KLEIN-GORDON EQUATION IN THE FESHBACH-VILLARS REPRESENTATION

When we write the Klein-Gordon equation in Schrödinger form, it takes the form of a matrix equation

\[
\left( i\hbar \frac{\partial}{\partial t} - H_f \right) \Psi = 0 \tag{1}
\]

where

\[
\Psi \equiv \begin{bmatrix} \phi \\ \chi \end{bmatrix} \tag{2}
\]

\[
H_f \equiv (\tau_3 + i\tau_2) \frac{p^2}{2m_0} + \tau_3 m_0 c^2 \tag{3}
\]

and the \( \tau_i \) are essentially the Pauli matrices

\[
\tau_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \tau_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \quad \tau_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \quad \mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4}
\]

In the non-relativistic limit, the \( \phi \) component represents a particle with positive charge and the \( \chi \) component a particle with negative charge.

\[
\Psi^{(+)} \rightarrow \frac{1}{\sqrt{L^3}} \begin{bmatrix} 1 \\ -\frac{v^2}{c^2} \end{bmatrix} \tag{5}
\]

\[
\Psi^{(-)} \rightarrow \frac{1}{\sqrt{L^3}} \begin{bmatrix} -\frac{v^2}{c^2} \\ 1 \end{bmatrix} \tag{6}
\]

The division into these two components isn’t exact, however, unless the particle’s velocity is actually zero. By applying yet another transformation, we can obtain a form of the wave function in which the separation is exact. This is known as the Feshbach-Villars representation. We define the operator
\[ U \equiv \frac{(m_0c^2 + E_p) - \tau_1 (m_0c^2 - E_p)}{\sqrt{4m_0c^2E_p}} \mathbf{1} \] (7)

Note that \( U \) is a \( 2 \times 2 \) matrix since it is defined in terms of the matrices in \( \mathbf{4} \).

We then apply the transformation on \( \Psi \) to get a new two-component vector \( \Phi \) as follows

\[ \Phi = U\Psi \] (8)

\[ \Phi^\dagger = \Psi^\dagger U^\dagger \] (9)

The operator \( U \) is not unitary, however, as \( U^\dagger \neq U^{-1} \). From its definition and the fact that \( \tau_1 \) and \( \mathbf{1} \) are both Hermitian, we see that \( U \) is Hermitian, so that \( U^\dagger = U \). Greiner shows that the inverse is given by

\[ U^{-1} = \tau_3 U \tau_3 \] (10)

We can work out \( U^{-1} \) explicitly by using the properties of the \( \tau_i \) matrices:

\[ \tau_i \tau_j = -\tau_j \tau_i = i \tau_k \quad i,j,k = 1,2,3 \text{ & cyclic perms.} \] (11)

\[ \tau_i^2 = \mathbf{1} \] (12)

In particular

\[ \tau_3 \tau_1 \tau_3 = i \tau_2 \tau_3 = -\tau_1 \] (13)

Therefore

\[ U^{-1} = \frac{(m_0c^2 + E_p) + \tau_1 (m_0c^2 - E_p)}{\sqrt{4m_0c^2E_p}} \mathbf{1} \] (14)

and the fact that \(UU^{-1} = \mathbf{1} \) can be verified by direct multiplication, shown in Greiner’s eqn 1.98.

We can now examine the behaviour of \( \Phi \) for a free particle. In the earlier Schrödinger form, the free particle with positive charge is represented by

\[ \Psi^{(+)} = \frac{1}{\sqrt{4m_0c^2 \sqrt{L^3E_p}}} \begin{bmatrix} m_0c^2 + E_p \\ m_0c^2 - E_p \end{bmatrix} \exp \left[ \frac{i}{\hbar} (\mathbf{p} \cdot \mathbf{x} - E_p t) \right] \] (15)

Working out \( \Phi^{(+)} = U\Psi^{(+)} \) by direct multiplication (Greiner’s eqn 1.99) yields

Greiner’s eqn 1.95 is incorrect; the \( \Phi \) on the RHS of each equation should be \( \Psi \).
\[
\Phi^{(+) \downarrow} = \frac{1}{\sqrt{L^3}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \exp \left[ \frac{i}{\hbar} (p \cdot x - E_p t) \right] \tag{16}
\]

Thus \( \Phi^{(+)} \), which represents a positively charged particle, is given by a vector entirely localized to the top component. This is true for all particle speeds, not just in the nonrelativistic limit, so it is true for a relativistic particle as well.

By doing a similar calculation (Greiner eqn 1.100) we get the result for the negative particle

\[
\Phi^{(- \downarrow)} = U \Psi^{(- \downarrow)} = \frac{1}{\sqrt{L^3}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \exp \left[ \frac{i}{\hbar} (p \cdot x + E_p t) \right] \tag{17}
\]

so \( \Phi^{(-)} \) is localized to the lower component.

As Greiner shows in eqns 1.101 to 1.102, the normalization condition on the original \( \Psi \) vectors remains unchanged when applied to the \( \Phi \) representation:

\[
\int \Psi^\dagger \tau_3 \Psi d^3 x = \int \Phi^\dagger \tau_3 \Phi d^3 x = \pm 1 \tag{18}
\]

We can define a generalized scalar product, or \('\Phi \text{ product}'\) as

\[
\langle \Psi | \Psi' \rangle_\Phi \equiv \int \Psi^\dagger \tau_3 \Psi' d^3 x \tag{19}
\]

This scalar product is the same as that for an ordinary two-component vector except for the insertion of the \( \tau_3 \) matrix within the integral. As with \( \Phi \), the \( \Phi \text{ product} \) is invariant under the transformation \( \Phi \) as can be shown by following the steps in Greiner’s eqn 1.101 for two different vectors. The result is

\[
\langle \Psi | \Psi' \rangle_\Phi = \langle \Phi | \Phi' \rangle_\Phi \tag{20}
\]

Suppose there is an operator \( A \) which satisfies the condition

\[
\langle \Psi | \Psi' \rangle_\Phi = \langle A \Phi | A \Phi' \rangle_\Phi \tag{21}
\]

In this case, we have

Greiner’s eqn 1.103 has a typo in that the second \( \Psi \) on the RHS should be \( \Psi' \).
\[ \langle \Psi | \Psi' \rangle_{\Phi} = \int \Psi^\dagger \tau_3 \Psi' d^3 x \quad (22) \]

\[ \langle A\Phi | A\Phi' \rangle_{\Phi} = \int (A\Phi)^\dagger \tau_3 A\Phi' d^3 x \]

\[ = \int \Phi^\dagger A^\dagger \tau_3 A\Phi' d^3 x \quad (24) \]

If we require \(21\) then from \(20\) we have

\[ \int \Phi^\dagger A^\dagger \tau_3 A\Phi' d^3 x = \langle \Psi | \Psi' \rangle_{\Phi} \quad (25) \]

\[ = \langle \Phi | \Phi' \rangle_{\Phi} \quad (26) \]

\[ = \int \Phi^\dagger \tau_3 \Phi' d^3 x \quad (27) \]

Therefore

\[ A^\dagger \tau_3 A = \tau_3 \quad (28) \]

Multiply by \(A^{-1}\) on the right and then by \(\tau_3\) on the left to get

\[ \tau_3 A^\dagger \tau_3 = A^{-1} \quad (29) \]

Thus \(A\) is not unitary in the ordinary sense (which would require \(A^\dagger = A^{-1}\)), but if \(A\) satisfies the condition \(29\) it is called '\(\Phi\) unitary'. Note that if \([\tau_3, A] = 0\), then \(A^\dagger = A^{-1}\) and \(A\) is a regular unitary operator.

Because of the normalization condition \(18\), the charge \(Q\) of a state \(\Psi\) is

\[ Q = e \int \Psi^\dagger \tau_3 \Psi d^3 x = e \int \Phi^\dagger \tau_3 \Phi d^3 x \quad (30) \]

This is generalized so that we can define the mean or expectation value of an operator \(L\) by

\[ \langle L \rangle = \int \Psi^\dagger \tau_3 L \Psi d^3 x \quad (31) \]

If \(L\) represents an observable, then \(\langle L \rangle\) must be a real number, so we must have

\[ \left( \int \Psi^\dagger \tau_3 L \Psi d^3 x \right)^\dagger = \int \Phi^\dagger L^\dagger \tau_3 \Phi d^3 x \]

\[ = \int \Psi^\dagger \tau_3 L \Psi d^3 x \quad (33) \]
Since $\tau_3^\dagger = \tau_3$ from 4, this gives the condition

$$L^\dagger \tau_3 = \tau_3 L$$

(34)

and since $\tau_3^2 = 1$ we can multiply on the left by $\tau_3$ to get

$$\tau_3 L^\dagger \tau_3 = L$$

(35)

With the definition

$$L^H \equiv \tau_3 L^\dagger \tau_3$$

(36)

the condition

$$L^H = L$$

(37)

is called the 'generalized hermiticity condition'. If $[L, \tau_3] = 0$, then $L^\dagger = L$ and $L$ is an ordinary Hermitian operator.