Why is spin an SU(2) Symmetry?

Chris Clark March 7, 2008

1 Introduction

In this paper, we will start from the following form of the Schrodinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \mu\boldsymbol{\sigma}\cdot\mathbf{B}\right)|\psi\rangle = i\hbar\partial_t|\psi\rangle$$

and show that $U|\psi\rangle$ represents the same physical state as $|\psi\rangle$ if $U \in SU(2)$.[1] This means that the solutions of this equation have an $SU(2)$ symmetry.

2 Lemmas

Let V_2 be the set of all hermitian traceless complex 2x2 matrices.

Lemma: Any $H \in V_2$ can be written as a real linear combination of the Pauli matrices, $x\sigma_x + y\sigma_y + z\sigma_z \equiv \mathbf{r}\cdot\boldsymbol{\sigma}$.
Proof: Assume $H \in V_2$. Then H is hermitian so

$$H = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}$$

Therefore a and d are real, and $b = c^*$. The traceless condition implies that $d = -a$.

$$H = \begin{pmatrix} a & c^* \\ c & -a \end{pmatrix} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} = \mathbf{r}\cdot\boldsymbol{\sigma}$$

where $x, y, z \in \mathbb{R}$ □

Lemma: V_2 is closed under the action of $U \in SU(2)$ defined by $U \cdot H = U^{-1}HU$ for all $H \in V_2$.
Proof: We only need to check that $U \cdot H$ is traceless and hermitian to show that it is in V_2.

$$\mathrm{Tr}(U^{-1}HU) = \mathrm{Tr}(HUU^{-1}) = \mathrm{Tr}(H) = 0$$

by the cyclic property of the trace and since H is traceless.

$$(U^{-1}HU)^\dagger = U^\dagger H^\dagger (U^{-1})^\dagger = U^{-1}HU$$

since $U^\dagger = U^{-1}$ by unitarity. Therefore $U^{-1}HU \in V_2$ □

Therefore we can write $U^{-1}(\mathbf{r}\cdot\boldsymbol{\sigma})U = \mathbf{r}'\cdot\boldsymbol{\sigma}$ because the second lemma assures us that the left hand side is in V_2 and the first lemma shows that this must then be some new real linear combination of the Pauli matrices, so we simply define $\mathbf{r}'$ to be the new set of coefficients.

Lemma: The transformation that takes $\mathbf{r}$ to $\mathbf{r}'$ is a linear transformation.
Proof:

$$(a\mathbf{r}_1 + b\mathbf{r}_2)'\cdot\boldsymbol{\sigma} = U^{-1}((a\mathbf{r}_1 + b\mathbf{r}_2)\cdot\boldsymbol{\sigma})U = U^{-1}(a\mathbf{r}_1\cdot\boldsymbol{\sigma} + b\mathbf{r}_2\cdot\boldsymbol{\sigma})U$$
$$= U^{-1}(a\mathbf{r}_1\cdot\boldsymbol{\sigma})U + U^{-1}(b\mathbf{r}_2\cdot\boldsymbol{\sigma})U = a\mathbf{r}_1'\cdot\boldsymbol{\sigma} + b\mathbf{r}_2'\cdot\boldsymbol{\sigma} = (a\mathbf{r}_1' + b\mathbf{r}_2')\cdot\boldsymbol{\sigma}$$

[1] We let $SU(2)$ represent the set of matrices mapped to by the defining representation of the group $SU(2)$.

Therefore $(a\mathbf{r}_1 + b\mathbf{r}_2)' = a\mathbf{r}_1' + b\mathbf{r}_2'$ since the Pauli matrices are linearly independent, so the transformation is linear $\square$

Therefore we can express the transformation as a 3×3 matrix $R(U)$, so that $\mathbf{r}' = R(U)\mathbf{r}$. We know that this matrix must be real because both $\mathbf{r}$ and $\mathbf{r}'$ are real vectors. This can be shown by setting $\mathbf{r} = \hat{e}_i$ for each of the three orthonormal basis vectors.

Lemma: $R(U)$ is a rotation matrix.
Proof: We already know that $R(U)$ is a real 3×3 matrix. We notice that

$$\det(H) = \begin{vmatrix} z & x - iy \\ x + iy & -z \end{vmatrix} = -z^2 - (x+iy)(x-iy) = -(x^2 + y^2 + z^2)$$

and similarly,

$$\det(H') = -(x'^2 + y'^2 + z'^2)$$

Then we have

$$\det(H') = \det(U^{-1}HU) = \det(U^{-1})\det(H)\det(U) = \det(U)^{-1}\det(H)\det(U) = \det(H)$$

so

$$x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2$$

Therefore lengths are preserved by the transformation between unprimed and primed coordinates. This means that $R(U)$ must be either a proper or improper rotation. The improper rotation case is ruled out by the fact that the parameters of U may be continuously varied to the identity. Therefore $R(U)$ is a rotation matrix $\square$

This means that acting elements of $SU(2)$ by conjugation on dot products of the form $\mathbf{r} \cdot \boldsymbol{\sigma}$ causes a spatial rotation of the vector $\mathbf{r}$.

3 Spin

Suppose we write the Schrodinger equation with spin, but we replace $|\psi\rangle$ with $U|\psi\rangle$ where $U \in SU(2)$,

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \mu\boldsymbol{\sigma}\cdot\mathbf{B}\right)(U|\psi\rangle) = i\hbar\partial_t(U|\psi\rangle)$$

Then multiply both sides by U^{-1}

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \mu U^{-1}\boldsymbol{\sigma}\cdot\mathbf{B}U\right)|\psi\rangle = i\hbar\partial_t|\psi\rangle$$

By the lemma this is just

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + \mu\boldsymbol{\sigma}\cdot[R(U)\mathbf{B}]\right)|\psi\rangle = i\hbar\partial_t|\psi\rangle$$

But this just says that $|\psi\rangle$ is the solution to the same equation in a rotated coordinate system, so we can say that $U|\psi\rangle = |\psi'\rangle$ where $|\psi'\rangle$ is the solution in a rotated coordinate system. But by the rotational invariance of the universe, we know that $|\psi\rangle \leftrightarrow |\psi'\rangle$, where the arrow means that the two sides correspond to the same physical state, even though they may not be equivalent element-by-element (recall that ket states are coordinate independent). Therefore $U|\psi\rangle \leftrightarrow |\psi\rangle$.

4 Two Electron System

We can repeat the derivation of the last system for a two-electron system if we use the Kronecker product to combine the state kets. The original two-particle Schrodinger equation looks like

$$\left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + \mu\boldsymbol{\sigma}_1\cdot\mathbf{B} + \mu\boldsymbol{\sigma}_2\cdot\mathbf{B}\right)(|\chi_1\rangle\otimes|\chi_2\rangle) = i\hbar\partial_t(|\chi_1\rangle\otimes|\chi_2\rangle)$$

We will proceed backwards, starting with a spatial rotation to find the affect on the state kets. The rotated form looks like

$$\left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + \mu\boldsymbol{\sigma}_1\cdot[R(U)\mathbf{B}] + \mu\boldsymbol{\sigma}_2\cdot[R(U)\mathbf{B}]\right)(|\chi_1'\rangle\otimes|\chi_2'\rangle) = i\hbar\partial_t(|\chi_1'\rangle\otimes|\chi_2'\rangle)$$

By the Lemma,

$$\left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + \mu U^{-1}(\boldsymbol{\sigma}_1\cdot\mathbf{B})U + \mu U^{-1}(\boldsymbol{\sigma}_2\cdot\mathbf{B})U\right)(|\chi_1'\rangle\otimes|\chi_2'\rangle) = i\hbar\partial_t(|\chi_1'\rangle\otimes|\chi_2'\rangle)$$

Now we use the fact that $\boldsymbol{\sigma_1}$ in this expression is a shorthand for $\boldsymbol{\sigma_1}\otimes I$ (similarly for $\boldsymbol{\sigma_2}$ and use the mixed product property of the Kronecker product i.e. $(A\otimes B)(C\otimes D) = AC\otimes BD$,

$$\left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2\right)(|\chi_1'\rangle\otimes|\chi_2'\rangle) + [\mu U^{-1}(\boldsymbol{\sigma}_1\cdot\mathbf{B})U|\chi_1'\rangle]\otimes|\chi_2'\rangle + |\chi_1'\rangle\otimes[\mu U^{-1}(\boldsymbol{\sigma}_2\cdot\mathbf{B})U|\chi_2'\rangle] = i\hbar\partial_t(|\chi_1'\rangle\otimes|\chi_2'\rangle)$$

Now we apply $U\otimes U$ to both sides and use the mixed product property,

$$\left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2\right)(U|\chi_1'\rangle\otimes U|\chi_2'\rangle) + [\mu(\boldsymbol{\sigma}_1\cdot\mathbf{B})U|\chi_1'\rangle]\otimes U|\chi_2'\rangle + U|\chi_1'\rangle\otimes[\mu(\boldsymbol{\sigma}_2\cdot\mathbf{B})U|\chi_2'\rangle] = i\hbar\partial_t(U|\chi_1'\rangle\otimes U|\chi_2'\rangle)$$

$$\left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + \mu\boldsymbol{\sigma}_1\cdot\mathbf{B} + \mu\boldsymbol{\sigma}_2\cdot\mathbf{B}\right)(U|\chi_1'\rangle\otimes U|\chi_2'\rangle) = i\hbar\partial_t(U|\chi_1'\rangle\otimes U|\chi_2'\rangle)$$

So we have found that

$$U|\chi_1'\rangle\otimes U|\chi_2'\rangle = (U\otimes U)(|\chi_1'\rangle\otimes|\chi_2'\rangle) = |\chi_1\rangle\otimes|\chi_2\rangle$$

If we let $U\to U^{-1}$ we get the final result that

$$(U\otimes U)|\chi_1\rangle\otimes|\chi_2\rangle = |\chi_1'\rangle\otimes|\chi_2'\rangle \leftrightarrow |\chi_1\rangle\otimes|\chi_2\rangle$$

Therefore there is a symmetry of two particle states under multiplication by matrices of the form $U\otimes U$, where $U\in SU(2)$. This set of matrices defines the $2\otimes 2$ representation of $SU(2)$, sometimes written as $\frac{1}{2}\otimes\frac{1}{2}$ using the relation $d = 2s+1$ to convert from dimension to spin. We still have an $SU(2)$ symmetry, but in the $2\otimes 2$ representation of $SU(2)$, the irreducible representations are the 3 and 1 representations. This is sometimes written as $2\otimes 2 = 3\oplus 1$ and it means that under a suitable similarity transformation we can block diagonalize $U\otimes U$ to look like

$$\begin{pmatrix} \square & \square & \square & 0 \\ \square & \square & \square & 0 \\ \square & \square & \square & 0 \\ 0 & 0 & 0 & \square \end{pmatrix}$$

with a 3×3 block and a 1×1 block. In this basis, three of the four independent functions in $|\chi_1\rangle\otimes|\chi_2\rangle$ are mixed together, and the fourth is independent. This mixing means that three of the fields must be physically indistinguishable, which is why the triplet states must have the same energy, but can have a different energy than the singlet state. When there is splitting between the energies of the triplet states it is due to higher order corrections from terms that do not possess the symmetries that we have shown here. It is interesting that we have neglected direct interactions between the two electrons by omitting the spin-spin terms in the Hamiltonian, yet there is still mixing between the states. This mixing must come from some mechanism that does not contribute to the energy of the system. Perhaps it is due to the fact that both particles are interacting with the same magnetic field, so when one interacts with a photon, that prevents the other from interacting with that photon. This idea comes from the fact that a rotation necessarily rotates the magnetic field in both terms because it is the same field that is interacting with both particles.

5 Annotated Bibliography

- Group Theory and Quantum Mechanics by Tinkham - The main Lemma derivation of this paper follows that given in section 5.4
- Principles of Quantum Mechanics by Shankar - Chapter 20 derives the Dirac equation with spin terms