SU(2) and Diagonalizations using the Pauli Basis

Ronak Ramachandran

Jul 24, 2024

1 Preliminaries: Properties of Unitary and Hermitian Matrices

Hermitian and unitary matrices are always diagonalizable, so we can express them in the form

$$\sum_j \alpha_j |j\rangle \langle j| . \tag{1.1}$$

In the case of unitary matrices, these α_j must be of the form $e^{i\theta_j}$ for some $\theta_j \in [0, 2\pi)$, since unitaries preserve norms. For every unitary matrix

$$U = \sum_j e^{i\theta_j} |j\rangle \langle j| , \tag{1.2}$$

there is a Hermitian matrix

$$H = \sum_j \theta_j |j\rangle \langle j| , \tag{1.3}$$

satisfying $U = e^{iH}$. We say this Hermitian matrix H is "the Hamiltonian generating U."

2 SU(2), su(2), and the Pauli Basis

$\mathsf{SU}(n)$ is the group consisting of all $n \times n$ unitary matrices of determinant 1 under the group operation of matrix multiplication. For a unitary $U = \sum_j e^{i\theta_j} |j\rangle \langle j|$ in $\mathsf{SU}(n)$, this means $\det(U) = \prod_j e^{i\theta_j} = 1$, so $\sum_j \theta_j = 0$. In other words, the Hamiltonian generating U must have trace 0. $\mathsf{SU}(n)$ is known as a Lie (pronounced "*LEE*") group. If we multiply the Hamiltonians generating $\mathsf{SU}(n)$ by i, they form what's called the Lie algebra[1] of $\mathsf{SU}(n)$, denoted by $\mathfrak{su}(n)$. We might succinctly say $\mathsf{SU}(n) = e^{\mathfrak{su}(n)}$.

Consider $\mathsf{SU}(2)$. Let's represent elements of $\mathsf{SU}(2)$ by the Hamiltonians generating them. How can we represent an arbitrary 2×2 Hermitian matrix H with trace 0? The defining property of a Hermitian matrix is that it is self adjoint, so H must have real entries along the diagonal and the off-diagonal entries must be complex conjugates. This means that for H to be trace-0, it must be of the form

$$H = \begin{bmatrix} c & a - bi \\ a + bi & -c. \end{bmatrix} \tag{2.1}$$

If we recall the Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.2}$$

we see that we could equivalently represent H as $\vec{v} \cdot \vec{\sigma}$, where $\vec{v} = (a, b, c) \in \mathbb{R}^3$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. Since any trace-0 Hermitian matrix can be represented this way, we say the Pauli matrices (after scaling by i) form a basis for $\mathfrak{su}(2)$. This also lets us see that for every $U \in \mathsf{SU}(2)$ there exists some $\vec{v}$ such that $U = e^{i\vec{v}\cdot\vec{\sigma}}$. For reasons we'll see shortly, a unique such $\vec{v}$ exists satisfying $\|\vec{v}\| \in [0, 2\pi)$.

[1]The details aren't important for this exposition, but we multiply by i to make the elements of $\mathfrak{su}(n)$ skew-Hermitian so that the Lie Algebra is closed under the Lie bracket, which in this case is the commutator $[A, B] = AB - BA$. See wiki.

2.1 Diagonalizing in the Pauli Basis

To see how $H := \vec{v} \cdot \vec{\sigma}$ depends on the vector $\vec{v}$, let's diagonalize it. The characteristic polynomial of H is $\lambda^2 - a^2 - b^2 - c^2$, so the eigenvalues of H are $\pm\sqrt{a^2+b^2+c^2} = \pm\|\vec{v}\|$. Let's denote the associated eigenvectors of H by $|v_\pm\rangle$ so $H|v_\pm\rangle = \pm\|\vec{v}\|\,|v_\pm\rangle$. We will solve for both eigenvectors simultaneously by carefully keeping track of these $\pm$'s. Since scaling an eigenvector by a complex number does not change whether it is an eigenvector, we will choose to make $|v_\pm\rangle$ of unit length, with its first entry real and non-negative.

First, note that if $a = b = 0$, then H is already diagonal with eigenvectors $|v_+\rangle = |0\rangle$ and $|v_-\rangle = |1\rangle$ having eigenvalues c and $-c$, respectively. For the rest of the analysis, we therefore assume $a \neq 0$ or $b \neq 0$, or more succinctly, $a + bi \neq 0$. Note that this implies $\|\vec{v}\| > c$.

Say

$$|v_\pm\rangle = \begin{bmatrix} x_\pm \\ y_\pm \end{bmatrix}. \tag{2.3}$$

For brevity, we will omit the subscripts on $x_\pm$ and $y_\pm$. Since $H|v_\pm\rangle = \pm\|\vec{v}\|\,|v_\pm\rangle$, we have

$$(H \mp \|\vec{v}\|I)\,|v_\pm\rangle = \begin{bmatrix} c \mp \|\vec{v}\| & a - bi \\ a + bi & -c \mp \|\vec{v}\| \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{2.4}$$

Solving, we get

$$y = \frac{a+bi}{c \pm \|\vec{v}\|}x = \pm\frac{a+bi}{\|\vec{v}\| \pm c}x, \tag{2.5}$$

where the rearrangement was done to ensure the denominator is positive. Note that since the matrix in equation 2.4 has determinant 0, $(a+bi)(a-bi) = a^2 + b^2 = (\|\vec{v}\| + c)(\|\vec{v}\| - c)$. We require $|v_\pm\rangle$ to be of unit length with its first entry, x, being real and non-negative, so

$$\langle v_\pm|v_\pm\rangle = x^*x + y^*y = \left(1 + \frac{a^2+b^2}{(\|\vec{v}\| \pm c)^2}\right)x^2 = \left(1 + \frac{\|\vec{v}\| \mp c}{\|\vec{v}\| \pm c}\right)x^2 = \left(\frac{2\|\vec{v}\|}{\|\vec{v}\| \pm c}\right)x^2 = 1, \tag{2.6}$$

meaning

$$x = \sqrt{\frac{\|\vec{v}\| \pm c}{2\|\vec{v}\|}} = \frac{1}{\sqrt{2}}\sqrt{1 \pm \frac{c}{\|\vec{v}\|}}, \tag{2.7}$$

and

$$y = \pm\frac{a+bi}{\|\vec{v}\| \pm c}\sqrt{\frac{\|\vec{v}\| \pm c}{2\|\vec{v}\|}} = \pm(a+bi)\sqrt{\frac{\|\vec{v}\| \mp c}{a^2+b^2}}\sqrt{\frac{1}{2\|\vec{v}\|}} = \pm\frac{1}{\sqrt{2}}\sqrt{1 \mp \frac{c}{\|\vec{v}\|}}\frac{a+bi}{\sqrt{a^2+b^2}}. \tag{2.8}$$

We conclude that for $\vec{v} = (a, b, c) \in \mathbb{R}^3$, $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, and $H := \vec{v} \cdot \vec{\sigma}$, we can diagonalize H as $H = \|\vec{v}\|\,|v_+\rangle\langle v_+| - \|\vec{v}\|\,|v_-\rangle\langle v_-|$, where either $a + bi = 0$ and $|v_+\rangle = |0\rangle$ and $|v_-\rangle = |1\rangle$, or $a + bi \neq 0$ and

$$|v_\pm\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{1 \pm \frac{c}{\|\vec{v}\|}} \\ \pm\sqrt{1 \mp \frac{c}{\|\vec{v}\|}}\frac{a+bi}{\sqrt{a^2+b^2}}. \end{bmatrix} \tag{2.9}$$

The unitary generated by the Hamiltonian H can be similarly diagonalized, that is, $U := e^{iH} = e^{i\|\vec{v}\|}\,|v_+\rangle\langle v_+| + e^{-i\|\vec{v}\|}\,|v_-\rangle\langle v_-|$. Note that scaling $\vec{v}$ by a constant doesn't change $|v_\pm\rangle$, and scaling $\vec{v}$ by a multiple of π simply multiplies U by -1.