

Now writing out 3.5.181 we have

$$\gamma^\mu - \frac{i}{4} \epsilon \omega^{\lambda\rho} [\gamma^\mu \sigma_{\lambda\rho} - \sigma_{\lambda\rho} \gamma^\mu] = \gamma^\mu + \epsilon \omega^{\mu \nu} \gamma^\nu.$$ \hspace{1cm} (3.5.183)

Therefore,

$$2i [g^\mu_\lambda \gamma_\lambda - g^\mu_\lambda \gamma_\nu] = [\gamma^\mu, \sigma_{\lambda \nu}].$$ \hspace{1cm} (3.5.184)

The best way to solve this equation is to guess the answer

$$\sigma_{\mu\nu} = \frac{i}{2} [\gamma_\mu, \gamma_\nu].$$ \hspace{1cm} (3.5.185)

Then,

$$S = 1 - \frac{i}{4} \epsilon \sigma_{\mu\nu} \omega^{\mu\nu} = 1 + \frac{\epsilon}{8} [\gamma_\mu, \gamma_\nu] \omega^{\mu\nu}.$$ \hspace{1cm} (3.5.186)

A finite Lorentz transformation $\Lambda^\nu_{\mu}$ is obtained from an infinitesimal by "exponentiation" from the infinitesimal transformation $\delta^\nu_{\mu} + \epsilon \omega^\nu_{\mu}$. We now do this. To begin, consider a boost with speed $v = \tanh \chi$ in the 1-direction. Then,

$$\Lambda^\nu_{\mu} = \begin{pmatrix} \cosh \chi & -\sinh \chi & 0 & 0 \\ -\sinh \chi & \cosh \chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$ \hspace{1cm} (3.5.187)
\[ \epsilon \omega_{\nu} = \epsilon \chi \left( \begin{array}{cccc} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) = \epsilon \chi M^\nu_{\nu} . \quad (3.5.188) \]

Clearly the “matrix” \( M = M^\nu_{\nu} \) satisfies

\[ M^2 = \left( \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \quad (3.5.189) \]

so that,

\[ M^3 = M . \quad (3.5.190) \]

Notice also that

\[ M^{\nu\mu} = \left( \begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) . \quad (3.5.191) \]

A finite Lorentz transformation can now be written

\[ x'^{\nu} = \Lambda^{\nu}_{\mu} x^{\mu} \]

\[ = \lim_{N \to \infty} \left( \delta + \frac{\chi}{N} M \right)^{\nu}_{\alpha_1} \left( \delta + \frac{\chi}{N} M \right)^{\alpha_2}_{\alpha_1} \ldots \left( \delta + \frac{\chi}{N} M \right)^{\alpha_N}_{\mu} x^{\mu} \]

\[ = (\exp(\chi M))^{\nu}_{\mu} x^{\mu} . \quad (3.5.192) \]

Expanding \( \exp(\chi M) \) and using \( M^3 = M \), \( M^4 = M^2 \) etc. we get

\[ \exp(\chi M) = 1 + \sum_{n=1}^{\infty} \frac{\chi^n M^n}{n!} \]

\[ = 1 + \frac{\chi M}{1!} + \frac{\chi^2 M^2}{2!} + \frac{\chi^3 M^3}{3!} + \frac{\chi^4 M^4}{4!} \ldots \]

\[ = 1 - M^2 + M^2 \sum_{n=0}^{\infty} \frac{\chi^{2n}}{(2n)!} + M \sum_{n=0}^{\infty} \frac{\chi^{2n+1}}{(2n+1)!} \]

\[ = 1 - M^2 + M^2 \cosh \chi + M \sinh \chi . \quad (3.5.193) \]

So, we recover the result we started with.

\[ \exp(\chi M) = \left( \begin{array}{cccc} \cosh \chi & -\sinh \chi & 0 & 0 \\ -\sinh \chi & \cosh \chi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \quad (3.5.194) \]
where, as stated at the beginning, we have

$$\cosh \chi = \frac{1}{\sqrt{1 - v^2/c^2}}, \quad \sinh \chi = \frac{v/c}{\sqrt{1 - v^2/c^2}}.$$  (3.5.195)

It is also fairly obvious that we can repeat this calculation for any direction.

Next we look at $S(\Lambda)$. We already found that for an infinitesimal Lorentz transformation

$$S(\Lambda) = 1 - \frac{i}{4} \epsilon \sigma_{\mu \nu} \omega^{\mu \nu}.$$  (3.5.196)

Therefore, for example, for a finite boost in the 1-direction we have

$$S(\Lambda) = \exp[-\frac{i}{4} \chi \sigma_{01}].$$  (3.5.197)

and for a rotation about the 3-axis by $\varphi$ we get

$$S(R_3(\varphi)) = \exp[i \frac{\varphi}{2} \sigma_{12}].$$  (3.5.198)

More generally, for a rotation about the axis $\hat{n}$ we find

$$S(R_n(\varphi)) = \exp[i \frac{\varphi}{2} \sigma \tilde{\Sigma} \cdot \hat{n}]$$  (3.5.199)

where

$$\tilde{\Sigma} = (\sigma_{23}, \sigma_{31}, \sigma_{12})$$  (3.5.200)

For rotations, $S$ is unitary since

$$\sigma_{ij} = \sigma_{ij}^\dagger.$$  (3.5.201)

However, for pure boosts this is not so and we have

$$S(\Lambda)^{-1} = \gamma^0 S^\dagger(\Lambda) \gamma^0.$$  (3.5.202)

This relation also holds for rotations since

$$[\gamma^0, \sigma_{ij}] = 0.$$  (3.5.203)

To see this we simply use the infinitesimal form. It then follows immediately that for proper orthochronous Lorentz transformations

$$S(\Lambda)^{-1} = 1 + \frac{i}{4} \epsilon \sigma_{\mu \nu} \omega^{\mu \nu}.$$  (3.5.204)

But, 

$$S(\Lambda)^\dagger = 1 + \frac{i}{4} \epsilon \sigma_{\mu \nu}^\dagger \omega^{\mu \nu}.$$  (3.5.205)

and therefore,

$$\gamma^0 S(\Lambda)^\dagger \gamma^0 = 1 + \frac{i}{4} \epsilon \gamma^0 \sigma_{\mu \nu}^\dagger \gamma^0 \omega^{\mu \nu} = 1 + \frac{i}{4} \epsilon \sigma_{\mu \nu} \omega^{\mu \nu} = S(\Lambda)^{-1}$$  (3.5.206)

as stated.
3.6 The Discrete Transformations

3.6.1 Charge Conjugation

Consider the transposed $\gamma$-matrices, $\gamma^{\nu t}$. Clearly they satisfy the same algebra as the original matrices

$$[\gamma^{\mu t}, \gamma^{\nu t}]_+ = 2g^{\mu\nu}. \quad (3.6.207)$$

This implies that there exists a matrix $B$ such that

$$\gamma^{\nu t} = B^{-1} \gamma^{\nu} B. \quad (3.6.208)$$

Or, taking the transpose of this equation

$$\gamma^{\nu} = B^t \gamma^{\mu t} B^{-1 t} = B^t B^{-1} \gamma^{\nu} B B^{-1 t}. \quad (3.6.209)$$

Hence, we find that

$$B^t B^{-1} \gamma^{\nu} = \gamma^{\nu} B^t B^{-1} \quad (3.6.210)$$

Therefore,

$$B^t B^{-1} = a \mathbf{1} \quad (3.6.211)$$

where $a$ is a constant. But,

$$\det(B^t B^{-1}) = 1 = a^4. \quad (3.6.212)$$

So, we see that $a = \pm 1$ or $\pm i$. Now, we choose the matrix $B$ to be unitary. Then,

$$B^\dagger = B^{-1} = B^{* t} \quad (3.6.213)$$

and since

$$B^t B^{-1} = a \mathbf{1} = B^{-1} B^t \quad (3.6.214)$$

we find, by taking the inverse, that

$$B B^{-1 t} = \frac{1}{a} \quad (3.6.215)$$

After complex conjugation this equation yields

$$B^* B^{-1 \dagger} = \frac{1}{a^*} \quad (3.6.216)$$

Furthermore, since

$$B^\dagger = B^{-1} \ , \ B^* = B^{-1 t} \quad (3.6.217)$$

we find that 3.6.216 becomes

$$B^* B = B^{-1 t} B = \frac{1}{a^*}. \quad (3.6.218)$$
Taking the inverse of this equation we obtain
\[ B^t B^{-1} = a^* = a \]  
(3.6.219)

So \( a \) is real and we have that \( a = \pm 1 \). We first consider the case \( a = -1 \) Then, \( B = -B^t \) and
\[ \gamma^\nu B = BB^{-1} \gamma^\nu B = B\gamma^\nu = -B^t \gamma^\nu = -(\gamma^\nu B)^t \]  
(3.6.220)

Also,
\[ \gamma^5 B = - (\gamma^5 B)^t \]  
(3.6.221)

This means that the matrices \( B, \gamma^5 B, \gamma^\mu B \) form a set of 6 antisymmetric matrices. Similarly, \( \gamma^5 \sigma^{\mu\nu} B, \sigma^{\mu\nu} B \) form a set of 10 symmetric matrices. (See problem 3.3a) This is consistent.

We next consider the case \( a = 1 \). Then, \( \gamma^5 \sigma^{\mu\nu} B, \sigma^{\mu\nu} B \) form a set of 10 antisymmetric matrices. (See problem 3.3b) They are linearly independent, but there are only six such antisymmetric \( 4 \times 4 \) matrices possible. Thus, this case is not consistent. Hence,
\[ B = -B^t \]  
(3.6.222)

Now, we define a matrix \( C \) by
\[ C = -\gamma^5 B \]  
(3.6.223)

This matrix satisfies
\[ C^{-1} \gamma^\mu C = -\gamma^\mu^t \]  
(3.6.224)

Also,
\[ C^t = -B^t \gamma^5 = BB^{-1} \gamma^5 B = \gamma^5 B = -C \]  
(3.6.225)

We now demonstrate the use of this matrix \( C \).

Consider the Dirac equation and its conjugate in the presence of an electromagnetic field.
\[
(-i \gamma^\mu \partial_\mu + m)\psi = e \gamma^\mu A_\mu \psi
\]
\[
(i \partial_\mu \bar \psi \gamma^\mu + m \bar \psi) = e A_\mu \bar \psi \gamma^\mu
\]  
(3.6.226)

If we transpose the second equation and use that \( \gamma^\mu^t = -C^{-1} \gamma^\mu C \) we find, after multiplying the resultant equation by \( C \), that
\[ (-i \gamma^\mu \partial_\mu + m)C\psi^t = -e \gamma^\mu A_\mu C\bar \psi^t \]  
(3.6.227)

Therefore, if we define
\[ \psi^C(x) = C\psi^t(x) \]  
(3.6.228)

we see that
\[ (-i \gamma^\mu \partial_\mu + m)\psi^C = -e \gamma^\mu A_\mu \psi^C \]  
(3.6.229)
Thus, $\psi^C$ is again a solution of the Dirac equation except that the sign of the charge and magnetic moment has been reversed.

Next we examine the Lorentz covariance of $\psi^C$. We want the equation
\[
\psi'(x') = S(\Lambda)\psi(x) \tag{3.6.230}
\]
to imply that
\[
\psi'^C(x') = S(\Lambda)\psi^C(x) \tag{3.6.231}
\]
But,
\[
\bar{\psi}'(x') = \pm \bar{\psi}(x)S(\Lambda)^{-1} \tag{3.6.232}
\]
The $\pm$ sign is to allow for the possibility of time reversal as we found before. So,
\[
\psi'^C(x') &= C \left( \bar{\psi}'(x') \right)^t \\
&= \pm C (\bar{\psi}(x)S(\Lambda)^{-1})^t \\
&= \pm CS(\Lambda)^{-1t} \bar{\psi}^t(x) \\
&= \pm CS(\Lambda)^{-1t} C^{-1} C \psi^t(x) \\
&= \pm CS(\Lambda)^{-1t} C^{-1} \psi^C(x) \tag{3.6.233}
\]
Thus, we need
\[
\pm CS(\Lambda)^{-1t} C^{-1} = S(\Lambda) \tag{3.6.234}
\]
or else
\[
S(\Lambda)^t = \pm C^{-1} S(\Lambda)^{-1} C \tag{3.6.235}
\]
Here, as before, the $+$ sign applies if the Lorentz transformation is orthochronous ($\Lambda^{00} > 0$) and the $-$ sign applies if we have time reversal ($\Lambda^{00} < 0$).

We can also use these results to limit the possible choices for the operators for parity and time reversal.

### 3.6.2 Parity

We are looking for a representative $S(P)$ of the parity operator $P$ such that for
\[
P(x^0, \vec{x}) = (x^0, -\vec{x}) \tag{3.6.236}
\]
we get
\[
S(P)^{-1} \gamma^0 S(P) = \gamma^0 \quad , \quad S(P)^{-1} \gamma^k S(P) = -\gamma^k \tag{3.6.237}
\]
where
\[
\text{det}[S(P)] = 1 \tag{3.6.238}
\]
So we can choose

\[ S(P) = \pm \gamma^0 \quad \text{or} \quad S(P) = \pm i\gamma^0 \quad . \quad (3.6.239) \]

With,

\[ C = -\gamma^5 B \quad (3.6.240) \]

we need, since \( \Lambda^{00} > 0 \) that

\[ S(P)^t = C^{-1}S(P)^{-1}C \quad . \quad (3.6.241) \]

If we try \( S(P) = \pm \gamma^0 \) we get

\[
C^{-1}S(P)^{-1}C = -B^{-1}\gamma^5(\pm \gamma^0)(-\gamma^5 B) \\
= \pm B^{-1}(-\gamma^0)B = \mp \gamma^0 t \quad . \quad (3.6.242)
\]

Therefore,

\[ C^{-1}S(P)^{-1}C \neq S(P)^t \quad . \quad (3.6.243) \]

So, we have a contradiction and we try \( S(P) = \pm i\gamma^0 \) Then,

\[
C^{-1}S(P)^{-1}C = -B^{-1}\gamma^5(\mp i\gamma^0)(-\gamma^5 B) \\
= \mp iB^{-1}(-\gamma^0)B = \pm i\gamma^0 t \\
= S(P)^t
\quad (3.6.244)
\]

as required and we conclude that this is consistent.

### 3.6.3 Time Reversal

Here we want a representative \( S(T) \) of the time reversal operator \( T \) such that if

\[ T(x^0, \bar{x}) = (-x^0, \bar{x}) \quad (3.6.245) \]

then,

\[ S(T)^{-1}\gamma^0 S(T) = -\gamma^0 \quad , \quad S(T)^{-1}\gamma^k S(T) = \gamma^k \quad (3.6.246) \]

with

\[ \det[S(T)] = 1 \quad . \quad (3.6.247) \]

Thus, we can choose either

\[ S(T) = \pm \gamma^5 \gamma^0 \quad \text{or else} \quad S(T) = \pm i\gamma^5 \gamma^0 \quad . \quad (3.6.248) \]

In either case we need

\[ S(T)^t = -C^{-1}S(T)^{-1}C \quad . \quad (3.6.249) \]

Now, suppose we use

\[ S(T) = \pm \gamma^5 \gamma^0 \quad . \quad (3.6.250) \]
So, with $C = -\gamma^5 B$ we get
\[
C^{-1}S(T)^{-1}C = -B^{-1}\gamma^5(\pm i\gamma^0\gamma^5)(-\gamma^5 B)
= \pm B^{-1}(\gamma^5\gamma^0)B = (\pm i\gamma^5\gamma^0)^t
= S(T)^t.
\] (3.6.251)

So, we have a contradiction. Therefore, we try
\[
S(T) = \pm i\gamma^5\gamma^0.
\] (3.6.252)

Then,
\[
C^{-1}S(T)^{-1}C = -B^{-1}\gamma^5(\mp i\gamma^0\gamma^5)(-\gamma^5 B)
= \mp iB^{-1}(\gamma^5\gamma^0)B = -(\pm i\gamma^5\gamma^0)^t
= -S(T)^t.
\] (3.6.253)

this yields a consistent solution.

For the sake of concreteness we illustrate these results using the definite representation in which
\[
\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \quad \gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\] (3.6.254)

With this choice we find that
\[
\gamma^0 t = \gamma^0 \quad \gamma^2 t = \gamma^2 \quad \gamma^4 t = -\gamma^1 \quad \gamma^3 t = -\gamma^3.
\] (3.6.255)

Also,
\[
(\gamma^0)^{-1} = \gamma^0 \quad (\gamma^2)^{-1} = -\gamma^2 \quad (\gamma^1)^{-1} = -\gamma^1 \quad (\gamma^3)^{-1} = -\gamma^3.
\] (3.6.256)

Thus, as is easily checked, a possible choice of the matrix $B$ is
\[
B = \gamma^1\gamma^3
\] (3.6.257)
and then
\[
C = -\gamma^5 B = -\gamma^5\gamma^1\gamma^3 = -i\gamma^0\gamma^2.
\] (3.6.258)

Again, we can choose
\[
S(P) = \pm i\gamma^0.
\] (3.6.259)

But, for $S(T)$ another possibility for a choice is
\[
S(T) = \pm i\gamma^5 C = \mp iB.
\] (3.6.260)

This works because we have
\[
C^{-1}S(T)^{-1}C = -i\gamma^2\gamma^0(\mp i\gamma^3\gamma^1)(-i\gamma^0\gamma^2)
= \pm i\gamma^2\gamma^0\gamma^3\gamma^1\gamma^0\gamma^2
= \mp i\gamma^3\gamma^1
= S(T)^t.
\] (3.6.261)
To summarize: For the representation chosen, a consistent set of discrete transformation matrices are given by

\[
\begin{align*}
C &= -\gamma^5 B = -i\gamma^0 \gamma^2 \\
S(P) &= \pm i\gamma^0 \\
S(T) &= \mp iB = \mp i\gamma^1 \gamma^3 .
\end{align*}
\] (3.6.262)

3.6.4 Charge Conjugate Spinors

Now, we look at the behaviour of plane wave solutions under these transformations. We have

\[
(\gamma \cdot p - m) u_+ (\vec{p}) = 0
\] (3.6.263)

as well as

\[
(-\gamma \cdot p - m) v (\vec{p}) = 0 .
\] (3.6.264)

Where, as always,

\[
v (\vec{p}) = u_- (\vec{p}) .
\] (3.6.265)

In addition we have

\[
\bar{v} (\vec{p}) (-\gamma \cdot p - m) = 0 .
\] (3.6.266)

After transposing this equation we get

\[
(-\gamma^t \cdot p - m) \bar{v}^t (\vec{p}) = 0 .
\] (3.6.267)

We next multiply by \( C \) from the left and get

\[
-C \gamma^t \cdot p C^{-1} C \bar{v}^t (\vec{p}) - C m \bar{v}^t (\vec{p}) = 0
\] (3.6.268)

or

\[
(\gamma \cdot p - m) C \bar{v}^t (\vec{p}) = 0 .
\] (3.6.269)

Therefore, if we define

\[
u^C (\vec{p}) = C \bar{v}^t (\vec{p})
\] (3.6.270)

then we find that

\[
(\gamma \cdot p - m) u^C (\vec{p}) = 0 .
\] (3.6.271)

Thus, \( u^C (\vec{p}) \) is a positive energy solution with momentum \( \vec{p} \). Furthermore, we find from the defining equation (see problem 3.4) that

\[
v (\vec{p}) = C [u^C (\vec{p})]^t .
\] (3.6.272)
3.7 Problems

3.1 Show that
\[ \text{Tr} \left( \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \right) = 4 \left( g^{\mu\sigma} g^{\nu\rho} + g^{\mu\nu} g^{\rho\sigma} - g^{\mu\rho} g^{\nu\sigma} \right) \]  \hfill (3.7.273)

3.2 For the representation
\[ \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \]  \hfill (3.7.274)

find \( \gamma^5 \) and a matrix \( B \) such that
\[ \gamma^\mu^t = B^{-1} \gamma^\mu B \]  \hfill (3.7.275)

Also find the operator
\[ C = -\gamma^5 B \]  \hfill (3.7.276)

and consistent parity and time-reversal operators \( S(P) \) and \( S(T) \).

3.3a) Show that if
\[ B^t = -B \]
then \( \sigma^{\mu\nu} B \) and \( \gamma^5 \sigma^{\mu\nu} B \) are symmetric matrices.

b) Show that if
\[ B^t = B \]
then \( \sigma^{\mu\nu} B \) and \( \gamma^5 \sigma^{\mu\nu} B \) are antisymmetric matrices.

3.4 Use the explicit representation for \( C \) given in the text to show that
\[ C^t = C^{-1} = -C \]
and use this to prove that
\[ \psi(\vec{p}) = C[\bar{\psi} C(\vec{p})] \]  .

3.5 Determine the Lorentz, parity, and time-reversal properties of
\[ \bar{\psi} \gamma^5 \psi, \bar{\psi} \gamma^\mu \gamma^5 \psi, \bar{\psi} \sigma^{\mu\nu} \gamma^5 \psi \] .

3.5 Compute
a) \( \gamma_\mu \gamma^\mu, \gamma_\mu \gamma_\alpha \gamma^\mu, \gamma_\mu \gamma_\alpha \gamma_\beta \gamma^\mu \) .

b) Simplify
\[ \psi \gamma^\mu \psi, \psi \gamma^\mu \gamma^\nu \psi \] .

c) Compute the following traces
\[ \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma, \]
\[ (\gamma \cdot p + m) \gamma^5 (\gamma \cdot p + \gamma \cdot q + m) \gamma^5 (\gamma \cdot p + \gamma \cdot q + m) \gamma^5 \]
where \( p, q \) are distinct four-vectors with \( p^2 = m^2 \) and \( q^2 = M^2 \).

3.6 In the Majorana representation all of the \( \gamma^\mu \) are pure imaginary. Find an explicit form of the Majorana representation.
Bibliography


Chapter 4

Structure of Dirac Particles

4.1 Electromagnetic Interaction

Just as in the case of the Klein-Gordon equation we introduce coupling to an electromagnetic field to obtain an interpretation of the structure of Dirac particles. Again the coupling is minimal and we simply replace

\[ p^\mu \rightarrow p^\mu - \frac{e}{c} A^\mu . \]  

(4.1.1)

So, the Dirac equation becomes

\[ \gamma^\mu \left( i \hbar \partial_\mu - \frac{e}{c} A_\mu (x) \right) \psi (x) = mc \psi (x) \]  

(4.1.2)

or

\[ \left( \not{p} - \frac{e}{c} \not{A} - mc \right) \psi = 0 \]  

(4.1.3)

To get an equation similar to the Klein-Gordon equation we set

\[ \psi = \frac{1}{mc} \left( \not{p} - \frac{e}{c} \not{A} + mc \right) \chi . \]  

(4.1.4)

We then have

\[ \left( \not{p} - \frac{e}{c} \not{A} \right)^2 \chi = m^2 c^2 \chi . \]  

(4.1.5)

Now, we rewrite the operator on the left hand side.

\[ \text{LHS} = \gamma^\mu \gamma^\nu \left( p_\mu - \frac{e}{c} A_\mu \right) \left( p_\nu - \frac{e}{c} A_\nu \right) \chi \]

\[ = \frac{1}{2} \left[ \gamma^\mu \gamma^\nu \left( p_\mu - \frac{e}{c} A_\mu \right) \left( p_\nu - \frac{e}{c} A_\nu \right) + \gamma^\nu \gamma^\mu \left( p_\nu - \frac{e}{c} A_\nu \right) \left( p_\mu - \frac{e}{c} A_\mu \right) \right] \chi \]

\[ = \frac{1}{2} \gamma^\mu \gamma^\nu \left[ p_\mu p_\nu \frac{e^2}{c^2} A_\mu A_\nu - \frac{e}{c} (p_\mu A_\nu + A_\mu p_\nu) \right] \chi \]
\[
\begin{align*}
&+ \frac{1}{2} \gamma^\nu \gamma^\mu \left[ p_\nu p_\mu + \frac{e^2}{c^2} A_\mu A_\nu - \frac{e}{c} (p_\nu A_\mu + A_\nu p_\mu) \right] \chi \\
&= \frac{1}{2} (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) \left[ p_\mu p_\nu + \frac{e^2}{c^2} A_\mu A_\nu - \frac{e}{c} (p_\mu A_\nu + A_\mu p_\nu) \right] \chi \\
&+ \frac{e}{2c} \gamma^\nu \gamma^\mu \left[ p_\mu A_\nu + A_\mu p_\nu - p_\nu A_\mu - A_\nu p_\mu \right] \chi \\
&= \left( p - \frac{e}{c} A \right)^2 \chi + \frac{e}{2c} i \hbar \gamma^\nu \gamma^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) \chi . \tag{4.1.6}
\end{align*}
\]

Next, we introduce the electromagnetic field tensor
\[
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{4.1.7}
\]
and rewrite
\[
\gamma^\mu \gamma^\nu = \frac{1}{2} \left[ (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) + (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \right] \\
= g^{\mu\nu} - i \sigma^{\mu\nu} \tag{4.1.8}
\]
Here we have introduced the antisymmetric tensor
\[
\sigma^{\mu\nu} = \frac{i}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \tag{4.1.9}
\]
that we already saw when we discussed Lorentz invariance. Using these results we find that the operator on the left hand side of 4.1.5 becomes
\[
\text{LHS} = \left( p - \frac{e}{c} A \right)^2 - \frac{e \hbar}{2c} \sigma^{\mu\nu} F_{\mu\nu} \tag{4.1.10}
\]
Thus, the equation for \( \chi \) reduces to
\[
\left[ (p_\mu - \frac{e}{c} A_\mu)(p^\mu - \frac{e}{c} A^\mu) - \frac{e \hbar}{2c} \sigma^{\mu\nu} F_{\mu\nu} \right] \chi = m^2 c^2 \chi . \tag{4.1.11}
\]
The term \((\hbar/2) \sigma^{\mu\nu} F_{\mu\nu}\) may be written more explicitly in terms of the electric field \(\vec{E}\) and the magnetic field \(\vec{B}\) to get.
\[
\frac{e \hbar}{2c} \sigma^{\mu\nu} F_{\mu\nu} = 2 \frac{e}{2c} \left( \frac{\hbar}{2} \vec{\Sigma} \cdot \vec{B} - \frac{e}{2} \vec{\alpha} \cdot \vec{E} \right) . \tag{4.1.12}
\]
We now see that this term consists of a magnetic dipole interaction \(\hbar \frac{1}{2} \vec{\Sigma} \cdot \vec{B}\) with a magnetic dipole of 1 Bohr magneton and a \(g\)-factor = 2 as is in fact experimentally observed. In addition to the magnetic dipole interaction there is the electric monopole term \(-i \frac{e}{2} \vec{\alpha} \cdot \vec{E}\). So, we have found an interpretation for part of the structure of a Dirac particle. However, we now have a second order differential equation and we still have four components. Thus, we have doubled the number of degrees of freedom. To remedy this situation we first notice that the only matrix operators appearing in equation 4.1.11 are \(\sigma^{\mu\nu}\) and these commute with \(\gamma^5\).
\[
[\sigma^{\mu\nu}, \gamma^5] = 0 \tag{4.1.13}
\]
Furthermore, $\gamma^5$ is hermitian with eigenvalues $\pm 1$ as is obvious from our standard representation \[ 3.3.72. \] So we can separate our solutions for $\chi$ into two classes: $\chi_+$ those with chiral eigenvalue (eigenvalue of $\gamma^5$) having a value $+1$ and $\chi_-$ those with chiral eigenvalue $-1$. If we write $\chi_{\pm}$ in terms of two-component spinors in the form

$$ \chi = \begin{pmatrix} a \\ b \end{pmatrix} $$

(4.1.14)

then we find, using our standard representation

$$ \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} , \quad \gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} $$

(4.1.15)

that

$$ \chi_+ = \begin{pmatrix} u \\ v \end{pmatrix} , \quad \chi_- = \begin{pmatrix} u \\ -v \end{pmatrix} . $$

(4.1.16)

The projection operators onto states of definite chirality are

$$ \frac{1}{2} (1 \pm \gamma^5) $$

since

$$ \left[ \frac{1}{2} (1 \pm \gamma^5) \right]^2 = \frac{1}{2} (1 \pm \gamma^5) $$

(4.1.17)

and

$$ \gamma^5 \frac{1}{2} (1 \pm \gamma^5) = \pm \frac{1}{2} (1 \pm \gamma^5) $$

(4.1.18)

so that

$$ \frac{1}{2} (1 \pm \gamma^5) \chi_{\pm} = \chi_{\pm} . $$

(4.1.19)

It is also clear that

$$ \frac{1}{2} (1 \pm \gamma^5) \frac{1}{2} (1 \mp \gamma^5) = 0 $$

(4.1.20)

while

$$ \frac{1}{2} (1 \pm \gamma^5) + \frac{1}{2} (1 \mp \gamma^5) = 1 . $$

(4.1.21)

These projection operators occur frequently in modern theories of elementary particles.

We now show that if we restrict ourselves to solutions of definite chirality then we get a one-one correspondence with solutions of the first order Dirac equation. That is, for every $\psi$ satisfying the Dirac equation there exists a unique solution of definite chirality of the second order equation and vice-versa.

Proof
We demonstrate the proof for positive chirality since for negative chirality the proof is identical. Clearly, given $\chi_+$ the Dirac solution $\psi$ is uniquely determined by
\[
\psi = \frac{1}{mc} \left[ \gamma - \frac{e}{c} A + mc \right] \chi_+ .
\] (4.1.22)

To show the converse we multiply this equation by $(1 + \gamma^5)$ to get
\[
(1 + \gamma^5)\psi = \frac{1}{mc} \left[ \left( \gamma - \frac{e}{c} A + mc \right) \chi_+ + \left( -\gamma + \frac{e}{c} A + mc \right) \gamma^5 \chi_+ \right]
\]
\[
= \frac{1}{mc} \left[ \left( \gamma - \frac{e}{c} A \right) (1 - \gamma^5) \chi_+ + mc(1 + \gamma^5) \chi_+ \right]
\]
\[
= \frac{1}{mc} mc 2\chi_+
\] (4.1.23)

So we find
\[
\chi_+ = \frac{1}{2} (1 + \gamma^5)\psi.
\] (4.1.24)

To get further insight into the second order equations we examine the matrix operators $\sigma_{\mu\nu}$ that occur in these equations. We begin with $\sigma_{30}$
\[
\sigma_{30} = \frac{1}{2i} (\gamma_3 \gamma_0 - \gamma_0 \gamma_3) = -\frac{1}{i} \gamma_0 \gamma_3
\] (4.1.25)

Now we use the fact that
\[
\gamma_1 \gamma_2 \gamma_1 \gamma_2 = -(\gamma_1)^2 (\gamma_2)^2 = -1
\] (4.1.26)
to write
\[
\sigma_{30} = (\gamma_1 \gamma_2)^2 \frac{1}{i} (\gamma_0 \gamma_3)
\]
\[
= -\frac{i}{i} \gamma_1 \gamma_2 \gamma_0 \gamma_1 \gamma_2 \gamma_3
\]
\[
= -\frac{i}{i} \gamma_1 \gamma_2 \gamma^5
\]
\[
= -i\sigma_{12} \gamma^5.
\] (4.1.27)

This means that when acting on $\chi_+ = \gamma^5 \chi_+$ we can replace
\[
\sigma_{30} \text{ by } -i\sigma_{12}
\]
\[
\sigma_{10} \text{ by } -i\sigma_{23}
\]
\[
\sigma_{20} \text{ by } -i\sigma_{31}.
\] (4.1.28)

If we now again use our standard representation 3.3.72
\[
\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\] (4.1.29)
we find that the equation for $\chi_+$ becomes

$$\left[ \left( i \hbar \partial_\mu - \frac{e}{c} A_\mu \right)^2 + \frac{e \hbar}{2c} \bar{\sigma} \cdot (\bar{B} + i \bar{E}) \right] u = m^2 c^2 u$$  \hspace{1cm} (4.1.30)

where

$$\chi_+ = \begin{pmatrix} u \\ u \end{pmatrix}.$$  \hspace{1cm} (4.1.31)

Similarly, when acting on $\chi_- = -\gamma^5 \chi_-$ we can replace

$$\sigma_{30} \text{ by } i \sigma_{12},$$

$$\sigma_{10} \text{ by } i \sigma_{23},$$

$$\sigma_{20} \text{ by } i \sigma_{31}.$$  \hspace{1cm} (4.1.32)

If we now again, as above, use our standard representation 3.3.72

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}$$

$$\gamma^5 = i \gamma^0 \gamma^1 \gamma^2 \gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$  \hspace{1cm} (4.1.33)

we find that the equation for $\chi$ becomes

$$\left[ \left( i \hbar \partial_\mu - \frac{e}{c} A_\mu \right)^2 + \frac{e \hbar}{2c} \bar{\sigma} \cdot (\bar{B} - i \bar{E}) \right] \chi = m^2 c^2 \chi$$  \hspace{1cm} (4.1.34)

where $\chi$ is either $\chi_+$ or $\chi_-$ and

$$\chi_+ = \begin{pmatrix} u \\ u \end{pmatrix}, \quad \chi_- = \begin{pmatrix} v \\ -v \end{pmatrix}.$$  \hspace{1cm} (4.1.35)

In addition to the gauge-invariant, minimal coupling $p_\mu \rightarrow p_\mu - (e/c) A_\mu$ that we have discussed it is possible to have gauge-invariant direct coupling to the electromagnetic field tensor

$$F^{\mu \nu} = \partial^\nu A^\mu - \partial^\mu A^\nu$$  \hspace{1cm} (4.1.36)

by so-called Pauli terms

$$\frac{e \hbar}{2 c} \sigma^{\mu \nu} F_{\mu \nu}.$$  

These terms add an anomalous magnetic moment

$$\mu \frac{e \hbar}{2mc}$$

to the usual magnetic moment so that the equation of motion becomes

$$\left[ \gamma^\mu \left( i \hbar \partial_\mu - \frac{e}{c} A_\mu \right) + \frac{e \hbar \mu}{2c} \sigma^{\mu \nu} F_{\mu \nu} \right] \psi = mc \psi.$$  \hspace{1cm} (4.1.37)
4.2 Constants of the Motion

Constants of the motion are those dynamical variables that commute with the hamiltonian. Now (with $\hbar = c = 1$) we have for a free particle

$$H = \vec{\alpha} \cdot \vec{p} + \beta m.$$  \hfill (4.2.38)

With the usual definition of the orbital angular momentum

$$\vec{L} = \vec{r} \times \vec{p} = (L_1, L_2, L_3)$$  \hfill (4.2.39)

we find that none of the components of $\vec{L}$ commute with $H$. Thus, the orbital angular momentum is not a constant of the motion. On the other hand if we use our standard representation with

$$\gamma^0 = \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$  \hfill (4.2.40)

we find that

$$\vec{J} = \vec{L} + \frac{1}{2} \vec{\Sigma}$$  \hfill (4.2.41)

with

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$$  \hfill (4.2.42)

does commute with $H$ and thus is a constant of the motion. We now prove this statement. To do this we simply verify that this result is true for $J_1$. By symmetry the result then also holds for $J_2$ and $J_3$. Clearly,

$$[J_1, H] = [J_1, \vec{\alpha} \cdot \vec{p}]$$

$$= [L_1, \vec{\alpha} \cdot \vec{p}] + \frac{1}{2} [\Sigma_1, \vec{\alpha} \cdot \vec{p}].$$  \hfill (4.2.43)

Now, we have that

$$\vec{\Sigma} = \gamma^5 \vec{\alpha} = \vec{\alpha} \gamma^5.$$  \hfill (4.2.44)

This implies that we can write

$$\vec{\alpha} = \vec{\Sigma} \gamma^5 = \gamma^5 \vec{\Sigma}.$$  \hfill (4.2.45)

Therefore,

$$\frac{1}{2} [\Sigma_1, \vec{\alpha} \cdot \vec{p}] = \frac{1}{2} [\gamma^5 \alpha_1, \vec{\alpha} \cdot \vec{p}]$$

$$= \frac{1}{2} \gamma^5 [\alpha_1, \vec{\alpha} \cdot \vec{p}].$$  \hfill (4.2.46)

Next, we use that

$$[\alpha_1, \alpha_2] = 2i \Sigma_3$$

$$[\alpha_2, \alpha_3] = 2i \Sigma_1$$

$$[\alpha_3, \alpha_1] = 2i \Sigma_2$$  \hfill (4.2.47)
to get that

\[
\frac{1}{2} [\Sigma_1, \vec{\alpha} \cdot \vec{p}] = i \gamma^5 (\Sigma_3 p_2 - \Sigma_2 p_3) \\
= i(p_2 \alpha_3 - p_3 \alpha_2) \\
= i(\vec{p} \times \vec{\alpha})_1 .
\]  

(4.2.48)

On the other hand,

\[
[L_1, \vec{\alpha} \cdot \vec{p}] = \vec{\alpha} \cdot [L_1, \vec{p}] \\
= i\vec{\alpha} \cdot (0, p_3, -p_2) \\
= i(\alpha_2 p_3 - \alpha_3 p_2) \\
= -i(\vec{p} \times \vec{\alpha})_1 .
\]  

(4.2.49)

Thus, we have shown that

\[
[J_1, \vec{\alpha} \cdot \vec{p}] = 0 .
\]  

(4.2.50)

By symmetry it now follows that quite generally

\[
[J, \vec{\alpha} \cdot \vec{p}] = 0 .
\]  

(4.2.51)

As a consequence we have that \(J^2\), \(J_3\), and \(H\) can be simultaneously diagonalized. Furthermore, since

\[
(\vec{\Sigma})^2 = 3 \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right)
\]  

(4.2.52)

we see that \((\vec{\Sigma})^2\) is also simultaneously diagonal. However, this is not possible for \((\vec{L})^2\) or for the spin-orbit coupling operator \(1/2 \vec{\Sigma} \cdot \vec{L}\). For this reason, and for later use we look for a different operator. A suitable candidate is

\[
K = \beta (\vec{\Sigma} \cdot \vec{L} + 1)
\]  

(4.2.53)

since

\[
[K, J] = [K, H] = 0 .
\]  

(4.2.54)

We also find that

\[
K^2 = L^2 + \Sigma \cdot L + 1 = \left( \vec{L} + \frac{1}{2} \vec{\Sigma} \right)^2 + \frac{1}{4} = J^2 + \frac{1}{4} .
\]  

(4.2.55)

Thus, the eigenvalues of \(K^2\) are simply

\[
j(j + 1) + \frac{1}{4} = \left( j + \frac{1}{2} \right)^2 .
\]  

(4.2.56)
4.3 Problems

4.1 If

\[ K = \beta(\vec{\Sigma} \cdot \vec{L} + \hbar) \]

show that

\[ K^2 = \left( \vec{L} + \frac{\hbar}{2} \vec{\Sigma} \right)^2 + \frac{\hbar^2}{2} . \]

Hint: Use the fact that

\[ (\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = \vec{a} \cdot \vec{b} + i \vec{\sigma} \cdot (\vec{a} \times \vec{b}) \]

and

\[ \vec{L} \times \vec{L} = i\hbar \vec{L} . \]

4.2 For a Dirac electron of mass \( m \) in an attractive electrostatic potential

\[ V(z) = \begin{cases} 
0 & z < 0, \ z > a \\
-V_0 & 0 < z < a 
\end{cases} \]

a) Find the energy levels.
b) Solve the problem of scattering of such an electron with momentum \( \vec{p} \) off this potential.

4.3 Find the exact energy eigenvalues and eigenfunctions for an electron in a uniform magnetic field.
Chapter 5

Dirac Equation: Central Potentials

5.1 Separation of Variables

The Dirac equation, in the case of interaction with an electromagnetic field, can
be solved exactly for several cases. The most important of these, in practice,
are:
1. The Coulomb potential
2. A homogeneous, static magnetic field. This, like the nonrelativistic case,
leads to Landau levels.
3. A plane electromagnetic wave, the so-called Volkov solutions.

Since the Coulomb potential has played such an important role, we treat
this case in detail. But first we discuss central potentials in general. With an
electrostatic potential \( V(r) \) the Dirac equation reads

\[
H = c\overrightarrow{\alpha} \cdot \overrightarrow{p} + \beta mc^2 + V(r)
\]  \hspace{1cm} (5.1.1)

It should be noted that \( V \) is the zero component of a four-vector and not a
scalar. If it were a scalar it would have appeared in the equation in the same
way as the mass term, namely as \( \beta V \). The constants of the motion of this
hamiltonian are:

\[
\overrightarrow{J} = \overrightarrow{r} \times \overrightarrow{p} + \frac{\hbar}{2} \overrightarrow{\Sigma}
\]

and \( P \)

where \( P \) is the parity operator. Also, we see that

\[
\overrightarrow{\Sigma} = \frac{1}{2i} \overrightarrow{\alpha} \times \overrightarrow{\alpha}.
\]  \hspace{1cm} (5.1.2)

Again using our standard representation of the \( \gamma \)-matrices

\[
\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\]  \hspace{1cm} (5.1.3)
we have
\[
\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}.
\] (5.1.4)

To solve the eigenvalue problem for the energies we put
\[
\Psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \quad \phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad \chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}.
\] (5.1.5)

Next, we look for simultaneous eigenstates of \(H\), \(J^2\), \(J_z\), \(P\).
\[
H \begin{pmatrix} \phi \\ \chi \end{pmatrix} = E \begin{pmatrix} \phi \\ \chi \end{pmatrix}
\] (5.1.6)
\[
J^2 \begin{pmatrix} \phi \\ \chi \end{pmatrix} = j(j + 1)\hbar^2 \begin{pmatrix} \phi \\ \chi \end{pmatrix}
\] (5.1.7)
\[
J_z \begin{pmatrix} \phi \\ \chi \end{pmatrix} = M\hbar \begin{pmatrix} \phi \\ \chi \end{pmatrix}
\] (5.1.8)

and
\[
\begin{pmatrix} \phi \\ \chi \end{pmatrix} = (-1)^{j + \tilde{\omega}/2} \begin{pmatrix} \phi \\ \chi \end{pmatrix}
\] (5.1.9)

where
\[
\tilde{\omega} = \begin{cases} 
+1 & \text{if parity is } (-1)^{j+1/2} \\
-1 & \text{if parity is } (-1)^{j-1/2}
\end{cases}
\] (5.1.10)

We add the orbital angular momentum \(L\) and the spin \(1/2\) to form states \(|jM\rangle\). The corresponding wave function has parity \((-1)^L\) and is written \(\mathcal{Y}^M_{lj}(\theta, \varphi)\).

The possible \(L\) values are:
\[
\mathcal{Y}^M_{lj}(\theta, \varphi) : \quad L = l + \frac{\varepsilon}{2} \quad P = (-1)^{j + \tilde{\omega}/2} = (-1)^l
\]
\[
\mathcal{Y}^M_{lj'}(\theta, \varphi) : \quad L' = l' = j - \frac{\varepsilon}{2} \quad P = (-1)^{j - \tilde{\omega}/2} = (-1)^{l'}
\] (5.1.11)

The two states just listed clearly have opposite parity. Explicit solutions for these states are
\[
j = l + \frac{1}{2}
\]
\[
\mathcal{Y}^M_{lj} = \begin{pmatrix} \sqrt{\frac{l+1/2+M}{2(l+1)}} Y_{l,M-1/2} \\ \sqrt{\frac{l+1/2-M}{2(l+1)}} Y_{l,M+1/2} \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{l+M}{2j}} Y_{j-1/2,M-1/2} \\ \sqrt{\frac{l-M}{2j}} Y_{j-1/2,M+1/2} \end{pmatrix}
\] (5.1.12)
\[
j = l - \frac{1}{2} \quad l > 0
\]
\[
\mathcal{Y}^M_{lj} = \begin{pmatrix} \sqrt{\frac{l-M+1}{2(j+1)}} Y_{j+1/2,M-1/2} \\ -\sqrt{\frac{l+M+1}{2(j+1)}} Y_{j+1/2,M+1/2} \end{pmatrix}
\] (5.1.13)
In both cases we have
\[ J^2 Y_{ij}^M = j(j+1)Y_{ij}^M \]  \hspace{1cm} (5.1.14)

Also,
\[ \vec{L} \cdot \vec{S} = 2\vec{L} \cdot \vec{s} = (J^2 - L^2 - 3/4) \]  \hspace{1cm} (5.1.15)

Next we write
\[ \phi = \frac{1}{r} F(r) Y_{ij}^M \]
\[ \chi = \frac{1}{r} G(r) Y_{ij}^M \]  \hspace{1cm} (5.1.16)

So, the state \( \Psi_{\omega,j}^M \) with angular momentum labels \( j, M \) and parity \( (-1)^{j+\omega/2} \) is given by
\[ \Psi_{\omega,j}^M = \frac{1}{r} \begin{pmatrix} F(r)Y_{ij}^M \\ G(r)Y_{ij}^M \end{pmatrix} \]  \hspace{1cm} (5.1.17)

We have now diagonalized all the constants of the motion except \( H \) itself. Thus, we still have to solve
\[ H \Psi_{\omega,j}^M = E \Psi_{\omega,j}^M \]  \hspace{1cm} (5.1.18)

To do this we go to spherical coordinates. We first realize that
\[ \vec{r} \cdot \vec{p} f = r \frac{1}{i} \frac{\partial f}{\partial r} \]  \hspace{1cm} (5.1.19)

and therefore define
\[ p_r f = \frac{1}{r} \frac{1}{i} \frac{\partial}{\partial r} (rf) = \frac{1}{r} (\vec{r} \cdot \vec{p} - i)f \]  \hspace{1cm} (5.1.20)

We also define
\[ \alpha_r = \frac{1}{r} \vec{\alpha} \cdot \vec{r} \]  \hspace{1cm} (5.1.21)

and recall the definition of
\[ K = \beta (\vec{\Sigma} \cdot \vec{L} + 1) \]  \hspace{1cm} (5.1.22)

Next, we prove the following lemma.

**Lemma**

If \( \vec{B} \) and \( \vec{C} \) are two matrix vector operators such that
\[ [\vec{\alpha}, \vec{B}] = [\vec{\alpha}, \vec{C}] = 0 \]  \hspace{1cm} (5.1.23)

then,
\[ (\vec{\alpha} \cdot \vec{B})(\vec{\alpha} \cdot \vec{C}) = \vec{B} \cdot \vec{C} + i\vec{\Sigma} \cdot (\vec{B} \times \vec{C}) \]  \hspace{1cm} (5.1.24)
Proof
First we note that since
\[
\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}
\] (5.1.25)

it follows that
\[
(\vec{\alpha} \cdot \vec{B})(\vec{\sigma} \cdot \vec{C}) = \begin{pmatrix} (\vec{\sigma} \cdot \vec{B})(\vec{\sigma} \cdot \vec{C}) & 0 \\ 0 & (\vec{\sigma} \cdot \vec{B})(\vec{\sigma} \cdot \vec{C}) \end{pmatrix}
\] (5.1.26)

and by explicit multiplication and the property of the Pauli matrices we find
\[
(\vec{\sigma} \cdot \vec{B})(\vec{\sigma} \cdot \vec{C}) = \vec{B} \cdot \vec{C} + i\vec{\sigma} \cdot (\vec{B} \times \vec{C})
\] (5.1.27)

thus proving the result. It therefore follows that
\[
(\vec{\alpha} \cdot \vec{r})(\vec{\alpha} \cdot \vec{p}) = \vec{r} \cdot \vec{p} + i\vec{\Sigma} \cdot \vec{L}
\] 
\[
= r \left( p_r + \frac{i}{r} \right) + i\vec{\Sigma} \cdot \vec{L}
\] (5.1.28)

Therefore,
\[
\alpha_r (\vec{\alpha} \cdot \vec{p}) = p_r + \frac{i}{r} + \frac{i}{r} \vec{\Sigma} \cdot \vec{L}
\] (5.1.29)

If we multiply this equation by \( \alpha_r \), from the left, and use the fact that \( \alpha_r^2 = 1 \) we find
\[
\vec{\alpha} \cdot \vec{p} = \alpha_r \left[ p_r + \frac{i}{r} \left( 1 + \vec{\Sigma} \cdot \vec{L} \right) \right]
\] 
\[
= \alpha_r \left[ p_r + \frac{i}{r} \beta K \right].
\] (5.1.30)

With these results we can rewrite the Hamiltonian as
\[
H = c\alpha_r p_r + \frac{i\hbar c}{r} \alpha_r \beta K + \beta mc^2 + V(r).
\] (5.1.31)

Before attacking this Hamiltonian we use the following two identities (See problem 5.1.)
\[
\vec{r} \times \vec{L} + \vec{L} \times \vec{r} = 2i\vec{r}
\]
\[
\vec{p} \times \vec{L} + \vec{L} \times \vec{p} = 2i\vec{p}
\] (5.1.32)

to show that the following operators commute
\[
[\beta, K] = 0
\]
\[
[\alpha_r, K] = 0
\]
\[
[p_r, K] = 0
\]
\[
[\alpha_r, p_r] = 0
\] (5.1.33)
Proof It is helpful to remember that

\[ K = \beta(\vec{\Sigma} \cdot \vec{L} + 1), \quad \vec{\Sigma} = \gamma^5 \vec{\alpha}. \]

Then,

\[
[\beta, K] = [\beta, \beta(\vec{\Sigma} \cdot \vec{L} + 1)] = \beta[\beta, \vec{\Sigma} \cdot \vec{L}]
= \beta[\beta, \gamma^5 \vec{\alpha} \cdot \vec{L}]
= \beta[\beta, \gamma^5 \vec{\alpha} \cdot \vec{L} + \beta \gamma^5 [\beta, \vec{\alpha}] \cdot \vec{L}]
= 2(\beta^2 \gamma^5 \vec{\alpha} \cdot \vec{L} + \beta \gamma^5 \beta \vec{\alpha} \cdot \vec{L}) = 0
\] (5.1.34)

\[
[\alpha_r, K] = [\alpha_r, \beta] + [\alpha_r, \beta \vec{\Sigma} \cdot \vec{L}]
= [\alpha_r, \beta] + \alpha_r \beta \gamma^5 \vec{\alpha} \cdot \vec{L} - \beta \gamma^5 \vec{\alpha} \cdot \vec{L} \alpha_r
= 2\alpha_r \beta - \beta \gamma^5 [(\vec{\alpha} \cdot \vec{r})(\vec{\alpha} \cdot \vec{L}) + (\vec{\alpha} \cdot \vec{L})(\vec{\alpha} \cdot \vec{r})]
= 2\alpha_r \beta - \beta \gamma^5 [\vec{r} \cdot \vec{L} + \vec{L} \cdot \vec{r} + i \vec{\Sigma} \cdot (\vec{r} \times \vec{L} + \vec{L} \times \vec{r})]
= 2\alpha_r \beta - \beta \gamma^5 i \vec{\Sigma} \cdot (2i \vec{r})
= 2(\alpha_r \beta + \beta \alpha_r) = 0.
\] (5.1.35)

\[
[p_r, K] = [p_r, \beta \vec{\Sigma} \cdot \vec{L}] = \beta \vec{\Sigma} \cdot [p_r, \vec{L}] = 0
\] (5.1.36)

and finally

\[
[\alpha_r, p_r] = 0
\] (5.1.37)

since

\[
\frac{\partial \vec{r}}{\partial r} = 0.
\] (5.1.38)

This means that the only noncommuting objects are \(p_r\) and functions of \(r\) as well as \(\alpha_r\) and \(\beta\). In fact, for these last two operators we have

\[
\alpha_r \beta + \beta \alpha_r = 0 \quad \alpha_r^2 = \beta^2 = 1.
\] (5.1.39)

Thus, we can use a 2 × 2 matrix representation for these two operators. A convenient one is

\[
\alpha_r = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\] (5.1.40)

To solve the Dirac equation for a spherically symmetric electrostatic potential we use this representation from now on.

As we have already seen, the eigenvalues of \(K\) are \(\pm(j + 1/2)\). We now try to specify these more precisely. First, we recall that since

\[ K = \beta(\vec{\Sigma} \cdot \vec{L} + 1) \] (5.1.41)

and

\[ \vec{J} = \vec{L} + \frac{1}{2} \vec{\Sigma} \] (5.1.42)
it follows that the eigenvalues of $\beta K$ are $j(j + 1) - l(l + 1) - 3/4 + 1$. Now consider the state $\Psi^M_{\omega j}$ and the action of $L^2$ on the upper component $\mathcal{Y}_{lj}^M$.

\[
L^2 \mathcal{Y}_{lj}^M = l(l + 1) \mathcal{Y}_{lj}^M \quad l = j + \frac{\omega}{2}
\]  

(5.1.43)

where, as always, $\omega = \pm 1$.

On the lower component $\mathcal{Y}_{lj'}^M$ we have

\[
L^2 \mathcal{Y}_{lj'}^M = l'(l' + 1) \mathcal{Y}_{lj'}^M \quad l' = j - \frac{\omega}{2}
\]

(5.1.44)

This means that acting on $\Psi^M_{\omega j}$ the operator $L^2$ is equivalent to

\[
\left( j + \beta \frac{\omega}{2} \right) \left( j + \beta \frac{\omega}{2} + 1 \right) = j(j + 1) + \frac{1}{4} + \frac{\omega}{2} (2j + 1) \beta
\]

(5.1.45)

since

\[
\left( j \pm \frac{\omega}{2} \right) \left( j \pm \frac{\omega}{2} + 1 \right) = j(j + 1) + \frac{1}{4} \pm \frac{\omega}{2} (2j + 1).
\]

(5.1.46)

But,

\[
1 + \vec{\Sigma} \cdot \vec{L} = J^2 - L^2 + \frac{1}{4}
\]

(5.1.47)

Therefore, acting on $\Psi^M_{\omega j}$

\[
1 + \vec{\Sigma} \cdot \vec{L} = j(j + 1) + \frac{1}{4} - [j(j + 1) + \frac{\omega}{2} (2j + 1) \beta]
\]

\[
= - \frac{1}{2} \omega (2j + 1) \beta
\]

(5.1.48)

so that

\[
K = \beta (1 + \vec{\Sigma} \cdot \vec{L}) = - \frac{1}{2} \omega (2j + 1) = - \omega (j + 1/2).
\]

(5.1.49)

Hence, with $\hbar = c = 1$ we can write

\[
H = \alpha_r \left[ p_r - \frac{i \omega (j + 1/2)}{r} \beta \right] + \beta m + V(r).
\]

(5.1.50)

and with

\[
\Psi^M_{\omega j} = \frac{1}{r} \begin{pmatrix} F(r) \mathcal{Y}_{lj}^M \\ G(r) \mathcal{Y}_{lj}^M \end{pmatrix}
\]

(5.1.51)

as well as, with our earlier choice for $\alpha_r$ and $\beta$, the energy eigenvalue equation becomes

\[
\begin{pmatrix}
\frac{d}{dr} & -\frac{d}{dr} \\
\frac{\omega (j + 1/2)}{r} & 0
\end{pmatrix}
+ \begin{pmatrix}
0 & \frac{\omega (j + 1/2)}{r} \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
F \\
G
\end{pmatrix}
= \begin{pmatrix}
E - m - V & 0 \\
0 & E + m - V
\end{pmatrix}
\begin{pmatrix}
F \\
G
\end{pmatrix}.
\]

(5.1.52)
Writing this out we get
\[
\left[-\frac{d}{dr} + \frac{\tilde{\omega}(j + 1/2)}{r}\right] G = (E - m - V) F \\
\left[-\frac{d}{dr} + \frac{\tilde{\omega}(j + 1/2)}{r}\right] F = (E + m - V) G.
\] (5.1.53)

The correct normalization for this requires
\[
\int_0^\infty (|F|^2 + |G|^2) \, dr = 1.
\] (5.1.54)

### 5.2 The Hydrogenic Atom

In this case we have
\[
V(r) = -\frac{Ze^2}{r}.
\] (5.2.55)

Asymptotically, for large \( r \) the solution must behave like
\[ e^{-\sqrt{m^2 - E^2} r}. \]

So, we set
\[
\kappa = \sqrt{m^2 - E^2} \quad \text{and} \quad \rho = \kappa r
\] (5.2.56)
as well as
\[
\tau = \tilde{\omega}(j + 1/2) \quad \text{and} \quad \nu = \sqrt{\frac{m - E}{m + E}}.
\] (5.2.57)

Then, the eigenvalue equations 5.1.53 become
\[
\left[-\frac{d}{d\rho} + \frac{\tau}{\rho}\right] G = \left(-\nu + \frac{Ze^2}{\rho}\right) F \\
\left[-\frac{d}{d\rho} + \frac{\tau}{\rho}\right] F = \left(\frac{1}{\nu} + \frac{Ze^2}{\rho}\right) G.
\] (5.2.58)

Here we clearly see that as \( \rho \to \infty \) both \( F \) and \( G \) behave as \( e^{-\rho} \). Therefore, we extract this asymptotic behaviour and set
\[
F(\rho) = e^{-\rho} f(\rho) \\
G(\rho) = e^{-\rho} g(\rho).
\] (5.2.59)

The pair of equations now becomes
\[
\rho \frac{df}{d\rho} = (\rho - \tau)f + (\rho/\nu + Ze^2)g \\
\rho \frac{dg}{d\rho} = (\rho + \tau)g + (\rho\nu - Ze^2)f
\] (5.2.60)
To solve this pair of equations we first set
\[ w = \begin{pmatrix} f \\ g \end{pmatrix} \]  
(5.2.61)

Then, with
\[ A = \begin{pmatrix} -\tau & Ze^2 \\ -Ze^2 & \tau \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1/\nu \\ \nu & 1 \end{pmatrix} \]  
(5.2.62)
we get
\[ \rho \frac{dw}{d\rho} = (A + \rho B)w \]  
(5.2.63)

We can diagonalize one of the two matrices \( A, B \). The eigenvalues of \( A \) are
\[ \pm \lambda = \pm \sqrt{\tau^2 - (Ze^2)^2} = \pm \sqrt{\left(j + \frac{1}{2}\right)^2 - \left(\frac{Ze^2}{hc}\right)^2} \]  
(5.2.64)

and the eigenvalues of \( B \) are 0, 2.

To solve the eigenvalue problem we use the Method of Frobenius and set
\[ w = \rho^\mu \sum_{s=0} \omega_s \rho^s = \sum_{s=0} \omega_s \rho^{s+\mu}. \]  
(5.2.65)
Here,
\[ \omega_s = \begin{pmatrix} f_s \\ g_s \end{pmatrix} \]  
(5.2.66)

Then,
\[ \rho \frac{dw}{d\rho} = \sum_{s=0} (s + \mu) \omega_s \rho^{s+\mu}, \]  
(5.2.67)
\[ Aw = \sum_{s=0} (Aw)_s \rho^{s+\mu}, \]  
(5.2.68)

and
\[ \rho Bw = \sum_{s=1} (Bw)_{s-1} \rho^{s+\mu}. \]  
(5.2.69)

Equating the coefficients of the different powers of \( \rho \) to zero we find.

\( s = 0 \) term
\[ \mu \omega_0 = (A\omega)_0. \]  
(5.2.70)

This means that \( \omega_0 \) is an eigenvector of \( A \) and \( \mu = \pm \lambda = \pm \sqrt{\tau^2 - (Ze^2)^2} \). Since the solution must be normalizable and the integral runs from 0 we must have \( \mu = +\lambda \). The corresponding eigenvector is given by
\[ \omega_0 = \begin{pmatrix} f_0 \\ g_0 \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} Ze^2 \\ \Delta^2 \end{pmatrix} \]  
(5.2.71)
where
\[ \Delta = 2\tau [\tau + \sqrt{\tau^2 - Z^2e^4}] . \]  
(5.2.72)

general s term
With \( \mu = +\lambda \) we have
\[ (s + \lambda - A)w_s = Bw_{s-1} \quad s \geq 1 . \]  
(5.2.73)

Now, for \( s > 0 \) we have
\[ \det(s + \lambda - A) \neq 0 \quad s > 0 . \]  
(5.2.74)

This means that \((s + \lambda - A)^{-1}\) exists and therefore we have
\[ w_s = \frac{1}{s} \left( 1 + \frac{\lambda - A}{s} \right)^{-1} Bw_{s-1} . \]  
(5.2.75)

Therefore,
\[ \| w_s \| \leq \frac{c}{s} \| w_{s-1} \| . \]  
(5.2.76)

So, the series converges. It is not difficult to check that the convergent series has asymptotic behaviour like \( e^{2p} \). Therefore, to have normalizable solutions the series must truncate. To impose this condition we take \( B \) in diagonal form.
\[ B' = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} = QBQ^{-1} . \]  
(5.2.77)

A possible choice for \( Q \) is
\[ Q = \begin{pmatrix} 1 & 1/\nu \\ 1 & -1/\nu \end{pmatrix} \quad Q^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ \nu & -\nu \end{pmatrix} . \]  
(5.2.78)

Then,
\[ w' = Qw =\begin{pmatrix} f' \\ g' \end{pmatrix} \]
\[ = \begin{pmatrix} f + \frac{1}{\nu} g \\ f - \frac{1}{\nu} g \end{pmatrix} \]
\[ w = Q^{-1}w' =\begin{pmatrix} f \\ g \end{pmatrix} \]
\[ = \frac{1}{2} \begin{pmatrix} f' + g' \\ \nu(f' - g') \end{pmatrix} \]  
(5.2.79)

and
\[ A' = QAQ^{-1} = \frac{1}{2} \begin{pmatrix} 1 & \frac{1}{\nu} \\ 1 & -\frac{1}{\nu} \end{pmatrix} \begin{pmatrix} -\tau & Z e^2 \\ -Z e^2 & \tau \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \nu & -\nu \end{pmatrix} \]
\[ = \frac{1}{2} \begin{pmatrix} (\nu - 1/\nu)Z e^2 & -2\tau - (\nu + 1/\nu)Z e^2 \\ -2\tau + (\nu + 1/\nu)Z e^2 & -(\nu - 1/\nu)Z e^2 \end{pmatrix} . \]  
(5.2.80)
This means that
\[
(s + \lambda - A') = \begin{pmatrix}
    s + \lambda - \frac{\nu^2 - 1}{2\nu} Ze^2 & \tau + \frac{\nu^2 + 1}{2\nu} Ze^2 \\
    -\tau + \frac{\nu^2 + 1}{2\nu} Ze^2 & s + \lambda - \frac{\nu^2 - 1}{2\nu} Ze^2
\end{pmatrix}.
\] (5.2.81)

and
\[
(s + \lambda - A')^{-1} = \frac{1}{s(s + 2\lambda)} \begin{pmatrix}
    s + \lambda + \frac{\nu^2 - 1}{2\nu} Ze^2 & -\tau - \frac{\nu^2 + 1}{2\nu} Ze^2 \\
    -\tau + \frac{\nu^2 + 1}{2\nu} Ze^2 & s + \lambda - \frac{\nu^2 - 1}{2\nu} Ze^2
\end{pmatrix}.
\] (5.2.82)

So, in general,
\[
w'_s = \begin{pmatrix}
f'_s \\
g'_s
\end{pmatrix} = \frac{2}{s(s + 2\lambda)} \begin{pmatrix}
    s + \lambda + \frac{\nu^2 - 1}{2\nu} Ze^2 \\
    -\tau + \frac{\nu^2 + 1}{2\nu} Ze^2
\end{pmatrix} f'_{s-1}.
\] (5.2.83)

Of course we also still have
\[
A' w'_0 = \lambda w'_0.
\] (5.2.84)

where
\[
w'_0 = \frac{1}{\Delta} \left( \frac{Ze^2 + \Delta^2}{Ze^2 - \Delta^2} \right).
\] (5.2.85)

So, \(w'\) is clearly determined as a power series. We now define an integer \(N\) such that \(w'_N \neq 0\) and \(w'_{N+1} = 0\). Then the power series becomes a polynomial of order \(N\). The necessary and sufficient conditions for this to occur are that
\[
f'_{N-1} \neq 0 \quad \text{and} \quad f'_N = 0.
\] (5.2.86)

This means that
\[
N + \lambda + \frac{\nu^2 - 1}{2\nu} Ze^2 = 0 \quad N \geq 1
\] (5.2.87)

or
\[
N + \lambda = \frac{1 - \nu^2}{2\nu} Ze^2 = \frac{2E}{m+E} \cdot \frac{1}{2} \sqrt{\frac{m+E}{m-E}} Ze^2
\]

where we have substituted the expression \(\sqrt{\frac{m-E}{m+E}}\) for \(\nu\).

Calling \(E/m = \epsilon\) we get
\[
N + \lambda = \frac{\epsilon}{\sqrt{1 - \epsilon^2}} Ze^2.
\] (5.2.88)

Solving for \(\epsilon^2\) we find
\[
\epsilon^2 = \frac{(N + \lambda)^2/(Ze^4)}{1 + (N + \lambda)^2/(Ze^4)} = \frac{1}{1 + \left(\frac{Ze^2}{N + \lambda}\right)^2}.
\] (5.2.89)
Now, introducing the fine structure constant $\alpha = e^2/(\hbar c)$ and substituting for $\lambda$ its value

$$\lambda = \sqrt{(j + \frac{1}{2})^2 - Z^2\alpha^2}$$

we get

$$\epsilon = \frac{E}{mc^2} = \frac{1}{\sqrt{1 + \frac{Z^2\alpha^2}{(N + \sqrt{(j + 1/2)^2 - Z^2\alpha^2})^2}}}.$$  \hspace{1cm} (5.2.90)

This very important result had already been obtained by Sommerfeld on the basis of the old quantum theory. It is also important to notice that

$$N + (j + 1/2) = 1, 2, 3, \ldots, n.$$  \hspace{1cm} (5.2.91)

Furthermore, the fine structure constant $\alpha < 1/137$ and therefore, in general,

$$Z\alpha << 1.$$  \hspace{1cm} (5.2.92)

Thus, we can expand the square root and obtain

$$\frac{E}{mc^2} \approx 1 - \frac{1}{2} \frac{Z^2\alpha^2}{(N + j + 1/2)^2} + \cdots$$

$$= 1 - \frac{1}{2} \frac{Z^2\alpha^2}{n^2} \left[ 1 + \frac{Z^2\alpha^2}{n^2} \left( \frac{n}{j + 1/2} - \frac{3}{4} \right) + \cdots \right]$$  \hspace{1cm} (5.2.93)

The first term after the 1 in the square brackets is called the fine structure. Also it is clear that to lowest order in $\alpha$ the nonrelativistic energy $E' = E - mc^2$ coincides with the Bohr term

$$E' = -\frac{1}{2} mc^2 \frac{Z^2\alpha^2}{n^2}.$$  \hspace{1cm} (5.2.94)

As a final note, it is worth mentioning that the polynomials that occur in this solution are again Laguerre polynomials, just like in the case of the Schrödinger equation.

### 5.3 Problems

#### 5.1 Prove the following identities

$$\vec{r} \times \vec{L} + \vec{L} \times \vec{r} = 2i\vec{r}$$

$$\vec{p} \times \vec{L} + \vec{L} \times \vec{p} = 2i\vec{p}$$  \hspace{1cm} (5.3.95)

#### 5.2 The matrices

$$\gamma'^\mu = (a - ib\gamma^5)\gamma^\mu \quad a^2 + b^2 = 1$$
obey the commutation rules

\[ [\gamma^\mu, \gamma^\nu]_+ = 2g^{\mu\nu} \]

Obtain the explicit representation of the similarity transformation \( S \) which relates \( \gamma'\mu \) and \( \gamma^\mu \).

\[ \gamma'\mu = S^{-1}\gamma^\mu S \]

5.3 a) Show that the solutions to the free Dirac equation for positive energies which are normalized to \( \bar{u}u = 1 \) may be written in the form

\[ u = \sqrt{\frac{E + M}{2M}} \left[ 1 - \frac{\gamma^0 \vec{\alpha} \cdot \vec{p}}{E + M} \right] u_0 \]

where

\[ u_0 = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \beta \\ 0 \end{pmatrix} \]

with \( \alpha \) and \( \beta \) the usual two-component Pauli spinors corresponding to spin up or spin down.

b) Obtain explicit values for the expectation value of the operators

\[ O_i = \gamma^\mu, \gamma^\mu \gamma^5, \gamma^5, \sigma^{\mu\nu} \]

in the states \( u(\vec{p}, s) \) and in the states \( v(\vec{p}, s) \). That is, compute

\[ \bar{u}(\vec{p}, s)O_i u(\vec{p}, s) \quad \text{and} \quad \bar{v}(\vec{p}, s)O_i v(\vec{p}, s) \]

Analyze your answers for the case that \( \vec{p} = (0, 0, p) \).

5.4 Obtain explicit representations for the matrix elements of the operators

\[ O_i = \gamma^\mu, \gamma^\mu \gamma^5, \gamma^5, \sigma^{\mu\nu} \]

between spinors \( u(\vec{p}, s) \) and \( u(\vec{q}, r) \). Analyze in detail the case that

\[ p = (E, 0, 0, p) \quad \text{and} \quad q = (E, 0, 0, -p) \]

That is, compute

\[ \bar{u}(\vec{p}, s)O_i u(\vec{q}, r) \]

5.5 a) Compute the following traces

\[ \text{Tr} \left( \gamma^\alpha \gamma^\beta \gamma^5 \right), \quad \text{Tr} \left( \gamma^\alpha \gamma^\beta \gamma^\rho \gamma^5 \right), \quad \text{Tr} \left( \gamma^\alpha \gamma^\beta \gamma^\delta \gamma^\rho \gamma^5 \right) \]

b) Show that

\[ \text{Tr}(\phi_1 \phi_2 \ldots \phi_n) = 4 \sum_p \delta_p (a_{p_1}a_{p_2})(a_{p_3}a_{p_4})\ldots(a_{p_{n-1}}a_{p_n}) \]

where the sum is taken over all different ways that one can dot the vectors into each other subject to the restriction \( p_1 < p_2, p_3 < p_4, \ldots \) and \( \delta_p = \pm 1 \) depending
on whether \((p_1, p_2, \ldots, p_n)\) is an even or odd permutation of \((1, 2, \ldots, n)\). Show that there are exactly \((2m)!/(2^m m!)\) terms.

5.6 The large and small components of the the solutions of the Dirac equation

\[
(\gamma \cdot p - m) \psi = V \psi
\]

are defined as

\[
\psi_\pm = \frac{1}{2} (1 \pm \gamma^0) \psi.
\]

Choose a representation in which \(\gamma^0 = \beta\) is diagonal.

a) Obtain the equation for \(\psi_+\) and show that when the kinetic energy \(T\) of the particle \((p^0 = m + T)\) is such that \(T \ll m\) then \(\psi_+\) reduces to the nonrelativistic Schrödinger wavefunction. Note however, that the “hamiltonian” in the exact equation for \(\psi_+\) is not necessarily hermitean. Why?

b) Obtain the nonrelativistic limit of the equation satisfied by \(\psi_+\) in the case that

(i) \(V = \gamma_\mu A^\mu\)  
(ii) \(V = \gamma_5 U\)  
(iii) \(V = i \gamma_\mu \gamma_5 U^\mu\)

(iv) \(V = \sigma_{\mu\nu} F^{\mu\nu}\)  
(v) \(V = \gamma_5 \sigma_{\mu\nu} G^{\mu\nu}\)  
(vi) \(V = i \gamma_\mu \gamma_5 \partial^\mu U\).

5.7 Solve, as far as you can, for the energy eigenvalues of a Dirac electron in the spherically symmetric electrostatic potential

\[
V(r) = \begin{cases} 
-V_0 & 0 < r < R \\
0 & r > R
\end{cases}
\]

5.8 Find a term to describe an electrostatic dipole interaction for a Dirac particle and discuss the invariance properties of such an interaction under Lorentz transformations, parity, and time reversal.
Chapter 6

The Weyl or Neutrino Equation

6.1 Derivation

In the Dirac equation, if we let the mass $m$ go to zero we get

$$i \gamma \cdot \partial \psi(x) = 0.$$  \hfill (6.1.1)

Now, if we use the representation

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0 & \bar{\sigma} \\ -\bar{\sigma} & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$  \hfill (6.1.2)

and write

$$\psi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$  \hfill (6.1.3)

we find that the equations for $\phi_1$ and $\phi_2$ decouple

$$i \partial_0 \phi_2 - \bar{\sigma} \cdot \nabla \phi_2 = 0$$
$$i \partial_0 \phi_1 + \bar{\sigma} \cdot \nabla \phi_1 = 0.$$  \hfill (6.1.4)

Furthermore, both $\phi_1$ and $\phi_2$ are eigenstates of the chirality operator $\gamma^5$ in the sense that

$$\gamma^5 \begin{pmatrix} \phi_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ 0 \end{pmatrix}, \quad \gamma^5 \begin{pmatrix} 0 \\ \phi_2 \end{pmatrix} = -\begin{pmatrix} 0 \\ \phi_2 \end{pmatrix}.$$  \hfill (6.1.5)

This means we can write two mass zero equations

$$i \hbar \partial_t \phi = c_i \hbar \bar{\sigma} \cdot \nabla \phi$$
$$i \hbar \partial_t \phi = -c_i \hbar \bar{\sigma} \cdot \nabla \phi$$  \hfill (6.1.6)
or
\[ i\hbar \partial_t \phi = \pm c \vec{\sigma} \cdot \vec{p} \phi \]  
(6.1.7)
if we put
\[ \sigma_0 = 1 \]  
(6.1.8)
then these equations (with \( h = c = 1 \)) are of the form
\[ \sigma^\mu p_\mu \phi = 0 . \]  
(6.1.9)
This equation is known as the Weyl equation for the neutrino. It is form invariant under Lorentz transformations if we can find a matrix \( S(\Lambda) \) such that
\[ S(\Lambda)^{-1} \sigma^\mu S(\Lambda) = \Lambda_\nu^\mu \sigma^\nu . \]  
(6.1.10)
The existence of such an \( S(\Lambda) \) is again demonstrated as in the case of the Dirac equation by using infinitesimal transformations. We leave the construction of this as an exercise. See problem (6.1).

### 6.2 Solutions

If we look for plane wave solutions of the form
\[ \psi(x) = e^{ip \cdot x} u(p) \]  
(6.2.11)
we find that the upper component (positive chirality) satisfies
\[ p^0 u(p) = \vec{\sigma} \cdot \vec{p} u(p) \]  
(6.2.12)
so that
\[ \left( p^0 - p^2 \right) u(p) = 0 , \quad p^0 = \pm |\vec{p}| \]  
(6.2.13)
The lower component (negative chirality) satisfies
\[ p^0 u(p) = -\vec{\sigma} \cdot \vec{p} u(p) . \]  
(6.2.14)
For this equation if \( p^0 = |\vec{p}| \) we call the solution a neutrino state. On the other hand if \( p^0 = -|\vec{p}| \) we call the solution an antineutrino state.

If we write
\[ \phi_1 = \begin{pmatrix} \phi_1 \\ 0 \end{pmatrix} \quad \phi_2 = \begin{pmatrix} 0 \\ \phi_2 \end{pmatrix} \]  
(6.2.15)
then with
\[ \phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \]  
(6.2.16)
we have that
\[ \phi_1 = \frac{1}{2} (1 + \gamma^5) \phi \quad \phi_2 = \frac{1}{2} (1 - \gamma^5) \phi \]  
(6.2.17)
The operators $\frac{1}{2}(1 \pm \gamma^5)$, as we have already seen, are called the *chirality projection operators*. They project onto left-handed and right-handed states respectively. The spin is then either parallel or anti-parallel to the momentum.

Experimentally it has been found that the neutrino is left-handed and the anti-neutrino is right-handed. See figure 6.1

![Neutrino and Antineutrino States](image)

**Figure 6.1**: Neutrino and Antineutrino states. In both cases $p^0 = |\vec{p}|$.

### 6.3 Problems

6.1 Obtain the representation of the Lorentz group $S(A)$ that is required to maintain the form invariance of the Weyl equations.  
**Hint**: Use the infinitesimal transformations.
Chapter 7

The Neutral Klein-Gordon Field

7.1 Introduction

In this section we develop canonical quantum field theory for the neutral Klein-Gordon field. For free fields the canonical formalism is simply an elegant way of writing states of bosons and fermions with the correct energy-momentum relation, spin and statistics. The power of the theory comes into play when interactions are considered.

Most of the advances in the theory of high energy particle theory was due to the development of quantum field theory and the canonical formalism in the 1930's. In spite of numerous proposed rival schemes, canonical quantum field theory remains the most successful and robust theory to date and with the use of perturbation theory à la Feynman has led to predictions experimentally verified to better than one part in $10^{10}$.

7.2 Classical Real Scalar Field

We begin with the classical, real, scalar field, the Klein-Gordon field which satisfies

$$ (\Box + m^2) \phi = 0 . \quad (7.2.1) $$

The lagrange density for this field is given by the scalar function

$$ \mathcal{L} = \frac{1}{2} (\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) . \quad (7.2.2) $$

The corresponding lagrangian is

$$ L = \int d^3 x \mathcal{L}(\phi, \partial^\mu \phi) \quad (7.2.3) $$
and the action is
\[ I = \int_{t_1}^{t_2} dx^0 L = \int_{t_1}^{t_2} d^4x \mathcal{L}(\phi, \partial^\mu \phi) \]. \tag{7.2.4}

If we vary the action in a manner such that the field at the end points remains fixed, that is such that the variations of the field at the end points vanishes
\[ \delta \phi(t_1, \vec{x}) = \delta \phi(t_2, \vec{x}) = 0 \] \tag{7.2.5}
we get
\[ \delta I = \int_{t_1}^{t_2} d^4x \left[ \frac{\delta \mathcal{L}}{\delta \phi} \delta \phi + \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \delta (\partial_\mu \phi) \right] \]. \tag{7.2.6}

Now, using that
\[ \delta (\partial_\mu \phi) = \partial_\mu (\delta \phi) \] \tag{7.2.7}
we can rewrite the last term in 7.2.6 and integrate by parts to get
\[ \delta I = \int_{t_1}^{t_2} d^4x \left[ \frac{\delta \mathcal{L}}{\delta \phi} \delta \phi - \partial_\mu \left( \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \right) \delta (\partial_\mu \phi) \right] + \int_{t_1}^{t_2} d^4x \partial_\mu \left( \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \delta \phi \right) \]. \tag{7.2.8}

The last term is an integral of a pure divergence and may be converted to a surface integral
\[ \int_{t_1}^{t_2} d^4x \partial_\mu \left( \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \delta \phi \right) = \int d\sigma(x) \tilde{n}_\mu \left( \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \delta \phi \right) \]. \tag{7.2.9}
where the spacelike surfaces over which this integral is to be evaluated are at \( t = t_1 \) and \( t = t_2 \). But at these points the variations \( \delta \phi \) have been chosen to vanish. Furthermore the fields are chosen to vanish as \( |\vec{x}| \to \infty \). Thus, the surface integral vanishes and we get
\[ \delta I = \int_{t_1}^{t_2} d^4x \left[ \frac{\delta \mathcal{L}}{\delta \phi} - \partial_\mu \left( \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \right) \right] \delta \phi \]. \tag{7.2.10}

If we now require that the action be an extremum so that its variations vanish, we find that, since the variations of the field \( \delta \phi \) are arbitrary, we must have that the field satisfies the Euler-Lagrange equation
\[ \frac{\delta \mathcal{L}}{\delta \phi} - \partial_\mu \left( \frac{\delta \mathcal{L}}{\delta (\partial_\mu \phi)} \right) = 0 \]. \tag{7.2.11}

We have now established the necessity of the Euler-Lagrange equation for the action to be an extremum. That this equation is also sufficient is obvious. If we write this equation out for the particular case of the Klein-Gordon lagrangian we simply recover the Klein-Gordon equation
\[ (\Box + m^2) \phi = 0 \]. \tag{7.2.12}
7.3 Particle Interpretation

If we take any complete set of real, orthonormal functions \( \{ u_n(\vec{x}) \} \) so that

\[
\int d^3x \; u_m(\vec{x}) u_n(\vec{x}) = \delta_{mn}
\]  

(7.3.13)

and

\[
\sum_n u_n(\vec{x}) u_n(\vec{y}) = \delta(\vec{x} - \vec{y})
\]  

(7.3.14)

then we can expand any solution of the Klein-Gordon equation in terms of these functions

\[
\phi(x) = \sum_n q_n(t) u_n(\vec{x})
\]  

(7.3.15)

The coefficients \( q_n \) are determined by

\[
q_n(t) = \int d^3x \; u_n(\vec{x}) \phi(x)
\]  

(7.3.16)

Also,

\[
\dot{\phi}(x) = \partial_0 \phi(x) = \sum_n \dot{q}_n(t) u_n(\vec{x})
\]  

(7.3.17)

and

\[
\dot{q}_n(t) = \int d^3x \; u_n(\vec{x}) \dot{\phi}(x)
\]  

(7.3.18)

We now again consider the lagrangian. Thus, we find

\[
\frac{\partial L}{\partial q_n} = \int d^3x \left[ \frac{\partial L}{\partial \phi} \frac{\partial \phi}{\partial q_n} + \sum_{k=1}^3 \frac{\partial L}{\partial (\partial_k \phi)} \frac{\partial (\partial_k \phi)}{\partial q_n} \right]
\]  

(7.3.19)

Now,

\[
\frac{\partial (\partial_k \phi)}{\partial q_n} = \frac{\partial}{\partial x^k} \left( \frac{\partial \phi}{\partial q_n} \right)
\]  

(7.3.20)

and

\[
\frac{\partial \phi}{\partial q_n} = u_n(\vec{x})
\]  

(7.3.21)

Again, we can integrate the last term by parts and convert the integrated term to a surface integral which vanishes to get

\[
\frac{\partial L}{\partial q_n} = \int d^3x \left[ \frac{\partial L}{\partial \phi} - \sum_{k=1}^3 \frac{\partial}{\partial x^k} \left( \frac{\partial L}{\partial (\partial_k \phi)} \right) \right] u_n(\vec{x})
\]  

(7.3.22)
Similarly, we find
\[
\frac{\partial L}{\partial q_n} = \int d^3 x \frac{\partial L}{\partial \phi} \frac{\partial \phi}{\partial q_n} = \int d^3 x \frac{\partial L}{\partial \phi} u_n(\vec{x}) .
\] (7.3.23)
Combining these results we get
\[
\frac{\partial L}{\partial q_n} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_n} = \int d^3 x \left[ \frac{\delta L}{\delta \phi} - \frac{\partial}{\partial x^\mu} \left( \frac{\delta L}{\delta (\partial_\mu \phi)} \right) \right] u_n(\vec{x}) .
\] (7.3.24)
If \( \phi \) satisfies the Euler-Lagrange equation we then find that the expression above vanishes and we get particle equations
\[
\frac{\partial L}{\partial q_n} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_n} = 0 .
\] (7.3.25)
The converse is clearly also true. This means that for any complete set of orthonormal functions we are able to get a particle interpretation. The price to be paid for this is that we have replaced the field \( \phi \) by an infinite set of particles.

The canonical momentum \( p_n \) conjugate to \( q_n \) is given by
\[
p_n(t) = \frac{\partial L}{\partial q_n} = \int d^3 x \frac{\partial L}{\partial \phi} u_n(\vec{x}) .
\] (7.3.26)

### 7.4 Quantization of the Field

Now, that we have classical particle equations, we can quantize in the usual manner by imposing the canonical commutation relations
\[
[q_n(t), p_m(t)] = i\hbar \delta_{n,m} .
\] (7.4.27)
For clarity we have included the factor \( \hbar \), but we immediately again set \( \hbar = 1 \). We also have that
\[
q_n(t) = \int d^3 x \ u_n(\vec{x}) \phi(x)
\] (7.4.28)
and
\[
p_m(t) = \int d^3 y \ \frac{\partial L}{\partial \phi} u_m(\vec{y}) = \int d^3 y \ \pi(y) u_m(\vec{y})
\] (7.4.29)
where we have defined the conjugate momentum density operator
\[
\pi(y) = \frac{\partial L}{\partial \phi(y)} .
\] (7.4.30)
For the Klein-Gordon field this is simply
\[
\pi(y) = \dot{\phi}(y) .
\] (7.4.31)
Replacing these expressions into the commutator 7.4.27 we get
\[
\int d^3 x d^3 y \ u_n(\vec{x}) u_m(\vec{y}) [\phi(x^0, \vec{x}), \pi(y^0, \vec{y})] = i\delta_{n,m} .
\] (7.4.32)
We now multiply this equation by \( u_n(\vec{x})u_m(\vec{y}) \), sum over \( n \) and \( m \), and use the completeness relation for the functions \( u_n(\vec{x}) \), \( u_m(\vec{y}) \) to get

\[
\sum_{n,m} \int d^3x d^3y \ u_n(\vec{x})u_m(\vec{y})u_n(\vec{x}')u_m(\vec{y})[\phi(x^0, \vec{x}), \pi(x^0, \vec{y})] \\
= i \sum_{n,m} u_n(\vec{x}')u_m(\vec{y}) \delta_{n,m} \\
= i\delta(\vec{x}' - \vec{y}) .
\]  

(7.4.33)

But the left hand side of this equation may be written

\[
LHS = \int d^3x d^3y \ \delta(\vec{x} - \vec{x}')\delta(\vec{y} - \vec{y})[\phi(x^0, \vec{x}), \pi(x^0, \vec{y})] \\
= [\phi(x^0, \vec{x}), \pi(x^0, \vec{y})] .
\]  

(7.4.34)

Thus, for the field operators, we have found commutation rules that are independent of the basis set of functions that we started with

\[
[\phi(x^0, \vec{x}), \pi(y^0, \vec{y})] = i\delta(\vec{x} - \vec{y}) .
\]  

(7.4.35)

Furthermore we find by direct computation and using that

\[
[q_n(t), p_m(t)] = [p_n(t), m_m(t)] = 0
\]  

(7.4.36)

that

\[
[\phi(x^0, \vec{x}), \phi(y^0, \vec{y})] = [\pi(x^0, \vec{x}), \pi(y^0, \vec{y})] = 0 .
\]  

(7.4.37)

The equal time commutators given by 7.4.35 and 7.4.37 are called canonical quantization. To employ canonical quantization it is clearly necessary that we have a lagrange density \( \mathcal{L} \) in which the field \( \phi \) and its time-derivative \( \dot{\phi} \) appear. Furthermore, these equal time commutators may be viewed as the initial conditions for the field \( \phi \) satisfying the Klein-Gordon equation 7.2.1.

### 7.5 Annihilation and Creation Operators

If we again start with the particle operators \( q_n(t), p_m(t) \) satifying

\[
[q_n(t), p_m(t)] = i\delta_{n,m}
\]  

(7.5.38)

and proceed as for the simple harmonic oscillator and define

\[
a_n(t) = \frac{1}{\sqrt{2}} (q_n(t) + ip_n(t)) \\
a^\dagger_n(t) = \frac{1}{\sqrt{2}} (q_n(t) - ip_n(t))
\]  

(7.5.39)
with the inverse relations

\[ p_n(t) = \frac{i}{\sqrt{2}} (a_n^\dagger(t) - a_n(t)) \]
\[ q_n(t) = \frac{1}{\sqrt{2}} (a_n^\dagger(t) + a_n(t)) \]  

(7.5.40)

then we find that

\[ [a_n(t), a_m^\dagger(t)] = \frac{1}{2} [q_n(t) + ip_n(t), q_m(t) - ip_m(t)] \]
\[ = -\frac{i}{2} [q_n(t), p_m(t)] + \frac{i}{2} [p_n(t), q_m(t)] \]
\[ = \delta_{n,m} \]  

(7.5.41)

Also, as a simple calculation shows,

\[ [a_n(t), a_m(t)] = [a_n^\dagger(t), a_m^\dagger(t)] = 0 \]  

(7.5.42)

So, these are just the same old creation and annihilation operators that we have come to know and love. These operators are still time dependent and therefore we recognize that we are working in the Heisenberg picture.

It is now straightforward to define a hilbert space (called Fock space) on which these operators are represented. One simply defines a state of zero quanta, the vacuum state \(|0\rangle\) by

\[ a_n(t)|0\rangle = 0 \]  

(7.5.43)

and proceeds from there by defining states of one particle, two particles, \cdots etc. by applying the creation operators. Before doing so explicitly we first examine the meaning of these annihilation and creation operators more closely.

### 7.6 The Hamiltonian

If we introduce the hamiltonian density \( \mathcal{H} \) by performing a Legendre transformation

\[ \mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{\pi}} \pi - \mathcal{L} = \dot{\phi} \pi - \mathcal{L} \]  

(7.6.44)

then the hamiltonian may be written

\[ H = \int d^3x \mathcal{H}(x) \]  

(7.6.45)

It is worth noticing that for the Klein-Gordon equation

\[ \dot{\phi}(x) = \pi(x) \]  

(7.6.46)
Using this fact we have for the Klein-Gordon field the Hamiltonian

\[
H = \int d^3x \left[ \pi^2(x) + (\nabla \phi)^2 + m^2 \phi^2 \right] \\
= \frac{1}{2} \int d^3x \sum_{n,m} \left[ p_n(t) u_n(\vec{x}) p_m(t) u_m(\vec{x}) + q_n(t) q_m(t) \nabla u_n(\vec{x}) \nabla u_m(\vec{x}) \right] \\
+ m^2 q_n(t) u_n(\vec{x}) q_m(t) u_m(\vec{x}) \right] \\
= \frac{1}{2} \sum_{n,m} \left[ p_n(t) p_m(t) + m^2 q_n(t) q_m(t) \right] \delta_{nm} \\
- \left[ q_n(t) q_m(t) \int d^3x \ u_n(\vec{x}) \nabla^2 u_m(\vec{x}) \right] \\
= \frac{1}{2} \sum_n \left[ p_n^2(t) + m^2 q_n^2(t) \right] + \frac{1}{2} \sum_{n,m} v_{nm} q_n(t) q_m(t) \\
(7.6.47)
\]

where we have introduced the "interaction term"

\[
v_{nm} = -u_n(\vec{x}) \nabla^2 u_m(\vec{x}) \quad . \qquad (7.6.48)
\]

In general, this is the best we can do for a general set of complete orthonormal functions. But, this is not very illuminating. We have no idea what the quanta or particles corresponding to these generalized coordinates and momenta are. However, this can be remedied by using the classical solutions of the field equation. These solutions give us an immediate physical interpretation and also diagonalize the Hamiltonian \( H \).

To see how the classical solutions play a role we consider the classical field equation

\[
(\Box + m^2) f(x) = 0 \quad (7.6.49)
\]

and for a plane wave solution \( e^{-i k \cdot x} \) we find

\[
-k^2 + m^2 = 0 \quad \text{so} \quad |k^0| = \omega(\vec{k}) = \sqrt{\vec{k}^2 + m^2} . \quad (7.6.50)
\]

This means that the general, classical real solution may be written

\[
f(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int_{k^0 > 0} \frac{d^3k}{k^0} \left[ a(\vec{k}) e^{-i k \cdot x} + a^*(\vec{k}) e^{i k \cdot x} \right] . \quad (7.6.51)
\]

We also immediately find

\[
f'(x) = \frac{-i}{\sqrt{2(2\pi)^3}} \int_{k^0 > 0} d^3k \left[ a(\vec{k}) e^{-i k \cdot x} - a^*(\vec{k}) e^{i k \cdot x} \right] . \quad (7.6.52)
\]

To invert these equations to find \( a(\vec{k}) \) and \( a^*(\vec{k}) \) we consider

\[
\int_{k^0 > 0} d^3x f(x) \ \delta_0 \ e^{i p \cdot x}
\]
\[\begin{align*}
&= \frac{i}{\sqrt{2(2\pi)^3}} \int_{k^0 > 0} d^3x \frac{d^3k}{k^0} \left\{ [a(\vec{k}) e^{-ikx} + a^*(\vec{k}) e^{ikx}] k^0 e^{ipx} \right. \\
&\left. + \left[ a(\vec{k}) e^{-ikx} \right. - \left. a^*(\vec{k}) e^{ikx} \right] k^0 e^{ipx} \right\} \\
&= \frac{i}{\sqrt{2(2\pi)^3}} \int_{k^0 > 0} d^3x \frac{d^3k}{k^0} \left\{ (p^0 + k^0) e^{-i(k^0 - p^0)x^0 + i(k - p)\vec{x}} a(\vec{k}) \right.
\\
&\left. + (p^0 - k^0) e^{i(k^0 + p^0)x^0 - i(k + p)\vec{x}} a^*(\vec{k}) \right\} \\
&= \frac{i}{\sqrt{2(2\pi)^3}} \int_{k^0 > 0} \frac{d^3k}{k^0} \left\{ (p^0 + k^0) e^{-i(k^0 - p^0)x^0} (2\pi)^3 \delta(\vec{k} - \vec{p}) a(\vec{k}) \right.
\\
&\left. + (p^0 - k^0) e^{i(k^0 + p^0)x^0} (2\pi)^3 \delta(\vec{k} + \vec{p}) a^*(\vec{k}) \right\}
\\
&= i\sqrt{2(2\pi)^3} a(\vec{p}) . \quad (7.6.53)
\end{align*}\]

Therefore, we have

\[a(\vec{p}) = -i \int d^3x f(x) \frac{e^{ipx}}{\sqrt{2(2\pi)^3}} \]

\[a^*(\vec{p}) = i \int d^3x f(x) \frac{e^{-ipx}}{\sqrt{2(2\pi)^3}} \quad (7.6.54)\]

where we have used \( f(x) = f^*(x) \), the fact that the field is real.

Having obtained the classical solution we now write the quantized field in the same manner by replacing the function \( f(x) \) by the field \( \phi(x) \) and the functions \( a(\vec{p}), a^*(\vec{p}) \) by operators \( a(\vec{p}), a^1(\vec{p}) \). To find the algebra of these operators we use canonical quantization.

\[\left[ \phi(x), \pi(y) \right]_{x^a = y^a} = i\delta(\vec{x} - \vec{y}) . \quad (7.6.55)\]

Then,

\[\begin{align*}
[a(\vec{p}), a^1(\vec{q})] &= \int_{x^a = y^a} d^3x d^3y \left[ \phi(x) \frac{e^{ipx}}{\sqrt{2(2\pi)^3}}, \phi(y) \frac{e^{-iqy}}{\sqrt{2(2\pi)^3}} \right] \\
&= \frac{1}{2(2\pi)^3} \int d^3x d^3y \left\{ [\phi(x), \phi(y)]_{x^a = y^a} p^0 q^0 + [\phi(x), \phi(y)]_{x^a = y^a} \right.
\\
&\left. + \left[ \phi(x), \phi(y) \right]_{x^a = y^a} iq^0 - ip^0[\phi(x), \phi(y)]_{x^a = y^a} \right\} e^{i(p - q)x} \\
&= \frac{1}{2(2\pi)^3} \int d^3x d^3y (p^0 + q^0) \delta(\vec{x} - \vec{y}) e^{i(p - q)x} . \quad (7.6.56)
\end{align*}\]

Thus, we find

\[\begin{align*}
[a(\vec{p}), a^1(\vec{q})] &= \frac{p^0 + q^0}{2(2\pi)^3} e^{-i(p^0 - q^0)x^0} (2\pi)^3 \delta(\vec{p} - \vec{q}) \\
&= q^0 \delta(\vec{p} - \vec{q}) . \quad (7.6.57)
\end{align*}\]

So, these are just annihilation and creation operators. Of course we still have to discover what it is that they annihilate and create. To answer this we reconsider the hamiltonian.
7.7 Normal Ordering

As we already found, the Hamiltonian density is given by

\[ H = \frac{1}{2} \left[ \pi^2 + (\nabla \phi)^2 + m^2 \phi^2 \right] \]  

(7.7.58)

The Hamiltonian is given by integrating this over a spacelike surface, say \( x^0 = \) constant. Now, using the expression for the field operator in terms of the creation and annihilation operators \( a(\vec{p}) \), \( a^{\dagger}(\vec{p}) \) we get

\[
H = \frac{1}{4(2\pi)^3} \int d^3x \int \frac{d^3k}{k^0} \frac{d^3q}{q^0} \left\{ -k^0q^0[a(\vec{k})e^{-ikx} - a^{\dagger}(\vec{k})e^{ikx}] \right\}
\]
\[
\times \left[ a(\vec{q})e^{-iqx} - a^{\dagger}(\vec{q})e^{iqx} \right]
\]
\[
+ (-\vec{k} \cdot \vec{q} + m^2)[a(\vec{k})e^{-ikx} + a^{\dagger}(\vec{k})e^{ikx}][a(\vec{q})e^{-iqx} + a^{\dagger}(\vec{q})e^{iqx}] \right\}
\]
\[
= \frac{(2\pi)^3}{4(2\pi)^3} \int \frac{d^3k}{k^0} \frac{d^3q}{q^0} \left\{ \left( -k^0q^0 - \vec{k} \cdot \vec{q} + m^2 \right)\delta(\vec{k} - \vec{q}) \right\}
\]
\[
\times \left[ e^{-i(k^0+q^0)x^0} a(\vec{k})a(\vec{q}) + e^{i(k^0+q^0)x^0} a^{\dagger}(\vec{k})a^{\dagger}(\vec{q}) \right]
\]
\[
+ (k^0q^0 + \vec{k} \cdot \vec{q} + m^2)\delta(\vec{k} - \vec{q}) \left[ a^{\dagger}(\vec{k})a(\vec{q}) + a(\vec{k})a^{\dagger}(\vec{q}) \right] \right\}
\]
\[
= \frac{1}{4} \int \frac{d^3k}{k^0k^0} \left\{ -k^{02} + k^2 + m^2 \right\}
\]
\[
\times \left[ e^{-2ik^0x^0} a(\vec{k})a(\vec{q}) + e^{2ik^0x^0} a^{\dagger}(\vec{k})a^{\dagger}(\vec{q}) \right]
\]
\[
+ \left[ (k^0)^2 + k^2 + m^2 \right] \left[ a^{\dagger}(\vec{k})a(\vec{q}) + a(\vec{k})a^{\dagger}(\vec{q}) \right] \right\}
\]
\[
= \frac{1}{2} \int \frac{d^3k}{k^0} \frac{d^3k}{k^0} \left[ a^{\dagger}(\vec{k})a(\vec{k}) + a(\vec{k})a^{\dagger}(\vec{k}) \right]. \]  

(7.7.59)

But, unfortunately

\[ a(\vec{k})a^{\dagger}(\vec{k}) = a^{\dagger}(\vec{k})a(\vec{k}) + \delta(0). \]  

(7.7.60)

This is the first infinity encountered in quantum field theory. It stems from the fact that we have an infinite number of degrees of freedom and a zero point energy of \( 1/2 k^0 \) associated with each of these degrees of freedom. This infinity is, however, only an artifact of our lack of properly ordering the field operators in defining the Hamiltonian density. After all, we want that the state of zero quanta, the vacuum state, should have zero energy.

\[ H|0\rangle = 0. \]  

(7.7.61)

Now, if this vacuum state is the Fock ground state, defined by

\[ a(\vec{k})|0\rangle = 0 \]  

(7.7.62)

then our goal is achieved by insisting that the Hamiltonian be written in such a manner that the annihilation operators appear to the right of the creation
operators. This assures that the vacuum carries zero energy. In this case we say that the hamiltonian is \textit{normal ordered} and we write it as: $H$. Thus, we have

\[
:H: = \frac{1}{2} \int \frac{d^3k}{k^0} k^0 2a^+(\vec{k})a(\vec{k}) = \int \frac{d^3k}{k^0} \omega(\vec{k})a^+(\vec{k})a(\vec{k}) .
\]  

(7.7.63)

This equation shows that the creation and annihilation operators $a(\vec{p})$, $a^+(\vec{p})$ annihilate and create quanta of energy

\[
E = \hbar \omega(\vec{k}) = \sqrt{c^2(h\vec{k})^2 + m^2c^4}
\]

(7.7.64)

as required.

From now on we always adhere to the following rule.

ALL OPERATORS MUST BE NORMAL ORDERED!

To get a further appreciation of the meaning of our field operators we now consider the energy-momentum tensor.

### 7.8 The Energy-Momentum Tensor

The energy-momentum tensor is given by

\[
T^{\mu\nu} = \partial^\mu \phi \frac{\partial L}{\partial (\partial_\nu \phi)} - g^{\mu\nu} L
\]

\[
= \partial^\mu \phi \partial^\nu \phi + \frac{1}{2} g^{\mu\nu} (m^2 \phi^2 - \partial_\lambda \phi \partial^\lambda \phi) .
\]

(7.8.65)

This tensor is symmetric

\[
T^{\mu\nu} = T^{\nu\mu}
\]

(7.8.66)

and conserved if the field $\phi$ satisfies the Euler-Lagrange equation since

\[
\partial_\nu T^{\mu\nu} = \partial^\mu \phi \frac{\partial}{\partial x^\nu} \left( \frac{\partial L}{\partial (\partial_\nu \phi)} \right) + (\partial_\nu \partial^\mu \phi) \frac{\partial L}{\partial (\partial_\nu \phi)} - \frac{\partial L}{\partial x_\mu}
\]

\[
= \partial^\mu \phi \frac{\partial}{\partial x^\nu} \left( \frac{\partial L}{\partial (\partial_\nu \phi)} \right) + (\partial_\nu \partial^\mu \phi) \frac{\partial L}{\partial (\partial_\nu \phi)} - \frac{\partial L}{\partial \phi} \frac{\partial^\mu \phi}{\partial \phi} - \frac{\partial L}{\partial (\partial_\nu \phi)} \partial_\nu \partial^\mu \phi
\]

\[
= \partial^\mu \phi \left[ \frac{\partial}{\partial x^\nu} \left( \frac{\partial L}{\partial (\partial_\nu \phi)} \right) - \frac{\partial L}{\partial \phi} \right] = 0 .
\]

(7.8.67)

So we have shown that

\[
\partial_\nu T^{\mu\nu} = 0 .
\]

(7.8.68)
Now consider
\[
\int d^3x \partial_\nu T^{\mu\nu} = \frac{d}{dx^0} \int d^3x T^{\mu0} + \int d^3x \frac{\partial}{\partial x^k} T^{\mu k} = 0.
\] (7.8.69)

The last integral can be converted into a surface integral and if we assume that the fields vanish at infinity we find that this integral also vanishes. Therefore, we get
\[
\frac{d}{dx^0} \int d^3x T^{\mu0} = 0.
\] (7.8.70)

This means that we can define the conserved quantity
\[
P^\mu = \int d^3x T^{\mu0},
\] (7.8.71)
and see that
\[
\frac{d}{dx^0} P^\mu = 0.
\] (7.8.72)

Examining this more closely we see that
\[
T^{00} = \int d^3x \left[ \phi \frac{\partial L}{\partial \phi} - L \right]
\]
\[
= \int d^3x \left[ \phi \pi - L \right]
\]
\[
= \int d^3x \mathcal{H}.
\] (7.8.73)

So, we have found that
\[
P^0 = H = \int_{\nu^+} \frac{d^3k}{k^0} \omega(\vec{k}) a^\dagger(\vec{k}) a(\vec{k})
\]
\[
= \text{the total hamiltonian}.
\] (7.8.74)

Also,
\[
P^k = \int d^3x (\partial^k \phi) \frac{\partial L}{\partial \phi} = \int d^3x (\partial^k \phi(x)) \pi(x)
\] (7.8.75)

We examine these three operators next. Of course, in keeping with our rules we have to normal order these operators. So,
\[
P^k = \int d^3x : (\partial^k \phi) \pi : (x).
\] (7.8.76)

Writing this out, we have
\begin{align}
\int d^3x \, (\partial^k \phi) \pi : (x) &= -\frac{1}{2(2\pi)^3} \int \frac{d^3p}{p^0} \frac{d^3q}{q^0} q^k p^0 \\
& \quad \left[ (a(q)a(p) e^{-i(q^0+p^0)x^0} + a^\dagger(q)a^\dagger(p) e^{i(q^0+p^0)x^0}) \delta(q + p) \right] \\
& - (a(q)a^\dagger(p) + a^\dagger(p)a(q)) \delta(q - p) \\
& = -\frac{1}{2} \int \frac{d^3q}{q^0} q^k \left[ a(q)a(-q) e^{-2iq^0x^0} + a^\dagger(q)a^\dagger(-q) e^{2iq^0x^0} \\
& - 2a^\dagger(q)a(q) \right].
\end{align}

(7.877)

Now, the terms
\[ a(q)a(-q) e^{-2iq^0x^0} + a^\dagger(q)a^\dagger(-q) e^{2iq^0x^0} \]
are even under the interchange of \( q \rightarrow -q \). However, \( q^k \) is odd. This means that when integrated over all values of \( q \) this integral vanishes. Therefore,
\begin{align}
P^k &= \int \frac{d^3q}{q^0} q^k a^\dagger(q)a(q) 
\end{align}
and
\begin{align}
P^\mu &= \int \frac{d^3q}{q^0} q^\mu a^\dagger(q)a(q).
\end{align}

(7.879)

To fully appreciate the meaning of these operators we consider a state
\begin{align}
|\vec{p}\rangle = a^\dagger(\vec{p})|0\rangle
\end{align}
(7.880)
and act on it with \( P^\mu \). Then,
\begin{align}
P^\mu|\vec{p}\rangle &= \int \frac{d^3k}{k^0} k^\mu a^\dagger(\vec{k})a(\vec{k})a^\dagger(\vec{p})|0\rangle \\
&= \int \frac{d^3k}{k^0} k^\mu a^\dagger(\vec{k}) \left[ a^\dagger(\vec{p})a(\vec{k}) + k^0 \delta(\vec{k} - \vec{p}) \right] |0\rangle \\
&= p^\mu a^\dagger(\vec{p})|0\rangle \\
&= p^\mu|\vec{p}\rangle
\end{align}
(7.881)
where
\begin{align}
p^\mu = (\omega(\vec{p}), \vec{p}) = (\sqrt{\vec{p}^2 + m^2}, \vec{p}).
\end{align}

(7.882)
so the state \( |\vec{p}\rangle = a^\dagger(\vec{p})|0\rangle \) is a one particle state with momentum \( \vec{p} \) and energy
\begin{align}
E = \omega(\vec{p}) = \sqrt{\vec{p}^2 + m^2}.
\end{align}
7.9 The Field Operators

In defining the Fock space vacuum we used the annihilation operator $a(p)$. We can also proceed directly from the field operator if we simply realize that the equation

$$a(p) = -i \int_{E^0=\text{constant}} \phi(x) \frac{\partial}{\partial \theta} \frac{e^{ipx}}{\sqrt{2(2\pi)^3}} d^3x$$

(7.9.83)

projects out the positive frequency part $\phi^+(x)$ of the field operator $\phi(x)$. So, if we make the Lorentz-invariant decomposition of $\phi(x)$ into positive and negative frequency parts

$$\phi(x) = \phi^+(x) + \phi^-(x)$$

(7.9.84)

where

$$\phi^+(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int_{k^0>0} d^3k \frac{k^0}{k^0} a(\vec{k}) e^{-ikx}$$

(7.9.85)

and

$$\phi^-(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int_{k^0>0} d^3k \frac{k^0}{k^0} a^+(\vec{k}) e^{ikx}$$

(7.9.86)

we can also define the vacuum state by

$$\phi^+(x|0\rangle = 0.$$  

(7.9.87)

So far we have only dealt with the equal-time commutators of the field operators. We are now ready to find their full algebra. To this end we define the full commutator $i\Delta(x-y)$ of two field operators $\phi(x)$ and $\phi(y)$ and evaluate this commutator.

$$[\phi(x), \phi(y)]$$

$$= \frac{1}{2(2\pi)^3} \int d^3k d^3q \frac{k^0}{q^0} \left[ a(\vec{k}) e^{-ikx} + a(\vec{q}) e^{i\bar{k}x} , a(\vec{q}) e^{i\bar{k}x} + a^+(\vec{q}) e^{i\bar{k}x} \right]$$

$$= \frac{1}{2(2\pi)^3} \int d^3k d^3q \frac{k^0}{q^0} \delta(\bar{k} - \bar{q}) \left[ e^{-i(kx-qy)} - e^{i(kx-qy)} \right]$$

$$= \frac{1}{2(2\pi)^3} \int d^3k \frac{k^0}{q^0} \left[ e^{-i(kx-y)} - e^{i(kx-y)} \right]$$

$$= i\Delta^+(x-y) + i\Delta^-(x-y)$$

$$= i\Delta(x-y).$$

(7.9.88)

This is the same function we encountered in section 2.6 where we discussed the normalization of the Klein-Gordon equation plane wave solutions. As we already saw there, $\Delta(x-y)$ vanishes whenever its argument $x-y$ is spacelike. This has direct physical implications and we now rederive this result by an explicit computation of $\Delta(x-y)$ in terms of known functions.
To explicitly evaluate the commutator we rewrite it as

\[
\Delta(x - y) = \frac{-1}{(2\pi)^3} \int \frac{d^3k}{k^0} e^{i\vec{k} \cdot (\vec{x} - \vec{y})} \sin k^0(x^0 - y^0). \tag{7.9.89}
\]

Next, we let

\[
r = |\vec{x} - \vec{y}| \quad \text{and} \quad |\vec{k}| = k. \tag{7.9.90}
\]

We can then carry out the two angular integrations and get

\[
\Delta(x - y) = \frac{-1}{2\pi^2} \int_0^\infty \frac{k^2 dk}{\sqrt{k^2 + m^2}} \frac{\sin kr \sin k^0(x^0 - y^0)}{kr} \left( \frac{1}{r} \frac{\partial}{\partial r} \right) F(r, x^0 - y^0). \tag{7.9.91}
\]

Here,

\[
F(r, x^0 - y^0) = \frac{1}{\pi} \int_0^\infty \frac{k^2 dk}{\sqrt{k^2 + m^2}} \cos kr \sin \sqrt{k^2 + m^2}(x^0 - y^0). \tag{7.9.92}
\]

If we now make a further change of variable to

\[
k = m \sinh z \tag{7.9.93}
\]

we find that this integral is just an integral representation for Bessel functions. So, we have

\[
F(r, x^0 - y^0) = \begin{cases} 
J_0(m \sqrt{(x^0 - y^0)^2 - (\vec{x} - \vec{y})^2}) & x^0 - y^0 > |\vec{x} - \vec{y}| \\
0 & |x^0 - y^0| < |\vec{x} - \vec{y}| \\
-J_0(m \sqrt{(x^0 - y^0)^2 - (\vec{x} - \vec{y})^2}) & x^0 - y^0 < -|\vec{x} - \vec{y}|
\end{cases}
\]

After carrying out the differentiation of this function we get

\[
\Delta(x) = -\frac{1}{2\pi} \epsilon(x^0) \left\{ \delta(x^2) - \frac{m^2}{2} \theta(x^2) \frac{J_1(m \sqrt{x^2})}{m \sqrt{x^2}} \right\}. \tag{7.9.94}
\]

From this expression we see explicitly the result that we proved earlier (section 2.6) that

\[
\Delta(x) = 0 \quad \text{if} \quad x^2 < 0.
\]

In other words, two fields commute if their arguments are spacelike separated. This means that two fields can only influence each other if they can be connected by a light signal.
7.9.1 Initial Value Problem

By using the fact that the field satisfies its field equation as well as the canonical equal time commutators it is possible to give an alternate definition of the commutator. Thus, writing

\[ i\Delta(x) = [\phi(x), \phi(0)] \]  \hspace{1cm} (7.9.95)

we immediately get

1. \hspace{1cm} (\Box + m^2)\Delta(x) = 0 \hspace{1cm} (7.9.96)

Furthermore, if we set \( x^0 = 0 \) we find

2. \hspace{1cm} \Delta(\vec{x}, 0) = 0 \hspace{1cm} (7.9.97)

and

3. \hspace{1cm} \left( \frac{\partial \Delta(x)}{\partial x^0} \right)_{x^0=0} = -\delta(\vec{x}) \hspace{1cm} (7.9.98)

since \( \partial_0 \phi(x) = \pi(x) \) and

\[ [\pi(x), \phi(0)]_{x^0=0} = -i\delta(\vec{x}) . \] \hspace{1cm} (7.9.99)

These three conditions define \( \Delta(x) \) uniquely. They also show that \( \Delta(x) \) can be used to solve initial value problems for the Klein-Gordon equation. To see this consider the following initial value problem.

Problem

Find a wave packet solution \( g(x) \) of the Klein-Gordon equation such that on the spacelike surface \( \sigma_0(x) \) the solution assumes the values \( g(x) = g_0(x) \) and the normal derivative assumes the values \( n_\mu \partial_\mu g(x) = f(x) \) where \( n_\mu(x) \) is the normal to the surface \( \sigma_0(x) \).

Solution

\[ g(x) = \int_{\sigma_0} \Delta(x-y) \overset{\uparrow}{n_\mu} g(y) \, d^3y \] \hspace{1cm} (7.9.100)

where \( \overset{\uparrow}{n_\mu} \) is a generalization of \( \partial_0 \). Writing this out we have

\[ g(x) = \int_{\sigma_0} \left[ \Delta(x-y)n_\mu(y)\partial^\mu g_0(y) - n_\mu(y)\partial^\mu \Delta(x-y)g_0(y) \right] \, d^3y . \] \hspace{1cm} (7.9.101)

To prove that this is indeed the required solution we use Gauss’ Theorem, namely

\[ \int_{\Omega} F_\mu(y) d\sigma^\mu(y) - \int_{\sigma_0} F_\mu(y) d\sigma^\mu(y) = \int_{\Omega} \overset{\uparrow}{\partial_0} F_\mu(y) \, d^4y \] \hspace{1cm} (7.9.102)

where \( \Omega \) is the spacetime volume enclosed between \( \sigma_0 \) and \( \sigma \). Here we have also tacitly assumed that

\[ \lim_{|y|\to\infty} F_\mu(y) = 0 . \] \hspace{1cm} (7.9.103)
Now, we choose
\[ F_\mu(y) = \Delta(x - y) \, \delta_\mu \, g(y) \]  
(7.9.104)
and recall that both \( \Delta \) and \( g \) satisfy the Klein-Gordon equation. Then, it is clear that
\[ \partial_\nu F_\mu(y) = 0. \]  
(7.9.105)
If we now choose the surface \( \sigma \) as the plane \( y^0 = x^0 = \text{constant} \) we have
\[ n_\mu(x) = (1, 0, 0, 0). \]  
(7.9.106)
Thus,
\[ \int_\sigma F_\mu(y) d\sigma^\mu(y) = \int_{x^0 = y^0} \Delta(x - y) \, \delta_0 \, g(y) \, dy \]
\[ = \int \Delta(\vec{x} - \vec{y}, 0) \partial_0 g(y) \, dy - \int \partial_0 \Delta(\vec{x} - \vec{y}, 0) g(y) \, dy \]
\[ = g(x). \]  
(7.9.107)
Here we have used that
\[ \Delta(\vec{x} - \vec{y}, 0) = 0 \quad \text{and} \quad \partial_0 \Delta(\vec{x} - \vec{y}, 0) = -\delta(\vec{x} - \vec{y}). \]  
(7.9.108)
Therefore,
\[ g(x) = \int_{\sigma_0} \Delta(x - y) \, \delta_0 \, \partial_\nu \, g_0(y) \, dy \]  
(7.9.109)
as stated.

7.10 Retarded and Advanced Green's Functions

We consider the following problem. Find a function \( G(x - y) \) such that
\[ (\Box + m^2)G(x - y) = \delta(x - y) \]  
(7.10.110)
subject to one of the following two conditions.
1.
\[ G(x - y) = 0 \quad \text{if} \quad x^0 < y^0. \]  
(7.10.111)
This is the retarded function \( \Delta_R(x - y) \).
2.
\[ G(x - y) = 0 \quad \text{if} \quad x^0 > y^0. \]  
(7.10.112)
This is the advanced function \( \Delta_A(x - y) \).
Below, we show that
\[ \Delta_R(x - y) = -\theta(x^0 - y^0) \Delta(x - y) \]  
(7.10.113)
and

\[ \Delta_A(x - y) = \theta(y^0 - x^0) \Delta(x - y). \] (7.10.114)

It is worth noting that since

\[ \theta(x^0 - y^0) + \theta(y^0 - x^0) = 1 \] (7.10.115)

we have that

\[ \Delta(x - y) = \Delta_A(x - y) - \Delta_R(x - y). \] (7.10.116)

This equation plays an important role for the quantized field.

We now derive the solutions quoted above. For this purpose we Fourier transform \( G(x - y) \) and write

\[ G(x - y) = -\frac{1}{(2\pi)^4} \int e^{-ik(x-y)} \tilde{G}(k) \, d^4k \] (7.10.117)

and recall that

\[ \delta(x - y) = \frac{1}{(2\pi)^4} \int e^{-ik(x-y)} \, d^4k. \] (7.10.118)

Therefore, the differential equation for \( G \) reduces to an algebraic equation for \( \tilde{G}(k) \).

\[ -(-k^2 + m^2) \tilde{G}(k) = 1 \] (7.10.119)

or

\[ \tilde{G}(k) = \frac{1}{k^2 - m^2}. \] (7.10.120)

Thus,

\[ G(x - y) = -\frac{1}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{k^2 - m^2} \, d^4k. \] (7.10.121)

If we now consider carrying out the integral over \( k^0 \), that is

\[ \int \frac{e^{-ik^0(x^0-y^0)}}{(k^0)^2 - \omega(\vec{k})^2} \, dk^0 \quad \omega(\vec{k})^2 = \vec{k}^2 + m^2 \] (7.10.122)

we see that the integrand has poles at \( \pm \omega(\vec{k}) \). So, we have a choice. We can either displace the contour of integration to miss the poles or we can shift the poles. In either case we can then close the contour with a large semi-circle. (See figure 7.1) This leads to two possibilities.

1. If \( x^0 - y^0 > 0 \) we can close the contour in the lower half plane. Now, in this case, we want the integral to vanish if we are evaluating \( \Delta_A(x - y) \), but not if we are evaluating \( \Delta_R(x - y) \). This is the situation depicted in 7.1. Here we have displaced the poles to lie just below the real axis. We could equally well have left the poles on the real axis and displaced the contour so that the horizontal line
lies just above the real axis. It has become conventional in physics to displace the poles rather than the axes.

2. If \( x^0 - y^0 < 0 \) we can close the contour in the upper half plane and in this case we want the integral to vanish if we are evaluating \( \Delta_R(x - y) \), but not if we are evaluating \( \Delta_A(x - y) \). In this case the two poles in 7.1 would have to lie above the real axis. After displacing the poles by \( \epsilon \) above or below the real axis we get

\[
\Delta_R(x - y) = -\frac{1}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{(k^0 + i\epsilon)^2 - \omega(k)^2} \, d^4k .
\] (7.10.123)

\[
\Delta_A(x - y) = -\frac{1}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{(k^0 - i\epsilon)^2 - \omega(k)^2} \, d^4k .
\] (7.10.124)

Carrying out first the integration over \( k^0 \) (see problem 7.1) now yields the results quoted above, namely

\[
\Delta_R(x - y) = -\theta(x^0 - y^0) \Delta(x - y)
\] (7.10.125)

and

\[
\Delta_A(x - y) = \theta(y^0 - x^0) \Delta(x - y)
\] (7.10.126)

as well as

\[
\Delta(x - y) = \Delta_A(x - y) - \Delta_R(x - y) .
\] (7.10.127)

### 7.11 \( c \)-number External Scalar Source

The simplest possible interaction for the quantized scalar field is with an external \( c \)-number source. The lagrange density in this case is

\[
\mathcal{L} = \frac{1}{2} \left( \partial^\mu \phi \partial_\mu \phi - m^2 \phi^2 \right) + g \rho \phi .
\] (7.11.128)
Here $\rho$ is a given $c$-number function which may be as smooth and well-behaved as we wish. It represents the external scalar source. The constant $g$ represents a coupling constant. The equation of motion for the Heisenberg field $\phi$ is

$$(\Box + m^2)\phi = g\rho(x) .$$  \hfill (7.11.129)

The field $\phi$ is also called the "interpolating field" for reasons that will soon become obvious. Also, the canonical momentum is

$$\pi(x) = \dot{\phi}(x) .$$  \hfill (7.11.130)

Using the Green's functions developed in the previous section we can write the solution as

$$\phi(x) = \phi_0(x) + g \int \Delta_{R,A}(x - y)\rho(y) \, d^4y .$$  \hfill (7.11.131)

Here, $\phi_0(x)$ is a free field solution

$$(\Box + m^2)\phi_0(x) = 0 .$$  \hfill (7.11.132)

So we have two separate ways of writing the solution depending on whether we use $\Delta_R$ or $\Delta_A$. We first consider the solution with $\Delta_R$.

Before proceeding we also define the solution in the remote past (before the source was turned on) by

$$\phi_{in}(x) = \lim_{x^0 \to -\infty} \phi(x) .$$  \hfill (7.11.133)

Now, for $x^0 \to -\infty$ we have that

$$\lim_{x^0 \to -\infty} \Delta_R(x - y) = 0$$  \hfill (7.11.134)

for all finite $y^0$. Therefore, since the source $\rho(y)$ vanishes both in the remote past and the remote future we have that

$$\phi_{in}(x) = \phi_0(x) \quad \text{a free field} .$$  \hfill (7.11.135)

Thus, we have the solution

$$\phi(x) = \phi_{in}(x) + g \int \Delta_R(x - y)\rho(y) \, d^4y .$$  \hfill (7.11.136)

Similarly, defining

$$\phi_{out}(x) = \lim_{x^0 \to +\infty} \phi(x) .$$  \hfill (7.11.137)

we find that

$$\phi(x) = \phi_{out}(x) + g \int \Delta_A(x - y)\rho(y) \, d^4y .$$  \hfill (7.11.138)

Here again, $\phi_{out}(x)$ is a free field.
Since \( \phi_{in}(x) \) is a free field we can expand it in terms of annihilation and creation operators

\[
\phi_{in}(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ a(\vec{k}) e^{-ikx} + a^{\dagger}(\vec{k}) e^{ikx} \right].
\] (7.11.139)

The operators \( a(\vec{k}) \) and \( a^{\dagger}(\vec{k}) \) annihilate and create "in" particles from the in-vacuum. That is, they annihilate and create particles of the type that, in the remote past, were sent into the beam tube to be scattered off the source \( \rho(x) \). The in-vacuum is the state with no incoming particles. Using the in-vacuum \( |0 : in\rangle \) and the creation operators \( a^{\dagger}(\vec{k}) \) we can build up the complete hilbert space for the incoming particles. This Fock space is called \( \mathcal{H}_{in} \). The in-vacuum is defined by

\[
a(\vec{k})|0 : in\rangle = 0.
\] (7.11.140)

The one-particle in-states are

\[
|\vec{k} : in\rangle = a^{\dagger}(\vec{k})|0 : in\rangle.
\] (7.11.141)

The two-particle in-states are

\[
|\vec{k}, \vec{q} : in\rangle = a^{\dagger}(\vec{k})a^{\dagger}(\vec{q})|0 : in\rangle
\] (7.11.142)

and so forth.

In a completely parallel manner we can expand the out-field in terms of annihilation and creation operators

\[
\phi_{out}(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ b(\vec{k}) e^{-ikx} + b^{\dagger}(\vec{k}) e^{ikx} \right]
\] (7.11.143)

and construct another Fock space \( \mathcal{H}_{out} \) in terms of the cyclic out-vacuum \( |0 : out\rangle \) defined by

\[
b(\vec{k})|0 : out\rangle = 0.
\] (7.11.144)

The one-particle out-states are again given by

\[
|\vec{k} : out\rangle = b^{\dagger}(\vec{k})|0 : out\rangle
\] (7.11.145)

The two-particle out-states are also given by

\[
|\vec{k}, \vec{p} : out\rangle = b^{\dagger}(\vec{k})b^{\dagger}(\vec{p})|0 : out\rangle
\] (7.11.146)

e etc. We next relate these two free fields to each other.
7.12 The S-Matrix and the S-Operator

The single particle out-states \(|q : out\rangle\) are related to the single particle in-states \(|k : in\rangle\) by the S-matrix.

\[
|q : out\rangle = \int S(q, k) |k : in\rangle \, d^3k .
\]

(7.12.147)

The S-matrix is the S-operator evaluated between single particle states. Furthermore, the S-operator maps the in-field onto the out-field according to

\[
\phi_{out} = S \phi_{in} S^\dagger
\]

(7.12.148)

By equating the two forms of the solution for the interpolating field \(\phi\) to each other we get

\[
\phi_{out}(x) + g \int \Delta_A(x - y) \rho(y) \, d^4y = \phi_{in}(x) + g \int \Delta_R(x - y) \rho(y) \, d^4y .
\]

(7.12.149)

This shows that the "interpolating field" is so called because it interpolates between the "in" and the "out" fields. We can now solve for \(\phi_{out}\) in terms of \(\phi_{in}\).

\[
\phi_{out}(x) = \phi_{in}(x) - g \int \left[ \Delta_A(x - y) - \Delta_R(x - y) \right] \rho(y) \, d^4y
\]

(7.12.150)

or

\[
\phi_{out}(x) = \phi_{in}(x) - g \int \Delta(x - y) \rho(y) \, d^4y .
\]

(7.12.151)

This can also be written

\[
\phi_{out}(x) = \phi_{in}(x) - gf(x) .
\]

(7.12.152)

where

\[
f(x) = \int \Delta(x - y) \rho(y) \, d^4y
\]

(7.12.153)

is just a c-number function. Notice that the function \(f(x)\) satisfies the Klein-Gordon equation. This allows us to expand this function in a Fourier integral in the same way as the field \(\phi_{in}(x)\) itself and leads to a somewhat more transparent form.

\[
f(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ c(\vec{k}) e^{-ikx} + c^*(\vec{k}) e^{ikx} \right] .
\]

(7.12.154)

Here, \(c(\vec{k})\) and \(c^*(\vec{k})\) are just c-number functions. Then,

\[
b(\vec{k}) = a(\vec{k}) - gc(\vec{k})
\]

\[
b^\dagger(\vec{k}) = a^\dagger(\vec{k}) - gc^*(\vec{k}) .
\]

(7.12.155)
The function \( c(\vec{k}) \) is explicitly given by

\[
c(\vec{k}) = -\frac{i}{\sqrt{2(2\pi)^3}} \int d^3x \, d^4y \left[ \Delta(x-y) \, \partial_x^0 \, e^{ikx} \right] \rho(y) \, .
\] (7.12.156)

But,

\[
\Delta(x-y) = -\frac{i}{2(2\pi)^3} \int \frac{d^3p}{p^0} \left[ e^{-ip(x-y)} - e^{ip(x-y)} \right] \, .
\] (7.12.157)

Therefore,

\[
c(\vec{k}) = \frac{-i}{[2(2\pi)^3]^{3/2}} \int d^3x \, d^4y \, \frac{d^3p}{p^0} \left\{ \left[ e^{-ip(x-y)} - e^{ip(x-y)} \right] k^0 e^{ikx} \\
+ \left[ e^{-ip(x-y)} + e^{ip(x-y)} \right] p^0 e^{ikx} \right\} \rho(y) \\
= \frac{-i}{[2(2\pi)^3]^{3/2}} \int d^4y \, \frac{d^3p}{p^0} \rho(y) \int d^3x \left\{ (k^0 + p^0) e^{-i(p-k)x+ipy} \\
- (k^0 - p^0) e^{i(p+k)x+ipy} \right\} \\
= \frac{-i}{[2(2\pi)^3]^{3/2}} \int d^4y \, \frac{d^3p}{p^0} \rho(y) \delta(\vec{p} - \vec{k}) \, e^{ipy} \\
= \frac{-i}{\sqrt{2(2\pi)^3}} \int d^4y \, \rho(y) \, e^{i(\omega(\vec{k})y^0 - \vec{k} \cdot \vec{y})} \\
= \frac{-i}{\sqrt{2(2\pi)^3}} \check{\rho}(\omega(\vec{k}), \vec{k}) \, .
\] (7.12.158)

Of course, \( c^*(\vec{k}) \) is simply the complex conjugate

\[
c^*(\vec{k}) = \frac{i}{\sqrt{2(2\pi)^3}} \check{\rho}^*(\omega(\vec{k}), \vec{k}) \, .
\] (7.12.159)

To find the S-matrix we now use that

\[
b(\vec{k}) = Sa(\vec{k})S^\dagger = a(\vec{k}) - gc(\vec{k})
\] (7.12.160)

is a simple translation. So, since \([A, B] = c\)-number we have the identity

\[
e^B A = A e^B - [A, B] e^B
\] (7.12.161)

we try the Ansatz

\[
S = \exp \left\{ i \int \frac{d^3q}{q^0} \left[ s(\vec{q}) a(\vec{q}) + s(\vec{q})^* a^*(\vec{q}) \right] \right\} \, .
\] (7.12.162)

Then,

\[
S^\dagger = \exp \left\{ -i \int \frac{d^3q}{q^0} \left[ s(\vec{q}) a(\vec{q}) + s(\vec{q})^* a^*(\vec{q}) \right] \right\}
\] (7.12.163)
and using our identity we find

\[
\begin{align*}
\alpha(\vec{k}) S^\dagger &= S^\dagger \alpha(\vec{k}) - i \int \frac{d^3q}{q^0} s^* (\vec{q}) q^0 \delta(\vec{k} - \vec{q}) S^\dagger \\
&= S^\dagger \alpha(\vec{k}) - is^* (\vec{k}) S^\dagger .
\end{align*}
\]

(7.12.164)

Therefore,

\[
\begin{align*}
Sa(\vec{k}) S^\dagger &= \alpha(\vec{k}) - is^* (\vec{k}) \\
&= \alpha(\vec{k}) - gc(\vec{k})
\end{align*}
\]

(7.12.165)

So, our Ansatz works if

\[
\begin{align*}
s^*(\vec{k}) &= -igc(\vec{k}) \\
s(\vec{k}) &= igc^* (\vec{k}) .
\end{align*}
\]

(7.12.166)

It now follows that

\[
S = \exp \left\{ - \int \frac{d^3q}{q^0} g \left[ c^* (\vec{q}) a(\vec{q}) - c(\vec{q}) a^\dagger (\vec{q}) \right] \right\} .
\]

(7.12.167)

We leave it as an exercise (Problem 7.2) to verify that

\[
SS^\dagger = S^\dagger S = 1 .
\]

(7.12.168)

It is also useful to rewrite the $S$-operator in normal ordered form. To do this we use the Campbell-Baker-Hausdorff formula which for $[A, B] = c$-number becomes

\[
e^{A+B} = e^A e^B e^{-1/2[A,B]} .
\]

(7.12.169)

Then,

\[
\begin{align*}
S &= \exp \left\{ \int \frac{d^3q}{q^0} gc(\vec{q}) a^\dagger (\vec{q}) \right\} \exp \left\{ - \int \frac{d^3q}{q^0} gc^* (\vec{q}) a(\vec{q}) \right\} \\
&\quad \times \exp \left\{ \frac{1}{2} \int \frac{d^3q}{q^0} \frac{d^3k}{k^0} gc^* (\vec{q}) gc(\vec{k}) [a(\vec{k}), a^\dagger (\vec{q})] \right\} \\
&= \exp \left\{ \int \frac{d^3q}{q^0} gc(\vec{q}) a^\dagger (\vec{q}) \right\} \exp \left\{ - \int \frac{d^3q}{q^0} gc^* (\vec{q}) a(\vec{q}) \right\} \\
&\quad \times \exp \left\{ \frac{1}{2} \int \frac{d^3q}{q^0} [gc(\vec{q})]^2 \right\} .
\end{align*}
\]

(7.12.170)

We can use this result to immediately obtain an explicit form for

\[
|0 : out \rangle = S |0 : in \rangle .
\]

(7.12.171)

Thus,

\[
|0 : out \rangle = \exp \left\{ \frac{1}{2} \int \frac{d^3q}{q^0} [gc(\vec{q})]^2 \right\} \exp \left\{ \int \frac{d^3q}{q^0} gc(\vec{q}) a^\dagger (\vec{q}) \right\} |0 : in \rangle .
\]

(7.12.172)
This shows that that $|0 : out\rangle$ is a coherent state of “in”-particles. One can now go on and relate any “in” and “out” state.

This result may be understood as follows. The time-dependent source produces particle-antiparticle pairs from the vacuum. Once the source is turned off, these pairs remain since they do not interact with each other and can not annihilate each other. This state now does not change in time and forms the lowest energy state. The state with all these particles is thus the vacuum for the field after the source is turned off (the out-field). This is the state we have called $|0 : out\rangle$.

### 7.13 Single Particle Matrix Elements

We first compute the probability amplitude that the in-vacuum remains a vacuum in the remote future. This means we compute

$$\langle 0 : out|0 : in \rangle = \langle 0 : in|S^\dagger|0 : in \rangle = \exp \left\{ \frac{1}{2} \int \frac{d^3q}{q^0} |gc(\vec{q})|^2 \right\}. \quad (7.13.173)$$

The probability amplitude for finding a one-particle state with momentum $\vec{k}$ in the outgoing beam when the incoming beam contained a particle of momentum $\vec{p}$ is

$$\langle \vec{k} : out|\vec{p} : in \rangle = \langle \vec{k} : in|S^\dagger|\vec{p} : in \rangle$$

$$= \langle 0 : out|b(\vec{k}) a^\dagger(\vec{p})|0 : in \rangle$$

$$= \langle 0 : in|S^\dagger[a(\vec{k}) - gc(\vec{k})]a^\dagger(\vec{p})|0 : in \rangle$$

$$= e^{\frac{i}{2} \int \frac{d^3q}{q^0} gc(\vec{q})^2} \langle 0 : in|e^{\int \frac{d^3q}{q^0} gc^*(\vec{q})a(\vec{q})[a(\vec{k}) - gc(\vec{k})]a^\dagger(\vec{p})|0 : in \rangle$$

$$= \exp \left\{ \frac{1}{2} \int \frac{d^3q}{q^0} |gc(\vec{q})|^2 \right\} \left[ \langle 0 : in| \exp \left\{ \int \frac{d^3q}{q^0} gc^*(\vec{q})a(\vec{q}) \right\} k^0 \delta(\vec{k} - \vec{q})|0 : in \rangleight.$$ 

$$- \langle 0 : in| \exp \left\{ \int \frac{d^3q}{q^0} gc^*(\vec{q})a(\vec{q}) \right\} a^\dagger(\vec{p})gc(\vec{k})|0 : in \rangle \right]$$

$$= \left[ k^0 \delta(\vec{k} - \vec{p}) - gc^*(\vec{p}) gc(\vec{k}) \right] e^{\frac{q^2}{2} \int \frac{d^3q}{q^0} gc(\vec{q})^2}. \quad (7.13.174)$$

So, combining this with our previous result we get

$$S(\vec{k}, \vec{p}) = \frac{\langle \vec{k} : out|\vec{p} : in \rangle}{\langle 0 : out|0 : in \rangle} = k^0 \delta(\vec{k} - \vec{p}) - g^2 c^*(\vec{p}) c(\vec{k}). \quad (7.13.175)$$

It is noteworthy that the single-particle S-matrix as expressed here is not unitary even though the S-operator is. The reason for this is that the completeness relation in Fock space is expressed by

$$1 = \sum_{n=0}^{\infty} \int \frac{d^3q_1}{q_1^0} \cdots \frac{d^3q_n}{q_n^0} |\vec{q}_1 \cdots \vec{q}_n : in\rangle \langle \vec{q}_1 \cdots \vec{q}_n : in|$$
Thus, in terms of matrix elements the unitarity condition reads
\[ 1 = \langle \vec{k} : in | S S^\dagger | \vec{p} : in \rangle \]
\[ = \sum_{n=0}^{\infty} \int \frac{d^3q_1}{q_1^0} \ldots \frac{d^3q_n}{q_n^0} \langle \vec{k} : in | S | \vec{q}_1 \ldots \vec{q}_n : out \rangle \langle \vec{q}_1 \ldots \vec{q}_n : out | \vec{p} : in \rangle \]
\[ = \sum_{n=0}^{\infty} \int \frac{d^3q_1}{q_1^0} \ldots \frac{d^3q_n}{q_n^0} \langle \vec{k} : in | \vec{q}_1 \ldots \vec{q}_n : out \rangle \langle \vec{q}_1 \ldots \vec{q}_n : out | \vec{p} : in \rangle \]
\[ = \langle \vec{k} : in | \vec{p} : in \rangle \]
\[ = k^0 \delta(\vec{k} - \vec{p}) \] \hspace{1cm} (7.13.177)

This is not the same as
\[ \int \frac{d^3q}{q^0} S(\vec{k}, \vec{q}) S^*(\vec{q}, \vec{p}) \] \hspace{1cm} .

There is also a physical reason why the single-particle S-matrix is not unitary. Particles are not conserved; as we saw, particle pairs are created out of the vacuum.

7.14 Problems

7.1 In the integrals
\[ \Delta_R(x - y) = -\frac{1}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{(k^0 + i\epsilon)^2 - \omega(k)^2} d^4k \]
\[ \Delta_A(x - y) = -\frac{1}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{(k^0 - i\epsilon)^2 - \omega(k)^2} d^4k \]

carry out the integrations over \( k^0 \) and use the results of section 7.9 to show that
\[ \Delta_R(x - y) = -\theta(x^0 - y^0) \Delta(x - y) \]

as well as
\[ \Delta_A(x - y) = \theta(y^0 - x^0) \Delta(x - y) \]

so that
\[ \Delta(x - y) = \Delta_A(x - y) - \Delta_R(x - y) \] \hspace{1cm} (7.14.182)

7.2 For the operator
\[ S = \exp \left\{ -\int \frac{d^3q}{q^0} \left[ c^*(\vec{q})a(\vec{q}) - c(\vec{q})a^\dagger(\vec{q}) \right] \right\} \]

verify that
\[ SS^\dagger = S^\dagger S = 1 \] \hspace{1cm} (7.14.184)
Chapter 8

The Charged Klein-Gordon Field

8.1 Gauge Invariance

To get a non-zero current for the Klein-Gordon field requires that the classical field be complex and that the corresponding quantized field be non-hermitian. Thus, we start with the lagrange density

\[ \mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi : \]  

(8.1.1)

By varying \( \phi^\dagger \) and \( \phi \) separately we get the equations of motion

\[ (\Box + m^2) \phi = 0 , \quad (\Box + m^2) \phi^\dagger = 0 . \]  

(8.1.2)

The conjugate momenta are

\[ \pi^\dagger \equiv \pi_{\phi^\dagger} = \dot{\phi} , \quad \pi \equiv \pi_\phi = \dot{\phi}^\dagger . \]  

(8.1.3)

The hamiltonian density is

\[ \mathcal{H} = \pi^\dagger \pi - \nabla \phi^\dagger \cdot \nabla \phi + m^2 \phi^\dagger \phi : \]  

(8.1.4)

Both the lagrange density and the hamiltonian density are invariant under a global gauge transformation of the form

\[ \phi(x) \to \phi'(x) = e^{i\alpha} \phi(x) \]

\[ \phi^\dagger(x) \to \phi'^\dagger(x) = e^{-i\alpha} \phi^\dagger(x) \]  

(8.1.5)

where \( \alpha \) is a constant. The fact that \( \alpha \) is a constant, and not a function of the spacetime coordinates \( x \), is what is meant by saying that this is a global rather than a local gauge transformation. We now use this invariance to obtain
a conserved current. To do this we perform the following variation of the fields.

\[ \delta \phi(x) \equiv \phi'(x) - \phi(x) \]

\[ = \alpha \frac{\partial \phi'(x)}{\partial \alpha} \bigg|_{\alpha=0} + O(\alpha^2) \]

\[ = i\alpha \phi(x) \]

\[ \delta \phi^\dagger(x) \equiv \phi'^\dagger(x) - \phi^\dagger(x) \]

\[ = -i\alpha \phi^\dagger(x) \quad (8.1.6) \]

Since, as already stated, the lagrange density \( \mathcal{L} \) is invariant under this global gauge transformation we have

\[ \delta \mathcal{L} = \mathcal{L}(\phi', \phi'^\dagger) - \mathcal{L}(\phi, \phi^\dagger) = 0 \quad (8.1.7) \]

Now, quite generally if we vary the lagrange density with respect to \( \phi \) and \( \phi^\dagger \) we get

\[ \delta \mathcal{L} = \left[ \frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x^\mu} \left( \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \right) \right] \delta \phi + \left[ \frac{\partial \mathcal{L}}{\partial \phi^\dagger} - \frac{\partial}{\partial x^\mu} \left( \frac{\partial \mathcal{L}}{\partial \phi^\dagger_{,\mu}} \right) \right] \delta \phi^\dagger \]

\[ + \frac{\partial}{\partial x^\mu} \left[ \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta \phi + \frac{\partial \mathcal{L}}{\partial \phi^\dagger_{,\mu}} \delta \phi^\dagger \right] \quad (8.1.8) \]

If, in addition, the action arising from this lagrangian is varied with the fields fixed at the end points so that

\[ \delta \phi(t_1) = \delta \phi(t_2) = \delta \phi^\dagger(t_1) = \delta \phi^\dagger(t_2) = 0 \quad (8.1.9) \]

and if the action is required to be an extremum, then the fields \( \phi \) and \( \phi^\dagger \) satisfy the Euler-Lagrange equations and we find that the change in the lagrange density computed above reduces to

\[ \delta \mathcal{L} = \partial_\mu J^\mu \quad (8.1.10) \]

where

\[ J^\mu = \left[ \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta \phi + \frac{\partial \mathcal{L}}{\partial \phi^\dagger_{,\mu}} \delta \phi^\dagger \right] \quad (8.1.11) \]

Therefore for the case in hand, where the lagrange density is invariant under this global gauge transformation, we have

\[ \delta \mathcal{L} = 0 \quad (8.1.12) \]

and therefore the current \( J^\mu \) is conserved. This is an example of the Noether Theorem. In this case the explicit form of \( J^\mu \) is

\[ J^\mu = i\alpha : [(\partial^\mu \phi^\dagger)\phi - (\partial^\mu \phi)\phi^\dagger] : \quad (8.1.13) \]
We choose to normalize this current so that it reads

\[
j^\mu(x) = i e : \[(\partial^\mu \phi^\dagger)\phi - (\partial^\mu \phi)\phi^\dagger : (x) . \quad (8.1.14)
\]

Up to here we might as well have been dealing with a classical field since the only indication that this field is quantized comes from the colons indicating normal ordering. To carry out the explicit quantization in terms of annihilation and creation operators we split the non-hermitian field \( \phi \) into two independent (that is commuting) hermitian fields \( \phi_1 \) and \( \phi_2 \) since we have already studied such fields in the previous chapter. To this end we write

\[
\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \quad \phi^\dagger = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2) . \quad (8.1.15)
\]

Then,

\[
\mathcal{L} = \frac{1}{2} \left[ \partial_\mu \phi_1 \partial^\mu \phi_1 - m^2 \phi_1^2 \right] + \frac{1}{2} \left[ \partial_\mu \phi_2 \partial^\mu \phi_2 - m^2 \phi_2^2 \right] . \quad (8.1.16)
\]

The equations of motion are

\[
(\Box + m^2) \phi_1 = 0 \\
(\Box + m^2) \phi_2 = 0 . \quad (8.1.17)
\]

Since both \( \phi_1 \) and \( \phi_2 \) are hermitian fields we quantize them in the same manner as before and write

\[
\phi_1(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ a(\vec{k}) e^{-ikx} + a^\dagger(\vec{k}) e^{ikx} \right]
\]

\[
\phi_2(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ b(\vec{k}) e^{-ikx} + b^\dagger(\vec{k}) e^{ikx} \right] . \quad (8.1.18)
\]

Also, since the two fields are independent they commute and we have

\[
[a(\vec{k}), b(\vec{q})] = [a(\vec{k}), b^\dagger(\vec{q})] = 0 \quad \text{etc.} . \quad (8.1.19)
\]

The complex field \( \phi \) can now be written

\[
\phi(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ \frac{a(\vec{k}) + ib(\vec{k})}{\sqrt{2}} e^{-ikx} + \frac{a^\dagger(\vec{k}) + ib^\dagger(\vec{k})}{\sqrt{2}} e^{ikx} \right] . \quad (8.1.20)
\]

If we now define new operators

\[
c(\vec{k}) = \frac{a(\vec{k}) + ib(\vec{k})}{\sqrt{2}} \quad c^\dagger(\vec{k}) = \frac{a^\dagger(\vec{k}) - ib^\dagger(\vec{k})}{\sqrt{2}}
\]

\[
d(\vec{k}) = \frac{a(\vec{k}) - ib(\vec{k})}{\sqrt{2}} \quad d^\dagger(\vec{k}) = \frac{a^\dagger(\vec{k}) + ib^\dagger(\vec{k})}{\sqrt{2}} \quad (8.1.21)
\]
we find that they also are just annihilation and creation operators

\[
[c(\vec{k}), c^\dagger(\vec{q})] = k^0 \delta(\vec{k} - \vec{q})
\]

\[
[d(\vec{k}), d^\dagger(\vec{q})] = k^0 \delta(\vec{k} - \vec{q})
\]

\[
[c(\vec{k}), d(\vec{q})] = 0
\]

\[
[c(\vec{k}), d^\dagger(\vec{q})] = 0.
\] (8.1.22)

In this case we can write

\[
\phi(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ c(\vec{k}) e^{-ikx} + d^\dagger(\vec{k}) e^{ikx} \right]
\]

\[
\phi^\dagger(x) = \frac{1}{\sqrt{2(2\pi)^3}} \int \frac{d^3k}{k^0} \left[ d(\vec{k}) e^{-ikx} + c^\dagger(\vec{k}) e^{ikx} \right].
\] (8.1.23)

To see what this means we first compute the electric charge for this field, namely

\[
Q = \int j^0 d^3x = ie \int d^3x : \phi^\dagger \partial^0 \phi : (x).
\] (8.1.24)

So, we find

\[
Q = ie \int \frac{d^3x}{2(2\pi)^3} \int \frac{d^3q d^3k}{q^0 k^0} \times
\]

\[
\left\{ [d(\vec{q}) e^{-iqx} + c^\dagger(\vec{q}) e^{iqx}] (ik^0) \left[ c(\vec{k}) e^{-ikx} - d^\dagger(\vec{k}) e^{ikx} \right] + iq^0 \left[ d(\vec{q}) e^{-iqx} - c^\dagger(\vec{q}) e^{iqx} \right] \left[ c(\vec{k}) e^{-ikx} + d^\dagger(\vec{k}) e^{ikx} \right] \right\} :
\]

\[
= e \int \frac{d^3q}{q^0} \frac{d^3k}{k^0} \int \frac{d^3x}{2(2\pi)^3} \times \left\{ k^0 \left[ d(\vec{q}) c(\vec{k}) e^{-i(k+q)x} - d(\vec{q}) d^\dagger(\vec{k}) e^{-i(q-k)x} \right]
\]

\[
+ c^\dagger(\vec{q}) c(\vec{k}) e^{i(q-k)x} - c^\dagger(\vec{q}) d^\dagger(\vec{k}) e^{i(q+k)x} \right]\right. 
\]

\[
- q^0 \left[ d(\vec{q}) c(\vec{k}) e^{-i(k+q)x} + d(\vec{q}) d^\dagger(\vec{k}) e^{-i(q-k)x} \right]
\]

\[
- c^\dagger(\vec{q}) c(\vec{k}) e^{i(q-k)x} - c^\dagger(\vec{q}) d^\dagger(\vec{k}) e^{i(q+k)x} \right\}:
\]

\[
= e \int \frac{d^3q}{q^0} \frac{d^3k}{k^0} \frac{d^3x}{q^0} \times \left\{ c^\dagger(\vec{q}) c(\vec{k}) - d^\dagger(\vec{k}) d(\vec{q}) \right\} : \delta(\vec{k} - \vec{q})
\]

\[
= e \int \frac{d^3k}{k^0} \left[ c^\dagger(\vec{k}) c(\vec{k}) - d^\dagger(\vec{k}) d(\vec{k}) \right].
\] (8.1.25)

If we now define the two particle number densities

\[
n(\vec{k}) = c^\dagger(\vec{k}) c(\vec{k})
\]

\[
\bar{n}(\vec{k}) = d^\dagger(\vec{k}) d(\vec{k})
\] (8.1.26)

then the charge operator may be written

\[
Q = e \int \frac{d^3k}{k^0} \left[ n(\vec{k}) - \bar{n}(\vec{k}) \right].
\] (8.1.27)
so the charge is the difference of two equal types of charge operators. We now interpret
\[ e \int \frac{d^3k}{k^0} n(\tilde{k}) \]
as the charge operator for particles and
\[ e \int \frac{d^3k}{k^0} \tilde{n}(\tilde{k}) \]
as the charge operator for antiparticles. For a real scalar field we clearly have \( Q = 0 \) since for such a field the particles are their own antiparticles and therefore \( n(\tilde{k}) = \tilde{n}(\tilde{k}) \). Notice, that by interpreting \( Q \) as a charge density we not only resolved the difficulty of negative probabilities, but we also were naturally led to the concept of antiparticle.

### 8.2 Interaction with an External Electromagnetic Field

If, as always, we do minimal coupling the lagrange density becomes
\[ \mathcal{L} =: (\partial_\mu + ieA_\mu)\phi^\dagger(\partial^\mu - ieA^\mu)\phi - m^2\phi^\dagger\phi : \]  
(8.2.28)

Before proceeding we first compute the current for this interacting field. To do so we again use the fact that the lagrange density is invariant under a global gauge transformation.

\[ \phi \rightarrow e^{ie\phi} \phi^\dagger \rightarrow e^{-ie}\phi^\dagger \]  
(8.2.29)

so that
\[ \delta \phi = ie\phi \quad \delta \phi^\dagger = -ie\phi^\dagger . \]  
(8.2.30)

Then, the Noether current is given by
\[ J^\mu := \frac{\partial \mathcal{L}}{\partial \phi^\dagger_\mu} \delta \phi + \frac{\partial \mathcal{L}}{\partial \phi_\mu} \delta \phi^\dagger : 
= ie : [(\partial^\mu + ieA^\mu)\phi^\dagger\phi - \phi^\dagger(\partial^\mu - ieA^\mu)\phi] : . \]  
(8.2.31)

Thus, the electric current is given by
\[ j^\mu(x) = ie : \phi^\dagger D^\mu \phi - (D^\mu \phi)^\dagger \phi : (x) = ie : \phi^\dagger \overset{\leftrightarrow}{D^\mu} \phi : (x) \]  
(8.2.32)

where
\[ D^\mu \phi = (\partial^\mu - ieA^\mu)\phi \]
\[ (D^\mu \phi)^\dagger = (\partial^\mu + ieA^\mu)\phi^\dagger . \]  
(8.2.33)
With this in mind we notice that we can rewrite the lagrange density as

\[ \mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I \]  \hspace{1cm} (8.2.34)

where \( \mathcal{L}_0 \) is the free lagrange density

\[ \mathcal{L}_0 =: \partial_\mu \phi^\dagger \partial^\mu \phi - m^2 \phi^\dagger \phi : \]  \hspace{1cm} (8.2.35)

and \( \mathcal{L}_I \) is the interaction part of the lagrange density

\[ \mathcal{L}_I = j^\mu (x) A_\mu (x) . \]  \hspace{1cm} (8.2.36)

The equations of motion, or field equations, are just the Euler-Lagrange equations.

\[ (\partial^\mu - ieA^\mu)(\partial_\mu - ieA_\mu)\phi + m^2 \phi = 0 . \]  \hspace{1cm} (8.2.37)

The canonical momenta are

\[ \pi_\phi = \frac{\partial \mathcal{L}}{\partial \phi} = (D_0 \phi)^\dagger = (\partial_0 + ieA_0)\phi^\dagger \]

\[ \pi_\phi^\dagger = \frac{\partial \mathcal{L}}{\partial \phi^\dagger} = D_0 \phi = (\partial_0 - ieA_0)\phi . \]  \hspace{1cm} (8.2.38)

The canonical equal time commutation relations are as before given by

\[ [\phi(x), \pi_\phi(y)]_{x^0 = y^0} = i\delta (\vec{x} - \vec{y}) . \]  \hspace{1cm} (8.2.39)

This means that

\[ [\phi(x), \phi^\dagger(y) + ieA_0(y)\phi^\dagger(y)]_{x^0 = y^0} = i\delta (\vec{x} - \vec{y}) . \]  \hspace{1cm} (8.2.40)

The problem is now completely specified. The question is how to proceed. The standard way is to treat the interaction as a perturbation on the free field. We sketch this in the next section.

### 8.3 Perturbation Theory

In order to proceed with perturbation theory we rewrite the field equations in the form

\[ (\Box + m^2)\phi = ie [A_\mu (\partial^\mu - ieA^\mu)\phi + (\partial^\mu A_\mu)\phi] \]

\[ = ieA_\mu \partial^\mu \phi + ie(\partial^\mu A_\mu)\phi + e^2 A_\mu A^\mu \phi . \]  \hspace{1cm} (8.3.41)

If we furthermore choose the gauge of the external field such that \( \partial \cdot A = 0 \), the so-called Lorentz gauge, then the right hand side of the equation above reduces to

\[ 2ieA_\mu \partial^\mu \phi + e^2 A^2 \phi . \]

Here we have used the shorthand

\[ A^2 = A_\mu A^\mu . \]  \hspace{1cm} (8.3.42)
In this form the Källén-Yang-Feldman equations read

$$\phi(x) = \phi_{in}(x) + \int \Delta_R(x-y)[2ieA_\mu(y)\partial^\mu + e^2A^2(y)]\phi(y)\,d^4y$$  \hfill (8.3.43)

or

$$\phi(x) = \phi_{out}(x) + \int \Delta_A(x-y)[2ieA_\mu(y)\partial^\mu + e^2A^2(y)]\phi(y)\,d^4y$$  \hfill (8.3.44)

These equations are still exact. If we now treat the right-hand side as a perturbation we get to lowest order in $e$

$$\phi(x) = \phi^{(0)}(x) + e\phi^{(1)}(x) + e^2\phi^{(2)}(x) + \ldots$$

$$= \phi_{in}(x) + 2ie \int \Delta_R(x-y)A_\mu(y)\partial^\mu\phi_{in}(y)\,d^4y$$

$$= \phi_{out}(x) + 2ie \int \Delta_A(x-y)A_\mu(y)\partial^\mu\phi_{out}(y)\,d^4y .$$  \hfill (8.3.45)

This is the beginning of the perturbation calculation. We leave this since in the next chapter we repeat these computations for the Dirac field in which case the results are of much greater interest.

### 8.4 Problems

1. For a quantized Klein-Gordon field solve the problem of scattering a particle of mass $m$ and momentum $\vec{p}$ off an electrostatic potential step

$$V(z) = \begin{cases} 
0 & z < 0 \\
V_0 & z > 0
\end{cases}$$

if the energy $E$ of the particle is

a) $E > V_0$

b) $E < V_0$. 
Chapter 9

The Dirac Field

9.1 The Free Dirac Field

We begin immediately with the lagrange density for the free Dirac field.

\[ \mathcal{L} = -\frac{1}{2} \bar{\psi} (i \gamma \cdot \partial + m) \psi - \frac{1}{2} (i \partial_\mu \bar{\psi} \gamma^\mu + m \bar{\psi}) \psi \ . \quad (9.1.1) \]

Here, \( \psi \) and \( \bar{\psi} \) are to be varied independently. Thus, by varying \( \bar{\psi} \) we get

\[ \frac{\partial \mathcal{L}}{\partial \bar{\psi}} = -\frac{1}{2} (-i \gamma \cdot \partial + m) \psi - \frac{1}{2} m \psi \quad (9.1.2) \]

and

\[ \frac{\partial}{\partial x^\mu} \left( \frac{\partial \mathcal{L}}{\partial \partial_{\mu} \psi} \right) = -\frac{1}{2} (i \gamma \cdot \partial) \psi \ . \quad (9.1.3) \]

So the Euler-Lagrange equation that results is just the Dirac equation

\[ (-i \gamma \cdot \partial + m) \psi = 0 \ . \quad (9.1.4) \]

The energy-momentum tensor is given by

\[ T^{\mu \nu} = \frac{\partial \mathcal{L}}{\partial \partial_{\nu} \psi} \frac{\partial \psi}{\partial x^\mu} + \frac{\partial \bar{\psi}}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \partial_{\nu} \bar{\psi}} - g^{\mu \nu} \mathcal{L} \quad (9.1.5) \]

Now if \( \psi \) and \( \bar{\psi} \) satisfy their respective Euler-Lagrange equations then we have that \( \mathcal{L} = 0 \) and so we get

\[ T^{\mu \nu} = i \frac{1}{2} (\bar{\psi} \gamma^\nu \partial^\mu \psi - \partial^\mu \bar{\psi} \gamma^\nu \psi) \quad (9.1.6) \]

This tensor is obviously not symmetrized but may be symmetrized. However, since this does not affect our results we do not bother to do so. On the other hand \( T^{\mu \nu} \) satisfies an equation of continuity (is conserved)

\[ \partial_\nu T^{\mu \nu} = 0 \quad (9.1.7) \]
The conserved four-momentum operator $P^\mu$ does not depend on the symmetrization and is given by

$$P^\nu = \int T^{0\nu} \, d^3x = \frac{i}{2} \int d^3x \left( \bar{\psi} \gamma^\nu \partial^0 \psi - \partial^0 \bar{\psi} \gamma^\nu \psi \right)$$

$$= i \int d^3x \left( \bar{\psi} \gamma^\nu \partial^0 \psi \right)$$

(9.1.8)

Also, in particular we find that the Hamiltonian $H = P^0$ is given by

$$H = P^0 = \int T^0 \, d^3x = i \int d^3x \left( \bar{\psi} \gamma^0 \partial^0 \psi \right)$$

$$= \int d^3x \bar{\psi}(x) \left( -i \gamma \cdot \vec{\nabla} + m \right) \psi(x)$$

$$= \int d^3x \psi^\dagger(x) \left( -i \alpha \cdot \vec{\nabla} + \beta m \right) \psi(x)$$

(9.1.9)

At this stage we would like to study the algebra of the field operators by imposing canonical quantization. However, there is a problem since $\bar{\pi}_\alpha$ and $\pi_\alpha$ are not independent as we now show.

$$\pi_\alpha = \frac{\partial L}{\partial \psi_\alpha} = \frac{i}{2} \bar{\psi} \gamma^0 = \frac{i}{2} \psi^\dagger$$

$$\bar{\pi}_\alpha = \frac{\partial L}{\partial \bar{\psi}_\alpha} = -\frac{i}{2} \gamma^0 \psi$$

(9.1.10)

To remedy this we consider the lagrange density

$$L = \bar{\psi} (i \gamma \cdot \partial - m) \psi.$$  

(9.1.11)

In this case we obtain the same Euler-Lagrange equation for $\psi$ if we vary only $\bar{\psi}$. Furthermore, the canonical momentum conjugate to $\psi_\alpha$ is

$$\pi_\alpha = \frac{\partial L}{\partial \dot{\psi}_\alpha} = i \psi^\dagger_\alpha.$$  

(9.1.12)

The Hamiltonian density is also the same as before.

$$H = \sum_\alpha \pi_\alpha \psi_\alpha - L$$

$$= \psi^\dagger i \frac{\partial \psi}{\partial x^0}$$

$$= \psi^\dagger \left( -i \alpha \cdot \vec{\nabla} + \beta m \right) \psi.$$  

(9.1.13)

### 9.2 Quantization

We want the quantization to be such that the Heisenberg equations of motion reproduce the Euler-Lagrange equations. In other words we want that

$$i \dot{\psi} = [\psi(\vec{x},t), H]$$
\[
\begin{align*}
&= \int d^3y \{\psi(\vec{x}, t), \psi^\dagger(\vec{y}, t)(-i\vec{\alpha} \cdot \vec{\nabla} + \beta m)\psi(\vec{y}, t)\} \\
&= (-i\vec{\alpha} \cdot \vec{\nabla} + \beta m)\psi(\vec{x}, t). 
\end{align*}
\]

At this stage we find that we have two possibilities: the fields can satisfy either commutation or else anticommutation relations. In either case we have the following expressions.

\[
\begin{align*}
[\psi(\vec{x}, t), \pi(\vec{y}, t)]_\pm &= \psi(\vec{x}, t)\pi(\vec{y}, t) \pm \pi(\vec{y}, t)\psi(\vec{x}, t) = i\delta(\vec{x} - \vec{y}) \\
[\psi(\vec{x}, t), \psi^\dagger(\vec{y}, t)]_\pm &= \psi(\vec{x}, t)\psi^\dagger(\vec{y}, t) \mp \psi^\dagger(\vec{y}, t)\psi(\vec{x}, t) = 0. 
\end{align*}
\]

(9.2.15)

Since \( \pi = i\psi^\dagger \), the first equation is equivalent to

\[ [\psi(\vec{x}, t), \psi^\dagger(\vec{y}, t)]_\pm = \delta(\vec{x} - \vec{y}). \]

(9.2.16)

In this case we leave the decision, as to whether to use commutators or anticommutators, until later when we shall be forced by the requirement that the energy of a free particle has to be positive, to choose the anticommutation relations.

Now, just as for the Klein-Gordon field, we expand the Dirac field in terms of the free field c-number solutions. Thus, we write

\[
\begin{align*}
\psi(x) &= \sum_s \int \frac{d^3p}{(2\pi)^3/2} \sqrt{\frac{m}{E(p)}} \left[ b(\vec{p}, s)u(\vec{p}, s) e^{-ipx} + d^\dagger(\vec{p}, s)v(\vec{p}, s) e^{ipx} \right] \\
\psi^\dagger(x) &= \sum_s \int \frac{d^3p}{(2\pi)^3/2} \sqrt{\frac{m}{E(p)}} \left[ b^\dagger(\vec{p}, s)u^\dagger(\vec{p}, s) e^{ipx} + d(\vec{p}, s)v^\dagger(\vec{p}, s) e^{-ipx} \right].
\end{align*}
\]

(9.2.17)

Here, the factor \( \sqrt{m/E(\vec{p})} \) is conventional and

\[ E(\vec{p}) = \sqrt{\vec{p}^2 + m^2}. \]

(9.2.18)

Also, as before,

\[
\begin{align*}
(\gamma \cdot p - m)u(\vec{p}) &= 0 \quad \bar{u}(\vec{p})(\gamma \cdot p - m) = 0 \\
(\gamma \cdot p + m)v(\vec{p}) &= 0 \quad \bar{v}(\vec{p})(\gamma \cdot p + m) = 0.
\end{align*}
\]

(9.2.19)

We also recall the orthogonality conditions.

\[
\begin{align*}
\bar{u}(\vec{p}, s)u(\vec{p}, s') &= -\bar{v}(\vec{p}, s)v(\vec{p}, s') = \delta_{ss'}, \\
\bar{u}^\dagger(\vec{p}, s)u(\vec{p}, s') &= \bar{v}^\dagger(\vec{p}, s)v(\vec{p}, s') = \frac{E(\vec{p})}{m}\delta_{ss'}, \\
\bar{v}(\vec{p}, s)u(\vec{p}, s') &= \bar{v}^\dagger(\vec{p}, s)v(\vec{p}, s') = 0.
\end{align*}
\]

(9.2.20)

The completeness relations are

\[
\sum_s [u_\alpha(\vec{p}, s)\bar{u}_\beta(\vec{p}, s) - v_\alpha(\vec{p}, s)v_\beta(\vec{p}, s)] = \delta_{\alpha\beta}.
\]

(9.2.21)
\[ [b(\vec{k}, r), b^\dagger(\vec{q}, s)]_\pm \]
\[ = \frac{1}{(2\pi)^3} \int d^3x \; d^3y \sqrt{\frac{m^2}{E(\vec{k})E(\vec{q})}} \; u^\dagger(\vec{k}, r) e^{i k z} [\psi(x), \psi^\dagger(y)]_\pm \big|_{x^0 = y^0} u(\vec{q}, s) e^{-i q y} \]
\[ = \frac{1}{(2\pi)^3} \int d^3x \; \frac{m}{\sqrt{E(\vec{k})E(\vec{q})}} e^{i (\vec{k} - \vec{q}).x} u^\dagger(\vec{k}, r) u(\vec{q}, s) \]
\[ = \frac{m}{E(\vec{k})} u^\dagger(\vec{k}, r) u(\vec{k}, s) \delta(\vec{k} - \vec{q}) \]
\[ = \delta_{rs} \delta(\vec{k} - \vec{q}) . \quad (9.2.26) \]

Similarly,
\[ [d(\vec{k}, r), d^\dagger(\vec{q}, r)]_\pm = \delta_{rs} \delta(\vec{k} - \vec{q}) . \quad (9.2.27) \]

All other commutators or anticommutators, as the case may be, vanish. We next settle the case as to whether we use commutators or anticommutators.

### 9.3 Positive Energy

We have the expression for the hamiltonian.
\[ H = \int d^3 x \psi^\dagger(x) [-i \vec{\alpha} \cdot \vec{\nabla} + \beta m] \psi(x) \quad (9.3.28) \]

Now, we substitute the expansions for \( \psi \) and \( \psi^\dagger \) and use the properties of the spinor solutions
\[ [-i \vec{\alpha} \cdot \vec{\nabla} + \beta m] u(\vec{p}, s) e^{-ipx} = E(\vec{p}) u(\vec{p}, s) e^{-ipx} \]
\[ [-i \vec{\alpha} \cdot \vec{\nabla} + \beta m] v(\vec{p}, s) e^{ipx} = -E(\vec{p}) v(\vec{p}, s) e^{ipx} . \quad (9.3.29) \]

Then,
\[ H = \int d^3 x \frac{d^3 p}{(2\pi)^3} \frac{d^3 k}{(2\pi)^3} \frac{m}{\sqrt{E(\vec{p})E(\vec{k})}} \]
\[ \times \sum_{rs} \left[ b^\dagger(\vec{p}, r) u^\dagger(\vec{p}, r) e^{ipx} + d(\vec{p}, r) u^\dagger(\vec{p}, r) e^{-ipx} \right] \]
\[ \times E(\vec{k}) \left[ b(\vec{k}, s) u(\vec{k}, s) e^{-ikx} + d^\dagger(\vec{k}, s) v(\vec{k}, s) e^{ikx} \right] \]
\[ = \frac{1}{(2\pi)^3} \int d^3 p \; d^3 k \; \frac{m}{\sqrt{E(\vec{p})E(\vec{k})}} \]
\[ \times \sum_{rs} \int d^3 x \left[ E(\vec{k}) b^\dagger(\vec{p}, r) b(\vec{k}, s) u^\dagger(\vec{p}, r) u(\vec{k}, s) e^{i(p-k)x} \right] \]
\[-E(\vec{k})d^{\dagger}(\vec{p}, r)d^\dagger(\vec{k}, s)v^\dagger(\vec{p}, r)v(\vec{k}, s) e^{-i(p-k)\cdot x} \]
\[+E(\vec{k})d(\vec{p}, r)b(\vec{k}, s)v^\dagger(\vec{p}, r)u(\vec{k}, s) e^{-i(p+k)\cdot x} \]
\[-E(\vec{k})b^\dagger(\vec{p}, r)d^\dagger(\vec{k}, s)u^\dagger(\vec{p}, r)v(\vec{k}, s) e^{i(p+k)\cdot x} \]

\[= \int d^3p \frac{m}{E(\vec{p})} \sum_s E(\vec{p}) \left[ b^\dagger(\vec{p}, s)b(\vec{p}, s) - d(\vec{p}, s)d^\dagger(\vec{p}, s) \right] \frac{E(\vec{p})}{m} \]
\[+ \int d^3p \frac{m}{E(\vec{p})} \sum_r E(\vec{p}) \left[ d(\vec{p}, r)b(\vec{p}, s)v^\dagger(\vec{p}, r)u(-\vec{p}, s) e^{-2i\vec{p} \cdot x^0} \right. \]
\[\left. - b^\dagger(\vec{p}, r)d^\dagger(\vec{p}, s)u^\dagger(\vec{p}, r)v(-\vec{p}, s) e^{2i\vec{p} \cdot x^0} \right] \]
\[= \int d^3p E(\vec{p}) \sum_s \left[ b^\dagger(\vec{p}, s)b(\vec{p}, s) - d(\vec{p}, s)d^\dagger(\vec{p}, s) \right] . \quad (9.3.30)\]

It is at this stage that we are forced to choose anticommutation relations since if we use commutators the energy is indefinite. However, if we choose anticommutators then

\[d(\vec{k}, s)d^\dagger(\vec{p}, r) = -d^\dagger(\vec{p}, r)d(\vec{k}, s) + \delta_{rs}\delta(\vec{k} - \vec{p}) \quad (9.3.31)\]

and

\[:d(\vec{k}, s)d^\dagger(\vec{p}, r): = -d^\dagger(\vec{p}, r)d(\vec{k}, s) \quad (9.3.32)\]

so that

\[H = \int d^3p E(\vec{p}) \sum_s \left[ b^\dagger(\vec{p}, s)b(\vec{p}, s) + d^\dagger(\vec{p}, s)d(\vec{p}, s) \right] \geq 0 . \quad (9.3.33)\]

Thus, for Dirac fields we use anticommutation relations. The necessity for this can proven in general and is the famous Spin and Statistics Theorem [1] of quantum field theory that states that integer spin fields satisfy Bose statistics and half odd integer spin fields satisfy Fermi-Dirac statistics.

### 9.4 Green’s Functions for Dirac Equation

We are looking for a matrix function \(S_{\alpha\beta}(x - y)\) that satisfies

\[\sum_{\rho} (i\gamma \cdot \partial_x - m)_{\alpha\rho} S_{\rho\beta}(x - y) = \delta(x - y)\delta_{\alpha\beta} \quad (9.4.34)\]

If we suppress the spinor indices, this equation reads

\[(i\gamma \cdot \partial_x - m) S(x - y) = \delta(x - y) \quad (9.4.35)\]

To solve this equation we set

\[S(x - y) = -(i\gamma \cdot \partial_x + m) F(x - y) \quad (9.4.36)\]
Then, we find that $F$ satisfies

$$\left(\Box_x + m^2\right) F(x - y) = \delta(x - y) \quad (9.4.37)$$

This means that $F$ is just one of the Klein-Gordon Green's functions. In other words it is one of $\Delta_R$ or $\Delta_A$ and of course we also have that $F = \Delta_A - \Delta_R$ satisfies the homogeneous Klein-Gordon equation. Thus, we have

$$S_R(x - y) = -(i\gamma \cdot \partial_x + m) \Delta_R(x - y)$$
$$S_A(x - y) = -(i\gamma \cdot \partial_x + m) \Delta_A(x - y) \quad (9.4.38)$$

and

$$S(x - y) = S_A(x - y) - S_R(x - y) = -(i\gamma \cdot \partial_x + m) \Delta(x - y) \quad (9.4.39)$$

It is now straightforward algebra to check that, from the anticommutation relations for the annihilation and creation operators, we get that (see problem 9.2)

$$[\psi_\alpha(x), \bar{\psi}_\beta(y)]_+ = -iS_{\alpha\beta}(x - y) \quad (9.4.40)$$

It is also worthwhile to note that

$$[\psi_\alpha(x), \bar{\psi}_\beta(y)]_+ |_{x^0 = y^0} = -i\gamma^0_{\alpha\beta} \partial_0 \Delta(x - y) |_{x^0 = y^0}$$
$$= \gamma^0_{\alpha\beta} \delta(\vec{x} - \vec{y}) \quad (9.4.41)$$

This immediately implies that

$$[\psi_\alpha(x), \psi^*_\beta(y)]_+ |_{x^0 = y^0} = \delta_{\alpha\beta} \delta(\vec{x} - \vec{y}) \geq 0 \quad (9.4.42)$$

as required.

Next, we write out the Fourier transforms of these Green's functions. To do this we note that

$$S_R(x) = -(i\gamma \cdot \partial + m)\Delta_R(x)$$
$$= (i\gamma \cdot \partial + m) \frac{1}{(2\pi)^4} \int d^4k \frac{e^{-ikx}}{(k^0 + i\epsilon)^2 - \vec{k}^2 - m^2}$$
$$= \frac{1}{(2\pi)^4} \int d^4k \frac{(\gamma \cdot k + m)e^{-ikx}}{(k^0 + i\epsilon)^2 - \vec{k}^2 - m^2} \quad (9.4.43)$$

For $S_A$ one simply replaces $+i\epsilon$ by $-i\epsilon$. It is also worth noting that

$$\frac{1}{(k^0 \pm i\epsilon)^2 - \vec{k}^2 - m^2} = \frac{1}{k^2 - m^2 \pm i\epsilon k^0}$$
$$= \frac{1}{k^2 - m^2} \mp i\epsilon(k^0)\delta(k^2 - m^2) \quad (9.4.44)$$

where, as always,

$$\epsilon(k^0) = \begin{cases} 1 & k^0 > 0 \\ -1 & k^0 < 0 \end{cases} \quad (9.4.45)$$
Also,
\[
\frac{(\gamma.k + m)}{k^2 - m^2} = \frac{(\gamma.k + m)}{(\gamma.k + m)(\gamma.k - m)} = \frac{1}{\gamma.k - m}.
\]  
(9.4.46)

We leave it as an exercise (problem 9.1) to show that
\[
\gamma^0 S_R^\dagger(x - y) \gamma^0 = S_A(x - y)
\]  
(9.4.47)

With these preliminaries out of the way we are ready to consider some interactions.

### 9.5 Perturbation of Electromagnetic Interaction

As a first step we consider the interaction of a Dirac field with a given external electromagnetic field. The equation of motion for minimal coupling now reads
\[
(i\gamma.\partial - m)\psi = e\gamma.A\psi
\]  
(9.5.48)

This can be converted into integral equations (Källén-Yang-Feldman equations) using the Green's functions developed above.

\[
\psi(x) = \psi_{in}(x) + e \int S_{R}(x - y)\gamma \cdot A(y) \psi(y) d^4 y
\]

\[
\psi(x) = \psi_{out}(x) + e \int S_{A}(x - y)\gamma \cdot A(y) \psi(y) d^4 y.
\]  
(9.5.49)

To solve these equations by perturbation theory we expand in powers of the dimensionless coupling constant \(e\) which, with \(\hbar\) and \(c\) written out, stands for \(1/\sqrt{\hbar c} \approx 1/\sqrt{137}\).

\[
\psi(x) = \psi^{(0)}(x) + e\psi^{(1)}(x) + e^2\psi^{(2)}(x) + \cdots
\]  
(9.5.50)

Then,
\[
\psi^{(0)}(x) = \psi_{in}(x)
\]  
(9.5.51)

and
\[
\psi^{(1)}(x) = \int S_{R}(x - y)\gamma \cdot A(y) \psi_{in}(y) d^4 y
\]

\[
\psi^{(1)}(x) = \int S_{A}(x - y)\gamma \cdot A(y) S_{A}(x - y) d^4 y
\]  
(9.5.52)

etc.

For these expressions to be meaningful requires that all integrals here converge. This means that for a static vector potential \(A(y)\) one should include a
regularizing, or convergence factor \( \exp(-\epsilon |y^0|) \) and at the end of the calculation let \( \epsilon \to 0 \). This "regularizing" procedure then yields expressions of the form

\[
\frac{n\epsilon}{(n\epsilon)^2 + p_0^2} \to \pi \delta(p_0)
\]  

(9.5.53)

as well as

\[
\frac{1}{n\epsilon + ip_0} \to -iP \frac{1}{p_0} + \pi \delta(p_0)
\]  

(9.5.54)

If we do not use such a regularization we encounter terms like

\[
\delta(p_0)\delta(p_0) = \delta(0)\delta(p_0)
\]

which require an interpretation to make them meaningful. We shall follow this more formal approach and make the appropriate interpretation when the time arises.

### 9.6 The S-Operator

We are looking for an operator \( S \) such that

\[
\psi_{out} = S\psi_{in} S^\dagger .
\]  

(9.6.55)

In keeping with the perturbation approach we expand \( S \) in a power series of \( \epsilon = \sqrt{\epsilon^2/(\hbar c)} \).

\[
S = 1 + \epsilon S^{(1)} + \epsilon^2 S^{(2)} + \ldots
\]  

(9.6.56)

Since \( S \) is unitary we have

\[
SS^\dagger = S^\dagger S = 1
\]  

(9.6.57)

To lowest order in \( \epsilon \) we therefore find

\[
(1 + \epsilon S^{(1)})(1 + \epsilon S^{(1)}) = (1 + \epsilon S^{(1)})(1 + \epsilon S^{(1)}\dagger) = 1
\]  

(9.6.58)

so that

\[
\epsilon (S^{(1)}\dagger + S^{(1)}) = 0
\]  

(9.6.59)

This means that we need

\[
S^{(1)}\dagger = -S^{(1)}
\]  

(9.6.60)

But,

\[
\psi_{out} = \psi_{in} - \epsilon \int S(x - y) A(y) \psi(y) d^4 y
\]  

(9.6.61)
and in lowest order, or Born, approximation we have

\[
\psi_{out} = \psi_{in} - e \int S(x - y) A(y) \psi_{in}(y) \, d^4y
\]
\[
= S \psi_{in} S^\dagger
\]
\[
= \psi_{in} - e \left[ S^{(1)\dagger} \psi_{in} - \psi_{in} S^{(1)\dagger} \right].
\] (9.6.62)

Thus, to lowest order in \( e \), we have

\[
S^{(1)\dagger} \psi_{in} - \psi_{in} S^{(1)\dagger} = \int S(x - y) A(y) \psi_{in}(y) \, d^4y.
\] (9.6.63)

So,

\[
[S^{(1)\dagger}, \psi_{in}] = \int S(x - y) A(y) \psi_{in}(y) \, d^4y.
\] (9.6.64)

Hence we get that

\[
S^{(1)\dagger} = -i \int : \bar{\psi}_{in}(z) A(z) \psi_{in}(z) : \, d^4z.
\] (9.6.65)

To verify this we simply compute the commutator, but for the moment we also drop the normal ordering.

\[
[S^{(1)\dagger}, \psi_{in} \gamma(z)]
\]
\[
= -i \int \psi_{in} \alpha(z) A_{\alpha\beta}(z) \psi_{in} \beta(z) \psi_{in} \gamma(z) \, d^4z
\]
\[
+ i \int \psi_{in} \gamma(z) \bar{\psi}_{in} \alpha(z) A_{\alpha\beta}(z) \psi_{in} \beta(z) \, d^4z
\]
\[
= i \int \{ \bar{\psi}_{in} \alpha(z) \psi_{in} \gamma(z) + \psi_{in} \gamma(z) \bar{\psi}_{in} \alpha(z) \} A_{\alpha\beta}(z) \psi_{in} \beta(z) \, d^4z
\]
\[
= \int S_{\gamma\alpha}(x - z) A_{\alpha\beta}(z) \psi_{in} \beta(z) \, d^4z
\]
\[
= \int (S(x - z) A(z) \psi_{in}(z))_\gamma \, d^4z
\] (9.6.66)

as required. Since \( S^{(1)\dagger} \) is defined by a commutator, it is specified only up to an irrelevant additive c-number which plays no role in the commutator. The irrelevance of the c-number comes about from the fact that unitarity requires that

\[
S^{(1)\dagger} = -S^{(1)}
\] (9.6.67)

so that the c-number has to be pure imaginary. But if we think of the entire \( S \)-operator then this c-number is just part of a phase factor \( e^{i\delta} \), \( \delta \) real) in the \( S \)-operator. So setting \( \delta = 0 \) is the same as setting the c-number = 0. This is simply a choice of the overall phase factor of the \( S \)-operator.
9.7 Scattering of an Electron by an External EM Field

We now apply the result obtained in the previous section. Let the initial state of an electron be

\[ |\tilde{q}s\rangle = b^\dagger (\tilde{q}, s)|\Omega\rangle \quad (9.7.68) \]

and the final state be

\[ |\tilde{p}r\rangle = b^\dagger (\tilde{p}, r)|\Omega\rangle . \quad (9.7.69) \]

We are interested in computing to lowest order

\[ \langle \tilde{p}r|S|\tilde{q}s\rangle = \delta(\tilde{p} - \tilde{q})\delta_{rs} + e\langle \tilde{p}r|S^{(1)}|\tilde{q}s\rangle \quad (9.7.70) \]

where, using the results obtained above, we have

\[ \langle \tilde{p}r|S^{(1)}|\tilde{q}s\rangle = i \int d^4x \langle \tilde{p}r|\tilde{\psi}_{in}(x)|\Omega\rangle \gamma^\mu \langle \Omega|\psi_{in}(x)|\tilde{q}s\rangle A_\mu (x) . \quad (9.7.71) \]

Also,

\[
\langle \Omega|\psi_{in}(x)|\tilde{q}s\rangle = \sum_{s'} \int \frac{d^3k}{(2\pi)^{3/2}} \sqrt{\frac{m}{E(k)}} \langle \Omega|b(k, s')|\tilde{q}s\rangle u(k, s') e^{-ikx} = \sqrt{\frac{m}{E(q)}} u(\tilde{q}, s) e^{-iqx} . \quad (9.7.72)
\]

Thus,

\[
e\langle \tilde{p}r|S^{(1)}|\tilde{q}s\rangle = ie \frac{m}{\sqrt{E(\tilde{p})E(\tilde{q})}} \frac{1}{(2\pi)^3} \int d^4x e^{i(x-q)x} \bar{u}(\tilde{p}r)\gamma^\mu u(\tilde{q}s) A_\mu (x) = ie \frac{m}{\sqrt{E(\tilde{p})E(\tilde{q})}} \frac{1}{(2\pi)^3} \bar{u}(\tilde{p}r)\gamma^\mu u(\tilde{q}s) \tilde{A}_\mu (p - q) . \quad (9.7.73)
\]

This is as far as one can go in general in obtaining the scattering amplitude. In the next section we specialize to an external Coulomb field to obtain the relativistic generalization of the Rutherford cross-section.

9.8 Cross-Section for Coulomb Scattering

We use the results of the previous section and specify the vector potential to be that corresponding to a Coulomb potential.

\[ A_0 = \frac{Ze}{r} \quad \tilde{A} = 0 . \quad (9.8.74) \]
Then,
\[
\tilde{A}(p - q) = \int d^4x \, e^{i(p - q) \cdot x} \, A_0(x) = 2\pi \delta(p^0 - q^0) \int d^3r \, e^{-i(p - q) \cdot r} \frac{Ze}{r} \\
= \frac{Ze}{|p - q|^2} 8\pi^2 \delta(p^0 - q^0). \quad (9.8.75)
\]
Thus,
\[
e\langle \vec{p}r | S^{(1)} | \vec{q}s \rangle = \frac{Ze^2}{\hbar c \pi} \sqrt{\frac{m^2}{E(\vec{p})E(\vec{q})}} \frac{1}{|p - q|^2} \gamma^0 u(\vec{q}s) \delta(E(\vec{p}) - E(\vec{q})). \quad (9.8.76)
\]
The square of the matrix element is given by
\[
\left| \langle \vec{p}r | eS^{(1)} | \vec{q}s \rangle \right|^2 = \frac{Z^2e^4}{\hbar^2c^2 \pi^2} \frac{m^2}{E(\vec{p})E(\vec{q})} \delta(E(\vec{p}) - E(\vec{q})) \delta(0) \left| \bar{u}(\vec{p})\gamma^0 u(\vec{q}s) \right|^2. \quad (9.8.77)
\]
The factor \(\delta(E(\vec{p}) - E(\vec{q}))\delta(0)\) is to be interpreted as
\[
\delta(E(\vec{p}) - E(\vec{q})) \delta(0) = \delta(E(\vec{p}) - E(\vec{q})) \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T/2}^{T/2} e^{i(E(\vec{p}) - E(\vec{q}))t} dt \\
= \delta(E(\vec{p}) - E(\vec{q})) \frac{1}{2\pi} \int_{-T/2}^{T/2} dt \\
= \delta(E(\vec{p}) - E(\vec{q})) \frac{T}{2\pi}. \quad (9.8.78)
\]
Here, \(T\) is typical of the time that the interaction is on.

On the other hand, if we had regularized and written
\[
A_0 = \frac{Ze}{r} e^{-\epsilon|t|} \quad (9.8.79)
\]
we would have found that
\[
\tilde{A}_0(p - q) = 4\pi \frac{Ze}{|p - q|^2} \frac{2\epsilon}{\epsilon^2 + (p_0 - q_0)^2}. \quad (9.8.80)
\]
Now,
\[
\left| \frac{2\epsilon}{\epsilon^2 + (p_0 - q_0)^2} \right|^2 = \frac{2\epsilon}{\epsilon^2 + (p_0 - q_0)^2} \cdot \frac{2\epsilon}{\epsilon^2 + (p_0 - q_0)^2} = 2\pi\delta(p_0 - q_0) \frac{2}{\epsilon} = 2\pi\delta(p_0 - q_0) \frac{T}{2\pi} \quad (9.8.81)
\]
where we have identified the typical time that the interaction is on by
\[ \frac{2}{\epsilon} = \frac{T}{2\pi} . \tag{9.8.82} \]

The transition probability per unit time is now given by dividing by the time \( T \) that the interaction is on and is
\[ \frac{dw}{dt} = \frac{Z^2 e^4}{2\pi^3} \left( \frac{m}{E(p)} \right)^2 \left| \bar{u}(\vec{p}r) \frac{1}{|\vec{p} - \vec{q}|^2} \gamma^0 u(\vec{q}s) \right|^2 \delta(E(\vec{p}) - E(\vec{q})) . \tag{9.8.83} \]

If the initial beam was unpolarized and our detectors do not detect the spin and, if furthermore, our detectors can only measure energy to withing a certain range then we are interested only in transitions to a final group of states with an energy density of final states
\[ \rho_f = \frac{dn_f}{dE} . \tag{9.8.84} \]

In this case we must
1. Sum over the final spin states.
2. Average over the initial spin states.
3. Multiply the result by \( \rho_f \).

Thus, we get
\[ \frac{dw}{dt} = \frac{4Z^2 e^4}{(2\pi)^3} \left( \frac{m}{E(\vec{p})} \right)^2 \frac{1}{|\vec{p} - \vec{q}|^2} \frac{1}{2} \left( \frac{\not{p} + m}{2m} \gamma^0 \frac{\not{q} + m}{2m} \gamma^0 \right) \rho_f . \tag{9.8.85} \]

Now,
\[ \frac{1}{2} \left( \frac{\not{p} + m}{2m} \gamma^0 \frac{\not{q} + m}{2m} \gamma^0 \right) \]
\[ = \frac{1}{8m^2} \left( \not{p} + m \gamma^0 (\not{q} + m) \gamma^0 \right) \]
\[ = \frac{1}{8m^2} \left( \not{p} \gamma^0 \not{q} \gamma^0 + m^2 \right) \]
\[ = \frac{1}{2m^2} \left( 2p^0 q^0 - p.q + m^2 \right) \]
\[ = \frac{1}{2m^2} \left( E^2 + \vec{p} \cdot \vec{q} + m^2 \right) . \tag{9.8.86} \]

Here we have used the fact that, according to 9.8.83, we have
\[ E = E(\vec{p}) = E(\vec{q}) \tag{9.8.87} \]

which implies
\[ |\vec{p}| = |\vec{q}| = p . \tag{9.8.88} \]

Then, we see that
\[ \vec{p} \cdot \vec{q} = p^2 \cos \theta = (E^2 - m^2) \cos \theta . \tag{9.8.89} \]
Therefore,
\[
\frac{1}{8m^2} \text{Tr} \left( \gamma^0 p \gamma^0 + m^2 \right) = \frac{1}{m^2} \left[ \frac{1 + \cos \theta}{2} (E^2 - m^2) + m^2 \right].
\]  \hfill (9.8.90)

But,
\[
E = \frac{m}{\sqrt{1 - v^2}}.
\]  \hfill (9.8.91)

So,
\[
\frac{m^2}{E^2} = 1 - v^2 \quad \text{or} \quad 1 - \frac{m^2}{E^2} = v^2.
\]  \hfill (9.8.92)

Thus,
\[
\frac{1}{8m^2} \text{Tr} \left( \gamma^0 \gamma^0 + m^2 \right) = \frac{E^2}{m^2} \left[ 1 - v^2 \sin^2 \theta/2 \right].
\]  \hfill (9.8.93)

Also,
\[
|\vec{p} - \vec{q}|^2 = p^2 + q^2 - 2\vec{p} \cdot \vec{q} = 2p^2(1 - \cos \theta) = 4p^2 \sin^2 \theta/2.
\]  \hfill (9.8.94)

To get the density of final states \(\rho_f\) we notice that on the single particle states
\[
\int d\Omega \left( b^\dagger (\vec{p}s) |\Omega\rangle \langle \Omega | b(\vec{q}s) = 1 \right.
\]  \hfill (9.8.95)

This means that \(dn_f = dp d\Omega\) and therefore that
\[
\rho_f = \frac{dn_f}{dE} = \frac{p^2 dp d\Omega}{dE}.
\]  \hfill (9.8.96)

Now using that \(p^2 + m^2 = E^2\) we find that \(p\,dp = E\,dE\) and therefore,
\[
\rho_f = Ep\,d\Omega.
\]  \hfill (9.8.97)

Thus,
\[
\frac{dw}{dt} = \frac{4Z^2e^4}{(2\pi)^3} \frac{1 - v^2 \sin^2 \theta/2}{16p^4 \sin^2 \theta/2} E p d\Omega.
\]  \hfill (9.8.98)

To get the differential cross-section we have to divide this result by the incident flux which may be obtained from the current density \(j_\mu(x)\) by taking its expectation value in the initial state \(|\vec{q}s\rangle = b^\dagger (\vec{q}s) |\Omega\rangle\). So, we compute
\[
\langle j_\mu(x) \rangle = -e\langle \Omega | b(\vec{q}s) : \bar{\psi}(x) \gamma_\mu \psi(x) : b^\dagger (\vec{q}s) |\Omega\rangle.
\]  \hfill (9.8.99)

As usual we put in intermediate states. However, we also notice that
\[
: \bar{\psi} \psi : \sim (b^\dagger + d)(b + d^\dagger) : \sim b^\dagger b - d^\dagger d + b^\dagger d^\dagger - bd.
\]  \hfill (9.8.100)

Then,
\[
\langle \vec{q}s | b^\dagger b - d^\dagger d + b^\dagger d^\dagger - bd |\vec{q}s\rangle = \langle \vec{q}s | b^\dagger b |\vec{q}s\rangle.
\]  \hfill (9.8.101)
Thus,

\[
\langle j_\mu(x) \rangle = -e \langle \Omega | b(\bar{q}s) \bar{\psi}^{(-)}(x) \gamma_\mu \psi^{(+)}(x) b^\dagger(\bar{q}s) | \Omega \rangle \\
= - \frac{e}{2\pi^3} \frac{mc^2}{E(q)} \bar{u}(\bar{q}s) \gamma_\mu u(qs) \\
= - \frac{e}{2\pi^3} \frac{mc^2}{E(q)} \frac{q_\mu}{mc^2} \\
= - \frac{e}{2\pi^3} \frac{q_\mu}{E(q)}. \tag{9.8.102}
\]

This means that the incoming flux is given by

\[
\frac{1}{(2\pi)^3} \frac{|q|}{E(q)} = \frac{1}{(2\pi)^3} \frac{p}{E}.
\]

Thus, finally

\[
\frac{d\sigma}{d\Omega} = \frac{Z^2e^4}{4p^2\sin^4 \theta/2} \left(1 - \frac{v^2}{c^2} \sin^2 \frac{\theta}{2} \right) \frac{E^2}{p^2} \\
= \frac{Z^2e^4}{4p^2v^2\sin^4 \theta/2} \left(1 - \frac{v^2}{c^2} \sin^2 \frac{\theta}{2} \right). \tag{9.8.103}
\]

This formula is known as the Mott cross-section and up to the factor

\[
\left(1 - \frac{v^2}{c^2} \sin^2 \frac{\theta}{2} \right)
\]

is identical with the Rutherford formula. This factor is due to the spin and comes from the trace calculation. Also for nonrelativistic energies we note that

\[
p^2v^2 = (mv^2)^2 = 4E^2. \tag{9.8.104}
\]

The lowest order perturbation term for Coulomb scattering or any scattering of electrons by an external electromagnetic field is pictorially represented by the Feynman diagram depicted in figure 9.1. As we shall see in section 11.8, this diagram gives an immediate mnemonic device for computing the corresponding scattering amplitude.

Figure 9.1: Electron Scattering by an External Field in Born Approximation
9.9 Pair Production in an External Field

Again we only consider the case of weak external electromagnetic field so that we may use perturbation theory. Then, to lowest order in perturbation theory this process is represented by the Feynman diagram in 9.2. In this case our initial state is the vacuum state $|\Omega\rangle$ and our final state is an electron-positron pair

$$|\tilde{p}\tilde{r}; \tilde{q}s\rangle = b^\dagger(\tilde{p}\tilde{r})d^\dagger(\tilde{q}s)|\Omega\rangle.$$  
(9.9.105)

The matrix element of interest is then

$$M = \langle\tilde{p}\tilde{r}; \tilde{q}s|S|\Omega\rangle.$$  
(9.9.106)

We have already computed $S$ to first order in $\epsilon$ and found

$$S = 1 + \epsilon S^{(1)}$$  
(9.9.107)

where

$$\epsilon S^{(1)} = i\epsilon \int :\bar{\psi}_{in}(x)\gamma^\mu\psi_{in}(x) : A_\mu(x) d^4x.$$  
(9.9.108)

So,

$$M = i\epsilon \int d^4x A_\mu(x)\langle\Omega|b(\tilde{p}\tilde{r})d(\tilde{q}s) : \bar{\psi}_{in}(x)\gamma^\mu\psi_{in}(x) : |\Omega\rangle$$

$$= \frac{i\epsilon}{(2\pi)^3} \int d^4x A_\mu(x)\sqrt{\frac{m^2}{E(\tilde{p})E(\tilde{q})}} \bar{u}(\tilde{p}\tilde{r})e^{ipx} \gamma^\mu v(\tilde{q}s)e^{iqx}$$

$$= \frac{i\epsilon}{(2\pi)^3} \sqrt{\frac{m^2}{E(\tilde{p})E(\tilde{q})}}\bar{u}(\tilde{p}\tilde{r})\gamma^\mu v(\tilde{q}s)\tilde{A}_\mu(p + q).$$  
(9.9.109)

Thus, we get that the square of the matrix element is given by

$$|M|^2 = \frac{\epsilon^2}{(2\pi)^6} \frac{m^2}{E(\tilde{p})E(\tilde{q})} \left|\bar{u}(\tilde{p}\tilde{r})\gamma^\mu v(\tilde{q}s)\tilde{A}_\mu(p + q)\right|^2.$$  
(9.9.110)
Now, the probability that a pair is produced with momenta between \( p \) and \( p + dp \) as well as \( k \) and \( k + dk \) and corresponding spins \( r, s \) is
\[
dw_{rs} = |M|_{rs}^2 d^3p d^3k .
\]
(9.9.111)
If the detectors are insensitive to the spin then we have to sum over both \( r \) and \( s \) to get
\[
dw = \sum_{rs} |M|_{rs}^2 d^3p d^3k .
\]
(9.9.112)
Writing this out we find
\[
dw = \frac{e^2}{(2\pi)^6} \frac{m^2}{E(\vec{p})E(\vec{k})} \tilde{A}_\mu(p + k) \tilde{A}_\nu^*(p + k) \text{Tr} \left[ \frac{m + \not{p}}{2m} \gamma^\mu \frac{m - \not{k}}{2m} \gamma^\nu \right] d^3p d^3k .
\]
(9.9.113)
But,
\[
\frac{1}{4m^2} \text{Tr} \left[ (m + \not{p}) \gamma^\mu (m - \not{k}) \gamma^\nu \right] = \frac{1}{4m^2} \text{Tr} \left[ m^2 \gamma^\mu \gamma^\nu - \not{p} \gamma^\mu \not{k} \gamma^\nu \right] = \frac{1}{m^2} \left[ (m^2 + p \cdot k) g^{\mu \nu} - p^\mu k^\nu - k^\mu p^\nu \right]
\]  
(9.9.114)
The total probability for producing a pair is therefore
\[
w = \frac{e^2}{(2\pi)^6} \int \frac{d^3p}{E(\vec{p})} \frac{d^3k}{E(\vec{k})} \tilde{A}_\mu(p + k) \tilde{A}_\nu^*(p + k) \left[ (m^2 + p \cdot k) g^{\mu \nu} - p^\mu k^\nu - k^\mu p^\nu \right] .
\]
(9.9.115)
We rewrite this as follows
\[
w = \frac{e^2}{(2\pi)^6} \int d^4Q \delta(Q - p - k) \int \frac{d^3p}{E(\vec{p})} \frac{d^3k}{E(\vec{k})} \tilde{A}_\mu(p + k) \tilde{A}_\nu^*(p + k)
\times \left[ (m^2 + p \cdot k) g^{\mu \nu} - p^\mu k^\nu - k^\mu p^\nu \right]
= \frac{e^2}{(2\pi)^6} \int d^4Q T^{\mu \nu}(Q) \tilde{A}_\nu^*(Q)
\]
(9.9.116)
where
\[
T^{\mu \nu}(Q) = \int \frac{d^3p}{E(\vec{p})} \frac{d^3k}{E(\vec{k})} \delta(Q - p - k)
\times \left[ (m^2 + p \cdot k) g^{\mu \nu} - p^\mu k^\nu - k^\mu p^\nu \right]
= 4 \int d^4p \delta(p^2 - m^2) \theta(p^0) d^4k \delta(k^2 - m^2) \theta(k^0) \delta(Q - p - k)
\times \left[ (m^2 + p \cdot k) g^{\mu \nu} - p^\mu k^\nu - k^\mu p^\nu \right]
= 4 \int d^4p \delta(p^2 - m^2) \theta(p^0) \delta((Q - p)^2 - m^2) \theta(Q^0 - p^0)
\times \left[ p \cdot Q g^{\mu \nu} + 2p^\mu p^\nu - p^\mu Q^\nu - Q^\mu p^\nu \right].
\]
(9.9.117)
The fact that $T^{\mu\nu}(Q)$ is a Lorentz tensor limits its most general form to
\[ T^{\mu\nu}(Q) = A(Q^2)g^{\mu\nu} + B(Q^2)Q^\mu Q^\nu \] (9.9.118)
where $A(Q^2)$ and $B(Q^2)$ are invariant functions of their arguments. By direct computation we find that
\[ Q_\mu T^{\mu\nu} = 4 \int d^4p \delta(p^2 - m^2)\theta(p^0)\delta((Q - p)^2 - m^2)\theta(Q^0 - p^0) \times [-Q^2p^\nu + 2p \cdot Q p^\nu] . \] (9.9.119)
But,
\[ p \cdot Q = \frac{1}{2} [Q^2 + p^2 - (Q - p)^2] = \frac{1}{2} Q^2 \] (9.9.120)
since by virtue of the delta functions
\[ p^2 = (Q - p)^2 = m^2 . \] (9.9.121)
Therefore, the integral vanishes and $T^{\mu\nu}$ is conserved. This also means that applying this result to the general form of $T^{\mu\nu}$ we get
\[ A(Q^2)Q^\nu + Q^2 B(Q^2)Q^\nu = 0 \] (9.9.122)
or
\[ A(Q^2) = -Q^2 B(Q^2) . \] (9.9.123)
Therefore,
\[ T^{\mu\nu} = B(Q^2) [Q^\mu Q^\nu - Q^2 g^{\mu\nu}] . \] (9.9.124)
But,
\[ T^{\mu}_\mu = B(Q^2) [Q^2 - 4Q^2] = -3Q^2 B(Q^2) . \] (9.9.125)
This means that we only need to evaluate the scalar function $B(Q^2)$, that is,
\[ B(Q^2) = -\frac{T^{\mu}_\mu}{3Q^2} . \] (9.9.126)
Writing this out we have
\[ B(Q^2) = -\frac{4}{3Q^2} \int d^4p \delta(p^2 - m^2)\theta(p^0) \delta((Q - p)^2 - m^2)\theta(Q^0 - p^0)2[m^2 + p \cdot Q] . \] (9.9.127)
The delta functions show that for $Q$ spacelike the integral vanishes. So, to evaluate the integral we choose $Q = 0$. Then, we get
\[ B(Q_0^2) = -\frac{8}{3Q_0^2} \int \frac{d^3p}{\sqrt{p^2 + m^2}} [m^2 + Q^0\sqrt{p^2 + m^2}] \times \delta(Q_0^2 - 2Q^0\sqrt{p^2 + m^2})\theta(Q^0 - p^0) \]
\[ = -\frac{32\pi}{3Q_0^2} \int_0^\infty \frac{d^3p}{\sqrt{p^2 + m^2}} [m^2 + Q^0\sqrt{p^2 + m^2}] \times \frac{\delta(Q_0^2/2 - \sqrt{p^2 + m^2})}{2|Q_0|} \theta(Q^0 - p^0) \] (9.9.128)
where
\[
\Pi^{(0)}(Q^2) = \frac{e^2}{12\pi^2} \left(1 + \frac{2m^2}{Q^2}\right) \sqrt{1 - \frac{4m^2}{Q^2} \theta(Q^2 - 4m^2) \theta(Q_0)} .
\] (9.9.138)

We again encounter this expression later on. For the moment we simply note that the function \(\theta(Q^2 - 4m^2)\) ensures that \(Q^2 > 4m^2\). This means that
\[
Q_0^2 > Q^2 + 4m^2 .
\] (9.9.139)

In other words, we need an energy in excess of \(2mc^2\) to produce an electron-positron pair by a single external "photon".

### 9.10 Vacuum Polarization

So far we have only looked at the S-matrix in simple cases in the Born approximation. Now we want, to the same order of approximation, to study the fields more closely. However, a quantity like the current is of higher order in perturbation theory and the corresponding term that yields vacuum polarization is represented by the Feynman diagram in figure 9.3. It is this quantity that we wish to study. The field equation (Heisenberg equation) for the Heisenberg field

\[
\psi(x)
\]

is
\[
(i\gamma \cdot \partial - m)\psi(x) = e A(x) \psi(x) .
\] (9.10.140)

Again to order \(e\) (first Born approximation) we have
\[
\psi(x) = \psi_{in}(x) + e \int d^4y \, S_R(x - y) A(y) \psi_{in}(y) .
\] (9.10.141)

In this case we want to compute the current corresponding to the field in this approximation. We saw that
\[
j^\mu(x) = e : \bar{\psi}(x) \gamma^\mu \psi(x) : .
\] (9.10.142)
There is another way to write this current that is more convenient.

If we consider the commutator \([\bar{\psi}(x), \gamma_\mu \psi(x)]\) and agree to mean by this

\[
\bar{\psi}_\alpha(x) \gamma^\mu_{\alpha\beta} \psi_\beta(x) - \psi_\beta(x) \gamma^\mu_{\alpha\beta} \bar{\psi}_\alpha(x)
\]

then we find that formally

\[
\frac{1}{2} [\bar{\psi}(x), \gamma_\mu \psi(x)] = \bar{\psi}(x) \gamma_\mu \psi(x) - \frac{1}{2} [\bar{\psi}(x), \gamma_\mu \psi(x)] +
\]

\[
= : \bar{\psi}(x) \gamma_\mu \psi(x) : + [\bar{\psi}^+(x), \gamma_\mu \psi^-(x)] + \frac{1}{2} [\bar{\psi}(x), \gamma_\mu \psi(x)] +
\]

\[
= : \bar{\psi}(x) \gamma_\mu \psi(x) : + 2g^{\mu 0} \delta(\vec{0}) - 2g^{0\mu} \delta(\vec{0})
\]

\[
= : \bar{\psi}(x) \gamma_\mu \psi(x) : .
\] (9.10.143)

This means that we can also write

\[
j^\mu(x) = \frac{e}{2} [\bar{\psi}(x), \gamma_\mu \psi(x)].
\] (9.10.144)

So, to Born approximation, we have

\[
j^\mu(x) = \frac{e}{2} [\bar{\psi}_{in}(x), \gamma_\mu \psi_{in}(x)] + \frac{e^2}{2} \int A_\nu(y) d^4 y \{ [\bar{\psi}_{in}(x), \gamma_\mu S_R(x - y) \gamma_\nu \psi_{in}(y)] + [\bar{\psi}_{in}(y) \gamma_\nu S_A(y - x) \gamma_\mu, \psi_{in}(x)] \} .
\] (9.10.145)

The interesting thing is that

\[
\langle \Omega | j^\mu(x) | \Omega \rangle = \int d^4 y K^{\mu\nu}(x - y) A_\nu(y) \neq 0 .
\] (9.10.146)

Here,

\[
K^{\mu\nu}(x - y) = \frac{e^2}{2} \langle \Omega | [\bar{\psi}_{in}(x), \gamma_\mu S_R(x - y) \gamma_\nu \psi_{in}(y)] | \Omega \rangle + \frac{e^2}{2} \langle \Omega | [\bar{\psi}_{in}(y) \gamma_\nu S_A(y - x) \gamma_\mu, \psi_{in}(x)] | \Omega \rangle .
\] (9.10.147)

The first term is

\[
K^{\mu\nu}_R(x - y) = \frac{e^2}{2} \langle \Omega | \bar{\psi}_{in}(x) \gamma_\mu S_R(x - y) \gamma_\nu \psi_{in}(y) | \Omega \rangle - \frac{e^2}{2} \langle \Omega | \gamma_\mu S_R(x - y) \gamma_\nu \psi_{in}(y) \bar{\psi}_{in}(x) | \Omega \rangle .
\] (9.10.148)

If we write out all the spinor indeces this become

\[
K^{\mu\nu}_R(x - y) = \frac{e^2}{2} \langle \Omega | \bar{\psi}_{in a}(x) \gamma^\mu_{\alpha\beta} S_R \beta\gamma(x - y) \gamma_\nu \delta_{\alpha\beta} \psi_{in}(y) | \Omega \rangle
\]
\[ -\frac{e^2}{2} \langle \Omega | \gamma^\mu_{\alpha \beta} S_{R \beta \gamma} (x - y) \gamma^\nu_{\gamma \delta} \psi_{in \delta} (y) \bar{\psi}_{in \alpha} (x) | \Omega \rangle \]
\[ = \frac{e^2}{2} \gamma^\mu_{\alpha \beta} S_{R \beta \gamma} (x - y) \gamma^\nu_{\gamma \delta} \times \]
\[ \times \left[ \langle \Omega | \psi_{in \alpha} (x) \psi_{in \delta} (y) | \Omega \rangle - \langle \Omega | \psi_{in \delta} (y) \bar{\psi}_{in \alpha} (x) | \Omega \rangle \right] \]
\[ = \frac{e^2}{2} \gamma^\mu_{\alpha \beta} S_{R \beta \gamma} (x - y) \gamma^\nu_{\gamma \delta} \sum_{\mathbf{r}, \mathbf{s}} \int \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \sqrt{\frac{m^2}{E(p)E(q)}} \]
\[ \left\{ \langle \Omega | \left[ b^\dagger (\vec{q}\vec{r}) u_\alpha (\vec{q}\vec{r}) e^{-ipx} + d(\vec{q}\vec{r}) \bar{v}_\alpha (\vec{q}\vec{r}) e^{-iqx} \right] \right\} \]
\[ \times \left[ b(\vec{p}\vec{s}) u_\delta (\vec{p}\vec{s}) e^{-ipy} + d^\dagger (\vec{p}\vec{s}) \bar{v}_\delta (\vec{p}\vec{s}) e^{ipy} \right] | \Omega \rangle \]
\[ - \langle \Omega | \left[ b(\vec{p}\vec{s}) u_\delta (\vec{p}\vec{s}) e^{-ipy} + d^\dagger (\vec{p}\vec{s}) \bar{v}_\delta (\vec{p}\vec{s}) e^{ipy} \right] \]
\[ \times \left[ b^\dagger (\vec{q}\vec{r}) u_\alpha (\vec{q}\vec{r}) e^{-iqx} + d(\vec{q}\vec{r}) \bar{v}_\alpha (\vec{q}\vec{r}) e^{-iqx} \right] | \Omega \rangle \right\} \]
\[ = \frac{e^2}{2} \gamma^\mu_{\alpha \beta} S_{R \beta \gamma} (x - y) \gamma^\nu_{\gamma \delta} \sum_{\mathbf{r}, \mathbf{s}} \int \frac{d^3p}{(2\pi)^3} \frac{m}{E(p)} \left\{ e^{-ip(x-y)} \bar{v}_\alpha (\vec{p}\vec{s}) v_\delta (\vec{p}\vec{s}) \right\} \]
\[ - e^{-ip(x-y)} \bar{u}_\alpha (\vec{p}\vec{s}) u_\delta (\vec{p}\vec{s}) \right\} \]
\[ = \frac{e^2}{2} \gamma^\mu_{\alpha \beta} S_{R \beta \gamma} (x - y) \gamma^\nu_{\gamma \delta} \int \frac{d^3p}{(2\pi)^3} \frac{m}{E(p)} \left\{ \left( \frac{m + \not{p}}{2m} \right) \alpha_\delta e^{ip(x-y)} \right\} \]
\[ + \left( \frac{m - \not{p}}{2m} \right) \alpha_\delta e^{-ip(x-y)} \right\}. \] (9.10.149)

The integral that occurs in this expression is
\[ S^{(1)} (x - y) = -\frac{1}{(2\pi)^3} \int \frac{d^3p}{E(p)} \left\{ (m + \not{p}) e^{ip(x-y)} + (m - \not{p}) e^{-ip(x-y)} \right\} \]
\[ = -(i\gamma_\tau \partial_\tau + m) \Delta^{(1)} (x - y) \] (9.10.150)

where
\[ \Delta^{(1)} (x - y) = \frac{1}{2(2\pi)^3} \int \frac{d^3p}{\omega(p)} \left\{ e^{-ip(x-y)} + e^{ip(x-y)} \right\} \]
\[ = \frac{1}{(2\pi)^3} \int d^4p e^{-ip(x-y)} \delta(p^2 - m^2). \] (9.10.151)

This equation shows that \( \Delta^{(1)} (x) = \Delta^{(1)} (-z) \) so that it is an even function. Therefore, we can also write
\[ S^{(1)} (x - y) = -(i\gamma_\tau \partial_\tau + m) \Delta^{(1)} (x - y) \]. (9.10.152)

So, we find that
\[ K^{\mu\nu} (x - y) = \frac{e^2}{2} \gamma^\mu_{\alpha \beta} S_{R \beta \gamma} (x - y) \gamma^\nu_{\gamma \delta} S^{(1)} (y - x) \]
\[ + \frac{e^2}{2} S^{(1)} (y - x) \gamma^\nu_{\alpha \beta} S_{A \beta \gamma} (x - y) \gamma^\mu_{\gamma \delta} \]
\[ \frac{e^2}{2} \left\{ \text{Tr} \left[ \gamma^\mu S_R(x - y) \gamma^\nu S^{(1)}(y - x) \right] + \text{Tr} \left[ \gamma^\mu S^{(1)}(x - y) \gamma^\nu S_A(y - x) \right] \right\}. \tag{9.10.153} \]

To study \( K^{\mu\nu} \) more closely we define its Fourier transform

\[ K^{\mu\nu}(x - y) = \frac{1}{(2\pi)^4} \int d^4 p \, e^{-ip(x - y)} \, \tilde{K}^{\mu\nu}(p). \tag{9.10.154} \]

Then,

\[ \tilde{K}^{\mu\nu}(p) = \int d^4 z \, e^{ipz} \, K^{\mu\nu}(z). \tag{9.10.155} \]

Or, writing it out

\[ \tilde{K}^{\mu\nu}(p) = \frac{e^2}{2} \int d^4 z \, e^{ipz} \left\{ \text{Tr} \left[ \gamma^\mu S_R(z) \gamma^\nu S^{(1)}(-z) \right] + \text{Tr} \left[ \gamma^\mu S^{(1)}(z) \gamma^\nu S_A(-z) \right] \right\}. \tag{9.10.156} \]

Now,

\[ S^{(1)}(z) = -(i\gamma.\partial + m) \frac{1}{(2\pi)^3} \int d^4 p \, \delta(p^2 - m^2) \, e^{-ipz} \tag{9.10.157} \]

whereas

\[ S_R(z) = -(i\gamma.\partial + m) \frac{1}{(2\pi)^4} \int d^4 k \, \frac{e^{-ikz}}{(k^0 + i\epsilon)^2 - \vec{k}^2 - m^2} \tag{9.10.158} \]

and

\[ S_A(z) = -(i\gamma.\partial + m) \frac{1}{(2\pi)^4} \int d^4 k \, \frac{e^{-ikz}}{(k^0 - i\epsilon)^2 - \vec{k}^2 - m^2}. \tag{9.10.159} \]

Therefore,

\[ \tilde{K}^{\mu\nu}(p) \left\{ \text{Tr} \left[ \gamma^\mu (\gamma.k + m) \gamma^\nu (\gamma.q + m) \right] \frac{1}{(k^0 + i\epsilon)^2 - \vec{k}^2 - m^2} \delta(k^2 - m^2) \right. \]

\[ + \left. \text{Tr} \left[ \gamma^\mu (\gamma.k + m) \gamma^\nu (\gamma.q + m) \right] \frac{1}{(q^0 - i\epsilon)^2 - \vec{q}^2 - m^2} \delta(q^2 - m^2) \right\} \]

\[ = \frac{e^2}{2(2\pi)^3} \int d^4 k \, d^4 q \, \delta(p - k + q) \left\{ \delta(q^2 - m^2) \left[ \frac{1}{k^2 - m^2} - i\pi\epsilon(k)\delta(k^2 - m^2) \right] \right. \]

\[ + \left. \delta(k^2 - m^2) \left[ \frac{1}{q^2 - m^2} + i\pi\epsilon(q)\delta(q^2 - m^2) \right] \right\}. \tag{9.10.160} \]
This integral differs from the one that we had in the case of pair production in that we have two integrations to perform and only one delta function. Thus, we are really summing over infinitely many intermediate states. Also, this integral as it stands is divergent. To be able to deal with this we use some physics to get general properties of this integral. To begin with we have

$$
\langle \Omega | j^\mu(x) | \Omega \rangle = \int d^4 y \ K^{\mu\nu}(x - y) A_\nu(y) .
$$

(9.10.161)

But,

$$
\partial_\mu j^\mu(x) = 0 \implies \partial_\mu K^{\mu\nu}(x - y) = 0
$$

(9.10.162)

and this in turn implies that

$$
p_\mu \tilde{K}^{\mu\nu}(p) = 0 .
$$

(9.10.163)

Therefore, $\tilde{K}^{\mu\nu}(p)$ must be of the form

$$
\tilde{K}^{\mu\nu}(p) = G(p) \left[ p^2 g^{\mu\nu} - p^\mu p^\nu \right] .
$$

(9.10.164)

Formally, as we now show, it is possible to use the integral expression that we have obtained for $\tilde{K}^{\mu\nu}(p)$ to verify that $p_\mu \tilde{K}^{\mu\nu}(p) = 0$.

$$
p_\mu \tilde{K}^{\mu\nu}(p)
= \frac{e^2}{2(2\pi)^3} \int d^4 k \ d^4 q \ \delta(p - k + q) \text{Tr} \left[ \gamma.p(\gamma.k + m)\gamma'^\nu(\gamma.q + m) \right] \\
\left\{ \delta(q^2 - m^2) \left[ P \frac{1}{k^2 - m^2} - i\pi \varepsilon(k) \delta(k^2 - m^2) \right] \\
+ \delta(k^2 - m^2) \left[ P \frac{1}{q^2 - m^2} + i\pi \varepsilon(q) \delta(q^2 - m^2) \right] \right\}
= \frac{e^2}{2(2\pi)^3} \int d^4 k \ d^4 q \ \delta(p - k + q) \text{Tr} \left[ (\gamma.q + m)(\gamma.p)(\gamma.k + m)\gamma'^\nu \right] \{\cdots\}
= \frac{e^2}{2(2\pi)^3} \int d^4 k \ d^4 q \ \delta(p - k + q) \text{Tr} \left[ (\gamma.q + m)(\gamma.p)(\gamma.k + m)\gamma'^\nu \right. \\
\left. - (\gamma.q + m)(\gamma.p)(\gamma.k + m)\gamma'^\nu \right] \{\cdots\}
= \frac{e^2}{2(2\pi)^3} \int d^4 k \ d^4 q \ \delta(p - k + q) \text{Tr} \left[ (\gamma.q)(\gamma.k + m)\gamma'^\nu(k^2 - m^2) \right. \\
\left. + (\gamma.k)\gamma'^\nu(q^2 - m^2) \right] \{\cdots\}
= \frac{e^2}{2(2\pi)^3} \int d^4 k \ d^4 q \ \delta(p - k + q) \left[ q^\nu(k^2 - m^2) \right. \\
\left. + k^\nu(q^2 - m^2) \right] \{\cdots\}
$$

(9.10.165)

Both terms vanish independently by symmetry since integrals of the type

$$
\int d^4 k k^\nu \delta(k^2 - m^2) = 0 .
$$

(9.10.166)
However, in actuality this integral is divergent and by changing variable of integration from \( k \) to \( k' = k - a \) we immediately obtain a result different from 0. Thus, it is only on purely physical grounds that we can require that \( p_\mu \tilde{K}^{\mu \nu}(p) = 0 \). It then follows that

\[
\langle \Omega | j^\mu(x) | \Omega \rangle = \frac{1}{(2\pi)^4} \int d^4p \ e^{ip(x-y)} G(p) [p^2g^{\mu \nu} - p^\mu p^\nu] A_\nu(y) \ d^4y \\
= \frac{1}{(2\pi)^4} \int d^4p \ d^4y \ e^{ip(x-y)} G(p) \left[ -\Box g^{\mu \nu} + \delta^\mu \delta^\nu \right] A_\nu(y) \\
= -\frac{1}{(2\pi)^4} \int d^4p \ d^4y \ e^{ip(x-y)} G(p) j_{\text{external}}^\mu(y) . \tag{9.10.167}
\]

On the other hand, just as in the computation of pair production we have

\[
G(p) = -\frac{1}{2p^3} \tilde{K}^\mu_\mu(p) \tag{9.10.168}
\]

or

\[
G(p) = -\frac{e^2}{6(2\pi)^3 p^2} \int d^4k \ d^4q \ \delta(p - k + q) \text{Tr} \left[ \gamma^\mu (\gamma \cdot k + m) \gamma^\mu (\gamma \cdot q + m) \right] \times \\
\left\{ \delta(q^2 - m^2) \left[ \frac{1}{k^2 - m^2} - i\pi\varepsilon(k)\delta(k^2 - m^2) \right] + \delta(k^2 - m^2) \left[ \frac{1}{q^2 - m^2} + i\pi\varepsilon(q)\delta(q^2 - m^2) \right] \right\} . \tag{9.10.169}
\]

The trace is easily evaluated and yields

\[
\text{Tr} \left[ \gamma^\mu (\gamma \cdot k + m) \gamma^\mu (\gamma \cdot q + m) \right] \\
= \text{Tr} \left[ \gamma^\mu \gamma_a \gamma^\mu \gamma_b \right] k^a q^b + m^2 \text{Tr} \left[ \gamma^\mu \gamma^\mu \right] \\
= 4[g^\mu_a g_a^\nu + g^\mu_a g^\nu_a - g^\mu_a g^\nu_a] k^a q^b + 16m^2 \\
= 8(2m^2 - k \cdot q) . \tag{9.10.170}
\]

### 9.11 Charge Renormalization

Experimentally one cannot observe the induced current \( \langle \Omega | j^\mu(x) | \Omega \rangle \) by itself, but only the total current

\[
j^\mu_{\text{observable}} = j^\mu_{\text{external}} + \langle \Omega | j^\mu(x) | \Omega \rangle \\
= \frac{1}{(2\pi)^4} \int d^4p \ d^4y \ e^{ip(x-y)} \left[ 1 - G(p) \right] j^\mu_{\text{external}}(y) . \tag{9.11.171}
\]

If \( G(p) \) were a constant we could account for this interaction between \( A_\mu \) and \( \psi \) by absorbing \( G(p) \) in a redefinition of the charge for \( j^\mu_{\text{external}} \). In this case the effect would not be observable. By the same argument we may add any constant to \( G(p) \) since such a constant is not observable in principle. This means that all that can be observed (measured) is that part of \( G(p) \) that varies
with \( p \). Thus, to uniquely compare this result with experiment we have to establish a convention for fixing the arbitrary constant in \( 1 - G(p) \). To do this we require (demand) that for very slowly varying external fields \( A_\mu(x) \) the quantity \( j^\mu_{\text{observable}} \) must coincide with \( j^\mu_{\text{external}} \). This simply means that we require that \( G(0) = 0 \). Therefore, we define

\[
j^\mu_{\text{observable}} = \frac{1}{(2\pi)^4} \int d^4 p \, d^4 y \, e^{ip(x-y)} [1-G(p)+G(0)] j^\mu_{\text{external}}(y). \tag{9.11.172}
\]

This manipulation of adding \( G(0) \) is called Charge Renormalization.

The imaginary part \( \Im[G(p)] \) of \( G(p) \) is quite easy to evaluate since it involves integrating with delta functions. Also, to proceed further we recall that

\[
\epsilon(k) = \frac{k^0}{|k^0|}. \tag{9.11.173}
\]

Thus,

\[
\begin{align*}
\Im[G(p)] & = \frac{e^2}{48\pi^2 p^2} \int d^4 k \, d^4 q \, \delta(p - k + q)[\epsilon(k) - \epsilon(q)] \\
& \times \delta(k^2 - m^2) \delta(q^2 - m^2) \delta(2m^2 - k.q) \\
& = \frac{e^2}{6\pi^2 p^2} \int d^4 k \, \delta(k^2 - m^2) \delta((k-p)^2 - m^2)(m^2 + k \cdot p) \\
& \times [\epsilon(k) + \epsilon(p - k)]. \tag{9.11.174}
\end{align*}
\]

The delta functions tell us that

\[
k^2 = m^2
\]

and

\[
p^2 = 2k.p.
\]

Therefore

\[
\begin{align*}
\Im[G(p)] & = \frac{e^2}{6\pi^2 p^2} (m^2 + p^2/2) \int d^4 k \delta(k^2 - m^2) \delta(p^2 - 2p.k)[\epsilon(k) + \epsilon(p - k)] \\
& = \frac{e^2}{6\pi^2 p^2} (m^2 + p^2/2) \int \frac{d^3 k}{\sqrt{k^2 + m^2}} \\
& \times \int dk^0 \left[ \delta(k^0 - \sqrt{k^2 + m^2}) + \delta(k^0 + \sqrt{k^2 + m^2}) \right] \\
& \times \delta(p^2 - 2k^0p^0 + 2k \cdot \vec{p})[\epsilon(k) + \epsilon(p - k)] \\
& = \frac{e^2}{6\pi^2 p^2} (m^2 + p^2/2) \int \frac{d^3 k}{\sqrt{k^2 + m^2}}
\end{align*}
\]
\[ \times \left\{ \delta(p^2 + 2\vec{k} \cdot \vec{p} - 2p^0 \sqrt{k^2 + m^2}) \left[ 1 + \frac{p^0 - \sqrt{k^2 + m^2}}{|p^0 - \sqrt{k^2 + m^2}|} \right] \\
+ \delta(p^2 + 2\vec{k} \cdot \vec{p} + 2p^0 \sqrt{k^2 + m^2}) \left[ -1 + \frac{p^0 + \sqrt{k^2 + m^2}}{|p^0 + \sqrt{k^2 + m^2}|} \right] \right\} . \]

(9.11.175)

This integral is clearly a Lorentz invariant quantity and vanishes for \((p^0)^2 < k^2 + m^2\) as can be seen from the two square brackets. Then, using the two delta functions in the first expression we find that \(p\) is a time-like vector. In fact we shall find that \(p^2 > 4m^2\). This means we can again choose a special frame in which \(\vec{p} = 0\) to evaluate this integral. Then,

\[ \Im[G(p)] = \frac{e^2 4\pi}{24\pi^2} \left( 1 + \frac{2m^2}{p_0^2} \right) \int_0^\infty \frac{k^2 dk}{\sqrt{k^2 + m^2}} \times \left\{ \delta(p_0^2 - 2p_0 \sqrt{k^2 + m^2}) 2\theta(p_0 - \sqrt{k^2 + m^2}) \\
+ \delta(p_0^2 + 2p_0 \sqrt{k^2 + m^2}) 2\theta(p_0 + \sqrt{k^2 + m^2}) \right\} \]

\[ = \frac{e^2}{3\pi} \left( 1 + \frac{2m^2}{p_0^2} \right) \int_0^\sqrt{p_0^2 - m^2} \frac{k^2 dk}{\sqrt{k^2 + m^2}} \frac{\delta(2|p_0|\sqrt{k^2 + m^2} - p_0^2)}{|p_0|} \]

\[ = \frac{e^2}{6\pi} \left( 1 + \frac{2m^2}{p_0^2} \right) \frac{1}{p_0} \sqrt{p_0^2/4 - m^2} \theta(|p_0| - 2m) \frac{p_0}{|p_0|} \]

\[ = \frac{e^2}{12\pi} \left( 1 + \frac{2m^2}{p_0^2} \right) \sqrt{1 - (4m^2)/p_0^2} \theta(p_0^2 - 4m^2) \frac{p_0}{|p_0|}. \]  

(9.11.176)

Rewriting this in terms of invariant quantities we get

\[ \Im[G(p)] = \frac{e^2}{12\pi} \left( 1 + \frac{2m^2}{p^2} \right) \sqrt{1 - (4m^2)/p^2} \theta(p^2 - 4m^2)e(p^0). \]  

(9.11.177)

Due to the theta function which forces \(p^2 > 4m^2\) we see that \(\Im[G(p)] = 0\) for \(|p^0| < 2m\). This means that

\[ \Im[G(0)] = 0. \]  

(9.11.178)

### 9.12 Dispersion Relations

To proceed from the previous section we could now evaluate \(\Re[G(p)]\) directly by integrating the expression that we have obtained. However, it is easier and more instructive to use causality. The relations we now derive are known as the Kramers-Kronig relations in optics, but in general are called dispersion relations.

If we consider the function \(K^{\mu\nu}(z)\) we see that it involves \(S_R(z)\) and \(S_A(-z)\). This means that

\[ K^{\mu\nu}(z) = 0 \text{ for } z^2 < 0 \]  

(9.12.179)
and also

\[ K^{\mu\nu}(z) = 0 \quad \text{for} \quad z^0 < 0. \quad (9.12.180) \]

It therefore follows that the function

\[ \tilde{G}(z) = \frac{1}{(2\pi)^4} \int d^4p \, e^{-i p^0} \, G(p) \quad (9.12.181) \]

must also vanish for \( z^0 < 0 \). This has the following implication for

\[ G(p) = \int d^4z \, e^{i p^0} \tilde{G}(z). \quad (9.12.182) \]

It can be analytically continued to complex \( p^0 \). Clearly, \( G(p^0 + i\eta, \vec{p}) \) is analytic for \( \eta > 0 \) since adding \( i\eta \) to \( p^0 \) can make an already convergent integral only more convergent. Therefore, we can write

\[ G(p^0 + i\eta, \vec{p}) = \frac{1}{2\pi i} \int \frac{G(\omega, \vec{p}) \, d\omega}{\omega - p^0 - i\eta} \quad (9.12.183) \]

where we choose a contour along the real axis with a semicircle closed in the upper complex \( \omega \)-plane. (See figure 9.4.) If we now let \( \eta \to 0^+ \) we get

\[ G(p^0 + i0, \vec{p}) = \frac{P}{2\pi i} \int_{-\infty}^{\infty} \frac{G(\omega, \vec{p}) \, d\omega}{\omega - p^0} + \frac{i\pi}{2\pi i} \int_{-\infty}^{\infty} \delta(\omega - p^0) G(\omega, \vec{p}) \, d\omega. \quad (9.12.184) \]

![Figure 9.4: The Contour Used to Evaluate \( G(p^0 + i\eta, \vec{p}) \).](image)

Therefore,

\[ G(p^0 + i0, \vec{p}) = \frac{1}{i\pi} P \int_{-\infty}^{\infty} \frac{G(\omega, \vec{p}) \, d\omega}{\omega - p^0}. \quad (9.12.185) \]

After taking the real and imaginary parts of this equation we find the desired dispersion relations

\[ \Re[G(p^0, \vec{p})] = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\Im[G(\omega, \vec{p})]}{\omega - p^0} \, d\omega \]

\[ \Im[G(p^0, \vec{p})] = \frac{-1}{\pi} P \int_{-\infty}^{\infty} \frac{\Re[G(\omega, \vec{p})]}{\omega - p^0} \, d\omega. \quad (9.12.186) \]
and this converges. Since, as stated, this is the physically interesting quantity we now evaluate it for $p^2 > 4m^2$

$$
\Pi^{(0)}(p^2) - \Pi^{(0)}(0) = \frac{e^2 p^2}{12\pi^2} \int_{4m^2}^{\infty} \left(1 + \frac{2m^2}{a}\right) \sqrt{1 - \frac{(4m^2)}{a}} \frac{da}{a(a - p^2)}
= \frac{e^2 p^2}{12\pi^2} \left\{ \frac{5}{3} + \frac{4m^2}{p^2} - \left(1 + \frac{2m^2}{p^2}\right) \sqrt{1 - \frac{4m^2}{p^2}} \ln \left[ \frac{1 + \sqrt{1 - \frac{4m^2}{p^2}}}{1 - \sqrt{1 - \frac{4m^2}{p^2}}} \right] \right\} .
$$

(9.12.195)

If $p^2 < 4m^2$ we have an arctan instead of the ln.

It is also worth noting that for

$$\left| \frac{p^2}{m^2} \right| << 1$$

we get

$$\Pi^{(0)}(p^2) - \Pi^{(0)}(0) \approx \frac{e^2 p^2}{60\pi^2 m^2} .$$

(9.12.196)

This causes the potential due to a point charge in vacuum to differ from the Coulomb potential so that, in effect, one has a potential

$$-Ze^2 \left[ \frac{1}{r} - \frac{4\pi e^2}{15\pi m^2} \delta(\vec{r}) \right] .$$

This is called the Uehling effect [2] and lowers the 2S$_{1/2}$ level in hydrogen relative to the 2P$_{1/2}$ level by about $2.7 \times 10^7$ Hz.

Thus, although the charge renormalization $\Pi^{(0)}(0)$ is infinite we get definite, finite predictions for experimentally observable effects. This renormalization (subtraction of $\Pi^{(0)}(0)$ ) which is absolutely necessary to get unambiguous predictions, also (happily) renders the expectation value of the current finite. It is in this manner that renormalization provides a satisfactory formulation of QED.

As this computation shows, the vacuum behaves like a polarizable medium with dielectric constant

$$\varepsilon(p^2) = 1 - \Pi^{(0)}(p^2) - \Pi^{(0)}(0) - i\epsilon(p^0)\Pi^{(0)}(p^2) .$$

(9.12.197)

### 9.13 Problems

9.1 Show that

$$\gamma^0 S^\dagger_R(x - y)\gamma^0 = S_A(y - x) .$$

9.2 Use the anticommutation relations for the Dirac annihilation and creation operators to show that

$$[\psi_\alpha(x), \bar{\psi}_\beta(y)]_+ = -iS_{\alpha\beta}(x - y) .$$

(9.13.198)
Bibliography

A detailed proof is also given in
R.F. Streater and A.S. Wightman, PCT, Spin And Statistics And All
That, W.A.Benjamin, Inc. (1964).

Chapter 10

Asymptotic Fields: LSZ Formulation

10.1 The Scalar Field

Suppose we have an interacting scalar field $\phi$ satisfying the Heisenberg field equation

\[ (\Box + m_0^2)\phi(x) = j(x) \]  \hspace{1cm} (10.1.1)

where $j$ is a functional of $\phi$ and the fields satisfy the canonical commutation relations.

\[ [\phi(t, \vec{x}), \pi(t, \vec{y})] = i\delta(\vec{x} - \vec{y}) \]
\[ [\phi(t, \vec{x}), \phi(t, \vec{y})] = 0 \]
\[ [\pi(t, \vec{x}), \pi(t, \vec{y})] = 0. \]  \hspace{1cm} (10.1.2)

If the interaction does not involve any derivative coupling then

\[ \pi(x) = \dot{\phi}(x). \]  \hspace{1cm} (10.1.3)

An example of such a case is provided by the lagrange density

\[ \mathcal{L} = \frac{1}{2} \left[ \partial_{\mu}\phi\partial^{\mu}\phi - m_0^2\phi^2 + \frac{\lambda}{2}\phi^4 \right]. \]  \hspace{1cm} (10.1.4)

In this case

\[ j(x) = \lambda\phi^3(x). \]  \hspace{1cm} (10.1.5)

We are interested in finding an asymptotic field operator $\phi_{in}(x)$ formed from the Heisenberg field $\phi(x)$ and having the following properties.

1) Under the Poincaré group $\phi_{in}(x)$ transforms in the same way as $\phi(x)$. Thus, for displacements we have

\[ [P_\mu, \phi_{in}(x)] = -i\frac{\partial \phi_{in}}{\partial x_\mu}. \]  \hspace{1cm} (10.1.6)
2) \( \phi_{in}(x) \) is a free Klein-Gordon field with the physical mass \( m \).

\[
(\Box + m^2) \phi_{in}(x) = 0 .
\] (10.1.7)

In general, \( m \neq m_0 \). We call \( m_0 \) the bare mass. The two conditions, we have imposed, imply that \( \phi_{in}(x) \) creates physical one-particle states from the physical vacuum \( |\Omega\rangle \).

Now consider a state \( |n\rangle \) such that

\[
P^\mu |n\rangle = p_n^\mu |n\rangle .
\] (10.1.8)

We then have that

\[
-i\partial_\mu \langle n | \phi_{in}(x) | \Omega \rangle = \langle n | [P^\mu, \phi_{in}(x)] | \Omega \rangle = p_n^\mu \langle n | \phi_{in}(x) | \Omega \rangle .
\] (10.1.9)

If we repeat this computation and use condition 2) we find

\[
(\Box + m^2) \langle n | \phi_{in}(x) | \Omega \rangle = (-p_n^2 + m^2) \langle n | \phi_{in}(x) | \Omega \rangle = 0 .
\] (10.1.10)

So, we can as usual write

\[
\phi_{in}(x) = \int \frac{d^3k}{\omega(k)} \left[ a_{in}(\vec{k}) f_{\vec{k}}(x) + a_{in}^†(\vec{k}) f_{\vec{k}}^*(x) \right]
\] (10.1.11)

where the inverse formula is

\[
a_{in}(\vec{k}) = -i \int d^3x f_{\vec{k}}^*(x) \partial_0 \phi_{in}(x) .
\] (10.1.12)

Here,

\[
f_{\vec{k}}(x) = \frac{1}{\sqrt{2(2\pi)^3}} e^{-ikx}
\] (10.1.13)

Furthermore,

\[
[P_\mu, a_{in}(\vec{k})] = -k_\mu a_{in}(\vec{k}) .
\] (10.1.14)

To express \( \phi_{in} \) in terms of \( \phi \) we rewrite the Heisenberg equation for \( \phi \) by adding a term \( \delta m^2 = m^2 - m_0^2 \) to both sides of the equation.

\[
(\Box + m^2) \phi(x) = j(x) + \delta m^2 \phi(x) = \tilde{j}(x) .
\] (10.1.15)

In that case we can write the integral equation

\[
\sqrt{Z} \phi_{in}(x) = \phi(x) - \int d^4y \Delta_R(x - y : m) \tilde{j}(y) .
\] (10.1.16)

Here we have included a normalization factor \( \sqrt{Z} \) so that we can properly normalize the single particle amplitude. For historical reasons it is called the wavefunction renormalization even though it renormalizes the field operator. The point is that we want

\[
\lim_{x^0 \to -\infty} \phi(x) = \sqrt{Z} \phi_{in}(x) .
\] (10.1.17)
Equation 10.1.17 cannot hold as an operator equation but can only as a weak limit, that is, for expectation values, as explained next.

Let $\ket{\alpha}, \ket{\beta}$ be two normalizable states and define

$$\phi(t, f) = i \int d^3 x \, f^* (t, \vec{x}) \, \partial_0 \, \phi \, (t, \vec{x})$$

(10.1.18)

where $f$ is a normalizable solution of

$$(\Box + m^2) f = 0$$

(10.1.19)

In this case the following limit is well defined.

$$\lim_{t \to -\infty} \langle \alpha | \phi(t, f) | \beta \rangle = \sqrt{Z} \langle \alpha | \phi_{in}(t, f) | \beta \rangle$$

(10.1.20)

The limit above is called a weak asymptotic condition.

In a similar manner we get another asymptotic free field $\phi_{out}(x)$.

$$\lim_{t \to \infty} \langle \alpha | \phi(t, f) | \beta \rangle = \sqrt{Z} \langle \alpha | \phi_{out}(t, f) | \beta \rangle$$

(10.1.21)

The $S$-matrix is the matrix of the $S$-operator formed between free "in" and "out" states.

$$S_{\beta \alpha} = \langle \beta \, \text{out} | S | \alpha \, \text{in} \rangle$$

(10.1.22)

This means that the $S$-operator transforms "in" states into "out" states

$$| \text{in} \rangle = S | \text{out} \rangle \quad | \text{out} \rangle = S^\dagger | \text{in} \rangle$$

(10.1.23)

In terms of the field operators we have

$$\phi_{in} = S \, \phi_{out} \, S^\dagger$$

(10.1.24)

We now show how to write the $S$-matrix elements as vacuum expectation values of products of field operators. This is known as the LSZ Formalism.

### 10.2 Reduction of the $S$-matrix: Scalar Field

The derivation that follows is for a real scalar field. Consider a general state $|\alpha; \vec{p}_n; \text{in}\rangle$ where

$$\alpha = \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_{n-1}$$

(10.2.25)

as well as another state $|\beta; \text{out}\rangle$ with

$$\beta = \vec{q}_1, \vec{q}_2, \ldots, \vec{q}_m$$

(10.2.26)

We want to write the amplitude $\langle \beta; \text{out}| \alpha; \vec{p}_n; \text{in}\rangle$ as the vacuum expectation value of a product of field operators by employing the asymptotic condition

$$\langle \beta; \text{out}| \alpha; \vec{p}_n; \text{in}\rangle = \langle \beta; \text{out}| a_{in}^\dagger (\vec{p}_n)| \alpha; \text{in}\rangle$$

$$= \langle \beta; \text{out}| a_{out}^\dagger (\vec{p}_n)| \alpha; \text{in}\rangle + \langle \beta; \text{out}| \left( a_{in}^\dagger (\vec{p}_n) - a_{out}^\dagger (\vec{p}_n) \right) | \alpha; \text{in}\rangle$$

$$= \langle \beta; \vec{p}_n; \text{out}| \alpha; \text{in}\rangle$$

$$- i \langle \beta; \text{out}| \int d^3 x \, f_{p_n}(x) \, \partial_0 \, \left( \phi_{in}(x) - \phi_{out}(x) \right) | \alpha; \text{in}\rangle$$

(10.2.27)
The symbol $\beta, \tilde{p}_n$ is to be interpreted as follows. If $\tilde{p}_n \in \beta$ then it means "remove this term". If $\tilde{p}_n \notin \beta$ it means that this state is annihilated. The first term $\langle \beta, \tilde{p}_n; out|\alpha; in \rangle$ contributes only to forward elastic scattering.

The next step is to replace $\phi_{in}$ and $\phi_{out}$ by limits.

\[
\langle \beta; out|\alpha, \tilde{p}_n; in \rangle = \langle \beta, \tilde{p}_n; out|\alpha; in \rangle + \frac{i}{\sqrt{Z}} \left( \lim_{x^0 \to +\infty} - \lim_{x^0 \to -\infty} \right) \int d^3x f_{p_n}(x) \partial_0^+ \langle \beta; out|\phi(x)|\alpha; in \rangle .
\]  

(10.2.28)

To get a more convenient expression we incorporate the $x^0 \to \pm \infty$ limits in integrals by using

\[
\left( \lim_{x^0 \to +\infty} - \lim_{x^0 \to -\infty} \right) \int d^3x f(x) \partial_0^+ g(x)
\]

\[
= \int_{-\infty}^{\infty} d^4x \frac{\partial}{\partial x_0} \left[ f(x) \partial_0^+ g(x) \right]
\]

\[
= \int_{-\infty}^{\infty} d^4x \left[ f(x)\partial_0^2 g(x) - \partial_0^2 f(x) g(x) \right].
\]  

(10.2.29)

Furthermore,

\[
\partial_0^2 f_{p_n}(x) = (\nabla^2 - m^2) f_{p_n}(x).
\]  

(10.2.30)

Thus, assuming that $\tilde{p}_n \notin \beta$, so that we can drop the elastic term, we get

\[
\langle \beta; out|\alpha, \tilde{p}_n; in \rangle = + \frac{i}{\sqrt{Z}} \int_{-\infty}^{\infty} d^4x \left[ f_{p_n}(x)(\partial_0^2 + m^2)\langle \beta; out|\phi(x)|\alpha; in \rangle - \nabla^2 f_{p_n}(x)\langle \beta; out|\phi(x)|\alpha; in \rangle \right]
\]  

(10.2.31)

Next, we integrate twice by parts to get for $\tilde{p}_n \notin \beta$

\[
\langle \beta; out|\alpha, \tilde{p}_n; in \rangle = \frac{i}{\sqrt{Z}} \int_{-\infty}^{\infty} d^4x f_{p_n}(x)(\square + m^2)\langle \beta; out|\phi(x)|\alpha; in \rangle .
\]  

(10.2.32)

With appropriate modifications we now repeat this procedure for the left side. First we write

\[
\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_m = \gamma, \tilde{q}_m
\]

(10.2.33)

and recall that

\[
\alpha = \tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{n-1}.
\]

(10.2.34)

Then,

\[
M = \langle \gamma, \tilde{q}_m; out|\phi(x)|\alpha; in \rangle = \langle \gamma; out|\phi(x)|\alpha, \tilde{q}_m; in \rangle
\]

\[
+ \langle \gamma; out|a_{out}(\tilde{q}_m)\phi(x) - \phi(x)a_{in}(\tilde{q}_m)|\alpha; in \rangle
\]

\[
= \langle \gamma; out|\phi(x)|\alpha, \tilde{q}_m; in \rangle
\]

\[
- i \int d^3y \langle \gamma; out|\phi_{out}(y)\phi(x) - \phi(x)\phi_{in}(y)|\alpha; in \rangle \partial_0 f_{\tilde{q}_m}^*(y) .
\]  

(10.2.35)
Now, we again use the asymptotic condition and rewrite the integral as a four-dimensional integration to get

\[
M = \frac{-i}{\sqrt{Z}} \left( \lim_{y^0 \to +\infty} - \lim_{y^0 \to -\infty} \right) \int d^3y \langle \gamma; \text{out} | T(\phi(y)\phi(x)) | \alpha; \text{in} \rangle \tag{10.2.36}
\]

where

\[
T(\phi(y)\phi(x)) = \phi(y)\phi(x)\theta(y^0 - x^0) + \phi(x)\phi(y)\theta(x^0 - y^0). \tag{10.2.37}
\]

So, if \( \vec{q}_m \not\in \alpha \) (no forward scattering) we find

\[
\langle \gamma, \vec{q}_m; \text{out} | \phi(x) | \alpha; \text{in} \rangle = \frac{i}{\sqrt{Z}} \int d^4y \left[ (\Box_y + m^2) \langle \gamma; \text{out} | T(\phi(y)\phi(x)) | \alpha; \text{in} \rangle \right] f^{*}_{q_m}(y) \tag{10.2.38}
\]

We repeat this procedure \((n - 1)\) and \((m - 1)\) times and assume that there is no forward scattering \((\vec{p}_i \not= \vec{q}_j \text{ for all } i, j)\). Then,

\[
\langle \gamma_1, \vec{q}_2, \ldots, \vec{q}_m; \text{out} | \vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n; \text{in} \rangle = \left( \frac{i}{\sqrt{Z}} \right)^{n+m} \prod_{i=1}^{n} \prod_{j=1}^{m} \int d^4x_id^4y_j f_{p_i}(x_i)f_{q_j}^*(y_j) (\Box_i + m^2) (\Box_j + m^2) \langle \Omega | T(\phi(y_1) \ldots \phi(y_m)\phi(x_1) \ldots \phi(x_n)) | \Omega \rangle. \tag{10.2.39}
\]

Comment
The expression

\[
\langle \Omega | T(\phi(z_1) \ldots \phi(z_r)) | \Omega \rangle
\]

represents the sum of all Feynman graphs with \(r\) particles created or destroyed at \(z_1 \ldots z_r\). This is all lines representing a meson and that start or end at \(z_1 \ldots z_r\). It is known as the Complete \(r\)-particle Green's Function and is shown in figure 10.1. After some rewriting, the result obtained above becomes the basis for all computations.

Next, we sketch the same computation for a Dirac field.
10.3 Reduction of the S-matrix: Dirac Field

We now repeat the computation of the previous section for a Dirac field. In this case a single particle state is labelled \((p_n s_n)\). Then, with an obvious generalization of the previous notation we have

\[
\begin{align*}
\langle \beta; \text{out}|\alpha, (p_n s_n); \text{in} \rangle &= \langle \beta, (p_n s_n); \text{out}|\alpha; \text{in} \rangle \\
&+ \langle \beta; \text{out}|b_{\text{in}}^\dagger (p_n s_n) - b_{\text{out}}^\dagger (p_n s_n)|\alpha; \text{in} \rangle \\
&= \langle \beta, (p_n s_n); \text{out}|\alpha; \text{in} \rangle \\
&+ \int d^3x \langle \beta; \text{out}|\psi_{\text{in}}^\dagger (x) - \psi_{\text{out}}^\dagger (x)|\alpha; \text{in} \rangle U_{(p_n s_n)}(x) .
\end{align*}
\]

(10.3.40)

Here, \(U_{(p_n s_n)}(x)\) stands for any positive energy solution of the free Dirac equation.

\[
\gamma^0 \partial^0 U_{(p_n s_n)}(x) = (\vec{\gamma} \cdot \nabla - im)U_{(p_n s_n)}(x) .
\]

(10.3.41)

So, these solutions are of the form

\[
U_{(p_n s_n)}(x) = \frac{1}{\sqrt{(2\pi)^3}} \sqrt{\frac{m}{E(p_n)}} e^{-i p_n \cdot x} u(p_n s_n) .
\]

(10.3.42)

Therefore,

\[
M = \langle \beta; \text{out}|\alpha, (p_n s_n); \text{in} \rangle = \langle \beta, (p_n s_n); \text{out}|\alpha; \text{in} \rangle \\
- \frac{1}{\sqrt{Z_2}} \int d^4x \langle \beta; \text{out}|\frac{\partial}{\partial x^\mu} (\bar{\psi}(x) \gamma^0 U_{(p_n s_n)}(x)) |\alpha; \text{in} \rangle .
\]

(10.3.43)

Here, \(Z_2\) is the "wavefunction renormalization" for the Dirac field. As before we assume that there is no forward scattering so that \((p_n s_n) \notin \beta\). Then,

\[
\begin{align*}
M &= - \frac{1}{\sqrt{Z_2}} \int d^4x \langle \beta; \text{out}| \left( \frac{\partial}{\partial x^\mu} \bar{\psi}(x) \gamma^0 \right) U_{(p_n s_n)}(x) \\
&+ \bar{\psi}(x) \left( \frac{\partial}{\partial x^\mu} \gamma^0 U_{(p_n s_n)}(x) \right) |\alpha; \text{in} \rangle \\
&= - \frac{1}{\sqrt{Z_2}} \int d^4x \langle \beta; \text{out}| \left( \frac{\partial}{\partial x^\mu} \bar{\psi}(x) \gamma^0 \right) U_{(p_n s_n)}(x) \\
&+ \bar{\psi}(x) \left( \vec{\gamma} \cdot \vec{\nabla} - im \right) U_{(p_n s_n)}(x) |\alpha; \text{in} \rangle .
\end{align*}
\]

(10.3.44)

Therefore,

\[
M = \frac{i}{\sqrt{Z_2}} \int d^4x \langle \beta; \text{out}| \left( \bar{\psi}(x) (i\vec{\nabla} + \vec{m}) \right) U_{(p_n s_n)}(x) |\alpha; \text{in} \rangle .
\]

(10.3.45)

If we had removed an antiparticle from the in-state we would have gotten

\[
\frac{i}{\sqrt{Z_2}} \int d^4x \bar{\psi}_{(p_n s_n)}(x)(i\vec{\nabla} - \vec{m}) \langle \beta; \text{out}|\psi(x) |\alpha; \text{in} \rangle .
\]
Similarly, if we had removed a particle from the out-state we would have gotten
\[
\frac{-i}{\sqrt{Z_2}} \int d^4x \tilde{U}_{(p_n s_n)}(x)(i\gamma - m)(\beta; out|\psi(x)|\alpha; in)
\]
and finally removing an antiparticle from the out-state would have yielded
\[
\frac{-i}{\sqrt{Z_2}} \int d^4x (\beta; out|\tilde{\psi}(x)|\alpha; in) (i\gamma + m) \tilde{V}_{(p_n s_n)}(x)
\]
The reader should now go through the necessary steps to verify these results.
Suppose we now remove a second particle from the right hand side of $M$. Then, with $\alpha' = (\alpha, (p_{n-1}s_{n-1}))$ and assuming no forward scattering ($(p_{n-1}s_{n-1}) \not\in \beta$), we get
\[
M = \langle \beta; out|\tilde{\psi}(x)|\alpha', (p_{n-1}s_{n-1}); in \rangle = \langle \beta; out|\tilde{\psi}(x)b^\dagger_{in}(p_{n-1}s_{n-1})|\alpha'; in \rangle
= \langle \beta; out|\tilde{\psi}(x)b^\dagger_{in}(p_{n-1}s_{n-1}) - b^\dagger_{out}(p_{n-1}s_{n-1})\tilde{\psi}(x)|\alpha'; in \rangle
= \int d^3y \langle \beta; out|\tilde{\psi}(x)\psi_{in}(y) - \psi_{out}(y)\tilde{\psi}(x)|\alpha'; in \rangle U_{(p_{n-1}s_{n-1})}(y) \quad (10.3.46)
\]
As before we use the asymptotic conditions to write
\[
\langle \beta|\tilde{\psi}_{in}(f,t)|\alpha \rangle = \lim_{t \to -\infty} \langle \beta|\tilde{\psi}(f,t)|\alpha \rangle
\]
\[
\langle \beta|\tilde{\psi}_{out}(f,t)|\alpha \rangle = \lim_{t \to +\infty} \langle \beta|\tilde{\psi}(f,t)|\alpha \rangle \quad (10.3.47)
\]
where
\[
\tilde{\psi}(f,t) = \int d^3x \tilde{\psi}(\vec{x}, t) f(\vec{x})
\]
\[
\psi(f,t) = \int d^3x f(\vec{x})\tilde{\psi}(\vec{x}, t). \quad (10.3.48)
\]
Therefore,
\[
M = \langle \beta; out|\tilde{\psi}(x)|\alpha', (p_{n-1}s_{n-1}); in \rangle
= -\frac{1}{\sqrt{Z_2}} \left( \lim_{y^0 \to +\infty} - \lim_{y^0 \to -\infty} \right)
\times \int d^3y \langle \beta; out|T(\tilde{\psi}(y)\gamma^0 \tilde{\psi}(x))|\alpha'; in \rangle U_{(p_{n-1}s_{n-1})}(y) \quad (10.3.49)
\]
Here,
\[
T(\tilde{\psi}_\alpha(x)\tilde{\psi}_\beta(y)) = \tilde{\psi}_\alpha(x)\tilde{\psi}_\beta(y)\theta(x^0 - y^0) - \tilde{\psi}_\beta(y)\tilde{\psi}_\alpha(x)\theta(y^0 - x^0) \quad (10.3.50)
\]
So, again incorporating the limits by going to an integral over $y^0$, we get
\[
\langle \beta|\tilde{\psi}_{in}(f,t)|\alpha \rangle
= -\frac{1}{\sqrt{Z_2}} \int d^4y \frac{\partial}{\partial y^0} \left[ \langle \beta; out|T(\tilde{\psi}(y)\gamma^0 \tilde{\psi}(x))|\alpha'; in \rangle U_{(p_{n-1}s_{n-1})}(y) \right]
= -\frac{i}{\sqrt{Z_2}} \int d^4y \langle \beta; out|T(\tilde{\psi}(y)\tilde{\psi}(x))|\alpha'; in \rangle (-i\gamma_0 - m) U_{(p_{n-1}s_{n-1})}(y) \quad (10.3.51)
\]
All of this generalizes as follows. If we start with an amplitude
\[ \langle \beta; \text{out} | T(\phi(x_1) \ldots \phi(x_n)\psi_{\alpha_1}(y_1) \ldots \psi_{\alpha_m}(y_m)\bar{\psi}_{\beta_1}(z_1) \ldots \bar{\psi}_{\beta_p}(z_p)) | \alpha, (ps); \text{in} \rangle \]
we get the following results for \((ps) \notin \beta\). If we remove a particle from the in-state the result is:
\[
\frac{-i}{\sqrt{Z_2}} \int d^4x \langle \beta; \text{out} | T(\phi(x_1) \ldots \phi(x_n)\psi_{\alpha_1}(y_1) \ldots \psi_{\alpha_m}(y_m)\bar{\psi}_{\beta_1}(z_1) \ldots \bar{\psi}_{\beta_p}(z_p)) | \alpha; \text{in} \rangle (-i\not\partial_x - m)_{\lambda \tau} U_{(ps, \tau)}(x).
\]
(10.3.52)

If we remove an antiparticle from the in-state the result is:
\[
\frac{i}{\sqrt{Z_2}} \int d^4x \tilde{\bar{V}}_{(ps, \tau)}(x) (i\not\partial_x - m)_{\tau \lambda} (-1)^{m+p} \langle \beta; \text{out} | T(\psi_{\lambda}(x)\phi(x_1) \ldots \phi(x_n)\psi_{\alpha_1}(y_1) \ldots \psi_{\alpha_m}(y_m)\bar{\psi}_{\beta_1}(z_1) \ldots \bar{\psi}_{\beta_p}(z_p)) | \alpha; \text{in} \rangle.
\]
(10.3.53)

This process may be continued until by repeated reduction we arrive at
\[ \langle \Omega | T(\phi(x_1) \ldots \psi_{\alpha_1}(y_1) \ldots \bar{\psi}_{\beta_1}(z_1) \ldots | \Omega \rangle. \]
This now involves nothing but vacuum expectation values (VEV)'s.

As an example of how this is used we consider the scattering of a meson of mass \(\mu\) from a proton of mass \(m\).

### 10.3.1 Proton-Meson Scattering

The Feynman diagram corresponding to this looks as in figure 10.2. The corresponding S-matrix element is
\[
S_{fi} = \langle \bar{q}' (\bar{p}' s'); \text{out} | \bar{q}, (p s); \text{in} \rangle = \delta_{fi} + \frac{1}{Z_2} \int d^4x d^4x' d^4y d^4y' (\delta_{x, x'}^2 + \mu^2)(\delta_{x, x'}^2 + \mu^2)f_q^*(x')f_q(x)\bar{U}_{p', s'}(y') \times \left[ (i\not\partial_{y'} - m)\langle \Omega | T(\psi(y')\bar{\psi}(y)\phi(x)\phi(x')) | \Omega \rangle \right] (-i\not\partial_{y} - m) U_{ps}(y).
\]
(10.3.54)
The term $\delta_{fi}$ is different from zero only for forward scattering. In this case if we choose plane waves, as one always does for practical calculations, we have

$$f_q(x) = \frac{1}{\sqrt{2(2\pi)^3}} e^{-iqx}$$

(10.3.55)

and

$$U_{ps}(x) = \frac{1}{\sqrt{(2\pi)^3}} \sqrt{\frac{M}{E(p)}} e^{-ipx}.$$  

(10.3.56)

Then,

$$\delta_{fi} = q^0 \delta(\vec{q} - \vec{q'}) \delta(\vec{p} - \vec{p'}) \delta_{ss'}.$$  

(10.3.57)

The next step is to derive the Interaction or Dirac Picture. This allows us to replace the the Heisenberg field operators $\phi(x)$, $\psi(x)$, $\tilde{\psi}(x)$ by the corresponding free field operators $\phi_{in}(x)$, $\psi_{in}(x)$, $\tilde{\psi}_{in}(x)$. Since these are known we arrive at an expression in terms of known fields.

It is in exactly this manner that Feynman diagrams arise in perturbation theory.

### 10.4 Problems

10.1 Verify the results that if we had removed an antiparticle from the in-state we would have gotten

$$\frac{i}{\sqrt{Z_2}} \int d^4x \tilde{V}_{(p\lambda s_n)}(x)(i\gamma - m)\langle \beta; out|\psi(x)|\alpha; in \rangle.$$  

Similarly, if we had removed a particle from the out-state we would have gotten

$$\frac{-i}{\sqrt{Z_2}} \int d^4x \tilde{U}_{(p\lambda s_n)}(x)(i\gamma - m)\langle \beta; out|\psi(x)|\alpha; in \rangle$$  

and finally removing an antiparticle from the out-state would have yielded

$$\frac{-i}{\sqrt{Z_2}} \int d^4x \langle \beta; out|\tilde{\psi}(x)|\alpha; in \rangle (i\gamma + m) V_{(p\lambda s_n)}(x).$$
Chapter 11

Perturbation Theory

11.1 The U-matrix

Since the Heisenberg field $\phi(x)$ and its canonical momentum $\pi(x)$ satisfy the same equal time commutation relations as the asymptotic free fields $\phi_{in}(x)$, $\pi_{in}(x)$ and since both $\phi(x)$, $\pi(x)$ and $\phi_{in}(x)$, $\pi_{in}(x)$ are complete sets of field operators, there has to exist an operator $U(t)$ such that

$$
\phi(\vec{x}, t) = U^{-1}(t) \phi_{in}(\vec{x}, t) U(t)
$$

$$
\pi(\vec{x}, t) = U^{-1}(t) \pi_{in}(\vec{x}, t) U(t)
$$

(11.1.1)

It is also straightforward to find an equation for $U$ since we have the Heisenberg equations of motion for $\phi(x)$, $\phi_{in}(x)$.

$$
\frac{\partial \phi_{in}(x)}{\partial t} = i[H_{in}(\phi_{in}, \pi_{in}), \phi_{in}]
$$

$$
\frac{\partial \pi_{in}(x)}{\partial t} = i[H_{in}(\phi_{in}, \pi_{in}), \pi_{in}]
$$

(11.1.2)

Here, $H_{in}(\phi_{in}, \pi_{in})$ is the hamiltonian for the free in-field which has a mass $m$.

The full Heisenberg field (interpolating field) $\phi$ which has a mass $m_0$ satisfies

$$
\frac{\partial \phi(x)}{\partial t} = i[H(\phi, \pi), \phi]
$$

$$
\frac{\partial \pi(x)}{\partial t} = i[H(\phi, \pi), \pi]
$$

(11.1.3)

It therefore follows that

$$
\frac{\partial \phi_{in}(\vec{x}, t)}{\partial t} = \frac{\partial}{\partial t} \left[ U(t) \phi(\vec{x}, t) U^{-1}(t) \right]
$$

$$
= \dot{U} \phi U^{-1} + U \dot{\phi} U^{-1} - U \phi \left( U^{-1} \dot{U} U^{-1} \right)
$$

$$
= \dot{U} U^{-1} \phi_{in} + iU[H(\phi, \pi), \phi] U^{-1} - \phi_{in} \dot{U} U^{-1}
$$

$$
= [\dot{U}(t) U^{-1}(t), \phi_{in}(\vec{x}, t)] + i[H(\phi_{in}, \pi_{in}), \phi_{in}(x)].
$$

(11.1.4)
Next, we write
\[ H(\phi_{in}, \pi_{in}) = H_{in}(\phi_{in}, \pi_{in}) + H_I(\phi_{in}, \pi_{in}) . \] (11.1.5)

Therefore,
\[ H_I(\phi_{in}, \pi_{in}) \equiv H_I(t) = H(\phi_{in}, \pi_{in}) - H_{in}(\phi_{in}, \pi_{in}) . \] (11.1.6)

Thus, we get
\[ \dot{\phi}_{in}(\vec{x}, t) = \phi_{in}(\vec{x}, t) + [\dot{U}(t)U^{-1}(t) + iH_I(t), \phi_{in}(\vec{x}, t)] . \] (11.1.7)

Similarly,
\[ \dot{\pi}_{in}(\vec{x}, t) = \pi_{in}(\vec{x}, t) + [\dot{U}(t)U^{-1}(t) + iH_I(t), \pi_{in}(\vec{x}, t)] . \] (11.1.8)

Therefore,
\[ \dot{U}(t)U^{-1}(t) + iH_I(t) = -iE_0(t) \text{ a c-number} . \] (11.1.9)

So,
\[ i\dot{U}(t) = (H_I(t) + E_0(t))U(t) . \] (11.1.10)

It is important to notice that \( H_I(t) \) contains all the interactions, including the mass counter term \( m - m_0 \).

Next, we define
\[ H'_I(t) = H_I(t) + E_0(t) . \] (11.1.11)

A solution for \( U(t) \) in terms of in-fields serves as a basis for perturbation expansions since the operator \( U \) allows us to write \( \langle \Omega | T(\phi(x_1) \ldots \phi(x_n)) | \Omega \rangle \) as an infinite series of terms of the form \( \langle \Omega | T(\phi_{in}(x_1) \ldots) | \Omega \rangle \).

To find \( U(t) \) we need initial conditions. In order to get these we consider the operator
\[ U(t, t') = U(t)U^{-1}(t') . \] (11.1.12)

This operator is also a solution of
\[ i\frac{\partial U(t, t')}{\partial t} = H'_I(t)U(t, t') . \] (11.1.13)

Furthermore, \( U(t, t') \) satisfies the initial condition
\[ U(t, t) = 1 . \] (11.1.14)

After converting 11.1.13 and the initial condition 11.1.14 to an integral equation we get
\[ U(t, t') = 1 - i \int_{t'}^{t} dt_1 H'_I(t_1) U(t_1, t') . \] (11.1.15)
If we iterate this equation we get

\[ U(t, t') = 1 - i \int_{t'}^{t} dt_1 H'_1(t_1) + (-i)^2 \int_{t'}^{t} dt_1 H'_1(t_1) \int_{t'}^{t_1} dt_2 H'_1(t_2) + \cdots \]

\[ + (-i)^n \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 \cdots \int_{t'}^{t_{n-1}} dt_n H'_1(t_1) \cdots H'_n(t_n) + \cdots \] (11.1.16)

Notice that in these integrals we always have

\[ t_1 \geq t_2 \geq \cdots t_n \geq \cdots \] (11.1.17)

This means we can write the product of interaction hamiltonians as

\[ T(H'_1(t_1) \cdots H'_1(t_n)) \] (11.1.18)

both for fermions and bosons without any change of sign for the fermion case since \( H'_1(t) \) always involves a product of an even number of fermions. We therefore have

\[ T(H'_1(t_1) \cdots H'_1(t_n)) = T(H'_1(t_{p_1}) \cdots H'_1(t_{p_n})) \] (11.1.19)

where \( p_1, \ldots, p_n \) is any permutation of 1, 2, \ldots, n . We can use this result to symmetrize the integrals. Thus, for example,

\[ \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 T(H'_1(t_1)H'_1(t_2)) \]

\[ = \int_{t'}^{t} dt_2 \int_{t'}^{t_2} dt_1 T(H'_1(t_1)H'_1(t_2)) \]

\[ = \frac{1}{2} \int_{t'}^{t} dt_1 \int_{t'}^{t} dt_2 T(H'_1(t_1)H'_1(t_2)) \] (11.1.20)

as is obvious from 11.1. In general we get a factor of \( 1/n! \) in front of the integrals.

---

**Figure 11.1:** Change of Integration Order
Thus,

\[
U(t, t') = 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{t'}^{t} dt_1 \cdots \int_{t'}^{t} dt_n T(H'_1(t_1) \cdots H'_n(t_n))
\]

\[
= T \left( \exp \left[ -i \int_{t'}^{t} d\tau H'_i(\tau) \right] \right)
\]

\[
= T \left( \exp \left[ -i \int_{t'}^{t} d^4x \mathcal{H}'_i(\phi_{\text{in}}(x)) \right] \right) .
\] (11.1.21)

The last two lines are just compact ways of writing what is really meant by the first line. One always calculates with the first line. It is worth noting that

\[
U(t, t') = U(t, t'')U(t'', t')
\] (11.1.22)

so that

\[
U^{-1}(t, t') = U(t', t) .
\] (11.1.23)

This means that we have

\[
\phi(x) = U^{-1}(t)\phi_{\text{in}}(x)U(t) .
\] (11.1.24)

### 11.2 Perturbation of VEV of T-Ordered Products

We now consider the vacuum expectation value of the time ordered product of \( n \) Heisenberg fields

\[
\tau(x_1, \ldots, x_n) = \langle \Omega | T(\phi(x_1) \cdots \phi(x_n)) | \Omega \rangle
\] (11.2.25)

and use the \( U \)-matrix to express this in terms of in-fields. Thus, we have

\[
\phi(x_1) = U^{-1}(t_1)\phi_{\text{in}}(x_1)U(t_1)
\]

\[
\phi(x_2) = U^{-1}(t_2)\phi_{\text{in}}(x_2)U(t_2)
\]

\[
\text{etc}
\] (11.2.26)

as well as

\[
U(t_1)U^{-1}(t_2) = U(t_1, t_2) .
\] (11.2.27)

Then,

\[
\tau(x_1, \ldots, x_n)
\]

\[
= \langle \Omega | T(U^{-1}(t_1)\phi_{\text{in}}(x_1)U(t_1, t_2) \cdots U(t_{n-1}, t_n)\phi_{\text{in}}(x_n)U(t_n)) | \Omega \rangle .
\] (11.2.28)
Next, we introduce a reference time $t$ where we later let $t \to \infty$ so that $t > \text{all other times (i.e. } t \text{ is later than all } t_j \text{ and } -t \text{ is earlier than all } t_j).$ Then,

$$\tau(x_1, \ldots, x_n) = \langle \Omega | T \left( U^{-1}(t) U(t, t_1) \phi_{in}(x_1) U(t_1, t_2) \cdots U(t_{n-1}, t_n) \phi_{in}(x_n) U(t_n, -t) U(-t) \right) | \Omega \rangle.$$  

(11.2.29)

After letting $t \to \infty$ we can now take the operators $U^{-1}(t)$ and $U(-t)$ outside the time-ordering to get

$$\tau(x_1, \ldots, x_n) = \langle \Omega | U^{-1}(t) T \left( \phi_{in}(x_1) \cdots \phi_{in}(x_n) U(t, t_1) U(t_1, t_2) \cdots U(t_{n-1}, t_n) U(t_n, -t) U(-t) \right) | \Omega \rangle$$

$$= \langle \Omega | U^{-1}(t) T \left( \phi_{in}(x_1) \cdots \phi_{in}(x_n) \exp \left[ -i \int_{-t}^{t} dt' H'_I(t') \right] \right) U(-t) | \Omega \rangle.$$  

(11.2.30)

Thus, except for the operators $U^{-1}(t)$ and $U(-t)$ we have expressed the function $\tau$ in terms of in-field operators. We next show that for $t \to \infty$ the state $|\Omega\rangle$ is an eigenstate of $U(-t) = U^{-1}(t)$ as well as $U(t)$. To see this, consider an arbitrary state $|\alpha, \vec{p}\rangle$ of Klein-Gordon particles containing at least one particle. In other words,

$$|\alpha, \vec{p}\rangle \neq |\Omega\rangle.$$  

(11.2.31)

Then,

$$\langle \alpha, \vec{p}; \text{in} | U(-t) | \Omega \rangle = \langle \alpha; \text{in} | a_{in}(\vec{p}) U(-t) | \Omega \rangle$$

$$= -i \int d^3 x f_p^* (\vec{x}, -t') \frac{\partial}{\partial t} \langle \alpha; \text{in} | \phi_{in}(\vec{x}, -t') U(-t) | \Omega \rangle$$

$$= -i \int d^3 x f_p^* (\vec{x}, -t') \frac{\partial}{\partial t'} \langle \alpha; \text{in} | U(-t') \phi(\vec{x}, -t') U^{-1}(-t') U(-t) | \Omega \rangle.$$  

(11.2.32)

We now carry out the indicated differentiation explicitly and then let $t = t' \to \infty$ to get

$$U^{-1}(-t') U(-t) = 1$$  

(11.2.33)

and

$$\lim_{t \to \infty} \phi(\vec{x}, -t) = \phi_{in}(\vec{x}, -t).$$  

(11.2.34)

So,

$$\langle \alpha, \vec{p}; \text{in} | U(-t) | \Omega \rangle = \langle \alpha; \text{in} | U(-t) a_{in}(\vec{p}) | \Omega \rangle$$

$$+ i \int d^3 x f_p^* (\vec{x}, -t) \left[ \langle \alpha; \text{in} | \hat{U}(-t) \phi(\vec{x}, -t) \right.$$  

$$+ U(-t) \phi(\vec{x}, -t) \hat{U}^{-1}(-t) U(-t) | \Omega \rangle \right].$$  

(11.2.35)
Now,
\[ a_{\text{in}}(\vec{p})|\Omega\rangle = 0 . \]  
(11.2.36)

Also,
\[
\dot{U} \phi + U \phi \dot{U}^{-1} U = \dot{U} U^{-1} \phi_{\text{in}} U + \phi_{\text{in}} U \dot{U}^{-1} U \\
= + \dot{U} U^{-1} \phi_{\text{in}} U - \phi_{\text{in}} U \dot{U}^{-1} U \\
= [\dot{U} U^{-1}, \phi_{\text{in}}]|U\rangle . 
\]  
(11.2.37)

But,
\[ \dot{U} U^{-1} = -i H_{\text{f}}(\phi_{\text{in}}) . \]  
(11.2.38)

Therefore,
\[
\dot{U} \phi + U \phi \dot{U}^{-1} U = -i[H_{\text{f}}(\phi_{\text{in}}), \phi_{\text{in}}]|U\rangle = 0 
\]  
(11.2.39)

since the commutator involved is an equal time commutator and we have assumed no derivative coupling. This means that
\[ \lim_{t \to \infty} \langle \alpha, \vec{p}; \text{in}|U(-t)|\Omega\rangle = 0 . \]  
(11.2.40)

This furthermore implies that \( U(-t)|\Omega\rangle \) is orthogonal to any state that contains one or more particles. Thus,
\[ \lim_{t \to \infty} U(-t)|\Omega\rangle = \lambda_-|\Omega\rangle \quad |\lambda_-| = 1 . \]  
(11.2.41)

In a very similar manner one finds
\[ \lim_{t \to \infty} U(t)|\Omega\rangle = \lambda_+|\Omega\rangle \quad |\lambda_+| = 1 . \]  
(11.2.42)

Now in our expression for
\[ \langle \alpha, \vec{p}; \text{in}|U(-t)|\Omega\rangle \]
we have found that we get an expression of the form
\[ \langle \Omega|U^{-1}(t)T\left( \phi_{\text{in}}(x_1) \ldots \phi_{\text{in}}(x_n) \exp \left[ -i \int_{-t}^{t} H_{\text{f}}(t') dt' \right] \right) U(-t)|\Omega\rangle . \]

This can therefore be rewritten in the limit as \( t \to \infty \) as
\[ \lim_{t \to \infty} \langle \Omega|U^{-1}(t)|\Omega\rangle \langle \Omega|T(\cdots)|\Omega\rangle \langle \Omega|U(-t)|\Omega\rangle \\
= \lambda_+^* \lambda_- \lim_{t \to \infty} \langle \Omega|T(\cdots)|\Omega\rangle 
\]  
(11.2.43)

Now,
\[ \lambda_+^* \lambda_- = \lim_{t \to \infty} \langle \Omega|U^{-1}(t)|\Omega\rangle \langle \Omega|U(-t)|\Omega\rangle \\
= \lim_{t \to \infty} \langle \Omega|U(-t)U^{-1}(t)|\Omega\rangle \\
= \lim_{t \to \infty} \langle \Omega|U(-t, t)|\Omega\rangle \\
= \lim_{t \to \infty} \langle \Omega|T\left( \exp \left[ i \int_{-t}^{t} H_{\text{f}}(t') dt' \right] \right) |\Omega\rangle \\
= \lim_{t \to \infty} \langle \Omega|T\left( \exp \left[ -i \int_{-t}^{t} H_{\text{f}}(t') dt' \right] \right) |\Omega\rangle^{-1} 
\]  
(11.2.44)
Thus,
\[
\tau(x_1, \ldots, x_n) = \lim_{t \to \infty} \frac{\langle \Omega | T \left( \phi_{in}(x_1) \ldots \phi_{in}(x_n) \exp \left[ -i \int_{-t}^{t} H_I(t') dt' \right] \right) | \Omega \rangle}{\langle \Omega | T \left( \exp \left[ -i \int_{-t}^{t} H_I(t') dt' \right] \right) | \Omega \rangle}
\]
(11.2.45)

As a last step we cancel the c-number factor \( \exp \left[ -i \int_{-t}^{t} E_0(t') dt' \right] \) and replace \( H_I \) by \( H_I \). Thus, the final result is
\[
\tau(x_1, \ldots, x_n) = \lim_{t \to \infty} \frac{\langle \Omega | T \left( \phi_{in}(x_1) \ldots \phi_{in}(x_n) \exp \left[ -i \int_{-t}^{t} H_I(t') dt' \right] \right) | \Omega \rangle}{\langle \Omega | T \left( \exp \left[ -i \int_{-t}^{t} H_I(t') dt' \right] \right) | \Omega \rangle}
\]

\[
= \frac{\sum_{r=0}^{\infty} \frac{(-i)^r}{r!} \int_{-\infty}^{\infty} d^4 y_1 \ldots d^4 y_r \langle \Omega | T \left( \phi_{in}(x_1) \ldots \phi_{in}(x_n) H_I(y_1) \ldots H_I(y_r) \right) | \Omega \rangle}{\sum_{r=0}^{\infty} \frac{(-i)^r}{r!} \int_{-\infty}^{\infty} d^4 y_1 \ldots d^4 y_r \langle \Omega | T \left( H_I(y_1) \ldots H_I(y_r) \right) | \Omega \rangle}
\]
(11.2.46)

Here we have used the notation that
\[
\mathcal{H}_I(y) = \mathcal{H}_I(\phi_{in}(y), \pi_{in}(y))
\]
(11.2.47)

We now have obtained an expression for the S-matrix
\[
S_{\beta \alpha} = \langle \beta; \text{out} | \alpha; \text{in} \rangle
\]
of the form
\[
S_{\beta \alpha} = \left( \frac{i}{\sqrt{Z}} \right)^{n+m} \prod_{i=1}^{n} \prod_{j=1}^{m} \int d^4 x_i d^4 y_j f_{p_i}(x_i) f_{q_j}^*(y_j)
\]
(\(\Box_i + m^2)(\Box_j + m^2)\langle \Omega | T (\phi(y_j) \phi(x_i)) | \Omega \rangle
\] + elastic term
(11.2.48)

Furthermore, we have found a way to express
\[
\langle \Omega | T \left( \prod_{i=1}^{n} \phi(x_i) \right) | \Omega \rangle
\]
in terms of
\[
\langle \Omega | T \left( \prod_{i=1}^{n} \phi_{in}(x_i) \right) | \Omega \rangle .
\]

To evaluate such expressions we rewrite them in terms of normal ordered products. In this manner we get Feynmann propagators \( \Delta_F \) and \( S_F \) which are nothing other than the VEV's of two free fields. Thus, one writes
\[
\langle \Omega | T (\phi_{in}(x) \phi_{in}(y)) | \Omega \rangle = i \Delta_F(x - y)
\]
\[
\langle \Omega | T (\psi_{in, \alpha}(x) \bar{\psi}_{in, \beta}(y)) | \Omega \rangle = i S_{F, \alpha \beta}(x - y).
\]
(11.2.49)

The final result for such a rewriting is called Wick's Theorem.
11.3 Wick’s Theorem

We first quote the theorem and then prove it by induction.

\[ T(\phi_{in}(x_1) \ldots \phi_{in}(x_n)) =: \phi_{in}(x_1) \ldots \phi_{in}(x_n) : \]
\[ + \bra{\Omega} T(\phi_{in}(x_1)\phi_{in}(x_2)) \ket{\Omega} : \phi_{in}(x_3) \ldots \phi_{in}(x_n) : + \text{ all permutations} \]
\[ + \bra{\Omega} T(\phi_{in}(x_1)\phi_{in}(x_2)) \ket{\Omega} \bra{\Omega} T(\phi_{in}(x_3)\phi_{in}(x_4)) \ket{\Omega} : \phi_{in}(x_5) \ldots \phi_{in}(x_n) : + \text{ all permutations} \]
\[ + \ldots \]
\[ + \bra{\Omega} T(\phi_{in}(x_1)\phi_{in}(x_2)) \ket{\Omega} \ldots \]
\[ \bra{\Omega} T(\phi_{in}(x_{n-1})\phi_{in}(x_n)) \ket{\Omega} \]
if \( n \) is even, or
\[ + \bra{\Omega} T(\phi_{in}(x_1)\phi_{in}(x_2)) \ket{\Omega} \ldots \]
\[ \bra{\Omega} T(\phi_{in}(x_{n-2})\phi_{in}(x_{n-1})) \ket{\Omega} \phi_{in}(x_n) \]
if \( n \) is odd \hspace{1cm} (11.3.50)

The vacuum expectation values (VEV’s) are also called “contractions” and are
also written as follows.

\[ \bra{\Omega} T(\phi_{in}(x)\phi_{in}(y)) \ket{\Omega} = \phi_{in}(x) \ldots \phi_{in}(y) \ldots \] \hspace{1cm} (11.3.51)

We now prove the theorem by induction.

**Proof**

The result is obvious for \( n = 1 \).

For \( n = 2 \) we have

\[ T(\phi_{in}(x_1)\phi_{in}(x_2)) =: \phi_{in}(x_1)\phi_{in}(x_2) : + \text{ c-number} \] \hspace{1cm} (11.3.52)

But, the expectation value of a normal ordered product of field operators van-
ishes. Thus, we get that

\[ \text{c-number} = \bra{\Omega} T(\phi_{in}(x_1)\phi_{in}(x_2)) \ket{\Omega}. \] \hspace{1cm} (11.3.53)

This c-number is the Feynman propagator \( i\Delta_F(x_1 - x_2) \). Therefore,

\[ T(\phi_{in}(x_1)\phi_{in}(x_2)) \]
\[ = : \phi_{in}(x_1)\phi_{in}(x_2) : + \bra{\Omega} T(\phi_{in}(x_1)\phi_{in}(x_2)) \ket{\Omega} \] \hspace{1cm} (11.3.54)

For fermions the argument is the same up to this point. Thus, for a Dirac
field \( \psi \) we have

\[ T(\psi_{in}(x_1)\bar{\psi}_{in}(x_2)) \]
\[ = : \psi_{in}(x_1)\bar{\psi}_{in}(x_2) : + \bra{\Omega} T(\psi_{in}(x_1)\bar{\psi}_{in}(x_2)) \ket{\Omega} \] \hspace{1cm} (11.3.55)
where again the last term on the right hand side is just the Feynman propagator $iS_F(x_1 - x_2)$. To continue with the proof we now assume that the theorem holds for the product of $n$ fields. Then, in the time-ordered product

$$T(\phi_{in}(x_1) \ldots \phi_{in}(x_{n+1}))$$

we can, since the product is already time-ordered, choose $t_{n+1} < t_j$ for all $j = 1, 2, \ldots, n$. In that case we have

$$T(\phi_{in}(x_1) \ldots \phi_{in}(x_{n+1})) = T(\phi_{in}(x_1) \ldots \phi_{in}(x_n)) \phi_{in}(x_{n+1})$$

$$= : \phi_{in}(x_1) \ldots \phi_{in}(x_n) : \phi_{in}(x_{n+1})$$

$$+ \sum \langle \Omega | T(\phi_{in}(x_1)\phi_{in}(x_2)) | \Omega \rangle : \phi_{in}(x_3) \ldots \phi_{in}(x_n) : \phi_{in}(x_{n+1})$$

$$+ \ldots$$

(11.3.56)

To complete the proof we need to insert $\phi_{in}(x_{n+1})$ inside the normal ordering. Now,

$$: \phi_{in}(x_3) \ldots \phi_{in}(x_n) := \sum_{A,B} \delta_P \prod_{i \in A} \phi_{in}^{(-)}(x_i) \prod_{j \in B} \phi_{in}^{(+)}(x_j).$$

(11.3.57)

Here, we have introduced the two index sets

$$A = \{p_1, p_2, \ldots, p_r\} \quad B = \{p_{r+1}, p_{r+2}, \ldots, p_n\}$$

(11.3.58)

where each index appears once and only once. Also the factor $\delta_P$ is given by

$$\delta_P = \begin{cases} 
1 & \text{for Bosons} \\
-1 & \text{for Fermions}
\end{cases}$$

(11.3.59)

We have to sum over all index sets $A$, $B$, that is over all permutations. Therefore,

$$: \phi_{in}(x_1) \ldots \phi_{in}(x_n) : \phi_{in}(x_{n+1}) = \sum_{A,B} \delta_P \prod_{i \in A} \phi_{in}^{(-)}(x_i) \prod_{j \in B} \phi_{in}^{(+)}(x_j)$$

$$\left[ \phi_{in}^{(+)}(x_{n+1}) + \phi_{in}^{(-)}(x_{n+1}) \right]$$

$$= \sum_{A,B} \delta_P \prod_{i \in A} \phi_{in}^{(-)}(x_i) \prod_{j \in B} \phi_{in}^{(+)}(x_j) \phi_{in}^{(+)}(x_{n+1})$$

$$+ \sum_{A,B} \delta_P \prod_{i \in A} \phi_{in}^{(-)}(x_i) \phi_{in}^{(-)}(x_{n+1}) \prod_{j \in B} \phi_{in}^{(+)}(x_j)$$

$$+ \sum_{A,B} \delta_P \prod_{i \in A} \phi_{in}^{(-)}(x_i) \prod_{j \in B, j \neq k} \phi_{in}^{(+)}(x_j) \langle \Omega | \phi_{in}^{(+)}(x_k) \phi_{in}^{(-)}(x_{n+1}) | \Omega \rangle .$$

(11.3.61)
Here, $\delta_{P^r}$ is the sign of the permutation appropriate for these terms. In absorbing $\phi_{in}(x_{n+1})$ into the normal ordered product we have obtained a string of terms with either commutators (for bosons) or anticommutators (for fermions) between $\phi_{in}^{(-)}(x_{n+1})$ and $\phi_{in}^{(+)}(x_k)$ $k \in B$. These (anti)commutators are c-numbers and we have written them as vacuum expectation values $\langle \Omega | \cdot | \Omega \rangle$. But,

$$\langle \Omega | \phi_{in}^{(+)}(x_k)\phi_{in}^{(-)}(x_{n+1}) | \Omega \rangle = \langle \Omega | \phi_{in}(x_k)\phi_{in}(x_{n+1}) | \Omega \rangle = \langle \Omega | T(\phi_{in}(x_k)\phi_{in}(x_{n+1})) | \Omega \rangle$$

(11.3.62)

since $t_{n+1} < t_k$. This proves the result both for bosons and fermions.

### 11.4 Graphical Representation

We have three types of terms.

1) Real scalar (or pseudoscalar) field with mass $\mu$. Then,

$$\langle \Omega | T(\phi_{in}(x)\phi_{in}(y)) | \Omega \rangle = i\Delta_F(x - y; \mu^2) .$$

(11.4.63)

This propagator is represented by figure 11.2.

2) Complex scalar (or pseudoscalar) field with mass $\mu$. Then,

$$\langle \Omega | T(\phi_{in}(x)\phi_{in}^\dagger(y)) | \Omega \rangle = i\Delta_F(x - y; \mu^2) .$$

(11.4.64)

This propagator is represented by figure 11.3.

3) Dirac field of mass $m$. Then,

$$\langle \Omega | T(\psi_{in,\alpha}(x)\bar{\psi}_{in,\beta}(y)) | \Omega \rangle = iS_F(x - y; m)_{\alpha,\beta} .$$

(11.4.65)

This propagator is represented by figure 11.4.

In using these expressions to evaluate the $S$-matrix only the contractions survive since $\langle \Omega | : \cdot : \Omega \rangle = 0$. Furthermore, $\mathcal{H}_f(\phi_{in}(x))$ contains a normal ordered product of field operators at the same point $x$. The propagators, that

```
\[ x \]
```

```
\[ y \]
```

Figure 11.2: Propagator for a Real Scalar or Pseudoscalar Field with mass $\mu$. 
arise from contractions of these operators with other operators, always have one point, namely \( x \), in common. They are joined at such a point, called a \textit{vertex}. We also have that

\[
\langle \Omega | \phi_{\text{in}}(x_1) \ldots \phi_{\text{in}}(x_n) | \Omega \rangle = 0 \text{ if } n \text{ is odd} \tag{11.4.66}
\]

or

\[
\langle \Omega | \phi_{\text{in}}(x_1) \ldots \phi_{\text{in}}(x_n) | \Omega \rangle =\sum_{\text{perms}} \prod \langle \Omega | \phi_{\text{in}}(x_i) \phi_{\text{in}}(x_j) | \Omega \rangle \text{ if } n \text{ is even} . \tag{11.4.67}
\]

To see how this works we consider some simple examples. Thus,

\[
\langle \Omega | T (\phi_{\text{in}}(x_1) \phi_{\text{in}}(x_2) : \phi_{\text{in}}^2(x) :) | \Omega \rangle = 2\langle \Omega | T (\phi_{\text{in}}(x_1) \phi_{\text{in}}(x_2)) | \Omega \rangle \langle \Omega | T (\phi_{\text{in}}(x) \phi_{\text{in}}(x_2)) | \Omega \rangle . \tag{11.4.68}
\]

This corresponds to the single diagram figure 11.5. There is no diagram of the form figure 11.6 since

\[
\langle \Omega | T (\phi_{\text{in}}(x_1) \phi_{\text{in}}(x_2)) | \Omega \rangle \langle \Omega | T (\phi_{\text{in}}^2(x) :) | \Omega \rangle = 0 \tag{11.4.69}
\]

due to the normal ordering.
Similarly, we have that

\[ \langle \Omega | T(\phi_{in}(x_1)\ldots\phi_{in}(x_4)) | \Omega \rangle \langle \Omega | T(\phi_{in}^4(x)) | \Omega \rangle = 0. \]  

But,

\[
\begin{align*}
\langle \Omega | T(\phi_{in}(x_1)\ldots\phi_{in}(x_4)) :\phi_{in}^4(x) : | \Omega \rangle & = 4! \langle \Omega | T(\phi_{in}(x_1)\phi_{in}(x)) | \Omega \rangle \ldots \langle \Omega | T(\phi_{in}(x_4)\phi_{in}(x)) | \Omega \rangle.
\end{align*}
\]

This corresponds to the diagram figure 11.7 and illustrates how a vertex is associated with an interaction term. It is also worth noting that due to the normal ordering we do not get any terms involving vacuum expectation values of time-ordered products of fields at the same point. These would be terms of the form

\[ \langle \Omega | T(\phi_{in}(x)\phi_{in}(x)) | \Omega \rangle \]

and lead to so-called "tadpole" diagrams of which one is illustrated in figure 11.8.
11.5 The Pion-Nucleon Interaction

The interaction hamiltonian density is given by

$$\mathcal{H}_I = G : \bar{\psi}_\text{in}(x) \gamma^5 \psi_\text{in}(x) : \phi_\text{in}(x).$$ (11.5.72)

The matrix element

$$\langle \Omega | T(\phi_\text{in}(x_1) \bar{\psi}_\text{in}(x_2) \psi_\text{in}(x_3) \mathcal{H}_I(x)) | \Omega \rangle.$$

is of order $G$ and represents the absorption or production of a pion. The vertex is shown in figure 11.9. To order $G^2$ we have a matrix element that represents

Figure 11.9: Vertex for Absorption or Production of a Pion
nucleon-nucleon scattering and is given by
\[ \langle \Omega | T(\bar{\psi}_i(x_1)\psi_i(x_2)\mathcal{H}_I(x)\mathcal{H}_I(y))|\Omega \rangle . \]
When written out with the contractions indicated we get
1. contractions between \( \bar{\psi}_i(x_1) \) and \( \psi_i(x) \).
2. contractions between \( \psi_i(x_1) \) and \( \bar{\psi}_i(y) \).
3. contractions between \( \bar{\psi}_i(x) \) and \( \psi_i(y) \).
4. contractions between \( \phi_i(x) \) and \( \phi_i(y) \).

So, the expression looks like
\[
G^2 \langle \Omega | T \left( \bar{\psi}_i(x_1) \gamma^\alpha_\beta : \psi_i(x) \bar{\psi}_i(y) : \gamma^\rho_\sigma \psi_i(y) \psi_i(x_2) \bar{\psi}_i(y) : \right) \times \phi_i(x) \phi_i(y) | \Omega \rangle .
\] (11.5.73)
The corresponding diagrams are figure 11.10.

![Figure 11.10: Second Order Pion Nucleon Scattering](image)

11.6 Self-Interacting Boson Field

Here the interaction hamiltonian density is given by
\[
\mathcal{H}_I = -\frac{1}{4}\lambda_0 : \phi_i^4(x) : -\frac{1}{2}\delta \mu^2 : \phi_i^2(x) : 
\] (11.6.74)
where
\[
\delta \mu^2 = \mu^2 - \mu_0^2
\] (11.6.75)
is the mass shift. The vertices corresponding to this interaction are shown in the figure below. A line leaving or entering a vertex represents a contraction (propagator).

\[-\frac{1}{4}\lambda_0 : \phi_{in}^4 : (x) \quad + \quad \frac{1}{2} \delta \mu^2 : \phi_{in}^2 : (x)\]

Figure 11.11: Interaction Vertices for Self-Interacting Boson Field

Then to second order in $\lambda_0$ and first order in $\delta \mu^2$ we get a term in $\tau(x_1, \ldots, x_4)$ of the form

\[
\frac{\lambda_0^3 \delta \mu^2}{2(4)^2} \langle \Omega | T(\phi_{in}(x_1)\phi_{in}(x_2)\phi_{in}(x_3)\phi_{in}(x_4)) \times : \phi_{in}^4(y_1) :: \phi_{in}^2(y_2) :: \phi_{in}^2(y_3) : \Omega \rangle
\]

(11.6.76)

So to draw the corresponding diagrams we need to have 4 lines ending at $x_1, x_2, x_3, x_4$ as indicated in figure 11.12. Between these four lines we need:

\[
\begin{align*}
\text{Figure 11.12: Meson-Meson Interaction} \\
\text{Four lines coming from } y_1 \text{ as in figure 11.13.} \\
\text{Four lines coming from } y_2 \text{ as in figure 11.14.} \\
\text{Two lines coming from } y_3 \text{ as in figure 11.15.}
\end{align*}
\]

This means that the possible graphs are as shown in figure 11.16 plus so-called disconnected graphs as shown in figure 11.17. These disconnected graphs contain vacuum diagrams. An example of such a vacuum diagram is shown in figure 11.18. The difference is that they have no external lines. Such bubble
or vacuum diagrams also occur if one expands the denominator in the expression for the $S$-matrix and they cancel exactly between the numerator and the denominator.

Any subgraph which is not connected in any way to an external line is called the disconnected part of the graph. A graph with no disconnected parts is called connected. Any graph in the numerator can be uniquely separated into a connected and a disconnected part. The same is true for the contributions to the $\tau$ function.

Consider all the graphs whose connected parts are of order $s$ in $\mathcal{H}_I$ and which occur in the numerator for $\tau$. The numerator then becomes
Figure 11.16: Possible Graphs for Self-Interacting Bose Field

Figure 11.17: Disconnected Graphs

Figure 11.18: Vacuum Graph

\[ N = \sum_{p=0}^{\infty} \frac{(-i)^p}{p!} \int d^4y_1 \ldots d^4y_p \langle \Omega | T(\phi_{in}(x_1) \ldots \phi_{in}(x_n) \mathcal{H}_1(y_1) \ldots \mathcal{H}_1(y_s)) | \Omega \rangle \]
\[ \times \left( \frac{p}{s} \right) \langle \Omega | T(\mathcal{H}_I(y_{s+1}) \ldots \mathcal{H}_I(y_p)) | \Omega \rangle \]  

(11.6.77)

Here,

\[ \left( \frac{p}{s} \right) = \frac{p!}{s!(p-s)!} \]  

(11.6.78)

is the number of ways of selecting \( s \) terms from a set of \( p \). Thus, letting \( r = p-s \) we get

\[
N = \frac{(-i)^s}{s!} \int d^4y_1 \ldots d^4y_s \langle \Omega | T(\phi_{in}(x_1) \ldots \phi_{in}(x_n) \mathcal{H}_I(y_1) \ldots \mathcal{H}_I(y_s)) | \Omega \rangle_{\text{connected}} \\
\times \sum_{r=0}^{\infty} \frac{(-i)^r}{r!} \int d^4z_1 \ldots d^4z_r \langle \Omega | T(\mathcal{H}_I(z_1) \ldots \mathcal{H}_I(z_r)) | \Omega \rangle .
\]  

(11.6.79)

This means that we have

\[ N = \text{(connected graphs)} \times \left( \sum \text{bubbles} \right) \]  

(11.6.80)

and the \( \left( \sum \text{bubbles} \right) \) cancels with the denominator. Symbolically, if we denote all graphs by \( G \), connected graphs by \( G^C \) and disconnected graphs by \( D \) we can write this as

\[
\tau = \frac{\sum_i G_i(x_1, \ldots, x_n)}{\sum_k D_k} \\
= \frac{\left( \sum_i G_i^C(x_1, \ldots, x_n) \right) \sum_k D_k}{\sum_k D_k} \\
= \sum_i G_i^C(x_1, \ldots, x_n).
\]  

(11.6.81)

Thus, \( \tau(x_1, \ldots, x_n) = \text{sum of the contributions of all connected Feynman graphs.} \)

We are now in a position to write down the Feynman rules for evaluating any graph.

### 11.7 Feynman Rules: External Field Problems

We describe the process of interest pictorially on a space-time diagram. An electron is represented by a solid line with an arrow pointing forward in time, that is up to the left or right. See figure 11.19. A positron is represented by a solid line with an arrow pointing backward in time, down and to the left or right. See figure 11.20 The external electromagnetic field is represented by a wiggly line terminating in an \( \times \) on one side and an electron-electron or positron-positron vertex at the other end as shown in figures 11.21 or else terminating in an electron-positron vertex as shown in figures 11.22. The first two graphs correspond to electron or positron scattering. The last two graphs correspond
to pair creation or annihilation. These various vertices may be pieced together to produce various other graphs.

The rules for computation are as follows:
1. For a process to be computed up to order $n$ draw all the connected graphs having the same number of vertices starting with zero vertices up to $n$ vertices.
2. For each vertex at a point $x$ insert a factor $-ie\gamma_\mu \int d^4x$
3. For each fermion line going from $x$ to $y$ insert a factor $iS_F(x-y)$.
   For bosons insert a factor $i\Delta_F(x-y)$.
4. For each initial external electron line starting at $x$ insert a factor $U_{ps}(x)$.
   For each external initial positron line starting at $x$ insert a factor $\tilde{V}_p(x)$. For
each final external electron line starting at $x$ insert a factor $\bar{U}_{ps}(x)$.
For each external final positron line starting at $x$ insert a factor $V_{ps}(x)$.
5. Insert a relative minus sign between two terms which correspond to the same
graph with an interchange of identical fermions as in the graphs in figure 11.23
6. For each closed fermion loop insert a factor $(-1)$.

\[ x_3, \bar{p}_3, s_3 \quad x_4, \bar{p}_4, s_4 \quad x_3, \bar{p}_3, s_3 \quad x_4, \bar{p}_4, s_4 \]
\[ x_1, p_1, s_1 \quad x_2, \bar{p}_2, s_2 \quad x_2, \bar{p}_2, s_2 \quad x_1, p_1, s_1 \]

Figure 11.23: Two Topologically Equivalent Feynman Diagrams

7. For each external electromagnetic field line connecting to a vertex at $x$ insert
a factor $A^\mu(x)$.

To illustrate these rules in lowest order we recompute some of the results we
obtained earlier.

### 11.8 Electron Scattered by External EM Field

The relevant Feynman diagram is, as we stated earlier, as shown in figure 11.24.
Using the rules above we immediately obtain that the amplitude for this process is

\[
-i e \int d^4x \sqrt{\frac{m^2}{E(\bar{p})E(q)}} \frac{\bar{u}(\bar{p}, r)}{\sqrt{(2\pi)^3}} e^{ipx} \gamma^\mu A_\mu \frac{u(q, r')}{\sqrt{(2\pi)^3}} e^{-iqx}
\]

\[
= -ie \frac{m^2}{(2\pi)^3} \sqrt{\frac{1}{E(\bar{p})E(q)}} \bar{u}(\bar{p}, r) \gamma^\mu u(q, r') \tilde{A}_\mu (p - q) .
\]  \hspace{1cm} (11.8.82)
This is the same result that we obtained by somewhat more tedious means in section 9.7. Using Feynman diagrams we were able to immediately write down the amplitude for this process. The same follows for positron scattering below.

The corresponding Feynman diagram for positron scattering is shown in figure 11.25. The amplitude for positron scattering can now be written down.

\[ -ie \int d^4x \sqrt{\frac{m^2}{E(p)E(q)}} \frac{\bar{v}(q,s)}{\sqrt{(2\pi)^3}} e^{-iqpx} \gamma^\mu A_\mu \frac{v(q,r)}{\sqrt{(2\pi)^3}} e^{ipx} \]

\[ = -\frac{ie}{(2\pi)^3} \sqrt{\frac{m^2}{E(p)E(q)}} \bar{v}(q,s) \gamma^\mu v(q,r) A_\mu(p-q) . \quad (11.83) \]

### 11.9 Pair Production

This corresponds to the calculation we did in section 9.9. The appropriate Feynman diagram is shown in figure 11.26. The corresponding S-matrix element is

\[ -ie \int d^4x \sqrt{\frac{m^2}{E(p)E(q)}} \frac{\bar{u}(p,r)}{\sqrt{(2\pi)^3}} e^{ipx} \gamma^\mu A_\mu \frac{u(q,s)}{\sqrt{(2\pi)^3}} e^{iqx} \]

\[ = -\frac{ie}{(2\pi)^3} \sqrt{\frac{m^2}{E(p)E(q)}} \bar{u}(p,r) \gamma^\mu v(q,s) A_\mu(p+q) . \quad (11.94) \]
Clearly, it follows from looking at several such processes that we can also write down Feynman rules in momentum space. We omit doing this.

11.10 Vacuum Polarization

In section 9.10 we carried out this calculation in detail. Here we can write down the expression for this process directly. However the tedious part of the computation still has to be done. The appropriate Feynman diagram is shown in figure 11.27 The corresponding $S$-matrix element can now be written down.

\[
S_{fi} = (-ie)^2 \int d^4x d^4y \gamma_\alpha^\mu S_{\mu\beta\gamma}(x - y) \gamma_\delta^\nu S_{\delta\alpha}(y - x) A_\mu(x) A_\nu(y). \tag{11.10.85}
\]

11.11 Second Order Scattering

As a final example we consider scattering to second order of an electron by an external electromagnetic field. This would have been a fairly tedious computation using the methods of chapter 9. Here we simply write down the amplitude. The appropriate diagram is shown in figure 11.28. The corresponding $S$-matrix
element is

\[
S_{fi} = (-ie)^2 \int d^4xd^4y \sqrt{\frac{m^2}{E(p)E(q)}} \frac{\bar{u}(\vec{p}, r)}{\sqrt{(2\pi)^3}} e^{ipx} \gamma^\mu A_\mu(x) \\
\times S_F(x - y) \gamma^\nu A_\nu(x) \frac{u(\vec{q}, r')}{\sqrt{(2\pi)^3}} e^{-iqy}
\]

\[
= -\frac{e^2}{(2\pi)^3} \sqrt{\frac{m^2}{E(p)E(q)}} \bar{u}(\vec{p}, r)\gamma^\mu \hat{S}_F(k)\gamma^\nu u(\vec{q}, r')\hat{A}_\mu(p - k)\hat{A}_\mu(k - q).
\]

(11.11.86)

### 11.12 Conclusion

The techniques we have developed suffice for computing the amplitudes for any process involving external fields. We have only touched on the subject of renormalization. In fact, we did not even consider the quantized electromagnetic field. The more modern approach to quantum electrodynamics is to view the renormalized perturbation series, rather than the Heisenberg equations with the canonical equal time (anti)commutators as the initial condition, as defining the theory. This pragmatic approach obviates many painful questions.

Another topic we omitted is regularization. A discussion of this would have taken us far afield.

Finally, the reader should be aware that many, if not most, computations these days begin with functional integral techniques. These path integral techniques provide a tremendously important computational tool. However, it is my opinion that for fundamental understanding a study of the canonical formalism will continue to provide a source of new breakthroughs.
11.13 Problems

1. Show that all 16 bilinear Dirac field operators \( \bar{\psi}(x)\Gamma\psi(x) \) commute for space-like separation.

2. For a free Dirac field compute

\[
\langle \Omega | \psi(x)\psi(y)\psi(z)|\Omega \rangle \quad \langle \Omega | \psi(x)\bar{\psi}(y)\bar{\psi}(z)\psi(w)|\Omega \rangle
\]

3. For a hermitean scalar field with lagrange density

\[
\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 + \frac{1}{4} g \phi^4 :
\]
compute the \( S \)-operator to order \( g \). Also compute the probability amplitude for the transition

\[ |\bar{\phi} \rangle \rightarrow |\bar{\tilde{\phi}}, \tilde{k}, \tilde{\sigma} \rangle \]

4. For a hermitean scalar field interacting with a scalar c-number source, compute the probability amplitude for the transition

\[ |\bar{\phi} \rangle \rightarrow |\bar{\tilde{\phi}}, \tilde{k} \rangle \]

5. Compute the differential cross-section for scattering of a positron off a Coulomb potential

\[ V(r) = -\frac{Ze}{r} \]

6. Find a term to describe an electrostatic dipole interaction for a quantized Dirac particle and discuss the invariance properties of such an interaction under Lorentz transformations, parity, and time reversal.

7. Find the configuration space as well as the momentum space representation of the Feynman propagator

\[ i\Delta_F(x - y) = \langle \Omega | T(\phi(x)\phi(y))|\Omega \rangle \]
for a free massive scalar field \( \phi \).

Hint: It may be useful to recall that

\[ i\Delta(x - y) = \langle \Omega | \phi(x)\phi(y)|\Omega \rangle = i\Delta^{(+)}(x - y) + i\Delta^{(-)}(x - y) \]

(11.13.87)

and first see that

\[ \Delta_F(x - y) = \theta(x - y)\Delta^{(+)}(x - y) - \theta(y - x)\Delta^{(-)}(x - y) \]

8. Find the configuration space as well as the momentum space representation of the Feynman propagator

\[ iS_F(x - y) = \langle \Omega | T(\psi(x)\bar{\psi}(y))|\Omega \rangle \]
for a free massive Dirac field \( \psi \). See the previous problem for hints.
Index

action
  scalar field, 81
antineutrino
  state, 76
antiparticle, 109
asymptotic condition
  weak, 145
asymptotic field operator, 143
asymptotic free field, 145

bare mass, 144
boost, 5, 43, 45

Campbell-Baker-Hausdorff formula, 102
canonical commutation relations, 83
canonical momentum, 83
  charged scalar field, 110
canonical quantization, 84
Casimir operators, 4, 5
causality, 138
charge conjugate solution, 16
charge operator, 108, 109
charge renormalization, 137
chiral eigenvalue, 56
chirality, 56, 57
  negative, 76
  operator, 75
  positive, 76
chirality projection operators, 77
coherent state, 103
commutation relations
  canonical, 83
conjugate momentum
  density operator, 83
contraction, 167
contractions, 160

current
  conserved, 26, 106
electric, 109
  four-vector, 10
  Noether, 109
cyclic property
  of trace, 27
Dirac adjoint, 28, 34
dispersion relations, 138
Duffin-Kemmer equation, 25
electric current
  four-vector, 10
electromagnetic field tensor, 55
energy-momentum tensor
  Dirac field, 112
  scalar field, 89
Euler-Lagrange equation, 81
field operator
  asymptotic, 143
  positive frequency, 92
finestructure constant, 21
Fock ground state, 88
Fock space, 85
Frobenius
  method of, 69
gamma matrices
  explicit representation, 32
gauge invariant, 15
gauge transformation
  global, 105, 109
  local, 106
Gauss' Theorem, 94
generators, 4
graph

177
connected, 168
  disconnected part, 168
graphs
  disconnected, 167

hamiltonian
  scalar field, 88
hamiltonian density
  Dirac field, 113
  scalar field, 88
Heisenberg field, 98
helicity, 35

interpolating field, 98, 100

Kramers-Kronig relations, 138
Källén-Yang-Feldman equations, 111, 119

lagrange density
  Dirac field, 112, 113
  scalar field, 80
lagrangian
  scalar field, 80
Landau levels, 62
Legendre transformation, 85
Levi-Civita pseudo-tensor, 4
Lorentz covariant
  normalization, 33
Lorentz gauge, 110
Lorentz invariant
  measure, 13
Lorentz transformation, 3
  orthochronous, 9
  proper, 9
Lorentz-invariant
  step function, 12

magnetic moment
  anomalous, 58
Majorana representation, 52
mass
  bare, 144
  counter term, 154
  physical, 144
mass shift, 167
minimal coupling, 15, 54

Mott cross-section, 126

neutrino
  state, 76
Noether
  current, 109
  Theorem, 106
normal ordered
  hamiltonian, 89
number density, 108

particle equation, 83
Pauli
  terms, 58
Pauli-Lubanski vector, 5
pseudoscalar field, 9

quantization
  canonical, 84

regularization, 120
regularizing factor, 120
renormalization
  wave function, 144
rest frame, 6, 31
Rutherford formula, 126

scalar field, 9
Sommerfeld, 72
spacelike, 14
spin, 6
Spin and Statistics Theorem, 117
spin operator, 32
state
  left-handed, 77
  right-handed, 77
step function
  Lorentz-invariant, 12
tachyons, 6
tadpole diagram, 164
transformation law
  Dirac spinor, 40

Uehling effect, 141

vacuum
RELATIVISTIC QUANTUM MECHANICS AND INTRODUCTION TO QUANTUM FIELD THEORY

This invaluable textbook is divided into two parts. The first part includes a detailed discussion on the discrete transformations for the Dirac equation, as well as on the central force problem for the Dirac equation. In the second part, the external field problem is examined; pair production and vacuum polarization leading to charge renormalization are treated in detail.

Relativistic Quantum Mechanics and Introduction to Quantum Field Theory has arisen from a graduate course which the author taught for several years at the University of Alberta to students interested in particle physics and field theory.