PAULI MATRICES: A USEFUL IDENTITY

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

Shankar, R. (1994), *Principles of Quantum Mechanics*, Plenum Press. Chapter 14, Exercise 14.3.4.

[If some equations are too small to read easily, use your browser's magnifying option (Ctrl + on Chrome, probably something similar on other browsers).]

The three components of the spin operator $\mathbf{S}$ for spin $\frac{1}{2}$ can be expressed in terms of the Pauli matrices

$$\sigma_x=\left[\begin{array}{cc}0 & 1\\1 & 0\end{array}\right];\quad\sigma_y=\left[\begin{array}{cc}0 & -i\\i & 0\end{array}\right];\quad\sigma_z=\left[\begin{array}{cc}1 & 0\\0 & -1\end{array}\right] \tag{0.1}$$

We can derive an identity involving the Pauli matrices:

$$\left(\mathbf{A}\cdot\sigma\right)\left(\mathbf{B}\cdot\sigma\right)=\left(\mathbf{A}\cdot\mathbf{B}\right)I+i\left(\mathbf{A}\times\mathbf{B}\right)\cdot\sigma \tag{0.2}$$

One way of proving this is to use the commutation relations for the Pauli matrices. We have

$$\left[\sigma_i,\sigma_j\right]_+ = 2\delta_{ij}I \tag{0.3}$$

$$\left[\sigma_i,\sigma_j\right] = 2i\sum_k\varepsilon_{ijk}\sigma_k \tag{0.4}$$

where ε_{ijk} is the Levi-Civita antisymmetric tensor.

We therefore have

$$\sigma_i\sigma_j = \frac{1}{2}\left(\left[\sigma_i,\sigma_j\right]_+ + \left[\sigma_i,\sigma_j\right]\right) \tag{0.5}$$

$$= \delta_{ij}I+i\sum_k\varepsilon_{ijk}\sigma_k \tag{0.6}$$

Using the summation convention where repeated indices are summed from 1 to 3 (that is, over x, y and z):

$$(\mathbf{A}\cdot\sigma)(\mathbf{B}\cdot\sigma) = A_i\sigma_i B_j\sigma_j \tag{0.7}$$
$$= A_i B_j \sigma_i \sigma_j \tag{0.8}$$
$$= A_i B_j \left(\delta_{ij} I + i\varepsilon_{ijk}\sigma_k\right) \tag{0.9}$$
$$= A_i B_i I + i\varepsilon_{ijk} A_i B_j \sigma_k \tag{0.10}$$
$$= (\mathbf{A}\cdot\mathbf{B}) I + i(\mathbf{A}\times\mathbf{B})\cdot\sigma \tag{0.11}$$

where the last term on the RHS follows from writing the vector cross product in terms of ε_{ijk}. [Note that in the second line, we've assumed that $\mathbf{B}$ commutes with σ.]

Another way of deriving this result is as follows. First, we add the 2×2 identity matrix I to the set of Pauli matrices, calling it $\sigma_0 \equiv I$. Then, because we have four independent matrices (Shankar shows they are linearly independent in his equations 14.3.40-41) each with 4 entries, we can write any 2×2 complex matrix as a linear combination of the σ_α (where a Greek subscript ranges from 0 to 3). That is, for a general 2×2 matrix M

$$M = \sum_\alpha m_\alpha \sigma_\alpha \tag{0.12}$$

From the trace identities

$$\mathrm{Tr}\left(\sigma_\alpha\sigma_\beta\right) = 2\delta_{\alpha\beta} \tag{0.13}$$

we can find m_α by right-multiplying by σ_β and taking the trace:

$$\mathrm{Tr}\left(M\sigma_\beta\right) = \sum_\alpha m_\alpha \mathrm{Tr}\left(\sigma_\alpha\sigma_\beta\right) \tag{0.14}$$
$$= 2\sum_\alpha m_\alpha \delta_{\alpha\beta} \tag{0.15}$$
$$= 2m_\beta \tag{0.16}$$

Thus

$$m_\alpha = \frac{1}{2}\mathrm{Tr}\left(M\sigma_\alpha\right) \tag{0.17}$$

Returning to 0.2, we can identify (again using the summation convention):

$$M = (\mathbf{A}\cdot\sigma)(\mathbf{B}\cdot\sigma) \tag{0.18}$$
$$= A_i\sigma_i B_j\sigma_j \tag{0.19}$$
$$= m_\alpha\sigma_\alpha \tag{0.20}$$

For $\alpha = 0$ we have

$$m_0 = \frac{1}{2}\mathrm{Tr}(M\sigma_0) \tag{0.21}$$
$$= \frac{1}{2}\mathrm{Tr}(M) \tag{0.22}$$
$$= \frac{1}{2}A_iB_j\mathrm{Tr}(\sigma_i\sigma_j) \tag{0.23}$$
$$= \frac{1}{2}A_iB_j(2\delta_{ij}) \tag{0.24}$$
$$= A_iB_i \tag{0.25}$$
$$= \mathbf{A}\cdot\mathbf{B} \tag{0.26}$$

where we used 0.13 to get the fourth line. This gives us the first term on the RHS of 0.2.

For the other three σ_i coefficients, we can use a similar argument. Consider σ_x.

$$m_x = \frac{1}{2}\mathrm{Tr}(M\sigma_x) \tag{0.27}$$
$$= \frac{1}{2}A_iB_j\mathrm{Tr}(\sigma_i\sigma_j\sigma_x) \tag{0.28}$$

From 0.6 we see that $\sigma_i\sigma_j$ can always be written as a single Pauli matrix σ_α. Thus the product of 3 Pauli matrices $\sigma_i\sigma_j\sigma_x$ can be reduced to a product of 2: $\pm\sigma_\alpha\sigma_x$ (the plus or minus sign is determined by the order in which we multiply the two matrices σ_i and σ_j). However, from 0.13, we see that the trace of $\sigma_\alpha\sigma_x$ is non-zero only if $\alpha = x$. The only way this can happen is if either $i = y$ and $j = z$ or $i = z$ and $j = y$. Therefore we have

$$m_x = \frac{1}{2}A_yB_z\mathrm{Tr}(\sigma_y\sigma_z\sigma_x) + \frac{1}{2}A_zB_y\mathrm{Tr}(\sigma_z\sigma_y\sigma_x) \tag{0.29}$$

(Repeated indices are *not* summed here!) From 0.3 we have

$$\sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x \tag{0.30}$$

Thus

$$\mathrm{Tr}\left(\sigma_y\sigma_z\sigma_x\right) = i\mathrm{Tr}\left(\sigma_x^2\right) = 2i$$

Therefore

$$m_x = i\left(A_yB_z - A_zB_y\right) \tag{0.31}$$

and m_x is the x component of $i\left(\mathbf{A}\times\mathbf{B}\right)$. A similar argument gives m_y and m_z, so putting everything together we again arrive at 0.2.

PINGBACKS

Pingback: General 2x2 matrix in terms of pauli matrices
Pingback: Projection operators for spin-1/2 + spin-1/2