G25.2666: Quantum Mechanics II

Notes for Lecture 5

I. REPRESENTING STATES IN THE FULL HILBERT SPACE

Given a representation of the states that span the spin Hilbert space, we now need to consider the problem of representing the the states the span the full Hilbert space:

$$\mathcal{H} = \mathcal{H}_r \bigotimes \mathcal{H}_s$$

We will work with the following complete set of commuting observables (CSCO): $\{X, Y, Z, S^2, S_z\}$, which means that the basis vectors which span the full Hilbert space must be simultaneous eigenvectors of these five operators. These will be represented as

$$|\mathbf{r}\ s\ m_s\rangle = |\mathbf{r}\rangle \bigotimes |s\ m_s\rangle$$

that is, they will be a tensor product of the usual coordinate eigenvector and the simultaneous eigenvector of S^2 and S_z. Thus, they will satisfy the eigenvalue equations:

$$\begin{aligned}
X|\mathbf{r}\ \ s\ m_s\rangle &= x|\mathbf{r}\ \ s\ m_s\rangle \\
Y|\mathbf{r}\ \ s\ m_s\rangle &= y|\mathbf{r}\ \ s\ m_s\rangle \\
Z|\mathbf{r}\ \ s\ m_s\rangle &= z|\mathbf{r}\ \ s\ m_s\rangle \\
S^2|\mathbf{r}\ \ s\ m_s\rangle &= s(s+1)\hbar^2|\mathbf{r}\ \ s\ m_s\rangle \\
S_z|\mathbf{r}\ \ s\ m_s\rangle &= m_s\hbar|\mathbf{r}\ \ s\ m_s\rangle
\end{aligned}$$

The basis vectors will also satisfy an orthogonality relation:

$$\rangle\mathbf{r}\ \ s\ m_s|\mathbf{r}'\ s\ m_s'\rangle = \delta_{m_s m_s'}\delta^{(3)}(\mathbf{r}-\mathbf{r}')$$

Any arbitrary vector $|\phi\rangle$ in the Hilbert space can be expanded in terms of these basis vectors:

$$|\phi\rangle = \sum_{m_s=-s}^{s} \int d\mathbf{r}\ \ |\mathbf{r}\ \ s\ m_s\rangle\langle\mathbf{r}\ \ s\ m_s|\phi\rangle$$

The expansion coefficients can, as usual, be designated as functions of $\mathbf{r}$:

$$\langle\mathbf{r}\ \ s\ m_s|\phi\rangle = \phi_{s,m_s}(\mathbf{r})$$

For the case of spin-1/2, the expansion takes the form

$$\begin{aligned}
|\phi\rangle &= \sum_{m_s=-1/2}^{1/2} \int d\mathbf{r}\ \left|\mathbf{r}\ \frac{1}{2}\ m_s\right\rangle\left\langle\mathbf{r}\ \frac{1}{2}\ m_s\right|\phi\rangle \\
&= \int d\mathbf{r}\ \left(\left|\mathbf{r}\ \frac{1}{2}\ -\frac{1}{2}\right\rangle\left\langle\mathbf{r}\ \frac{1}{2}\ -\frac{1}{2}\right|\phi\rangle\ +\ \left|\mathbf{r}\ \frac{1}{2}\ \frac{1}{2}\right\rangle\left\langle\mathbf{r}\ \frac{1}{2}\ \frac{1}{2}\right|\phi\rangle\right)
\end{aligned}$$

The coefficients are designated by

$$\begin{aligned}
\left\langle\mathbf{r}\ \frac{1}{2}\ \frac{1}{2}\right|\phi\rangle &= \phi_{\frac{1}{2}\ \frac{1}{2}}(\mathbf{r}) \quad \text{or} \quad \phi_{\frac{1}{2}}(\mathbf{r}) \quad \text{or} \quad \phi_{\uparrow}(\mathbf{r}) \\
\left\langle\mathbf{r}\ \frac{1}{2}\ -\frac{1}{2}\right|\phi\rangle &= \phi_{\frac{1}{2}\ -\frac{1}{2}}(\mathbf{r}) \quad \text{or} \quad \phi_{-\frac{1}{2}}(\mathbf{r}) \quad \text{or} \quad \phi_{\downarrow}(\mathbf{r})
\end{aligned}$$

Then, since the basis vectors are:

$$\left|\mathbf{r}\ \frac{1}{2}\ \frac{1}{2}\right\rangle = |\mathbf{r}\rangle \bigotimes \left|\frac{1}{2}\ \frac{1}{2}\right\rangle = |\mathbf{r}\rangle \bigotimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\left|\mathbf{r}\ \frac{1}{2}\ -\frac{1}{2}\right\rangle = |\mathbf{r}\rangle \bigotimes \left|\frac{1}{2}\ -\frac{1}{2}\right\rangle = |\mathbf{r}\rangle \bigotimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

the expansion can be written as

$$\begin{aligned}
|\phi\rangle &= \int d\mathbf{r}\ \left(|\mathbf{r}\rangle \bigotimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \phi_{-\frac{1}{2}}(\mathbf{r}) + |\mathbf{r}\rangle \bigotimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \phi_{\frac{1}{2}}(\mathbf{r})\right) \\
&= \int d\mathbf{r}\ |\mathbf{r}\rangle \bigotimes \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} \phi_{-\frac{1}{2}}(\mathbf{r}) + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \phi_{\frac{1}{2}}(\mathbf{r})\right] \\
&= \int d\mathbf{r}\ |\mathbf{r}\rangle \bigotimes \begin{pmatrix} \phi_{\frac{1}{2}}(\mathbf{r}) \\ \phi_{-\frac{1}{2}}(\mathbf{r}) \end{pmatrix}
\end{aligned}$$

The vector

$$\begin{pmatrix} \phi_{\frac{1}{2}}(\mathbf{r}) \\ \phi_{-\frac{1}{2}}(\mathbf{r}) \end{pmatrix}$$

is called a two-component spinor. Note that

$$\begin{aligned}
\langle\phi|\phi\rangle &= \int d\mathbf{r} \int d\mathbf{r}'\ \left(\phi^*_{\frac{1}{2}}(\mathbf{r}')\quad \phi^*_{-\frac{1}{2}}(\mathbf{r}')\right) \begin{pmatrix} \phi_{\frac{1}{2}}(\mathbf{r}) \\ \phi_{-\frac{1}{2}}(\mathbf{r}) \end{pmatrix} \langle\mathbf{r}'|\mathbf{r}\rangle \\
&= \int d\mathbf{r} \int d\mathbf{r}'\ \left[\phi_{\frac{1}{2}}(\mathbf{r}')\phi_{\frac{1}{2}}(\mathbf{r}') + \phi_{-\frac{1}{2}}(\mathbf{r})\phi_{-\frac{1}{2}}(\mathbf{r})\right] \delta^{(3)}(\mathbf{r}-\mathbf{r}') \\
&= \int d\mathbf{r}\ \left(|\phi_{\frac{1}{2}}(\mathbf{r})|^2 + |\phi_{-\frac{1}{2}}(\mathbf{r})|^2\right)
\end{aligned}$$

Example: If we have a spin-independent Hamiltonian that is also spherically symmetric, then the quantum numbers that characterize the states will be n, l, m, s, m_s. Thus, for the hydrogen atom,

$$H = \left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{e^2}{r}\right]$$

which is spin independent. The ground state will, therefore, be twofold degenerate with the two eigenfunctions being:

$$\begin{aligned}
\psi_{100\frac{1}{2}\ \frac{1}{2}}(r,\theta,\varphi) &= \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-r/a_0} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\psi_{100\frac{1}{2}\ -\frac{1}{2}}(r,\theta,\varphi) &= \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-r/a_0} \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}$$

II. ROTATIONS IN SPIN SPACE

Given two types of angular momentum, orbital and spin, it is possible to define a *total* angular momantum

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

$\mathbf{J}$ plays a special role in quantum mechanics. Not only is it often a constant of the motion even when H is spin-dependent, but it is the *generator of rotations* in the Hilbert space.

To see what this means, consider a simpler situation with the total linear momentum P. The linear momentum is known as the *generator* of translations in the Hilbert space. By this, we mean that the operator

$$T_{\mathbf{a}} = e^{-i\mathbf{P}\cdot\mathbf{a}/\hbar}$$

which is a function of P produces translations in space by an amount $\mathbf{a}$. Thus, its action on an arbitrary function of $\mathbf{r}$ is

$$T_{\mathbf{a}}\psi(\mathbf{r}) = \psi(\mathbf{r} - \mathbf{a})$$

To see that this is true, consider the one-dimensional version of this operator

$$T_a = e^{-iPa/\hbar}$$

Using the fact that $P = (\hbar/i)(d/dx)$, the action of T_a on an arbitrary function $\psi(x)$ is

$$T_a\psi(x) = e^{-ad/dx}\psi(x)$$

This can be evaluated by a Taylor series:

$$\begin{aligned} e^{-ad/dx}\psi(x) &= \left[1 - a\frac{d}{dx} + \frac{1}{2!}a^2\frac{d^2}{dx^2} - \frac{1}{3!}a^3\frac{d^3}{dx^3} + \cdots\right]\psi(x) \\ &= \psi(x) - a\psi'(x) + \frac{1}{2!}a^2\psi''(x) - \cdots \\ &= \psi(x-a) \end{aligned}$$

That is, the next to last line is just the Taylor expansion of $\psi(x-a)$ about $a = 0$. $\mathbf{P}$ is, therefore, called the generator of the *translation group.*

By analogy and by similar reasoning, it can be shown that $\mathbf{J}$ is the generator of rotations of vectors in the Hilbert space via the operator:

$$R_\alpha(\mathbf{n}) = \exp\left[-\frac{i}{\hbar}\alpha\mathbf{J}\cdot\hat{\mathbf{n}}\right]$$

which produces rotations of a vector by an angle α about an axis defined by the unit vector $\hat{\mathbf{n}}$. $\mathbf{J}$ is called the *generator of the rotation group.*

Since $\mathbf{L}$ and $\mathbf{S}$ commute (they act in different Hilbert spaces), the rotation operator can be written as

$$\begin{aligned} R_\alpha(\mathbf{n}) &= \exp\left[-\frac{i}{\hbar}\alpha(\mathbf{L}+\mathbf{S})\cdot\hat{\mathbf{n}}\right] \\ &= \exp\left[-\frac{i}{\hbar}\alpha\mathbf{L}\cdot\hat{\mathbf{n}}\right]\exp\left[-\frac{i}{\hbar}\alpha\mathbf{S}\cdot\hat{\mathbf{n}}\right] \\ &= R_\alpha^{(r)}(\mathbf{n})R_\alpha^{(s)}(\mathbf{n}) \end{aligned}$$

Thus, a particle whose state vector is separable into spatial and spin components according to

$$|\psi\rangle = |\phi\rangle \bigotimes |\chi\rangle$$

will be transformed according to

$$\begin{aligned} |\psi'\rangle = R_\alpha(\mathbf{n})|\psi\rangle &= \left[\exp\left(-\frac{i}{\hbar}\alpha\mathbf{L}\cdot\hat{\mathbf{n}}\right)|\phi\rangle\right] \bigotimes \left[\exp\left(-\frac{i}{\hbar}\alpha\mathbf{S}\cdot\hat{\mathbf{n}}\right)|\chi\rangle\right] \\ &= |\phi'\rangle \bigotimes |\chi'\rangle \end{aligned}$$

Let us focus on the spin part of this equation, which transform $|\chi\rangle \longrightarrow |\chi'\rangle$ by

$$|\chi'\rangle = \exp\left(-\frac{i}{\hbar}\alpha\mathbf{S}\cdot\hat{\mathbf{n}}\right)|\chi\rangle$$

Since $\mathbf{S} = (\hbar/2)\,\vec{\sigma}$, where $\vec{\sigma}$ is the vector of Pauli matrices, the spin rotation operator becomes

$$R_\alpha^{(s)}(\mathbf{n}) = \exp\left[-i\frac{\alpha}{2}\,\vec{\sigma}\cdot\hat{\mathbf{n}}\right]$$

Thus, the generators of the spin-1/2 rotation group are just the $2\times$ 2 Pauli matrices.

The spin-1/2 rotation group has a special name. It is known as SU(2). SU(2) is the group of 2× 2 unitary matrices with unit determinant. The representation of such a matrix as

$$\exp\left(-i\frac{\alpha}{2}\vec{\sigma}\cdot\hat{\mathbf{n}}\right)$$

shows that there are an infinite number of such matrices, since the parameters α and $\hat{\mathbf{n}}$, which constitutes three parameters (remember $\hat{\mathbf{n}}$, being a unit vector, has only two independent components), and thus, SU(2), is an example of a continuous *Lie group* (because the generators satisfy a Lie algebra).

In general, SU(n) is the group of $n \times n$ unitary matrices with unit determinant. The number of generators belonging to SU(n) is $n^2 - 1$. Thus, for SU(2), there should be $2^2 - 1 = 3$ generators, which is, indeed, the number of Pauli matrices. SU(3), for example, should have $3^2 - 1 = 8$ generators. (Since SU(3) is the group in terms of which quantum chromodynamics, the theory of quarks, is formulated, the eight generators correspond to the eight gluons in the theory.)

Note that it is possible to *represent* the group in terms of matrices of higher dimension then n, so long as the number of generators and independent parameters remains the same. For example, the group SU(2) and the group SO(3) (SO(n) is the group of $n \times n$ orthogonal matrices with unit determinant), which is used to generate rotations of vectors in ordinary Cartesian space, have the same number of generators and independent parameters. Thus, SU(2) is said to be *isomorphic* to SO(3), and, therefore, there should be a *representation* of SU(2) in terms of 3× 3 matrices. This will be true of any group to which SU(2) is isomorphic.

In order to generate a representation of SU(2), we need to determine the generators of that representation. This can be accomplished by knowing the action of the raising and lowering operators and the operator S_z on the spin states. The general relations are:

$$\begin{aligned}
S_z|s\ \ m_s\rangle &= m_s\hbar|s\ \ m_s\rangle \\
S_+|s\ \ m_s\rangle &= \sqrt{(s-m_s)(s+m_s+1)}|s\ \ m_s+1\rangle \\
S_-|s\ \ m_s\rangle &= \sqrt{(s+m_s)(s-m_s+1)}|s\ \ m_s-1\rangle
\end{aligned}$$

Note that for spin-1/2, this reduces to the relations we wrote down before. This are *general* relations that we will need later when we consider addition of angular momentum.

From these relations, we can construct a representation of SU(2). Consider the case of a spin-1/2 particle. It is clear that the operator S_z is diagonal, and its eigenvalues must be $\pm\hbar/2$, so we can write down the form of S_z immediately, using the fact that it is diagonal in the basis we are working with:

$$S_z = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix}$$

In order to get S_x and S_y, note that the raising and lowering operators must satisfy

$$\begin{aligned}
S_+\begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \hbar\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
S_-\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \hbar\begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}$$

The matrix forms for S_+ and S_- that produce this action on the spin states must be

$$\begin{aligned}
S_+ &= \hbar\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\
S_- &= \hbar\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}
\end{aligned}$$

Then, S_x and S_y are given by

$$\begin{aligned}
S_x &= \frac{1}{2}(S_+ + S_-) = \begin{pmatrix} 0 & \frac{\hbar}{2} \\ \frac{\hbar}{2} & 0 \end{pmatrix} \\
S_y &= \frac{1}{2i}(S_+ - S_-) = \begin{pmatrix} 0 & -\frac{i\hbar}{2} \\ \frac{i\hbar}{2} & 0 \end{pmatrix}
\end{aligned}$$

The same can be done, for example, for a spin-1 particle, which will yield the 3× 3 representation of the group generators.

B. Explicit form of the spin-1/2 rotation operator

For spin-1/2, the rotation operator

$$R_\alpha^{(s)}(\mathbf{n}) = \exp\left(-i\frac{\alpha}{2}\vec{\sigma}\cdot\hat{\mathbf{n}}\right)$$

can be written as an explicit 2× 2 matrix. This is accomplished by expanding the exponential into a Taylor series:

$$\exp\left(-i\frac{\alpha}{2}\vec{\sigma}\cdot\hat{\mathbf{n}}\right) = 1 - \frac{i\alpha}{2}\vec{\sigma}\cdot\hat{\mathbf{n}} + \frac{1}{2!}\left(\frac{i\alpha}{2}\right)^2(\vec{\sigma}\cdot\hat{\mathbf{n}})^2 - \frac{1}{3!}\left(\frac{i\alpha}{2}\right)^3(\vec{\sigma}\cdot\hat{\mathbf{n}})^3 + \frac{1}{4!}\left(\frac{i\alpha}{2}\right)^4(\vec{\sigma}\cdot\hat{\mathbf{n}})^4 - \cdots$$

Note that

$$(\vec{\sigma}\cdot\hat{\mathbf{n}})^2 = (\vec{\sigma}\cdot\hat{\mathbf{n}})(\vec{\sigma}\cdot\hat{\mathbf{n}}) = \hat{\mathbf{n}}\cdot\hat{\mathbf{n}} + i\sigma(\hat{\mathbf{n}}\times\hat{\mathbf{n}}) = 1$$

Thus, the Taylor series becomes:

$$\begin{aligned}
\exp\left(-i\frac{\alpha}{2}\vec{\sigma}\cdot\hat{\mathbf{n}}\right) &= 1 - \frac{i\alpha}{2}\vec{\sigma}\cdot\hat{\mathbf{n}} + \frac{1}{2!}\left(\frac{i\alpha}{2}\right)^2 - \frac{1}{3!}\left(\frac{i\alpha}{2}\right)^3(\vec{\sigma}\cdot\hat{\mathbf{n}}) + \frac{1}{4!}\left(\frac{i\alpha}{2}\right)^4 - \cdots \\
&= \left[1 - \frac{1}{2!}\left(\frac{\alpha}{2}\right)^2 + \frac{1}{4!}\left(\frac{\alpha}{2}\right)^4 + \cdots\right] - i\,\vec{\sigma}\cdot\hat{\mathbf{n}}\left[\left(\frac{\alpha}{2}\right) - \frac{1}{3!}\left(\frac{\alpha}{2}\right)^3 + \cdots\right] \\
&= \cos\left(\frac{\alpha}{2}\right) - i\,\vec{\sigma}\cdot\hat{\mathbf{n}}\sin\left(\frac{\alpha}{2}\right)
\end{aligned}$$

Thus,

$$R_\alpha^{(s)}(\mathbf{n}) = \cos\left(\frac{\alpha}{2}\right) - i\,\vec{\sigma}\cdot\hat{\mathbf{n}}\sin\left(\frac{\alpha}{2}\right)$$

As a 2×2 matrix,

$$\vec{\sigma}\cdot\hat{\mathbf{n}} = \sigma_x n_x + \sigma_y n_y + \sigma_z n_z = \begin{pmatrix} 0 & n_x \\ n_x & 0 \end{pmatrix} + \begin{pmatrix} 0 & -in_y \\ in_y & 0 \end{pmatrix} + \begin{pmatrix} n_z & 0 \\ 0 & -n_z \end{pmatrix} = \begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix}$$

so that the rotation operator becomes

$$R_\alpha^{(s)}(\mathbf{n}) = \begin{pmatrix} \cos\left(\frac{\alpha}{2}\right) - in_z\sin\left(\frac{\alpha}{2}\right) & (-in_x - n_y)\sin\left(\frac{\alpha}{2}\right) \\ (-in_x + n_y)\sin\left(\frac{\alpha}{2}\right) & \cos\left(\frac{\alpha}{2}\right) + in_z\sin\left(\frac{\alpha}{2}\right) \end{pmatrix}$$

Now consider the example of $\alpha = 2\pi$. In this case, it is easy to see that the rotation operator reduces to

$$R_{2\pi}^{(s)}(\hat{\mathbf{n}}) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathrm{I}$$

Interestingly, a rotation through an angle 2π of a spin state returns the state to its original value but causes it to pick up an overall phase factor

$$-1 = e^{i\pi}$$

While this phase factor cannot affect any physical property, it is, nevertheless observable in the experiment depicted below:

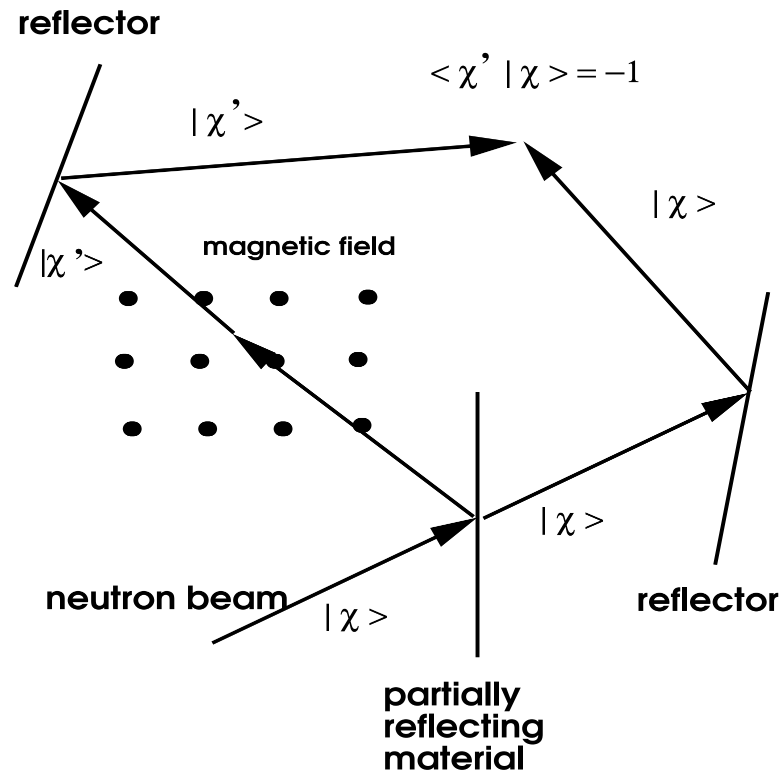

FIG. 1.

A beam of neutral spin-1/2 particles, such as neutrons, initially prepared in a definite spin state $|\chi\rangle$, is split by a partially reflecting material into two beams. One of these is sent through a magnetic field region tuned to generate a rotation by $\alpha = 2\pi$ of the spin state, so that the new state is $|\chi'\rangle$. The beams are then brought back together and allowed to interfere. The overlap, $\langle\chi'|\chi\rangle = -1$ is measured, which will yield the over phase factor -1.

III. INTRODUCTION TO THE DIRAC EQUATION

In 1928, P.A.M. Dirac proposed a relativistic formulation of the quantum mechanics of the electron from which spin emerges as a natural consequence of the relativistic treatment. Dirac's relativistic formulation of the electron becomes necessary to employ when one is interested in the low lying (core) states of heavy atoms, where, because of the large Coulomb forces (Z is large), the speed of electrons close to the nucleus approaches the speed of light. In addition, Dirac's theory is the basis for modern quantum electrodynamics, one of the most accurate quantum theories to date.

The problem with trying to marry quantum mechanics to Einstein's special theory of relativity is the fact that the relativistic energy of a free particle of mass m and momentum, $\mathbf{p}$ is given by

$$E = \sqrt{\mathbf{p}^2c^2 + m^2c^4}$$

where c is the speed of light. Note than when $\mathbf{p} = 0$, this reduces to Einstein's formula for the rest mass energy of a particle of mass m:

$$E = mc^2$$

Note that, when $pc \ll mc^2$, the non-relativistic limit is approached. In this case, the energy formula can be expanded about $|\mathbf{p}| = 0$, to give

$$E = \sqrt{m^2c^4\left(1 + \frac{\mathbf{p}^2c^2}{m^2c^4}\right)}$$

$$
\begin{aligned}
&= mc^2\sqrt{1+\frac{\mathbf{p}^2c^2}{m^2c^4}} \\
&\approx mc^2\left(1+\frac{\mathbf{p}^2c^2}{2m^2c^4}\right) \\
&= mc^2+\frac{\mathbf{p}^2}{2m} \equiv mc^2+E_s
\end{aligned}
$$

where E_s is defined to be the energy relative to the rest mass energy. Thus, it can be seen that when the rest mass energy is large, the kinetic energy $\mathbf{p}^2/2m$ is simply added on to the rest mass energy. Generally, in the non-relativistic theory, we define all energies relative to the rest mass energy.

The problem with formulating a relativistic Schrödinger equation is the energy expression, itself. If we naively try to generate a Hamiltonian by promoting the classical variable $\mathbf{p}$ to a quantum operator $\mathbf{P}$, then we would have a Hamiltonian of the form:

$$
H = \sqrt{\mathbf{P}^2c^2+m^2c^4}
$$

and we have no way to interpret the square root of an operator.

Various attempts were made to circumvent this problem. One such attempt involved simply squaring the Hamiltonian in the Schrödinger equation, so that one would have

$$
H^2|\psi(t)\rangle = -\hbar^2|\psi(t)\rangle
$$

This generates a kind of wave equation, called the Klein-Gordon equation, that has two solutions of the general form

$$
\begin{aligned}
|\psi(t)\rangle &= e^{-iHt/\hbar}|\psi(0)\rangle \\
|\psi(t)\rangle &= e^{iHt/\hbar}|\psi(0)\rangle
\end{aligned}
$$

i.e., both forward and backward propagating solutions. It was later suggested that the backward propagating solutions should correspond to anti-particle solutions. Feynman's proposal was that anti-particles should be viewed as particles traveling backward in time, and this notion remains even today.

The problem with the Klein-Gordon equation is that it does not incorporate spin and thus will only work for spinless particles. The idea of Dirac was to demand that there be Hamiltonian that is linear in $\mathbf{P}$ such the square of H would give the required formula

$$
H^2 = \mathbf{P}^2c^2+m^2c^4
$$

He took a general Hamiltonian of the form

$$
H = c\,\vec{\alpha}\cdot\mathbf{P}+\beta mc^2
$$

where $\vec{\alpha}=(\alpha_x,\alpha_y,\alpha_z)$ and β are parameters to be determined by the H^2 condition. But look at H^2:

$$
\begin{aligned}
H^2 &= c^2\left(\vec{\alpha}\cdot\mathbf{P}\right)^2+mc^3\left[\beta\left(\vec{\alpha}\cdot\mathbf{p}\right)+\left(\vec{\alpha}\cdot\mathbf{p}\right)\beta\right]+\beta^2m^2c^4 \\
&= c^2\left(\alpha_xP_x+\alpha_yP_y+\alpha_zP_z\right)^2+mc^3\left[\beta\left(\vec{\alpha}\cdot\mathbf{p}\right)+\left(\vec{\alpha}\cdot\mathbf{p}\right)\beta\right]+\beta^2m^2c^4 \\
&= c^2\left[\alpha_x^2P_x^2+\alpha_y^2P_y^2+\alpha_z^2P_z^2\right. \\
&+\left.\left(\alpha_x\alpha_y+\alpha_y\alpha_x\right)P_xP_y+\left(\alpha_x\alpha_z+\alpha_z\alpha_x\right)P_xP_z+\left(\alpha_y\alpha_z+\alpha_z\alpha_y\right)P_yP_z\right] \\
&+mc^3\left[\beta\left(\alpha_xP_x+\alpha_yP_y+\alpha_zP_z\right)+\left(\alpha_xP_x+\alpha_yP_y+\alpha_zP_z\right)\beta\right] \\
&+\beta^2m^2c^4 \\
&= |\mathbf{P}|^2c^2+m^2c^4
\end{aligned}
$$

Thus, we see that the required condition is satisfied if $\vec{\alpha}$ and β satisfy the following:

$$
\alpha_x^2=\alpha_y^2=\alpha_z^2=1
$$

$$
\alpha_x\alpha_y+\alpha_y\alpha_x=0
$$

$$\alpha_x\alpha_z + \alpha_z\alpha_x = 0$$
$$\alpha_y\alpha_z + \alpha_z\alpha_y = 0$$

$$\beta\alpha_x + \alpha_x\beta = 0$$
$$\beta\alpha_y + \alpha_y\beta = 0$$
$$\beta\alpha_z + \alpha_z\beta = 0$$

$$\beta^2 = 1$$

These conditions can *only* be satisfied if $\vec{\alpha}$ and β are matrices! Indeed, we need a total of four anticommuting matrices, none of which is the identity matrix. In addition, we can show that the matrices must all be traceless. To see this, note that because

$$\beta^2 = I \qquad \Rightarrow \qquad \beta = \beta^{-1}$$

and similarly, it can be see that $\alpha_x = \alpha_x^{-1}$ and the same for α_y and α_z. Thus, using the fact that

$$\beta\alpha_x = -\alpha_x\beta \qquad \Rightarrow \qquad \alpha_x = \beta^{-1}\alpha_x\beta$$

and taking the trace of both sides, we find that

$$\begin{aligned}\mathrm{Tr}(\alpha_x) &= -\mathrm{Tr}\left(\beta^{-1}\alpha_x\beta\right)\\ &= -\mathrm{Tr}\left(\alpha_x\beta\beta^{-1}\right)\\ &= -\mathrm{Tr}(\alpha_x)\end{aligned}$$

Thus, since $\mathrm{Tr}(\alpha_x) = -\mathrm{Tr}(\alpha_x)$, it follows that $\mathrm{Tr}(\alpha_x) = 0$. The same argument can be applied to α_y, α_z and β. Thus, we need a set of four traceless, anticommuting matrices. It turns out that the minimum dimension needed to satisfy these conditions is 4, and, therefore, $\vec{\alpha}$ and β are 4×4 matrices. One possible representation of the matrices is in terms of the Pauli matrices and the identity and takes the form:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} \mathrm{I} & 0 \\ 0 & -\mathrm{I} \end{pmatrix}$$

where each element is a 2×2 sub-block of the 4×4 matrix.