Let's look at plane-wave solutions of the K-G equation,

\[ \psi = e^{i (\mathbf{p} \cdot \mathbf{r} - E t) / \hbar} \]  

(E, \mathbf{p} = c-numbers, parameters of wave).

Plugging into the K-G eqn, we find a solution of

\[ E = \pm \sqrt{c^2 p^2 + m^2 c^4}. \]

This of course is just the classical energy-momentum relation, but with the possibility of a minus sign, that is, \( E < 0 \). The KG eqn possesses solutions of negative energy. These have no analog in classical relativity theory or in the NR Sch. theory of QM, so we don't know what they mean. We ignore this interpretational difficulty and push on.

In the next step we introduce covariant notation. Let the space-time 4-vector be

\[ x^\mu = (ct, \mathbf{r}) \]

i.e., \( x^0 = ct \), \( \mu = 0, 1, 2, 3 \).

The momentum 4-vector is

\[ p^\mu = \left( \frac{E}{c}, \mathbf{p} \right). \]

We use the metric

\[ g_{\mu \nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \]

to raise and lower indices, so time-like vectors have + norm, and space-like vectors have - norm. Then
\( x_\mu = (ct, \vec{x}) \)

\( p_\mu = \left( \frac{E}{c}, -\vec{p} \right) \)

and

\[ p_\mu p^\mu = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2. \]

We also introduce the derivative operators,

\[ \partial_\mu = \frac{\partial}{\partial x^\mu} = \left( \frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \]

\[ \partial^\mu = \frac{\partial}{\partial x_\mu} = \left( \frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right). \]

Notice the positions of the indices. So far this is all classical.

Now we can write the de-Broglie - Einstein relations in covariant form,

\[ \frac{E}{c} \rightarrow \text{it} \frac{\partial}{\partial (ct)} = \text{it} \partial_0 \]

\[ \vec{p} \rightarrow -\text{it} \nabla = \text{it} \partial^i, \quad i = 1, 2, 3. \]

That is,

\[ p_\mu \rightarrow \text{it} \partial^\mu \]

In terms of the 4-momentum operators, the KG eqn can now be written

\[ p^\mu p_\mu \psi = m^2 c^2 \psi \]

or

\[ \partial^\mu \partial_\mu \psi = -\left( \frac{mc}{\hbar} \right)^2 \psi \]

Here of course the \( p^\mu \) are operators, not c-numbers.
Next we look at the probability density and current. For the NR Sch. eqn., these are

\[ \rho = |\psi|^2 \]
\[ \mathcal{J} = \frac{-i\hbar}{2m} (\psi^* \nabla \psi - \psi \nabla \psi^*) \quad \text{(when } \hbar = 0) \]

If \( \psi \) is a soln of the Sch. eqn, then

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

Sometimes the continnity eqn can be put into covariant form. We define

\[ J^\mu = (c\rho, \mathcal{J}) \]

whereupon

\[ \frac{\partial \rho}{\partial t} + \nabla \cdot \mathcal{J} = \frac{\partial J^\mu}{\partial x^\mu} = \partial_\mu J^\mu. \]

We should not use such notation for the NR Sch. eqn. versions of \( \rho \) and \( \mathcal{J} \), however. Just because we define write down 4 quantities and put a Greek index on them does not mean that they form a 4-vector. Instead they must transform as a 4-vector under Lorentz transformations. The quantities \( (c\rho, \mathcal{J}) \) defined for the NR Sch. eqn. do not transform as a 4-vector, even though they do satisfy the continuity eqn. Instead, they (and the continuity eqn) are specific to one Lorentz frame.

The KG eqn however does possess a conserved 4-current.
\[ J^\mu = \frac{i\hbar}{2m} \left( \psi^* \partial^\mu \psi - \psi \partial^\mu \psi^* \right). \]

To show it is conserved, compute

\[ \partial_\mu J^\mu = \frac{i\hbar}{2m} \left[ (\partial_\mu \psi^*)(\partial^\mu \psi) + \psi^* (\partial_\mu \partial^\mu \psi) - (\partial_\mu \psi)(\partial^\mu \psi^*) - \partial_\mu (\partial^\mu \psi^*) \right] \]

\[ \quad \text{cancel.} \]

The 1st + 3rd terms cancel. So do the 2nd and 4th if we use the KG eqn:

\[ \partial_\mu \partial^\mu \psi = -\left( \frac{mc}{\hbar} \right)^2 \psi \]
\[ \partial_\mu \partial^\mu \psi^* = -\left( \frac{mc}{\hbar} \right)^2 \psi^* \]

So, \[ \partial_\mu J^\mu = 0. \]

Moreover, this eqn is covariant, since \( J^\mu \) transforms as a 4-vector. This follows if we assume \( \psi \) transforms as a scalar (the simplest assumption for a scalar field).

Unfortunately, there is a problem with this 4-current. The density

\[ J^0 = \frac{i\hbar}{2mc^2} \left( \psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \]

is not positive definite. It leads to "negative probabilities." We will not go into this in detail, except to remark that the negative probabilities can be seen to be related to the negative energy solutions, which are also related to the fact that the KG eqn is 2nd order in time.
For these reasons, the KG eqn. was abandoned for several years after its introduction, and regarded as unsatisfactory. Later it was revived, and it is now considered a respectable wave eqn. describing spin-0 bosons. But we need several new concepts to understand this. In fact, in this course we will not have much more to say about the KG eqn. for lack of time.

This was the state of affairs when Dirac began his investigations of relativistic wave eqns. To avoid negative probabilities, Dirac considered an eqn. that is first order in time,

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

where $H$ is some Hamiltonian operator to be determined. At first we take the case of a free particle.

Since the wave eqn. is 1st order in time, and since relativity theory is supposed to treat space and time on an equal footing, Dirac reasoned that the wave eqn. should be first order in space as well. (Of course the Sch. eqn. is 2nd order in space.) So using the momentum operators $p_k = -i\hbar \frac{\partial}{\partial x_k}$,
\( k = 1, 2, 3 \) and introducing coefficients \( \alpha_k, k = 1, 2, 3 \) and another coefficient \( \beta \), Dirac wrote

\[
\frac{i\hbar}{\partial t} \frac{\partial \psi}{\partial t} = -i\hbar c \sum_{k=1}^{3} \alpha_k \frac{\partial \psi}{\partial x_k} + mc^2 \beta \psi.
\]

The \( c \) in the \( \alpha \)-term and the \( mc^2 \) in the \( \beta \) term are introduced to make \( \alpha_k \) and \( \beta \) dimensionless. This is equivalent to the Hamiltonian,

\[
H = c \vec{\alpha} \cdot \vec{p} + mc^2 \beta \quad \text{(free particle Dirac Hamiltonian)}.
\]

The quantities \( \alpha_k \) or \( \vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \) cannot be ordinary numbers because if they were they would specify a vector in ordinary space that would pick out a particular direction, which we don't expect for a free particle. Instead, Dirac assumed that \( \psi \) is a multi-component "spinor" wave function with \( N \) components,

\[
\psi = \begin{bmatrix}
\psi_1 \\
\vdots \\
\psi_N
\end{bmatrix}
\]

where \( N \) is to be determined. Then \( \alpha_k \) and \( \beta \) are interpreted as \( N \times N \) matrices. We think of \( \vec{\alpha} \) as a "vector of matrices"
much like the Pauli matrices $\vec{\sigma}$. The matrices $\sigma_\alpha$ and $\beta$ must be constants (independent of $\vec{x}, t$) because the free particle Hamiltonian must be invariant under space or time translations. Thus $\vec{A} \cdot \vec{p} = \vec{p} \cdot \vec{A}$ (where $\vec{p} = -i\hbar \nabla$).

Henceforth when looking at Dirac formalism you must remember that $\psi$ is a spinor and $\vec{A}, \beta$ are matrices.

To determine the form of the $\vec{A}, \beta$ matrices Dirac required that every free particle solution of the Dirac eqn. should also be a solution of the KG eqn. This is so that the classical energy-momentum relations $E^2 = c^2 p^2 + m^2 c^4$ would be satisfied. Thus we take

$$i\hbar \frac{\partial \psi}{\partial t} = H \psi = \left( c \sum_k \alpha_k p_k + mc \beta \right) \psi$$

and apply $i\hbar \frac{\partial}{\partial t}$ to both sides, getting

$$i\hbar \frac{\partial}{\partial t} \left( i\hbar \frac{\partial \psi}{\partial t} \right) = H \left( i\hbar \frac{\partial \psi}{\partial t} \right) = H^2 \psi = \left( c \sum_k \alpha_k p_k + mc \beta \right)^2 \psi$$

$$= -\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = \left( c^2 \sum_k \alpha_k \alpha_k p_k p_k + mc^2 \sum_k \alpha_k \beta + \beta \alpha_k + mc^2 \beta \right) \psi.$$

To get the KG eqn, we must have:
\[ \beta^2 = 1 \quad (\text{the } N \times N \text{ identity matrix}) \]

\[ \alpha_k \beta + \beta \alpha_k = 0 \]

\[ \alpha_k^2 = 1, \quad k = 1, 2, 3 \]

\[ \alpha_k \alpha_2 = -\alpha_2 \alpha_k \quad (k \neq 2). \]

If these are satisfied then

\[ -\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -\hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi \quad (\text{kg equ}). \]

The above equations constitute the Dirac algebra of the \( \gamma, \beta \)
matrices. The Dirac algebra is more conveniently written in terms of anticommutators:

\[
\begin{align*}
\{ \alpha_k, \alpha_l \} &= 2 \delta_{kl} \\
\{ \alpha_k, \beta^a \} &= 0 \\
\{ \beta^a, \beta^b \} &= 2 \delta^{ab}.
\end{align*}
\]

Dirac algebra, anticommutator version.

To find the matrices that satisfy the Dirac algebra, first note that they must be Hermitian if we expect \( H = c \beta \cdot \vec{p} + m^2 \beta \) to be Hermitian. Next, since \( \alpha_k^2 = \beta^2 = 1 \), their eigenvalues must be \( \pm 1 \).

Next, take \( \alpha_k \beta + \beta \alpha_k = 0 \), multiply by \( \beta \), use \( \beta^2 = 1 \), and take traces, so that

\[ \text{tr} (\beta \alpha_k \beta) + \text{tr} (\alpha_k) = 0. \]
But \( \text{tr} (\beta \gamma e \beta) = \text{tr} (\beta^2 e) = \text{tr} (e) \), so

\[ 2 \text{tr} (e) = 0 \quad \Rightarrow \quad \text{tr} e = 0. \]

Similarly we can show that \( \text{tr} \beta = 0 \). But since the trace is the sum of the eigenvalues, which must be \pm 1, the number of +1's and -1's must be equal, so \( N = \text{even} \).

Thus the simplest case to try is \( N = 2 \). In this case all Hermitian matrices can be represented as a linear combination of \((I, \gamma^2)\), and with some computation one can show that it is impossible to satisfy the Dirac algebra with \( N = 2 \). Instead we try \( N = 4 \).

At \( N = 4 \), Dirac was able to find matrices that satisfy his algebra. These are

\[ \gamma = \begin{pmatrix} 0 & \gamma^3 \\ \gamma^7 & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{(Dirac–Pauli representation)} \]

Here all 4x4 matrices are partitioned into 4 2x2 matrices, and we use the Pauli matrices \( \gamma \) for the 2x2 subblocks. Also, \( I \) means identity matrix.

When we have a set of algebraic relations that some objects satisfy, such as the Dirac algebra, and we find a concrete set of matrices that satisfy these relations, then we say that we have a representation of those relations. Thus, Dirac found a 4x4
matrix representation of his algebra. (the Dirac-Pauli representation).

However, that representation is not unique, because if we conjugate
the matrices \( \gamma^i, \beta \) by any fixed unitary matrix,

\[
\begin{align*}
\gamma^i & \to U \gamma^i U^+ \\
\beta & \to U \beta U^+
\end{align*}
\]

then the new \( \gamma^i, \beta \) are still Hermitian and still satisfy the same
algebra. Any two representations that differ by such a conjugation
(a change of basis) are said to be equivalent.

Another 4x4 representation of the Dirac algebra, equivalent to
the Dirac-Pauli representation, is

\[
\gamma^0 = \begin{pmatrix} 0 & 0 \\
0 & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} 0 & -1 \\
-1 & 0 \end{pmatrix} \quad \text{(Weyl representation)}
\]

One should note that this differs physically. Any calculation that can
be carried out in the Dirac-Pauli rep's can also be carried out in
the Weyl rep's, with the same answers from a physical standpoint.
But the calculation may be more or less convenient in one or the
other. The Dirac-Pauli rep's is especially useful when studying
the NR limit of the Dirac eqn, while the Weyl rep's is useful for
ultra relativistic or massless particles (like neutrinos).
There is another equivalent representation of the Dirac algebra, due to Majorana, that is commonly used. We will not quote it, however.

It can be shown that all 4x4 representations of the Dirac algebra are equivalent to the one found by Dirac. Thus, apart from a change of basis (unitary conjugation), Dirac's solution by 4x4 matrices is unique.

The Dirac matrices are 4-dimensional, but this has nothing to do with the fact that there are 4 space-time dimensions. Instead, the Dirac matrices act on "spin space", while usual tensors in relativity theory, $F_{\mu\nu}$ etc., act on space-time components.
The free particle Dirac eqn is
\[ \frac{i\hbar}{\partial t} \psi = -i\hbar c \nabla \psi + mc^2 \beta \psi = H \psi \]
where \( H = c \nabla^2 \beta + mc^2 \beta \). To incorporate the interaction with the EM field, we use the minimal coupling prescription. This is the simplest method of coupling an otherwise free particle with the EM field which is Lorentz covariant. It amounts to replacing the 4-momentum of the particle \( p^\mu \) by
\[ p^\mu \rightarrow p^\mu - \frac{q}{c} A^\mu, \]
where \( q = \) charge and
\[ A^\mu = (\Phi, \vec{A}) \]
is the 4-vector potential (\( \Phi = \) scalar potential, \( \vec{A} = \) 3-vector potential).

In QM, the 4-momentum \( p^\mu \) becomes the operator \( i\hbar \partial^\mu \), so the minimal coupling prescription gives
\[ i\hbar \frac{\partial}{\partial t} \rightarrow i\hbar \frac{\partial}{\partial t} - q \Phi, \]
\[ -i\hbar \nabla \rightarrow -i\hbar \nabla - \frac{q}{c} \vec{A}. \]
With this change, the Dirac eqn becomes
\[ \frac{i\hbar}{\partial t} \psi = -i\hbar c \nabla \psi - \frac{q}{c} (\nabla \vec{A}) \psi + mc^2 \beta \psi + q \Phi \psi. \]
or
\[ i\hbar \frac{\partial \psi}{\partial t} = H\psi, \]

where
\[ H = C\tilde{\alpha} \cdot (\overrightarrow{p} - \frac{q}{c} \overrightarrow{A}) + mc^2\beta + q\Phi. \]

It is a guess that this is correct. There are other Lorentz covariant couplings to the EM field, but these are not as simple as the minimal one, and they lead to more complicated equations.

Let us now look for a conserved probability current. Following the steps used in the NR Sch. equ., we write down the wave eqn for \( \psi \) and then take the Hermitian conjugate, which maps the column spinor \( \psi = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_n \end{bmatrix} \) into the row spinor \( \psi^+ = [\phi_1^* ... \phi_n^*] \). Since matrices \( \tilde{\alpha}, \beta \) are Hermitian, they are not affected by the Herm. conj., but we must reverse the order of the spin matrix-spinor products:

\[ i\hbar \frac{\partial \psi}{\partial t} = -i\hbar c \overrightarrow{\nabla} \psi - \frac{q}{c} (\overrightarrow{A} \cdot \overrightarrow{\nabla}) \psi + mc^2\beta \psi + q\Phi \psi \]

\[ -i\hbar \frac{\partial \psi^+}{\partial t} = i\hbar c \nabla \psi^+ \cdot \overrightarrow{A} - \frac{q}{c} \psi^+(\overrightarrow{A} \cdot \overrightarrow{\nabla}) + mc^2 \psi^+ \beta + q\Phi \psi^+ \]

Now multiply 1st eqn by \( \psi^+ \) from the left and the 2nd by \( -\psi \) from the right. Then all spin indices are contracted and we have a scalar insofar as spin indices are concerned. The quantities still depend on \( \mathbb{R}, t \). The result is

\[ i\hbar \frac{\partial (\psi^+ \psi)}{\partial t} = -i\hbar c \overrightarrow{\nabla} (\psi^+ \overrightarrow{\partial} \psi), \]
\[ \frac{\partial \rho}{\partial t} + \nabla \cdot \bar{J} = 0 \]

where

\[ \rho = \psi^* \psi, \quad \bar{J} = \mathcal{E} (\psi^* \nabla \psi). \]

To write this out explicitly, using \( a, b = 1, \ldots, 4 \) for spinor indices,

\[ \rho = \sum_{a=1}^{4} |\psi_a|^2 \]

\[ \bar{J}_a = \mathcal{E} \sum_{ab} \psi_a^* (\psi_b)_{ab} \psi_b. \]

We see that \( \rho \geq 0 \), so \( \rho \) may be interpreted as a probability density. Dirac felt that in this way he had overcome the main difficulty of the Klein-Gordon eqn (negative probabilities). Thus we define \( \bar{J}^H = \mathcal{E} (cp, \bar{J}) \), so the continuity eqn is \( \partial \rho \bar{J}^H = 0. \) Of course we must show that \( \bar{J}^H \) transforms as a 4-vector to show that this eqn is really covariant. That is a fairly large project that we defer until we have a better understanding of the physical meaning of the Dirac equation.

To develop the physics of the Dirac eqn, we consider the simplest possible exact solution, that of a free particle at rest, for which \( \nabla \psi = 0. \)