The Theoretical Minimum
Quantum Mechanics - Solutions

M. Bivert

April 5, 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 serie.

Contents

1 Systems and Experiments

1.1 Inner Products

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

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

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

---

a) Let us recall the two axioms in question:

Axiom 1.

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

Axiom 2.

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

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

Let us recall also that if

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

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

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

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

We thus have:

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

b) Mainly from the second axiom:

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

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

Let us recall the two axioms in question:

Axiom 3.

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

Axiom 4.

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

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

And let us recall Eq. 1.2 of the book:

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

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

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

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

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

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

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

Hence, regarding the second axiom:

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

2 Quantum States

2.1 Along the x Axis

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

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

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

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

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

For instance:

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

Or, similarly:

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

2.2 Along the y Axis

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

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

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

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

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

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

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

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

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

And to make better sense of those equations, let us recall that $\alpha_u^*\alpha_u = \langle A|u\rangle \langle u|A\rangle$ is the probability of a state vector $|A\rangle = \alpha_u|u\rangle + \alpha_d|d\rangle$ to be measured in the state $|u\rangle$.
For Eq. 2.7, we have

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

So this means the first term is multiplied by

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

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

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

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

$$\boxed{\text{The proposed solution is \textit{not} unique [up to a phase factor].}}$$

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

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

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

a) Use Eqs. 2.8 *to show that*

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

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

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

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

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

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

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

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

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

---

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

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

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

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

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

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

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

Because of the complex conjugation rule, we have

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

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

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

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

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

---

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:

$$\begin{aligned}
\alpha^*\beta \underbrace{\langle u|r\rangle}_{=1/\sqrt{2}} \underbrace{\langle r|d\rangle}_{=1/\sqrt{2}} + \beta^*\alpha \underbrace{\langle d|r\rangle}_{=1/\sqrt{2}} \underbrace{\langle r|u\rangle}_{=1/\sqrt{2}} &= 0 \\
\Leftrightarrow \frac{1}{2}\alpha^*\beta + \frac{1}{2}\beta^*\alpha &= 0 \\
\Leftrightarrow \boxed{\alpha^*\beta + \beta^*\alpha = 0} \quad \square
\end{aligned}$$

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:

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

---

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:

$$\left(zw\right)^* = z^*w^*$$

Indeed:

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

Hence:

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

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:

$$\alpha^*\beta + \left(\alpha^*\beta\right)^* = 0$$

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

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 some space, I will assume that you're familiar with linear algebra.[1]. 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*[2].

Now, I'm *not* going to prove the the infinite dimension version here. There's a good reason why quantum mechanics courses often start with spins: they don't require the generalized results, which demands heavy mathematical machinery (a copious amount of functional analysis, and in some formulation at least, the Lebesgue integral, hence portions of measure theory). You may want to refer to F. Schuller YouTube lectures on quantum mechanics[3] for a thorough development.

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

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

[1] I'm preparing some notes thoroughly covering the necessary background, but they won't be ready anytime soon.

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

[3] `https://www.youtube.com/watch?v=GbqA9Xn_iM0&list=PLPH7f_7ZlzxQVx5jRjbfRGEzWY_upS5K6`; see also the lectures notes (.pdf) made by a student (Simon Rea): `https://drive.google.com/file/d/1nchF1fRGSY3R3rP1QmjUg7fe28tAS428/view`

Theorem 1. *Let $H : V \to V$ be a Hermitian operator on a finite-dimensional vector space V, equipped with an inner-product*[4].

$$\boxed{\text{Then, the eigenvectors of } H \text{ form an orthonormal basis of } V.}$$

Saying it otherwise, it means that H is diagonalizable, and that two eigenvectors associated with distinct eigenvalues are orthogonal.

For clarity, let's recall a few definitions.

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

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

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

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

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

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

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 $\boldsymbol{v} \neq \mathbf{0}_V$[5].

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

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

where:

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

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 $(\alpha_0, \alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$ which are not all zero, such that:

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

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

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

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

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

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

[5] 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.

This allows you to graph the polynomial and so forth:

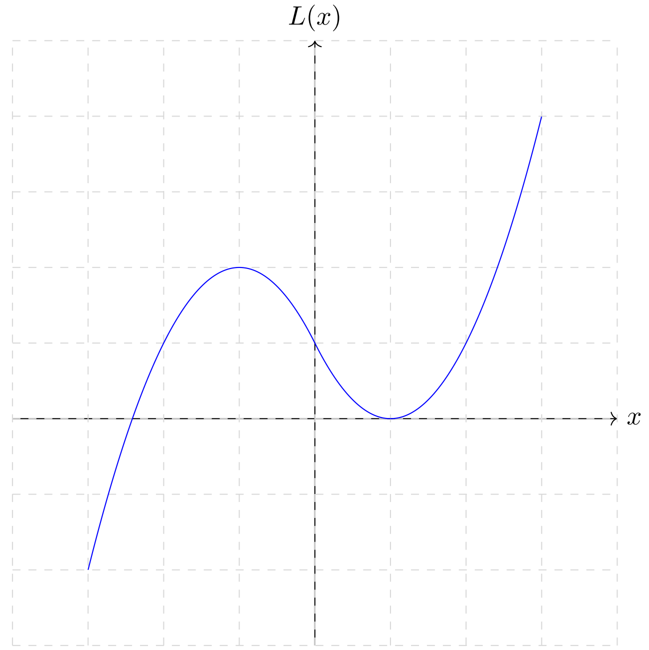

Figure 1: $L(x) = x^2 - 2x + 1$

But that's all kindergarten polynomials. The more "correct" polynomials are *not* functions of a real variable. Rather, we say that $L(x)$ is a polynomial of a single variable/indeterminate[6] 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 is sufficient to make it clear, and perhaps most importantly, because an truly unambiguous notation is disturbingly 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*[7] *of polynomials of a single indeterminate over a field*[8] $\mathbb{F}$. This means that mathematicians have defined a way This means that $X^2 - 2X + 1$ is actually a shortcut for:

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

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

[7] `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.

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

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, in particular, the *fundamental theorem of algebra.*

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

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

Our goal is to transform this expression so that it involves a polynomial in $\mathbb{C}[L]$[9].

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 of complex numbers to a scalar multiplication on a function.

The conventional/natural way (the simplest consistent way) to do so, is to define them pointwise[10] 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:

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

This means that our P is well defined, and that we can indeed pull the $\boldsymbol{v}$ out.

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

But if you look closely:

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

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

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

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 intent and purposes, it "is" a polynomial, and we can manipulate it as such.
So we can apply the fundamental theorem of algebra[11], we know that we can always factorize polynomials with complex coefficient as such:

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

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

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

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

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

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 all real.

Saying it otherwise, it means that the 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 up an basis for the space. Again, you may want to refer yourself to a linear algebra course for more details.

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

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

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.

This means it really remains 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 $\forall n \in \mathbb{N}\backslash\{0\}$.

$\boxed{\text{n=1}}$ 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 every 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 $\{\boldsymbol{v}_1, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{k+1}\} \subset V$ so that $\{\boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{k+1}\}$ is a basis of V[12].

Apply the Gram-Schmidt procedure[13] to extract from it an orthonormal basis $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_{k+1}\}$ 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[14].

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:

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

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

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

Rewrite D_H accordingly, and break it into blocks:

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

[12] 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.

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

[14] $H(\boldsymbol{b}_1) = H(\boldsymbol{v}/\|\boldsymbol{v}\|)$, 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, $A^\dagger = (0\ 0\ \ldots\ 0) = A$, i.e:

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

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

□

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 $(\sigma_z)_{22}$:

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

3.5 A Common Misconception

3.6 3-Vector Operators Revisited

3.7 Reaping the Results