The Theoretical Minimum

Quantum Mechanics - Solutions

M. Bivert

April 20, 2023

Abstract

Below are solution proposals to the exercises of *The Theoretical Minimum - Quantum Mechanics*, written by Leonard Susskind and Art Friedman. An effort has been so as to recall from the book all the referenced equations, and to be rather verbose regarding mathematical details, rather in line with the general tone of the series.

Contents

1 Systems and Experiments

1.1 Inner Products

Exercise 1. *a) Using the axioms for inner products, prove*

$$\Big(\langle A| + \langle B|\Big)|C\rangle = \langle A|C\rangle + \langle B|C\rangle$$

b) Prove $\langle A|A\rangle$ is a real number.

---

a) Let us recall the two axioms in question:

Axiom 1.

$$\langle C|\Big(|A\rangle + |B\rangle\Big) = \langle C|A\rangle + \langle C|B\rangle$$

Axiom 2.

$$\langle B|A\rangle = \langle A|B\rangle^*$$

Where z^* is the complex conjugate of $z \in \mathbb{C}$

Let us recall also that if

- $\langle A|$ is the bra of $|A\rangle$
- $\langle B|$ is the bra of $|B\rangle$

Then $\langle A| + \langle B|$ is the bra of $|A\rangle + |B\rangle$.

Let us also observe that for $(a, b) = (x_a + iy_a, x_b + iy_b) \in \mathbb{C}^2$:

$$\begin{aligned}(a+b)^* &= (x_a + iy_a + x_b + iy_b)^* \\ &= x_a - iy_a + x_b - iy_b \\ &= a^* + b^*\end{aligned}$$

We thus have:

$$\begin{aligned}\Big(\langle A| + \langle B|\Big)|C\rangle &= \langle C|\Big(|A\rangle + |B\rangle\Big)^* \\ &= \Big(\langle C|A\rangle + \langle C|B\rangle\Big)^* \\ &= \langle C|A\rangle^* + \langle C|B\rangle^* \\ &= \langle A|C\rangle + \langle B|C\rangle \quad \square\end{aligned}$$

b) Mainly from the second axiom:

$$\begin{aligned} x + iy &= \langle A|A\rangle \\ &= \langle A|A\rangle^* \\ &= x - iy \\ &\Rightarrow 2iy = 0 \\ &\Rightarrow y = 0 \\ &\Rightarrow \langle A|A\rangle = x \in \mathbb{R} \quad \square \end{aligned}$$

Exercise 2. *Show that the inner product defined by Eq.* 1.2 *satisfies all the axioms of inner products.*

Let us recall the two axioms in question:

Axiom 3.

$$\langle C|\Big(|A\rangle + |B\rangle\Big) = \langle C|A\rangle + \langle C|B\rangle$$

Axiom 4.

$$\langle B|A\rangle = \langle A|B\rangle^*$$

Where z^* is the complex conjugate of $z \in \mathbb{C}$

And let us recall Eq. 1.2 of the book:

$$\begin{aligned} \langle B|A\rangle &= \begin{pmatrix} \beta_1^* & \beta_2^* & \beta_3^* & \beta_4^* & \beta_5^* \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix} \\ &= \beta_1^*\alpha_1 + \beta_2^*\alpha_2 + \beta_3^*\alpha_3 + \beta_4^*\alpha_4 + \beta_5^*\alpha_5 \end{aligned}$$

For the first axiom, considering $\langle C| = (\gamma_i^*)$:

$$\begin{aligned} \langle C|\Big(|A\rangle + |B\rangle\Big) &= \begin{pmatrix} \gamma_1^* & \gamma_2^* & \gamma_3^* & \gamma_4^* & \gamma_5^* \end{pmatrix} \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \alpha_3 + \beta_3 \\ \alpha_4 + \beta_4 \\ \alpha_5 + \beta_5 \end{pmatrix} \\ &= \gamma_1^*(\alpha_1 + \beta_1) + \gamma_2^*(\alpha_2 + \beta_2) + \gamma_3^*(\alpha_3 + \beta_3) + \gamma_4^*(\alpha_4 + \beta_4) + \gamma_5^*(\alpha_5 + \beta_5) \\ &= \big(\gamma_1^*\alpha_1 + \gamma_2^*\alpha_2 + \gamma_3^*\alpha_3 + \gamma_4^*\alpha_4 + \gamma_5^*\alpha_5\big) + \big(\gamma_1^*\beta_1 + \gamma_2^*\beta_2 + \gamma_3^*\beta_3 + \gamma_4^*\beta_4 + \gamma_5^*\beta_5\big) \\ &= \begin{pmatrix} \gamma_1^* & \gamma_2^* & \gamma_3^* & \gamma_4^* & \gamma_5^* \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix} + \begin{pmatrix} \gamma_1^* & \gamma_2^* & \gamma_3^* & \gamma_4^* & \gamma_5^* \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{pmatrix} \\ &= \langle C|A\rangle + \langle C|B\rangle \quad \square \end{aligned}$$

Before diving into the second axiom, let us observe that for $(a, b) = (x_a + iy_a, x_b + iy_b) \in \mathbb{C}^2$:

$$\begin{aligned} (ab)^* &= \Big((x_a + iy_a) \times (x_b + iy_b)\Big)^* \\ &= \Big(x_a x_b - y_a y_b + i(x_b y_a + x_a y_b)\Big)^* \\ &= x_a x_b - y_a y_b - i(x_b y_a + x_a y_b) \\ &= (x_a - iy_a) \times (x_b - iy_b) \\ &= a^* b^* \end{aligned}$$

Or, perhaps more simply using complex numbers' exponential's form:

$$\begin{aligned}(ab)^* &= \left(r_a r_b e^{i(\theta_a+\theta_b)}\right)^* \\ &= r_a r_b e^{-i(\theta_a+\theta_b)} \\ &= a^* b^*\end{aligned}$$

Hence, regarding the second axiom:

$$\begin{aligned}\langle B|A\rangle &= \Big(\big(\langle B|A\rangle\big)^*\Big)^* \\ &= \Big(\big(\beta_1^*\alpha_1 + \beta_2^*\alpha_2 + \beta_3^*\alpha_3 + \beta_4^*\alpha_4 + \beta_5^*\alpha_5\big)^*\Big)^* \\ &= \Big(\beta_1\alpha_1^* + \beta_2\alpha_2^* + \beta_3\alpha_3^* + \beta_4\alpha_4^* + \beta_5\alpha_5^*\Big)* \\ &= \Big(\alpha_1^*\beta_1 + \alpha_2^*\beta_2 + \alpha_3^*\beta_3 + \alpha_4^*\beta_4 + \alpha_5^*\beta_5\Big)* \\ &= \left(\begin{pmatrix}\alpha_1^* & \alpha_2^* & \alpha_3^* & \alpha_4^* & \alpha_5^*\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_5\end{pmatrix}\right)^* \\ &= \langle A|B\rangle^* \quad \square\end{aligned}$$

2 Quantum States

2.1 Along the x Axis

Exercise 3. *Prove that the vector $|r\rangle$ in Eq. 2.5 is orthogonal to vector $|l\rangle$ in Eq. 2.6.*

Let us recall respectively Eq. 2.5 and Eq. 2.6:

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

Orthogonality can be detected with the inner-product: $|l\rangle$ and $|r\rangle$ are orthogonals $\Leftrightarrow \langle r|l\rangle = \langle l|r\rangle = 0$.

Remark 1.
The nullity of either inner-product is sufficient, because of the $\langle A|B\rangle = \langle B|A\rangle^*$ axiom.

For instance:

$$\begin{aligned}\langle l|r\rangle &= \begin{pmatrix}\lambda_u^* & \lambda_d^*\end{pmatrix}\begin{pmatrix}\rho_u\\ \rho_d\end{pmatrix} \\ &= \begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \\ &= 0 \quad \square\end{aligned}$$

Or, similarly:

$$\begin{aligned}\langle r|l\rangle &= \begin{pmatrix}\rho_u^* & \rho_d^*\end{pmatrix}\begin{pmatrix}\lambda_u \\ \lambda_d\end{pmatrix} \\ &= \begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}}\end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \\ &= 0 \quad \square\end{aligned}$$

2.2 Along the y Axis

Exercise 4. *Prove that $|i\rangle$ and $|o\rangle$ satisfy all of the conditions in Eqs.* 2.7, 2.8 *and* 2.9. *Are they unique in that respect?*

Let us recall, in order, Eqs. 2.7, 2.8, 2.9, 2.10, which defines $|i\rangle$ and $|o\rangle$, and both 2.5 and 2.6 which defines $|r\rangle$ and $|l\rangle$:

$$\langle i|o\rangle = 0$$

$$\langle o|u\rangle\langle u|o\rangle = \frac{1}{2} \qquad \langle o|d\rangle\langle d|o\rangle = \frac{1}{2}$$
$$\langle i|u\rangle\langle u|i\rangle = \frac{1}{2} \qquad \langle i|d\rangle\langle d|i\rangle = \frac{1}{2}$$

$$\langle o|r\rangle\langle r|o\rangle = \frac{1}{2} \qquad \langle o|l\rangle\langle l|o\rangle = \frac{1}{2}$$
$$\langle i|r\rangle\langle r|i\rangle = \frac{1}{2} \qquad \langle i|l\rangle\langle l|i\rangle = \frac{1}{2}$$

$$|i\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle \qquad |o\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle$$

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

For clarity, let us recall that $\langle u|A\rangle$ is the component of $|A\rangle$ on the orthonormal vector $|u\rangle$. This is because in a $(|i\rangle)_{i\in F}$ *orthonormal* basis we have:

$$\begin{aligned}|A\rangle &= \sum_{i\in F}\alpha_i|i\rangle \\ \Rightarrow \langle j|A\rangle &= \langle j|\sum_{i\in F}\alpha_i|i\rangle = \sum_{i\in F}\alpha_i\langle j|i\rangle = \alpha_j\end{aligned}$$

And to make better sense of those equations, let us recall that $\alpha_u^*\alpha_u = \langle A|u\rangle\langle u|A\rangle$ is the probability of a state vector $|A\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle$ to be measured in the state $|u\rangle$.

For Eq. 2.7, we have

$$\begin{aligned}\langle i|o\rangle &= \begin{pmatrix}\iota_u^* & \iota_d^*\end{pmatrix}\begin{pmatrix}o_u\\ o_d\end{pmatrix}\\ &= \iota_u^* o_u + \iota_d^* o_d\\ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{-i}{\sqrt{2}}\frac{-i}{\sqrt{2}} = \frac{1}{2} - \frac{1}{2} = 0 \quad \square\end{aligned}$$

For Eqs. 2.8, we can rely on the projection on an orthonormal vector:

$$\langle o|u\rangle\langle u|o\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} = \frac{1}{2} \quad \square \qquad \langle o|d\rangle\langle d|o\rangle = \frac{i}{\sqrt{2}}\frac{-i}{\sqrt{2}} = \frac{1}{2} \quad \square$$

$$\langle i|u\rangle\langle u|i\rangle = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} = \frac{1}{2} \quad \square \qquad \langle i|d\rangle\langle d|i\rangle = \frac{-i}{\sqrt{2}}\frac{i}{\sqrt{2}} = \frac{1}{2} \quad \square$$

For Eqs. 2.9, we need to rely on the column form of the inner-product:

$$\begin{aligned}\langle o|r\rangle\langle r|o\rangle &= \begin{pmatrix}o_u^* & o_d^*\end{pmatrix}\begin{pmatrix}\rho_u\\ \rho_d\end{pmatrix}\begin{pmatrix}\rho_u^* & \rho_d^*\end{pmatrix}\begin{pmatrix}o_u\\ o_d\end{pmatrix}\\ &= (\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\frac{1}{\sqrt{2}})(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\frac{-i}{\sqrt{2}})\\ &= (\frac{1}{2} + \frac{i}{2})(\frac{1}{2} - \frac{i}{2})\\ &= \frac{1}{4}(1+i)(1-i)\\ &= \frac{1}{4}(1+i-i+1) = \frac{1}{2} \quad \square\end{aligned}$$

$$\begin{aligned}\langle o|l\rangle\langle l|o\rangle &= \begin{pmatrix}o_u^* & o_d^*\end{pmatrix}\begin{pmatrix}\lambda_u\\ \lambda_d\end{pmatrix}\begin{pmatrix}\lambda_u^* & \lambda_d^*\end{pmatrix}\begin{pmatrix}o_u\\ o_d\end{pmatrix}\\ &= (\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\frac{-1}{\sqrt{2}})(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{-1}{\sqrt{2}}\frac{-i}{\sqrt{2}})\\ &= (\frac{1}{2} - \frac{i}{2})(\frac{1}{2} + \frac{i}{2})\\ &= \frac{1}{4}(1-i)(1+i)\\ &= \frac{1}{4}(1-i+i+1) = \frac{1}{2} \quad \square\end{aligned}$$

$$\begin{aligned}\langle i|r\rangle\langle r|i\rangle &= \begin{pmatrix}\iota_u^* & \iota_d^*\end{pmatrix}\begin{pmatrix}\rho_u\\ \rho_d\end{pmatrix}\begin{pmatrix}\rho_u^* & \rho_d^*\end{pmatrix}\begin{pmatrix}\iota_u\\ \iota_d\end{pmatrix}\\ &= (\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{-i}{\sqrt{2}}\frac{1}{\sqrt{2}})(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\frac{i}{\sqrt{2}})\\ &= (\frac{1}{2} - \frac{i}{2})(\frac{1}{2} + \frac{i}{2})\\ &= \frac{1}{4}(1-i)(1+i)\\ &= \frac{1}{4}(1+i+i+1) = \frac{1}{2} \quad \square\end{aligned}$$

$$\begin{aligned}\langle i|l\rangle\langle l|i\rangle &= \begin{pmatrix}\iota_u^* & \iota_d^*\end{pmatrix}\begin{pmatrix}\lambda_u\\ \lambda_d\end{pmatrix}\begin{pmatrix}\lambda_u^* & \lambda_d^*\end{pmatrix}\begin{pmatrix}\iota_u\\ \iota_d\end{pmatrix}\\ &= (\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{-i}{\sqrt{2}}\frac{-1}{\sqrt{2}})(\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} + \frac{-1}{\sqrt{2}}\frac{i}{\sqrt{2}})\\ &= (\frac{1}{2} + \frac{i}{2})(\frac{1}{2} - \frac{i}{2})\\ &= \frac{1}{4}(1+i)(1-i)\\ &= \frac{1}{4}(1+i-i+1) = \frac{1}{2} \quad \square\end{aligned}$$

Regarding the unicity of $|i\rangle, |o\rangle$, as for $|r\rangle, |l\rangle$, there definitely is a phase ambiguity, meaning, we can multiply either $|i\rangle$ or $|o\rangle$ by a *phase factor*, say $e^{i\theta}$, without disturbing any of the constraints: orthogonality, probabilities, and the resulting vectors are still unitary.

But as stated by the authors for $|r\rangle, |l\rangle$, measurable quantities are independant of any phase factors. So up to it, they seem to be unique so far.

However, let's try to change the i's place for instance in $|i\rangle$:

$$|i\rangle = \frac{i}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle$$

The vector is still unitary, we still have orthogonality with $|o\rangle$, and if you try to compute $\langle i|u\rangle\langle u|i\rangle$, $\langle i|d\rangle\langle d|i\rangle$, $\langle i|r\rangle\langle r|i\rangle$ or $\langle i|l\rangle\langle l|i\rangle$, you'll still have the same probablities.
Now the question is, is this "swapping" of the i a phase factor? Meaning, can encode this transformation as a multiplication by some $e^{i\theta}$, for some $\theta \in \mathbb{R}$?

Well, the first term of $|i\rangle$ is multiplied by i; recall the definition of the complex exponential:

$$e^{i\theta} = \cos\theta + i\sin\theta$$

So this means the first term is multiplied by

$$\exp(i\frac{\pi}{2}) = 0 + i$$

The second term though, is multiplied by $-i$, this means, multiplied by:

$$\exp(-i\frac{\pi}{2}) = 0 + i \times (-1)$$

So we've found a variant of $|i\rangle$, that cannot be obtained by multiplying $|i\rangle$ by a phase factor, and hence:

The proposed solution is *not* unique [up to a phase factor].

Remark 2. *It may be interesting/possible to classify all such variants, meaning, see how much variety there is / how much structure they share and so forth.*

Exercise 5. *For the moment, forget that Eqs.* 2.10 *give us working definitions for* $|i\rangle$ *and* $|o\rangle$ *in terms of* $|u\rangle$ *and* $|d\rangle$*, and assume that the components* α, β, γ *and* δ *are unknown:*

$$|o\rangle = \alpha|u\rangle + \beta|d\rangle \qquad\qquad |i\rangle = \gamma|u\rangle + \delta|d\rangle$$

a) Use Eqs. 2.8 *to show that*

$$\alpha^*\alpha = \beta^*\beta = \gamma^*\gamma = \delta^*\delta = \frac{1}{2}$$

b) Use the above results and Eqs. 2.9 *to show that*

$$\alpha^*\beta + \alpha\beta^* = \gamma^*\delta + \gamma\delta^* = 0$$

c) Show that $\alpha^*\beta$ *and* $\gamma^*\delta$ *must each be pure imaginary.*

If $\alpha^*\beta$ *is pure imaginary, then* α *and* β *cannot both be real. The same reasoning applies to* $\gamma^*\delta$*.*

Let's start by recalling Eqs. 2.8, 2.9 and 2.10, which are respectively:

$$\begin{aligned} \langle o|u\rangle\langle u|o\rangle = \frac{1}{2} \qquad \langle o|d\rangle\langle d|o\rangle = \frac{1}{2} \\ \langle i|u\rangle\langle u|i\rangle = \frac{1}{2} \qquad \langle i|d\rangle\langle d|i\rangle = \frac{1}{2} \end{aligned} \tag{1}$$

$$\begin{aligned} \langle o|r\rangle\langle r|o\rangle = \frac{1}{2} \qquad \langle o|l\rangle\langle l|o\rangle = \frac{1}{2} \\ \langle i|r\rangle\langle r|i\rangle = \frac{1}{2} \qquad \langle i|l\rangle\langle l|i\rangle = \frac{1}{2} \end{aligned} \tag{2}$$

$$|i\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle \qquad |o\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle \tag{3}$$

---

a) Let's start by recalling that the inner-product in a Hilbert space is defined between a bra and a ket, and that it should satisfy the following two axioms:

$$\langle C|\{|A\rangle + |B\rangle\} = \langle C|A\rangle + \langle C|B\rangle \text{ (linearity)}$$

$$\langle B|A\rangle = \langle A|B\rangle^* \text{ (complex conjugation)}$$

Furthermore, the scalar-multiplication of a ket is linear:

$$z \in \mathbb{C}, \qquad |zA\rangle = z|A\rangle$$

Then we can multiply $|o\rangle = \alpha|u\rangle + \beta|d\rangle$ to the left by $\langle u|$ to compute $\langle u|o\rangle$, using the linearity of the inner-product/scalar multiplication, and the fact that $|u\rangle$ and $|d\rangle$ are, by definition, unitary orthogonal vectors (meaning, $\langle u|d\rangle = 0$ and $\langle u|u\rangle = \langle d|d\rangle = 1$)

$$\langle u|o\rangle = \alpha \langle u|u\rangle + \beta \langle u|d\rangle = \alpha$$

Because of the complex conjugation rule, we have

$$\langle o|u\rangle = \langle u|o\rangle^* = \alpha^*$$

And so by Eqs. 2.8 and the previous computation we have

$$\frac{1}{2} = \underbrace{\langle o|u\rangle}_{\alpha} \underbrace{\langle u|o\rangle}_{\alpha^*} = \alpha\alpha^* \quad \square$$

The process is very similar to prove $\beta^*\beta = \gamma^*\gamma = \delta^*\delta = \frac{1}{2}$:

$$\begin{aligned}
\frac{1}{2} &= \langle o|d\rangle \langle d|o\rangle \\
&= (\langle d|o\rangle)^* \langle d|o\rangle \\
&= \Big(\langle d|\{\alpha|u\rangle + \beta|d\rangle\}\Big)^* \Big(\langle d|\{\alpha|u\rangle + \beta|d\rangle\}\Big) \\
&= \Big(\alpha \underbrace{\langle d|u\rangle}_{=0} + \beta \underbrace{\langle d|d\rangle}_{=1}\Big)^* \Big(\alpha \underbrace{\langle d|u\rangle}_{=0} + \beta \underbrace{\langle d|d\rangle}_{=1}\Big) \\
&= \beta^*\beta \quad \square
\end{aligned}$$

$$\begin{aligned}
\frac{1}{2} &= \langle i|u\rangle \langle u|i\rangle \\
&= (\langle u|i\rangle)^* \langle u|i\rangle \\
&= \Big(\langle u|\{\gamma|u\rangle + \delta|d\rangle\}\Big)^* \Big(\langle u|\{\gamma|u\rangle + \delta|d\rangle\}\Big) \\
&= \Big(\gamma \underbrace{\langle u|u\rangle}_{=1} + \delta \underbrace{\langle u|d\rangle}_{=0}\Big)^* \Big(\gamma \underbrace{\langle u|u\rangle}_{=1} + \delta \underbrace{\langle u|d\rangle}_{=0}\Big) \\
&= \gamma^*\gamma \quad \square
\end{aligned}$$

$$\begin{aligned}
\frac{1}{2} &= \langle i|d\rangle \langle d|i\rangle \\
&= (\langle d|i\rangle)^* \langle d|i\rangle \\
&= \Big(\langle d|\{\gamma|u\rangle + \delta|d\rangle\}\Big)^* \Big(\langle d|\{\gamma|u\rangle + \delta|d\rangle\}\Big) \\
&= \Big(\gamma \underbrace{\langle d|u\rangle}_{=0} + \delta \underbrace{\langle d|d\rangle}_{=1}\Big)^* \Big(\gamma \underbrace{\langle d|u\rangle}_{=0} + \delta \underbrace{\langle d|d\rangle}_{=1}\Big) \\
&= \delta^*\delta \quad \square
\end{aligned}$$

---

b) I don't think we can conclude here without recalling the definition of $|r\rangle$:

$$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle$$

Let's start with a piece from Eqs. 2.9, arbitrarily (we could use $\langle i|l\rangle \langle l|i\rangle = \frac{1}{2}$, but I think we'd still need the previous definition of $|r\rangle$):

$$\langle i|r\rangle \langle r|i\rangle = \frac{1}{2}$$

But:

$$\langle r|i\rangle = \langle r|\{\alpha + |u\rangle + \beta|d\rangle\} = \alpha \langle r|u\rangle + \beta \langle r|d\rangle$$

And:

$$\langle i|r\rangle = (\langle r|i\rangle)^* = (\alpha \langle r|u\rangle + \beta \langle r|d\rangle)^* = \alpha^* \langle u|r\rangle + \beta^* \langle d|r\rangle$$

So

$$\langle i|r\rangle \langle r|i\rangle = \frac{1}{2}$$

$$\Leftrightarrow \Big(\alpha^* \langle u|r\rangle + \beta^* \langle d|r\rangle\Big)\Big(\alpha \langle r|u\rangle + \beta \langle r|d\rangle\Big) = \frac{1}{2}$$

$$\Leftrightarrow \underbrace{\alpha^*\alpha}_{=1/2} \langle u|r\rangle \langle r|u\rangle + \alpha^*\beta \langle u|r\rangle \langle r|d\rangle + \beta^*\alpha \langle d|r\rangle \langle r|u\rangle + \underbrace{\beta^*\beta}_{=1/2} \langle d|r\rangle \langle r|d\rangle = \frac{1}{2}$$

$$\Leftrightarrow \frac{1}{2}\Big(\langle u|r\rangle \langle r|u\rangle + \langle d|r\rangle \langle r|d\rangle\Big) + \alpha^*\beta \langle u|r\rangle \langle r|d\rangle + \beta^*\alpha \langle d|r\rangle \langle r|u\rangle = \frac{1}{2}$$

Now if $|r\rangle = \rho_u|u\rangle + \rho_d|d\rangle$, then

$$\langle u|r\rangle \langle r|u\rangle + \langle d|r\rangle \langle r|d\rangle = \rho_u\rho_u^* + \rho_d\rho_d^* = 1$$

As $\rho_u\rho_u^*$ would be the probability of $|r\rangle$ to be up, and $\rho_d\rho_d^*$ would the probability of $|r\rangle$ to be down, which are two orthogonal states in a two-states setting, and so the sum of their probability must be 1.

Hence the previous expression becomes:

$$\alpha^*\beta \langle u|r\rangle \langle r|d\rangle + \beta^*\alpha \langle d|r\rangle \langle r|u\rangle = 0$$

Note that so far, we haven't needed the expression of $|r\rangle$, but I think we don't have a choice but to use it to conclude:

$$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle$$

So, as the coefficient are real numbers:

$$\langle u|r\rangle = \frac{1}{\sqrt{2}} = \langle r|u\rangle\,; \qquad \langle d|r\rangle = \frac{1}{\sqrt{2}} = \langle r|d\rangle$$

Replacing in the previous expression we have:

$$\alpha^*\beta \underbrace{\langle u|r\rangle}_{=1/\sqrt{2}} \underbrace{\langle r|d\rangle}_{=1/\sqrt{2}} + \beta^*\alpha \underbrace{\langle d|r\rangle}_{=1/\sqrt{2}} \underbrace{\langle r|u\rangle}_{=1/\sqrt{2}} = 0$$

$$\Leftrightarrow \frac{1}{2}\alpha^*\beta + \frac{1}{2}\beta^*\alpha = 0$$

$$\Leftrightarrow \boxed{\alpha^*\beta + \beta^*\alpha = 0} \quad \square$$

The process is very similar to prove $\gamma^*\delta + \gamma\delta^* = 0$; one has to start again from a Eqs. 2.9, but this time, from another piece involving o, arbitrarily:

$$\langle o|r\rangle \langle r|o\rangle = \frac{1}{2}$$

$$\Leftrightarrow \Big(\langle r|o\rangle\Big)^* \langle r|o\rangle = \frac{1}{2}$$

$$\Leftrightarrow \Big(\langle r|\{\gamma|u\rangle + \delta|d\rangle\}\Big)^* \Big(\langle r|\{\gamma|u\rangle + \delta|d\rangle\}\Big) = \frac{1}{2}$$

$$\Leftrightarrow \Big(\gamma^* \langle u|r\rangle + \delta^* \langle d|r\rangle\Big)\Big(\gamma \langle r|u\rangle + \delta \langle r|d\rangle\Big) = \frac{1}{2}$$

$$\Leftrightarrow \underbrace{\gamma^*\gamma}_{=1/2} \langle u|r\rangle \langle r|u\rangle + \gamma^*\delta \langle u|r\rangle \langle r|d\rangle + \delta^*\gamma \langle d|r\rangle \langle r|u\rangle + \underbrace{\delta^*\delta}_{=1/2} \langle d|r\rangle \langle r|d\rangle = \frac{1}{2}$$

$$\Leftrightarrow \frac{1}{2}\Big(\underbrace{\langle u|r\rangle \langle r|u\rangle + \langle d|r\rangle \langle r|d\rangle}_{=1}\Big) + \gamma^*\delta \langle u|r\rangle \langle r|d\rangle + \delta^*\gamma \langle d|r\rangle \langle r|u\rangle = \frac{1}{2}$$

$$\Leftrightarrow \gamma^*\delta \underbrace{\langle u|r\rangle \langle r|d\rangle}_{=1/2} + \delta^*\gamma \underbrace{\langle d|r\rangle \langle r|u\rangle}_{=1/2} = 0$$

$$\Leftrightarrow \boxed{\gamma^*\delta + \delta^*\gamma = 0} \quad \square$$

---

c) Let's assume $\alpha\beta^*$ is a complex number of the form:

$$\alpha\beta^* = a + ib, \qquad (a, b) \in \mathbb{R}^2$$

But then:

$$\Big(\alpha\beta^*\Big)^* = a - ib = \alpha^*\beta$$

That's because, for two complex numbers $z = a + ib$ and $w = x + iy$, we have:

$$\Big(zw\Big)^* = z^*w^*$$

Indeed:

$$zw = (a + ib)(x + iy) = (ax - by) + i(bx + ya)$$

Hence:

$$\Big(zw\Big)^* = (ax - by) - i(bx + ya)$$

But:

$$z^*w^* = (a - ib)(x - iy) = (ax - by) - i(bx + ya)$$

Hence the result. Back to our α and β, we established in b) that:

$$\alpha^*\beta + \alpha\beta^* = 0$$

Which is equivalent from our previous little proof to:

$$\alpha^*\beta + \Big(\alpha^*\beta\Big)^* = 0$$

$$\Leftrightarrow (a + ib) + (a - ib) = 0 \Leftrightarrow 2a = 0 \Leftrightarrow \boxed{a = 0}$$

Which is the same as saying that the real part of $\alpha^*\beta$ is zero, or that it's a pure imaginary number. The exact same argument applies for $\gamma^*\delta$.

3 Principles of Quantum Mechanics

3.1 Mathematical Interlude: Linear Operators

3.1.1 Hermitian Operators and Orthonormal Bases

Exercise 6. *Prove the following: If a vector space in N-dimensional, an orthonormal basis of N vectors can be constructed from the eigenvectors of a Hermitian operator.*

We're here asked to prove a portion of an important theorem. I'm going to be somehow thorough in doing so, but to save space, I'll assume familiarity with linear algebra, up to diagonalization. Let's start with some background.

This exercise is about proving one part of what the authors call the *Fundamental theorem*, also often called in the literature the *(real) Spectral theorem*. So far, we've been working more or less explicitly in finite-dimensional spaces, but this result in particular has a notorious analogue in infinite-dimensional Hilbert spaces, called the *Spectral theorem*[1].

Now, I'm *not* going to prove the the infinite dimension version here. There's a good reason why quantum mechanics courses often start with spins: they don't require the generalized results, which demands heavy mathematical machinery (a copious amount of functional analysis, and in some formulation at least, the Lebesgue integral, hence portions of measure theory). You may want to refer to F. Schuller YouTube lectures on quantum mechanics[2] for a thorough development.

Finally, I'm going to use a mathematically inclined approach here (definitions/theorems/proofs), and as we won't need it, I won't be using the bra-ket notation.

To fix things, here's the theorem we're going to prove (I'll slightly restate it with minor adjustments later on):

Theorem 1. *Let $H : V \to V$ be a Hermitian operator on a finite-dimensional vector space V, equipped with an inner-product*[3]*.*

$$\boxed{\textit{Then, the eigenvectors of } H \textit{ form an orthonormal basis}}$$

Saying it otherwise, it means that a matrix representation M_H of H is diagonalizable, and that two eigenvectors associated with distinct eigenvalues are orthogonal.

For clarity, let's recall a few definitions.

Definition 1. *Let $L : V \to V$ be a linear operator on a vector space V over a field $\mathbb{F}$. We say that a non-zero $\boldsymbol{p} \in U$ is an eigenvector for L, with associated eigenvalue $\lambda \in \mathbb{F}$ whenever:*

$$L(\boldsymbol{p}) = \lambda \boldsymbol{p}$$

Remark 3. *As this can be a source of confusion later on, note that the definition of eigenvector/eigenvalue does not depend on the diagonalizability of L.*

Remark 4. *Note also that while eigenvectors must be non-zero, no such restrictions are imposed on the eigenvalues.*

Definition 2. *Two vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ from a vector space V over a field $\mathbb{F}$ equipped with an inner product $\langle ., . \rangle$ are said to be orthogonal (with respect to the inner-product) whenever:*

$$\boxed{\langle \boldsymbol{p}, \boldsymbol{q} \rangle = 0_{\mathbb{F}}}$$

[1] See `https://ncatlab.org/nlab/show/spectral+theorem` and `https://en.wikipedia.org/wiki/Spectral_theorem`

[2] `https://www.youtube.com/watch?v=GbqA9Xn_iMO&list=PLPH7f_7ZlzxQVx5jRjbfRGEzWY_upS5K6`; see also the lectures notes (.pdf) made by a student (Simon Rea): `https://drive.google.com/file/d/1nchF1fRGSY3R3rP1QmjUg7fe28tAS428/view`

[3] Remember, we need it to be able to talk about orthogonality.

The following lemma will be of great use later on. Don't let yourself be discouraged by the length of the proof: it can literally be be shorten to just a few lines, but I'm going to be very precise, hence very explicit, as to make the otherwise simple underlying mathematical constructions as clear as I can.

Lemma 1. *A linear operator $L : V \to V$ on a $n \in \mathbb{N}$ dimensional vector space V over the complex numbers has at least one eigenvalue.*

Proof. Let's take a $\boldsymbol{v} \in V$. We assume V is not trivial, that is, V isn't reduced to its zero vector $\mathbf{0}_V$, and so we can always choose $\boldsymbol{v} \neq \mathbf{0}_V$[4].

Consider the following set of $n+1$ vectors:

$$\{\boldsymbol{v}, L(\boldsymbol{v}), L^2(\boldsymbol{v}), \dots, L^n(\boldsymbol{v})\}$$

where:

$$L^0 := \mathrm{id}_V; \qquad L^i := \underbrace{L \circ L \circ \dots \circ L}_{i \in \mathbb{N} \text{ times}}$$

It's a set of $n+1$ vectors, but the space is n dimensional, so its vectors are *not* all linearly independent. This means there's a set of $(\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{C}^n$ which are not all zero, such that:

$$\sum_{i=0}^{n} \alpha_i L^i(\boldsymbol{v}) = \mathbf{0}_V \tag{4}$$

Here's the "subtle" part. You remember what a polynomial is right, something like:

$$x^2 - 2x + 1$$

You know it's customary to then consider this a function of a single variable x, which for instance, can range through the reals:

$$L : \begin{pmatrix} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2 - 2x + 1 \end{pmatrix}$$

This allows you to graph the polynomial and so forth:

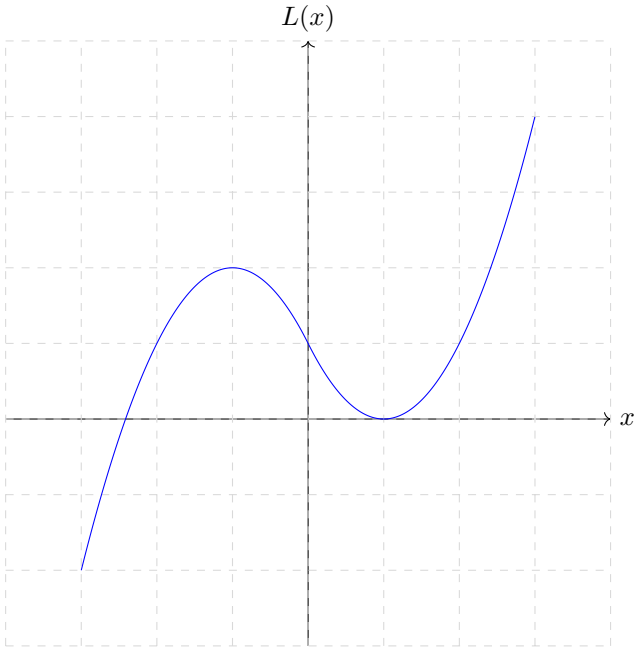

Figure 1: $L(x) = x^2 - 2x + 1$

[4]Note that if V is trivial, because an eigenvalue is always associated to a non-zero vector, there are no eigenvalues/eigenvectors, and the result is trivial.

But that's kindergarten polynomials. The more "correct" polynomials are *not* functions of a real variable. Rather, we say that $L(x)$ or L is a polynomial of a single variable/indeterminate[5] x, where x stands for an abstract symbol.

The reason is that, when you say that x is a real number (or a complex number, or whatever), you tacitly assume that you can for instance add, subtract or multiply various occurrences of x, but when mathematicians study polynomials, they want to do so without requiring additional (mathematical) structure on x.

Hence, x is just a placeholder, an abstract symbol.

The set of polynomials of a single variable X with coefficient in a field $\mathbb{F}$ is denoted $\mathbb{F}[X]$. For instance, $\mathbb{C}[f]$ is the set of all polynomials with complex coefficient of a single variable f, say, $P(f) = (3+2i)f^3 + 5f \in \mathbb{C}[f]$.

Now you'd tell me, wait a minute: if I have a $P(X) = X^2 - 2X + 1$, am I not then adding a polynomial $X^2 - 2X$ with an element from the field, 1?

Well, you'd be somehow right: the notation *is* ambiguous, in part inherited from the habits of kindergarten polynomials, in part because the context often makes things clear, and perhaps most importantly, because a truly unambiguous notation is unpractically verbose. Actually, $X^2 - 2X + 1$ is a shortcut notation for $X^2 - 2X^1 + 1X^0$. So no: all the $+$ here are between polynomials.

What does this mean that the $+$ are between polynomials? Well, most often when you encounter $\mathbb{F}[X]$, it's actually a shortcut for $(\mathbb{F}[X], +_{\mathbb{F}[X]}, \cdot_{\mathbb{F}[X]})$, which is a *ring*[6] *of polynomials of a single indeterminate over a field*[7] $\mathbb{F}$. This means that mathematicians have defined a way This means that $X^2 - 2X + 1$ is actually a shortcut for:

$$1_{\cdot_{\mathbb{F}[X]}} X^2 +_{\mathbb{F}[X]} (-2)_{\cdot_{\mathbb{F}[X]}} X^1 +_{\mathbb{F}[X]} (1)_{\cdot_{\mathbb{F}[X]}} X^0$$

Awful, right? Hence why we often use ambiguous notations and reasonable syntactical shortcuts.

The main takeaway though is that mathematicians have defined a set of precise rules (addition, scalar multiplication, exponentiation of an indeterminate), and that by cleverly combining such rules and only such rules, they have obtain a bunch of interesting results, and we want to use one of them in particular.

Let's get back to our equation (4); let me add some parenthesis for clarity:

$$\sum_{i=0}^{n} \left(\alpha_i L^i(\boldsymbol{v})\right) = \mathbf{0}_V$$

Our goal is to transform this expression so that it involves a polynomial in $\mathbb{C}[L]$[8].

Let's start by pulling out the $\boldsymbol{v}$ on the left-hand side as such:

$$\left(\underbrace{\sum_{i=0}^{n} \alpha_i L^i}_{=:P(L)} \right) (\boldsymbol{v}) = \mathbf{0}_V$$

What's P ? It's a function which takes a linear operator on V and returns ... A polynomial? But then, we don't know how to evaluate a polynomial on a vector $\boldsymbol{v} \in V$ so there's an problem somewhere.

[5] `https://en.wikipedia.org/wiki/Indeterminate_(variable)`

[6] `https://en.wikipedia.org/wiki/Ring_(mathematics)`. Note that there is no notion of subtraction in a ring: the minus signs actually are part of the coefficients.

[7] `ttps://en.wikipedia.org/wiki/Field_(mathematics)`

[8] Remember, this means a polynomial of a single variable L, with coefficient in $\mathbb{C}$.

P actually returns a new *linear operator on* V:

$$P : \begin{pmatrix} (V \to V) & \to & (V \to V) \\ L & \mapsto & \sum_{i=0}^{n} \alpha_i L^i \end{pmatrix}$$

But this means that while in (4) the $\sum$ was a sum of complex numbers, it's now a sum of functions, and that $\alpha_i L_i$ went from a multiplication between complex numbers to a scalar multiplication on a function.

The natural way, that is, the simplest consistent way, to do so, is to define them pointwise[9] for two functions $f, g : X \to Y$, we define $(f + g) : X \to Y$ by:

$$(\forall x \in X),\ (f + g)(x) := f(x) + g(x)$$

The process is similar for scalar multiplication:

$$(\forall x \in X), (\forall y \in Y),\ (yf)(e) := yf(e)$$

We *equip* the space of (linear) functions (on V) with additional laws. All in all, P is well defined[10], and that we can indeed pull the $\boldsymbol{v}$ out.

How then can we go from such a weird "meta" function P to a polynomial? Well, as we stated earlier, polynomials are defined by a set of specific rules: addition, scalar multiplication, and exponentiation of the indeterminate.

But if you look closely:

- Our point-wise addition has the same property as the additions on polynomial (symmetric, existence of inverse elements, neutral element, etc.)
- Similarly for our scalar multiplication;
- And our rules of exponentiation on function by repeated application also follows the rules of exponentiation for an indeterminate variable.

This mean that if we squint a little, if we only look at the expression $P(L)$ as having nothing but those properties, then it behaves exactly as a polynomial. Hence, for all intents and purposes, it "is" a polynomial, and we can manipulate it as such.
So we can apply the fundamental theorem of algebra[11], we know that we can always factorize polynomials with complex coefficient as such:

$$(\exists (c, \lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n+1}, c \neq 0),\ P(L) = c \prod_{i=0}^{n} (L - \lambda_i)$$

But don't we have a problem here? L is an abstract symbol, and we're "subtracting" it a scalar? Well, there are a few implicit elements:

$$P(L) = c \prod_{i=0}^{n} (L^1 + (-\lambda_i) L^0)$$

Let's replace this new expression for $P(L)$ in our previous equation, which we can do essentially re-using our previous argument: the rules (addition, scalar multiplication, etc.) to manipulate polynomials are "locally" consistent with the rules to manipulate our (linear) functions:

$$\left(c \prod_{i=0}^{n} (L^1 - \lambda_i L^0) \right) (\boldsymbol{v}) = \mathbf{0}_V$$

[9] `https://en.wikipedia.org/wiki/Pointwise`

[10] Meaning, the laws we introduce on functions are consistent with the results we would otherwise get without using them; you can check this out if you want

[11] `https://en.wikipedia.org/wiki/Fundamental_theorem_of_algebra`

Note that L^0 becomes the identity function, and by using the previous point-wise operations, we can reduce it to:

$$c\prod_{i=0}^{n}(L(\boldsymbol{v})-\lambda_i \mathrm{id}_V(\boldsymbol{v}))=c\prod_{i=0}^{n}(L(\boldsymbol{v})-\lambda_i\boldsymbol{v})=\mathbf{0}_V$$

Now, $c\neq 0$ by the fundamental theorem of algebra. So we must have:

$$\prod_{i=0}^{n}(L(\boldsymbol{v})-\lambda_i\boldsymbol{v})=\mathbf{0}_V$$

Which implies that there's at least a λ_j for which

$$L(\boldsymbol{v})-\lambda_j\boldsymbol{v}=\mathbf{0}_V \Leftrightarrow L(\boldsymbol{v})=\lambda_j\boldsymbol{v}$$

But we've selected $\boldsymbol{v}$ to be non-zero: λ_j is then an eigenvalue λ_j associated to the eigenvector $\boldsymbol{v}$.

□

OK; let me adjust the fundamental theorem a little bit, and let's prove it.

Theorem 2. *Let $H : V \to V$ be a Hermitian operator on a finite, n-dimensional vector space V, equipped with an inner-product $\langle .,. \rangle$.*

Then, the eigenvectors of H form an orthogonal basis of V, and the associated eigenvalues are real.

Saying it otherwise, it means that a matrix representation M_H of H is diagonalizable, and that two eigenvectors associated with distinct eigenvalues are orthogonal.

Proof. I'm assuming that this is clear for you that the eigenvectors associated to the eigenvalues of a diagonalizable matrix makes a basis for the vector space. Again, refer to a linear algebra course for more.

Furthermore, you can refer to the book for a proof of orthogonality of the eigenvectors associated to distinct eigenvalues[12].

Note that I've included a mention to characterize the eigenvalues as real numbers: there's already a proof in the book, but it comes with almost no effort with the present proof, so I've included it anyway.

Remains then to prove that the matrix representation M_H of H is diagonalizable (and that the eigenvalues are real). Let's prove this by induction on the dimension of the vector space. If you're not familiar with proofs by induction, the idea is as follow:

- Prove that the result is true, say, for $n = 1$;
- Then, prove that if the result is true for $n = k$, then the result must be true for $n = k + 1$.
- If the two previous points hold, then you can combine them: if the first point hold then by applying the second point, the result must be true $n = 1 + 1 = 2$. But then by applying the second point again, it must be true that the result holds for $n = 2 + 1 = 3$.
- And so on: the result is true $\forall n \in \mathbb{N}\backslash\{0\}$.

$\boxed{n=1}$ Then, H is reduced to a 1×1 matrix, containing a single element h. This is trivially diagonal already, and because H is assumed to be Hermitian, the only eigenvalue $h = h^*$ is real.

$\boxed{\text{Induction}}$ Assume the result holds for any Hermitian operator $H : W \to W$ on a k-dimensional vector space W over $\mathbb{C}$.

Let V be a $k + 1$-dimensional vector space over $\mathbb{C}$. By our previous lemma, $H : V \to V$ must have at least one eigenvalue $\lambda \in \mathbb{C}$ associated to an eigenvector $\boldsymbol{v} \in V$.

[12] I'm not doing it here, as I've avoided the bra-ket notation, and this would force me to talk about dual spaces, and so on.

Pick $\{\boldsymbol{v}_1, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k+1}\} \subset V$ so that $\{\boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k+1}\}$ is an (ordered) basis of V[13].

Apply the Gram-Schmidt procedure[14] to extract from it an (ordered) orthonormal basis $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{k+1}\}$ of V; note that by construction:

$$\boldsymbol{b}_1 = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$$

That's to say, $\boldsymbol{b}_1$ is still an eigenvector for λ[15].

Now we're trying to understand what's the matrix representation D_H of H, in this orthonormal basis. If you've taken the blue pill, you know how to "read" a matrix:

$$D_H = \begin{pmatrix} \begin{pmatrix} | \\ H(\boldsymbol{b}_1) \\ | \end{pmatrix} & \begin{pmatrix} | \\ H(\boldsymbol{b}_2) \\ | \end{pmatrix} & \ldots & \begin{pmatrix} | \\ H(\boldsymbol{b}_{k+1}) \\ | \end{pmatrix} \end{pmatrix}$$

OK; let's start by what we know: $\boldsymbol{b}_1$ is an eigenvector for H associated to λ, meaning:

$$H(\boldsymbol{b_1}) = \lambda \boldsymbol{b}_1 = \lambda \boldsymbol{b}_1 + \sum_{i=2}^{k+1} 0 \times \boldsymbol{e}_i = \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Rewrite D_H accordingly, and break it into blocks:

$$D_H = \begin{pmatrix} \lambda & \begin{pmatrix} | \\ H(\boldsymbol{b}_2) \\ | \end{pmatrix} & \ldots & \begin{pmatrix} | \\ H(\boldsymbol{b}_{k+1}) \\ | \end{pmatrix} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{pmatrix} = \left(\begin{array}{c|ccc} \lambda & & A & \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right)$$

Where A is a $1 \times k$ matrix (a row vector), and C a $k \times k$ matrix. But then H is Hermitian, which means its matrix representation obeys:

$$D_H = (D_H^T)^* = D_H^\dagger$$

This implies first that $\lambda = \lambda^*$, i.e λ is real, and we'll see shortly, can be considered an eigenvalue, as we can transform D_H in a diagonal matrix with λ on the diagonal.

Second, $A^\dagger = (0\ 0\ \ldots\ 0) = A$, i.e:

$$\left(\begin{array}{c|ccc} \lambda & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right)$$

Third, $C = C^\dagger$. But then, C is a $k \times k$ Hermitian matrix, corresponding to a Hermitian operator in a k-dimensional vector space. Using the induction assumption, it is diagonalizable, with real valued eigenvalues. Hence D_H is diagonalizable, and all its eigenvalues are real.

□

[13] Start with $W = \{\boldsymbol{v}\}$, and progressively augment it with elements of V so that all elements in W are linearly independent. If we can't select such elements no more, this mean we've got a basis. Ordering naturally follows from the iteration steps.

[14] `https://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process`

[15] $H(\boldsymbol{b}_1) = H(\boldsymbol{v}/\|\boldsymbol{v}\|)$, by linearity of H, this is equal to $\frac{1}{\|\boldsymbol{v}\|} H(\boldsymbol{v})$. But $\boldsymbol{v}$ is an eigenvector for an eigenvalue λ, so this is equal to $\frac{\lambda}{\|v\|}\boldsymbol{v} = \lambda \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = \lambda \boldsymbol{b_1}$

3.1.2 The Gram-Schmidt Procedure

3.2 The Principles

3.3 An Example: Spin Operators

3.4 Constructing Spin Operators

Exercise 7. *Prove that Eq.* 3.16 *is the unique solution to Eqs.* 3.14 *and* 3.15.

Let's recall all the equations, 3.14, 3.15 and 3.16

$$\begin{pmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{5}$$

$$\begin{pmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6}$$

$$\begin{pmatrix} (\sigma_z)_{11} & (\sigma_z)_{12} \\ (\sigma_z)_{21} & (\sigma_z)_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7}$$

By developing the matrix product and identifying the vectors components, the first two equations make a system of four equations involving four unknowns $(\sigma_z)_{11}$, $(\sigma_z)_{12}$, $(\sigma_z)_{21}$ and $(\sigma_z)_{22}$:

$$\begin{cases} 1(\sigma_z)_{11} + 0(\sigma_z)_{12} &= 1 \\ 1(\sigma_z)_{21} + 0(\sigma_z)_{22} &= 0 \\ 0(\sigma_z)_{11} + 1(\sigma_z)_{12} &= 0 \\ 0(\sigma_z)_{21} + 1(\sigma_z)_{22} &= -1 \end{cases} \Leftrightarrow \begin{cases} (\sigma_z)_{11} &= 1 \\ (\sigma_z)_{21} &= 0 \\ (\sigma_z)_{12} &= 0 \\ (\sigma_z)_{22} &= -1 \end{cases} \Leftrightarrow \boxed{\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \quad \square \tag{8}$$

Remark 5. *Observe that we are (were) trying to build a Hermitian operator with eigenvalues* $+1$ *and* -1*. The fundamental theorem / real spectral theorem, assures us that Hermitian operators are diagonalizable, hence there exists a basis in which the operator can be represented by a* 2×2 *matrix containing the eigenvalues on its diagonal:*

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Which is exactly the matrix we've found.

But now of course, you'd be wondering: wait a minute, right after this exercise, we're trying to build σ_x*, which also has those same eigenvalues* $+1$ *and* -1*, what's the catch?*

Well, remember the diagonalization process: M *diagonalizable means that there's a basis where it's diagonal. That is, there's a change of basis, which is an invertible linear function, which has a matrix representation* P*, such that the linear operation represented by* M *in a starting basis is now represented by a diagonal matrix* D*:*

$$M = PDP^{-1}$$

Furthermore:

- *The elements on the diagonal of* D *are the eigenvalues;*
- *The columns of* P *are the corresponding eigenvectors*

So regarding σ_x*, we still have a*

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

But the catch is that before for σ_z*,* P *was the identity matrix* I_2 *(because of our choice for* $|u\rangle$ *and* $|d\rangle$*). But now, given our values for* $|r\rangle$ *and* $|l\rangle$*, we have:*

$$|r\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{and} \quad |l\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \quad \Rightarrow \quad P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Note that the column order matters: the first column of P must be $|r\rangle$, and the first column of D must contain the eigenvalue associated to $|r\rangle$. But:

$$\sigma_x = PDP^{-1} \Leftrightarrow \sigma_x P = PD(\underbrace{P^{-1}P}_{:=I_2}) = PD = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$$

Hence,

$$\sigma_x P = \begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \Leftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix}(\sigma_x)_{11} & (\sigma_x)_{12}\\ (\sigma_x)_{21} & (\sigma_x)_{22}\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$$

Solving for the components of σ_x:

$$\Leftrightarrow \begin{cases}(\sigma_x)_{11} + (\sigma_x)_{12} = 1\\ (\sigma_x)_{11} - (\sigma_x)_{12} = -1\\ (\sigma_x)_{21} + (\sigma_x)_{22} = 1\\ (\sigma_x)_{21} - (\sigma_x)_{22} = 1\end{cases}$$

Which indeed yields the expected Pauli matrix, as described in the book, and computed by the authors using a different approach:

$$\boxed{\sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}}$$

And obviously, the same can be done for σ_y : that's to say that, reassuringly, we reach the same results using pure linear algebra.

3.5 A Common Misconception

3.6 3-Vector Operators Revisited

3.7 Reaping the Results

Exercise 8. *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 6. *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:[16]

$$\sigma_n = PDP^{-1} \quad \Leftrightarrow \quad \sigma_n P = PD$$

$$\Leftrightarrow \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix} = \begin{pmatrix}\lambda_1 a & \lambda_2 b\\ \lambda_1 c & \lambda_2 d\end{pmatrix}$$

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:

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

[16] 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

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:

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

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[17]).

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;

[17] 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`

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 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[18]. 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[19] 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

[18] Source: `https://physics.stackexchange.com/a/720025`
[19] L03E01.pdf

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 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 7. *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[20]:

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

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

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:

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

Exercise 9. *Let $n_z = \cos\theta$, $n_x = \sin\theta\cos\phi$ and $n_y = \sin\theta\sin\phi$. Angles θ and ϕ are defined according to the usual conventions for spherical coordinates (Fig. 3.2). Compute the eigenvalues and eigenvectors for the matrix of Eq. 3.23.*

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

Let's recall Eq. 3.23, which is general form of the spin 3-vector operator:

$$\sigma_n = \begin{pmatrix} n_z & (n_x - in_y) \\ (n_x + in_y) & -n_z \end{pmatrix} = \begin{pmatrix} \cos\theta & (\sin\theta\cos\phi - i(\sin\theta\sin\phi)) \\ (\sin\theta\cos\phi + i(\sin\theta\sin\phi)) & -\cos\theta \end{pmatrix}$$

Observe (e.g. from the trigonometric circle) that:

$$\cos\theta = \cos(-\theta); \qquad \sin\theta = -\sin(-\theta)$$

Hence:

$$\exp(-i\theta) := \cos(-\theta) + i\sin(-\theta) = \cos\theta - i\sin\theta$$

And we can simplify our previous expression of σ_n to:

$$\sigma_n = \begin{pmatrix} \cos\theta & \exp(-i\phi)\sin\theta \\ \exp(i\phi)\sin\theta & -\cos\theta \end{pmatrix}$$

Note that as we're now in the general case, we indeed have two degrees of freedom, encoded by the two angles θ and ϕ; the *why* has been explicited in subsection 2.5.

We're still confronted to a spin operator: we expect the eigenvalues to be $+1$ and -1[21]. But let's check this first: an eigenvector $|\lambda\rangle$ associated to an eigenvalue λ must obey:

$$\sigma_n|\lambda\rangle = \lambda|\lambda\rangle$$

$$\Leftrightarrow \sigma_n|\lambda\rangle - \lambda|\lambda\rangle = 0 \Leftrightarrow (\sigma_n - I_2\lambda)|\lambda\rangle = 0$$

But eigenvectors are non-zero, hence, again with 0_2 being the 2×2 zero matrix:

$$\Leftrightarrow \sigma_n - I_2\lambda = 0_2$$

And so this matrix $\sigma_n - I_2\lambda$ cannot be invertible[22]. This translates to a condition on the determinant:

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

The remaining difficulty is then in finding the eigenvectors. We can use the following argument[23].

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:

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

[21] Remember from the real spectral theorem, or as the authors call it, the *fundamental theorem*, that because we have a Hermitian matrix, we know it's diagonalizable, that its eigenvalues are real, and that the corresponding eigenvectors form a orthogonal basis

[22] Again for otherwise, as recalled in L03E03, multiply both sides of the equation by its inverse, get an identity on the left-hand-side and still the zero matrix on the right-hand-side

[23] `https://physics.stackexchange.com/a/720025`

We can then ignore the general phase factor $\exp(i\phi_1)$, e.g. set $\phi_1 = 0$. Furthermore, we want the vector to be normalized (this is an eigenvector associated to the eigenvalue of a Hermitian operator: it must be normalized per the real spectral theorem), 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 $\alpha \in \mathbb{R}$. Finally, let's rename $\phi_2 = \phi_\alpha$[24], which brings us to consider eigenvectors of the form:

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

As for the previous exercise, we can use two different parameter α and β for each eigenvector. Again, because of the diagonalization process, we have the following relation

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

But the columns of P must contain our eigenvectors, so this is equivalent to:

$$\underbrace{\begin{pmatrix} \cos\theta & \exp(-i\phi)\sin\theta \\ \exp(i\phi)\sin\theta & -\cos\theta \end{pmatrix}}_{=\sigma_n}\underbrace{\begin{pmatrix} \cos\alpha & \cos\beta \\ \exp(i\phi_\alpha)\sin\alpha & \exp(i\phi_\beta)\sin\beta \end{pmatrix}}_{=P} = \begin{pmatrix} \cos\alpha & \cos\beta \\ \exp(i\phi_\alpha)\sin\alpha & \exp(i\phi_\beta)\sin\beta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\alpha & -\cos\beta \\ \exp(i\phi_\alpha)\cos\alpha & -\exp(i\phi_\beta)\sin\beta \end{pmatrix}$$

Let's perform the matrix multiplication on the left:

$$\begin{pmatrix} \cos\theta\cos\alpha + \exp(i(\phi_\alpha - \phi))\sin\theta\sin\alpha & \cos\theta\cos\beta + \exp(i(\phi_\beta - \phi))\sin\theta\sin\beta \\ \exp(i\phi)\sin\theta\cos\alpha - \exp(i\phi_\alpha)\cos\theta\sin\alpha & \exp(i\phi)\sin\theta\cos\beta - \exp(i\phi_\beta)\cos\theta\sin\beta \end{pmatrix}$$

$$= \begin{pmatrix} \cos\alpha & -\cos\beta \\ \exp(i\phi_\alpha)\cos\alpha & -\exp(i\phi_\beta)\sin\beta \end{pmatrix}$$

From which we can extract the following system of equations:

$$\begin{cases} \cos\theta\cos\alpha + \exp(i(\phi_\alpha - \phi))\sin\theta\sin\alpha &= \cos\alpha \\ \cos\theta\cos\beta + \exp(i(\phi_\beta - \phi))\sin\theta\sin\beta &= -\cos\beta \end{cases}$$

Remark 8. *As for the previous exercise, I leave it to you to check that the solution we'll find for this system also solve the two other omitted equations.*

It's tempting to set $\phi = \phi_\alpha = \phi_\beta$, but can we do so? Well, we know the two eigenvectors will have to be orthogonal: this adds an additional constraint, which decrease our degrees of freedom by one, meaning there's one superfluous variable in $\{\alpha, \beta, \phi_\alpha, \phi_\beta\}$. We can *choose* to implement this constraint by setting $\phi_\alpha = \phi_\beta$.

From there, we can indeed set $\phi_\alpha = \phi_\beta = \phi$, as this allows us to solve the equation for α and β more easily:

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

Which is exactly the same system we had for the previous exercise, which was solved by:

$$\begin{cases} \alpha &= \theta/2 \\ \beta &= \frac{1}{2}(\theta + \pi) \end{cases}$$

[24] Note that I'm not yet identifying ϕ_α with ϕ; this will come naturally later on

With the same trigonometric identities as for the previous exercise:

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

We reach the following eigenvectors

$$\boxed{\begin{cases} |+1\rangle &= \begin{pmatrix} \cos\alpha \\ \exp(i\phi)\sin\alpha \end{pmatrix} = \begin{pmatrix} \cos(\theta/2) \\ \exp(i\phi)\sin(\theta/2) \end{pmatrix} \\ |-1\rangle &= \begin{pmatrix} \cos\beta \\ \exp(i\phi)\sin\beta \end{pmatrix} = \begin{pmatrix} -\sin(\theta/2) \\ \exp(i\phi)\cos(\theta/2) \end{pmatrix} \end{cases}}$$

Alright, let's make the same verifications the authors did in the book after the previous exercise. First, we get the expected eigenvalues $+1$, -1, which are the only two eigenvalues we have for a spin operator.

Then the two eigenvectors must be orthogonal, indeed (I only do it one way; the other is trivially similar):

$$\langle +1|-1\rangle = \begin{pmatrix} \cos(\theta/2) & \exp(-i\phi)\sin(\theta/2) \end{pmatrix} \begin{pmatrix} -\sin(\theta/2) \\ \exp(i\phi)\cos(\theta/2) \end{pmatrix}$$

$$= -\cos(\theta/2)\sin(\theta/2) + \exp(-i\phi + i\phi)\cos(\theta/2)\sin(\theta/2) = 0$$

Finally, if we prepare a spin along the z-axis in the up state $|u\rangle$, then rotate our apparatus to lie along the $\hat{n}$ axis, which *is not* restricted to the xz-plane anymore, we have according to the fourth principle[25]:

$$P(+1) = |\langle u|+1\rangle|^2 = \cos^2(\theta/2)$$

$$P(-1) = |\langle u|-1\rangle|^2 = \sin^2(\theta/2)$$

Which then lead to the exact same computation regarding the expected value for the measurement:

$$\langle\sigma_n\rangle = \sum_i \lambda_i P(\lambda_i) = (+1)\cos^2(\theta/2) + (-1)\sin^2(\theta/2) = \boxed{\cos\theta}$$

Note also that $P(+1) + P(-1) = 1$.

Exercise 10. *Suppose that a spin is prepared so that $\sigma_m = +1$. The apparatus is then rotated to the $\hat{n}$ direction and σ_n is measured. What is the probability that the result is $+1$? Note that $\sigma_m = \sigma \cdot \hat{m}$, using the same convention we used for σ_n.*

There are essentially two ways of solving the issue.

The first one, and the simplest, is to observe that if we consider $\hat{n}$ in a frame of reference where $\hat{m}$ acts as our $z - axis$, then we're essentially in the case of our previous exercise: we've prepared a spin in the "up" state (now corresponding to a state where $\sigma_m = +1$), we've moved our apparatus away from $\hat{m}$ by a a certain angle θ[26], and we know from the previous exercise that the probability of measuring a $+1$ after aligning our apparatus with the $\hat{n}$ axis is now

$$\boxed{P(+1) = \cos^2\frac{\theta}{2}}$$

Which is exactly what we wanted to show (the answer is given in the book by the authors, after the exercise).

---

[25] Don't hesitate to get back to the definition of $|u\rangle$ and that of the inner-product if this isn't clear enough.

[26] θ really is the angle between $\hat{m}$ and $\hat{n}$, not some angle between $\hat{n}$ and the "real" z-axis

I'll only draft the second approach, as I expect it to be more time consuming[27]. The idea is not to rely on the previous observation, and to consider that we've prepared to spin so that $\sigma_m = +1$, which means the state of the system is the eigenvector corresponding to this eigenvalue, which we know from the previous exercise, with θ_m the angle between the z-axis and $\hat{m}$, and ϕ_m the angle between the x-axis and the projection of $\hat{m}$ on the xy-plane:

$$|+1_m\rangle = \begin{pmatrix} \cos(\theta_m/2) \\ \exp(i\phi_m)\sin(\theta_m/2) \end{pmatrix}$$

If we then align the apparatus in the $\hat{n}$ direction, with corresponding θ_n / ϕ_n angles, *which are relative to the z-axis, not* $\hat{m}$, we now, by the same result, that the eigenvector corresponding to the probability of measuring a $+1$ in the $\hat{n}$ direction is:

$$|+1_n\rangle = \begin{pmatrix} \cos(\theta_n/2) \\ \exp(i\phi_n)\sin(\theta_n/2) \end{pmatrix}$$

Then, the probability to measure a $+1$ is given, again by using the fourth principle:

$$P(+1) = |\langle +1_m|+1_n\rangle|^2$$

We would then need to develop the inner-product between the two state vectors, and find a way to identify it with the half-angle between $\hat{n}$ and $\hat{m}$.

All the difficulty is then in expressing this half-angle in terms of our four angles (θ_m, ϕ_m, θ_n, ϕ_n). I *suppose* we get some insightful elements by cleverly:

- Expressing $\hat{m}$ and $\hat{n}$ both in rectangular coordinates;
- Observing that by the regular 3-vector dot product, $\hat{n}\cdot\hat{m} = \|\hat{n}\|\|\hat{m}\|\cos\theta_{mn} = \cos\theta_{mn}$ (where θ_{mn} is the angle between $\hat{m}$ and $\hat{n}$
- Observing that $\cos\frac{\theta_{mn}}{2} = \frac{1}{\sqrt{2}}\hat{n}\cdot(\hat{n}+\hat{m})$ (again from the regular 3-vector dot product, as $\hat{n}+\hat{m}$ will be a (non-unitary) vector bisecting θ_{mn}[28])

3.8 The Spin-Polarization Principle

4 Time and Change

4.1 A Classical Reminder

4.2 Unitarity

4.3 Determinism in Quantum Mechanics

4.4 A Closer Look at $U(t)$

Exercise 11. *Prove that if U is unitary, and if $|A\rangle$ and $|B\rangle$ are any two state-vectors, then the inner product of $U|A\rangle$ and $U|B\rangle$ is the same as the inner product of $|A\rangle$ and $|B\rangle$. One could call this the conservation of overlaps. It expresses the fact that the logical relation between states is preserved with time.*

The inner-product has been defined as the product of a bra and a ket. So the inner-product of $U|A\rangle$ and $U|B\rangle$ is the product of e.g. the bra associated to $U|A\rangle$ and $U|B\rangle$. But in section 3.1.5 of the book, we've established that:

$$|C\rangle = M|D\rangle \Leftrightarrow \langle C| = \langle D|M^\dagger$$

Hence the inner-product we're looking for is:

$$\langle A|\underbrace{U^\dagger U}_{I}|B\rangle = \langle A|B\rangle \quad \square$$

Remark 9. *The terminology is a bit confusing: we're talking about the inner-product of two kets, while we've defined the inner-product to be an operation between a bra and a ket. Overall, the bra-ket notation makes things a little more complicated than just having to deal with an inner-product space.*

[27] And hopefully, valid...

[28] `https://math.stackexchange.com/a/2285989`: the parallelogram involved in the sum of two vectors in a rhombus.

4.5 The Hamiltonian

4.6 What Ever Happened to $\hbar$?

4.7 Expectation Values

4.8 Ignoring the Phase-Factor

4.9 Connections to Classical Mechanics

Exercise 12. *Prove that if M and L are both Hermitian, $i[M,L]$ is also Hermitian. Note that the i is important. The commutator is, by itself, not Hermitian.*

$$\begin{aligned}
(i[M,L])^\dagger &= (i(ML-LM))^\dagger \\
&= (iML - iLM)^\dagger \\
&= \left((iML-iLM)^T\right)^* && (\dagger\text{'s definition}) \\
&= \left((iML)^T - (iLM)^T\right)^* && \left((A+B)^T = A^T + B^T\right) \\
&= \left(iL^TM^T - iM^TL^T\right)^* && \left((AB)^T = B^TA^T\right) \\
&= \left(-i(L^TM^T)^* + i(M^TL^T)^*\right) && ((zt)^* = z^*t^*; (z+t)* = z^* + t^*)) \\
&= i\left(M^\dagger L^\dagger - L^\dagger M^\dagger\right) && (\dagger\text{'s definition}) \\
&= i\left(ML - LM\right) && \left(L = L^\dagger; M = M^\dagger\right) \\
&= i[M,L] && \square
\end{aligned}$$

Exercise 13. *Go back to the definition of Poisson brackets in Volume I and check that the identification in Eq.* 4.21 *is dimensionally consistent. Show that without the factor $\hbar$, it would not be.*

Let's recall first Eq. 4.21, where $[.,.]$ is the commutator and $\{.,.\}$ the Poisson brackets:

$$[F,G] \Longleftrightarrow i\hbar\{F,G\}$$

The Poisson brackets are defined in *Volume I*, Eq. (9) at the end of Lecture 9 (The Phase Space Fluid and the Gibbs-Liouville Theorem), as:

$$\{F,G\} := \sum_i \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right)$$

Where the p_i are the generalized momentum, and q_i are the generalized coordinates. Recall that a momentum is typically defined as a mass in motion, while the coordinates are simply distances to an origin:

$$[p_i] = \text{kg.m.s}^{-1}; \qquad [q_i] = \text{m}$$

For clarity, let's rewrite one of those partial derivative in terms of a limit:

$$\frac{\partial F}{\partial q_i} = \lim_{\epsilon\to 0}\frac{F(q_i+\epsilon) - F(q_i)}{\epsilon}$$

First ϵ must be of the same dimension than q_i is this case, for otherwise $q_i+\epsilon$ is ill-defined; more generally it'll have the same dimension that the dimension of the differentiation variable.

Second, observe that, again because otherwise we'd be adding carrots and potatoes:

$$\left[\sum_i \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right)\right] = \left[\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right], \quad \text{for any arbitrary } i \text{ that is}$$

But then,

$$[i\hbar\{F,G\}] = \left[\hbar\left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right)\right] = [\hbar]\left[\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i}\right] - [\hbar]\left[\frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right]$$

We know $[\hbar] = \text{kg.m}^2.\text{s}^{-1} = [q_i p_i]$, and if we make the limits explicit as we did before, it remains from the previous expression:

$$[i\hbar\{F, G\}] = [FG]$$

On the other side:

$$[[F, G]] = [FG - GF]$$

For $FG - GF$ to be well defined, it must be that $[FG] = [GF]$. And so we're done:

$$\boxed{[[F, G]] = [FG] = [i\hbar\{F, G\}]} \quad \square$$

4.10 Conservation of Energy

4.11 Spin in a Magnetic Field

Exercise 14. *Verify the commutation relations of Eqs.* 4.26*.*

Let's first recall this set of equations:

$$[\sigma_x, \sigma_y] = 2i\sigma_z;$$
$$[\sigma_y, \sigma_z] = 2i\sigma_x;$$
$$[\sigma_z, \sigma_x] = 2i\sigma_y;$$

For clarity, let's recall the commutator's definition, for two observable F and G:

$$[F, G] = FG - GF$$

And finally, let's recall the Pauli matrices σ_x, σ_y and σ_z (from Eqs. 3.20, at the end of section 3.4)

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

Then this is just elementary matrix multiplication.

$$[\sigma_x, \sigma_y] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix} = 2i\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\sigma_z} \quad \square$$

$$[\sigma_y, \sigma_z] = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2i \\ 2i & 0 \end{pmatrix} = 2i\underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\sigma_x} \quad \square$$

$$[\sigma_z, \sigma_x] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} = 2i\underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\sigma_y} \quad \square$$

4.12 Solving the Schrödinger Equation

Exercise 15. *Take any unit* 3*-vector* $\boldsymbol{n}$ *and form the operator*

$$H = \frac{\hbar\omega}{2}\sigma \cdot \boldsymbol{n}$$

Find the energy eigenvalues and eigenvectors by solving the time-independent Schrödinger equation. Recall that Eq. 3.23 *gives* $\sigma \cdot \boldsymbol{n}$ *in component form.*

Let's recall Eq. 3.23, which is general form of the spin 3-vector operator:

$$\sigma_n = \sigma \cdot \boldsymbol{n} = \begin{pmatrix} n_z & (n_x - in_y) \\ (n_x + in_y) & -n_z \end{pmatrix}$$

And the *time-independent Schrödinger equation*[29]:

$$H|E_j\rangle = E_j|E_j\rangle$$

[29] That's quite a fancy name for describing the eigenvectors of an operator, by comparison with the "iconic" Schrödinger equation...

In an earlier exercise (L03E04), we actually diagonalized σ_n: this gave us two eigenvalues $+1$ and -1, and two eigenvectors:

$$|+1\rangle = \begin{pmatrix} \cos(\theta/2) \\ \exp(i\phi)\sin(\theta/2) \end{pmatrix}; \qquad |-1\rangle = \begin{pmatrix} -\sin(\theta/2) \\ \exp(i\phi)\cos(\theta/2) \end{pmatrix}$$

Where $\boldsymbol{n}$ was a regular unitary 3-vector expressed in spherical coordinates:

$$\boldsymbol{n} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}$$

Let's see how we can leverage this previous work to our advantage: such an $\boldsymbol{n}$ vector still fit our purpose here. Furthermore, we know that the eigenvalues of σ_n are the only solutions to:

$$\sigma_n|F_j\rangle = F_j|F_j\rangle$$

But if we multiply both sides of this equation by $\dfrac{\hbar\omega}{2}$, we get exactly the equation we want to solve:

$$\underbrace{\frac{\hbar\omega}{2}\sigma_n}_{H}|F_j\rangle = \left(\frac{\hbar\omega}{2}F_j\right)|F_j\rangle$$

Multiplying the equation by a constant doesn't change the eigenvectors: they still are the only solutions, but the associated eigenvalues are now different:

$$\boxed{\lambda_1 = \frac{\hbar\omega}{2}; \qquad |\lambda_1\rangle = \begin{pmatrix} \cos(\theta/2) \\ \exp(i\phi)\sin(\theta/2) \end{pmatrix}}$$

$$\boxed{\lambda_2 = -\frac{\hbar\omega}{2}; \qquad |\lambda_2\rangle = \begin{pmatrix} -\sin(\theta/2) \\ \exp(i\phi)\cos(\theta/2) \end{pmatrix}}$$

4.13 Recipe for a Schrödinger Ket

Exercise 16. *Carry out the Schrödinger Ket recipe for a single spin. The Hamiltonian is $H = \dfrac{\omega\hbar}{2}\sigma_z$ and the final observable is σ_x. The initial state is given as $|u\rangle$ (the state in which $\sigma_z = +1$).*

After time t, an experiment is done to measure σ_y. What are the possible outcomes and what are the probabilities for those outcomes?

Congratulations! You have now solved a real quantum mechanics problem for an experiment that can actually be carried out in the laboratory. Feel free to pat yourself on the back.

Remark 10. *There's a typo in the statement of this exercise: the final observable is said first to be σ_x and then σ_y. The French version of the book uses σ_y for both, so that's what I'll do here.*

1. *Derive, look up, guess, borrow, or steal the Hamiltonian operator H;*
   Well, let's take it from the authors:
   $$H = \frac{\omega\hbar}{2}\sigma_z = \frac{\omega\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
2. *Prepare an initial state $|\Psi(0)\rangle$;*
   Again, from the exercise statement, let's prepare an up state:
   $$|\Psi(0)\rangle = |u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

3. *Find the eigenvalues and eigenvectors of H by solving the time-independent Schrödinger equation,*

$$H|E_j\rangle = E_j|E_j\rangle$$

I don't recall us already diagonalizing σ_z before, so let's do it, but I'll be shorter than usual. The eigenvalues are given by the non-invertibility condition of $H - I\lambda$, as the solutions of

$$\det(H - I\lambda) = (\frac{\omega\hbar}{2} - \lambda)(\lambda - \frac{\omega\hbar}{2}) = 0$$

Hence the two eigenvalues:

$$E_1 = \frac{\omega\hbar}{2}; \qquad E_2 = -\frac{\omega\hbar}{2}$$

From which we can derive the two eigenvectors:

$$\underbrace{\frac{\omega\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}}_{H}|E_1\rangle = \frac{\omega\hbar}{2}|E_1\rangle$$

Assuming an eigenvector of a general form $(a\ b)^T$ yields the following system:

$$\Leftrightarrow \begin{cases} a = a \\ -b = b \end{cases}$$

So $b = 0$; furthermore, as $|E_1\rangle$ must be unitary (from the fundamental theorem/real spectral theorem, we know the eigenvectors of a Hermitian operator, which H most definitely is, are unitary, because the eigenvectors make an orthonormal basis), we must have $a = \pm 1$; let's chose more or less arbitrarily $a = 1$. Hence:

$$\boxed{|E_1\rangle = \begin{pmatrix}1\\0\end{pmatrix}}$$

Similarly for $|E_2\rangle$, assume a general form of $(c\ d)^T$, this yields the following system:

$$\Leftrightarrow \begin{cases} c = -c \\ -d = -d \end{cases}$$

By a similar argument, as before we find:

$$\boxed{|E_2\rangle = \begin{pmatrix}0\\1\end{pmatrix}}$$

Remark 11. *I'm not sure why we have an extra degree of freedom via the signs on the non-zero component of the eigenvectors; I can't think of an extra constraint.*

4. *Use the initial state-vector $|\Psi(0)\rangle$, along with the eigenvectors $|E_j\rangle$ from step 3, to calculate the initial coefficients $\alpha_j(0)$:*

$$\alpha_j(0) = \langle E_j|\Psi(0)\rangle$$

That's an elementary computation:

$$\alpha_1(0) = 1; \qquad \alpha_2(0) = 0$$

5. *Rewrite $|\Psi(0)\rangle$ in terms of the eigenvectors $|E_j\rangle$ and the initial coefficients $\alpha_j(0)$:*

$$|\Psi(0)\rangle = \sum_j \alpha_j(0)|E_j\rangle$$

Again, quite elementary given the quantities involved:

$$|\Psi(0)\rangle = 1|E_1\rangle = |u\rangle = \begin{pmatrix}1\\0\end{pmatrix}$$

6. *In the above equation, replace each* $\alpha_j(0)$ *with* $\alpha_j(t)$ *to capture its time-dependence. As a result,* $|\Psi(0)\rangle$ *becomes* $|\Psi(t)\rangle$*:*

$$|\Psi(t)\rangle = \sum_j \alpha_j(t)|E_j\rangle$$

Naturally:

$$|\Psi(t)\rangle = \alpha_1(t)|E_1\rangle + \alpha_2(t)|E_2\rangle$$

7. *Using Eq.* 4.30[30]*, replace each* $\alpha_j(t)$ *with* $\alpha_j(0)\exp(-\frac{i}{\hbar}E_j t)$*:*

$$|\Psi(t)\rangle = \sum_j \alpha_j(0)\exp(-\frac{i}{\hbar}E_j t)|E_j\rangle$$

Because $\alpha_2(0) = 0$, it only remains:

$$\boxed{|\Psi(t)\rangle = \exp(-\frac{i}{\hbar}t)|u\rangle}$$

OK, then the idea is that if we have an observable L, the probability to measure λ (where λ is then an eigenvalue of L) is given by:

$$P_\lambda(t) = |\,\langle\lambda|\Psi(t)\rangle\,|^2$$

The authors are asking us to consider as an observable $L = \sigma_y$. Recall:

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

This is a matrix corresponding to the spin observable following the y-axis: we *must* expect its eigenvalues to be ± 1 and its eigenvectors to be $|i\rangle$ and $|o\rangle$, but let's compute them all anyway for practice:

$$\det(\sigma_y - I\lambda) = \lambda^2 + i^2 = 0 \Leftrightarrow \lambda = \pm 1$$

For the eigenvectors, again we can assume a general form and solve the corresponding system of equations:

$$\underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\sigma_y}\begin{pmatrix} a \\ b \end{pmatrix} = (+1)\begin{pmatrix} a \\ b \end{pmatrix} \Leftrightarrow \begin{cases} -ib = a \\ ia = b \end{cases}$$

Both equations are actually equivalent (multiply the first one by i to get the second). We furthermore have an additional constraint as the eigenvectors are supposed to be unitary, which yields:

$$|E_1\rangle = \begin{pmatrix} a \\ ia \end{pmatrix} \text{ and } a^2 + (ia)(-ia) = 1 \Leftrightarrow |E_1\rangle = \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix} = |i\rangle$$

Similarly:

$$\underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\sigma_y}\begin{pmatrix} c \\ d \end{pmatrix} = (-1)\begin{pmatrix} c \\ d \end{pmatrix} \Leftrightarrow \begin{cases} -id = -c \\ ic = -d \end{cases}$$

Again, the two equations are equivalent (multiply the first by $-i$ to get the second one), but we have an additional constraint, as the vector must be unitary. In the end, this yields:

$$|E_2\rangle = \begin{pmatrix} c \\ -ic \end{pmatrix} \text{ and } c^2 + (ic)(-ic) = 1 \Leftrightarrow |E_1\rangle = \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{pmatrix} = |o\rangle$$

We may now apply our previous probability formula (Principle 4):

$$P_{+1}(t) = |\,\langle i|\Psi(t)\rangle\,|^2 = |\frac{1}{\sqrt{2}}\exp(-\frac{it}{\hbar})|^2 = \boxed{\frac{1}{2}}$$

And either because the sum of probabilities must be 1, or by explicit computation:

$$P_{-1}(t) = |\,\langle o|\Psi(t)\rangle\,|^2 = |\frac{1}{\sqrt{2}}\exp(-\frac{it}{\hbar})|^2 = \boxed{\frac{1}{2}}$$

[30] This equation corresponds exactly to what this step describes

4.14 Collapse

5 Uncertainty and Time Dependence

5.1 Mathematical Interlude: Complete Sets of Commuting Variables

5.1.1 States That Depend On More Than One Measurable

5.1.2 Wave Functions

5.1.3 A Note About Terminology

5.2 Measurement

Exercise 17. *Verify this claim.*

The claim being that any 2×2 Hermitian matrix can be represented as a linear combination of:

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

The general form of a 2×2 Hermitian matrix is:

$$(\forall (r, r', w) \in \mathbb{R}^2 \times \mathbb{C}), \quad \begin{pmatrix} r & w \\ w^* & r' \end{pmatrix}$$

Recall indeed that because for a Hermitian matrix L we have $L = L^\dagger := (L^*)^T$, hence the diagonal elements must then be real.

Compare then with the general form for a linear combination of the four matrices above:

$$(\forall (a, b, c, d) \in \mathbb{R}^4), \quad a\sigma_x + b\sigma_y + c\sigma_z + dI = \begin{pmatrix} c+d & a-ib \\ a+ib & c-d \end{pmatrix}$$

Clearly we can identify $w \in \mathbb{C}$ with $a - ib$: this is a general form for a complex number, and this naturally identifies w^* with $a + ib$, as expected.

Regarding the remaining parameters, we have on one side two real parameters, and on the other side, two non-equivalent equations involving two parameters, meaning, two degrees of freedom on both sides. So there's room to identify r with $c+d$ and r' with $c-d$. More precisely, given two arbitrary $(r, r') \in \mathbb{R}^2$, we can always find $(c, d) \in \mathbb{R}^2$ such that $r = c + d$ and $r' = c - d$:

$$\begin{cases} r = c + d \\ r' = c - d \end{cases} \Leftrightarrow \begin{cases} c = r - d \\ d = c - r' \end{cases} \Leftrightarrow \begin{cases} c = r - (c - r') \\ d = (r - d) - r' \end{cases} \Leftrightarrow \begin{cases} c = \frac{r+r'}{2} \\ d = \frac{r-r'}{2} \end{cases}$$

Remark 12. *Note that (real) linear combinations of those 4 matrices are isomorphic to* $\mathbb{Q}$[31].

5.3 The Uncertainty Principle

5.4 The Meaning of Uncertainty

5.5 Cauchy-Schwarz Inequality

5.6 The Triangle Inequality and the Cauchy-Schwarz Inequality

5.7 The General Uncertainty Principle

6 Combining Systems: Entanglement

6.1 Mathematical Interlude: Tensor Products

6.1.1 Meet Alice and Bob

[31] https://en.wikipedia.org/wiki/Quaternion