Pauli Spin Matrices

It is a bit awkward to picture the wavefunctions for electron spin because – the electron isn't spinning in normal 3D space, but in some internal dimension that is "rolled up" inside the electron. We have invented abstract states "α" and "β" that represent the two possible orientations of the electron spin, but because there isn't a classical analog for spin we can't draw "α" and "β" wavefunctions. This situation comes up frequently in chemistry. We will often deal with molecules that are large and involve many atoms, each of which has a nucleus and many electrons … and it will simply be impossible for us to picture the wavefunction describing all the particles at once. Visualizing one-particle has been hard enough! In these situations, it is most useful to have an **abstract** way of manipulating operators and wavefunctions without looking explicitly at what the wavefunction or operator looks like in real space. The wonderful tool that we use to do this is called **Matrix Mechanics** (as opposed to the wave mechanics we have been using so far). We will use the simple example of spin to illustrate how matrix mechanics works.

The basic idea is that we can write any electron spin state as a linear combination of the two states α and β:

$$\psi \equiv c_\alpha \alpha + c_\beta \beta$$

Note that, for now, we are ignoring the spatial part of the electron wavefunction (e.g. the angular and radial parts of ψ). You might ask how we can be sure that *every* state can be written in this fashion. To assure yourself that this is true, note that for this state, the probability of finding the electron with spin "up" ("down") is $|c_\alpha|^2$ ($|c_\beta|^2$). If there was a state that could not be written in this fashion, there would have to be some other spin state, γ, so that

$$\psi \equiv c_\alpha \alpha + c_\beta \beta + c_\gamma \gamma$$

In this case, however, there would be a probability $|c_\gamma|^2$ of observing the electron with spin γ – which we know experimentally is impossible, as the electron only has two observable spin states.

The basic idea of matrix mechanics is then to **replace the wavefunction with a vector**:

$$\psi \equiv c_\alpha \alpha + c_\beta \beta \quad \rightarrow \quad \mathbf{\psi} \equiv \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix}$$

Note that this is not a vector in physical (x,y,z) space but just a convenient way to arrange the coefficients that define ψ. In particular, this is a nice way to put a wavefunction into a computer, as computers are very adept at dealing with vectors.

Now, our goal is to translate everything that we might want to do with the wavefunction ψ into something we can do to the vector $\mathbf{\psi}$. By going through this step-by-step, we arrive at a few rules

Integrals are replaced with dot products. We note that the overlap between any two wavefunctions can be written as a modified dot product between the vectors. For example, if $\phi \equiv d_\alpha \alpha + d_\beta \beta$ then:

$$\int \phi^* \psi \, d\sigma = d_\alpha^* c_\alpha \underbrace{\int \alpha^* \alpha \, d\sigma}_{1} + d_\alpha^* c_\beta \underbrace{\int \alpha^* \beta \, d\sigma}_{0} + d_\beta^* c_\alpha \underbrace{\int \beta^* \alpha \, d\sigma}_{0} + d_\beta^* c_\beta \underbrace{\int \beta^* \beta \, d\sigma}_{1}$$

$$= d_\alpha^* c_\alpha + d_\beta^* c_\beta$$

$$= \begin{pmatrix} d_\alpha^* & d_\beta^* \end{pmatrix} \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix}$$

$$\equiv \mathbf{\phi}^\dagger \cdot \mathbf{\psi}$$

Where, on the last line, we defined the **adjoint** of a vector:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^\dagger \equiv \begin{pmatrix} x_1^* & x_2^* \end{pmatrix}.$$

Thus, complex conjugation of the wavefunction is replaced by taking the adjoint of a vector. Note that we must take the transpose of the vector and not just its complex conjugate. Otherwise when we took the overlap we would have (column) x (column) and the number of rows and columns would not match.

These rules lead us to natural definitions of normalization and orthogonality of the wavefunctions in terms of the vectors:

$$\int \psi^* \psi \, d\sigma = 1 \quad \rightarrow \quad \mathbf{\psi}^\dagger \cdot \mathbf{\psi} = \begin{pmatrix} c_\alpha^* & c_\beta^* \end{pmatrix} \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix} = |c_\alpha|^2 + |c_\beta|^2 = 1$$

and

$$\int \phi^* \psi \, d\sigma = 0 \quad \rightarrow \quad \boldsymbol{\phi} \cdot \boldsymbol{\psi} = \begin{pmatrix} d_\alpha^* & d_\beta^* \end{pmatrix} \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix} = d_\alpha^* c_\alpha + d_\beta^* c_\beta = 0$$

Finally, we'd like to be able to act operators on our states in matrix mechanics, so that we can compute average values, solve eigenvalue equations, etc. To understand how we should represent operators, we note that in wave mechanics operators turn a wavefunction into another wavefunction. For example, the momentum operator takes one wavefunction and returns a new wavefunction that is the derivative of the original one:

$$\psi(x) \xrightarrow{\hat{p}} \frac{d\psi}{dx}(x)$$

Thus, in order for operators to have the analogous behavior in matrix mechanics, operators must turn vectors into vectors. As it turns out this is the most basic property of a matrix: it turns vectors into vectors. Thus we have our final rule: **Operators are represented by matrices.**

As an illustration, the $\hat{S}_z$ operator will be represented by a 2x2 matrix in spin space:

$$\hat{S}_z \rightarrow \mathbf{S_z} \equiv \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix}$$

We have yet to determine what the elements of this matrix are (hence the question marks) but we can see that this has all the right properties: vectors are mapped into vectors:

$$\hat{S}_z \psi = \psi' \quad \rightarrow \quad \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix} = \begin{pmatrix} c_\alpha' \\ c_\beta' \end{pmatrix}$$

and average values are mapped into numbers

$$\int \psi^* \hat{S}_z \psi \, d\sigma \quad \rightarrow \quad \begin{pmatrix} c_\alpha^* & c_\beta^* \end{pmatrix} \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} \begin{pmatrix} c_\alpha \\ c_\beta \end{pmatrix}$$

Now, the one challenge is that we have to determine which matrix belongs to a given operator – that is, we need to fill in the question marks above. As an exercise, we will show how this is done by working out the matrix representations of $\hat{S}^2, \hat{S}_z$, $\hat{S}_x$ and $\hat{S}_y$.

We'll start with $\hat{S}^2$. We know how $\hat{S}^2$ acts on the α and β wavefunctions:

$$\hat{S}^2 \alpha = \frac{3}{4} \hbar^2 \alpha \qquad \qquad \hat{S}^2 \beta = \frac{3}{4} \hbar^2 \beta$$

Now represent $\hat{S}^2$ as a matrix with unknown elements.

$$\mathbf{S^2} = \begin{pmatrix} c & d \\ e & f \end{pmatrix}$$

In wave mechanics, operating $\hat{S}^2$ on α gives us an eigenvalue back, because α is and eigenfunction of $\hat{S}^2$ (with eigenvalue $\frac{3}{4}\hbar^2$). Translating this into matrix mechanics, when we multiply the matrix $\mathbf{S^2}$ times the vector $\vec{\alpha}$, we should get the same eigenvalue back times $\vec{\alpha}$:

$$\hat{S}^2\alpha = \frac{3}{4}\hbar^2\alpha \quad \rightarrow \quad \mathbf{S^2}\vec{\alpha} = \frac{3}{4}\hbar^2\vec{\alpha}$$

In this way, instead of talking about eigen*functions*, in matrix mechanics we talk about eigen**vectors** and we say that $\vec{\alpha}$ **is an eigenvector of** $\mathbf{S^2}$.

We can use this information to obtain the unknown constants in the matrix $\mathbf{S^2}$:

$$\hat{S}^2\alpha = \frac{3}{4}\hbar^2\alpha \quad \rightarrow \quad \mathbf{S^2}\vec{\alpha} = \frac{3}{4}\hbar^2\vec{\alpha} \Rightarrow \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \tfrac{3}{4}\hbar^2\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} c \\ e \end{pmatrix} = \begin{pmatrix} \frac{3}{4}\hbar^2 \\ 0 \end{pmatrix}$$

so

$$c = \tfrac{3}{4}\hbar^2 \quad e = 0$$

Operating on β

$$\hat{S}^2\beta = \frac{3}{4}\hbar^2\beta \quad \rightarrow \quad \mathbf{S^2}\vec{\alpha} = \frac{3}{4}\hbar^2\vec{\alpha} \Rightarrow \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \tfrac{3}{4}\hbar^2\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} d \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{3}{4}\hbar^2 \end{pmatrix}$$

Thus:

$$d = 0 \quad f = \tfrac{3}{4}\hbar^2$$

So for $\mathbf{S^2}$ we have

$$\boxed{\mathbf{S^2} = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}$$

We can derive $\mathbf{S}_z$ in a similar manner. We know its eigenstates as well:

$$\hat{S}_z\alpha = \tfrac{\hbar}{2}\alpha \qquad\qquad \hat{S}_z\beta = -\tfrac{\hbar}{2}\beta$$

So that the matrix must satisfy:

$$\hat{S}_z \alpha = \frac{\hbar}{2} \alpha \quad \rightarrow \quad \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} c \\ e \end{pmatrix} = \begin{pmatrix} \frac{\hbar}{2} \\ 0 \end{pmatrix}$$

which gives us

$$c = \frac{\hbar}{2} \quad e = 0$$

Analogous operations on β give

$$\hat{S}_z \beta = -\frac{\hbar}{2} \beta \quad \rightarrow \quad \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} d \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{\hbar}{2} \end{pmatrix}$$

$$d = 0 \quad f = -\frac{\hbar}{2}$$

Compiling these results, we arrive at the matrix representation of $\mathbf{S}_z$:

$$\boxed{\mathbf{S_z} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}$$

Now, we need to obtain $\mathbf{S}_x$ and $\mathbf{S}_y$, which turns out to be a bit more tricky.

We remember from our operator derivation of angular momentum that we can re-write the $\mathbf{S}_x$ and $\mathbf{S}_y$ in terms of raising and lowering operators:

$$\mathbf{S_x} = \frac{1}{2}(\mathbf{S_+} + \mathbf{S_-}) \qquad \mathbf{S_y} = \frac{1}{2i}(\mathbf{S_+} - \mathbf{S_-})$$

where we know that

$$\hat{S}_+ \beta = c_+ \alpha \qquad \hat{S}_+ \alpha = 0 \qquad \text{and} \qquad \hat{S}_- \alpha = c_- \beta \qquad \hat{S}_- \beta = 0$$

where c_+ and c_- are constants to be determined. Therefore for the raising operator we have

$$\hat{S}_+ \beta = c_+ \alpha \quad \rightarrow \quad \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c_+ \\ 0 \end{pmatrix} \qquad\qquad \hat{S}_+ \alpha = 0 \quad \rightarrow \quad \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow d = c_+ \ f = 0 \qquad\qquad\qquad\qquad\qquad\qquad \Rightarrow c = 0 \ e = 0$$

And for the lowering operator

$$\hat{S}_- \alpha = c_- \beta \quad \rightarrow \quad \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ c_- \end{pmatrix} \qquad\qquad \hat{S}_- \beta = 0 \quad \rightarrow \quad \begin{pmatrix} c & d \\ e & f \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow c = 0 \ e = c_- \qquad\qquad\qquad\qquad\qquad\qquad \Rightarrow d = 0 \ f = 0$$

Thus

$$\mathbf{S_+} = c_+ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \mathbf{S_-} = c_- \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Therefore we find for S_x and S_y

$$S_x = \tfrac{1}{2}(S_+ + S_-) = \frac{1}{2}\begin{pmatrix} 0 & c_+ \\ c_- & 0 \end{pmatrix} \qquad S_y = \tfrac{1}{2i}(S_+ - S_-) = \frac{1}{2}\begin{pmatrix} 0 & -ic_+ \\ ic_- & 0 \end{pmatrix}$$

Thus, we just need to determine the constants: c_+ and c_- We do this by using two bits of information. First, we note that S_x and S_y are observable, so they must be Hermitian. Thus,

$$\int \alpha^* \hat{S}_x \beta\, d\sigma = \left(\int \beta^* \hat{S}_x \alpha\, d\sigma \right)^* \quad \rightarrow \quad \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \tfrac{1}{2}c_+ \\ \tfrac{1}{2}c_- & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \left(\begin{pmatrix} 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & \tfrac{1}{2}c_+ \\ \tfrac{1}{2}c_- & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)^*$$

$$\Rightarrow c_+ = c_-^{\,*}$$

This result reduces our work to finding one constant (c_+) in terms of which we have

$$S_x = \frac{1}{2}\begin{pmatrix} 0 & c_+ \\ c_+^{\,*} & 0 \end{pmatrix} \qquad S_y = \frac{1}{2}\begin{pmatrix} 0 & -ic_+ \\ ic_+^{\,*} & 0 \end{pmatrix}$$

Note that S_x and S_y have the interesting property that they are equal to their adjoints:

$$S_x^{\dagger} = \frac{1}{2}\begin{pmatrix} 0 & c_+ \\ c_+^{\,*} & 0 \end{pmatrix}^{\dagger} = \frac{1}{2}\begin{pmatrix} 0 & c_+ \\ c_+^{\,*} & 0 \end{pmatrix} = S_x$$

$$S_y^{\dagger} = \frac{1}{2}\begin{pmatrix} 0 & -ic_+ \\ ic_+^{\,*} & 0 \end{pmatrix}^{\dagger} = \frac{1}{2}\begin{pmatrix} 0 & -ic_+ \\ ic_+^{\,*} & 0 \end{pmatrix} = S_y$$

This property turns out to be true in general: **Hermitian operators are represented by matrices that are equal to their own adjoint.** These self-adjoint matrices are typically called Hermitian matrices for this reason, and the adjoint operation is sometimes called Hermitian conjugation.

To determine the remaining constant, we use the fact that $S^2 = S_x^{\,2} + S_y^{\,2} + S_z^{\,2}$. Plugging in our matrix representations for S_x, S_y, S_z and S^2 we find:

$$\frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & c_+ \\ c_+^{\,*} & 0 \end{pmatrix}\begin{pmatrix} 0 & c_+ \\ c_+^{\,*} & 0 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0 & -ic_+ \\ ic_+^{\,*} & 0 \end{pmatrix}\begin{pmatrix} 0 & -ic_+ \\ ic_+^{\,*} & 0 \end{pmatrix}$$

$$= \frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{|c_+|^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{|c_+|^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Rightarrow \frac{3}{4}\hbar^2 = \frac{\hbar^2}{4} + \frac{|c_+|^2}{4} + \frac{|c_+|^2}{4}$$

$$\Rightarrow |c_+|^2 = \hbar^2$$

$$\Rightarrow c_+ = \hbar$$

where on the last line, we have made an arbitrary choice of the sign of c_+. Thus we arrive at the final expressions for S_x and S_y

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

In summary, then, the matrix representations of our spin operators are:

$$\boxed{S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad S^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}$$

Given our representations of S_x, S_y and S_z now we can work out any property we like by doing some linear algebra. For example, we can work out the commutator of S_x and S_y:

$$[S_x, S_y] = S_x S_y - S_y S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$= \frac{\hbar^2}{4}\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right]$$

$$= \frac{\hbar^2}{4}\left[\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\right] = \frac{\hbar^2}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$\Rightarrow [S_x, S_y] = i\hbar S_z$$

It is comforting to see that the matrices that represent the angular momentum operators obey the commutation relations for angular momentum! It is similarly possible to work out the eigenvalues of $\hat{S}^2, \hat{S}_z$ and $\hat{S}_x$ by computing the eigenvalues of S_x, S_y and S_z; average values can be obtained by taking (row)*x(matrix)x(column) products. We are thus in a position to compute anything we want for these operators. As it turns out, any operator in this 2x2 space can be written as a linear combination of S_x, S_y, S_z and S^2. So, in some sense, we now have all the information we could possibly want about spin-$\frac{1}{2}$ systems.

In honor of Pauli, it is conventional to define the dimensionless versions of S_x, S_y, S_z (i.e. matrices without the leading factor of $\hbar/2$) as Pauli Spin matrices:

$$\boxed{\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}$$

One of the reasons the Pauli spin matrices are useful (or equivalently the matrix representations of S_x, S_y and S_z) is that they are the simplest

possible place to practice our skills at using matrix mechanics. We will see later on that more complicated systems can also be described by matrix-vector operations. However, in these cases the matrices will have many more elements – sometimes even infinite numbers of elements! Therefore, it is best to get some practice on these simple 2x2 matrices before taking on the more complicated cases.