The Theoretical Minimum

Quantum Mechanics - Solutions

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

M. Bivert

September 15, 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:

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

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 relevant axioms:

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 $\langle C| = (\gamma_1^*\ \gamma_2^*\ \gamma_3^*\ \gamma_4^*\ \gamma_5^*)$:

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

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

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

Remark 1. *We could have derived it using complex numbers' exponential's form:*

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

Hence, regarding the second axiom:

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

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 2.
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 $|A\rangle$ along the orthonormal vector $|u\rangle$. This is because in an *orthonormal* basis $(|i\rangle)_{i\in F}$ we have:

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

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. Thus, so far, there seems to be unicity, up to such a phase factor.

Remark 3. *I think some sort of dimensional argument might be required to rigorously prove that indeed there's no way to extract more than three pairs of mutually orthogonal vectors which have a inner-product to* $1/2$*, in a* $\mathbb{C}$*-vector space setting.*

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 at least the following 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 $\beta^*\beta = \gamma^*\gamma = \delta^*\delta = \frac{1}{2}$:

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

---

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 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. You may want to refer to F. Schuller YouTube lectures on quantum mechanics[2] for a deeper mathematical 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.

[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`

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 4. *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 5. *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:*

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

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, something like:

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

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

[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.

You know it's customary to then consider this a function of a single variable x, which for instance, can be a real number:

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

This allows you to graph the polynomial and so forth:

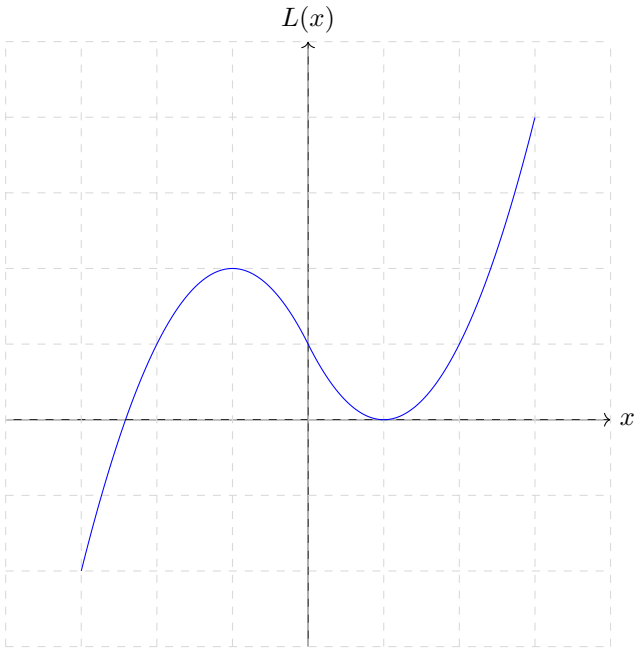


Figure 1: $L(x) = x^2 - 2x + 1$

But that's "kindergarten" polynomials so to speak. "Advanced" 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*

[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.

over a field[7] $\mathbb{F}$. This means that $X^2 - 2X + 1$ is actually a shortcut for:

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

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:

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

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.

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.)

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

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

[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

- 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 coefficients 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:

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

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].

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

[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.

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 $\boldsymbol{v} \in V$.

Pick $\{\boldsymbol{v}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k+1}\} \subset V$ so that $\{\boldsymbol{v}, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{k+1}\}$ 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:

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

Rewrite D_H accordingly, and break it into blocks:

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

[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 $\frac{\lambda}{\|v\|}\boldsymbol{v} = \lambda \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = \lambda \boldsymbol{b_1}$

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

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

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.

□

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 6. *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:

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

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 7. *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:

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

[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

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:

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

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]).

[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`

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 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

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

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 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 8. *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

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

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

[19] L03E01.pdf

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

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.*

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:

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

[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

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:

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

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 9. *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.*

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

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

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:

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

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

[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.

"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).

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*

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

[27] And hopefully, valid...

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

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 10. *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.*

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]:

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

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.

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

Remark 11. *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 12. *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:

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

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

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:

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

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 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 13. *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

Exercise 18. *1) Show that* $\Delta A^2 = \langle \bar{A}^2 \rangle$ *and* $\Delta B^2 = \langle \bar{B}^2 \rangle$

2) Show that $[\bar{A}, \bar{B}] = [A, B]$

3) Using these relations, show that

$$\Delta A \ \Delta B \geq \frac{1}{2} \langle \Psi | [A, B] | \Psi \rangle$$

OK, let's as usual recall the context: A and B are two observables. We defined the expectation value of an observable C with eigenvalues labelled as c to be:

$$\langle C \rangle := \langle \Psi | C | \Psi \rangle = \sum_c cP(c)$$

We construct from C a new observable $\bar{C}$:

$$\bar{C} := C - \langle C \rangle I$$

Where the identity I is sometimes implicit. The eigenvalues of $\bar{C}$ are denoted $\bar{c}$ and can be expressed in terms of C's eigenvalues, denoted c:

$$\bar{c} = c - \langle C \rangle$$

From there, we defined the *standard deviation*, or the square of the uncertainty of C, assuming a "well-behaved" probability distribution P, by:

$$(\Delta C)^2 := \sum_c \bar{c}^2 P(c)$$

Let's first quickly prove that $\bar{c} = c - \langle C \rangle$ are indeed the eigenvalues of $\bar{C} = C - \langle C \rangle I$. Consider an eigenvalue c of C, with associated eigenvector $|c\rangle$. It follows that:

$$\begin{aligned} & C|c\rangle = && c|c\rangle \\ \Leftrightarrow \quad & C|c\rangle - \langle C \rangle |c\rangle = && c|c\rangle - \langle C \rangle |c\rangle \\ \Leftrightarrow \quad & (C - \langle C \rangle I)|c\rangle = && (c - \langle C \rangle)|c\rangle \\ \Leftrightarrow \quad & \bar{C}|c\rangle = && (c - \langle C \rangle)|c\rangle \end{aligned}$$

Meaning, $|c\rangle$ is still an eigenvector of $\bar{C}$, but now associated to the eigenvalue $c - \langle C \rangle$. The $|c\rangle$ still make an orthonormal basis of the state space, so there are no other eigenvectors (there can't be more eigenvectors than the dimension of the surrounding state-space). $\square$

Similarly, we can prove that c^2 are the eigenvalues associated to C^2, for an observable C: again start from an eigenvalue c of C, associated to an eigenvector $|C\rangle$:

$$C|c\rangle = c|c\rangle \Leftrightarrow C(C|c\rangle) = C(c|c\rangle) \Leftrightarrow C^2|c\rangle = c(\underbrace{C|c\rangle}_{c|c\rangle}) \Leftrightarrow C^2|c\rangle = c^2|c\rangle) \quad \square$$

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

---

1) We'll prove the fact for an arbitrary observable C: it'll naturally hold for both A and B.

$$\begin{aligned}(\Delta C)^2 := &\quad \sum_c \bar{c}^2 P(c) \\ = &\quad \sum_c (c - \langle c \rangle)^2 P(c) \quad \text{(definition of } \bar{c}\text{)} \\ = &\quad \langle \Psi | \bar{C}^2 | \Psi \rangle =: \left\langle \bar{C}^2 \right\rangle \quad \text{(two previous properties)} \quad \square\end{aligned}$$

---

2) This is an elementary calculation:

$$\begin{aligned}[\bar{A}, \bar{B}] := &\quad \bar{A}\bar{B} - \bar{B}\bar{A} && \text{(commutator's definition)} \\ = &\quad (A - \langle A \rangle I)(B - \langle B \rangle I) - (B - \langle B \rangle I)(A - \langle A \rangle I) && \text{(definition of } \bar{C}\text{)} \\ = &\quad \Big(AB - \langle A \rangle B - \langle B \rangle A + \langle A \rangle \langle B \rangle I\Big) - \Big(BA - \langle B \rangle A - \langle A \rangle B + \langle B \rangle \langle A \rangle I\Big) \\ = &\quad AB - BA \\ =: &\quad [A, B] && \text{(commutator's definition)} \quad \square\end{aligned}$$

Remember, $\langle A \rangle$ and $\langle B \rangle$ are real numbers (their multiplication is then commutative).

---

3) This is now just about following the reasoning preceding the exercise in the book, as suggested by the authors, by replacing A and B with $\bar{A}$ and $\bar{B}$.

So let:

$$|X\rangle = \bar{A}|\Psi\rangle = (A - \langle A \rangle I)|\Psi\rangle; \qquad |Y\rangle = i\bar{B}|\Psi\rangle = i(B - \langle B \rangle I)|\Psi\rangle$$

Recall the general form of Cauchy-Schwarz for a complex vector space[32]:

$$2|X||Y| \geq |\langle X|Y\rangle + \langle Y|X\rangle|$$

Where the norm is defined from the inner-product:

$$|X| = \sqrt{\langle X|X\rangle}$$

Injecting our two vectors in such a Cauchy-Schwarz equation yields:

$$\begin{aligned}2\sqrt{\left\langle \bar{A}^2 \right\rangle \left\langle \bar{B}^2 \right\rangle} \geq &\quad |i(\langle \Psi|\bar{A}\bar{B}|\Psi\rangle - \langle \Psi|\bar{B}\bar{A}|\Psi\rangle)| \\ \geq &\quad |\langle \Psi|[\bar{A}, \bar{B}]|\Psi\rangle| && \text{(commutator definition)} \\ \geq &\quad |\langle \Psi|[A, B]|\Psi\rangle| && \text{(from 2), } [\bar{A}, \bar{B}] = [A, B])\end{aligned}$$

But from 1), we know that

$$2\sqrt{\left\langle \bar{A}^2 \right\rangle \left\langle \bar{B}^2 \right\rangle} = 2\sqrt{(\Delta A)^2 (\Delta B)^2} = 2\Delta A \Delta B$$

Note that the $\sqrt{.}$ can be removed "safely" as the ΔC^2 are defined as a sum of positive terms (no absolute values necessary).

Putting the two together yields the expected, *general uncertainty principle*:

$$\boxed{\Delta A \Delta B \geq |\langle \Psi|[A, B]|\Psi\rangle|} \quad \square$$

[32] I'm sticking to the authors' terminology and notations.

6 Combining Systems: Entanglement

6.1 Mathematical Interlude: Tensor Products

6.1.1 Meet Alice and Bob

6.1.2 Representing the Combined System

6.2 Classical Correlation

Exercise 19. *Prove that if $P(a, b)$ factorizes, then the correlation between a and b is zero.*

Let's assume that $P(a, b)$ factorizes, meaning, let's assume than there are two functions P_A and P_B such that:

$$P(a, b) = P_A(a)P_B(b)$$

Recall that the authors have defined the (statistical) correlation the quantity[33]:

$$\langle \sigma_A \sigma_B \rangle - \langle \sigma_A \rangle \langle \sigma_B \rangle$$

Where $\langle \sigma_C \rangle$ is the average value of C's observations, also known as the expected value[34], and was defined earlier as:

$$\langle \sigma_C \rangle = \sum_c cP(c)$$

How should we understand $\langle \sigma_A \sigma_B \rangle$? We're trying to find a way to express it as we just did for $\langle \sigma_C \rangle$. It's defined as the average of the product of σ_A and σ_B, meaning, the sum of all possible products of a and b, weighted by some probability distribution, but which one? Well, we don't really know its form specifically, but if for $\langle \sigma_C \rangle$ it was a function of c, then we can guess it must now be a function of a and b: this is the $P(a, b)$ from the exercise statement:

$$\langle \sigma_A \sigma_B \rangle = \sum_a \left(\sum_b abP(a, b) \right)$$

From there, it's just a matter of developing the computation, using our assumption that $P(a, b)$ factorizes:

$$\begin{aligned}
\langle \sigma_A \sigma_B \rangle = & \sum_a \left(\sum_b abP(a, b) \right) \\
= & \sum_a \sum_b (abP_A(a)P_B(b)) \\
= & \sum_a \sum_b ((aP_A(a))(bP_B(b))) \\
= & \left(\sum_a aP_A(a) \right) \left(\sum_b bP_B(b) \right) \\
= & \langle \sigma_A \rangle \langle \sigma_B \rangle \\
\Leftrightarrow & \boxed{\langle \sigma_A \sigma_B \rangle - \langle \sigma_A \rangle \langle \sigma_B \rangle = 0} \quad \square
\end{aligned}$$

Remark 14. *If you're uncertain about the $\sum$ manipulations, you may want to rewrite them as explicit sum over a small number of terms to convince you of their correctness.*

6.3 Combining Quantum Systems

6.4 Two Spins

6.5 Product States

Exercise 20. *Show that if the two normalization conditions of Eqs.* 6.4 *are satisfied, then the state-vector of Eq.* 6.5 *is automatically normalized as well. In other words, show that for this product state, normalizing the overall state-vector does not put any additional constraints on the α's and the β's.*

[33] Precise mathematical formulations are more involved, see for instance `https://en.wikipedia.org/wiki/Correlation`

[34] `https://en.wikipedia.org/wiki/Expected_value`

Recall that we're in the context of two distinct state-spaces, each of them referring to a full-blown spin. Spin states for the first space (Alice's) are denoted:

$$\alpha_u|u\} + \alpha_d|d\}, \quad (\alpha_u, \alpha_d) \in \mathbb{C}^2$$

While spin states for the second space (Bob's) are denoted:

$$\beta_u|u\rangle + \beta_d|d\rangle, \quad (\beta_u, \beta_d) \in \mathbb{C}^2$$

Such states are, as usual, normalized: this is the condition referred to by Eqs. 6.4:

$$\alpha_u^*\alpha_u + \alpha_d^*\alpha_d = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1$$

The two underlying state spaces (complex space, but really, Hilbert spaces) are glued by a tensor product: this allows the creation of new state space, called the *product state space*, which states can refer to both Alice's and Bob's state in a single expression.

Remark 15. *I encourage you to have a look at how Mathematicians formalize the notion of a tensor product of vector spaces: there is for instance a great introductory YouTube video*[35] *by Michael Penn on the topic.*

The core idea is to start with what is called a formal product of vector spaces, which is a new space built from the span of purely "syntactical" combinations of elements of two (or more) vector spaces. Equivalence classes are then used to constrain this span to be a vector space.

For instance, the three following elements would be distinct elements in the formal product of $\mathbb{R}^2$ *and* $\mathbb{R}^3$*:*

$$2\begin{pmatrix}1\\2\end{pmatrix} * \begin{pmatrix}3\\4\\5\end{pmatrix}; \qquad \begin{pmatrix}2\\4\end{pmatrix} * \begin{pmatrix}3\\4\\5\end{pmatrix}; \qquad \begin{pmatrix}1\\2\end{pmatrix} * \begin{pmatrix}6\\8\\10\end{pmatrix}$$

But they would be identified by equivalence classes so as to be the same element in the tensor product of $\mathbb{R}^2$ *and* $\mathbb{R}^3$*. We can keep identifying elements likewise until the operations (sum, scalar product) on the formal product space respect the properties the corresponding operations in a vector space.*

Here's Eq. 6.5, the general form for such a product state, living in the tensor product space created from Alice's and Bob's state spaces (I've just named it Ψ so as to refer to it later on):

$$|\Psi >= \alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle + \alpha_d\beta_u|du\rangle + \alpha_d\beta_d|dd\rangle$$

The claim we have to prove is that this vector is naturally normalized, from the normalization constraints imposed on the individual state spaces.
Let's start by computing the norm of product state (assuming an ordered basis $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$):

$$|\Psi|^2 = \langle\Psi|\Psi\rangle = \begin{pmatrix}(\alpha_u\beta_u)^* & (\alpha_u\beta_d)^* & (\alpha_d\beta_u)^* & (\alpha_d\beta_d)^*\end{pmatrix} \begin{pmatrix}\alpha_u\beta_u\\ \alpha_u\beta_d\\ \alpha_d\beta_u\\ \alpha_d\beta_d\end{pmatrix}$$

We can develop it further, using the fact that for $(a, b) \in \mathbb{C}$, $(ab)^* = a^*b^*$:

$$\begin{aligned}
|\Psi|^2 = & \quad \alpha_u^*\beta_u^*\alpha_u\beta_u + \alpha_u^*\beta_d^*\alpha_u\beta_d + \alpha_d^*\beta_u^*\alpha_d\beta_u + \alpha_d^*\beta_d^*\alpha_d\beta_d \\
= & \quad \alpha_u^*\alpha_u\underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1} + \alpha_d^*\alpha_d\underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1} \\
= & \quad \underbrace{\alpha_u^*\alpha_u + \alpha_d^*\alpha_d}_{=1} \\
= & \quad 1
\end{aligned}$$

But the norm is axiomatically positively defined (i.e. $(\forall\Psi \in \mathcal{H})$, $|\Psi| \geq 0$ with equality iff $\Psi = 0_{\mathcal{H}}$) so:

$$\boxed{|\Psi| = 1} \quad \square$$

[35] `https://www.youtube.com/watch?v=K7f2pCQ3p3U`

6.6 Counting Parameters for the Product State

6.7 Entangled States

Exercise 21. *Prove that the state $|sing\rangle$ cannot be written as a product state.*

Let's recall the definition of the so-called *singlet* state $|\text{sing}\rangle$:

$$|\text{sing}\rangle = \frac{1}{\sqrt{2}}\left(|ud\rangle - |du\rangle\right)$$

As for the previous exercise, we're still in the context of combining two state spaces: Alice's and Bob's, each representing the states of a spin, where the general form of Alice's state vectors is:

$$\alpha_u|u\} + \alpha_d|d\}, \quad (\alpha_u, \alpha_d) \in \mathbb{C}^2$$

While spin states for the second space (Bob's) are denoted:

$$\beta_u|u\rangle + \beta_d|d\rangle, \quad (\beta_u, \beta_d) \in \mathbb{C}^2$$

In this context, let's clarify the difference between a product state and a general composite state, with potential entanglement:

Product state obtained by developing a product between two states from Alice and Bob's state spaces, which yield something along the form:

$$\alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle + \alpha_d\beta_u|du\rangle + \alpha_d\beta_d|dd\rangle$$

Remember from the previous exercise that such a state vector is naturally normalized, as a consequence of the normalization of the underlying vectors from Alice and Bob's space states;

General state for a 2-spins system obtained by linear combination of the vectors from the ordered basis $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$:

$$\Psi = \psi_{uu}|uu\rangle + \psi_{ud}|ud\rangle + \psi_{du}|du\rangle + \psi_{dd}|dd\rangle$$

And impose a normalization condition on the scalar factors:

$$|\Psi| = 1 \Leftrightarrow \psi^*_{uu}\psi_{uu} + \psi^*_{ud}\psi_{ud} + \psi^*_{du}\psi_{du} + \psi^*_{dd}\psi_{dd} = 1$$

Clearly, $|\text{sing}\rangle$ is normalized: it's at least a general state for a 2-spins system. Assume it is a product state. Then there exists $(\alpha_u, \alpha_d, \beta_u, \beta_d) \in \mathbb{C}^4$ such that:

$$\begin{cases} \alpha_u\beta_d = & \frac{1}{\sqrt{2}} \\ \alpha_d\beta_u = & -\frac{1}{\sqrt{2}} \\ \alpha_u\beta_u = & 0 \\ \alpha_d\beta_d = & 0 \end{cases}$$

But now, if $\alpha_u\beta_u = 0$, then at least either $\alpha_u = 0$ or $\beta_u = 0$. Assume $\alpha_u = 0$. But then, we can't have $\alpha_u\beta_d = 1/\sqrt{2}$. Assume then $\beta_u = 0$. Yet in this case, we can't have $\alpha_d\beta_u = -1/\sqrt{2}$.

So the system isn't solvable and our previous assumption can't hold. Hence, there's no such $(\alpha_u, \alpha_d, \beta_u, \beta_d) \in \mathbb{C}^4$, and $|\text{sing}\rangle$ is not a product state. □

6.8 Alice and Bob's Observables

Exercise 22. *Use the matrix forms of σ_z, σ_x, and σ_y and the column vectors for $|u\}$ and $|d\}$ to verify Eqs.* 6.6. *Then, use Eqs.* 6.6 *and* 6.7 *to write the equations that were left out of Eqs.* 6.8. *Use the appendix to check your answers.*

As usual, let's recall our Pauli matrices:

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

The base vectors $|u\}$ and $|d\}$ are the canonical basis vectors for $\mathbb{R}^2$:

$$|u\} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \qquad |d\} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We're trying to understand how for instance an operator σ_x define on Alice's state spaces can be extended to work on a state vector, taken from a combined state space involving Alice's.

The core idea is that the operator will only act on the "component" of the vector that is related to Alice's state space, while leaving the components involving other state spaces untouched.

Eqs. 6.6 (first column below) simply encode how the spin operators act on the basis vectors, in Alice's state space; Eqs. 6.7 (second column below) are identical, but for Bob's state space:

$$\begin{array}{rrrcr}
\sigma_z|u\} = & |u\}; & \tau_z|u\rangle & = & |u\rangle \\
\sigma_z|d\} = & -|d\}; & \tau_z|d\rangle & = & -|d\rangle \\
\sigma_x|u\} = & |d\}; & \tau_x|u\rangle & = & |d\rangle \\
\sigma_x|d\} = & |u\}; & \tau_x|d\rangle & = & |u\rangle \\
\sigma_y|u\} = & i|d\}; & \tau_y|u\rangle & = & i|d\rangle \\
\sigma_y|d\} = & -i|u\}; & \tau_y|d\rangle & = & -i|u\rangle
\end{array}$$

Now verifying that the matrix products indeed evaluates as such is child's play (matrix $\times$ vector products), there's no use of being more explicit here.

For similar reasons, I'll just write a completed 6.8 here, but won't develop the computations: one just have to follow the aforementioned rule: act with the operator on the correct component, extract the eventual scalar factor, and generally update the corresponding vector component. This yields, in agreement with the appendix:

$$\begin{array}{rrrcr}
\sigma_z|uu\rangle = & |uu\rangle; & \tau_z|uu\rangle & = & |uu\rangle \\
\sigma_z|ud\rangle = & |ud\rangle; & \tau_z|ud\rangle & = & -|ud\rangle \\
\sigma_z|du\rangle = & -|du\rangle; & \tau_z|du\rangle & = & |du\rangle \\
\sigma_z|dd\rangle = & -|dd\rangle; & \tau_z|dd\rangle & = & -|dd\rangle \\
\hline
\sigma_x|uu\rangle = & |du\rangle; & \tau_x|uu\rangle & = & |ud\rangle \\
\sigma_x|ud\rangle = & |dd\rangle; & \tau_x|ud\rangle & = & |uu\rangle \\
\sigma_x|du\rangle = & |uu\rangle; & \tau_x|du\rangle & = & |dd\rangle \\
\sigma_x|dd\rangle = & |ud\rangle; & \tau_x|dd\rangle & = & |du\rangle \\
\hline
\sigma_y|uu\rangle = & i|du\rangle; & \tau_y|uu\rangle & = & i|ud\rangle \\
\sigma_y|ud\rangle = & i|dd\rangle; & \tau_y|ud\rangle & = & -i|uu\rangle \\
\sigma_y|du\rangle = & -i|uu\rangle; & \tau_y|du\rangle & = & i|dd\rangle \\
\sigma_y|dd\rangle = & -i|ud\rangle; & \tau_y|dd\rangle & = & -i|du\rangle
\end{array}$$

Exercise 23. *Prove the following theorem:*

When any of Alice's or Bob's spin operators acts on a product state, the result is still a product state.

Show that in a product state, the expectation value of any component of $\boldsymbol{\sigma}$ or $\boldsymbol{\tau}$ is exactly the same as it would be in the individual single-spin states.

Remark 16. *This is a bit long, but fairly straightforward.*

As usual, let's recall the context. We have two state spaces, one for Alice, and one for Bob, each sufficient to describe a spin.

Spin states for Alice's and Bob's spaces are respectively denoted:

$$\alpha_u|u\} + \alpha_d|d\}, \quad (\alpha_u, \alpha_d) \in \mathbb{C}^2; \qquad \beta_u|u\rangle + \beta_d|d\rangle, \quad (\beta_u, \beta_d) \in \mathbb{C}^2$$

Such states are normalized:

$$\alpha_u^*\alpha_u + \alpha_d^*\alpha_d = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1$$

We use a tensor product to join the two spaces. Among all the possible linear combination from the resulting product space, which is a vector space, product states are those of the form (where the αs and βs are constrained by the previous normalization conditions):

$$|\Psi >= \alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle + \alpha_d\beta_u|du\rangle + \alpha_d\beta_d|dd\rangle$$

Now, we want to act on such a product state with an operator from either Alice's state space ($\boldsymbol{\sigma}$) or Bob's ($\boldsymbol{\tau}$), which, as we've saw earlier, can naturally be extended from the individual spaces to the product spaces. Recall that the operators's definition in their own respective state spaces are identical

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

However, when acting on a product state (and more generally, on a vector from the product space), each will respectively only act on the corresponding part of the tensor product gluing basis vectors, for instance:

$$\sigma_x(\gamma|ab\rangle) = \gamma\sigma_x(|a\} \otimes |b\rangle) = \gamma|(\sigma_x(a))b\rangle$$
$$\tau_x(\gamma|ab\rangle) = \gamma\tau_x(|a\} \otimes |b\rangle) = \gamma|a(\tau_x(b))\rangle$$

Because the computation will be exactly symmetric, we're only going to do the work for Alice's operators.

Remark 17. *It would be interesting to see under which circumstances the result generalizes to arbitrary observables (Hermitian operators). It seems we would need for such an operator σ to transform the basis vectors $|u\rangle$ and $|d\rangle$ in such a way that the induced rotation and scaling to reach $\sigma|u\rangle$ and $\sigma|d\rangle$, would somehow balance, so as to preserve the product state constraint. In particular, $\sigma|u\rangle$ and $\sigma|d\rangle$ should be orthogonal.*

This is exactly what happens below, for the spin operators.

Note that:

$$\sigma_x|u\} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |d\}; \qquad \sigma_x|d\} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |u\}$$

Then:

$$\begin{aligned}
\sigma_x|\Psi\rangle = \quad & \alpha_u\beta_u\Big(\underbrace{(\sigma_x|u\})}_{|d\}}\otimes|u\rangle\Big) + \alpha_u\beta_d\Big(\underbrace{(\sigma_x|u\})}_{|d\}}\otimes|d\rangle\Big) + \alpha_d\beta_u\Big(\underbrace{(\sigma_x|d\})}_{|u\}}\otimes|u\rangle\Big) + \alpha_d\beta_d\Big(\underbrace{(\sigma_x|d\})}_{|u\}}\otimes|d\rangle\Big) \\
= \quad & \alpha_u\beta_u|du\rangle + \alpha_u\beta_d|dd\rangle + \alpha_d\beta_u|uu\rangle + \alpha_d\beta_d|ud\rangle \\
= \quad & \alpha_d\beta_u|uu\rangle + \alpha_d\beta_d|ud\rangle + \alpha_u\beta_u|du\rangle + \alpha_u\beta_d|dd\rangle \\
= \quad & \gamma_u\delta_u|uu\rangle + \gamma_u\delta_d|ud\rangle + \gamma_d\delta_u|du\rangle + \gamma_d\delta_d|dd\rangle
\end{aligned}$$

Where, for the last step, we've just introduced some renaming (it'll be made explicit in a moment). Such a state will be a product state if the following hold:

$$\gamma_u^*\gamma_u + \gamma_d^*\gamma_d = 1; \quad \delta_u^*\delta_u + \delta_d^*\delta_d = 1$$

Let's transcribe this in terms of αs and βs:

$$\alpha_d^*\alpha_d + \alpha_u^*\alpha_u = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1$$

Which are but the normalization conditions underlying $|\Psi\rangle$:

$$\alpha_u^*\alpha_u + \alpha_d^*\alpha_d = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1$$

Hence, $\sigma_x|\Psi\rangle$ is a state product. $\square$

We'll now do similar computations, but for σ_y and σ_z. Starting with σ_y, note that:

$$\sigma_y|u\} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} = i|d\}; \qquad \sigma_y|d\} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix} = -i|u\}$$

Then:

$$\begin{aligned}
\sigma_y|\Psi\rangle = \quad & \alpha_u\beta_u\Big(\underbrace{(\sigma_y|u\})}_{i|d\}}\otimes|u\rangle\Big) + \alpha_u\beta_d\Big(\underbrace{(\sigma_y|u\})}_{i|d\}}\otimes|d\rangle\Big) + \alpha_d\beta_u\Big(\underbrace{(\sigma_y|d\})}_{-i|u\}}\otimes|u\rangle\Big) + \alpha_d\beta_d\Big(\underbrace{(\sigma_y|d\})}_{-i|u\}}\otimes|d\rangle\Big) \\
= \quad & i\alpha_u\beta_u|du\rangle + i\alpha_u\beta_d|dd\rangle - i\alpha_d\beta_u|uu\rangle - i\alpha_d\beta_d|ud\rangle \\
= \quad & -i\alpha_d\beta_u|uu\rangle - i\alpha_d\beta_d|ud\rangle + i\alpha_u\beta_u|du\rangle + i\alpha_u\beta_d|dd\rangle \\
= \quad & \gamma_u\delta_u|uu\rangle + \gamma_u\delta_d|ud\rangle + \gamma_d\delta_u|du\rangle + \gamma_d\delta_d|dd\rangle
\end{aligned}$$

Where again, for the last step, we've performed some renaming (again, made explicit in a few lines). For this to be a product state, the following must hold:

$$\gamma_u^*\gamma_u + \gamma_d^*\gamma_d = 1; \quad \delta_u^*\delta_u + \delta_d^*\delta_d = 1$$

Again, transcribed in terms of αs and βs this yields:

$$\begin{gathered}
(-i\alpha_d)^*(-i\alpha_d) + (i\alpha_u)^*(i\alpha_u) = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1 \\
\Leftrightarrow ((i\alpha_d^*)(-i\alpha_d) + (-i\alpha_u^*)(i\alpha_u) = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1) \\
\Leftrightarrow (\alpha_d^*\alpha_d + \alpha_u^*\alpha_u = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1)
\end{gathered}$$

Which again, is the normalization conditions for $|\Psi\rangle$. Hence, $\sigma_y|\Psi\rangle$ is a product state. $\square$

One last time for σ_z, start by observing:

$$\sigma_y|u\} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |u\}; \qquad \sigma_y|d\} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -|d\}$$

Then:

$$\begin{aligned}
\sigma_z|\Psi\rangle = \quad & \alpha_u\beta_u\Big(\underbrace{(\sigma_z|u\})}_{|u\}}\otimes|u\rangle\Big) + \alpha_u\beta_d\Big(\underbrace{(\sigma_z|u\})}_{|u\}}\otimes|d\rangle\Big) + \alpha_d\beta_u\Big(\underbrace{(\sigma_z|d\})}_{-|d\}}\otimes|u\rangle\Big) + \alpha_d\beta_d\Big(\underbrace{(\sigma_z|d\})}_{-|d\}}\otimes|d\rangle\Big) \\
= \quad & \alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle - \alpha_d\beta_u|du\rangle - \alpha_d\beta_d|dd\rangle \\
= \quad & \gamma_u\delta_u|uu\rangle + \gamma_u\delta_d|ud\rangle + \gamma_d\delta_u|du\rangle + \gamma_d\delta_d|dd\rangle
\end{aligned}$$

The renaming is much simpler this time. Let's recall one last time the product state condition:

$$\gamma_u^*\gamma_u + \gamma_d^*\gamma_d = 1; \quad \delta_u^*\delta_u + \delta_d^*\delta_d = 1$$

Or, transcribed in terms of αs and βs:

$$\begin{gathered}
\alpha_u^*\alpha_u + (-\alpha_d)^*(-\alpha_d) = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1 \\
\Leftrightarrow (\alpha_u^*\alpha_u + \alpha_d^*\alpha_d = 1; \quad \beta_u^*\beta_u + \beta_d^*\beta_d = 1)
\end{gathered}$$

Which again, is but the condition for $|\Psi\rangle$ to be a state product. Hence, $\sigma_z|\Psi\rangle$ is a state product. $\square$

It remains to establish the last part of the exercise, namely, that the expectation is unchanged. Recall that for an observable A, given a state $|\Psi\rangle$, the expected value is defined as:

$$\langle A\rangle := \langle\Psi|A|\Psi\rangle$$

Now, we've been computing $A|\Psi\rangle$ in the previous section for all "component" of Alice's spin; so we just have to take a product with $\langle\Psi|$ to get the expected value.

Now remember, we consider an ordered basis $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$ to create column/row vectors, for instance:

$$|\Psi >= \alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle + \alpha_d\beta_u|du\rangle + \alpha_d\beta_d|dd\rangle = \begin{pmatrix}\alpha_u\beta_u\\ \alpha_u\beta_d\\ \alpha_d\beta_u\\ \alpha_d\beta_d\end{pmatrix}$$

We previously established that:

$$\sigma_x|\Psi\rangle = \alpha_d\beta_u|uu\rangle + \alpha_d\beta_d|ud\rangle + \alpha_u\beta_u|du\rangle + \alpha_u\beta_d|dd\rangle$$

Hence:

$$\begin{aligned}
\langle\sigma_x\rangle = &\quad \langle\Psi|(\sigma_x|\Psi\rangle)\\
= &\quad \begin{pmatrix}\alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^*\end{pmatrix}\begin{pmatrix}\alpha_d\beta_u\\ \alpha_d\beta_d\\ \alpha_u\beta_u\\ \alpha_u\beta_d\end{pmatrix}\\
= &\quad \alpha_u^*\beta_u^*\alpha_d\beta_u + \alpha_u^*\beta_d^*\alpha_d\beta_d + \alpha_d^*\beta_u^*\alpha_u\beta_u + \alpha_d^*\beta_d^*\alpha_u\beta_d\\
= &\quad \beta_d^*\beta_d(\alpha_u^*\alpha_d + \alpha_d^*\alpha_u) + \beta_u^*\beta_u(\alpha_u^*\alpha_d + \alpha_d^*\alpha_u)\\
= &\quad \underbrace{(\beta_d^*\beta_d + \beta_u^*\beta_u)}_{=1}(\alpha_u^*\alpha_d + \alpha_d^*\alpha_u)\\
= &\quad \alpha_u^*\alpha_d + \alpha_d^*\alpha_u
\end{aligned}$$

I don't think we've already computed $\langle\Psi|\sigma_x|\Psi\rangle$ in terms of αs and βs before (we did earlier in L03E04 computed it in terms of θ, an angle between two states), so let's do it (I'll use σ_x^A to indicate that we're using σ_x restricted to Alice's space; for clarity, I'll be using the *ordered* basis $\{|u\}, |d\}\}$):

$$\begin{aligned}
\langle\sigma_x^A\rangle = &\quad \{\Psi|\sigma_x^A|\Psi\}\\
= &\quad \begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\alpha_u\\ \alpha_d\end{pmatrix}\\
= &\quad \begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix}\begin{pmatrix}\alpha_d\\ \alpha_u\end{pmatrix}\\
= &\quad \alpha_u^*\alpha_d + \alpha_d^*\alpha_u\\
= &\quad \langle\sigma_x\rangle \quad \square
\end{aligned}$$

Let's do the same thing for $\langle\sigma_y\rangle$; recall that we've computed earlier.

$$\sigma_y|\Psi\rangle = -i\alpha_d\beta_u|uu\rangle - i\alpha_d\beta_d|ud\rangle + i\alpha_u\beta_u|du\rangle + i\alpha_u\beta_d|dd\rangle$$

Hence,

$$\begin{aligned}
\langle\sigma_y\rangle = &\quad \langle\Psi|(\sigma_y|\Psi\rangle) \\
= &\quad \begin{pmatrix}\alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^*\end{pmatrix}\begin{pmatrix}-i\alpha_d\beta_u\\ -i\alpha_d\beta_d\\ i\alpha_u\beta_u\\ i\alpha_u\beta_d\end{pmatrix} \\
= &\quad i(-\alpha_u^*\beta_u^*\alpha_d\beta_u - \alpha_u^*\beta_d^*\alpha_d\beta_d + \alpha_d^*\beta_u^*\alpha_u\beta_u + \alpha_d^*\beta_d^*\alpha_u\beta_d) \\
= &\quad i\Big(\beta_u^*\beta_u(\alpha_d^*\alpha_u - \alpha_u^*\alpha_d) + \beta_d^*\beta_d(\alpha_d^*\alpha_u - \alpha_u^*\alpha_d)\Big) \\
= &\quad i\underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1}(\alpha_d^*\alpha_u - \alpha_u^*\alpha_d) \\
= &\quad i(\alpha_d^*\alpha_u - \alpha_u^*\alpha_d)
\end{aligned}$$

On the other hand:

$$\begin{aligned}
\langle\sigma_y^A\rangle = &\quad \{\Psi|\sigma_y^A|\Psi\} \\
= &\quad \begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}\alpha_u\\ \alpha_d\end{pmatrix} \\
= &\quad \begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix}\begin{pmatrix}-i\alpha_d\\ i\alpha_u\end{pmatrix} \\
= &\quad i(\alpha_d^*\alpha_u - \alpha_u^*\alpha_d) \\
= &\quad \langle\sigma_y\rangle \quad \square
\end{aligned}$$

Finally for $\langle\sigma_z\rangle$, recall:

$$\sigma_z|\Psi\rangle = \alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle - \alpha_d\beta_u|du\rangle - \alpha_d\beta_d|dd\rangle$$

Hence,

$$\begin{aligned}
\langle\sigma_z\rangle = &\quad \langle\Psi|(\sigma_z|\Psi\rangle) \\
= &\quad \begin{pmatrix}\alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^*\end{pmatrix}\begin{pmatrix}\alpha_u\beta_u\\ \alpha_u\beta_d\\ -\alpha_d\beta_u\\ -\alpha_d\beta_d\end{pmatrix} \\
= &\quad \alpha_u^*\beta_u^*\alpha_u\beta_u + \alpha_u^*\beta_d^*\alpha_u\beta_d - \alpha_d^*\beta_u^*\alpha_d\beta_u - \alpha_d^*\beta_d^*\alpha_d\beta_d \\
= &\quad \beta_u^*\beta_u(\alpha_u^*\alpha_u - \alpha_d^*\alpha_d) + \beta_d^*\beta_d(\alpha_u^*\alpha_u - \alpha_d^*\alpha_d) \\
= &\quad \underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1}(\alpha_u^*\alpha_u - \alpha_d^*\alpha_d) \\
= &\quad \alpha_u^*\alpha_u - \alpha_d^*\alpha_d
\end{aligned}$$

And on the other hand:

$$\begin{aligned}
\langle\sigma_z^A\rangle = &\quad \{\Psi|\sigma_z^A|\Psi\} \\
= &\quad \begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\alpha_u\\ \alpha_d\end{pmatrix} \\
= &\quad \begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix}\begin{pmatrix}\alpha_u\\ -\alpha_d\end{pmatrix} \\
= &\quad \alpha_u^*\alpha_u - \alpha_d^*\alpha_d \\
= &\quad \langle\sigma_y\rangle \quad \square
\end{aligned}$$

6.9 Composite Observables

Exercise 24. *Assume Charlie has prepared the two spins in the singlet state. This time, Bob measures τ_y and Alice measure σ_x. What is the expectation value of $\sigma_x\tau_y$?*

What does this say about the correlation between the two measurements?

Let's recall the Pauli matrices involved:

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

And the singlet state:

$$|\text{sing}\rangle = \frac{1}{\sqrt{2}} \left(|ud\rangle - |du\rangle \right)$$

Recall:

$$\sigma_x |u\} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} =: |d\}; \qquad \sigma_x |d\} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} =: |u\}$$

$$\tau_y |u\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} =: i|d\rangle; \qquad \tau_y |d\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix} =: -i|u\rangle$$

The expectation value of $\sigma_x \tau_y$ is then:

$$\begin{aligned}
\langle \sigma_x \tau_y \rangle :=& \quad \langle \text{sing} | \sigma_x \tau_y | \text{sing} \rangle \\
=& \quad \langle \text{sing} | \sigma_x \tau_y \frac{1}{\sqrt{2}} \left(|ud\rangle - |du\rangle \right) \\
=& \quad \frac{1}{\sqrt{2}} \langle \text{sing} | \left(\underbrace{(\sigma_x |u\})}_{|d\}} \otimes \underbrace{(\tau_y |d\rangle)}_{-i|u\rangle} - \underbrace{(\sigma_x |d\})}_{|u\}} \otimes \underbrace{(\tau_y |u\rangle)}_{i|d\rangle} \right) \\
=& \quad \frac{-i}{\sqrt{2}} \langle \text{sing} | \left(|du\rangle + |ud\rangle \right) \\
=& \quad \frac{-i}{2} (\langle ud| - \langle du|)(|du\rangle + |ud\rangle) \\
=& \quad \frac{-i}{2} \left(\underbrace{\langle ud|du\rangle}_{0} + \underbrace{\langle ud|ud\rangle}_{1} - \underbrace{\langle du|du\rangle}_{1} - \underbrace{\langle du|ud\rangle}_{0} \right) \\
=& \quad \boxed{0}
\end{aligned}$$

Remember for the last step that $|du\rangle$ and $|ud\rangle$ are orthonormal basis vectors.

On to the correlation between the two measurements: recall from section 6.2 that the statistical correlation between an observable σ_A in Alice's space and an observable σ_B in Bob's space was defined as the quantity:

$$\langle \sigma_A \sigma_B \rangle - \langle \sigma_A \rangle \langle \sigma_B \rangle$$

Hence in our case, the correlation between the two measurements is (the authors previously computed $\langle \sigma_x \rangle = 0$ and $\langle \sigma_y \rangle = 0$: the computation of $\langle \tau_y \rangle$ would be identical as for the latter)

$$\langle \sigma_x \tau_y \rangle - \langle \sigma_x \rangle \langle \tau_y \rangle = 0$$

Hence, we can conclude that the two measurements aren't correlated at all.

Exercise 25. *Next, Charlie prepares the spins in a different state, called $|T_1\rangle$, where*

$$|T_1\rangle = \frac{1}{\sqrt{2}} \left(|ud\rangle + |du\rangle \right)$$

In these examples, T stands for triplet. These triplet states are completely different from the states in the coin and die examples. What are the expectation values of the operators $\sigma_z \tau_z$, $\sigma_x \tau_x$, and $\sigma_y \tau_y$?

What a difference a sign can make!

This is the same kind of computations there were done in the previous exercise, and earlier in the book. As usual, recall the Pauli matrices:

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

Also recall, from L06E04, the rules for acting on composite state vectors[36]:

$$\begin{array}{rrrrrr}
\sigma_z|uu\rangle = & |uu\rangle; & \tau_z|uu\rangle & = & |uu\rangle \\
\sigma_z|ud\rangle = & |ud\rangle; & \tau_z|ud\rangle & = & -|ud\rangle \\
\sigma_z|du\rangle = & -|du\rangle; & \tau_z|du\rangle & = & |du\rangle \\
\sigma_z|dd\rangle = & -|dd\rangle; & \tau_z|dd\rangle & = & -|dd\rangle \\
\hline
\sigma_x|uu\rangle = & |du\rangle; & \tau_x|uu\rangle & = & |ud\rangle \\
\sigma_x|ud\rangle = & |dd\rangle; & \tau_x|ud\rangle & = & |uu\rangle \\
\sigma_x|du\rangle = & |uu\rangle; & \tau_x|du\rangle & = & |dd\rangle \\
\sigma_x|dd\rangle = & |ud\rangle; & \tau_x|dd\rangle & = & |du\rangle \\
\hline
\sigma_y|uu\rangle = & i|du\rangle; & \tau_y|uu\rangle & = & i|ud\rangle \\
\sigma_y|ud\rangle = & i|dd\rangle; & \tau_y|ud\rangle & = & -i|uu\rangle \\
\sigma_y|du\rangle = & -i|uu\rangle; & \tau_y|du\rangle & = & i|dd\rangle \\
\sigma_y|dd\rangle = & -i|ud\rangle; & \tau_y|dd\rangle & = & -i|du\rangle
\end{array}$$

We now have everything we need to compute the expectation values.

$$\begin{aligned}
\langle\sigma_z\tau_z\rangle := & \langle T_1|\sigma_z\tau_z|T_1\rangle \\
= & \frac{1}{\sqrt{2}}\langle T_1|\sigma_z\tau_z\left(|ud\rangle + |du\rangle\right) \\
= & \frac{1}{\sqrt{2}}\langle T_1|\sigma_z\left(-|ud\rangle + |du\rangle\right) \\
= & -\frac{1}{\sqrt{2}}\langle T_1|\left(|ud\rangle + |du\rangle\right) \\
= & -\frac{1}{2}(\langle ud| + \langle du|)(|ud\rangle + |du\rangle) \\
= & -\frac{1}{2}\left(\underbrace{\langle ud|ud\rangle}_{1} + \underbrace{\langle ud|du\rangle}_{0} + \underbrace{\langle du|ud\rangle}_{0} + \underbrace{\langle du|du\rangle}_{1}\right) \\
= & \boxed{-1}
\end{aligned}$$

For the last step, remember, as for the previous exercise, that $|du\rangle$ and $|ud\rangle$ are orthonormal basis vectors.

$$\begin{aligned}
\langle\sigma_x\tau_x\rangle := & \langle T_1|\sigma_x\tau_x|T_1\rangle \\
= & \frac{1}{\sqrt{2}}\langle T_1|\sigma_x\tau_x\left(|ud\rangle + |du\rangle\right) \\
= & \frac{1}{\sqrt{2}}\langle T_1|\sigma_x\left(|uu\rangle + |dd\rangle\right) \\
= & \frac{1}{\sqrt{2}}\langle T_1|\left(|du\rangle + |ud\rangle\right) \\
= & \frac{1}{2}(\langle ud| + \langle du|)(|du\rangle + |ud\rangle) \\
= & -\frac{1}{2}\left(\underbrace{\langle ud|du\rangle}_{0} + \underbrace{\langle ud|ud\rangle}_{1} + \underbrace{\langle du|du\rangle}_{1} + \underbrace{\langle du|ud\rangle}_{0}\right) \\
= & \boxed{+1}
\end{aligned}$$

[36] You have the same in the book's appendix

$$
\begin{aligned}
\langle \sigma_y \tau_y \rangle := \quad & \langle T_1 | \sigma_y \tau_y | T_1 \rangle \\
= \quad & \frac{1}{\sqrt{2}} \langle T_1 | \sigma_y \tau_y \left(|ud\rangle + |du\rangle \right) \\
= \quad & \frac{1}{\sqrt{2}} \langle T_1 | \sigma_y \left(-i|uu\rangle + i|dd\rangle \right) \\
= \quad & \frac{i}{\sqrt{2}} \langle T_1 | \left(-i|du\rangle - i|ud\rangle \right) \\
= \quad & \frac{1}{2} (\langle ud| + \langle du|)(|du\rangle + |ud\rangle) \\
= \quad & -\frac{1}{2} \left(\underbrace{\langle ud|du\rangle}_{0} + \underbrace{\langle ud|ud\rangle}_{1} + \underbrace{\langle du|du\rangle}_{1} + \underbrace{\langle du|ud\rangle}_{0} \right) \\
= \quad & \boxed{+1}
\end{aligned}
$$

Exercise 26. *Do the same for the other two entangled triplet states,*

$$|T_2\rangle = \frac{1}{\sqrt{2}} \left(|uu\rangle + |dd\rangle \right)$$

$$|T_3\rangle = \frac{1}{\sqrt{2}} \left(|uu\rangle - |dd\rangle \right)$$

As for previous exercise, this is just about crunching numbers. We won't be using the Pauli matrices explicitly here; instead, we'll use the multiplication table from L06E04

$$
\begin{array}{rrrcr}
\sigma_z|uu\rangle = & |uu\rangle; & \tau_z|uu\rangle & = & |uu\rangle \\
\sigma_z|ud\rangle = & |ud\rangle; & \tau_z|ud\rangle & = & -|ud\rangle \\
\sigma_z|du\rangle = & -|du\rangle; & \tau_z|du\rangle & = & |du\rangle \\
\sigma_z|dd\rangle = & -|dd\rangle; & \tau_z|dd\rangle & = & -|dd\rangle \\
\hline
\sigma_x|uu\rangle = & |du\rangle; & \tau_x|uu\rangle & = & |ud\rangle \\
\sigma_x|ud\rangle = & |dd\rangle; & \tau_x|ud\rangle & = & |uu\rangle \\
\sigma_x|du\rangle = & |uu\rangle; & \tau_x|du\rangle & = & |dd\rangle \\
\sigma_x|dd\rangle = & |ud\rangle; & \tau_x|dd\rangle & = & |du\rangle \\
\hline
\sigma_y|uu\rangle = & i|du\rangle; & \tau_y|uu\rangle & = & i|ud\rangle \\
\sigma_y|ud\rangle = & i|dd\rangle; & \tau_y|ud\rangle & = & -i|uu\rangle \\
\sigma_y|du\rangle = & -i|uu\rangle; & \tau_y|du\rangle & = & i|dd\rangle \\
\sigma_y|dd\rangle = & -i|ud\rangle; & \tau_y|dd\rangle & = & -i|du\rangle
\end{array}
$$

As the computations are fairly similar, and to save space, I'll be computing the expectation values for T_2 and T_3 in parallel, distinguishing them by a subscript number.

Let's start with $\langle\sigma_z\tau_z\rangle$:

$$\begin{aligned}
\langle\sigma_z\tau_z\rangle_2 &:= \langle T_2|\sigma_z\tau_z|T_2\rangle \\
&= \frac{1}{\sqrt{2}}\langle T_2|\sigma_z\tau_z\left(|uu\rangle + |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_2|\sigma_z\left(|uu\rangle - |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_2|\left(|uu\rangle + |dd\rangle\right) \\
&= \frac{1}{2}(\langle uu| + \langle dd|)(|uu\rangle + |dd\rangle) \\
&= \frac{1}{2}\left(\underbrace{\langle uu|uu\rangle}_{1} + \underbrace{\langle uu|dd\rangle}_{0} + \underbrace{\langle dd|uu\rangle}_{0} + \underbrace{\langle dd|dd\rangle}_{1}\right) \\
&= \boxed{+1}
\end{aligned}$$

$$\begin{aligned}
\langle\sigma_z\tau_z\rangle_3 &:= \langle T_3|\sigma_z\tau_z|T_3\rangle \\
&= \frac{1}{\sqrt{2}}\langle T_3|\sigma_z\tau_z\left(|uu\rangle - |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_3|\sigma_z\left(|uu\rangle + |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_3|\left(|uu\rangle - |dd\rangle\right) \\
&= \frac{1}{2}(\langle uu| - \langle dd|)(|uu\rangle - |dd\rangle) \\
&= \frac{1}{2}\left(\underbrace{\langle uu|uu\rangle}_{1} - \underbrace{\langle uu|dd\rangle}_{0} - \underbrace{\langle dd|uu\rangle}_{0} + \underbrace{\langle dd|dd\rangle}_{1}\right) \\
&= \boxed{+1}
\end{aligned}$$

Moving on to $\langle\sigma_x\tau_x\rangle$:

$$\begin{aligned}
\langle\sigma_x\tau_x\rangle_2 &:= \langle T_2|\sigma_x\tau_x|T_2\rangle \\
&= \frac{1}{\sqrt{2}}\langle T_2|\sigma_x\tau_x\left(|uu\rangle + |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_2|\sigma_x\left(|ud\rangle + |du\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_2|\left(|dd\rangle + |uu\rangle\right) \\
&= \frac{1}{2}(\langle uu| + \langle dd|)(|dd\rangle + |uu\rangle) \\
&= \frac{1}{2}\left(\underbrace{\langle uu|dd\rangle}_{0} + \underbrace{\langle uu|uu\rangle}_{1} + \underbrace{\langle dd|dd\rangle}_{1} + \underbrace{\langle dd|uu\rangle}_{0}\right) \\
&= \boxed{+1}
\end{aligned}$$

$$\begin{aligned}
\langle\sigma_x\tau_x\rangle_3 &:= \langle T_3|\sigma_x\tau_x|T_3\rangle \\
&= \frac{1}{\sqrt{2}}\langle T_3|\sigma_x\tau_x\left(|uu\rangle - |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_3|\sigma_x\left(|ud\rangle - |du\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_3|\left(|dd\rangle - |uu\rangle\right) \\
&= \frac{1}{2}(\langle uu| - \langle dd|)(|dd\rangle - |uu\rangle) \\
&= \frac{1}{2}\left(\underbrace{\langle uu|dd\rangle}_{0} - \underbrace{\langle uu|uu\rangle}_{1} - \underbrace{\langle dd|dd\rangle}_{1} + \underbrace{\langle dd|uu\rangle}_{0}\right) \\
&= \boxed{-1}
\end{aligned}$$

Finally for $\langle\sigma_y\tau_y\rangle$:

$$\begin{aligned}
\langle\sigma_y\tau_y\rangle_2 &:= \langle T_2|\sigma_y\tau_y|T_2\rangle \\
&= \frac{1}{\sqrt{2}}\langle T_2|\sigma_y\tau_y\left(|uu\rangle + |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_2|\sigma_y\left(i|ud\rangle - i|du\rangle\right) \\
&= \frac{i}{\sqrt{2}}\langle T_2|\left(i|dd\rangle + i|uu\rangle\right) \\
&= -\frac{1}{2}(\langle uu| + \langle dd|)(|dd\rangle + |uu\rangle) \\
&= \frac{-1}{2}\left(\underbrace{\langle uu|dd\rangle}_{0} + \underbrace{\langle uu|uu\rangle}_{1} + \underbrace{\langle dd|dd\rangle}_{1} + \underbrace{\langle dd|uu\rangle}_{0}\right) \\
&= \boxed{-1}
\end{aligned}$$

$$\begin{aligned}
\langle\sigma_y\tau_y\rangle_3 &:= \langle T_3|\sigma_y\tau_y|T_3\rangle \\
&= \frac{1}{\sqrt{2}}\langle T_3|\sigma_y\tau_y\left(|uu\rangle - |dd\rangle\right) \\
&= \frac{1}{\sqrt{2}}\langle T_3|\sigma_y\left(i|ud\rangle + i|du\rangle\right) \\
&= \frac{i}{\sqrt{2}}\langle T_3|\left(i|dd\rangle - i|uu\rangle\right) \\
&= -\frac{1}{2}(\langle uu| - \langle dd|)(|dd\rangle - |uu\rangle) \\
&= \frac{-1}{2}\left(\underbrace{\langle uu|dd\rangle}_{0} - \underbrace{\langle uu|uu\rangle}_{1} - \underbrace{\langle dd|dd\rangle}_{1} + \underbrace{\langle dd|uu\rangle}_{0}\right) \\
&= \boxed{+1}
\end{aligned}$$

We can conclude, from those expectation values alone, that whenever:

- The expectation value is -1, Bob and Alice measure a spin pointing in different directions;
- The expectation value is $+1$, Bob and Alice measure a spin pointing in the same direction.

I just want to spend a few more lines to make something clear. Recall the definition of $|\text{sing}\rangle$:

$$|\text{sing}\rangle = \frac{1}{\sqrt{2}}(|ud\rangle - |du\rangle)$$

The argument of the authors was that, the reason for $\langle \tau_z \sigma_z \rangle$ to be -1 was that $|\text{sing}\rangle$ is built from two spins, one of which is always up while the other is down, and we're measuring both spin alongside the axis on which they are either up or down.

However, in the case of e.g. $\langle \tau_x \sigma_x \rangle$, the answer was not as obviously, because we're in this case measuring the spins alongside the x-axis, and it's not immediate from the expression of $|\text{sing}\rangle$ what kind of balance we have alongside the x-axis.

Let's do a little experiment. Recall the definition of the "basis vectors" for the x-axis, left and right:

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

We want to express, say, T_3 in terms of $|l\rangle$ and $|r\rangle$, to see if indeed, when expressed as such, T_3 is created from two spins such that when one is left, the other is right, which would be concordant with the idea that $\langle \sigma_x \tau_x \rangle_3 = -1$. Let's start by rewriting $|u\rangle$ and $|d\rangle$ in terms of $|r\rangle$ and $|l\rangle$:

$$\begin{cases} |r\rangle = \frac{1}{\sqrt{2}}(|u\rangle + |d\rangle) \\ |l\rangle = \frac{1}{\sqrt{2}}(|u\rangle - |d\rangle) \end{cases} \Leftrightarrow \begin{cases} |u\rangle = \sqrt{2}|r\rangle - |d\rangle \\ |d\rangle = -\sqrt{2}|l\rangle + |u\rangle \end{cases} \Leftrightarrow \begin{cases} |u\rangle = \frac{\sqrt{2}}{2}(|r\rangle + |l\rangle) \\ |d\rangle = \frac{\sqrt{2}}{2}(|r\rangle - |l\rangle) \end{cases}$$

Let's now rewrite T_3 in the $|r\rangle$, $|l\rangle$ basis:

$$\begin{aligned} |T_3\rangle = & \frac{1}{\sqrt{2}}(|uu\rangle - |dd\rangle) \\ = & \frac{1}{\sqrt{2}}(|u\} \otimes |u\rangle - |d\} \otimes |d\rangle) \\ = & \frac{1}{\sqrt{2}}\left(\frac{2}{4}(|r\} + |l\})(|r\rangle + |l\rangle) - \frac{2}{4}(|r\} - |l\})(|r\rangle - |l\rangle)\right) \\ = & \frac{1}{2\sqrt{2}}(|rr\rangle + |rl\rangle + |lr\rangle + |ll\rangle - (|rr\rangle - |rl\rangle - |lr\rangle + |ll\rangle)) \\ = & \frac{1}{\sqrt{2}}(|rl\rangle + |lr\rangle) \end{aligned}$$

And indeed, as expected, T_3 is built from two spins such that when one is left, the other is right. Let's do another one to be sure: consider T_2 on the x-axis: this gives us a $+1$, so we expect a normalized linear combination of $|rr\rangle$ and $|ll\rangle$.

$$\begin{aligned} |T_2\rangle = & \frac{1}{\sqrt{2}}(|uu\rangle + |dd\rangle) \\ = & \frac{1}{\sqrt{2}}(|u\} \otimes |u\rangle + |d\} \otimes |d\rangle) \\ = & \frac{1}{\sqrt{2}}\left(\frac{2}{4}(|r\} + |l\})(|r\rangle + |l\rangle) + \frac{2}{4}(|r\} - |l\})(|r\rangle - |l\rangle)\right) \\ = & \frac{1}{2\sqrt{2}}(|rr\rangle + |rl\rangle + |lr\rangle + |ll\rangle + (|rr\rangle - |rl\rangle - |lr\rangle + |ll\rangle)) \\ = & \frac{1}{\sqrt{2}}(|rr\rangle + |ll\rangle) \end{aligned}$$

Exercise 27. *Prove that the four vectors $|sing\rangle$, $|T_1\rangle$, $|T_2\rangle$, and $|T_3\rangle$ are eigenvectors of $\boldsymbol{\sigma} \cdot \boldsymbol{\tau}$. What are their eigenvalues?*

Recall the definition of those four vectors:

$$\begin{aligned} |\text{sing}\rangle = & \frac{1}{\sqrt{2}}(|ud\rangle - |du\rangle); \qquad |T_1\rangle = \frac{1}{\sqrt{2}}(|ud\rangle + |du\rangle) \\ |T_2\rangle = & \frac{1}{\sqrt{2}}(|uu\rangle + |dd\rangle); \qquad |T_3\rangle = \frac{1}{\sqrt{2}}(|uu\rangle - |dd\rangle) \end{aligned}$$

And the definition of $\boldsymbol{\sigma} \cdot \boldsymbol{\tau}$:

$$\boldsymbol{\sigma} \cdot \boldsymbol{\tau} = \sigma_x \tau_x + \sigma_y \tau_y + \sigma_z \tau_z$$

Again for this exercise, we won't need to explicitly use the Pauli matrices σ_i/τ_j. But actually, we won't even need the multiplication table either, as we've already done most of the work in earlier exercises. Indeed, if we want to prove that $|\Psi\rangle$ is an eigenvector for $\boldsymbol{\sigma}\cdot\boldsymbol{\tau}$, we expect to be able to carry some computation following this pattern:

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|\Psi\rangle &= (\sigma_x\tau_x+\sigma_y\tau_y+\sigma_z\tau_z)|\Psi\rangle\\ &= (\sigma_x\tau_x)|\Psi\rangle+(\sigma_y\tau_y)|\Psi\rangle+(\sigma_z\tau_z)|\Psi\rangle\\ &= \dots\\ &= \lambda_\Psi|\Psi\rangle\end{aligned}$$

But we know from the book that:

$$\sigma_x\tau_x|\text{sing}\rangle=\sigma_y\tau_y|\text{sing}\rangle=\sigma_z\tau_z|\text{sing}\rangle=-|\text{sing}\rangle$$

From L06E07 that

$$\begin{aligned}\sigma_x\tau_x|T_1\rangle &= \frac{1}{\sqrt{2}}(|du\rangle+|ud\rangle) &=:\quad T_1;\\ \sigma_y\tau_y|T_1\rangle &= \frac{1}{\sqrt{2}}(|du\rangle+|ud\rangle) &=:\quad T_1;\\ \sigma_z\tau_z|T_1\rangle &= -\frac{1}{\sqrt{2}}(|du\rangle+|ud\rangle) &=:\quad -T_1;\end{aligned}$$

And from L06E08 that:

$$\begin{aligned}\sigma_x\tau_x|T_2\rangle &= \frac{1}{\sqrt{2}}(|uu\rangle+|dd\rangle) &=:\quad T_2; \quad \sigma_x\tau_x|T_3\rangle &= \frac{1}{\sqrt{2}}(|dd\rangle-|uu\rangle) &=:\quad -T_3;\\ \sigma_y\tau_y|T_2\rangle &= -\frac{1}{\sqrt{2}}(|uu\rangle+|dd\rangle) &=:\quad -T_2; \quad \sigma_y\tau_y|T_3\rangle &= \frac{1}{\sqrt{2}}(|uu\rangle-|dd\rangle) &=:\quad T_3;\\ \sigma_z\tau_z|T_2\rangle &= \frac{1}{\sqrt{2}}(|uu\rangle+|dd\rangle) &=:\quad T_2; \quad \sigma_z\tau_z|T_3\rangle &= \frac{1}{\sqrt{2}}(|uu\rangle-|dd\rangle) &=:\quad T_3.\end{aligned}$$

It follows that:

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|\text{sing}\rangle &= (\sigma_x\tau_x)|\text{sing}\rangle+(\sigma_y\tau_y)|\text{sing}\rangle+(\sigma_z\tau_z)|\text{sing}\rangle &= \boxed{-3}|\text{sing}\rangle\\ (\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|T_1\rangle &= (\sigma_x\tau_x)|T_1\rangle+(\sigma_y\tau_y)|T_1\rangle+(\sigma_z\tau_z)|T_1\rangle &= \boxed{+1}|T_1\rangle\\ (\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|T_2\rangle &= (\sigma_x\tau_x)|T_2\rangle+(\sigma_y\tau_y)|T_2\rangle+(\sigma_z\tau_z)|T_2\rangle &= \boxed{+1}|T_2\rangle\\ (\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|T_3\rangle &= (\sigma_x\tau_x)|T_3\rangle+(\sigma_y\tau_y)|T_3\rangle+(\sigma_z\tau_z)|T_3\rangle &= \boxed{+1}|T_3\rangle\end{aligned}$$

Hence, as foretold by the authors after this exercise, the triplets share a degenerate eigenvalue $(+1)$, while the singlet is associated to a unique eigenvalue (-3), which justifies *a posteriori* their names.

Exercise 28. *A system of two spins has the Hamiltonian*

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

What are the possible energies of the system, and what are the eigenvectors of the Hamiltonian?

Suppose the system starts in the state $|uu\rangle$. What is the state at any later time? Answer the same question for initial states $|ud\rangle$, $|du\rangle$, and $|dd\rangle$.

The first part of the question essentially is about diagonalizing the Hamiltonian: the eigenvalues correspond to the measurable values for the energy. More generally, the exercise is about repeating what we've done earlier in chapter 4, in particular in exercise L04E06, meaning, applying what the authors call the *recipe for a Schrödinger Ket (section* 4.13*)*:

1. Derive, look up, guess, borrow, or steal the Hamiltonian operator H;
2. Prepare an initial state $|\Psi(0)\rangle$;
3. Find the eigenvalues and eigenvectors of H by solving the time-independent Schrödinger equation,

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

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

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

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

7. Using Eq. 4.30[37], replace each $\alpha_j(t)$ with $\alpha_j(0)\exp(-\frac{i}{\hbar}E_j t)$:
$$|\Psi(t)\rangle = \sum_j \alpha_j(0)\exp(-\frac{i}{\hbar}E_j t)|E_j\rangle$$

We'll start by diagonalizing $\boldsymbol{H}$, and then, by loosely applying the rest of the procedure with the various proposed initial states. Recall from the previous exercise that we've found 4 eigenvectors for $\boldsymbol{\sigma}\cdot\boldsymbol{\tau}$:

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|\text{sing}\rangle &= -3|\text{sing}\rangle\\ (\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|T_1\rangle &= +1|T_1\rangle\\ (\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|T_2\rangle &= +1|T_2\rangle\\ (\boldsymbol{\sigma}\cdot\boldsymbol{\tau})|T_3\rangle &= +1|T_3\rangle\end{aligned}$$

Let's recall the expression of those 4 vectors in the up/down basis:

$$\begin{aligned}|\text{sing}\rangle &= \frac{1}{\sqrt{2}}\left(|ud\rangle - |du\rangle\right); \qquad |T_1\rangle = \frac{1}{\sqrt{2}}\left(|ud\rangle + |du\rangle\right)\\ |T_2\rangle &= \frac{1}{\sqrt{2}}\left(|uu\rangle + |dd\rangle\right); \qquad |T_3\rangle = \frac{1}{\sqrt{2}}\left(|uu\rangle - |dd\rangle\right)\end{aligned}$$

It is immediate to check that those eigenvectors all have norm 1, and that they are orthogonal pairwise[38].

Furthermore, we know that $\boldsymbol{\sigma}\cdot\boldsymbol{\tau}$ is an operator in a 4 dimensional vector space $A\otimes B$[39]. And we know from the spectral theorem (aka, the fundamental theorem, proved in L03E01) that the eigenvectors of a Hermitian operator (i.e. an observable) make an orthonormal basis for the surrounding vector space.

Hence we can conclude that our 4 eigenvectors $|\text{sing}\rangle$, $|T_1\rangle$, $|T_2\rangle$, and $|T_3\rangle$ are *the* eigenvectors of $\boldsymbol{\sigma}\cdot\boldsymbol{\tau}$: there are no others, for we've reached the dimension of our vector space $A\otimes B$. By scaling our operator by $\omega/2$, we find back our Hamiltonian $\boldsymbol{H}$, for which we then have the same eigenvectors, only the eigenvalues now need to be shifted likewise:

$$\begin{aligned}\boldsymbol{H}|\text{sing}\rangle &= \frac{-3\omega}{2}|\text{sing}\rangle; \qquad \boldsymbol{H}|T_1\rangle = \frac{+\omega}{2}|T_1\rangle\\ \boldsymbol{H}|T_2\rangle &= \frac{+\omega}{2}|T_2\rangle; \qquad \boldsymbol{H}|T_3\rangle = \frac{+\omega}{2}|T_3\rangle\end{aligned}$$

Hence, we can only measure two values for the energy:

$$\boxed{E_{\text{sing}} = \frac{-3\omega}{2}; \qquad E_{T_1} = E_{T_2} = E_{T_3} = \frac{+\omega}{2}}$$

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

[38] If unsure, compute respectively the norm, which is derived from the inner-product: $||\Psi\rangle| := \sqrt{\langle\Psi|\Psi\rangle}$, and that the same inner-product between two vectors is zero iff said vectors are orthogonal

[39] If this is unclear, you can refer to the beginning on this Chapter (6), where we explore how the combine vector space was built

And our eigenvectors are:

$$\boxed{|\text{sing}\rangle, \quad |T_1\rangle, \quad |T_2\rangle, \quad |T_3\rangle}$$

At this point, we've reached the end of step 3. of the *recipe for a Schrödinger cat* recalled earlier. We're now ready to follow through the other steps, by varying the initial state. Let's start as suggested with $|\Psi_{uu}(0)\rangle = |uu\rangle$: we're trying to rewrite this initial vector state in the basis corresponding to the eigenvectors of our observable (our Hamiltonian).

To this effect, we start by computing the coefficient $\alpha_j(0)$:

$$\begin{aligned}
\alpha_{\text{sing}}(0) &:= \langle \text{sing}|\Psi_{uu}(0)\rangle \\
&= \langle \text{sing}|uu\rangle \\
&= \frac{1}{\sqrt{2}}(\langle ud| - \langle du|)|uu\rangle \\
&= \boxed{0}
\end{aligned}
\qquad
\begin{aligned}
\alpha_{T_1}(0) &:= \langle T_1|\Psi_{uu}(0)\rangle \\
&= \langle T_1|uu\rangle \\
&= \frac{1}{\sqrt{2}}(\langle ud| + \langle du|)|uu\rangle \\
&= \boxed{0}
\end{aligned}$$

$$\begin{aligned}
\alpha_{T_2}(0) &:= \langle T_2|\Psi_{uu}(0)\rangle \\
&= \langle T_2|uu\rangle \\
&= \frac{1}{\sqrt{2}}(\langle uu| + \langle dd|)|uu\rangle \\
&= \boxed{\frac{1}{\sqrt{2}}}
\end{aligned}
\qquad
\begin{aligned}
\alpha_{T_3}(0) &:= \langle T_3|\Psi_{uu}(0)\rangle \\
&= \langle T_3|uu\rangle \\
&= \frac{1}{\sqrt{2}}(\langle uu| - \langle dd|)|uu\rangle \\
&= \boxed{\frac{1}{\sqrt{2}}}
\end{aligned}$$

Hence we can rewrite (step 5.) $|\Psi_{uu}(0)\rangle = |uu\rangle$ in the eigenbase:

$$|\Psi_{uu}(0)\rangle = |uu\rangle = \sum_j \alpha_j(0)|E_j\rangle = \frac{1}{\sqrt{2}}(|T_2\rangle + |T_3\rangle)$$

And from a previous equation (4.30) we can find the evolution over time of our state:

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

That is:

$$\boxed{|\Psi_{uu}(t)\rangle = \frac{1}{\sqrt{2}}\exp(-\frac{\omega i}{2\hbar}t)(|T_2\rangle + |T_3\rangle)}$$

Let's repeat the exact same process, but this time with an initial state $|\Psi_{ud}(0)\rangle = |ud\rangle$. I'll just perform the computation, you can refer to the previous steps if need be.

$$\begin{aligned}
\alpha_{\text{sing}}(0) &:= \langle \text{sing}|\Psi_{ud}(0)\rangle \\
&= \langle \text{sing}|ud\rangle \\
&= \frac{1}{\sqrt{2}}(\langle ud| - \langle du|)|ud\rangle \\
&= \boxed{\frac{1}{\sqrt{2}}}
\end{aligned}
\qquad
\begin{aligned}
\alpha_{T_1}(0) &:= \langle T_1|\Psi_{ud}(0)\rangle \\
&= \langle T_1|ud\rangle \\
&= \frac{1}{\sqrt{2}}(\langle ud| + \langle du|)|ud\rangle \\
&= \boxed{\frac{1}{\sqrt{2}}}
\end{aligned}$$

$$\begin{aligned}
\alpha_{T_2}(0) &:= \langle T_2|\Psi_{ud}(0)\rangle \\
&= \langle T_2|ud\rangle \\
&= \frac{1}{\sqrt{2}}(\langle uu| + \langle dd|)|ud\rangle \\
&= \boxed{0}
\end{aligned}
\qquad
\begin{aligned}
\alpha_{T_3}(0) &:= \langle T_3|\Psi_{ud}(0)\rangle \\
&= \langle T_3|ud\rangle \\
&= \frac{1}{\sqrt{2}}(\langle uu| - \langle dd|)|ud\rangle \\
&= \boxed{0}
\end{aligned}$$

But:

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

So:

$$\boxed{|\Psi_{ud}(t)\rangle = \frac{1}{\sqrt{2}}\left(\exp(\frac{3\omega i}{2\hbar}t)|\text{sing}\rangle + \exp(-\frac{\omega i}{2\hbar}t)|T_1\rangle\right)}$$

Let's do it more time, with an initial state of $|\Psi_{du}(0)\rangle = |du\rangle$.

$$\begin{aligned}
\alpha_{\text{sing}}(0) &:= \langle\text{sing}|\Psi_{du}(0)\rangle & \alpha_{T_1}(0) &:= \langle T_1|\Psi_{du}(0)\rangle \\
&= \langle\text{sing}|du\rangle & &= \langle T_1|du\rangle \\
&= \frac{1}{\sqrt{2}}(\langle ud| - \langle du|)|du\rangle & &= \frac{1}{\sqrt{2}}(\langle ud| + \langle du|)|du\rangle \\
&= \boxed{-\frac{1}{\sqrt{2}}} & &= \boxed{\frac{1}{\sqrt{2}}} \\
\alpha_{T_2}(0) &:= \langle T_2|\Psi_{du}(0)\rangle & \alpha_{T_3}(0) &:= \langle T_3|\Psi_{du}(0)\rangle \\
&= \langle T_2|du\rangle & &= \langle T_3|du\rangle \\
&= \frac{1}{\sqrt{2}}(\langle uu| + \langle dd|)|du\rangle & &= \frac{1}{\sqrt{2}}(\langle uu| - \langle dd|)|du\rangle \\
&= \boxed{0} & &= \boxed{0}
\end{aligned}$$

But:

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

So:

$$\boxed{|\Psi_{du}(t)\rangle = \frac{1}{\sqrt{2}}\left(\exp(-\frac{\omega i}{2\hbar}t)|T_1\rangle - \exp(\frac{3\omega i}{2\hbar}t)|\text{sing}\rangle\right)}$$

One last time, starting from $|\Psi_{dd}(0)\rangle = |dd\rangle$.

$$\begin{aligned}
\alpha_{\text{sing}}(0) &:= \langle\text{sing}|\Psi_{dd}(0)\rangle & \alpha_{T_1}(0) &:= \langle T_1|\Psi_{dd}(0)\rangle \\
&= \langle\text{sing}|dd\rangle & &= \langle T_1|dd\rangle \\
&= \frac{1}{\sqrt{2}}(\langle ud| - \langle du|)|dd\rangle & &= \frac{1}{\sqrt{2}}(\langle ud| + \langle du|)|dd\rangle \\
&= \boxed{0} & &= \boxed{0} \\
\alpha_{T_2}(0) &:= \langle T_2|\Psi_{dd}(0)\rangle & \alpha_{T_3}(0) &:= \langle T_3|\Psi_{dd}(0)\rangle \\
&= \langle T_2|dd\rangle & &= \langle T_3|dd\rangle \\
&= \frac{1}{\sqrt{2}}(\langle uu| + \langle dd|)|dd\rangle & &= \frac{1}{\sqrt{2}}(\langle uu| - \langle dd|)|dd\rangle \\
&= \boxed{\frac{1}{\sqrt{2}}} & &= \boxed{-\frac{1}{\sqrt{2}}}
\end{aligned}$$

But:

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

So:

$$\boxed{|\Psi_{dd}(t)\rangle = \frac{1}{\sqrt{2}}\exp(-\frac{\omega i}{\hbar}t)\left(|T_2\rangle - |T_3\rangle\right)}$$

7 More on Entanglement

7.1 Mathematical Interlude: Tensor Products in Component Form

7.1.1 Building Tensor Product Matrices from Basic Principles

7.1.2 Building Tensor Product Matrices from Component Matrices

Exercise 29. *Write the tensor product $I \otimes \tau_x$ as a matrix, and apply that matrix to each of the $|uu\rangle$, $|ud\rangle$, $|du\rangle$, and $|dd\rangle$ column vectors. Show that Alice's half of the state-vector is unchanged in each case. Recall that I is the 2×2 unit matrix.*

Recall that τ_x is a Pauli matrix, while I really is the identity matrix:

$$\tau_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We saw two different ways of building $I \otimes \tau_x$. Let's start with the first one: consider the usual ordered basis of the underlying composite space: $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$. Then, the elements of the matrix representation of $I \otimes \tau_x$ in this basis are given by:

$$(I \otimes \tau_x)_{ab,cd} = \langle ab|(I \otimes \tau_x)|cd\rangle$$

We can then use the multiplication table from either the appendix or from L06E04, where, remember, τ_x in this multiplication table was a shortcut notation for $I \otimes \tau_x$.

$$\begin{aligned} \tau_x|uu\rangle &= |ud\rangle; & \tau_x|ud\rangle &= |uu\rangle \\ \tau_x|du\rangle &= |dd\rangle; & \tau_x|dd\rangle &= |du\rangle \end{aligned}$$

And we're now ready to evaluate the operator's matrix form:

$$\begin{aligned} I \otimes \tau_x\text{"} = \text{"} & \begin{pmatrix} \langle uu|(I \otimes \tau_x)|uu\rangle & \langle uu|(I \otimes \tau_x)|ud\rangle & \langle uu|(I \otimes \tau_x)|du\rangle & \langle uu|(I \otimes \tau_x)|dd\rangle \\ \langle ud|(I \otimes \tau_x)|uu\rangle & \langle ud|(I \otimes \tau_x)|ud\rangle & \langle ud|(I \otimes \tau_x)|du\rangle & \langle ud|(I \otimes \tau_x)|dd\rangle \\ \langle du|(I \otimes \tau_x)|uu\rangle & \langle du|(I \otimes \tau_x)|ud\rangle & \langle du|(I \otimes \tau_x)|du\rangle & \langle du|(I \otimes \tau_x)|dd\rangle \\ \langle dd|(I \otimes \tau_x)|uu\rangle & \langle dd|(I \otimes \tau_x)|ud\rangle & \langle dd|(I \otimes \tau_x)|du\rangle & \langle dd|(I \otimes \tau_x)|dd\rangle \end{pmatrix} \\ \text{"} = \text{"} & \begin{pmatrix} \langle uu|ud\rangle & \langle uu|uu\rangle & \langle uu|dd\rangle & \langle uu|du\rangle \\ \langle ud|ud\rangle & \langle ud|uu\rangle & \langle ud|dd\rangle & \langle ud|du\rangle \\ \langle du|ud\rangle & \langle du|uu\rangle & \langle du|dd\rangle & \langle du|du\rangle \\ \langle dd|ud\rangle & \langle dd|uu\rangle & \langle dd|dd\rangle & \langle dd|du\rangle \end{pmatrix} \\ \text{"} = \text{"} & \boxed{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}} \end{aligned}$$

Let's move on to the second way, which consists in using Eq. 7.6 of the book:

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix}$$

Which then yields:

$$\begin{aligned} I \otimes \tau_x\text{"} = \text{"} & \begin{pmatrix} 1 \times \tau_x & 0 \times \tau_x \\ 0 \times \tau_x & 1 \times \tau_x \end{pmatrix} \\ \text{"} = \text{"} & \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} \\ \text{"} = \text{"} & \boxed{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}} \end{aligned}$$

Which is exactly what we've found earlier, albeit less tediously.

In our usual ordered basis $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$, the column representations of the basis vectors are as follow:

$$|uu\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}; \quad |ud\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}; \quad |du\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}; \quad |dd\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

Remark 18. *Remember than the column notation is merely a syntactical shortcut over linear combinations of the basis vectors:*

$$\begin{pmatrix}a\\b\\c\\d\end{pmatrix} := a|uu\rangle + b|ud\rangle + c|du\rangle + d|dd\rangle$$

Remark 19. *Note that we could also have used, as the authors did in the book, Eq.* 7.6 *to derive them.*

Then it's just a matter of computing some elementary matrix×vector products. As a shortcut, one can also recall from one's linear algebra class than such products, when they involve basis vectors, are simply a matter of extracting the columns of the matrix (which is fairly trivial to see):

$$(I \otimes \tau_x)|uu\rangle = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} = |ud\rangle; \quad (I \otimes \tau_x)|ud\rangle = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} = |uu\rangle;$$

$$(I \otimes \tau_x)|du\rangle = \begin{pmatrix}0\\0\\0\\1\end{pmatrix} = |dd\rangle; \quad (I \otimes \tau_x)|dd\rangle = \begin{pmatrix}0\\0\\1\\0\end{pmatrix} = |du\rangle$$

Remark 20. *Naturally, this is consistent with the multiplication table we've recalled earlier; and Alice's part of the state is indeed kept unchanged, as expected.*

Exercise 30. *Calculate the matrix elements of $\sigma_z \otimes \tau_x$ by forming inner products as we did in Eq.* 7.2*.*

This is essentially the same exercise as the previous one, but with a different composite operator. To check for errors, I'll still do the computation using the two approaches.

We'll start with the approach suggested in the exercise's statement: let's first start by recalling the portion of interest from the multiplication table computed in L06E04:

$$\begin{aligned}
\sigma_z|uu\rangle &= |uu\rangle; & \tau_x|uu\rangle &= |ud\rangle \\
\sigma_z|ud\rangle &= |ud\rangle; & \tau_x|ud\rangle &= |uu\rangle \\
\sigma_z|du\rangle &= -|du\rangle; & \tau_x|du\rangle &= |dd\rangle \\
\sigma_z|dd\rangle &= -|dd\rangle; & \tau_x|dd\rangle &= |du\rangle
\end{aligned}$$

Then, Eq. 7.2 applied to $\sigma_z \otimes \tau_x$ will give:

$$\sigma_z \otimes \tau_x\text{"} = \text{"} \begin{pmatrix} \langle uu|(\sigma_z \otimes \tau_x)|uu\rangle & \langle uu|(\sigma_z \otimes \tau_x)|ud\rangle & \langle uu|(\sigma_z \otimes \tau_x)|du\rangle & \langle uu|(\sigma_z \otimes \tau_x)|dd\rangle \\ \langle ud|(\sigma_z \otimes \tau_x)|uu\rangle & \langle ud|(\sigma_z \otimes \tau_x)|ud\rangle & \langle ud|(\sigma_z \otimes \tau_x)|du\rangle & \langle ud|(\sigma_z \otimes \tau_x)|dd\rangle \\ \langle du|(\sigma_z \otimes \tau_x)|uu\rangle & \langle du|(\sigma_z \otimes \tau_x)|ud\rangle & \langle du|(\sigma_z \otimes \tau_x)|du\rangle & \langle du|(\sigma_z \otimes \tau_x)|dd\rangle \\ \langle dd|(I \otimes \tau_x)|uu\rangle & \langle dd|(\sigma_z \otimes \tau_x)|ud\rangle & \langle dd|(\sigma_z \otimes \tau_x)|du\rangle & \langle dd|(\sigma_z \otimes \tau_x)|dd\rangle \end{pmatrix}$$

$$\text{"} = \text{"} \begin{pmatrix} \langle uu|\sigma_z|ud\rangle & \langle uu|\sigma_z|uu\rangle & \langle uu|\sigma_z|dd\rangle & \langle uu|\sigma_z|du\rangle \\ \langle ud|\sigma_z|ud\rangle & \langle ud|\sigma_z|uu\rangle & \langle ud|\sigma_z|dd\rangle & \langle ud|\sigma_z|du\rangle \\ \langle du|\sigma_z|ud\rangle & \langle du|\sigma_z|uu\rangle & \langle du|\sigma_z|dd\rangle & \langle du|\sigma_z|du\rangle \\ \langle dd|\sigma_z|ud\rangle & \langle dd|\sigma_z|uu\rangle & \langle dd|\sigma_z|dd\rangle & \langle dd|\sigma_z|du\rangle \end{pmatrix}$$

$$\text{"} = \text{"} \begin{pmatrix} \langle uu|ud\rangle & \langle uu|uu\rangle & -\langle uu|dd\rangle & -\langle uu|du\rangle \\ \langle ud|ud\rangle & \langle ud|uu\rangle & -\langle ud|dd\rangle & -\langle ud|du\rangle \\ \langle du|ud\rangle & \langle du|uu\rangle & -\langle du|dd\rangle & -\langle du|du\rangle \\ \langle dd|ud\rangle & \langle dd|uu\rangle & -\langle dd|dd\rangle & -\langle dd|du\rangle \end{pmatrix}$$

$$\text{"} = \text{"} \boxed{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}}$$

Let's verify our computation using the second approach, relying on Eq. 7.6 of the book:

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix}$$

Recall the Pauli matrices:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \qquad \tau_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Which then yields:

$$\sigma_z \otimes \tau_x\text{"} = \text{"} \begin{pmatrix} 1 \times \tau_x & 0 \times \tau_x \\ 0 \times \tau_x & -1 \times \tau_x \end{pmatrix}$$

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

$$\text{"} = \text{"} \boxed{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}}$$

Which agrees with our previous result.

Exercise 31. *a) Rewrite Eq.* 7.10 *in component form, replacing the symbols A, B, a, and b with the matrices and column vectors from Eqs.* 7.7 *and* 7.8*.*

b) Perform the matrix multiplications Aa and Bb on the right-hand side. Verify that each result is a 4×1 *matrix.*

c) Expand all three Kronecker products.

d) Verify the row and column sizes of each Kronecker product:

- $A \otimes B$ *:* 4×4
- $a \otimes b$ *:* 4×1
- $Aa \otimes Bb$ *:* 4×1

e) Perform the matrix multiplication on the left-hand side, resulting in a 4×1 column vector. Each row should be the sum of four separate terms.

f) Finally, verify that the resulting column vectors on the left and right sides are identical.

Recall Eq. 7.10

$$(A \otimes B)(a \otimes b) = (Aa \otimes Bb)$$

And Eq. 7.7 and 7.8:

$$A \otimes B = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}; \qquad \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \otimes \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} \\ a_{11}b_{21} \\ a_{21}b_{11} \\ a_{21}b_{21} \end{pmatrix}$$

Our goal is to prove Eq. 7.10 by following all the recommended steps. It's a bit tedious, but otherwise presents no major difficulties.

---

a) Let's rewrite the equation (that's still to be proved) in component form:

$$(A \otimes B)(a \otimes b) = (Aa \otimes Bb)$$

$$\Leftrightarrow \left(\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \otimes \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\right)\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \otimes \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}\right) = \left(\left(\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}\right) \otimes \left(\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}\right)\right)$$

---

b) Let's expand Aa and Bb:

$$Aa = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \begin{pmatrix} A_{11}a_{11} + A_{12}a_{21} \\ A_{21}a_{11} + A_{22}a_{21} \end{pmatrix}; \quad Bb = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} = \begin{pmatrix} B_{11}b_{11} + B_{12}b_{21} \\ B_{21}b_{11} + B_{22}b_{21} \end{pmatrix};$$

From Eqs. 7.7 and 7.8, we can see that all Kronecker products indeed expand to 4×1 matrices. Equation 7.10 is then equivalent to:

$$\left(\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \otimes \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}\right)\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \otimes \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}\right) = \left(\begin{pmatrix} A_{11}a_{11} + A_{12}a_{21} \\ A_{21}a_{11} + A_{22}a_{21} \end{pmatrix} \otimes \begin{pmatrix} B_{11}b_{11} + B_{12}b_{21} \\ B_{21}b_{11} + B_{22}b_{21} \end{pmatrix}\right)$$

---

c), d), e), f) I'll be mixing all those steps together, because this is fairly trivial. First, $A \otimes B$ and $a \otimes b$ are respectively Eqs. 7.7 and 7.8. This gives us already:

$$\begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}\begin{pmatrix} a_{11}b_{11} \\ a_{11}b_{21} \\ a_{21}b_{11} \\ a_{21}b_{21} \end{pmatrix} = \left(\begin{pmatrix} A_{11}a_{11} + A_{12}a_{21} \\ A_{21}a_{11} + A_{22}a_{21} \end{pmatrix} \otimes \begin{pmatrix} B_{11}b_{11} + B_{12}b_{21} \\ B_{21}b_{11} + B_{22}b_{21} \end{pmatrix}\right)$$

It remains to expand the last Kronecker product, for which we can use 7.8:

$$\begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}\begin{pmatrix} a_{11}b_{11} \\ a_{11}b_{21} \\ a_{21}b_{11} \\ a_{21}b_{21} \end{pmatrix} = \begin{pmatrix} (A_{11}a_{11} + A_{12}a_{21})(B_{11}b_{11} + B_{12}b_{21}) \\ (A_{11}a_{11} + A_{12}a_{21})(B_{21}b_{11} + B_{22}b_{21}) \\ (A_{21}a_{11} + A_{22}a_{21})(B_{11}b_{11} + B_{12}b_{21}) \\ (A_{21}a_{11} + A_{22}a_{21})(B_{21}b_{11} + B_{22}b_{21}) \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}B_{11}a_{11}b_{11} + A_{11}B_{12}a_{11}b_{21} + A_{12}B_{11}a_{21}b_{11} + A_{12}B_{12}a_{21}b_{21} \\ A_{11}B_{21}a_{11}b_{11} + A_{11}B_{22}a_{11}b_{21} + A_{12}B_{21}a_{21}b_{11} + A_{12}B_{22}a_{21}b_{21} \\ A_{21}B_{11}a_{11}b_{11} + A_{21}B_{12}a_{11}b_{21} + A_{22}B_{11}a_{21}b_{11} + A_{22}B_{12}a_{21}b_{21} \\ A_{21}B_{21}a_{11}b_{11} + A_{21}B_{22}a_{11}b_{21} + A_{22}B_{21}a_{21}b_{11} + A_{22}B_{22}a_{21}b_{21} \end{pmatrix}$$

And it's now trivial to verify that this holds, as expected. $\square$

7.2 Mathematical Interlude: Outer Products

7.3 Density Matrices: A New Tool

7.4 Entanglement and Density Matrices

7.5 Entanglement for Two Spins

Exercise 32. *Calculate the density matrix for:*

$$|\Psi\rangle = \alpha|u\rangle + \beta|d\rangle$$

Answer:

$$\begin{aligned} \psi(u) = &\quad \alpha; \qquad \psi^*(u) = \quad \alpha^* \\ \psi(d) = &\quad \beta; \qquad \psi^*(d) = \quad \beta^* \end{aligned}$$

$$\rho_{a'a} = \begin{pmatrix} \alpha^*\alpha & \alpha^*\beta \\ \beta^*\alpha & \beta^*\beta \end{pmatrix}$$

Now try plugging in some numbers for α and β. Make sure they are normalized to 1. For example, $\alpha = \frac{1}{\sqrt{2}}$, $\beta = \frac{1}{\sqrt{2}}$.

Start by recalling the definition of the density matrix for a single spin in a known state:

$$\rho_{aa'} = \psi^*(a')\psi(a)$$

Now we have no wave function ψ in the exercise statement (the answer set aside), but we can find it by identification with general form of $|\Psi\rangle$:

$$|\Psi\rangle = \sum_{a,b,c,\ldots} \psi(a,b,c,\ldots)|a,b,c,\ldots\rangle$$

Hence, $\psi(u)$ is the component of $|\Psi\rangle$ following the $|u\rangle$ axis, and $\psi(d)$ the one on the $|d\rangle$ axis:

$$\psi(u) = \langle u|\Psi\rangle = \alpha; \qquad \psi(d) = \langle d|\Psi\rangle = \beta;$$

Immediately:

$$\psi^*(u) = \alpha^*; \qquad \psi^*(d) = \beta^*;$$

Then it's just about packaging all the $\rho_{aa'}$ in a matrix: the basis is ordered ($\{|u\rangle, |d\rangle\}$) hence:

$$\rho = \begin{pmatrix} \rho_{uu} & \rho_{ud} \\ \rho_{du} & \rho_{dd} \end{pmatrix} = \begin{pmatrix} \psi^*(u)\psi(u) & \psi^*(d)\psi(u) \\ \psi^*(u)\psi(d) & \psi^*(d)\psi(d) \end{pmatrix} = \boxed{\begin{pmatrix} \alpha^*\alpha & \beta^*\alpha \\ \alpha^*\beta & \beta^*\beta \end{pmatrix}}$$

Remark 21. *We could also use the fact that the density operator is defined as a linear combination of of projectors corresponding to the potential states of the system, each scaled by a probability, and so that the sum of those probabilities is 1, e.g.:*

$$\rho = \sum_i P_i|\psi_i\rangle\langle\psi_i|; \qquad \textit{where: } \sum_i P_i = 1$$

As we're in the case of a single spin in a known state $|\Psi\rangle$, this reduces to

$$\rho = 1|\Psi\rangle\langle\Psi| = |\Psi\rangle\langle\Psi|$$

Assuming again the ordered basis $\{|u\rangle, |d\rangle\}$, we can write $\langle\Psi|$ and $|\Psi\rangle$ in column form, and perform the outer-product:

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

This allows us to double-check our previous result: it seems there's a typo in the exercise statement.

Let's compute a few density matrices for well-known states:

$$|u\rangle = 1|u\rangle + 0|d\rangle \quad \Rightarrow \quad \rho_{|u\rangle} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$|d\rangle = 0|u\rangle + 1|d\rangle \quad \Rightarrow \quad \rho_{|d\rangle} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$$|r\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{1}{\sqrt{2}}|d\rangle \quad \Rightarrow \quad \rho_{|r\rangle} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$|l\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{1}{\sqrt{2}}|d\rangle \quad \Rightarrow \quad \rho_{|l\rangle} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

$$|i\rangle = \frac{1}{\sqrt{2}}|u\rangle + \frac{i}{\sqrt{2}}|d\rangle \quad \Rightarrow \quad \rho_{|i\rangle} = \begin{pmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{pmatrix}$$

$$|o\rangle = \frac{1}{\sqrt{2}}|u\rangle - \frac{i}{\sqrt{2}}|d\rangle \quad \Rightarrow \quad \rho_{|o\rangle} = \begin{pmatrix} 1/2 & i/2 \\ -i/2 & 1/2 \end{pmatrix}$$

The French version of this exercise[40] is a bit more interesting, there are a few additional questions. We can for instance check that ρ is Hermitian:

$$\rho^\dagger = (\rho^*)^T = \begin{pmatrix} (\alpha^*\alpha)^* & (\beta^*\alpha)^* \\ (\alpha^*\beta)^* & (\beta^*\beta)^* \end{pmatrix}^T = \begin{pmatrix} \alpha\alpha^* & \beta\alpha^* \\ \alpha\beta^* & \beta\beta^* \end{pmatrix}^T = \begin{pmatrix} \alpha^*\alpha & \beta^*\alpha \\ \alpha^*\beta & \beta^*\beta \end{pmatrix} =: \rho \quad \square$$

Or that its trace is 1, because of the normalization condition on $|\Psi\rangle$:

$$\mathrm{Tr}(\rho) = \alpha^*\alpha + \beta^*\beta = 1 \quad \square$$

Finally, we can check that ρ projects to $|\Psi\rangle$. Consider a vector which has a component perpendicular to $|\Psi\rangle$, that is, in the direction of $|\Psi^\perp\rangle$, and a component in the direction of $|\Psi\rangle$

$$|\Phi\rangle = \gamma|\Psi^\perp\rangle + \delta|\Psi\rangle$$

By linearity:

$$\rho|\Phi\rangle = \gamma\rho|\Psi^\perp\rangle + \delta\rho|\Psi\rangle$$

Using the fact that $\rho = |\Psi\rangle\langle\Psi|$, we see, by associativity on the products, and by the orthogonality condition between $|\Psi\rangle$ and $|\Psi^\perp\rangle$:

$$\rho|\Psi^\perp\rangle = (|\Psi\rangle\langle\Psi|)|\Psi^\perp\rangle = |\Psi\rangle(\underbrace{\langle\Psi|\Psi^\perp\rangle}_{=0}) = 0$$

On the other hand, by the normalization condition on $|\Psi\rangle$:

$$\rho|\Psi\rangle = (|\Psi\rangle\langle\Psi|)|\Psi^\rangle = |\Psi\rangle(\underbrace{\langle\Psi|\Psi\rangle}_{=1}) = |\Psi\rangle$$

By injecting the two previous results in the one before, it follows that indeed that ρ projects a vector on the $|\Psi\rangle$ direction:

$$\rho|\Phi\rangle = \delta|\Psi\rangle$$

Exercise 33. *a) Show that*

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}^2 = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}$$

b) Now, suppose

$$\rho = \begin{pmatrix} 1/3 & 0 \\ 0 & 2/3 \end{pmatrix}$$

[40] See `https://leminimumtheorique.jimdofree.com/le%C3%A7on-7/exercice-7-4/` for a relevant excerpt, which by the way seems to confirm the typo hypothesis.

Calculate

$$\rho^2$$

$$Tr(\rho)$$

$$Tr(\rho^2)$$

c) If ρ is a density matrix, does it represent a pure state or a mixed state?

The exercise is fairly trivial.

a)

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}^2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \boxed{\begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}} \quad \square$$

b) By application of the previous result,

$$\rho^2 = \begin{pmatrix} 1/3 & 0 \\ 0 & 2/3 \end{pmatrix}^2 = \begin{pmatrix} (1/3)^2 & 0 \\ 0 & (2/3)^2 \end{pmatrix} = \boxed{\begin{pmatrix} 1/9 & 0 \\ 0 & 4/9 \end{pmatrix}}$$

Recall that there's a result alluded to by the authors in a footnote page 195 (section 7.2) that the trace of an operator is the sum of the diagonal elements of any matrix representation of this operator. Hence:

$$\mathrm{Tr}(\rho) = \frac{1}{3} + \frac{2}{3} = \boxed{1}; \qquad \mathrm{Tr}(\rho^2) = \frac{1}{9} + \frac{4}{9} = \boxed{\frac{5}{9}}$$

c) We just saw in the book some properties of density matrices. In particular, for a pure state, and a density matrix ρ, we *must* have:

$$\rho^2 = \rho \text{ and } \mathrm{Tr}(\rho)^2 = 1$$

While for a mixed state, we *must* have:

$$\rho^2 \neq \rho \text{ and } \mathrm{Tr}(\rho)^2 < 1$$

Clearly, in our case, $\boxed{\rho \text{ represents a mixed state.}}$.

Exercise 34. *Use Eq. 7.22 to show that if ρ is a density matrix, then*

$$Tr(\rho) = 1.$$

Eq. 7.22 is the following:

$$P(a) = \rho_{aa}$$

Where $P(a)$ is the probability for an observable L tied to Alice's state space, extended to act on a composite state-space made from Alice's and Bob's, to be measured with the eigenvalue a. On the other hand, ρ_{aa} corresponds to the diagonal elements of Alice's density matrix, expressed in Alice state space.

Well, there will be one $P(a)$ for each eigenvalue, and thus by Eq. 7.22, there is a systematic correspondence with the diagonal elements of the density matrix. But the trace of an operator is defined as the sum of the diagonal elements of a matrix representation of this operator, and it so happens that this value is unique up to a change of basis (meaning, the trace of an operator is the same for all matrix representation of this operator).

Hence because the eigenvalues a represent all the potential measurement values, we know that $\sum_a P(a) = 1$, which by our previous reasoning implies indeed that $\boxed{\mathrm{Tr}(\rho) := \sum_a \rho_{aa} = \sum_a P(a) = 1}$ $\square$

7.6 A Concrete Example: Calculating Alice's Density Matrix

Exercise 35. *Use Eq. 7.24 to calculate ρ^2. How does this result confirm that ρ represents an entangled state? We'll soon discover that there are other ways to check for entanglement.*

Here's Eq. 7.24:

$$\rho = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

From there it's trivial to see that:

$$\rho^2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}^2 = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix}$$

The authors demonstrated earlier a criteria to determine whether a density matrix corresponds to an entangled state or not, at the end of section 7.5: for a pure state, and a density matrix ρ, we *must* have:

$$\rho^2 = \rho \text{ and } \mathrm{Tr}(\rho)^2 = 1$$

While for a mixed or entangled state, we *must* have:

$$\rho^2 \neq \rho \text{ and } \mathrm{Tr}(\rho)^2 < 1$$

Hence, $\boxed{\rho \text{ represents an entangled state}}$.

Exercise 36. *Consider the following states*

$$\begin{aligned} |\psi_1\rangle &= \frac{1}{2}\left(|uu\rangle + |ud\rangle + |du\rangle + |dd\rangle\right) \\ |\psi_2\rangle &= \frac{1}{\sqrt{2}}\left(|uu\rangle + |dd\rangle\right) \\ |\psi_3\rangle &= \frac{1}{5}\left(3|uu\rangle + 4|ud\rangle\right) \end{aligned}$$

For each one, calculate Alice's density matrix, and Bob's density matrix. Check their properties.

Let's recall first the definition of the matrix elements for Alice's density matrix, and second, by symmetry, Bob's:

$$\rho_{a'a} = \sum_b \psi^*(a,b)\psi(a',b); \qquad \rho_{b'b} = \sum_a \psi^*(a,b)\psi(a,b')$$

Let's start with $|\psi_1\rangle$. We know Alice's matrix must be of the form:

$$\rho_A = \begin{pmatrix} \rho_{uu} & \rho_{ud} \\ \rho_{du} & \rho_{dd} \end{pmatrix}$$

And so must be Bob's actually. Filling in with our previous formulas, we obtain:

$$\begin{aligned} \rho_{1A} &= \begin{pmatrix} \psi_1^*(u,u)\psi_1(u,u) + \psi_1^*(u,d)\psi_1(u,d) & \psi_1^*(d,u)\psi_1(u,u) + \psi_1^*(d,d)\psi_1(u,d) \\ \psi_1^*(u,u)\psi_1(d,u) + \psi_1^*(u,d)\psi_1(d,d) & \psi_1^*(d,u)\psi_1(d,u) + \psi_1^*(d,d)\psi_1(d,d) \end{pmatrix} \\ &= \begin{pmatrix} (1/2)(1/2) + (1/2)(1/2) & (1/2)(1/2) + (1/2)(1/2) \\ (1/2)(1/2) + (1/2)(1/2) & (1/2)(1/2) + (1/2)(1/2) \end{pmatrix} \\ &= \boxed{\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}} \end{aligned}$$

Where, remember, the wave function's values correspond to the basis vector coefficients, which are all 1/2 here. By symmetry, we would obtain exactly the same matrix for Bob:

$$\rho_{1B} = \boxed{\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}}$$

Let's check the density matrices properties:

- Clearly, $\rho_{1A} = \rho_{1B}$ is Hermitian;
- Its trace is $1/2 + 1/2 = 1$, as expected;
- Let's compute its square:

$$\rho_{1A}^2 = \rho_{1B}^2 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \rho_{1A} = \rho_{1B}$$

  And $\mathrm{Tr}(\rho_{1A}^2) = \mathrm{Tr}(\rho_{1B}^2) = 1$, from which we can conclude that $\boxed{\psi_1 \text{ is a pure state.}}$
- Without having to compute them explicitly, this implies that its eigenvalues must be 0 and 1.

Let's compute the eigenvalues by partially diagonalizing the matrix anyway for practice: an eigenvector $|\lambda\rangle$ is tied to an eigenvalue λ by:

$$\rho_{1A}|\lambda\rangle = \lambda|\lambda\rangle \Leftrightarrow \rho_{1A}|\lambda\rangle - \lambda|\lambda\rangle = 0 \Leftrightarrow (\rho_{1A} - \lambda I)|\lambda\rangle = 0$$

Because an eigenvector is by definition non-zero, this implies that $\rho_{1A} - \lambda I$ must be non-invertible[41]. This implies that:

$$\det(\rho_{1A} - \lambda I) = 0 \Leftrightarrow 0 = \begin{vmatrix} 1/2 - \lambda & 1/2 \\ 1/2 & 1/2 - \lambda \end{vmatrix} = \left(\frac{1}{2} - \lambda\right)^2 - \frac{1}{2}^2 = (\frac{1}{2} - \lambda - \frac{1}{2})(\frac{1}{2} - \lambda + \frac{1}{2}) = \lambda(\lambda - 1)$$

$$\Leftrightarrow \boxed{\begin{cases} \lambda = 0 \\ \lambda = 1 \end{cases}}$$

As expected.

Let's move on to ψ_2: by a similar reasoning as before we have:

$$\begin{aligned} \rho_{2A} = & \begin{pmatrix} \psi_2^*(u,u)\psi_2(u,u) + \psi_2^*(u,d)\psi_2(u,d) & \psi_2^*(d,u)\psi_2(u,u) + \psi_2^*(d,d)\psi_2(u,d) \\ \psi_2^*(u,u)\psi_2(d,u) + \psi_2^*(u,d)\psi_2(d,d) & \psi_2^*(d,u)\psi_2(d,u) + \psi_2^*(d,d)\psi_2(d,d) \end{pmatrix} \\ = & \begin{pmatrix} (1/\sqrt{2})(1/\sqrt{2}) + (0)(0) & (0)(1/\sqrt{2}) + (1/\sqrt{2})(0) \\ (1/\sqrt{2})(0) + (0)(1/\sqrt{2}) & (0)(0) + (1/\sqrt{2})(1/\sqrt{2}) \end{pmatrix} \\ = & \boxed{\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}} \end{aligned}$$

Again, by a symmetry argument, we can already conclude that $\rho_{2B} = \rho_{2A}$ (the idea is that you can swap the labels corresponding to Bob and Alice in the description of the state ψ_2 and by reordering the terms, you see that the state is unchanged).

Finally, let's check the density matrices properties:

1. Clearly Hermitian;
2. $\mathrm{Tr}(\rho_{2A}) = 1/2 + 1/2 = 1$;
3. Let's compute the square to determine the state quality:

$$\rho_{2A}^2 = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} \neq \rho_{2A}$$

   and $\mathrm{Tr}(\rho_{2A}^2) = 1/2 < 1$: $\boxed{\psi_2 \text{ is a mixed state}}$;
4. The matrix is diagonal: clearly, all its eigenvalue (there's a single degenerate eigenvalue $1/2$) are positive and ≤ 1.

[41] For otherwise, multiply both sides of the equation by its inverse: LHS is equal to $|\lambda\rangle$ while the RHS is still equal to 0

Moving on to the last one. Observe that this time, there is not symmetry between Alice and Bob matrices, so we'll have to compute them both.

$$\begin{aligned}
\rho_{3A} = & \begin{pmatrix} \psi_3^*(u,u)\psi_3(u,u) + \psi_3^*(u,d)\psi_3(u,d) & \psi_3^*(d,u)\psi_3(u,u) + \psi_3^*(d,d)\psi_3(u,d) \\ \psi_3^*(u,u)\psi_3(d,u) + \psi_3^*(u,d)\psi_3(d,d) & \psi_3^*(d,u)\psi_3(d,u) + \psi_3^*(d,d)\psi_3(d,d) \end{pmatrix} \\
= & \begin{pmatrix} (3/5)(3/5) + (4/5)(4/5) & (0)(3/5) + (0)(4/5) \\ (3/5)(0) + (4/5)(0) & (0)(0) + (0)(0) \end{pmatrix} \\
= & \begin{pmatrix} 9/25 + 16/25 & 0 \\ 0 & 0 \end{pmatrix} \\
= & \boxed{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}
\end{aligned}$$

Regarding density matrices properties:

1. Hermitian;
2. $\mathrm{Tr}(\rho_{3A}) = 1 + 0 = 1$;
3. $\rho_{3A}^2 = \rho_{3A}$: $\boxed{\psi_3 \text{ is a pure state}}$;
4. This is confirmed by the eigenvalues 1 and 0 (matrix trivially diagonal).

Remains Bob's matrix!

$$\begin{aligned}
\rho_{3B} = & \begin{pmatrix} \psi_3^*(u,u)\psi_3(u,u) + \psi_3^*(d,u)\psi_3(d,u) & \psi_3^*(u,u)\psi_3(u,d) + \psi_3^*(d,u)\psi_3(d,d) \\ \psi_3^*(u,d)\psi_3(u,u) + \psi_3^*(d,d)\psi_3(d,u) & \psi_3^*(u,d)\psi_3(u,d) + \psi_3^*(d,d)\psi_3(d,d) \end{pmatrix} \\
= & \begin{pmatrix} (3/5)(3/5) + (0)(0) & (3/5)(4/5) + (0)(0) \\ (4/5)(3/5) + (0)(0) & (4/5)(4/5) + (0)(0) \end{pmatrix} \\
= & \boxed{\frac{1}{25}\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}}
\end{aligned}$$

One last time, let's check its density matrices properties:

1. Clearly Hermitian;
2. $\mathrm{Tr}(\rho_{3B}) = 9/25 + 16/25 = 1$;
3. Let's square it to determine the state quality:

$$\begin{aligned}
\rho_{3B}^2 = & \frac{1}{25^2}\begin{pmatrix} 9\times 9 + 12\times 12 & 9\times 12 + 12\times 16 \\ 12\times 9 + 16\times 12 & 12\times 12 + 16\times 16 \end{pmatrix} \\
= & \frac{1}{25^2}\begin{pmatrix} 81 + 100 + 40 + 4 & 90 + 18 + 100 + 80 + 12 \\ 90 + 18 + 100 + 80 + 12 & 100 + 40 + 4 + 100 + 120 + 36 \end{pmatrix} \\
= & \frac{1}{25^2}\begin{pmatrix} 225 & 300 \\ 300 & 400 \end{pmatrix} \\
= & \frac{1}{25^2}\begin{pmatrix} (4\times 2 + 1)\times 25 & 3\times 4\times 25 \\ 3\times 4\times 25 & 4\times 4\times 25 \end{pmatrix} \\
= & \frac{1}{25}\begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix} = \rho_{3B}
\end{aligned}$$

   Thus $\mathrm{Tr}(\rho_{3B}^2) = \mathrm{Tr}(\rho_{3B}) = 1$ and $\boxed{\psi_3 \text{ is a pure state}}$;
4. This implies again that its eigenvalues must be 0 and 1

Let's compute the eigenvalues for practice, going a bit faster this time:

$$\begin{vmatrix} 9/25 - \lambda & 12/25 \\ 12/25 & 16/25 - \lambda \end{vmatrix} = 0 \Leftrightarrow \left(\left(\frac{9}{25} - \lambda\right)\left(\frac{16}{25} - \lambda\right) - \left(\frac{12}{25}\right)^2\right) = 0$$

$$\Leftrightarrow \lambda^2 - \lambda + \frac{9 \times 16}{25^2} - \left(\frac{12}{25}\right)^2 = 0$$

$$\Leftrightarrow \lambda^2 - \lambda + \frac{3 \times 3 \times 4 \times 4}{25^2} - \frac{3 \times 4 \times 3 \times 4}{25^2} = 0$$

$$\Leftrightarrow \lambda(\lambda - 1) = 0$$

$$\Leftrightarrow \boxed{\begin{cases} \lambda = 0 \\ \lambda = 1 \end{cases}}$$

7.7 Tests for Entanglement

7.7.1 The Correlation Test for Entanglement

Exercise 37. *Given any Alice observable* $\boldsymbol{A}$ *and Bob observable* $\boldsymbol{B}$*, show that for a product state, the correlation* $C(A, B)$ *is zero.*

Recall that we're in the context of a composite system S_{AB} made from two state spaces, one corresponding to Alice, S_A, and one corresponding to Bob, S_B, mathematically tied by a tensor product.

The correlation $C(\boldsymbol{A}, \boldsymbol{B})$ between two observables $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined as[42]:

$$C(\boldsymbol{A}, \boldsymbol{B}) := \langle \boldsymbol{AB} \rangle - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle$$

Remember that the authors proved[43] that the expected value $\langle \boldsymbol{L} \rangle$ of an observable $\boldsymbol{L}$ being in a state $|\Psi\rangle$ is:

$$\langle \boldsymbol{L} \rangle = \langle \Psi | \boldsymbol{L} | \Psi \rangle$$

Remark 22. *There's an issue in this first derivation, reported by Jannis Koeckeritz; I've left it so you can "have fun" trying to find it on your own; the solution is in this footnote*[44]*.*

Here's a first derivation, where we use the following formula[45] defined for an observable $\boldsymbol{L}$, and a system described by a density matrix ρ:

$$\langle \boldsymbol{L} \rangle = \mathrm{Tr}(\rho \boldsymbol{L})$$

Recall[46] that for any operator $\boldsymbol{A}$ and $\boldsymbol{B}$, in particular, where $\boldsymbol{AB} \neq \boldsymbol{BA}$, we still have:

$$\mathrm{Tr}(\boldsymbol{AB}) = \mathrm{Tr}(\boldsymbol{BA})$$

We also know[47] that, because we're dealing with *a* product state, this can't be a mixed state (it cannot be expressed as a weighted sum of multiple states), i.e if we name $|\Psi\rangle$ that (pure) product state:

$$\rho = |\Psi\rangle\langle\Psi|$$

Finally[48], again because that product state is pure, we have $\rho^2 = \rho$, which should be clear from the previous expression of ρ, as $|\Psi\rangle$ is normalized ($\sqrt{\langle\Psi|\Psi\rangle} = 1$) and the "product(s)" being associative.

---

[42] The authors are a bit irregular in their use of boldface for operators; I'll try to do better, but things should be clear from the context

[43] p106, section 4.7 - *Expectation values*

[44] The trace is invariant *only* under cyclic permutations: we can't jump from $\mathrm{Tr}(\rho^2 \boldsymbol{AB})$ to $\mathrm{Tr}(\rho \boldsymbol{A} \rho \boldsymbol{B})$ knowing only $\mathrm{Tr}(\boldsymbol{AB}) = \mathrm{Tr}(\boldsymbol{BA})$.

[45] p206, section 7.5 - *Entanglement for two spins*

[46] p209, section 7.5 - *Entanglement for two spins*

[47] p202, section 7.5 - *Entanglement for two spins*

[48] p207, section 7.5 - *Entanglement for two spins*

It follows that:

$$\begin{aligned} C(A,B) := &\quad \langle \boldsymbol{AB} \rangle - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \\ = &\quad Tr(\rho \boldsymbol{AB}) - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \\ = &\quad Tr(\rho^2 \boldsymbol{AB}) - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \\ = &\quad Tr(\rho(\boldsymbol{A}\rho\boldsymbol{B})) - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \\ = &\quad \langle \boldsymbol{A}\rho\boldsymbol{B} \rangle - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \\ = &\quad \langle \Psi | \boldsymbol{A}\rho\boldsymbol{B} | \Psi \rangle - \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \\ = &\quad \langle \Psi | \boldsymbol{A}\rho\boldsymbol{B} | \Psi \rangle - \langle \Psi | \boldsymbol{A} \underbrace{|\Psi\rangle\langle\Psi|}_{\rho} \boldsymbol{B} | \Psi \rangle \\ = &\quad \boxed{0} \quad \square \end{aligned}$$

Here's a second solution, rephrased from Michel Rennes's approach.

We start by expressing the expectation value in terms of an inner-product again, assuming we start in the state $|\Psi\rangle$:

$$\langle \boldsymbol{AB} \rangle = \langle \Psi | \boldsymbol{AB} | \Psi \rangle$$

Then, recall that $\boldsymbol{A}$ and $\boldsymbol{B}$ are two observables respectively from Alice and Bob' state spaces, which have been extended, as previously studied, so as to be able to act on a state vector $|\Psi\rangle$, taken from the composite system S_{AB}.

We definitely need this to be able to express the correlation $C(\boldsymbol{A}, \boldsymbol{B})$ in terms of those inner-products, for otherwise, the second terms in the equation below applying $\boldsymbol{A}$ or $\boldsymbol{B}$ to $|\Psi\rangle$ wouldn't make any sense:

$$C(\boldsymbol{A}, \boldsymbol{B}) = \langle \Psi | \boldsymbol{AB} | \Psi \rangle - \langle \Psi | \boldsymbol{A} | \Psi \rangle \langle \Psi | \boldsymbol{B} | \Psi \rangle$$

Hence there's an abuse of notation: with $\boldsymbol{I}_X$ being the identity operator on the space S_X:

$$\boldsymbol{A} \text{ " = " } \boldsymbol{A} \otimes \boldsymbol{I}_B; \quad \boldsymbol{B} \text{ " = " } \boldsymbol{I}_A \otimes \boldsymbol{B}$$

For clarity, I'll note $\boldsymbol{A}_A$ the observable $\boldsymbol{A}$ expressed in the system S_A, and similarly for $\boldsymbol{B}_B$:

$$\boldsymbol{A} = \boldsymbol{A}_A \otimes \boldsymbol{I}_B; \quad \boldsymbol{B} = \boldsymbol{I}_A \otimes \boldsymbol{B}_B$$

Regarding $|\Psi\rangle$, this is a product state, and we know[49] that it can be expressed as a tensor product of a state in S_A and of a state in S_B:

$$|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle$$

We can then rewrite:

$$\begin{aligned} \langle \boldsymbol{AB} \rangle = &\quad \langle \Psi | \boldsymbol{AB} | \Psi \rangle \\ = &\quad (\langle \psi | \otimes \langle \phi |) \, \boldsymbol{AB} \, (|\psi\rangle \otimes |\phi\rangle) \\ = &\quad (\langle \psi | \otimes \langle \phi |) \, \boldsymbol{A} \, ((\boldsymbol{I}_A \otimes \boldsymbol{B}_B) \, (|\psi\rangle \otimes |\phi\rangle)) \\ = &\quad (\langle \psi | \otimes \langle \phi |) \, \boldsymbol{A} \left(\underbrace{\boldsymbol{I}_A |\psi\rangle}_{|\psi\rangle} \otimes \boldsymbol{B}_B |\phi\rangle \right) \\ = &\quad (\langle \psi | \otimes \langle \phi |) \, (\boldsymbol{A}_A |\psi\rangle \otimes \boldsymbol{B}_B |\phi\rangle) \end{aligned}$$

Where I've skipped the development for the application of $\boldsymbol{A}$ (same procedure as for applying $\boldsymbol{B}$). Then, observe[50] that $\langle \psi |$ is an operator defined on S_A, and similarly for $\langle \phi |$ being an operator defined on S_B. Their tensor product is then an operator defined on S_{AB} and the usual rules for applying this combined operator hold:

$$\begin{aligned} \langle \boldsymbol{AB} \rangle = &\quad (\langle \psi | \boldsymbol{A}_A | \psi \rangle) \otimes (\langle \phi | \boldsymbol{B}_B | \phi \rangle) \\ = &\quad \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle \end{aligned}$$

[49] p164, section 6.5 - *Product states*

[50] It would be interesting to formalized that more thoroughly. If I'm not mistaken the idea is that the bras of $S_A \otimes S_B$ can be expressed as a combination of one bra from S_A and one bra from S_B. More precisely, the bras being elements of the dual spaces, it's because of the following (canonical) isomorphism: $S^*_{AB} = (S_A \otimes S_B)^* \cong S^*_A \otimes S^*_B$, see for instance `https://planetmath.org/tensorproductofdualspacesisadualspaceoftensorproduct`

Hence clearly, $\boxed{C(\boldsymbol{A},\boldsymbol{B}) := \langle \boldsymbol{AB}\rangle - \langle \boldsymbol{A}\rangle \langle \boldsymbol{B}\rangle = 0}$.

For completeness, here's one last solution, rephrased from Filip Van Lijsebetten's approach (p52), which relies on the probabilistic definition of the average value.

Remember that the average value of an observable $\boldsymbol{L}$ is (mathematically) defined[51] as:

$$\langle \boldsymbol{L}\rangle := \sum_i \lambda_i P(\lambda_i)$$

Hence:

$$\langle \boldsymbol{AB}\rangle = \sum_{ab} \lambda_{ab} P(\lambda_{ab}); \quad \langle \boldsymbol{A}\rangle = \sum_a \lambda_a P(\lambda_a); \quad \langle \boldsymbol{B}\rangle = \sum_b \lambda_b P(\lambda_b)$$

Recall that the ab corresponds to all labels created by concatenating all potential values for a and b. This means that we'll have $\sum_{ab} = \sum_{a,b} := \sum_a \sum_b$. Let's rewrite the correlation $C(\boldsymbol{A},\boldsymbol{B})$:

$$\begin{aligned}
C(\boldsymbol{A},\boldsymbol{B}) := \quad & \langle \boldsymbol{AB}\rangle - \langle \boldsymbol{A}\rangle \langle \boldsymbol{B}\rangle \\
= \quad & \left(\sum_{ab} \lambda_{ab} P(\lambda_{ab})\right) - \left(\sum_a \lambda_a P(\lambda_a)\right)\left(\sum_b \lambda_b P(\lambda_b)\right) \\
= \quad & \left(\sum_{ab} \lambda_{ab} P(\lambda_{ab})\right) - \left(\sum_a \lambda_a P(\lambda_a)\left(\sum_b \lambda_b P(\lambda_b)\right)\right) \\
= \quad & \left(\sum_{ab} \lambda_{ab} P(\lambda_{ab})\right) - \left(\sum_a \sum_b \lambda_a P(\lambda_a)\lambda_b P(\lambda_b)\right) \\
= \quad & \left(\sum_{ab} \lambda_{ab} P(\lambda_{ab})\right) - \left(\sum_{a,b} \lambda_a \lambda_b P(\lambda_a) P(\lambda_b)\right) \\
= \quad & \sum_{a,b} \Big(\lambda_{ab} P(\lambda_{ab}) - \lambda_a \lambda_b P(\lambda_a) P(\lambda_b)\Big)
\end{aligned}$$

Now the notation is a bit confusing[52], but recall than λ_{ab} corresponds to the value we get for our combined state (which occurs with a probability of $P(\lambda_{ab})$). And this precisely corresponds the fact that we have λ_a in the subspace S_A and λ_b in the subspace S_B: so we can read it like $\lambda_{ab} \simeq \lambda_a \lambda_b$. Hence this factors as:

$$C(\boldsymbol{A},\boldsymbol{B}) = \sum_{a,b} \lambda_{ab} \big(P(\lambda_{ab}) - P(\lambda_a) P(\lambda_b)\big)$$

Remark 23. *So far, we've essentially just restated with a different notation what we did in L06E01*

Now by definition for a product state, there is independence between the two "events": the measurement of either A or B doesn't affect the other one. That is, $P(\lambda_{ab}) = P(\lambda_a)P(\lambda_b)$[53], hence the correlation really is zero. □

7.7.2 The Density Matrix Test for Entanglement

7.8 The Process of Measurement

Exercise 38. *Verify that the state-vector in* 7.30 *represents a completely untangled state.*

Let's recall the state-vector from 7.30, and let's call it $|\Psi\rangle$.

$$|\Psi\rangle = \alpha_u |u,b\rangle + \alpha_d |d,b\rangle$$

[51] p105, section 4.7 - *Expectation values*

[52] I could have made things a bit clearer: for instance, we really have three different probability distributions, one for each state involved, but they are all denoted very similarly.

[53] This is the definition of independence of events in ordinary probability theory: `https://en.wikipedia.org/wiki/Independence_(probability_theory)#For_events`

As I've found this confusing, let me start by recalling a bit of vocabulary[54]. A quantum state can be either **pure** or **mixed**: either its a single state, or a *convex combination*[55] of pure states. This is true for a "regular" state space, as for a state space built via a tensor products of two (or finitely many, by induction) other state spaces.

Now there's a second qualification, that is only applicable for states which are taken from a state space made by glueing two (again, or finitely many) other state spaces: **entangled** states, and **disentangled** states.

Mixed and *entangled* are definitely not synonymous: you can have a non-mixed (i.e. pure) entangled state for example.

Example 1. *The state vector from* 7.30 *is a pure state: this is not a convex combination of states. But this tells us absolutely nothing regarding whether it's an entangled state. We know however that it makes sense to talk about it being entangled or not, as we're dealing with a combined system involving* (i) *an apparatus and* (ii) *a spin to be measured by said apparatus.*

We could test this purity by computing the density matrix ρ*, and checking whether* $\rho^2 = \rho$ *or* $Tr(\rho) = 1$*.*

Let's clarify the vocabulary one step further: a **completely untangled state** *is* a **product state**: that's a state where measurements on one subsystem affect in no ways the other subsystem(s).

From there, we have a few different ways of proceeding.

The simplest approach is to remember that a state is a product state when it can be expressed via two components (well, or more, but we're in the case where there are two subsystems here: the apparatus, and the spin to be measured with the apparatus), one for each subsystem. Recall that $|a, \alpha\rangle$ really is a shortcut for $|a\rangle \otimes |\alpha\}$. This means, the prepared state really is:

$$\alpha_u |u\rangle \otimes |b\} + \alpha_d |d\rangle \otimes |b\}$$

But the tensor product distributes[56], hence this simplifies as:

$$(\alpha_u |u\rangle + \alpha_d |d\rangle) \otimes |b\}$$

As the combined state is normalized, we must have $\sqrt{\alpha_u^2 + \alpha_d^2} = 1$, which implies that the sub-state corresponding to the spin is also normalized. Trivially, the sub-state corresponding to the apparatus is also normalized. Hence, we've expressed our combined state as a tensor product of two normalized state, one for each subsystems: this is a product state.

A slightly more involved (calculus-wise) variant of this approach would be to rely on the general form of the product state[57] and to evaluate whether our state vector can be expressed in such a way. The general form can be computed, again using the distributive nature of the tensor product:

$$\begin{aligned}
|\text{product state}\rangle = & \left\{\alpha_u |u\rangle + \alpha_d |d\rangle\right\} \otimes \left\{\beta_b |b\} + \beta_+ |+1\} + \beta_- |-1\}\right\} \\
= & \ \alpha_u |u\rangle \otimes \left\{\beta_b |b\} + \beta_+ |+1\} + \beta_- |-1\}\right\} + \alpha_d |d\rangle \left\{\beta_b |b\} + \beta_+ |+1\} + \beta_- |-1\}\right\} \\
= & \ \alpha_u \beta_d |u, b\rangle + \alpha_u \beta_+ |u, +1\rangle + \alpha_u \beta_- |u, -1\rangle + \alpha_d \beta_d |d, b\rangle + \alpha_d \beta_+ |d, +1\rangle + \alpha_d \beta_- |d, -1\rangle
\end{aligned}$$

[54] See for instance: `https://www.researchgate.net/post/What-is-difference-between-mixed-state-and-mixed-entangled-state`

[55] A fancy term you may find here and there: a linear combination of elements, where the scalars factors sums to 1; see `https://en.wikipedia.org/wiki/Convex_combination`

[56] As is common in most Physics-centered introduction to Quantum Mechanics, the tensor product introduction is a bit hand-wavy. For a more rigorous development, see for instance this video by F. Schuller. Some subtleties such as the fact that the equivalence classes respect addition and scalar multiplication have been left as homework; there's a set of notes which contains the "missing" proofs.

[57] p164, section 6.5 - *Product states*

By setting:

$$\beta_d = 1; \beta_+ = \beta_- = 0$$

We found back the state vector from 7.30, retrospectively justifying the notation for α_u and α_d. Because the subsystem states, must be normalized, the resulting combined state is also normalized.

However, and perhaps this is more in line with the author's intent, we've just saw[58] two tests to check whether the state corresponding to a given a wave-function for a composite system is entangled or not.

For the first criteria, we'd need to take any two arbitrary observables from each subsystem, say observable $\boldsymbol{A}$ and $\boldsymbol{B}$, and prove that their correlation $C(\boldsymbol{A}, \boldsymbol{B})$ is zero. But essentially, the proof will end up relying on the tensor product distributivity, rely on the density matrix (see just after), or essentially mimick the proofs of L07E09.

I don't think there's added value to develop it further here.

The second technique is slightly more original: the idea is that, for any product state, the density matrix has exactly one non-zero eigenvalue, and that eigenvalue is exactly 1.

Recall that in L07E04 we've already determined Alice's density matrix:

$$\rho = \begin{pmatrix} \alpha_u^* \alpha_u & \alpha_u^* \alpha_d \\ \alpha_d^* \alpha_u & \alpha_d^* \alpha_d \end{pmatrix}$$

Let's diagonalize it: as usual, we have the eigenvector/eigenvalue relationship:

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

Which implies that $\rho - I_2\lambda$ isn't invertible[59], which translates to its determinant being equal to zero:

$$\begin{aligned} \begin{vmatrix} \alpha_u^* \alpha_u - \lambda & \alpha_u^* \alpha_d \\ \alpha_d^* \alpha_u & \alpha_d^* \alpha_d - \lambda \end{vmatrix} &= \Big((\alpha_u^* \alpha_u - \lambda)(\alpha_d^* \alpha_d - \lambda) \Big) - \Big(\alpha_u^* \alpha_d \alpha_d^* \alpha_u \Big) \\ &= \Big(\alpha_u^* \alpha_u \alpha_d^* \alpha_d - \lambda \underbrace{(\alpha_u^* \alpha_u + \alpha_d^* \alpha_d)}_{=\langle\Psi|\Psi\rangle=1} + \lambda^2 \Big) - \alpha_u^* \alpha_d \alpha_d^* \alpha_u \\ &= \lambda(1 - \lambda) \end{aligned}$$

Clearly, we have one non-zero eigenvalue which is exactly one: the criteria indeed applies, and the state must be non-entangled.

7.9 Entanglement and Locality

7.10 The Quantum Sim: An Introduction to Bell's Theorem

7.11 Entanglement Summary

Exercise 39. *Calculate Alice's density matrix for σ_z for the "near-singlet" state.*

The "near-singlet" state is characterized by the following state-vector:

$$|\Psi\rangle = \sqrt{0.6}|ud\rangle - \sqrt{0.4}|du\rangle$$

[58] p212 and onward, section 7.7 *Tests for Entanglement*

[59] Multiply both side of the equation by the inverse, and use the fact that an eigenvector cannot be the zero vector

Alice's density matrix is defined by its components in Eq. 7.20[60]:

$$\rho_{a'a} = \sum_b \psi^*(a,b)\psi(a',b)$$

Where $\psi(a,b)$ is the wave-function of the composite system, that we can extract from $|\Psi\rangle$:

$$|\Psi\rangle = \psi(u,u)|uu\rangle + \psi(u,d)|ud\rangle + \psi(d,u)|du\rangle + \psi(d,d)|dd\rangle \Rightarrow \begin{cases} \psi(u,u) = \psi(d,d) & = 0 \\ \psi(u,d) & = \sqrt{0.6} \\ \psi(d,u) & = -\sqrt{0.4} \end{cases}$$

Hence:

$$\rho = \begin{pmatrix} \rho_{uu} & \rho_{ud} \\ \rho_{du} & \rho_{dd} \end{pmatrix} = \begin{pmatrix} \psi^*(u,u)\psi(u,u) + \psi^*(u,d)\psi(u,d) & \psi^*(d,u)\psi(u,u) + \psi^*(d,d)\psi(u,d) \\ \psi^*(u,u)\psi(d,u) + \psi^*(u,d)\psi(d,d) & \psi^*(d,u)\psi(d,u) + \psi^*(d,d)\psi(d,d) \end{pmatrix} = \boxed{\begin{pmatrix} 0.6 & 0 \\ 0 & 0.4 \end{pmatrix}} \quad \square$$

Remark 24. *I'm not sure what the authors expect regarding σ_z; we're asked to verify all numerical values in the next exercise, which likely should cover pretty much every intepretation (we'll even have to compute Alice's density matrix again, so as to check $\rho^2/Tr(\rho^2)$).*

Exercise 40. *Verify the numerical values in each rap sheet.*

Remark 25. *This is a long "solution": I'm taking the time to rederive some results that were previously established either in the course, or in earlier exercises, while, clarifying what I think is a important source of confusion: the elements of an Hilbert space (the so-called "state-space") aren't the states of a quantum system. More on that when exploring the singlet state.*

Some numerical results have been automatically computed by a R script, inlined at the end of this exercise.

Let's start by recalling a few things before diving in.

We are, as usual, in the case of two spin systems, one for Alice, S_A, one for Bob, S_B. A composite system S_{AB} is then created from those two via a tensor product.

When relevant, we'll be using the usual ordered bases:

- $\{|u\rangle, |d\rangle\}$ for S_A;
- $\{|u\}, |d\}\}$ for S_B;
- and $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$ for S_{AB}.

Then, let's recall the Pauli matrices corresponding to the observables associated to the spin "components": (from Eqs. 3.20, at the end of section 3.4)

$$\boldsymbol{\tau}_{\boldsymbol{z}}^B = \boldsymbol{\sigma}_{\boldsymbol{z}}^A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \boldsymbol{\tau}_{\boldsymbol{x}}^B = \boldsymbol{\sigma}_{\boldsymbol{x}}^A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \boldsymbol{\tau}_{\boldsymbol{y}}^B = \boldsymbol{\sigma}_{\boldsymbol{y}}^A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Implicitly, the matrices are expressed in the ordered basis of the corresponding sub-system, S_A for $\boldsymbol{\sigma}_{\boldsymbol{i}}^1$ and S_B for $\boldsymbol{\tau}_{\boldsymbol{i}}^B$[61]. We will also need the matrix forms for the lifted operators from the sub-systems to the composite system, where $\boldsymbol{I}^A$ is the identity operator on S_A and $\boldsymbol{I}^B$ the identity operator on S_B, again implicitly relying on S_{AB}'s usual ordered basis.

Remark 26. *We could have used the equivalent tables provided in annexes in the book, or the results of previous exercises.*

[60] p205, section 7.5 *Entanglement for Two Spins*

[61] Strictly speaking, the operators aren't equal; their matrix representation in their respective-basis are. The equals signs are to be understood in this context.

Remark 27. *Note that the final result appears twice below: this isn't an error. The first occurence is a manual computation, while the second has been automatically performed in R.*

$$\boldsymbol{\sigma_z} := \quad \boldsymbol{\sigma_z^A} \otimes \boldsymbol{I^B} \quad = \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\boldsymbol{\tau_z} := \quad \boldsymbol{I^A} \otimes \boldsymbol{\tau_z^B} \quad = \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\boldsymbol{\sigma_x} := \quad \boldsymbol{\sigma_x^A} \otimes \boldsymbol{I^B} \quad = \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\tau_x} := \quad \boldsymbol{I^A} \otimes \boldsymbol{\tau_x^B} \quad = \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{\sigma_y} := \quad \boldsymbol{\sigma_y^A} \otimes \boldsymbol{I^B} \quad = \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \quad \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\tau_y} := \quad \boldsymbol{I^A} \otimes \boldsymbol{\tau_y^B} \quad = \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \quad \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}$$

We'll also need a few more "combined" observables. I'll skip the manual matrix multiplication here: those have been automatically computed by R:

$$\boldsymbol{\sigma_z \tau_z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \qquad \boldsymbol{\tau_z \sigma_z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\tau_x \sigma_x} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}; \qquad \boldsymbol{\tau_y \sigma_y} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

Finally, let's recall one more time the formula for the expectation value $\langle \boldsymbol{L} \rangle$ of an observable $\boldsymbol{L}$ of a system in a state $|\Psi\rangle$:

$$\langle \boldsymbol{L} \rangle = \langle \Psi | \boldsymbol{L} | \Psi \rangle$$

Product state

We're starting from the following state-vector for the composite system, and using various properties of the tensor product of vector states we progressively reach:

$$\begin{aligned} |\Psi\rangle = \quad & \alpha_u \beta_u |uu\rangle + \alpha_u \beta_d |ud\rangle + \alpha_d \beta_u |du\rangle + \alpha_d \beta_d |dd\rangle \\ = \quad & \alpha_u \beta_u |u\rangle \otimes |u\} + \alpha_u \beta_d |u\rangle \otimes |d\} + \alpha_d \beta_u |d\rangle \otimes |u\} + \alpha_d \beta_d |d\rangle \otimes |d\} \\ = \quad & (\alpha_u |u\rangle) \otimes (\beta_u |u\}) + (\alpha_u |u\rangle) \otimes (\beta_d |d\}) + (\alpha_d |d\rangle) \otimes (\beta_u |u\}) + (\alpha_d |d\rangle) \otimes (\beta_d |d\}) \\ = \quad & \alpha_u |u\rangle \otimes (\beta_d |d\} + \beta_u |u\}) + \alpha_d |d\rangle \otimes (\beta_u |u\} + \beta_d |d\}) \\ = \quad & \underbrace{(\alpha_u |u\rangle + \alpha_d |d\rangle)}_{=:|\phi\rangle} \otimes \underbrace{(\beta_d |d\} + \beta_u |u\})}_{=:|\psi\}} \end{aligned}$$

We've verified that this particular composite state is a state product: it can be expressed as the tensor product of two states, $|\phi\rangle \in S_A$ and $|\psi\} \in S_B$.

The normalization condition yields:

$$\begin{aligned}
\||\Psi\rangle\| := \sqrt{\langle\Psi|\Psi\rangle} = 1 \Leftrightarrow \quad & \sqrt{\alpha_u^*\beta_u^*\alpha_u\beta_u + \alpha_u^*\beta_d^*\alpha_u\beta_d + \alpha_d^*\beta_u^*\alpha_d\beta_u + \alpha_d^*\beta_d^*\alpha_d\beta_d} = 1 \\
\Leftrightarrow \quad & \sqrt{\alpha_u^*\alpha_u(\beta_u^*\beta_u + \beta_d^*\beta_d) + \alpha_d^*\alpha_d(\beta_u^*\beta_u + \beta_d^*\beta_d)} = 1 \\
\Leftrightarrow \quad & \sqrt{(\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)(\beta_u^*\beta_u + \beta_d^*\beta_d)} = 1 \\
\Leftrightarrow \quad & (\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)(\beta_u^*\beta_u + \beta_d^*\beta_d) = 1 \qquad (\forall\gamma \in \mathbb{C}),\ \gamma^*\gamma =: |\gamma| > 0 \\
\Leftrightarrow \quad & \begin{cases} \alpha_u^*\alpha_u + \alpha_d^*\alpha_d &= 1 \\ \beta_u^*\beta_u + \beta_d^*\beta_d &= 1 \end{cases} \\
\Leftrightarrow \quad & \begin{cases} \sqrt{\alpha_u^*\alpha_u + \alpha_d^*\alpha_d} &= 1 \\ \sqrt{\beta_u^*\beta_u + \beta_d^*\beta_d} &= 1 \end{cases} \qquad (\forall\gamma \in \mathbb{C}),\ \gamma^*\gamma =: |\gamma| > 0 \\
\Leftrightarrow \quad & \begin{cases} \||\phi\rangle\| &= 1 \\ \||\psi\}\| &= 1 \end{cases}
\end{aligned}$$

We've verified consistency of the normalization condition between the composite system, and the sub-systems: our composite state vector is normalized iff the individual vectors from the sub-systems are normalized.

Moving on to the density matrices, let's do the reasoning for Alice's state only, as the same argument applies to Bob's. Let's first stay in S_A. We know that the subsystem's state is a *pure* state $|\phi\rangle$: it's a convex combination[62] with a single term. The density matrix ρ^A is thus:

$$\rho^A = 1|\phi\rangle\langle\phi| = \begin{pmatrix}\alpha_u \\ \alpha_d\end{pmatrix}\begin{pmatrix}\alpha_u^* & \alpha_d^*\end{pmatrix} = \begin{pmatrix}\alpha_u\alpha_u^* & \alpha_u\alpha_d^* \\ \alpha_d\alpha_u^* & \alpha_d\alpha_d^*\end{pmatrix}$$

It is immediate to check that ρ^A is Hermitian ($\rho^A = (\rho^A)^\dagger$, as $(\alpha_u\alpha_d^*)^* = \alpha_d\alpha_u^*$). Furthermore:

$$\mathrm{Tr}(\rho^A) = \alpha_u\alpha_u^* + \alpha_d\alpha_d^* = 1 \text{ (because of the normalization condition)}$$

We also have:

$$\begin{aligned}
(\rho^A)^2 = \quad & \begin{pmatrix}\alpha_u\alpha_u^* & \alpha_u\alpha_d^* \\ \alpha_d\alpha_u^* & \alpha_d\alpha_d^*\end{pmatrix}\begin{pmatrix}\alpha_u\alpha_u^* & \alpha_u\alpha_d^* \\ \alpha_d\alpha_u^* & \alpha_d\alpha_d^*\end{pmatrix} \\
= \quad & \begin{pmatrix}(\alpha_u\alpha_u^*)(\alpha_u\alpha_u^*) + (\alpha_u\alpha_d^*)(\alpha_d\alpha_u^*) & (\alpha_u\alpha_u^*)(\alpha_u\alpha_d^*) + (\alpha_u\alpha_d^*)(\alpha_d\alpha_d^*) \\ (\alpha_d\alpha_u^*)(\alpha_u\alpha_u^*) + (\alpha_d\alpha_d^*)(\alpha_d\alpha_u^*) & (\alpha_d\alpha_u^*)(\alpha_u\alpha_d^*) + (\alpha_d\alpha_d^*)(\alpha_d\alpha_d^*)\end{pmatrix} \\
= \quad & \begin{pmatrix}(\alpha_u\alpha_u^*)\underbrace{(\alpha_u\alpha_u^* + \alpha_d^*\alpha_d)}_{=1} & (\alpha_u\alpha_d^*)\underbrace{(\alpha_u\alpha_u^* + \alpha_d^*\alpha_d)}_{=1} \\ (\alpha_d\alpha_u^*)\underbrace{(\alpha_u\alpha_u^* + \alpha_d^*\alpha_d)}_{=1} & (\alpha_d\alpha_d^*)\underbrace{(\alpha_u\alpha_u^* + \alpha_d^*\alpha_d)}_{=1}\end{pmatrix} \\
= \quad & \rho^A
\end{aligned}$$

And naturally, $\mathrm{Tr}((\rho^A)^2) = \mathrm{Tr}(\rho) = 1$. Those last two conditions are indeed we expect for a pure state.

Let's move on to diagonalizing ρ^A. As usual, eigenvectors $|\lambda\rangle$ are tied to their corresponding eigenvalues

[62] `https://en.wikipedia.org/wiki/Convex_combination`

λ via:

$$\begin{aligned}
\rho^A|\lambda\rangle = \lambda|\lambda\rangle \Leftrightarrow (\rho^A - \boldsymbol{I}^A\lambda)|\lambda\rangle = 0 \Leftrightarrow \quad & |\rho^A - \boldsymbol{I}^A\lambda| = 0 \\
\Leftrightarrow \quad & \begin{vmatrix} \alpha_u\alpha_u^* - \lambda & \alpha_u\alpha_d^* \\ \alpha_d\alpha_u^* & \alpha_d\alpha_d^* - \lambda \end{vmatrix} \\
\Leftrightarrow \quad & (\alpha_u\alpha_u^* - \lambda)(\alpha_d\alpha_d^* - \lambda) - \alpha_d\alpha_u^*\alpha_u\alpha_d^* = 0 \\
\Leftrightarrow \quad & \lambda^2 - \underbrace{(\alpha_u\alpha_u^* + \alpha_d\alpha_d^*)}_{=1}\lambda = 0 \\
\Leftrightarrow \quad & \lambda(\lambda - 1) = 0 \\
\Leftrightarrow \quad & \begin{cases} \lambda &= 0 \\ \lambda &= 1 \end{cases}
\end{aligned}$$

Let's verify that the eigenvector $|\lambda_1\rangle$ associated to $\lambda = 1$ is indeed the wave-function associated to Alice's sub-system, i.e with components α_u and α_d

$$\rho^A|\lambda_1\rangle = \begin{pmatrix} \alpha_u\alpha_u^* & \alpha_u\alpha_d^* \\ \alpha_d\alpha_u^* & \alpha_d\alpha_d^* \end{pmatrix}\begin{pmatrix} \alpha_u \\ \beta_d \end{pmatrix} = \begin{pmatrix} \alpha_u\alpha_u^*\alpha_u + \alpha_u\alpha_d^*\alpha_d \\ \alpha_d\alpha_u^*\alpha_u + \alpha_d\alpha_d^*\alpha_d \end{pmatrix} = \begin{pmatrix} \alpha_u \underbrace{(\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)}_{=1} \\ \alpha_d \underbrace{(\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)}_{=1} \end{pmatrix} = 1|\lambda_1\rangle$$

Again, by symmetry, we immediately have for S_B:

$$\rho_B = \begin{pmatrix} \beta_u\beta_u^* & \beta_u\beta_d^* \\ \beta_d\beta_u^* & \beta_d\beta_d^* \end{pmatrix}; \qquad \rho_B^2 = \rho_B; \qquad \mathrm{Tr}(\rho_B^2) = \mathrm{Tr}(\rho_B) = 1$$

Same eigenvalues/eigenvectors.

What about S_{AB}? Well, $|\Psi\rangle$ still is a pure state in S_{AB}, meaning, its density matrix again is a convex combination involving a single term; expanding it as a matrix in S_{AB}'s usual ordered basis yields:

$$\begin{aligned}
\rho = 1|\Psi\rangle\langle\Psi| = \quad & \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix}\begin{pmatrix} (\alpha_u\beta_u)^* & (\alpha_u\beta_d)^* & (\alpha_d\beta_u)^* + (\alpha_d\beta_d)^* \end{pmatrix} \\
= \quad & \begin{pmatrix} \alpha_u\beta_u\alpha_u^*\beta_u^* & \alpha_u\beta_u\alpha_u^*\beta_d^* & \alpha_u\beta_u\alpha_d^*\beta_u^* & \alpha_u\beta_u\alpha_d^*\beta_d^* \\ \alpha_u\beta_d\alpha_u^*\beta_u^* & \alpha_u\beta_d\alpha_u^*\beta_d^* & \alpha_u\beta_d\alpha_d^*\beta_u^* & \alpha_u\beta_d\alpha_d^*\beta_d^* \\ \alpha_d\beta_u\alpha_u^*\beta_u^* & \alpha_d\beta_u\alpha_u^*\beta_d^* & \alpha_d\beta_u\alpha_d^*\beta_u^* & \alpha_d\beta_u\alpha_d^*\beta_d^* \\ \alpha_d\beta_d\alpha_u^*\beta_u^* & \alpha_d\beta_d\alpha_u^*\beta_d^* & \alpha_d\beta_d\alpha_d^*\beta_u^* & \alpha_d\beta_d\alpha_d^*\beta_d^* \end{pmatrix}
\end{aligned}$$

Because this is again a pure state, this time in the S_{AB} system, we expect the usual formulas to hold. Let's check them for good measure:

$$\begin{aligned}
\mathrm{Tr}(\rho) = \quad & \alpha_u\beta_u\alpha_u^*\beta_u^* + \alpha_u\beta_d\alpha_u^*\beta_d^* + \alpha_d\beta_u\alpha_d^*\beta_u^* + \alpha_d\beta_d\alpha_d^*\beta_d^* \\
= \quad & \alpha_u\alpha_u^*\underbrace{(\beta_u\beta_u^* + \beta_d\beta_d^*)}_{=1} + \alpha_d\alpha_d^*\underbrace{(\beta_u\beta_u^* + \beta_d\beta_d^*)}_{=1} \\
= \quad & 1
\end{aligned}$$

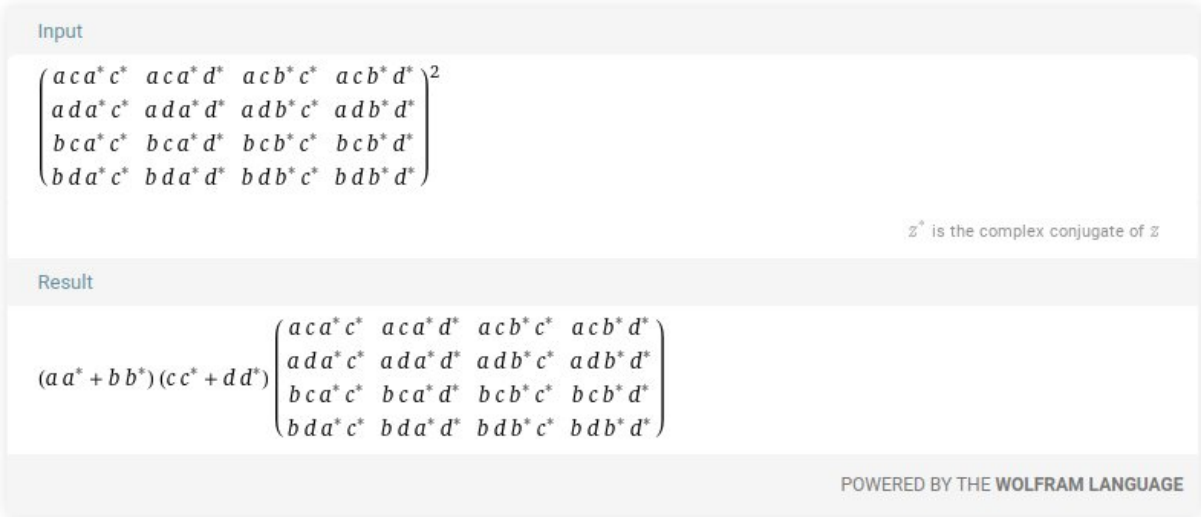
The following is going to be too tedious to do by hand:

$$\begin{aligned}
\rho^2 = \quad & \begin{pmatrix} \alpha_u\beta_u\alpha_u^*\beta_u^* & \alpha_u\beta_u\alpha_u^*\beta_d^* & \alpha_u\beta_u\alpha_d^*\beta_u^* & \alpha_u\beta_u\alpha_d^*\beta_d^* \\ \alpha_u\beta_d\alpha_u^*\beta_u^* & \alpha_u\beta_d\alpha_u^*\beta_d^* & \alpha_u\beta_d\alpha_d^*\beta_u^* & \alpha_u\beta_d\alpha_d^*\beta_d^* \\ \alpha_d\beta_u\alpha_u^*\beta_u^* & \alpha_d\beta_u\alpha_u^*\beta_d^* & \alpha_d\beta_u\alpha_d^*\beta_u^* & \alpha_d\beta_u\alpha_d^*\beta_d^* \\ \alpha_d\beta_d\alpha_u^*\beta_u^* & \alpha_d\beta_d\alpha_u^*\beta_d^* & \alpha_d\beta_d\alpha_d^*\beta_u^* & \alpha_d\beta_d\alpha_d^*\beta_d^* \end{pmatrix}\begin{pmatrix} \alpha_u\beta_u\alpha_u^*\beta_u^* & \alpha_u\beta_u\alpha_u^*\beta_d^* & \alpha_u\beta_u\alpha_d^*\beta_u^* & \alpha_u\beta_u\alpha_d^*\beta_d^* \\ \alpha_u\beta_d\alpha_u^*\beta_u^* & \alpha_u\beta_d\alpha_u^*\beta_d^* & \alpha_u\beta_d\alpha_d^*\beta_u^* & \alpha_u\beta_d\alpha_d^*\beta_d^* \\ \alpha_d\beta_u\alpha_u^*\beta_u^* & \alpha_d\beta_u\alpha_u^*\beta_d^* & \alpha_d\beta_u\alpha_d^*\beta_u^* & \alpha_d\beta_u\alpha_d^*\beta_d^* \\ \alpha_d\beta_d\alpha_u^*\beta_u^* & \alpha_d\beta_d\alpha_u^*\beta_d^* & \alpha_d\beta_d\alpha_d^*\beta_u^* & \alpha_d\beta_d\alpha_d^*\beta_d^* \end{pmatrix} \\
= \quad & \dots \\
= \quad & \rho
\end{aligned}$$

With some text reformatting and a few tweak, we can convert the LaTeXcode in a Wolfram Alpha matrix:

```
{{a c Conjugate(a) Conjugate(c),a c Conjugate(a) Conjugate(d),
    a c Conjugate(b) Conjugate(c),
    a c Conjugate(b) Conjugate(d)},
{a d Conjugate(a) Conjugate(c),
    a d Conjugate(a) Conjugate(d),
    a d Conjugate(b) Conjugate(c),
    a d Conjugate(b) Conjugate(d)},
{b c Conjugate(a) Conjugate(c),
    b c Conjugate(a) Conjugate(d),
    b c Conjugate(b) Conjugate(c),
    b c Conjugate(b) Conjugate(d)},
{b d Conjugate(a) Conjugate(c),
    b d Conjugate(a) Conjugate(d),
    b d Conjugate(b) Conjugate(c),
    b d Conjugate(b) Conjugate(d)}}
```

And use for instance the web interface (see `https://www.wolframalpha.com/input?i2d=true&i=Power%5B%7B%7Ba+c+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Ca+c+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%2Ca+c+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Ca+c+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%7D%2C%7Ba+d+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Ca+d+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%2Ca+d+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Ca+d+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%7D%2C%7Bb+c+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Cb+c+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%2Cb+c+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Cb+c+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%7D%2C%7Bb+d+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Cb+d+Conjugate%5C%2840%29a%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%2Cb+d+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29c%5C%2841%29%2Cb+d+Conjugate%5C%2840%29b%5C%2841%29+Conjugate%5C%2840%29d%5C%2841%29%7D%7D%2C2%5D`) to make sure it works:



Remark 28. *The leading factors correspond to the norms of the sub-system states: it's a* 1 *in disguise.*

It follows that $\mathrm{Tr}(\rho^2) = \mathrm{Tr}(\rho) = 1$.

We've already brushed upon it, but let's make things crystal clear regarding the wave-functions: we have one wave function for each sub-systems:

$$\begin{aligned} \psi^A(|u\rangle) = \quad & \alpha_u; \qquad \psi^B(|u\}) = \beta_u; \\ \psi^A(|d\rangle) = \quad & \alpha_d; \qquad \psi^B(|d\}) = \beta_d. \end{aligned}$$

And a wave-function for the composite system, which indeed factorize as a product of the sub-systems wave-functions:

$$\psi : \begin{cases} |uu\rangle & \to \alpha_u\beta_u = \psi^A(|u\rangle)\psi^B(|u\}) \\ |ud\rangle & \to \alpha_u\beta_d = \psi^A(|u\rangle)\psi^B(|d\}) \\ |du\rangle & \to \alpha_d\beta_u = \psi^A(|d\rangle)\psi^B(|u\}) \\ |dd\rangle & \to \alpha_d\beta_d = \psi^A(|d\rangle)\psi^B(|d\}) \end{cases} \Leftrightarrow \psi(a,b) = \psi^A(a)\psi^B(b)$$

Finally, we can crunch some numbers using the usual expectation value formula, using the matrices for the spin observables we've re-established earlier. This could have been automated, but I haven't looked much into how to perform symbolic computation with R.

$$\begin{aligned}
\langle \boldsymbol{\sigma_x} \rangle = \langle \Psi | \boldsymbol{\sigma_x} | \Psi \rangle = & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} \alpha_d\beta_u \\ \alpha_d\beta_d \\ \alpha_u\beta_u \\ \alpha_u\beta_d \end{pmatrix} \\
= & \alpha_u^*\beta_u^*\alpha_d\beta_u + \alpha_u^*\beta_d^*\alpha_d\beta_d + \alpha_d^*\beta_u^*\alpha_u\beta_u + \alpha_d^*\beta_d^*\alpha_u\beta_d \\
= & \beta_u^*\beta_u(\alpha_u^*\alpha_d + \alpha_d^*\alpha_u) + \beta_d^*\beta_d(\alpha_u^*\alpha_d + \alpha_d^*\alpha_u) \\
= & (\alpha_u^*\alpha_d + \alpha_d^*\alpha_u)\underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1} \\
= & \alpha_u^*\alpha_d + \alpha_d^*\alpha_u \\
\Rightarrow \langle \boldsymbol{\sigma_x} \rangle^2 = & (\alpha_u^*\alpha_d + \alpha_d^*\alpha_u)^2 \\
= & (\alpha_u^*\alpha_d)^2 + 2\alpha_u^*\alpha_d\alpha_d^*\alpha_u + (\alpha_d^*\alpha_u)^2
\end{aligned}$$

$$\begin{aligned}
\langle \boldsymbol{\sigma_y} \rangle = \langle \Psi | \boldsymbol{\sigma_y} | \Psi \rangle = & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & i \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} -\alpha_d\beta_u \\ -\alpha_d\beta_d \\ \alpha_u\beta_u \\ \alpha_u\beta_d \end{pmatrix} \\
= & i(-\alpha_u^*\beta_u^*\alpha_d\beta_u - \alpha_u^*\beta_d^*\alpha_d\beta_d + \alpha_d^*\beta_u^*\alpha_u\beta_u + \alpha_d^*\beta_d^*\alpha_u\beta_d) \\
= & i(-\alpha_u^*\alpha_d + \alpha_d^*\alpha_u)\underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1} \\
= & i(-\alpha_u^*\alpha_d + \alpha_d^*\alpha_u) \\
\Rightarrow \langle \boldsymbol{\sigma_y} \rangle^2 = & (i(-\alpha_u^*\alpha_d + \alpha_d^*\alpha_u))^2 \\
= & -(\alpha_d^*\alpha_u)^2 + 2\alpha_d^*\alpha_u\alpha_u^*\alpha_d - (\alpha_u^*\alpha_d)^2
\end{aligned}$$

$$
\begin{aligned}
\langle \boldsymbol{\sigma_z} \rangle = \langle \Psi | \boldsymbol{\sigma_z} | \Psi \rangle = & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ -\alpha_d\beta_u \\ -\alpha_d\beta_d \end{pmatrix} \\
= & \alpha_u^*\beta_u^*\alpha_u\beta_u + \alpha_u^*\beta_d^*\alpha_u\beta_d - \alpha_d^*\beta_u^*\alpha_d\beta_u - \alpha_d^*\beta_d^*\alpha_d\beta_d \\
= & \beta_u^*\beta_u(\alpha_u^*\alpha_u - \alpha_d^*\alpha_d) + \beta_d^*\beta_d(\alpha_u^*\alpha_u - \alpha_d^*\alpha_d) \\
= & (\alpha_u^*\alpha_u - \alpha_d^*\alpha_d)\underbrace{(\beta_u^*\beta_u + \beta_d^*\beta_d)}_{=1} \\
= & \alpha_u^*\alpha_u - \alpha_d^*\alpha_d \\
\Rightarrow \langle \boldsymbol{\sigma_z} \rangle^2 = & (\alpha_u^*\alpha_u - \alpha_d^*\alpha_d)^2 \\
= & (\alpha_u^*\alpha_u)^2 - 2\alpha_u^*\alpha_u\alpha_d^*\alpha_d + (\alpha_d^*\alpha_d)^2
\end{aligned}
$$

We can now compute:

$$
\begin{aligned}
\langle \boldsymbol{\sigma_x} \rangle^2 + \langle \boldsymbol{\sigma_y} \rangle^2 + \langle \boldsymbol{\sigma_z} \rangle^2 = & (\alpha_u^*\alpha_d)^2 + 2\alpha_u^*\alpha_d\alpha_d^*\alpha_u + (\alpha_d^*\alpha_u)^2 \\
& - (\alpha_d^*\alpha_u)^2 + 2\alpha_d^*\alpha_u\alpha_u^*\alpha_d - (\alpha_u^*\alpha_d)^2 \\
& + (\alpha_u^*\alpha_u)^2 - 2\alpha_u^*\alpha_u\alpha_d^*\alpha_d + (\alpha_d^*\alpha_d)^2 \\
= & (\alpha_u^*\alpha_u)^2 + 2\alpha_u^*\alpha_u\alpha_d^*\alpha_d + (\alpha_d^*\alpha_d)^2 \\
= & (\underbrace{\alpha_u^*\alpha_u + \alpha_d^*\alpha_d}_{=1})^2 \\
= & \boxed{1}
\end{aligned}
$$

Moving on to the other spin "components":

$$
\begin{aligned}
\langle \boldsymbol{\tau_x} \rangle = \langle \Psi | \boldsymbol{\tau_x} | \Psi \rangle = & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} \alpha_u\beta_d \\ \alpha_u\beta_u \\ \alpha_d\beta_d \\ \alpha_d\beta_u \end{pmatrix} \\
= & \alpha_u^*\beta_u^*\alpha_u\beta_d + \alpha_u^*\beta_d^*\alpha_u\beta_u + \alpha_d^*\beta_u^*\alpha_d\beta_d + \alpha_d^*\beta_d^*\alpha_d\beta_u \\
= & \alpha_u^*\alpha_u(\beta_u^*\beta_d + \beta_d^*\beta_u) + \alpha_d^*\alpha_d(\beta_u^*\beta_d + \beta_d^*\beta_u) \\
= & (\beta_u^*\beta_d + \beta_d^*\beta_u)\underbrace{(\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)}_{=1} \\
= & \beta_u^*\beta_d + \beta_d^*\beta_u
\end{aligned}
$$

Remark 29. *At this stage, it's important to observe than up to a renaming ($\beta \leftarrow \alpha$), this is the same expression we had for $\langle \boldsymbol{\sigma_x} \rangle$. And it makes sense given how symmetrical the "physical" situation is. We expect to find the same thing for the two other components: if this is the case, we could then directly conclude $\langle \boldsymbol{\tau_x} \rangle^2 + \langle \boldsymbol{\tau_y} \rangle^2 + \langle \boldsymbol{\tau_z} \rangle^2 = 1$, without additional computations.*

$$\begin{aligned}
\langle \boldsymbol{\tau_y} \rangle = \langle \Psi | \boldsymbol{\tau_y} | \Psi \rangle = & \quad \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \quad i \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} -\alpha_u\beta_d \\ \alpha_u\beta_u \\ -\alpha_d\beta_d \\ \alpha_d\beta_u \end{pmatrix} \\
= & \quad i(-\alpha_u^*\beta_u^*\alpha_u\beta_d + \alpha_u^*\beta_d^*\alpha_u\beta_u - \alpha_d^*\beta_u^*\alpha_d\beta_d + \alpha_d^*\beta_d^*\alpha_d\beta_u) \\
= & \quad i(\alpha_u^*\alpha_u(-\beta_u^*\beta_d + \beta_d^*\beta_u) + \alpha_d^*\alpha_d(-\beta_u^*\beta_d + \beta_d^*\beta_u)) \\
= & \quad i(-\beta_u^*\beta_d + \beta_d^*\beta_u)\underbrace{(\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)}_{=1} \\
= & \quad i(-\beta_u^*\beta_d + \beta_d^*\beta_u) \\
=_{\beta\leftarrow\alpha} & \quad \langle \boldsymbol{\sigma_y} \rangle
\end{aligned}$$

$$\begin{aligned}
\langle \boldsymbol{\tau_z} \rangle = \langle \Psi | \boldsymbol{\tau_z} | \Psi \rangle = & \quad \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \quad \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ -\alpha_u\beta_d \\ \alpha_d\beta_u \\ -\alpha_d\beta_d \end{pmatrix} \\
= & \quad \alpha_u^*\beta_u^*\alpha_u\beta_u - \alpha_u^*\beta_d^*\alpha_u\beta_d\alpha_d^*\beta_u^*\alpha_d\beta_u - \alpha_d^*\beta_d^*\alpha_d\beta_d \\
= & \quad \alpha_u^*\alpha_u(\beta_u^*\beta_u - \beta_d^*\beta_d) + \alpha_d^*\alpha_d(\beta_u^*\beta_u - \beta_d^*\beta_d) \\
= & \quad (\beta_u^*\beta_u - \beta_d^*\beta_d)\underbrace{(\alpha_u^*\alpha_u + \alpha_d^*\alpha_d)}_{=1} \\
= & \quad \beta_u^*\beta_u - \beta_d^*\beta_d \\
=_{\beta\leftarrow\alpha} & \quad \langle \boldsymbol{\sigma_z} \rangle
\end{aligned}$$

Hence by our previous remark, indeed:

$$\boxed{\langle \boldsymbol{\tau_x} \rangle^2 + \langle \boldsymbol{\tau_y} \rangle^2 + \langle \boldsymbol{\tau_z} \rangle^2 =_{\beta\leftarrow\alpha} \langle \boldsymbol{\sigma_x} \rangle^2 + \langle \boldsymbol{\sigma_y} \rangle^2 + \langle \boldsymbol{\sigma_z} \rangle^2 = 1}$$

Moving on to the correlation:

$$\begin{aligned}
\langle \boldsymbol{\sigma_z\tau_z} \rangle = \langle \Psi | \boldsymbol{\sigma_z\tau_z} | \Psi \rangle = & \quad \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ \alpha_u\beta_d \\ \alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \quad \begin{pmatrix} \alpha_u^*\beta_u^* & \alpha_u^*\beta_d^* & \alpha_d^*\beta_u^* & \alpha_d^*\beta_d^* \end{pmatrix} \begin{pmatrix} \alpha_u\beta_u \\ -\alpha_u\beta_d \\ -\alpha_d\beta_u \\ \alpha_d\beta_d \end{pmatrix} \\
= & \quad \alpha_u^*\beta_u^*\alpha_u\beta_u - \alpha_u^*\beta_d^*\alpha_u\beta_d - \alpha_d^*\beta_u^*\alpha_d\beta_u + \alpha_d^*\beta_d^*\alpha_d\beta_d \\
= & \quad \alpha_u^*\alpha_u(\beta_u^*\beta_u - \beta_d^*\beta_d) - \alpha_d^*\alpha_d(\beta_u^*\beta_u - \beta_d^*\beta_d) \\
= & \quad \underbrace{(\alpha_u^*\alpha_u - \alpha_d^*\alpha_d)}_{=\langle \boldsymbol{\sigma_z} \rangle}\underbrace{(\beta_u^*\beta_u - \beta_d^*\beta_d)}_{=\langle \boldsymbol{\tau_z} \rangle} \\
\Leftrightarrow & \quad \boxed{\langle \boldsymbol{\sigma_z\tau_z} \rangle - \langle \boldsymbol{\sigma_z} \rangle \langle \boldsymbol{\tau_z} \rangle = 0}
\end{aligned}$$

Singlet state

The singlet state is characteristic of a maximally entangled state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|ud\rangle - |du\rangle\right)$$

This means that we *won't* be able to express this state as a tensor product of two states from S_A and S_B, as we just did for a product state.

Let's start with the wave function and normalization, for the composite space: the general form of a state vector in this space is:

$$|\Psi\rangle = \psi_{uu}|uu\rangle + \psi_{ud}|ud\rangle + \psi_{du}|du\rangle + \psi_{dd}|dd\rangle$$

While the normalization condition translates to:

$$\||\Psi\rangle\| = 1 \Leftrightarrow \sqrt{\langle\Psi|\Psi\rangle} = \sqrt{\psi^*_{uu}\psi_{uu} + \psi^*_{ud}\psi_{ud} + \psi^*_{du}\psi_{du} + \psi^*_{dd}\psi_{dd}} = 1$$

But because each individual term under the square root is positive, this is equivalent to:

$$\psi^*_{uu}\psi_{uu} + \psi^*_{ud}\psi_{ud} + \psi^*_{du}\psi_{du} + \psi^*_{dd}\psi_{dd} = 1$$

For the singlet state, the wave function is:

$$\psi : \begin{cases} |uu\rangle \rightarrow & \psi_{uu} = 0 \\ |ud\rangle \rightarrow & \psi_{ud} = 1/\sqrt{2} \\ |du\rangle \rightarrow & \psi_{du} = -1/\sqrt{2} \\ |dd\rangle \rightarrow & \psi_{dd} = 0 \end{cases}$$

It's trivial to check that it's normalized.

What's the wave function for each subsystem state? Well, think about it: if there's a wave function for each subsystem, then there's a *pure*, normalized state for each subsystem, and then the composite state can be expressed as a tensor product between those two. Meaning, this composite state *would be* a product state. But, it's claimed here than the composite state is entangled, meaning, it's *not* a product state, and so we shouldn't be able to find such wave-functions for the isolated subsystems.

We've already studied in L06E03 why this particular singlet state isn't a product state, let me recall you how it went: the idea is to identify the general form of a composite state:

$$|\Psi\rangle = \psi_{uu}|uu\rangle + \psi_{ud}|ud\rangle + \psi_{du}|du\rangle + \psi_{dd}|dd\rangle$$

With the general form of a product state:

$$|\Phi\rangle = \alpha_u\beta_u|uu\rangle + \alpha_u\beta_d|ud\rangle + \alpha_d\beta_u|du\rangle + \alpha_d\beta_d|dd\rangle$$

Which yields the particular following equations systems for the singlet state:

$$\begin{cases} \psi_{ud} = \alpha_u\beta_d = & \frac{1}{\sqrt{2}} \\ \psi_{du} = \alpha_d\beta_u = & -\frac{1}{\sqrt{2}} \\ \psi_{uu} = \alpha_u\beta_u = & 0 \\ \psi_{dd} = \alpha_d\beta_d = & 0 \end{cases}$$

But consider for example the third equation, which implies that at least either $\alpha_u = 0$ or $\beta_u = 0$. In the former case, the first equation can't hold, while in the latter, the second equation can't hold. Hence the system is inconsistent, and cannot be solved.

This proves that the composite system's wave-function cannot be factorized.

So what does this mean regarding the states of the subsystems? Surely it conceptually makes sense to still talk about the existence of such states? Yes, obviously is does and that's precisely where the notion

of density matrix becomes most useful[63]: to express *impure* states.

So, let's move on to density matrices then. Starting with the easiest: the composite system's: it's a pure state so we have (again, the vector/matrix representation depends implicitly on the usual ordered basis):

$$\rho = |\Psi\rangle\langle\Psi| = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Let's verify the usual matrix properties for pure states:

$$\rho^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/4+1/4 & -1/4-1/4 & 0 \\ 0 & -1/4-1/4 & 1/4+1/4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \rho$$

And trivially, $\mathrm{Tr}(\rho^2) = \mathrm{Tr}(\rho) = 1/2 + 1/2 = 1$.

Moving on to the density matrices of the subsystems, that is, on the most accurate state description we can provide to each subsystems. In the book, we derived a formula[64]: we first introduced an arbitrary observable $\boldsymbol{L}^A$, acting on S_A, and upgraded it to the composite system: $\boldsymbol{L} = \boldsymbol{L}^A \otimes \boldsymbol{I}^B$.

Let me rework the proof, while being a bit more explicit. First, component-wise, we have:

$$\boldsymbol{L} = \boldsymbol{L}^A \otimes \boldsymbol{I}^B = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & L_{12} & 0 \\ 0 & L_{11} & 0 & L_{12} \\ L_{21} & 0 & L_{22} & 0 \\ 0 & L_{21} & 0 & L_{22} \end{pmatrix} \Leftrightarrow L_{a'b',ab} = L^A_{a'a}\delta_{b'b}$$

Now assume we're in a composite state:

$$|\Psi\rangle = \begin{pmatrix} \psi_{11} \\ \psi_{12} \\ \psi_{21} \\ \psi_{22} \end{pmatrix}; \text{ thus: } \langle\Psi| = \begin{pmatrix} \psi^*_{11} & \psi^*_{12} & \psi^*_{21} & \psi^*_{22} \end{pmatrix}$$

We can then compute the expectation value for $\boldsymbol{L}$:

$$\begin{aligned}
\langle \boldsymbol{L} \rangle = \quad & \langle\Psi|\boldsymbol{L}|\Psi\rangle && = \sum_{a,b,a',b'} \psi^*_{a'b'} L_{a'b',ab} \psi_{ab} \\
= \quad & \sum_{a',b,a} \psi^*_{a'b'} L^A_{a',a} \psi_{ab} && = \sum_{a',a} \underbrace{\left(\sum_b \psi^*_{a'b'} \psi_{ab} \right)}_{\rho_{a'a}} L^A_{a',a} \\
= \quad & \sum_{a'} \left(\sum_a \rho_{a',a} L^A_{a',a} \right) && = \sum_{a'} \left(\sum_a \langle a|\rho^A|a'\rangle\langle a'|\boldsymbol{L}^A|a\rangle \right) \\
= \quad & \sum_{a'} \left(\sum_a \langle a'|\boldsymbol{L}^A|a\rangle\langle a|\rho^A|a'\rangle \right) && = \sum_{a'} \left(\langle a'|\boldsymbol{L}^A \underbrace{\left(\sum_a |a\rangle\langle a| \right)}_{=\boldsymbol{I}^A} \rho^A|a'\rangle \right) \\
= \quad & \sum_{a'} \left(\langle a'|\boldsymbol{L}^A\rho^A|a'\rangle \right) && =: \ \mathrm{Tr}(\boldsymbol{L}^A\rho^A) = \mathrm{Tr}(\rho^A\boldsymbol{L}^A)
\end{aligned}$$

Where the last equality is an usual property of the trace operator.

[63] For pure states, the density matrix ρ is isomorphic to the state $|\Psi\rangle$: $\rho = |\Psi\rangle\langle\Psi|$

[64] p204+, section 7.5 *Entanglement for two spins*

And we've already demonstrated in the book, using similar arguments, that

$$\langle \boldsymbol{L} \rangle = \langle \Psi | \boldsymbol{L} | \Psi \rangle = \mathrm{Tr}(\underbrace{|\Psi\rangle \langle \Psi|}_{=:\rho} \boldsymbol{L}) = \mathrm{Tr}(\rho \boldsymbol{L})$$

So what we've proved in the end is that:

$$\langle \boldsymbol{L} \rangle = \mathrm{Tr}(\rho \boldsymbol{L}) = \mathrm{Tr}(\rho^A \boldsymbol{L}^A) = \langle \boldsymbol{L}^A \rangle$$

That is, we've reduced the density matrix ρ on the composite system to a density matrix ρ^A on Alice's subsystem.

I'll come back to this density matrix ρ^A in a moment, but before moving on any further, let me emphasize a subtle point that I think could have been made clearer in the book. Suppose we're in a mixed state in Alice's subsystem. This means that there's some amount of chance we're in this state, or some amount we're in this other state, and so on, something like:

$$|\psi\rangle = P_1 |\psi_1\rangle + P_2 |\psi_2\rangle + \dots$$

But wait a minute, each of those $|\psi_i\rangle$ is an element of the state space (the Hilbert space), so they can all be expressed as a linear combination of its basis vectors. In the context of a spin:

$$\begin{aligned} |\psi\rangle = & \quad P_1(\alpha_1 |u\rangle + \beta_1 |d\rangle) + P_2(\alpha_2 |u\rangle + \beta_2 |d\rangle) + \dots \\ = & \quad \left(\sum_i P_i \alpha_i \right) |u\rangle + \left(\sum_i P_i \beta_i \right) |d\rangle \end{aligned}$$

Assuming we renormalize that last state vector if need be, haven't we just found a wave-function describing Alice's state? But haven't we just stated that we cannot find a wave-function for Alice's state because it's a mixed state?

You could push this thinking one step further: can't we do the same thing for Bob's space, and join the two resulting states with a tensor product?

Well, there's one considerable issue with the previous reasoning, and I don't think it's clear from the book. So let me emphasize it:

Elements of the so called state-space **aren't** states!.

Meaning, $|\psi\rangle$ **isn't** a state! What we have is the following:

The set of all *pure states* is isomorphic to $S_A = \{|\psi\rangle\}$.

The most careful definition of quantum system states is[65]:

The states of a (quantum) system are all positive, trace-class, linear maps $\rho : S_A \to S_A$ for which $\mathrm{Tr}\rho = 1$

The previous definition accounts for some refinements that will be introduced in the next chapter of Susskind's book. For example, a *trace-class* map refers to a map who has a finite trace: it's always the case in a finite dimension vector space, but divergence may occur in infinite dimension vector spaces, which are mandatory to express position observables for example.

To simplify, such maps corresponds to our density matrices, which as we've saw, can encode both pure and mixed states. We could argue on terminology regarding what a state is: a "mixed state" may only corresponds to the information we have about a state, and not to an actual, physical state, which

[65] See `https://youtu.be/GbqA9Xn_iM0?t=4453`

would justify the less strict terminology, while introducing some confusion on the use of the term "state".

It just so happen than for every pure state, there's a $1:1$ correspondance with the elements of the Hilbert "state" space, as $\rho = |\psi\rangle\langle\psi|$.

So we can't just create a convex combination of "pure states" (well, something that's isomorphic to a pure state in our modern terminology). That's why when density matrices were introduced in Susskind's book, the convex combination was performed over projection operators built from pure states:

$$\rho = P_1 |\psi_1\rangle\langle\psi_1| + P_2 |\psi_2\rangle\langle\psi_2| + \dots$$

And not as I've just show you, directly over elements of the Hilbert space.

Now that we have a clear definition of what a state is, we can check that our matrix ρ^A really is a state: we want to prove that it's a positive, trace-class linear map such that $\mathrm{Tr}(\rho^A) = 1$.

It's clearly **trace-class**, because we're in finite dimension: we have a matrix, the trace is a finite sum, it always converges.

Let me write the matrix in component form:

$$\rho^A = \begin{pmatrix} \sum_b \psi^*_{1b}\psi_{1b} & \sum_b \psi^*_{1b}\psi_{2b} \\ \sum_b \psi^*_{2b}\psi_{1b} & \sum_b \psi^*_{2b}\psi_{2b} \end{pmatrix} = \begin{pmatrix} \psi^*_{11}\psi_{11} + \psi^*_{12}\psi_{12} & \psi^*_{11}\psi_{21} + \psi^*_{12}\psi_{22} \\ \psi^*_{21}\psi_{11} + \psi^*_{22}\psi_{12} & \psi^*_{21}\psi_{21} + \psi^*_{22}\psi_{22} \end{pmatrix}$$

Let's compute its trace:

$$\mathrm{Tr}(\rho^A) = \psi^*_{11}\psi_{11} + \psi^*_{12}\psi_{12} + \psi^*_{21}\psi_{21} + \psi^*_{22}\psi_{22} = \langle\Psi|\Psi\rangle = \boxed{1} \text{ (as } |\Psi\rangle \text{ is normalized)}$$

Lastly, as every component of the matrix is a positive real number, I guess this is enough to prove that ρ^A is **positive**.

We're ready to move on to verify that $(\rho^A)^2 \neq \rho^A$. After a quick check, I don't think we reach anything conclusive by carrying the computation symbolically, so let's do it numerically:

$$(\rho^A)^2 = \begin{pmatrix} 0\times 0 + 1/\sqrt{2}\times 1/\sqrt{2} & 0\times(-1\sqrt{2}) + 1/\sqrt{2}\times 0 \\ (-1/\sqrt{2})\times 0 + 0\times(1/\sqrt{2}) & -1/\sqrt{2}\times 1/\sqrt{2} + 0\times 0 \end{pmatrix}^2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}^2 = \frac{1}{4}\boldsymbol{I}^A \neq \rho^A$$

Clearly, $\mathrm{Tr}((\rho^A)^2) = 2(1/4) = 1/2 < 1$, as expected.

Finally, let's crunch some numbers, using our R script:

$$\langle\boldsymbol{\sigma_z}\rangle = \langle\Psi|\boldsymbol{\sigma_z}|\Psi\rangle = 0; \qquad \langle\boldsymbol{\sigma_x}\rangle = \langle\Psi|\boldsymbol{\sigma_x}|\Psi\rangle = 0; \qquad \langle\boldsymbol{\sigma_y}\rangle = \langle\Psi|\boldsymbol{\sigma_y}|\Psi\rangle = 0$$

$$\langle\boldsymbol{\tau_z}\rangle = \langle\Psi|\boldsymbol{\tau_z}|\Psi\rangle = 0; \qquad \langle\boldsymbol{\tau_x}\rangle = \langle\Psi|\boldsymbol{\tau_x}|\Psi\rangle = 0; \qquad \langle\boldsymbol{\tau_y}\rangle = \langle\Psi|\boldsymbol{\tau_y}|\Psi\rangle = 0$$

$$\langle \boldsymbol{\tau_z}\boldsymbol{\sigma_z}\rangle = \langle\Psi|\boldsymbol{\tau_z}\boldsymbol{\sigma_z}|\Psi\rangle = \begin{pmatrix}0 & 1/\sqrt{2} & -1/\sqrt{2} & 0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\\0\end{pmatrix} = -1$$

$$\langle \boldsymbol{\tau_x}\boldsymbol{\sigma_x}\rangle = \langle\Psi|\boldsymbol{\tau_x}\boldsymbol{\sigma_x}|\Psi\rangle = \begin{pmatrix}0 & 1/\sqrt{2} & -1/\sqrt{2} & 0\end{pmatrix}\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}\begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\\0\end{pmatrix} = -1$$

$$\langle \boldsymbol{\tau_y}\boldsymbol{\sigma_y}\rangle = \langle\Psi|\boldsymbol{\tau_y}\boldsymbol{\sigma_y}|\Psi\rangle = \begin{pmatrix}0 & 1/\sqrt{2} & -1/\sqrt{2} & 0\end{pmatrix}\begin{pmatrix}0&0&0&-1\\0&0&1&0\\0&1&0&0\\-1&0&0&0\end{pmatrix}\begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\\0\end{pmatrix} = -1$$

$$\langle \boldsymbol{\sigma_z}\boldsymbol{\tau_z}\rangle = \langle\Psi|\boldsymbol{\sigma_z}\boldsymbol{\tau_z}|\Psi\rangle = \begin{pmatrix}0 & 1/\sqrt{2} & -1/\sqrt{2} & 0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\\0\end{pmatrix} = -1$$

The last one served to verify the following correlation:

$$\langle \boldsymbol{\sigma_z}\boldsymbol{\tau_z}\rangle - \langle \boldsymbol{\sigma_z}\rangle\langle \boldsymbol{\tau_z}\rangle = -1$$

"Near-singlet" state
Starting from the following composite state:

$$|\Psi\rangle = \sqrt{0.6}|ud\rangle - \sqrt{0.4}|du\rangle$$

We can easily identify the wave-function for the composite system:

$$\psi : \begin{cases}|uu\rangle & \to \psi_{11} = 0\\ |ud\rangle & \to \psi_{12} = \sqrt{0.6}\\ |du\rangle & \to \psi_{21} = -\sqrt{0.4}\\ |dd\rangle & \to \psi_{22} = 0\end{cases}$$

The density matrix for the composite system naturally follows for the definition of the state:

$$\rho = |\Psi\rangle\langle\Psi| = \begin{pmatrix}0\\\sqrt{0.6}\\-\sqrt{0.4}\\0\end{pmatrix}\begin{pmatrix}0 & \sqrt{0.6} & -\sqrt{0.4} & 0\end{pmatrix}$$

$$= \begin{pmatrix}0&0&0&0\\0&0.6&-\sqrt{0.6\times0.4}&0\\0&-\sqrt{0.6\times0.4}&0.4&0\\0&0&0&0\end{pmatrix} = \begin{pmatrix}0&0&0&0\\0&0.6&-\sqrt{0.24}&0\\0&-\sqrt{0.24}&0.4&0\\0&0&0&0\end{pmatrix}$$

Let's square it:

$$\rho^2 = \begin{pmatrix}0&0&0&0\\0&0.6\times0.6+(-\sqrt{0.24})^2&-0.6\sqrt{0.24}-0.4\sqrt{0.24}&0\\0&-0.6\sqrt{0.24}-0.4\sqrt{0.24}&0.4\times0.4+(-\sqrt{0.24})^2&0\\0&0&0&0\end{pmatrix} = \begin{pmatrix}0&0&0&0\\0&0.36+0.24&-\sqrt{0.24}&0\\0&-\sqrt{0.24}&0.16+0.24&0\\0&0&0&0\end{pmatrix} = \rho$$

Immediately, $\mathrm{Tr}(\rho^2) = \mathrm{Tr}(\rho) = 1$.

Again, this is an entangled state: by the same (abstract) reasoning as before, we can't find a wave-function for the subsystems, and we must look for a density matrix. I'll skip the details this time:

$$\rho^A = \begin{pmatrix}\psi_{11}^*\psi_{11}+\psi_{12}^*\psi_{12} & \psi_{11}^*\psi_{21}+\psi_{12}^*\psi_{22}\\ \psi_{21}^*\psi_{11}+\psi_{22}^*\psi_{12} & \psi_{21}^*\psi_{21}+\psi_{22}^*\psi_{22}\end{pmatrix} = \begin{pmatrix}0.6&0\\0&0.4\end{pmatrix}$$

Let's square it:

$$(\rho^A)^2 = \begin{pmatrix} 0.36 & 0 \\ 0 & 0.16 \end{pmatrix} \neq \rho^A$$

Clearly, $\mathrm{Tr}((\rho^A)^2) = 0.36 + 0.16 = 0.52 < 1$.

Let's crunch some numbers again[66]; again, this has been automatically computed by the *R* script:

$$\langle \boldsymbol{\sigma_z} \rangle = \langle \Psi | \boldsymbol{\sigma_z} | \Psi \rangle = 0.2; \qquad \langle \boldsymbol{\sigma_x} \rangle = \langle \Psi | \boldsymbol{\sigma_x} | \Psi \rangle = 0; \qquad \langle \boldsymbol{\sigma_y} \rangle = \langle \Psi | \boldsymbol{\sigma_y} | \Psi \rangle = 0$$

$$\langle \boldsymbol{\tau_z} \rangle = \langle \Psi | \boldsymbol{\tau_z} | \Psi \rangle = -0.2; \qquad \langle \boldsymbol{\tau_x} \rangle = \langle \Psi | \boldsymbol{\tau_x} | \Psi \rangle = 0; \qquad \langle \boldsymbol{\tau_y} \rangle = \langle \Psi | \boldsymbol{\tau_y} | \Psi \rangle = 0$$

$$\langle \boldsymbol{\tau_z}\boldsymbol{\sigma_z} \rangle = \langle \Psi | \boldsymbol{\tau_z}\boldsymbol{\sigma_z} | \Psi \rangle = \begin{pmatrix} 0 & \sqrt{0.6} & -\sqrt{0.4} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \sqrt{0.6} \\ -\sqrt{0.4} \\ 0 \end{pmatrix} = -1$$

$$\langle \boldsymbol{\tau_x}\boldsymbol{\sigma_x} \rangle = \langle \Psi | \boldsymbol{\tau_x}\boldsymbol{\sigma_x} | \Psi \rangle = \begin{pmatrix} 0 & \sqrt{0.6} & -\sqrt{0.4} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \sqrt{0.6} \\ -\sqrt{0.4} \\ 0 \end{pmatrix} \simeq -0.9797959 \simeq -2\sqrt{0.24}$$

$$\langle \boldsymbol{\tau_y}\boldsymbol{\sigma_y} \rangle = \langle \Psi | \boldsymbol{\tau_y}\boldsymbol{\sigma_y} | \Psi \rangle = \begin{pmatrix} 0 & \sqrt{0.6} & -\sqrt{0.4} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \sqrt{0.6} \\ -\sqrt{0.4} \\ 0 \end{pmatrix} = -1$$

$$\langle \boldsymbol{\sigma_z}\boldsymbol{\tau_z} \rangle = \langle \Psi | \boldsymbol{\sigma_z}\boldsymbol{\tau_z} | \Psi \rangle = \begin{pmatrix} 0 & \sqrt{0.6} & -\sqrt{0.4} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \sqrt{0.6} \\ -\sqrt{0.4} \\ 0 \end{pmatrix} = -1$$

The last one served to verify the following correlation:

$$\langle \boldsymbol{\sigma_z}\boldsymbol{\tau_z} \rangle - \langle \boldsymbol{\sigma_z} \rangle \langle \boldsymbol{\tau_z} \rangle = -1 - 0.2 \times (-0.2) = -0.96$$

For completeness, here's the aforementioned, self-contained *R* script. You may want to look at the LaTeXsource file. I've wrote a separate article (.html) showcasing various R features used in this script.

```
#!/bin/Rscript

# Quirky, but does the job.

# Computes expectation values for a composite
# system built from two quantum spins.
#
# Used with either with 1 or 5 arguments:
#   <operator> [wave-function (uu, ud, du, dd)]
#
# Each argument is parsed and evaluated, so you can
# use "tau_x %*% sigma_x" as an operator for example.
#
# If only an operator is provided (1 arg), its 4x4 matrix form
# is displayed as LaTeX on stdout.
#
# Otherwise, evaluates the expectation value for the operator
# for a system in a state described by the wave-function.
```

[66] There are a few more than asked

```
tmp <- "/tmp/R/lib"
.libPaths(tmp)
dir.create(tmp, recursive = TRUE, mode = "0755")

loadpkg <- function(p) {
    if (!require(p, character.only = TRUE)) {
        install.packages(p, repos='http://cran.us.r-project.org', lib=tmp)
        library(p, character.only = TRUE)
    }
}

# 2x2 identity and Pauli matrices
id2 = matrix(c(1,  0,  0,  1), 2, 2)
pz  = matrix(c(1,  0,  0, -1), 2, 2)
px  = matrix(c(0,  1,  1,  0), 2, 2)
py  = matrix(c(0, 1i, -1i, 0), 2, 2)

# Ugprade subsystem spin component operators to the composite system
sigma_z = kronecker(pz, id2)
tau_z   = kronecker(id2, pz)

sigma_x = kronecker(px, id2)
tau_x   = kronecker(id2, px)

sigma_y = kronecker(py, id2)
tau_y   = kronecker(id2, py)

# Expectation value computation
avg <- function(L, psi) {
    return((Conj(t(psi)) %*% L %*% psi)[1])
}

# Evaluate CLI arguments (e.g. "interpret" "tau_z %*% sigma_z"
# as the corresponding observable)
#
# Note that we need a list (lapply), as some of the arguments will be
# vectors already.
args <- lapply(
    commandArgs(trailingOnly = TRUE),
    function(x) { return(eval(parse(text=x))) }
)

# LaTeX export
loadpkg("xtable")

# See https://tales.mbivert.com/on-exporting-r-complex-matrix-latex/
xtable <- function(x, ...) {
    if (class(x[[1]]) == "complex") {
        z <- sapply(x, function(y) {
            if (y == 0)      return(as.character(0))
            if (Im(y) == 0) return(as.character(Re(y)))

            t <- ""
            if (Re(y) != 0) t <- as.character(Re(y))

            if (Im(y) == 1) {
                if (Re(y) == 0) t <- "i"
```

```
                else              t <- paste(t, "+i")
            } else if (Im(y) == -1)
                t <- paste(t, "-i")
            else {
                if (Re(y) == 0)      t <- paste(Im(y), "i")
                else if (Im(y) > 0) t <- paste(t, "+", Im(y), "i")
                else                 t <- paste(t, Im(y), "i")
            }
            return(t)
        })
        dim(z) <- dim(x)
        xtable::xtable(z, ...)
    } else
        xtable::xtable(x, ...)
}

if (length(args) == 1) {
    x <- xtable(args[[1]], align=rep("",ncol(args[[1]])+1))

    # 1.00 -> 1, in our peculiar case
    digits(x) <- xdigits(x)
    print(x,
        floating = FALSE, tabular.environment = "pmatrix",
        hline.after=NULL, include.rownames=FALSE, include.colnames=FALSE
    )
    q()
} else if (length(args) != 5) stop("Incomplete wave function")

x <- avg(args[[1]], unlist(args[2:5]))

# avoids some 0+0i; refinable
if (x == 0) {cat(0, "\n")} else {cat(x, "\n")}
```

8 Particles and Waves

8.1 Mathematical Interlude: Working with Continuous Functions

8.1.1 Wave Function review

8.1.2 Functions as Vectors

8.1.3 Integration by Parts

8.1.4 Linear Operators

Exercise 41. *Prove that $\boldsymbol{X}$ and $\boldsymbol{D}$ are linear operators.*

The two operators are defined on a Hilbert space $\mathcal{H}$ by:

$$\boldsymbol{X} : \begin{pmatrix} \mathcal{H} & \to & \mathcal{H} \\ \psi & \mapsto & (x \mapsto x\psi(x)) \end{pmatrix}; \qquad \boldsymbol{D} : \begin{pmatrix} \mathcal{H} & \to & \mathcal{H} \\ \psi & \mapsto & (x \mapsto \frac{d}{dx}\psi(x)) \end{pmatrix}$$

Generally speaking, an operator $\boldsymbol{L} : \mathcal{H} \to \mathcal{H}$ is said to be linear if those two axioms are verified:

$$(\forall \alpha \in \mathbb{C}),\ \boldsymbol{L}(\alpha\psi) = \alpha\boldsymbol{L}(\psi)$$

$$(\forall (\psi, \phi) \in \mathcal{H}^2),\ \boldsymbol{L}(\psi + \phi) = \boldsymbol{L}(\psi) + \boldsymbol{L}(\phi)$$

Remark 30. *It's customary in quantum-mechanics to drop the parentheses when apply an operator, i.e. $\boldsymbol{L}\psi := \boldsymbol{L}(\psi)$; I'll keep them here for clarity.*

Furthermore, while the authors often use "$\psi(x)$" to denote a function, I'll be using ψ instead, and reserve $\psi(x)$ to the result of the application of ψ to the variable x, as is usual in mathematics.

Finally, recall[67] *that addition and scalar-multiplication are defined pointwise*[68] *on functions:*

$$(\forall(\psi,\phi)\in\mathcal{H}),\quad \psi+\phi := \Big(x\mapsto(\psi+\phi)(x) := \psi(x)+\phi(x)\Big)$$

$$(\forall(\psi,\alpha)\in\mathcal{H}\times\mathbb{C}),\quad \alpha\psi := \Big(x\mapsto(\alpha\psi)(x) := \alpha\psi(x)\Big)$$

To ease notation, I'll use the same symbols for e.g. the addition of complex numbers and the (pointwise) addition of functions. Don't hesitate to label them in your mind in case of doubt.

Starting with $\boldsymbol{X}$ and the first axiom; let $x\in\mathbb{R}$, $\alpha\in\mathbb{C}$ and $\psi\in\mathcal{H}$:

$$\begin{aligned}(\boldsymbol{X}(\alpha\psi))(x) &= x\alpha\psi(x)\\ &= \alpha(x\psi(x))\\ &= \alpha(\boldsymbol{X}\psi)(x)\end{aligned}$$

As this is true for any x, we can conclude:

$$\boxed{\boldsymbol{X}(\alpha\psi)=\alpha\boldsymbol{X}(\psi)}$$

Moving on to the second axiom; let $x\in\mathbb{R}$ and $(\psi,\phi)\in\mathcal{H}^2$:

$$\begin{aligned}(\boldsymbol{X}(\psi+\phi))(x) &= x(\psi(x)+\phi(x))\\ &= x\psi(x)+x\phi(x)\\ &= (\boldsymbol{X}(\psi))(x)+(\boldsymbol{X}(\phi))(x)\\ &= (\boldsymbol{X}(\psi)+\boldsymbol{X}(\phi))(x)\end{aligned}$$

Again, this is true for any x and thus:

$$\boxed{\boldsymbol{X}(\psi+\phi)=\boldsymbol{X}(\psi)+\boldsymbol{X}(\phi)}$$

The second operator is a little more interesting; let $x\in\mathbb{R}$, $\alpha\in\mathbb{C}$ and $\psi\in\mathcal{H}$:

$$\begin{aligned}(\boldsymbol{D}(\alpha\psi))(x) &= (\frac{d}{dx}\alpha\psi)(x)\\ &= \alpha(\frac{d}{dx}\psi)(x)\\ &= \alpha(\boldsymbol{D}\psi)(x)\end{aligned}$$

You may be wondering why we're allowed to shift the α outside of the differential operator. Let me clarify this a little. The (real) differentiation operator is defined as a limit:

$$\frac{d}{dx}\psi(x) := \lim_{\epsilon\to0}\frac{\psi(x+\epsilon)-\psi(x)}{\epsilon} =: \psi'(x)$$

So we can develop our previous equation as[69]:

$$\frac{d}{dx}(\alpha\psi(x)) = \lim_{\epsilon\to0}\frac{\alpha\psi(x+\epsilon)-\alpha\psi(x)}{\epsilon} = \lim_{\epsilon\to0}\alpha\frac{\psi(x+\epsilon)-\psi(x)}{\epsilon}$$

And thus all the difficulty is in knowing wether the α can "jump" outside of the limit. And the answer is yes[70] *assuming the remaining limit exists.* Meaning we can as long as the following limit exists (it must converges to some fixed point in $\mathbb{C}$, or equivalently, is must not diverge to $\pm\infty$):

$$\lim_{\epsilon\to0}\frac{\psi(x+\epsilon)-\psi(x)}{\epsilon}$$

[67] The authors did it a bit quickly a few pages earlier, but I've done it more carefully in L03E01

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

[69] Remember the pointwise definition of the scalar multiplication of a function.

[70] For a proof, have a look at `https://tutorial.math.lamar.edu/classes/calci/limitproofs.aspx`.

This is equivalent to saying that we can do it as long as ψ is differentiable. In a physics context, functions are often always assumed to be differentiable everywhere. Hence the first axiom indeed holds for $\boldsymbol{D}$:

$$\boxed{\boldsymbol{D}(\alpha\psi) = \alpha\boldsymbol{D}(\psi)}$$

There's an analogue reasoning for the second axiom: let $x \in \mathbb{R}$ and $(\psi, \phi) \in \mathcal{H}^2$:

$$\begin{aligned}(\boldsymbol{D}(\psi+\phi))(x) = &\quad \frac{d}{dx}(\psi+\phi)(x)\\ = &\quad (\frac{d}{dx}\psi + \frac{d}{dx}\phi)(x)\\ = &\quad (\boldsymbol{D}(\psi) + \boldsymbol{D}(\phi))(x)\end{aligned}$$

Again we can rewrite the "questionable" line by expanding the differentiation as a limit while unwrapping the pointwise addition of functions:

$$\begin{aligned}(\frac{d}{dx}(\psi+\phi))(x) = &\quad \lim_{\epsilon\to 0}\frac{(\psi+\phi)(x+\epsilon) - (\psi+\psi)(x)}{\epsilon}\\ = &\quad \lim_{\epsilon\to 0}\frac{\psi(x+\epsilon) + \phi(x+\epsilon) - (\psi(x)+\phi(x))}{\epsilon}\\ = &\quad \lim_{\epsilon\to 0}\frac{\Big(\psi(x+\epsilon) - \psi(x)\Big) + \Big(\phi(x+\epsilon) - \phi(x)\Big)}{\epsilon}\\ = &\quad \lim_{\epsilon\to 0}\Big(\frac{\psi(x+\epsilon) - \psi(x)}{\epsilon} + \frac{\phi(x+\epsilon) - \phi(x)}{\epsilon}\Big)\end{aligned}$$

Again, we can split the limit of a sum to a sum of limits[71], *as long as both limits converge.* Hence the second axioms holds as long as ψ and ϕ are differentiable:

$$\boxed{\boldsymbol{D}(\psi+\phi) = \boldsymbol{D}(\psi) + \boldsymbol{D}(\phi)}$$

8.2 The State of a Particle

8.2.1 The Eigenvalues and Eigenvectors of Position

8.2.2 Momentum and Its Eigenvectors

8.3 Fourier Transforms and the Momentum Basis

8.3.1 Resolving the Identity

8.4 Commutators and Poisson Brackets

8.5 The Heisenberg Uncertainty Principle

9 Particle Dynamics

9.1 A Simple Example

9.2 Nonrelativistic Free Particles

9.3 Time-Independent Schrödinger Equation

Exercise 42. *Derive Eq.* 9.7 *by plugging Eq.* 9.6 *into Eq.* 9.5.

Let's recall in order, Eq. 9.7, Eq. 9.6 and Eq. 9.5:

$$E = p^2/2m; \qquad \psi(x) = \exp(ipx/\hbar); \qquad -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} = E\psi(x)$$

In that last equation, the RHS could be rewritten as $\boldsymbol{H}|\Psi\rangle$, where $\boldsymbol{H}$ is the "quantized" classical Hamiltonian corresponding to a free particle, that is, a particle not affected by a potential energy: the

[71] There's a proof on the same website as before

Hamiltonian is then built solely from the "quantized" kinetic energy.

Eq. 9.6 (the middle one) is a solution proposal to the ODE yielded by Eq. 9.5 (the last one). Let's see how it goes:

$$\begin{aligned}
& -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2} = && E\psi(x) \\
\Leftrightarrow \quad & -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\exp\left(ipx/\hbar\right) = && E\psi(x) \\
\Leftrightarrow \quad & -\frac{\hbar^2}{2m}\left(\frac{ip}{\hbar}\right)^2 \underbrace{\exp\left(ipx/\hbar\right)}_{=:\psi(x)} = && E\psi(x) \\
\Leftrightarrow \quad & \frac{p^2}{2m}\psi(x) = && E\psi(x)
\end{aligned}$$

And so indeed, at least as long as $\psi(x) \neq 0$:

$$\boxed{E = p^2/2m} \quad \square$$

9.4 Velocity and Momentum

Exercise 43. *Prove Eq.* 9.10 *by expanding each site and comparing the results.*

Eq. 9.10 is:

$$[\boldsymbol{P}^2, \boldsymbol{X}] = \boldsymbol{P}[\boldsymbol{P}, \boldsymbol{X}] + [\boldsymbol{P}, \boldsymbol{X}]\boldsymbol{P}$$

We're trying to compute the "velocity" of a free particle, meaning, the time derivative of its position operator $\boldsymbol{X}$, which can be performed thanks to an earlier formula derived in section 4.9 - *Connections to classical mechanics*:

$$\frac{d}{dt}\langle \boldsymbol{X} \rangle = \frac{i}{\hbar}\langle [\boldsymbol{H}, \boldsymbol{X}] \rangle$$

To push the computation forward, we then need to compute $[\boldsymbol{H}, \boldsymbol{X}]$, which is, for a free particle, up to a constant, equivalent to $[\boldsymbol{P}^2, \boldsymbol{X}]$.

Recall that the *commutator* measures how much two operators defined on a Hilbert space commute:

$$[\boldsymbol{A}, \boldsymbol{B}] := \boldsymbol{AB} - \boldsymbol{BA}$$

Let's progressively expand both sides, keeping all expansions strictly equivalent:

$$\begin{aligned}
& [\boldsymbol{P}^2, \boldsymbol{X}] = && \boldsymbol{P}[\boldsymbol{P}, \boldsymbol{X}] + [\boldsymbol{P}, \boldsymbol{X}]\boldsymbol{P} \\
\Leftrightarrow \quad & \boldsymbol{P}^2\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P}^2 = && \boldsymbol{P}(\boldsymbol{PX} - \boldsymbol{XP}) + (\boldsymbol{PX} - \boldsymbol{XP})\boldsymbol{P} \\
\Leftrightarrow \quad & \boldsymbol{PPX} - \boldsymbol{XPP} = && \boldsymbol{PPX} - \boldsymbol{PXP} + \boldsymbol{PXP} - \boldsymbol{XPP} \\
\Leftrightarrow \quad & \boldsymbol{PPX} - \boldsymbol{XPP} = && \boldsymbol{PPX} - \boldsymbol{XPP} \\
\Leftrightarrow \quad & 0 = && 0 \\
\Leftrightarrow \quad & \text{true} && \square
\end{aligned}$$

9.5 Quantization

9.6 Forces

Exercise 44. *Show that the right-hand side of Eq.* 9.17 *simplifies to the right-hand side of Eq.* 9.16. *Hint: First expand the second term by taking the derivative of the product. Then look for cancellations.*

Eq. 9.17 and Eq. 9.16 are respectively:

$$[\boldsymbol{V}(x), \boldsymbol{P}]\psi(x) = V(x)(-i\hbar\frac{d}{dx})\psi(x) - (-i\hbar\frac{d}{dx})V(x)\psi(x)$$

$$[\boldsymbol{V}(x), \boldsymbol{P}] = i\hbar\frac{dV(x)}{dx}$$

Let's start as suggested by expanding the second term of Eq. 9.17, ignoring the $-i\hbar$ factor for now: this is a basic product rule application:

$$(\frac{d}{dx})(V(x)\psi(x)) = \frac{dV(x)}{dx}\psi(x) + V(x)\frac{d\psi(x)}{dx}$$

The second term will cancel with the first term of the RHS of Eq. 9.17:

$$\begin{aligned}
[\boldsymbol{V}(x), \boldsymbol{P}]\psi(x) = &\quad V(x)(-i\hbar\frac{d}{dx})\psi(x) - (-i\hbar\frac{d}{dx})V(x)\psi(x) \\
= &\quad -i\hbar V(x)\frac{d\psi(x)}{dx} + i\hbar\left(\frac{dV(x)}{dx}\psi(x) + V(x)\frac{d\psi(x)}{dx}\right) \\
= &\quad i\hbar\frac{dV(x)}{dx}\psi(x)
\end{aligned}$$

As long as $\psi(x) \neq 0$, we can divide by $\psi(x)$, and indeed establish Eq. 9.16:

$$\boxed{[\boldsymbol{V}(x), \boldsymbol{P}] = i\hbar\frac{dV(x)}{dx}} \quad \square$$

9.7 Linear Motion and the Classical Limit

9.8 Path integrals

10 The Harmonic Oscillator

10.1 The Classical Description

Exercise 45. *Find the second time derivative of x in Eq.* 10.9*, and thereby show that it solves Eq.* 10.8*.*

Eq. 10.9 and 10.8 respectively are:

$$x(t) = A\cos(\omega t) + B\sin(\omega t); \qquad -\omega^2 x(t) = \ddot{x}(t)$$

Where I've systematically made the time-dependence explicit by replacing x with $x(t)$. This is an elementary differentiation exercise that I think has already been performed in the previous volume on classical mechanics. Nevertheless:

$$\begin{aligned}
\dot{x}(t) = &\quad -A\omega\sin(\omega t) + B\omega\cos(\omega t) \\
\ddot{x}(t) = &\quad -A\omega^2\cos(\omega t) - B\omega^2\sin(\omega t) \\
= &\quad -\omega^2\underbrace{(A\cos(\omega t) + B\sin(\omega t))}_{=:x(t)} \\
= &\quad -\omega^2 x(t) \quad \square
\end{aligned}$$

10.2 The Quantum Mechanical Description

10.3 The Schrödinger Equation

10.4 Energy Levels

10.5 The Ground State

10.6 Creation and Annihilation Operators

10.7 Back to the Wave Functions

10.8 The Importance of Quantization