The Theoretical Minimum

Quantum Mechanics - Solutions

L03E03

Last version: tales.mbivert.com/on-the-theoretical-minimum-solutions/ or github.com/mbivert/ttm

M. Bivert

May 10, 2023

Exercise 1. *Calculate the eigenvectors and eigenvalues of σ_n. Hint: Assume the eigenvector λ_1 has the form:*

$$\begin{pmatrix}\cos\alpha \\ \sin\alpha\end{pmatrix},$$

where α is an unknown parameter. Plug this vector into the eigenvalue equation and solve for α in terms of θ. Why did we use a single parameter α? Notice that our suggested column vector must have unit length.

Let's recall the context: we're trying to build an operator that allows us to measure the spin of a particle. We've started by building the components of such an operator, each representing our ability to measure the spin along any of the $3D$ axes: σ_x, σ_y and σ_z. Each of them was built from the behavior of the spin we "measured": we extracted from the observed behavior a set of constraints, which allowed us to determine the components of the spin operator:

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

Those are individually fine to measure the spin components along the 3 main axis, but we'd like to measure spin components along an arbitrary axis $\hat{n}$. Such a measure can be performed by an operator constructed as a linear combination of the previous three matrices:

$$\sigma_n = \boldsymbol{\sigma} \cdot \hat{n} = \begin{pmatrix}\sigma_x \\ \sigma_y \\ \sigma_z\end{pmatrix} \cdot \begin{pmatrix}n_x \\ n_y \\ n_z\end{pmatrix} = n_x\sigma_x + n_y\sigma_y + n_z\sigma_z$$

Remark 1. *Remember from your linear algebra courses that matrices can be added and scaled: they form a vector space.*

The present exercise involves an arbitrary spin vector, that is, a linear combination of σ_x, σ_y and σ_z that is of the form:

$$\begin{aligned}
\sigma_n &= \sin\theta\sigma_x + \cos\theta\sigma_z \\
&= \sin\theta\begin{pmatrix}0 & 1 \\ 1 & 0\end{pmatrix} + \cos\theta\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix} \\
&= \begin{pmatrix}\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta\end{pmatrix}
\end{aligned}$$

We're then asked to look for the eigenvalues/eigenvectors of that matrix, that is, we want to understand what kind of spin (states) can be encoded by such a matrix, and which values they can take.

Let's recall that to find the eigenvalues/eigenvectors, we need to diagonalize the matrix: assuming it can be diagonalized, it means that there's a basis where it can be expressed as a diagonal matrix; the change of basis is encoded by a linear map, thus a matrix, and so we must be able to find an invertible matrix P and a diagonal matrix D such that:[1]

$$\sigma_n = PDP^{-1} \quad \Leftrightarrow \quad \sigma_n P = PD$$
$$\Leftrightarrow \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 a & \lambda_2 b \\ \lambda_1 c & \lambda_2 d \end{pmatrix}$$

Where λ_1 and λ_2 would be the eigenvalues, associated to the two eigenvectors:

$$|\lambda_1\rangle = \begin{pmatrix} a \\ c \end{pmatrix}; \qquad |\lambda_2\rangle = \begin{pmatrix} b \\ d \end{pmatrix}$$

Note that the previous equation implies that we must have:

$$(\forall i \in \{1,2\}),\ \sigma_n|\lambda_i\rangle = \lambda_i|\lambda_i\rangle$$

Which is equivalent to saying, where 0_2 is the zero 2×2 matrix, and I_2 the 2×2 identity matrix:

$$\sigma_n|\lambda_i\rangle - \lambda_i|\lambda_i\rangle = 0_2 \quad \Leftrightarrow \quad (\sigma_n - I_2\lambda_i)|\lambda_i\rangle = 0_2$$

If we want a non-trivial solution (i.e. $|\lambda_i\rangle \neq \mathbf{0}$), then it follows that we must have:

$$\sigma_n - I_2\lambda_i = 0_2$$

This means that the matrix $\sigma_n - I_2\lambda_i$ cannot be invertible (for otherwise multiplying it by its inverse would yield, by the rule of invertibility I_2, but on the other side, from the matrix's definition, it would yield 0_2, hence a contradiction, hence it's not invertible).

Non-invertibility of a matrix translates to their determinant being zero, which means the λ_i solves the following equation for λ:

$$\begin{aligned}
\det(\sigma_n - I_2\lambda) = 0 \quad &\Leftrightarrow \quad \begin{vmatrix} \cos\theta - \lambda & \sin\theta \\ \sin\theta & -\cos\theta - \lambda \end{vmatrix} = 0 \\
&\Leftrightarrow \quad -(\cos\theta - \lambda)(\cos\theta + \lambda) - \sin^2\theta = 0 \\
&\Leftrightarrow \quad -(\cos^2\theta - \lambda^2) - \sin^2\theta = 0 \\
&\Leftrightarrow \quad \lambda^2 - \underbrace{(\sin^2\theta + \cos^2\theta)}_{=1} = 0 \\
&\Leftrightarrow \quad \lambda^2 = 1 \\
&\Leftrightarrow \quad \lambda = \begin{cases} 1 & = \lambda_1 \\ -1 & = \lambda_2 \end{cases}
\end{aligned}$$

Now that we have our eigenvalues, we can use them to determine the associated eigenvectors, as, remember, they are linked by:

$$(\forall i \in \{1,2\}),\ \sigma_n|\lambda_i\rangle = \lambda_i|\lambda_i\rangle$$

And so:

$$\begin{aligned}
\sigma_n|\lambda_1\rangle = \lambda_1|\lambda_1\rangle \quad &\Leftrightarrow \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \\
&\Leftrightarrow \quad \begin{cases} a\cos\theta + c\sin\theta & = a \\ a\sin\theta - c\cos\theta & = c \end{cases} \\
&\Leftrightarrow \quad \begin{cases} a(\cos\theta - 1) + c\sin\theta & = 0 \\ a\sin\theta + c(-\cos\theta - 1) & = 0 \end{cases}
\end{aligned}$$

[1] This is "basic" linear algebra; the authors assume that you're already familiar with it to some degree (e.g. matrix product); don't hesitate to refer to a more thorough course on the subject for more. I'll quickly review here how diagonalization works

Consider the first equation of this system: we're left with two main choices, depending on whether $\cos\theta = 1$ or not. If it is, let's take $\theta = 0$ for instance, but this would true modulo π, then we must have $\sin\theta = 0$, and the first equations gives us nothing of value. The second then simplifies to $c = 0$, thus $a = 0$.

Let's now consider the case where $\cos\theta \neq 1$. The system can be rewritten as:

$$\begin{cases} a &= \dfrac{-c\sin\theta}{\cos\theta - 1} \\ a\sin\theta + c(-\cos\theta - 1) &= 0 \end{cases}$$

We can inject the first equation in the second to yield:

$$\begin{aligned} \frac{-c\sin\theta}{\cos\theta - 1}\sin\theta + c(-\cos\theta - 1) = 0 \quad &\Leftrightarrow \quad \frac{-c\sin\theta}{\cos\theta - 1}\sin\theta + \frac{\cos\theta - 1}{\cos\theta - 1}c(-\cos\theta - 1) = 0 \\ &\Leftrightarrow \quad \frac{c\big(-\sin^2\theta - (\cos\theta - 1)(\cos\theta + 1)\big)}{\cos\theta - 1} = 0 \\ &\Rightarrow \quad c(-\sin^2\theta - (\cos^2\theta - 1)) = 0 \\ &\Rightarrow \quad c(-\underbrace{(\sin^2\theta + \cos^2\theta)}_{=1} - 1) = 0 \\ &\Rightarrow \quad c = 0 \Rightarrow a = 0 \end{aligned}$$

That's a struggle; we don't seem to be able to extract anything but the trivial solution; maybe there's some trigonometric trick to find the general solution[2]).

Instead, let's try to use and understand the authors' hint, which is to look for eigenvectors of the form:

$$\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$$

Why is this a reasonable choice? Let's start by answering why we need a single parameter α: it corresponds to the single degree of freedom we have in this case. Let's recall the two equivalent ways of counting the number of degree of freedom that were given in subsection 2.5:

1. First, point the apparatus in any direction in the xz-plane (remember for comparison that in subsection 2.5, we were allowed to take a direction in the xyz-space). A single angle is sufficient to encode this single direction (2 were needed in the xyz space). Furthermore, note that this angle would have has its coordinate in the xz-plane $\cos\alpha$ and $\sin\alpha$, respectively in the x and z directions.

   Note that we're really capturing *directions*: a point in $\mathbb{R}^2$ contains too much information, as we want to identify all the points which share the same direction;

2. The second approach was to say that the general form of the spin state in xyz-space was given by a (complex) linear combination $\alpha_u|u\rangle + \alpha_d|d\rangle$. But, recall the definition of $|l\rangle$ and $|r\rangle$, the vectors associated with the x-direction:

   $$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle; \qquad |l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle$$

   They didn't involved complex numbers. We started to need, and have proven in exercise L02E03 that this was mandatory once we had enough constraints to cover the three spatial directions (i.e., when dealing with $|i\rangle$ and $|o\rangle$, after having already established the two other pairs of orthogonal vectors).

   That's to say, we don't need complex numbers when we only have two directions, so actually, the general form of a spin in a plane is a *real* linear combination, which cuts down the number of

[2] There definitely is one, see for instance `https://www.wolframalpha.com/input?i=diagonalize+%7B%7Bcos+x%2C+sin+x%7D%2C%7Bsin+x%2C-cos+x%7D%7D`

degrees of freedom to 2.

Normalization adds yet another constraint, which cuts us down to a single degree of freedom. But, shouldn't the phase ambiguity brings us to ... *zero* degree of freedom? What are we missing?

Well, the idea of phase ambiguity was that we could multiply the vectors by a $\exp(i\theta) = cos\theta + i\sin\theta$, for $\theta \in \mathbb{R}$. But we saw that we actually don't need complex numbers when we're in a $2D$-plane, which means $\sin\theta = 0$, and thus forces $\cos\theta = 1$, so the phase ambiguity doesn't impact the number of degrees of freedom;

3. Here's a third argument that we'll re-use in the next exercise[3]. Consider as a first guess an eigenvector of the form

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}; \qquad (z_1, z_2) \in \mathbb{C}^2$$

We can put both complex numbers in exponential form:

$$\begin{pmatrix} r_1 \exp(i\phi_1) \\ r_2 \exp(i\phi_2 \end{pmatrix} = \exp(i\phi_1) \begin{pmatrix} r_1 \\ r_2 \exp(i(\phi_2 - \phi_1)) \end{pmatrix}; \qquad (r_1, r_2, \phi_1, \phi_2) \in \mathbb{R}^4$$

We can then ignore the general phase factor $\exp(i\phi_1)$, e.g. choose $\phi_1 = 0$. Furthermore, we'll want the (eigen)vector to be normalized (remember, the eigenvector associated to the eigenvalues of of a Hermitian operator make an orthonormal basis), i.e.:

$$|r_1|^2 + |r_2 \exp(i\phi_2)|^2 = 1 \Leftrightarrow |r_1|^2 + |r_2|^2 = 1$$

But we're then losing a degree of freedom, meaning, r_1 and r_2 are not independent from each other: we can express them both in term of a single parameter, as long as the previous equation is satisfied. We can choose, as it'll make computation easier, $r_1 = \cos\alpha$, $r_2 = \sin\alpha$, with $\alpha \in \mathbb{R}$. Which brings us to:

$$\begin{pmatrix} \cos\alpha \\ \exp(i\phi_2)\sin\alpha \end{pmatrix}$$

If ϕ_2 varies, then our eigenvector isn't restricted to a plane. But, because our eigenvector will be a eigenvector of a Hermitian matrix, we know by the real spectral theorem[4] that it must be a (an orthonormal basis) vector of the xz-plane. So we can choose $\phi_2 = 0$ to restrict it to a plane.

Note that the form of this vector is naturally normalized ($\cos^2\alpha + \sin^2\alpha = 1$). Recall that it *must* be normalized because this column vector actually corresponds to:

$$\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \cos\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sin\alpha \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \cos\alpha|u\rangle + \sin\alpha|d\rangle$$

And the square of the magnitude of $\cos\alpha$ encodes the probability for the measured value to correspond to $|u\rangle$ while the square of the magnitude of $\sin\alpha$ encodes the probability of the system to be measured in state $|d\rangle$, and both states are orthogonal: the total probability must be 1.

Alright, let's get to actually finding the eigenvectors associated to our eigenvalues. We can use the same trick as in the previous exercise L03E02.pdf: because of the diagonalization process, we have the following relation:

$$\sigma_n = PDP^{-1} \Leftrightarrow \sigma_n P = PD(\underbrace{PP^{-1}}_{:=I_2}) = PD = P\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\Leftrightarrow \underbrace{\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}}_{=\sigma_n} \underbrace{\begin{pmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{pmatrix}}_{=P} = \begin{pmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\cos\beta \\ \cos\alpha & -\sin\beta \end{pmatrix}$$

Where the columns of P are the eigenvectors associated to the eigenvalues 1 and -1. Both have the same "form", as previously explained. We could have used the same approach as in the book (see the

[3]Source: `https://physics.stackexchange.com/a/720025`
[4]L03E01.pdf

previous exercise), but you'll get with the same (kind?) of system in the end. Let's perform the matrix multiplication on the left and extract two equations from the four we can get by identifying the matrix components:

$$\begin{pmatrix} \cos\theta\cos\alpha + \sin\theta\sin\alpha & \cos\theta\cos\beta + \sin\theta\sin\beta \\ \sin\theta\cos\alpha - \cos\theta\sin\alpha & \sin\theta\cos\beta - \cos\theta\sin\beta \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\cos\beta \\ \cos\alpha & -\sin\beta \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} \cos\theta\cos\alpha + \sin\theta\sin\alpha = \cos\alpha \\ \cos\theta\cos\beta + \sin\theta\sin\beta = -\cos\beta \end{cases}$$

Remark 2. *Strictly speaking, we don't really know if this is equivalent so far, as we're just extracting two equations from potentially four distinct equations. For correctness' sake, we could (I won't out of laziness) verify that the solution we find for those two equations also solve the two other remaining equations.*

The following trigonometric identities[5]:

$$\cos\theta\cos\alpha = \frac{1}{2}(\cos(\theta-\alpha) + \cos(\theta+\alpha)); \qquad \sin\theta\sin\alpha = \frac{1}{2}(\cos(\theta-\alpha) - \cos(\theta+\alpha))$$

$$\cos(\alpha - \pi) = -\cos\alpha$$

Allows us to rewrite the previous system as

$$\Leftrightarrow \begin{cases} \frac{1}{2}\Big(\big(\cos(\theta-\alpha) + \cos(\theta+\alpha)\big) + \big(\cos(\theta-\alpha) - \cos(\theta+\alpha)\big)\Big) = \cos\alpha \\ \frac{1}{2}\Big(\big(\cos(\theta-\beta) + \cos(\theta+\beta)\big) + \big(\cos(\theta-\beta) - \cos(\theta+\beta)\big)\Big) = \cos(\beta-\pi) \end{cases}$$

$$\Leftrightarrow \begin{cases} \cos(\theta-\alpha) = \cos\alpha \\ \cos(\theta-\beta) = \cos(\beta-\pi) \end{cases}$$

And with the following identities:

$$\cos(\alpha + \frac{\pi}{2}) = -\sin\alpha; \qquad \sin(\alpha + \frac{\pi}{2}) = \cos\alpha$$

We reach:

$$\Rightarrow \begin{cases} \theta - \alpha = \alpha \\ \theta - \beta = \beta - \pi \end{cases} \quad \Rightarrow \begin{cases} \alpha = \frac{\theta}{2} \\ \beta = \frac{1}{2}(\theta + \pi) \end{cases} \quad \Rightarrow \boxed{\begin{cases} |+1\rangle = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \\ |-1\rangle = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix} = \begin{pmatrix} \cos(\theta/2+\pi/2) \\ \sin(\theta/2+\pi/2) \end{pmatrix} = \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix} \end{cases}}$$

[5]Look around for the proofs if needed; formulas can be found on Wikipedia