L and M are simultaneously measurable

$$\Leftrightarrow \quad 0 = [M, L] = A\sigma_x\sigma_y + B\sigma_y\sigma_z + C\sigma_z\sigma_a$$

$$= A\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + B\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + C\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} Ai & Bi+C \\ Bi-C & -Ai \end{pmatrix}$$

$$\Leftrightarrow \quad A = B = C = 0$$

$$\Leftrightarrow \quad a_1 b_2 = a_2 b_1,\ b_1 c_2 = b_2 c_1, \text{ and } c_1 a_2 = c_2 a_1 \quad \blacksquare$$

Corollary. No two of the components σ_x, σ_y, and σ_z of the spin 3-vector operator $\vec{\sigma}$ of Alice's system can be simultaneously measured.

Proof. Let $L = \sigma_x$, $M = \sigma_y$, and $N = \sigma_z$. For L and M, $a_1 b_2 = 1$ and $a_2 b_1 = 0$, so the simultaneous equations are not satisfied. Similarly for L & N, and M & N. $\blacksquare$

We now investigate the topic of uncertainty in an observation. Let an observable L have eigenvalues $\{\lambda\}_{\lambda \in \Lambda}$. Let $|\Psi\rangle = \sum_{\lambda \in \Lambda} \lambda |\lambda\rangle$ be the normalized state vector for L. Then

(i) $\langle L \rangle \overset{(4.13)}{=} \langle \Psi | L | \Psi \rangle \overset{(4.11)}{=} \sum_{\lambda} \lambda \mathrm{P}_L(\lambda)$.

Define

(ii) $\bar{L} = L - \langle L \rangle \boldsymbol{I}$

(iii) $\bar{\lambda} = \lambda - \langle L \rangle$

and

(iv) $|\bar{\lambda}\rangle = |\lambda\rangle$.

Then

(v) $\{\bar{\lambda}\}$ and $\{|\bar{\lambda}\rangle\}$ are the eigenvalues and eigenvectors of $\bar{L}$:

$$\bar{L}|\bar{\lambda}\rangle \overset{(iv)}{=} \bar{L}|\lambda\rangle \overset{(ii)}{=} (L - \langle L \rangle \boldsymbol{I})|\lambda\rangle = L|\lambda\rangle - \langle L \rangle \boldsymbol{I}|\lambda\rangle = \lambda|\lambda\rangle - \langle L \rangle |\lambda\rangle$$

$$= (\lambda - \langle L \rangle)|\lambda\rangle \overset{(iii,\ iv)}{=} \bar{\lambda}|\bar{\lambda}\rangle \quad ✔$$

Notice that $\langle \bar{L} \rangle \overset{(ii)}{=} \langle L \rangle - \langle L \rangle \langle \boldsymbol{I} \rangle = 0$. In fact the distribution of $\bar{L}$ is simply the distribution of L shifted left or right to have a mean of zero. Therefore

(vi) $\mathrm{P}_{\bar{L}}(\bar{\lambda}) = \mathrm{P}_L(\lambda)$.

Definition. The **uncertainty of an observable L** is its standard deviation, $\boldsymbol{\Delta L}$.

Theorem. $\left(\Delta L\right)^2 = \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}_L\left(\lambda\right)$ (5.3)

Proof. $\left(\Delta L\right)^2 = \text{Variance} = \sum_\lambda \left(\lambda - \left\langle L \right\rangle\right)^2 \mathrm{P}_L\left(\lambda\right)$ (5.4)

$$\overset{\text{(iii)}}{=} \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}_L\left(\lambda\right) \qquad \blacksquare$$

$$\left\langle \bar{L}^2 \right\rangle \overset{\text{(Defn)}}{=} \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}_{\bar{L}^2}\left(\bar{\lambda}^2\right) \overset{\text{(Defn)}}{=} \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}\left(\bar{L}^2 = \bar{\lambda}^2\right) = \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}\left(\bar{L} = \bar{\lambda}\right)$$
$$\overset{\text{(Defn)}}{=} \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}_{\bar{L}}\left(\bar{\lambda}\right) \overset{\text{(vi)}}{=} \sum_\lambda \bar{\lambda}^2 \, \mathrm{P}_L\left(\lambda\right) \overset{\text{(5.3)}}{=} \left(\Delta L\right)^2$$

In case $\left\langle L \right\rangle = 0$ then $\bar{L} \overset{\text{(ii)}}{=} L$ and so this formula takes the simpler form

$$\left(\Delta L\right)^2 = \left\langle L^2 \right\rangle \overset{\text{(i)}}{=} \left\langle \Psi \middle| L^2 \middle| \Psi \right\rangle \tag{5.4b}$$

In other words the square of the uncertainty is the mean value of the operator L^2.

For the next two theorems Susskind reverts to standard vector notation, $\vec{X}$.
Keep in mind that the vector dot product is just the bra-ket inner product:

$$\vec{X} \bullet \vec{Y} = \left\langle X \middle| Y \right\rangle \tag{5.4c}$$

Theorem. **(Cauchy-Schwarz Inequality for real vectors)**. If $\vec{X}$ and $\vec{Y}$ are vectors in $\mathbb{R}^n$ then

$$\left|\vec{X}\right|\left|\vec{Y}\right| \geq \left|\vec{X} \bullet \vec{Y}\right| \tag{5.7}$$

Proof. From the Triangle Inequality,

$$\left|\vec{X}\right| + \left|\vec{Y}\right| \geq \left|\vec{X} + \vec{Y}\right|. \tag{i}$$

By definition,

$$\left|\vec{X}\right|^2 \overset{(1.24)}{=} \vec{X} \bullet \vec{X}. \tag{ii}$$

So

$$\left|\vec{X} + \vec{Y}\right|^2 \overset{\text{(ii)}}{=} \left(\vec{X} + \vec{Y}\right) \bullet \left(\vec{X} + \vec{Y}\right)$$
$$= \vec{X} \bullet \vec{X} + \vec{Y} \bullet \vec{Y} + 2\vec{X} \bullet \vec{Y} \overset{\text{(ii)}}{=} \left|\vec{X}\right|^2 + \left|\vec{Y}\right|^2 + 2\vec{X} \bullet \vec{Y}. \tag{iii}$$

Hence

$$\left|\vec{X}\right|^2+\left|\vec{Y}\right|^2+2\left|\vec{X}\right|\left|\vec{Y}\right|=\left(\left|\vec{X}\right|+\left|\vec{Y}\right|\right)^2 \overset{(i)}{\geq}\left|\vec{X}+\vec{Y}\right|^2 \overset{(iii)}{=}\left|\vec{X}\right|^2+\left|\vec{Y}\right|^2+2\vec{X}\bullet\vec{Y}$$

$$\Rightarrow \quad \left|\vec{X}\right|\left|\vec{Y}\right|\geq\vec{X}\bullet\vec{Y} \tag{5.6}$$

Since also $\left|\vec{X}\right|\left|\vec{Y}\right|=\left|\vec{X}\right|\left|-\vec{Y}\right|\geq\vec{X}\bullet\left(-\vec{Y}\right)=-\vec{X}\bullet\vec{Y}$

we have $\left|\vec{X}\right|\left|\vec{Y}\right|\geq\left|\vec{X}\bullet\vec{Y}\right|$ ■

Note 1. For real vectors, $\vec{X}\bullet\vec{Y}=\left|\vec{X}\right|\left|\vec{Y}\right|\cos\theta$ where θ is the angle between the vectors. Equation 5.6 follows easily from this fact, but the above approach is useful as a model for the complex case.

Note 2. For real vectors $\vec{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ and $\vec{Y}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$, $\vec{X}\bullet\vec{Y}=\sum_{k=1}^{n}x_k\,y_k$.

However, for complex vectors, bra-ket notation makes it clear that

$$\vec{X}\bullet\vec{Y}=\sum_{k=1}^{n}x_k^{\,*}\,y_k: \tag{5.7a}$$

$$\vec{X}\bullet\vec{Y}\overset{(5.4c)}{=}\left\langle X\middle|Y\right\rangle\overset{(1.14)}{=}\left(\left\langle X\right|\right)\left(\left|Y\right\rangle\right)\overset{(1.07)}{=}\begin{pmatrix}x_1^* & \cdots & x_n^*\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\sum_{k=1}^{n}x_k^{\,*}\,y_k$$

Consequently, in general $\left\langle X\middle|Y\right\rangle\neq\left\langle Y\middle|X\right\rangle$.

Theorem. **(Cauchy-Schwarz Inequality for complex vectors).** If $\vec{X}$ and $\vec{Y}$ are vectors in $\mathbb{C}^n$ then

$$\left|X\right|\left|Y\right|\geq\frac{1}{2}\left|\left\langle X\middle|Y\right\rangle+\left\langle Y\middle|X\right\rangle\right| \tag{5.9}$$

Proof. We begin with the Triangle Inequality:

$$\left|X\right|+\left|Y\right|\geq\left|X+Y\right|. \tag{i}$$

Then

$$\left|X+Y\right|^2\overset{(1.24)}{=}\left\langle X+Y\middle|X+Y\right\rangle\overset{(1.21,\ 1.21b,\ 1,24)}{=}\left|X\right|^2+\left|Y\right|^2+\left\langle X\middle|Y\right\rangle+\left\langle Y\middle|X\right\rangle. \tag{ii}$$

Therefore

$$|X|^2+|Y|^2+2|X||Y|=\left(|X|+|Y|\right)^2\overset{\text{(i)}}{\geq}|X+Y|^2\overset{\text{(ii)}}{=}|X|^2+|Y|^2+\langle X|Y\rangle+\langle Y|X\rangle$$

which implies

$$2|X||Y|\geq\langle X|Y\rangle+\langle Y|X\rangle. \tag{iii}$$

Since

$$2|X||Y|=2|X||-Y|\overset{\text{(iii)}}{\geq}\langle X|-Y\rangle+\langle -Y|X\rangle=-\left(\langle X|Y\rangle+\langle Y|X\rangle\right) \tag{iv}$$

we have that

$$2|X||Y|\overset{\text{(iii, iv)}}{\geq}\left|\langle X|Y\rangle+\langle Y|X\rangle\right|. \quad \blacksquare$$

We now use the Cauchy-Schwarz formula (5.9) to develop the Uncertainty Principle. This formula is primarily for Case A in which L and M are Alice observables. While it is true for Case B (i.e., Alice and Bob observables), it does not have much substance since the commutator is zero. In case B the observables are independent and are not encumbered with uncertainty.

Theorem. **(The Uncertainty Principle)**. Let L and M be observables and $|\Psi\rangle$ any normalized simultaneous eigenvector of L and M. Then

$$\Delta L\,\Delta M\geq\frac{1}{2}\left|\langle\Psi|\,[L,M]\,|\Psi\rangle\right| \tag{5.13}$$

Proof. Define

$$|X\rangle=L|\Psi\rangle \quad\Rightarrow\quad \langle X|=\langle\Psi|L^\dagger\overset{\text{(Hermitian)}}{=}\langle\Psi|L \tag{a}$$

and

$$|Y\rangle=iM|\Psi\rangle \quad\Rightarrow\quad \langle Y|=-i\langle\Psi|M^\dagger=-i\langle\Psi|M\,. \tag{b}$$

Then

$$|X|^2\overset{(1.24)}{=}\langle X|X\rangle\overset{\text{(a)}}{=}\left(\langle\Psi|L\right)\left(L|\Psi\rangle\right)=\langle\Psi|L^2|\Psi\rangle\overset{(4.14)}{=}\langle L^2\rangle$$

$$\Rightarrow\qquad |X|=\sqrt{\langle L^2\rangle} \tag{c}$$

and

$$|Y|^2=\langle Y|Y\rangle\overset{\text{(b)}}{=}-i^2\langle\Psi|M^2|\Psi\rangle\overset{(4.14)}{=}\langle M^2\rangle$$

$$\Rightarrow \qquad |Y| = \sqrt{\langle M^2 \rangle}\,. \tag{d}$$

Thus,

$$\langle X | Y \rangle \overset{(a,\, b)}{=} \big(\langle \Psi | L \rangle (iM | \Psi \rangle) \big) = i \langle \Psi | LM | \Psi \rangle \tag{e}$$

and

$$\langle Y | X \rangle \overset{(a,\, b)}{=} -i \langle \Psi | ML | \Psi \rangle. \tag{f}$$

Claim: $\sqrt{\langle L^2 \rangle \langle M^2 \rangle} \geq \frac{1}{2} \left| \langle \Psi | [L, M] | \Psi \rangle \right|$: (5.12)

$$\begin{aligned} 2\sqrt{\langle L^2 \rangle \langle M^2 \rangle} &\overset{(c,\, d)}{=} 2|X||Y| \overset{(5.9)}{\geq} \left| \langle X | Y \rangle + \langle Y | X \rangle \right| \\ &\overset{(e,\, f)}{=} |i| \left| \langle \Psi | LM | \Psi \rangle - \langle \Psi | ML | \Psi \rangle \right| = \left| \langle \Psi | [L, M] | \Psi \rangle \right| \quad ✔ \end{aligned}$$

If $\langle L \rangle = 0 = \langle M \rangle$ then

$$(\Delta L)^2 \overset{(5.4b)}{=} \langle L^2 \rangle \tag{g}$$

and

$$(\Delta M)^2 \overset{(5.4b)}{=} \langle M^2 \rangle. \tag{h}$$

So

$$\Delta L \, \Delta M \overset{(g,\, h)}{=} \sqrt{\langle L^2 \rangle \langle M^2 \rangle} \overset{(5.12)}{\geq} \frac{1}{2} \left| \langle \Psi | [L, M] | \Psi \rangle \right| \quad \blacksquare$$

Exercise 5.2 proves that the Uncertainty Principle holds even without the assumption that $\langle L \rangle = 0 = \langle M \rangle$.

Corollary. In a discrete quantum system, $\Delta L \, \Delta M \geq \frac{1}{2} \langle [L, M] \rangle$

Proof. $\Delta L \, \Delta M \overset{(5.13)}{\geq} \frac{1}{2} \left| \langle \Psi | [L, M] | \Psi \rangle \right| \overset{(4.13)}{=} \frac{1}{2} \langle [L, M] \rangle$. $\blacksquare$

The Uncertainty Principle: Equation (5.13): The product of the uncertainties of observables *L* and *M* cannot be smaller than half the magnitude of the expected value of their commutator.

This quantifies the fact that if the commutator of *L* and *M* is not zero then both observables cannot be simultaneously certain.

Wave Function

Definition. Let $\left\{\left|a\right\rangle\right\}$ be an orthonormal basis for an observable L. If $\left|\Psi\right\rangle = \sum_a \psi_j \left|a\right\rangle$ is the state vector of L then the set of coefficients $\left\{\psi_j\right\}$ is known as a **wave function.** We sometimes write ψ_j as $\psi(j)$. If there are multiple observables then the state vector would be

$$\left|\Psi\right\rangle = \sum_{a,b,c,\cdots} \psi_{a,b,c,\cdots} \left|a,b,c,\cdots\right\rangle = \sum_{a,b,c,\cdots} \psi(a,b,c,\cdots)\left|a,b,c,\cdots\right\rangle$$

and the wave function is the set of coefficients $\left\{\ \psi(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\cdots) = \psi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\cdots}\ \right\}$.

Note. Since a function has both a domain and a range, it would be more mathematically rigorous to define the wave function as $\left\{\left(j,\psi_j\right)\right\}$.

Chapter 6. Combining Systems: Entanglement

Convention. If Alice and Bob have quantum systems, we use $\mathbf{S}_{\mathbf{A}}$ to represent **Alice's state space**, the **vector space of all her states.** We use $\mathbf{S}_{\mathbf{B}}$ to represent **Bob's state space**. We often write Alice's basis states as $\left|A\right\}$ and Bob's as $\left|B\right\rangle$. The use of braces for Alice's states is to remind us that we cannot add, subtract, multiply, or otherwise mathematically combine Alice's and Bob's states. Instead, we define a new object, $\left|AB\right\rangle$, called the combined state, which is the tensor product (i.e., juxtaposition) of $\left|A\right\}$ and $\left|B\right\rangle$.

Definitions. The **tensor product** of vectors $\left|A\right\}$ and $\left|B\right\rangle$ is denoted $\left|\boldsymbol{A}\right\}\otimes\left|\boldsymbol{B}\right\rangle$ and is usually written in shorthand notation as $\left|\boldsymbol{AB}\right\rangle$ or sometimes as $\left|\boldsymbol{A}\right\}\left|\boldsymbol{B}\right\rangle$. This **combined state** simply means that Alice's state is $\left|A\right\}$ and Bob's is $\left|B\right\rangle$.

Let $\left\{\left|a\right\}\right\}$ and $\left\{\left|b\right\rangle\right\}$ be bases for S_{A} and S_{B}, respectively. We define the **product state space** (or, **composite space**) $\mathbf{S}_{\mathbf{AB}}$ as the vector space whose basis is the set of tensor product states $\left|ab\right\rangle$. S_{AB} is the tensor product of Alice's and Bob's states: $\mathbf{S}_{\mathbf{AB}}=\mathbf{S}_{\mathbf{A}}\otimes\mathbf{S}_{\mathbf{B}}$.

Let $n_{\mathrm{A}}=\dim\left(\mathrm{S}_{\mathrm{A}}\right)$ and $n_{\mathrm{B}}=\dim\left(\mathrm{S}_{\mathrm{B}}\right)$. Then there are $n_{\mathrm{A}}\,n_{\mathrm{B}}$ basis vectors $\left|ab\right\rangle$ in S_{AB}. The dimension of S_{AB} is $\dim\left(\mathrm{S}_{\mathrm{AB}}\right)=n_A\,n_B=\dim\left(\mathrm{S}_{\mathrm{A}}\right)\dim\left(\mathrm{S}_{\mathrm{B}}\right)$.

If $\left|A\right\rangle=\sum_a\alpha_a\left|a\right\}$ and $\left|B\right\rangle=\sum_b\beta_b\left|b\right\rangle$ are state vectors in S_{A} and S_{B}, respectively, then we call

$$\left|\boldsymbol{AB}\right\rangle=\left|\boldsymbol{A}\right\rangle\otimes\left|\boldsymbol{B}\right\rangle=\sum_{\boldsymbol{a}}\sum_{\boldsymbol{b}}\boldsymbol{\alpha}_{\boldsymbol{a}}\boldsymbol{\beta}_{\boldsymbol{b}}\left|\boldsymbol{ab}\right\rangle \tag{6.01}$$

a **product state vector (**or just **product state** for short). The general form for a member of the product state space S_{AB} is the **composite state vector** (or just **composite state)**

$$\left|\Psi\right\rangle=\sum_{\boldsymbol{a}}\sum_{\boldsymbol{b}}\psi_{\boldsymbol{ab}}\left|\boldsymbol{ab}\right\rangle. \tag{6.02}$$

We'll see shortly that S_{AB} is in general much larger than the set of product state vectors.

It will be useful first to develop a couple of composite state inner product formulas. The book does not do this.

Theorem. Let $\left|\Psi\right\rangle = \sum_{ab} \psi_{ab} \left|ab\right\rangle$ and $\left|\Phi\right\rangle = \sum_{ab} \phi_{ab} \left|ab\right\rangle$ be composite states. Then

$$\left\langle \Psi \middle| \Phi \right\rangle = \sum_{ab} \psi_{ab}^{*} \phi_{ab} \ . \tag{6.03}$$

Proof.

$$\left\langle \Psi \middle| \Phi \right\rangle = \left(\left\langle \Psi \right| \right)\left(\left| \Phi \right\rangle \right) = \left(\sum_{ab} \left\langle ab \right| \psi_{ab}^{*} \right)\left(\sum_{a'b'} \phi_{a'b'} \left| a'b' \right\rangle \right) = \sum_{ab} \psi_{ab}^{*} \sum_{a'b'} \phi_{a'b'} \left\langle ab \middle| a'b' \right\rangle$$

$$\overset{(1.15)}{=} \sum_{ab} \psi_{ab}^{*} \sum_{a'b'} \phi_{a'b'} \delta_{aba'b'} = \sum_{ab} \psi_{ab}^{*} \sum_{a'b'} \phi_{a'b'} \delta_{aa'} \delta_{bb'} = \sum_{ab} \psi_{ab}^{*} \phi_{ab} \quad \blacksquare$$

Corollary. The inner product of product states is "**factorizable**":

$$\left\langle A'B' \middle| AB \right\rangle = \left\langle A' \middle| A \right\rangle \left\langle B' \middle| B \right\rangle . \tag{6.04}$$

Proof. Set $\left|\Psi\right\rangle = \left|A'B'\right\rangle$ and $\left|\Phi\right\rangle = \left|AB\right\rangle$. Then $\psi_{ab} = \alpha'_a \beta'_b$ and $\phi_{ab} = \alpha_a \beta_b$. So

$$\left\langle A'B' \middle| AB \right\rangle = \left\langle \Psi \middle| \Phi \right\rangle \overset{(6.03)}{=} \sum_{ab} \alpha'^{*}_a \beta'^{*}_b \alpha_a \beta_b = \left(\sum_a \alpha'^{*}_a \alpha_a \right)\left(\sum_b \beta'^{*}_b \beta_b \right) = \left\langle A' \middle| A \right\rangle \left\langle B' \middle| B \right\rangle \quad \blacksquare$$

Let σ_A and σ_B be observables of Alice's and Bob's systems, respectively. The amount of **correlation** between Alice's and Bob's systems is

$$\mathbf{Corr}\left(\boldsymbol{A,B}\right) \equiv \left\langle \sigma_A \sigma_B \right\rangle - \left\langle \sigma_A \right\rangle \left\langle \sigma_B \right\rangle \tag{6.2}$$

Correlation is a number between -1 and +1, with +1 meaning **perfectly correlated**, -1 **perfectly anti-correlated**, and 0 **completely independen**t (i.e., **uncorrelated**).

Example. Suppose Alice's and Bob's systems are such that every time Alice experiences a +1 measurement, Bob experiences -1, and vice-versa. Thus, for any observation, $\sigma_A \sigma_B = -1$. Hence $\left\langle \sigma_A \sigma_B \right\rangle = -1$. If +1 and -1 are equally likely for Alice and Bob, then $\left\langle \sigma_A \right\rangle = 0 = \left\langle \sigma_B \right\rangle$. This means

$\mathrm{Corr}\left(A,B\right) \overset{(6.2)}{=} -1 - (0)(0) = -1$; that is, the systems are perfectly anti-correlated.

Exercise 6.1. If $\mathrm{P}(a,b) = \mathrm{P}_A(a)\mathrm{P}_B(b)$ then $\mathrm{Corr}(A,B) = 0$; i.e., uncorrelated.

Exercise 6.2. If Alice's basis vectors are orthonormal and Bob's basis vectors are orthonormal, then so are the combined basis vectors. For example, if Alice has 2 basis vectors $\left|a'\right\}$ and $\left|a\right\}$ and Bob has 2 basis vectors $\left|b'\right\rangle$ and $\left|b\right\rangle$ then the combined basis vectors $|ab\rangle$ and $|a'b'\rangle$ are orthogonal unit vectors. We can express this as $\langle a'b'|ab\rangle = \delta_{a'a}\,\delta_{b'b}$.

Definition. Let $\{\,|a\}\,\}$ and $\{\,|b\rangle\,\}$ be bases for S_A and S_B, respectively. Suppose that ***M* is a linear operator acting on the states of the combined system S_{AB}.** *M* has $n_A n_B$ rows and $n_A n_B$ columns. **The element of *M* in row $a'b'$ and column ab is expressed as the inner product $m_{a'b',ab} = \langle a'b'|M|ab\rangle$.**

To understand this expression, recall that $\langle a'b'|M$ is a row vector and it multiplies the column vector $|ab\rangle$ resulting in a single complex term. To be more precise, $\langle a'b'|M$ is a pair of juxtaposed row vectors that are not mathematically combined, and $|ab\rangle$ is a pair of juxtaposed column vectors, and so $m_{a'b',ab}$ is a pair of juxtaposed complex numbers:

$$m_{a'b',ab} = m_{a'a} \otimes m_{b'b} = (m_{a'a}, m_{b'b}). \qquad (6.05)$$

Example: Product of 2 Spin Systems.

Alice's and Bob's spin states

- Basis $\{\,|u\rangle, |d\rangle\,\}$ for each.
- System State:
  - Alice prepares her system in spin state $|A\rangle = \alpha_u |u\} + \alpha_d |d\}$
  - Bob prepares his system in spin state $|B\rangle = \beta_u |\,u\rangle + \beta_d |\,d\rangle$
  - Assuming each state is normalized: $\begin{cases} \alpha_u^* \alpha_u + \alpha_d^* \alpha_d = 1 \\ \beta_u^* \beta_u + \beta_d^* \beta_d = 1 \end{cases}$ (6.4)

Spin System S_{AB}

- Basis $\{\,|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\,\}$
- System State:

$$|AB\rangle \overset{(6.01)}{=} |A\rangle \otimes |B\rangle = \alpha_u \beta_u\,|uu\rangle + \alpha_u \beta_d\,|ud\rangle + \alpha_d \beta_u\,|du\rangle + \alpha_d \beta_d\,|dd\rangle \qquad (6.5)$$

- General composite spin state
  - $|\Psi\rangle \overset{(6.02)}{=} \psi_{uu}|uu\rangle + \psi_{ud}|ud\rangle + \psi_{du}|du\rangle + \psi_{dd}|dd\rangle$
- Assuming the composite spin state is normalized
  - $\langle \Psi|\Psi\rangle \overset{(6.03)}{=} \psi_{uu}^* \psi_{uu} + \psi_{ud}^* \psi_{ud} + \psi_{du}^* \psi_{du} + \psi_{dd}^* \psi_{dd} = 1$ (6.06)

Definition The **singlet state** is

$$\left|\mathbf{sing}\right\rangle = \frac{1}{\sqrt{2}}\left(\left|ud\right\rangle - \left|du\right\rangle\right) \tag{6.07}$$

and the **triplet states** are

$$\begin{aligned} \left|T_1\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|ud\right\rangle + \left|du\right\rangle\right), \\ \left|T_2\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|uu\right\rangle + \left|dd\right\rangle\right), \\ \left|T_3\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|uu\right\rangle - \left|dd\right\rangle\right). \end{aligned} \tag{6.08}$$

The singlet and triplet states are clearly normalized. For example,

$$\begin{aligned} \left\langle \text{sing}\,\middle|\,\text{sing}\right\rangle &\overset{(6.06)}{=} \left(\left\langle ud\right|\frac{1}{\sqrt{2}} - \left\langle du\right|\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\left|ud\right\rangle - \frac{1}{\sqrt{2}}\left|du\right\rangle\right) \\ &= \frac{1}{2}\left(\left\langle ud\middle|ud\right\rangle - \left\langle ud\middle|du\right\rangle - \left\langle du\middle|ud\right\rangle + \left\langle du\middle|du\right\rangle \right) \\ &\overset{(6.03)}{=} \frac{1}{2}\left(1-0-0+1\right) = 1 \end{aligned}$$

Exercise 6.3. The singlet state cannot be written as a product state.

This proves the earlier claim that S_{AB} is larger than the set of product states. That is, there are states in S_{AB} that are not product states.

In fact we can show that most states are not product states by comparing the number of degrees of freedom (dof) for a normalized product state vector and a normalized composite state vector.

Alice's state $\left|A\right\rangle$ requires 2 complex numbers $\left(\alpha_u \text{ and } \alpha_d\right)$ as does Bob's state $\left|B\right\rangle$, equating to 8 real parameters for product state $\left|AB\right\rangle$. The normalization (6.4) reduces this to 6. Since phase has no physical significance, dof of the product state $\left|AB\right\rangle$ is reduced to 4:

> That is, $\left|A\right\rangle = \alpha_u\left|u\right\rangle + \alpha_d\left|d\right\rangle$ where $\alpha_u = r\,e^{i\theta}$ and $\alpha_d = s\,e^{i\phi}$ are both complex numbers. We can find an equivalent vector with one real and one complex coefficient, say $\left|A'\right\rangle = e^{-i\theta}\left|A\right\rangle$, thus using phase to reduce dof by one. In general we cannot convert both complex numbers to real. However, we can use phase to also reduce dof of $\left|B\right\rangle$ by one, for a total phase reduction of two.

The composite state vector $|\Psi\rangle$ has 4 complex parameters but only one normalization condition and one phase to ignore, for an equivalent of 6 real parameters; i.e., dof = 6:

$|\Psi\rangle = \psi_{uu}|uu\rangle + \psi_{ud}|ud\rangle + \psi_{du}|du\rangle + \psi_{dd}|dd\rangle$ where $\psi_{uu} = r_1 e^{i\theta_1}$, $\psi_{ud} = r_2 e^{i\theta_2}$, $\psi_{du} = r_3 e^{i\theta_3}$, $\psi_{dd} = r_4 e^{i\theta_4}$ are all complex numbers. We can find an equivalent vector with one real coefficient and three complex coefficients, say $|\Psi'\rangle = e^{-i\theta_1}|\Psi\rangle$, thus using phase to reduce the dof by one. That is the best we can do with phase. We cannot force more than one coefficient to be real.

Thus, the state space S_{AB} is richer than just the set of product states. The difference is called **entanglement.** An entangled state $|\Psi\rangle$ is a complete description of the combined system. However, there are degrees of entanglement.

Exercise 6.4. Using the matrix expressions (3.20) for σ_x, σ_y, and σ_z and the vector expressions for $|u\}$ (2.11) and $|d\rangle$ (2.12) show that Alice's single-spin states are

$$\sigma_z|u\} = |u\}, \quad \sigma_z|d\} = -|d\}, \quad \sigma_x|u\} = |d\},$$
$$\sigma_x|d\} = |u\}, \quad \sigma_y|u\} = i|d\}, \quad \sigma_y|d\} = -i|u\}.$$

Bob's states are similar:

$$\tau_z|u\rangle = |u\rangle, \quad \tau_z|d\rangle = -|d\rangle, \quad \tau_x|u\rangle = |d\rangle$$
$$\tau_x|d\rangle = |u\rangle, \quad \tau_y|u\rangle = i|d\rangle, \quad \tau_y|d\rangle = -i|u\rangle$$

We can calculate how Alice's and Bob's operators act on the product basis vectors by remembering that Alice's *s* operators act only on the first element and Bob's act only on the 2nd. Examples are

$\sigma_z|du\rangle = -|du\rangle$ and $\tau_z|du\rangle = |du\rangle$.

By way of explanation, the first equation could be written

$$(\sigma_z \otimes \boldsymbol{I})\left(|d\} \otimes |u\rangle\right) = \sigma_z|d\} \otimes \boldsymbol{I}|u\rangle = -d\} \otimes |u\rangle.$$

All of the spin operator results are tabulated in the Appendix Tables 1-3, below.

TABLE 1. Up-Down Basis

	2-Spin Eigenvectors			
	$\|uu\rangle$	$\|ud\rangle$	$\|du\rangle$	$\|dd\rangle$
σ_z	$\|uu\rangle$	$\|ud\rangle$	$-\|du\rangle$	$-\|dd\rangle$
σ_x	$\|du\rangle$	$\|dd\rangle$	$\|uu\rangle$	$\|ud\rangle$
σ_y	$i\|du\rangle$	$i\|dd\rangle$	$-i\|uu\rangle$	$-i\|ud\rangle$
τ_z	$\|uu\rangle$	$-\|ud\rangle$	$\|du\rangle$	$-\|dd\rangle$
τ_x	$\|ud\rangle$	$\|uu\rangle$	$\|dd\rangle$	$\|du\rangle$
τ_y	$i\|ud\rangle$	$-i\|uu\rangle$	$i\|dd\rangle$	$-i\|du\rangle$

TABLE 2. Right-Left Basis

	2-Spin Eigenvectors			
	$\|rr\rangle$	$\|r\,\ell\rangle$	$\|\ell\,r\rangle$	$\|\ell\ell\rangle$
σ_z	$\|\ell r\rangle$	$\|\ell\ell\rangle$	$\|r\,r\rangle$	$\|r\ell\rangle$
σ_x	$\|rr\rangle$	$\|r\,\ell\rangle$	$-\|\ell\,r\rangle$	$-\|\ell\ell\rangle$
σ_y	$-i\|\ell r\rangle$	$-i\|\ell\ell\rangle$	$i\|r\,r\rangle$	$i\|r\ell\rangle$
τ_z	$\|r\ell\rangle$	$\|r\,r\rangle$	$\|\ell\ell\rangle$	$\|\ell r\rangle$
τ_x	$\|rr\rangle$	$-\|r\,\ell\rangle$	$\|\ell r\rangle$	$-\|\ell\ell\rangle$
τ_y	$-i\|r\ell\rangle$	$i\|r\,r\rangle$	$-i\|\ell\ell\rangle$	$i\|r\ell\rangle$

TABLE 3. In-Out Basis

	2-Spin Eigenvectors			
	$\|i\,i\rangle$	$\|i\,o\rangle$	$\|o\,i\rangle$	$\|oo\rangle$
σ_z	$\|oi\rangle$	$\|oo\rangle$	$\|i\,i\rangle$	$\|i\,o\rangle$
σ_x	$i\|oi\rangle$	$i\|oo\rangle$	$-\|i\,i\rangle$	$-\|i\,o\rangle$
σ_y	$\|i\,i\rangle$	$\|i\,o\rangle$	$-\|o\,i\rangle$	$-\|oo\rangle$
τ_z	$\|io\rangle$	$\|i\,i\rangle$	$\|oo\rangle$	$\|o\,i\rangle$
τ_x	$i\|io\rangle$	$-i\|i\,i\rangle$	$i\|oo\rangle$	$-i\|o\,i\rangle$
τ_y	$\|i\,i\rangle$	$-\|io\rangle$	$\|o\,i\rangle$	$-\|oo\rangle$

Let $|\Psi\rangle$ be a normalized composite state and L an observable in S_{AB}. Then $\langle L\rangle \overset{(4.13)}{=} \langle\Psi|L|\Psi\rangle$. If $L = L_A \otimes I$ is the product observable that represent Alice's observable L_A, then

$$\langle L_A \otimes I\rangle \overset{(4.13)}{=} \langle\Psi|L_A \otimes I|\Psi\rangle$$

Since we know that L_A operates only on Alice's part of $|\Psi\rangle = \sum_{ab} \psi_{ab}|ab\rangle$, we will usually shorten this expression to

$$\langle L_A\rangle = \langle\Psi|L_A|\Psi\rangle. \tag{6.09}$$

For example, $\langle\sigma_z\rangle = \langle \text{sing}|\sigma_z|\text{sing}\rangle$ means $\langle\sigma_z\otimes I\rangle = \langle \text{sing}|\sigma_z\otimes I|\text{sing}\rangle$.

Exercise 6.5 asks us to prove that every prediction about Alice's half of the system is the same as it would be in the single-state theory, and similarly for Bob. Thus, in the 2-state product system the following still holds:

$$\langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 = 1 \tag{6.11}$$

This means that the Spin-Polarization Principle still holds for the product system: there is some state for which the spin is +1. We claim that this principle does not hold for the entangled state $|\text{sing}\rangle$:

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

$$\overset{(\text{Table 1})}{=} \frac{1}{2}\left(\langle ud| - \langle du|\right)\left(|ud\rangle + |du\rangle\right)$$

$$= \frac{1}{2}\left(\langle ud|ud\rangle + \langle ud|du\rangle - \langle du|ud\rangle - \langle du|du\rangle\right) = (1+0-0-1) = 0.$$

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

$$\overset{(\text{Table 1})}{=} \frac{1}{2}\left(\langle ud| - \langle du|\right)\left(|dd\rangle - |uu\rangle\right)$$

$$= \frac{1}{2}\left[\langle ud|dd\rangle - \langle ud|uu\rangle - \langle du|dd\rangle + \langle du|uu\rangle\right]$$

$$= \frac{1}{2}\left(\cancel{\langle u|d\rangle}^{0}\langle d|d\rangle - \langle u|u\rangle\cancel{\langle d|u\rangle}^{0} - \langle d|d\rangle\cancel{\langle u|d\rangle}^{0} + \cancel{\langle d|u\rangle}^{0}\langle u|u\rangle\right)$$

$$= 0.$$

Similarly, $\langle\sigma_y\rangle = 0$, and so (6.11) is not true. ✔

$\langle\sigma_x\rangle = \langle\sigma_y\rangle = \langle\sigma_z\rangle = 0$ means that +1 and -1 are equally likely for each operator. Thus, the individual outcomes are completely uncertain even though we know the exact state of vector $|\text{sing}\rangle$.

Exercise 6.9. The operator $\vec{\sigma}\otimes\vec{\tau} = \sigma_x\tau_x + \sigma_y\tau_y + \sigma_z\tau_z$ has eigenvectors $|\text{sing}\rangle$, $|T_1\rangle$, $|T_2\rangle$, and $|T_3\rangle$ with eigenvalues -3, +1, +1, and +1, respectively. That is, +1 is an eigenvalue with degeneracy 3 which is why $|T_1\rangle$, $|T_2\rangle$, and $|T_3\rangle$ are called triplets and $|\text{sing}\rangle$ a singlet.

Theorem. Let u = x, y, or z and v = x, y, or z. For $\left|\Psi\right\rangle = \left|\text{sing}\right\rangle$

$$\left\langle \sigma_u \tau_v \right\rangle = \begin{cases} -1 & \text{if } u = v \\ 0 & \text{otherwise} \end{cases} \tag{6.12}$$

That is,

$$\left\langle \sigma_x \tau_x \right\rangle = \left\langle \sigma_y \tau_y \right\rangle = \left\langle \sigma_z \tau_z \right\rangle = -1$$

and

$$\left\langle \sigma_x \tau_y \right\rangle = \left\langle \sigma_x \tau_z \right\rangle = \left\langle \sigma_y \tau_x \right\rangle = \left\langle \sigma_y \tau_z \right\rangle = \left\langle \sigma_z \tau_x \right\rangle = \left\langle \sigma_z \tau_y \right\rangle = 0 .$$

Proof: $\left\langle \sigma_x \tau_x \right\rangle = \left\langle \sigma_y \tau_y \right\rangle = \left\langle \sigma_z \tau_z \right\rangle = -1$ was shown during the proof of Exercise 6.9. We compute just one mixed case here since the others are similar.

$$\begin{aligned}
\left\langle \sigma_x \tau_y \right\rangle &= \left\langle \text{sing} \middle| \sigma_x \tau_y \middle| \text{sing} \right\rangle = \left\langle \text{sing} \right| \left[\sigma_x \tau_y \frac{1}{\sqrt{2}} \left(\left|ud\right\rangle - \left|du\right\rangle \right) \right] \\
&\overset{\text{(Table 1)}}{=} \left\langle \text{sing} \right| \left[\sigma_x \frac{1}{\sqrt{2}} \left(-i\left|uu\right\rangle - i\left|dd\right\rangle \right) \right] \\
&\overset{\text{(Table 1)}}{=} \frac{1}{2} \left(\left\langle ud \right| - \left\langle du \right| \right) \left(-i\left|du\right\rangle - i\left|ud\right\rangle \right) \\
&= -\frac{i}{2} \left(\cancelto{0}{\left\langle ud \middle| du \right\rangle} + \cancelto{1}{\left\langle ud \middle| ud \right\rangle} - \cancelto{1}{\left\langle du \middle| du \right\rangle} - \cancelto{0}{\left\langle du \middle| ud \right\rangle} \right) \\
&= -\frac{i}{2} (1 - 1) = 0
\end{aligned}$$

■

Example. Show $\left|\text{sing}\right\rangle$ represents two particles entangled with opposite spins.

Solution. Suppose that Alice and Bob measure in the $\hat{n} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$ direction.

$$(\vec{\sigma}\vec{\tau})_n \equiv \sigma_n \tau_n \overset{(3.22)}{=} (\vec{\sigma} \bullet \hat{n})(\vec{\tau} \bullet \hat{n}) = (n_x \sigma_x + n_y \sigma_y + n_z \sigma_z)(n_x \tau_x + n_y \tau_y + n_z \tau_z)$$
$$\sum_{u=x,y,z} \; \sum_{v=x,y,z} n_u n_v \sigma_u \tau_v .$$

$$\left\langle (\vec{\sigma}\vec{\tau})_n \right\rangle = \sum_{u=x,y,z} \; \sum_{v=x,y,z} n_u n_v \left\langle \sigma_u \tau_v \right\rangle \overset{(6.12)}{=} -\left(n_x^2 + n_y^2 + n_z^2 \right) = -1.$$

Thus, whenever Alice measures +1 spin then Bob measures -1, and vice-versa

■

Chapter 7. More on Entanglement

Thus far we have defined the tensor product of two vectors, say $|A\rangle = \sum_{a=1}^{n} \alpha_a |a\rangle$ and $|B\rangle = \sum_{b=1}^{n} \beta_b |b\rangle$, as $|A\rangle \otimes |B\rangle = \sum_{a,b=1}^{n} \alpha_a \beta_b |ab\rangle$. For $n = 2$, $|A\rangle$ and $|B\rangle$ have 2 components each, so $|A\rangle \otimes |B\rangle$ has 4 components: $\alpha_1\beta_1$, $\alpha_1\beta_2$, $\alpha_2\beta_1$ and $\alpha_2\beta_2$. We know how to write $|A\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ and $|B\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ as arrays, and we have defined $|A\rangle \otimes |B\rangle$, but we haven't yet specified how to write $|A\rangle \otimes |B\rangle$ as an array. Moreover, we have not yet defined the tensor product $L \otimes M$ of two matrices nor how to express it as an array. We do these things now for 2-dimensional systems S_A and S_B. The definitions and notation easily extend to n dimensions.

Definition. Let Alice's and Bob's systems S_A and S_B have respective bases $\{|a_1\rangle, |a_2\rangle\}$ and $\{|b_1\rangle, |b_2\rangle\}$. Then S_{AB} has composite basis $\{ |a_1 b_1\rangle, |a_1 b_2\rangle, |a_2 b_1\rangle, |a_2 b_2\rangle \}$. Let $|A\rangle = \alpha_1|a_1\rangle + \alpha_2|a_2\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ and $|B\rangle = \beta_1|b_1\rangle + \beta_2|b_2\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ be vectors, and $L = \begin{pmatrix} \ell_{11} & \ell_{12} \\ \ell_{21} & \ell_{22} \end{pmatrix}$ and $M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$ matrices expressed in terms of the composite basis. We define the **tensor product** $\boldsymbol{L \otimes M}$ by its action on Alice's and Bob's vectors $|A\rangle$ and $|B\rangle$:

$$\boldsymbol{L \otimes M} \ |\boldsymbol{A}\rangle \otimes |\boldsymbol{B}\rangle \equiv \boldsymbol{L}|\boldsymbol{A}\rangle \otimes \boldsymbol{M}|\boldsymbol{B}\rangle. \tag{7.10}$$

In terms of matrix multiplication, we know how to express

$$L|A\rangle = \begin{pmatrix} \ell_{11} & \ell_{12} \\ \ell_{21} & \ell_{22} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \ell_{11}\alpha_1 + \ell_{12}\alpha_2 \\ \ell_{21}\alpha_1 + \ell_{22}\alpha_2 \end{pmatrix} \tag{I}$$

and

$$M|B\rangle = \begin{pmatrix} m_{11}\beta_1 + m_{12}\beta_2 \\ m_{21}\beta_1 + m_{22}\beta_2 \end{pmatrix} \tag{II}$$

but not yet how to express $L \otimes M$ nor $L|A\rangle \otimes M|B\rangle$. We define these now using a pattern definition.

$$\text{Definition. } L \otimes M = \begin{pmatrix} \ell_{11} & \ell_{12} \\ \ell_{21} & \ell_{22} \end{pmatrix} \otimes M = \begin{pmatrix} \ell_{11}M & \ell_{12}M \\ \ell_{21}M & \ell_{22}M \end{pmatrix} \tag{7.6}$$

$$L \otimes M = \begin{pmatrix} \ell_{11}m_{11} & \ell_{11}m_{12} & | & \ell_{12}m_{11} & \ell_{12}m_{12} \\ \ell_{11}m_{21} & \ell_{11}m_{22} & | & \ell_{12}m_{21} & \ell_{12}m_{22} \\ --- & --- & + & --- & --- \\ \ell_{21}m_{11} & \ell_{21}m_{12} & | & \ell_{22}m_{11} & \ell_{22}m_{12} \\ \ell_{21}m_{21} & \ell_{21}m_{22} & | & \ell_{22}m_{21} & \ell_{22}m_{22} \end{pmatrix} \tag{7.7}$$

This pattern definition can be extended to any size array, larger or smaller. In particular, replacing matrix L by vector $|A\rangle$ and matrix M by vector $|B\rangle$ gives

$$|A\rangle \otimes |B\rangle \overset{(7.6)}{=} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \otimes |B\rangle \overset{(7.6)}{=} \begin{pmatrix} \alpha_1 |B\rangle \\ \alpha_2 |B\rangle \end{pmatrix} \overset{(7.7)}{=} \begin{pmatrix} \alpha_1\beta_1 \\ \alpha_1\beta_2 \\ -- \\ \alpha_2\beta_1 \\ \alpha_2\beta_2 \end{pmatrix}. \tag{7.8}$$

Exercise 7.3 (below) proves that the matrix pattern definition (7.7) is in agreement with the tensor definition (7.10). It shows that the matrix expressions for LHS and RHS of (7.10) are equal.

Exercise 7.3. Show in matrix notation that $L \otimes M \; |A\rangle \otimes |B\rangle = L|A\rangle \otimes M|B\rangle$.

Solution. Multiplying the right-hand sides of (7.7) and (7.8) yields

$$L \otimes M |A\rangle \otimes |B\rangle = \begin{pmatrix} \ell_{11}m_{11}\alpha_1\beta_1 + \ell_{11}m_{12}\alpha_1\beta_2 + \ell_{12}m_{11}\alpha_2\beta_1 + \ell_{12}m_{12}\alpha_2\beta_2 \\ \ell_{11}m_{21}\alpha_1\beta_1 + \ell_{11}m_{22}\alpha_1\beta_2 + \ell_{12}m_{21}\alpha_2\beta_1 + \ell_{12}m_{22}\alpha_2\beta_2 \\ \ell_{21}m_{11}\alpha_1\beta_1 + \ell_{21}m_{12}\alpha_1\beta_2 + \ell_{22}m_{11}\alpha_2\beta_1 + \ell_{22}m_{12}\alpha_2\beta_2 \\ \ell_{21}m_{21}\alpha_1\beta_1 + \ell_{21}m_{22}\alpha_1\beta_2 + \ell_{22}m_{21}\alpha_2\beta_1 + \ell_{22}m_{22}\alpha_2\beta_2 \end{pmatrix}. \tag{a}$$

Replacing $|A\rangle$ by $L|A\rangle$ and $|B\rangle$ by $M|B\rangle$ in the pattern definition (7.8) yields

$$L|A\rangle \otimes M|B\rangle \overset{(\mathrm{I})}{=} \begin{pmatrix} \ell_{11}\alpha_1 + \ell_{12}\alpha_2 \\ \ell_{21}\alpha_1 + \ell_{22}\alpha_2 \end{pmatrix} \otimes M|B\rangle \overset{(7.8)}{=} \begin{pmatrix} (\ell_{11}\alpha_1 + \ell_{12}\alpha_2)M|B\rangle \\ (\ell_{21}\alpha_1 + \ell_{22}\alpha_2)M|B\rangle \end{pmatrix}$$

$$\overset{(\mathrm{ii})}{=} \begin{pmatrix} (\ell_{11}\alpha_1 + \ell_{12}\alpha_2)(m_{11}\beta_1 + m_{12}\beta_2) \\ (\ell_{11}\alpha_1 + \ell_{12}\alpha_2)(m_{21}\beta_1 + m_{22}\beta_2) \\ ----------------- \\ (\ell_{21}\alpha_1 + \ell_{22}\alpha_2)(m_{11}\beta_1 + m_{12}\beta_2) \\ (\ell_{21}\alpha_1 + \ell_{22}\alpha_2)(m_{21}\beta_1 + m_{22}\beta_2) \end{pmatrix}$$

$$= \begin{pmatrix} \ell_{11}\alpha_1 m_{11}\beta_1 + \ell_{11}\alpha_1 m_{12}\beta_2 + \ell_{12}\alpha_2 m_{11}\beta_1 + \ell_{12}\alpha_2 m_{12}\beta_2 \\ \ell_{11}\alpha_1 m_{21}\beta_1 + \ell_{11}\alpha_1 m_{22}\beta_2 + \ell_{12}\alpha_2 m_{21}\beta_1 + \ell_{12}\alpha_2 m_{22}\beta_2 \\ \ell_{21}\alpha_1 m_{11}\beta_1 + \ell_{21}\alpha_1 m_{12}\beta_2 + \ell_{22}\alpha_2 m_{11}\beta_1 + \ell_{22}\alpha_2 m_{12}\beta_2 \\ \ell_{21}\alpha_1 m_{21}\beta_1 + \ell_{21}\alpha_1 m_{22}\beta_2 + \ell_{22}\alpha_2 m_{21}\beta_1 + \ell_{22}\alpha_2 m_{22}\beta_2 \end{pmatrix}$$

$$= \begin{pmatrix} \ell_{11}m_{11}\alpha_1\beta_1 + \ell_{11}m_{12}\alpha_1\beta_2 + \ell_{12}m_{11}\alpha_2\beta_1 + \ell_{12}m_{12}\alpha_2\beta_2 \\ \ell_{11}m_{21}\alpha_1\beta_1 + \ell_{11}m_{22}\alpha_1\beta_2 + \ell_{12}m_{21}\alpha_2\beta_1 + \ell_{12}m_{22}\alpha_2\beta_2 \\ \ell_{21}m_{11}\alpha_1\beta_1 + \ell_{21}m_{12}\alpha_1\beta_2 + \ell_{22}m_{11}\alpha_2\beta_1 + \ell_{22}m_{12}\alpha_2\beta_2 \\ \ell_{21}m_{21}\alpha_1\beta_1 + \ell_{21}m_{22}\alpha_1\beta_2 + \ell_{22}m_{21}\alpha_2\beta_1 + \ell_{22}m_{22}\alpha_2\beta_2 \end{pmatrix}$$

$$\overset{(\mathrm{a})}{=} L|A\rangle \otimes M|B\rangle \quad ✔ \qquad ■$$

In a 2-spin system, an operator is represented by a 4x4 matrix with rows $\langle uu|$, $\langle ud|$, $\langle du|$, and $\langle dd|$ and columns $|uu\rangle$, $|ud\rangle$, $|du\rangle$, and $|dd\rangle$. The next example shows how to use the tensor pattern definition (7.7) to compute the matrix of a 2-spin operator and then confirms that the answer agrees with the inner product definition (6.05).

Example. Compute $\sigma_z \otimes \boldsymbol{I}$ using both the matrix (7.7) and inner-product (6.05) definitions to confirm that they are consistent. Then show that $\sigma_z \otimes \boldsymbol{I}$ acts only on Alice's vector components while leaving Bob's vector components alone.

Solution.

$$\sigma_z \overset{(3.20)}{=} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{(III)}$$

and

$$\boldsymbol{I} = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{IV}$$

Thus,

$$\sigma_z \otimes \boldsymbol{I} \overset{(7.6)}{=} \begin{pmatrix} \sigma_{11}\boldsymbol{I} & \sigma_{12}\boldsymbol{I} \\ \sigma_{21}\boldsymbol{I} & \sigma_{22}\boldsymbol{I} \end{pmatrix}$$

$$\sigma_z \otimes \boldsymbol{I} \overset{(7.7)}{=} \begin{pmatrix} \sigma_{11}I_{11} & \sigma_{11}I_{12} & | & \sigma_{12}I_{11} & \sigma_{12}I_{12} \\ \sigma_{11}I_{21} & \sigma_{11}I_{22} & | & \sigma_{12}I_{21} & \sigma_{12}I_{22} \\ - & - & + & - & - \\ \sigma_{21}I_{11} & \sigma_{21}I_{12} & | & \sigma_{22}I_{11} & \sigma_{22}I_{12} \\ \sigma_{21}I_{21} & \sigma_{21}I_{22} & | & \sigma_{22}I_{21} & \sigma_{22}I_{22} \end{pmatrix}$$

$$\overset{(\text{III, IV})}{=} \begin{pmatrix} 1\bullet 1 & 1\bullet 0 & | & 0 & 0 \\ 1\bullet 0 & 1\bullet 1 & | & 0 & 0 \\ - & - & + & - & - \\ 0 & 0 & | & -1\bullet 1 & -1\bullet 0 \\ 0 & 0 & | & -1\bullet 0 & -1\bullet 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Recall the inner product definition (6.05) for an element of a linear operator M : $m_{a'b',ab} = \langle a'b' | m | ab \rangle$. Letting $m = \sigma_z \otimes I$ yields

$$\sigma_z \otimes \boldsymbol{I} = \begin{pmatrix} \langle uu|\sigma_z \boldsymbol{I}|uu\rangle & \langle uu|\sigma_z \boldsymbol{I}|ud\rangle & \langle uu|\sigma_z \boldsymbol{I}|du\rangle & \langle uu|\sigma_z \boldsymbol{I}|dd\rangle \\ \langle ud|\sigma_z \boldsymbol{I}|uu\rangle & \langle ud|\sigma_z \boldsymbol{I}|ud\rangle & \langle ud|\sigma_z \boldsymbol{I}|du\rangle & \langle ud|\sigma_z \boldsymbol{I}|dd\rangle \\ \langle du|\sigma_z \boldsymbol{I}|uu\rangle & \langle du|\sigma_z \boldsymbol{I}|ud\rangle & \langle du|\sigma_z \boldsymbol{I}|du\rangle & \langle du|\sigma_z \boldsymbol{I}|dd\rangle \\ \langle dd|\sigma_z \boldsymbol{I}|uu\rangle & \langle dd|\sigma_z \boldsymbol{I}|ud\rangle & \langle dd|\sigma_z \boldsymbol{I}|du\rangle & \langle dd|\sigma_z \boldsymbol{I}|dd\rangle \end{pmatrix} \tag{7.2}$$

$$= \begin{pmatrix} \left(\langle uu|\sigma_z\right)\left(\boldsymbol{I}|uu\rangle\right) & \left(\langle uu|\sigma_z\right)\left(\boldsymbol{I}|ud\rangle\right) & \left(\langle uu|\sigma_z\right)\left(\boldsymbol{I}|du\rangle\right) & \left(\langle uu|\sigma_z\right)\left(\boldsymbol{I}|dd\rangle\right) \\ \left(\langle ud|\sigma_z\right)\left(\boldsymbol{I}|uu\rangle\right) & \left(\langle ud|\sigma_z\right)\left(\boldsymbol{I}|ud\rangle\right) & \left(\langle ud|\sigma_z\right)\left(\boldsymbol{I}|du\rangle\right) & \left(\langle ud|\sigma_z\right)\left(\boldsymbol{I}|dd\rangle\right) \\ \left(\langle du|\sigma_z\right)\left(\boldsymbol{I}|uu\rangle\right) & \left(\langle du|\sigma_z\right)\left(\boldsymbol{I}|ud\rangle\right) & \left(\langle du|\sigma_z\right)\left(\boldsymbol{I}|du\rangle\right) & \left(\langle du|\sigma_z\right)\left(\boldsymbol{I}|dd\rangle\right) \\ \left(\langle dd|\sigma_z\right)\left(\boldsymbol{I}|uu\rangle\right) & \left(\langle dd|\sigma_z\right)\left(\boldsymbol{I}|ud\rangle\right) & \left(\langle dd|\sigma_z\right)\left(\boldsymbol{I}|du\rangle\right) & \left(\langle dd|\sigma_z\right)\left(\boldsymbol{I}|dd\rangle\right) \end{pmatrix}$$

$$\overset{\text{(Table 1)}}{=} \begin{pmatrix} \langle uu|uu\rangle & \langle uu|ud\rangle & \langle uu|du\rangle & \langle uu|dd\rangle \\ \langle ud|uu\rangle & \langle ud|ud\rangle & \langle ud|du\rangle & \langle ud|dd\rangle \\ -\langle du|uu\rangle & -\langle du|ud\rangle & -\langle du|du\rangle & -\langle du|dd\rangle \\ -\langle dd|uu\rangle & -\langle dd|ud\rangle & -\langle dd|du\rangle & -\langle dd|dd\rangle \end{pmatrix}$$

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

We next confirm that $\sigma_z \otimes I$ acts only on Alice's vector components and that Bob's components are unchanged.

Since $|u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|d\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, the matrix pattern definition yields

$$|uu\rangle = |u\rangle \otimes |u\rangle = \begin{pmatrix} 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |ud\rangle = |u\rangle \otimes |d\rangle = \begin{pmatrix} 1\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$|du\rangle = |d\rangle \otimes |u\rangle = \begin{pmatrix} 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |dd\rangle = |d\rangle \otimes |d\rangle = \begin{pmatrix} 0\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 1\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad (7.9)$$

So

$$\sigma_z \otimes \boldsymbol{I}|uu\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = |uu\rangle = \sigma_z|u\rangle \otimes |u\rangle \quad ✔$$

$$\sigma_z \otimes \boldsymbol{I}|ud\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = |ud\rangle = \sigma_z|u\rangle \otimes |d\rangle \quad ✔$$

$$\sigma_z \otimes \boldsymbol{I}\left|du\right\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} = -\left|du\right\rangle = \sigma_z\left|d\right\rangle \otimes \left|u\right\rangle \quad ✔$$

$$\sigma_z \otimes \boldsymbol{I}\left|dd\right\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} = -\left|dd\right\rangle = \sigma_z\left|d\right\rangle \otimes \left|d\right\rangle \quad ■$$

Exercise 7.1 is similar, demonstrating that Alice's half of the state vector $\boldsymbol{I} \otimes \tau_x$ is unchanged while τ_x works on Bob's half.

In the next example we show that the tensor and inner product definitions agree for a more complex tensor product.

Exercise 7.2. Compute $\sigma_z \otimes \tau_x$ using the tensor and inner product definitions. The tensor definition yields

$$\sigma_z \otimes \tau_x \overset{(3.20)}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & | & 0 & 0 \\ 1 & 0 & | & 0 & 0 \\ - & - & + & - & - \\ 0 & 0 & | & 0 & -1 \\ 0 & 0 & | & -1 & 0 \end{pmatrix}.$$

The inner product definition $m_{a'b',ab} \overset{(6.05)}{=} \left\langle a'b'\right|m\left|ab\right\rangle$ yields

$$\sigma_z \otimes \tau_x = \begin{pmatrix} \left\langle uu\right|\sigma_z\tau_x\left|uu\right\rangle & \left\langle uu\right|\sigma_z\tau_x\left|ud\right\rangle & \left\langle uu\right|\sigma_z\tau_x\left|du\right\rangle & \left\langle uu\right|\sigma_z\tau_x\left|dd\right\rangle \\ \left\langle ud\right|\sigma_z\tau_x\left|uu\right\rangle & \left\langle ud\right|\sigma_z\tau_x\left|ud\right\rangle & \left\langle ud\right|\sigma_z\tau_x\left|du\right\rangle & \left\langle ud\right|\sigma_z\tau_x\left|dd\right\rangle \\ \left\langle du\right|\sigma_z\tau_x\left|uu\right\rangle & \left\langle du\right|\sigma_z\tau_x\left|ud\right\rangle & \left\langle du\right|\sigma_z\tau_x\left|du\right\rangle & \left\langle du\right|\sigma_z\tau_x\left|dd\right\rangle \\ \left\langle dd\right|\sigma_z\tau_x\left|uu\right\rangle & \left\langle dd\right|\sigma_z\tau_x\left|ud\right\rangle & \left\langle dd\right|\sigma_z\tau_x\left|du\right\rangle & \left\langle dd\right|\sigma_z\tau_x\left|dd\right\rangle \end{pmatrix}$$

$$\overset{(\text{Table 1})}{=} \begin{pmatrix} \left\langle uu|ud\right\rangle & \left\langle uu|uu\right\rangle & \left\langle uu|dd\right\rangle & \left\langle uu|du\right\rangle \\ \left\langle ud|ud\right\rangle & \left\langle ud|uu\right\rangle & \left\langle ud|dd\right\rangle & \left\langle ud|du\right\rangle \\ -\left\langle du|ud\right\rangle & -\left\langle du|uu\right\rangle & -\left\langle du|dd\right\rangle & -\left\langle du|du\right\rangle \\ -\left\langle dd|ud\right\rangle & -\left\langle dd|uu\right\rangle & -\left\langle dd|dd\right\rangle & -\left\langle dd|du\right\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

■

Definition. Given a bra $\langle\Phi|$ and a ket $|\Psi\rangle$, their **outer product** is the linear operator $|\boldsymbol{\Psi}\rangle\langle\boldsymbol{\Phi}|$ defined by its **action on ket vectors** $|A\rangle$:

$$|\boldsymbol{\Psi}\rangle\langle\boldsymbol{\Phi}|\;|A\rangle \equiv |\Psi\rangle\;\langle\Phi|A\rangle. \tag{7.01}$$

Applying the Bra-ket interchange rule to (7.01) yields

$$\langle A|\;|\Phi\rangle\langle\Psi| = \langle A|\Phi\rangle\;\langle\Psi|.$$

Exchanging Φ and Ψ, we generate the outer product **action on bra vectors**:

$$\langle A|\;|\boldsymbol{\Psi}\rangle\langle\boldsymbol{\Phi}| \equiv \langle A|\Psi\rangle\;\langle\Phi|. \tag{7.02}$$

Notation. Suppose $|\Psi\rangle = \begin{pmatrix}\psi_1\\ \vdots \\ \psi_n\end{pmatrix}$ and $|\Phi\rangle = \begin{pmatrix}\phi_1\\ \vdots \\ \phi_n\end{pmatrix}$. To write the outer product linear operator $L = |\Psi\rangle\langle\Phi|$ as a matrix, set $L = (\ell_{ij})$. Then

$$\ell_{ij} = \psi_i\phi_j^* : \tag{7.03}$$

$$\ell_{ij} \overset{(3.01)}{=} \langle i|L|j\rangle \overset{(3.05)}{=} \langle i|\big(L|j\rangle\big) = \langle i|\Big[\big(|\Psi\rangle\langle\Phi|\big)|j\rangle\Big] \overset{(7.01)}{=} \langle i|\;\big(|\Psi\rangle\;\langle\Phi|j\rangle\big)$$

$$= \langle\Phi|j\rangle\;\langle i|\big(|\Psi\rangle\big) = \langle i|\Psi\rangle\;\langle\Phi|j\rangle \overset{(1.5,\ 1.5b)}{=} \psi_i\phi_j^*$$

Thus,

$$|\Psi\rangle\langle\Phi| = \begin{pmatrix}\psi_1\phi_1^* & \cdots & \psi_1\phi_n^* \\ \vdots & & \vdots \\ \psi_n\phi_1^* & \cdots & \psi_n\phi_n^*\end{pmatrix}. \tag{7.04}$$

Observe that the pattern approach used to write tensors in matrix form also works for outer products. Since $\langle\Phi| = (\phi_1^* \cdots \phi_n^*)$,

$$|\Psi\rangle\langle\Phi| = \begin{pmatrix}\psi_1\\ \vdots \\ \psi_n\end{pmatrix}\langle\Phi| = \begin{pmatrix}\psi_1\Phi\\ \vdots \\ \psi_n\Phi\end{pmatrix} = \begin{pmatrix}\psi_1\phi_1^* & \cdots & \psi_1\phi_n^* \\ \vdots & & \vdots \\ \psi_n\phi_1^* & \cdots & \psi_n\phi_n^*\end{pmatrix}. \quad ✔$$

Example. Confirm that the pattern expression (7.04) and the tensor-product definition (7.01) for outer product are in agreement.

Solution. Let $\left|A\right\rangle = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$. Then

$$\left|\Psi\right\rangle\left\langle\Phi\right| \ \left|A\right\rangle \overset{(7.01)}{=} \left\langle\Phi\middle|A\right\rangle \left|\Psi\right\rangle \overset{(1.2)}{=} \sum_{j=1}^{n} \phi_j^{\,*}\alpha_j \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} = \begin{pmatrix} \psi_1 \sum_j \phi_j^{\,*}\alpha_j \\ \vdots \\ \psi_n \sum_j \phi_j^{\,*}\alpha_j \end{pmatrix}$$

and

$$\left|\Psi\right\rangle\left\langle\Phi\right| \ \left|A\right\rangle \overset{(7.04)}{=} \begin{pmatrix} \psi_1\phi_1^{\,*} & \cdots & \psi_1\phi_n^{\,*} \\ \vdots & & \vdots \\ \psi_n\phi_1^{\,*} & \cdots & \psi_n\phi_n^{\,*} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \psi_1 \sum_j \phi_j^{\,*}\alpha_j \\ \vdots \\ \psi_n \sum_j \phi_j^{\,*}\alpha_j \end{pmatrix} \quad \blacksquare$$

Definition. If $\left|\Psi\right\rangle$ is a normalized vector, $\left|\boldsymbol{\Psi}\right\rangle\left\langle\boldsymbol{\Psi}\right|$ is called **a projection operator**.

Observe that if $\left|A\right\rangle$ is a vector then $\left|\Psi\right\rangle\left\langle\Psi\right| \ \left|A\right\rangle = \left|\Psi\right\rangle \ \left\langle\Psi\middle|A\right\rangle$ is a vector proportional to $\left|\Psi\right\rangle$. We say $\left|\boldsymbol{\Psi}\right\rangle\left\langle\boldsymbol{\Psi}\right|$ **projects** $\left|\boldsymbol{A}\right\rangle$ **onto** $\left|\boldsymbol{\Psi}\right\rangle$**.**

Definition. Let L be an operator. The **trace of** $\boldsymbol{L}$ is

$$\mathbf{Tr}\, \boldsymbol{L} = \sum_i \left\langle i\middle|L\middle|i\right\rangle, \tag{7.05}$$

the sum of the diagonal elements.

Theorem 7.1. Let $L = \left|\Psi\right\rangle\left\langle\Psi\right|$ be a projection operator.

(a) L is Hermitian

(b) $\left|\Psi\right\rangle$ is an eigenvector of L with eigenvalue +1

(c) If $\left|\Phi\right\rangle \perp \left|\Psi\right\rangle$ then $\left|\Phi\right\rangle$ is an eigenvector of L with eigenvalue 0.

(d) The only eigenvector with unit eigenvalue is $\left|\Psi\right\rangle$.

(e) $L^2 = L$.

(f) $\mathrm{Tr}\, L = 1$.

(g) $\sum_i \left|i\right\rangle\left\langle i\right| = I$ (7.11)

(i.e., the sum of all projection operators for a given basis is I)

(h) $\langle \Psi | M | \Psi \rangle = \text{Tr}\left(|\Psi\rangle\langle\Psi| \; M \right)$ (7.12)

(the expected value of an operator M prepared in normalized state $|\Psi\rangle$ is the trace of the product of the projection operator and the operator M.)

Proof.

(a) By 7.03, $L = (\ell_{ij})$ where $\ell_{ij} = \psi_i \psi_j^*$. So $\ell_{ji} = \psi_j \psi_i^*$ and $\ell_{ji}^* = \psi_i \psi_j^* = \ell_{ij}$. ✔

(b) Since $\langle \Psi | \Psi \rangle = 1$, $L|\Psi\rangle = |\Psi\rangle \langle \Psi|\Psi\rangle = |\Psi\rangle$. ✔

(c) Since $\langle \Psi | \Phi \rangle = 0$, $L|\Phi\rangle = |\Psi\rangle \langle \Psi|\Phi\rangle = 0$. ✔

(d) If $|\Phi\rangle$ is a vector such that $L|\Phi\rangle = |\Phi\rangle$, then

$|\Phi\rangle = L|\Phi\rangle = |\Psi\rangle\langle\Psi| \; |\Phi\rangle = |\Psi\rangle \; \langle\Psi|\Phi\rangle \equiv \alpha|\Psi\rangle$ where $\alpha \in \mathbb{C}$. That is, $|\Psi\rangle$ and $|\Phi\rangle$ are the same eigenvector because any multiple of an eigenvector is considered to be the same eigenvector. ✔

(e) $L^2 = \left(|\Psi\rangle\langle\Psi| \right)\left(|\Psi\rangle\langle\Psi| \right) = |\Psi\rangle \; \langle\Psi|\Psi\rangle \; |\Psi\rangle = |\Psi\rangle\langle\Psi| = L$ since $|\Psi\rangle$ is normalized. ✔

(f) $\text{Tr}\, L = \sum_i \psi_i \psi_i^* = \langle \Psi | \Psi \rangle = 1$. ✔

(g) Let $|A\rangle = \sum_i \alpha_i |i\rangle$ be any vector. Then

$$\left(\sum_i |i\rangle\langle i| \right) |A\rangle = \sum_i \left(|i\rangle \; \langle i|A\rangle \right) = \sum_i \alpha_i |i\rangle = |A\rangle$$

and similarly $\langle A| \left(\sum_i |i\rangle\langle i| \right) = \langle A|$. Therefore $\sum_i |i\rangle\langle i| = I$. ✔

(h) Let $N = |\Psi\rangle\langle\Psi|M = (n_{ij})$. Then $n_{ij} = \langle i|N|j\rangle = \langle i|\Psi\rangle\langle\Psi|M|j\rangle$. Therefore

$$\text{Tr}\left(|\Psi\rangle\langle\Psi|M \right) = \text{Tr}\, N = \sum_i n_{ii} = \sum_i \langle i|\Psi\rangle\langle\Psi|M|i\rangle = \sum_i \langle\Psi|M|i\rangle\langle i|\Psi\rangle$$

$$= \langle\Psi|M \left(\sum_i |i\rangle\langle i| \right) |\Psi\rangle \overset{(g)}{=} \langle\Psi|M|\Psi\rangle$$ ■

Definition. Suppose there are several states $|\Psi_k\rangle$, $k = 1, \cdots, r$ and that Alice has prepared her system in state $|\Psi_k\rangle$ with probability $\text{P}(k)$ where $\sum_{k=1}^{r} \text{P}(k) = 1$. We define **Alice's density matrix** as the operator

$$\rho = \sum_{k=1}^{r} \mathbf{P}(\boldsymbol{k}) \left| \Psi_k \right\rangle \left\langle \Psi_k \right| \tag{7.12a}$$

and we say **Alice's system is mixed** or the **density matrix represents a mixed state**. When the density matrix represents a single state $\left|\Psi_A\right\rangle$, ρ is simply the projection operator,

$$\rho = \left|\Psi_A\right\rangle \left\langle \Psi_A \right| \tag{7.12b}$$

and we say that **Alice's system is pure**. Notice that "pure" can be considered a special case of "mixed" where r = 1. When we wish to exclude "pure" we will say that **Alice's system is entangled.** (The book does not make this distinction.)

Theorem 7.2. Let L_A be any observable of Alice's system and suppose her system is mixed.

(a) Show that Alice's expectation is

$$\left\langle L_A \right\rangle = \mathrm{Tr}\left(\rho L_A\right). \tag{7.13}$$

(b) Let $\left\{\left|a\right\rangle\right\}_{a=1}^{n_A}$ be a basis for the Hilbert space of Alice's states, and denote $L_A = \left(\ell_{a'a}\right)$ and $\rho = \left(\rho_{a'a}\right)$ in this basis. Show that $\left\langle L_A \right\rangle$ can be expressed

$$\left\langle L_A \right\rangle = \sum_{a'=1}^{n_A} \sum_{a=1}^{n_A} \rho_{a'a} \ell_{aa'} \tag{7.14}$$

Proof. Had Alice prepared her system in a single state $\left|\Psi_k\right\rangle$ then she would have computed $\left\langle L_A \right\rangle \overset{(4.13)}{=} \left\langle \Psi_k \middle| L_A \middle| \Psi_k \right\rangle$. So, for a mixed state, she computes

$$\left\langle L_A \right\rangle = \sum_{k=1}^{r} \mathrm{P}(k) \left\langle \Psi_k \middle| L_A \middle| \Psi_k \right\rangle \overset{(7.12)}{=} \sum_{k=1}^{r} \mathrm{P}(k) \mathrm{Tr}\left(\left|\Psi_k\right\rangle \left\langle \Psi_k \right| L_A \right)$$

$$= \mathrm{Tr}\left[\left(\sum_{k=1}^{r} \mathrm{P}(k) \left|\Psi_k\right\rangle \left\langle \Psi_k \right| \right) L_A \right] \overset{(7.12a)}{=} \mathrm{Tr}\left(\rho L_A\right) \checkmark$$

In the basis $\left\{\left|a\right\rangle\right\}_{a=1}^{n_A}$, $\rho L_A = \left(\rho_{a'a}\right)\left(\ell_{aa''}\right) = \left(\sum_{a=1}^{n_A} \rho_{a'a} \ell_{aa''} \right)$. That is, the element of ρL_A in row a' and column a'' is $\sum_{a=1}^{n_A} \rho_{a'a} \ell_{aa''}$. The diagonal element in row a' and column a' is $\sum_{a=1}^{n_A} \rho_{a'a} \ell_{aa'}$. Thus, $\mathrm{Tr}\left(\rho L_A\right) = \sum_{a'=1}^{n_A} \sum_{a=1}^{n_A} \rho_{a'a} \ell_{aa'}$ is the sum of the diagonal elements of ρL_A. Finally, $\left\langle L_A \right\rangle \overset{(7.13)}{=} \mathrm{Tr}\left(\rho L_A\right) = \sum_{a'=1}^{n_A} \sum_{a=1}^{n_A} \rho_{a'a} \ell_{aa'}$ ■

Note: The $n_A \times n_A$ matrix ρ is a sum of r terms.

$$\rho = \left(\rho_{a'a}\right) \overset{(7.12a)}{=} \sum_{k=1}^{r} \mathsf{P}(k)\left|\Psi_k\right\rangle\left\langle\Psi_k\right| \overset{(7.04)}{=} \sum_{k=1}^{r} \mathsf{P}(k)\Big(\left(\psi_k\right)_a^* \left(\psi_k\right)_{a'} \Big)$$

$$= \begin{pmatrix} & \vdots & \\ \cdots & \sum_{k=1}^{r} \mathsf{P}(k)\left(\psi_k\right)_a^*\left(\psi_k\right)_{a'} & \cdots \\ & \vdots & \end{pmatrix} \tag{7.12c}$$

Thus each term $\rho_{a'a}$ in (7.14) is a sum of *r* terms:

$$\rho_{a'a} \overset{(7.12c)}{=} \sum_{k=1}^{r} \mathsf{P}(k)\left(\psi_k\right)_a^*\left(\psi_k\right)_{a'} \tag{7.12d}$$

Definition. **A pure composite system** is one in which there is a state vector $\left|\Psi\right\rangle$ (or, equivalently, a wave function $\left\{\psi_{ab}\right\}$). If there is a distribution of possible starting states, we say that AB is **mixed.** If AB is pure we say that **Alice knows the wave function** $\left\{\psi_{ab}\right\}$ (and also the composite system state $\left|\Psi\right\rangle$). Because $\left|\Psi\right\rangle$ exists we can define a **density matrix** ρ**:**

$$\boldsymbol{\rho_{aa'} = \sum_{b} \psi^*_{a'b}\psi_{ab}} \,. \tag{7.17}$$

Exercise. Suppose Alice and Bob have single-spin systems and Alice selects an observable L_A of her subsystem A. L_A has no effect on Bob's subsystem B. Find the operator L in the composite system AB that represents L_A. If AB is pure, find Alice's expected outcome $\left\langle L_A\right\rangle$.

Solution. We use the σ_z basis $\left\{ \left|u\right\rangle, \left|d\right\rangle \right\}$. An element of L_A in row *a*' and column *a* is $\ell_{a'a} = \left\langle a'\right|L_A\left|a\right\rangle$ where *a'* and *a* range over *u* and *d*. Thus,

$$L_A = \begin{pmatrix} \ell_{uu} & \ell_{ud} \\ \ell_{du} & \ell_{dd} \end{pmatrix}.$$

Similarly, an element of *L* in row *a'b'* and column *ab* is $\ell_{a'b'ab} = \left\langle a'b'\right|L\left|ab\right\rangle$. Thus, any operator on the composite system has the form

$$L = \begin{pmatrix} \ell_{uuuu} & \ell_{uuud} & \ell_{uudu} & \ell_{uudd} \\ \ell_{uduu} & \ell_{udud} & \ell_{uddu} & \ell_{uddd} \\ \ell_{duuu} & \ell_{duud} & \ell_{dudu} & \ell_{dudd} \\ \ell_{dduu} & \ell_{ddud} & \ell_{dddu} & \ell_{dddd} \end{pmatrix}.$$

Keep in mind that the 1st and 3rd subscripts pertain to Alice and the 2nd and 4th pertain to Bob.

If L is the operator on the composite system that represents L_A, we must have

$$\langle L_A \rangle = \langle L \rangle \ .$$

In addition, the half of the elements, like ℓ_{uduu}, that represent a change in Bob's state must equal zero and the remaining elements that leave Bob's states unchanged must equal the corresponding L_A element. For example, ℓ_{duuu} equates to ℓ_{du}. We represent this mathematically as

$$\ell_{a'b'ab} = \ell_{a'a}\delta_{b'b} \tag{7.18}$$

where on the RHS $\ell_{a'a}$ captures Alice's observations and $\delta_{b'b}$ captures whether or not Bob changes states, "0" if he does, "1" if not.

We replace the elements of L using (7.18), carefully managing the subscripts because they are in a different order on each side of (7.18):

$$L = \begin{pmatrix} \ell_{uu}\delta_{uu} & \ell_{uu}\delta_{ud} & \ell_{ud}\delta_{uu} & \ell_{ud}\delta_{ud} \\ \ell_{uu}\delta_{du} & \ell_{uu}\delta_{dd} & \ell_{ud}\delta_{du} & \ell_{ud}\delta_{dd} \\ \ell_{du}\delta_{uu} & \ell_{du}\delta_{ud} & \ell_{dd}\delta_{uu} & \ell_{dd}\delta_{ud} \\ \ell_{du}\delta_{du} & \ell_{du}\delta_{dd} & \ell_{dd}\delta_{du} & \ell_{dd}\delta_{dd} \end{pmatrix} = \begin{pmatrix} \ell_{uu} & 0 & \ell_{ud} & 0 \\ 0 & \ell_{uu} & 0 & \ell_{ud} \\ \ell_{du} & 0 & \ell_{dd} & 0 \\ 0 & \ell_{du} & 0 & \ell_{dd} \end{pmatrix}.$$

To confirm that L represents L_A, we set

$$\left(\delta_{b'b}\right) = \begin{pmatrix} \delta_{uu} & \delta_{ud} \\ \delta_{du} & \delta_{dd} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

and we use the pattern definition for matrix representation of tensor product to get

$$L_A \otimes I \overset{(7.7)}{=} \begin{pmatrix} \ell_{uu}I & \ell_{ud}I \\ \ell_{du}I & \ell_{dd}I \end{pmatrix} \overset{(7.7)}{=} \left(\begin{array}{cc|cc} \ell_{uu} & 0 & \ell_{ud} & 0 \\ 0 & \ell_{uu} & 0 & \ell_{ud} \\ \hline \ell_{du} & 0 & \ell_{dd} & 0 \\ 0 & \ell_{du} & 0 & \ell_{dd} \end{array}\right) = L\,. \quad \checkmark$$

Now, suppose AB is pure. Then there is a unit composite state vector $\left|\Psi\right\rangle$. Let $\left\{\left|ab\right\rangle\right\}$ be a set of normalized eigenvectors of L. Then $\left\{\left|ab\right\rangle\right\}$ forms an orthonormal basis for AB by the Fundamental Theorem and also by Exercise 3.1. Hence, we can express $\left|\Psi\right\rangle$ as

$$\left|\Psi\right\rangle = \sum_{a,b} \psi_{ab}\left|ab\right\rangle.$$

So

$$\left\langle\Psi\right| = \sum_{a,b} \left\langle ab\right| \psi^*_{ab},$$

and

$$\left\langle L\right\rangle \overset{(4.14)}{=} \left\langle\Psi\right|L\left|\Psi\right\rangle.$$

We are now set up to compute Alice's expected outcome. To compute $\left\langle L\right\rangle$, recall $\left\langle A\right|M\left|A\right\rangle \overset{(3.05)}{=} \sum_{i,j} \alpha_i^* m_{ij}\alpha_j$. By letting $i = a'b'$, $j = ab$, $A = \Psi$, and $M = L$ we get

$$\left\langle L_A\right\rangle = \left\langle L\right\rangle \overset{(4.14)}{=} \left\langle\Psi\right|L\left|\Psi\right\rangle \overset{(3.05)}{=} \sum_{ab,a'b'} \psi^*_{a'b'}\ell_{a'b'ab}\psi_{ab} \tag{7.15}$$

$$\left\langle L_A\right\rangle \overset{(7.18)}{=} \sum_{a,b,a',b'} \psi^*_{a'b'}\ell_{a'a}\delta_{b'b}\psi_{ab}$$

$$\left\langle L_A\right\rangle = \sum_{a,b,a'} \psi^*_{a'b}\ell_{a'a}\psi_{ab} \tag{7.19}$$

$$\left\langle L_A\right\rangle = \sum_{a,a'} \ell_{a'a} \sum_{b} \psi^*_{a'b}\psi_{ab}\,. \tag{7.19b}$$

$$\left\langle L_A\right\rangle \overset{(7.17)}{=} \sum_{a,a'} \rho_{aa'}\ell_{a'a} = \sum_{a',a} \rho_{a'a}\ell_{aa'} \quad ✔ \tag{7.21}$$

This is the same as equation (7.14) in which ρ represented Alice's density matrix for a mixed system S_A. Thus, even though AB is pure, A is described as a mixed state. We will quantify just how "mixed" A is shortly.

To generate the other expression, (7.13), for $\left\langle L_A\right\rangle$, we denote Alice's density matrix as $\rho = \left(\rho_{aa'}\right) = \begin{pmatrix} \rho_{uu} & \rho_{ud} \\ \rho_{du} & \rho_{dd} \end{pmatrix}$. Then

$$\rho L_A = \begin{pmatrix} \rho_{uu} & \rho_{ud} \\ \rho_{du} & \rho_{dd} \end{pmatrix}\begin{pmatrix} \ell_{uu} & \ell_{ud} \\ \ell_{du} & \ell_{dd} \end{pmatrix} = \begin{pmatrix} \rho_{uu}\ell_{uu} + \rho_{ud}\ell_{du} & \rho_{uu}\ell_{ud} + \rho_{ud}\ell_{dd} \\ \rho_{du}\ell_{uu} + \rho_{dd}\ell_{du} & \rho_{du}\ell_{ud} + \rho_{dd}\ell_{dd} \end{pmatrix} \tag{7.20b}$$

and

$$\langle L_A \rangle \overset{(7.21)}{=} \sum_{a',a} \rho_{a'a}\,\ell_{aa'} = \rho_{uu}\ell_{uu} + \rho_{ud}\ell_{du} + \rho_{du}\ell_{ud} + \rho_{dd}\ell_{dd}.$$

Noticing that this expression is the sum of the diagonal of ρL_A, as shown in (7.20b), we obtain $\langle L_A \rangle = \mathrm{Tr}(\rho L_A)$. ✔ ■

Susskind says that in order to calculate Alice's density matrix ρ she may have had need of the composite wave function, but once she has ρ she can forget where it came from and use it compute anything about her observations. I believe this is because the diagonal of ρ contains the probabilities of all of Alice's possible states. But I find that sometimes one has to be very careful with the notation involved. (See "Caution" below). Thus, I have formalized Susskind's comment in next theorem. The last section of the proof to the Corollary to Theorem 7.5, coming soon, is an example of when one must be careful with the notation.

Theorem 7.3. Let L_A be an observable of Alice's system $\mathrm{S_A}$ having orthonormal eigenvectors $\{|a\rangle\}$ and associated eigenvalues $\{a\}$. The collection $\{|a\rangle\}$ constitutes an orthonormal basis for S_A. Let L_B be an observable of Bob's system S_B having orthonormal eigenvectors $\{|b\rangle\}$. Suppose the composite system $\mathrm{S_{AB}}$ is prepared in the state $|\Psi\rangle = \sum_{ab} \psi_{ab} |ab\rangle$. Let $\rho = (\rho_{aa'})$ be Alice's density matrix per equation (7.17). After a measurement L_A is performed, the probability of being in state $|a\rangle$ is

$$\mathrm{P}_{L_A}(a) = \rho_{aa} \tag{7.22}$$

Poof. From Principle 4, $\mathrm{P}_L(a,b) \overset{(3.11)}{=} \langle \Psi | ab \rangle \langle ab | \Psi \rangle = \psi^*_{ab}\psi_{ab}$. So

$$\mathrm{P}_{L_A}(a) = \sum_b \mathrm{P}_L(a,b) = \sum_b \psi^*_{ab}\psi_{ab} \overset{(7.17)}{=} \rho_{aa}.$$ ■

Note. ρ_{aa} is a diagonal element of ρ. So the diagonal elements are probabilities.

Caution. The subscripts a and b are not 1, 2, …, n as in previous formulas. For example, let L_A be the observable "energy state E_i of an atom". Then $|a\rangle = |E_i\rangle$ and $a = E_i$. Thus the subscript "a" ranges over E_1, E_2, …, E_n, not over 1, 2, …, n. As another example, let $L_A = \sigma_x$. Then $|a\rangle$ ranges over $|r\rangle$ and $|\ell\rangle$, and "a" ranges over λ_r and λ_ℓ, the eigenvalues of σ_x.

Theorem 7.4. Alice's and Bob's systems are pure iff the composite wave function ψ_{ab} factorizes (i.e., $\psi_{ab} = \psi_a \phi_b$). In this case Alice's wave function can be expressed

$$\rho_{aa'} = \psi_{a'}^* \psi_a \tag{7.25}$$

Proof. Suppose $\psi_{ab} = \psi_a \phi_b$. Define $\left|\Psi_A\right\rangle = \sum_a \psi_a \left|a\right\rangle$ and $\left|\Phi_B\right\rangle = \sum_b \phi_b \left|b\right\rangle$, states of Alice and Bob, respectively. Then

$$\rho_{aa'} \overset{(7.17)}{=} \sum_b \psi_{a'b}^* \psi_{ab} = \psi_{a'}^* \psi_a \sum_b \phi_b^* \phi_b = \psi_{a'}^* \psi_a \left\langle \Phi_B \middle| \Phi_B \right\rangle = \psi_{a'}^* \psi_a .$$

That is, $\rho \overset{(7.04)}{=} \left|\Psi_A\right\rangle\left\langle\Psi_A\right|$ and, by definition (7.12b), Alice's system is pure. ✔
Similarly for Bob. ✔

Conversely, suppose A and B are pure. Then there are state vectors $\left|\Psi_A\right\rangle = \sum_a \psi_a \left|a\right\rangle$ and $\left|\Phi_B\right\rangle = \sum_b \phi_b \left|b\right\rangle$. So

$$\rho \overset{(7.12b)}{=} \left|\Psi_A\right\rangle\left\langle\Psi_A\right|$$

$$\Rightarrow \quad \rho_{aa'} = \psi_{a'}^* \psi_a = \psi_{a'}^* \psi_a \left\langle \Phi_B \middle| \Phi_B \right\rangle = \sum_b \psi_a^* \phi_b^* \psi_{a'} \phi_b$$

By definition (7.17), for a composite system,

$$\rho_{aa'} \overset{(7.17)}{=} \sum_b \psi_{a'b}^* \psi_{ab} .$$

So

$$\sum_b \psi_{a'b}^* \psi_{ab} = \sum_b \psi_a^* \phi_b^* \psi_{a'} \phi_b \;\; \forall a, a' .$$

This represents n^2 equations (i.e., $a, a' = 1, \cdots, n$) in n^2 unknowns (i.e., ψ_{ab} where $a, b = 1, \cdots, n$). Clearly $\psi_{ab} = \psi_a \phi_b \;\forall a, b$ is a solution. Thus, ψ_{ab} factorizes. ■

Definition. If the composite wave function does *not* factorize, we say that AB is (and Alice and Bob are) **entangled.** AB is **maximally entangled** if all of Alice's (and Bob's) probabilities P(k) in equation (7.12a) are equal.

We can restate Theorem 7.4 using this definition: If Alice and Bob are entangled then Alice does not have a state space. That is, AB is entangled implies A is entangled.
Summary – What Alice knows about her system and the composite system

when AB is pure

- Alice has complete knowledge about the composite system because she knows composite state vector $|\Psi\rangle$. She knows everything that can be known.
- As for her own system:
    - At one extreme, if $|\Psi\rangle$ factorizes, then from Theorem 7.4 her system is pure. She knows her state vector $|\Psi_A\rangle$. Alice has complete knowledge of her system.
    - At the other extreme, if AB is maximally entangled, she knows nothing of her state (because all outcomes are equally likely) though she has maximum information about AB.
        - Einstein had difficulty accepting this.

Example. If the composite state vector is a singlet or a triplet, then AB is maximally entangled.

Proof. We must show that all of Alice's probabilities are equal; that is, $\mathrm{P}(u)=\mathrm{P}(d)=\frac{1}{2}$.

Suppose $|\Psi\rangle=|\text{sing}\rangle=\frac{1}{\sqrt{2}}\left(|ud\rangle-|du\rangle\right)$. Then $\langle\Psi|=\frac{1}{\sqrt{2}}\left(\langle ud|-\langle du|\right)$. Thus,

$$\psi_{ud}=\psi^*_{ud}=\frac{1}{\sqrt{2}},\ \psi_{du}=\psi^*_{du}=-\frac{1}{\sqrt{2}}\ \text{ and }\ \psi_{uu}=\psi^*_{uu}=\psi_{dd}=\psi^*_{dd}=0.$$

Using $\rho_{aa'}\overset{(7.17)}{=}\sum_b \psi^*_{a'b}\psi_{ab}$ we get

$$\rho_{uu}=\psi^*_{uu}\psi_{uu}+\psi^*_{ud}\psi_{ud}=\frac{1}{2},$$

$$\rho_{dd}=\psi^*_{du}\psi_{du}+\psi^*_{dd}\psi_{dd}=\frac{1}{2},$$

$$\rho_{ud}=\psi^*_{du}\psi_{uu}+\psi^*_{dd}\psi_{ud}=0,$$

$$\rho_{du}=\psi^*_{uu}\psi_{du}+\psi^*_{ud}\psi_{dd}=0.$$

So $\rho=\begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix}\overset{(7.22)}{=}\begin{pmatrix}\mathrm{P}(u) & 0\\ 0 & \mathrm{P}(d)\end{pmatrix}\ \Rightarrow\ \mathrm{P}(u)=\mathrm{P}(d)=\frac{1}{2}$. Thus, AB is maximally entangled. ✔

We showed in Exercise 7.8(b) for $\left|\Psi\right\rangle = \left|T_1\right\rangle = \frac{1}{\sqrt{2}}\left(\left|ud\right\rangle + \left|du\right\rangle\right)$ that

$$\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \quad ✔$$

and we can similarly show this for $\left|T_2\right\rangle$ and $\left|T_3\right\rangle$. ✔ ■

We next prove the footnote on p. 108 of the book that states that any Hermitian matrix can be diagonalized by a change of basis. There are insights to be gained, including the corollary to the theorem.

Theorem 7.5. If M is an n x n Hermitian matrix, there is a unitary matrix U such that $U^\dagger M U$ is a diagonal matrix having the eigenvalues of M on its diagonal.

Proof. The matrix M represents an operator in some basis $\{\left|i\right\rangle\}$. Because M is Hermitian there is an orthonormal basis $\{\left|\lambda_i\right\rangle\}$ of eigenvectors of M and a set $\{\lambda_i\}$ of corresponding eigenvalues. Write $\left|\lambda_i\right\rangle = \sum_j \lambda_{ij}\left|j\right\rangle$, and let U be the matrix whose columns are the eigenvectors $\left|\lambda_j\right\rangle$:

$$U = \begin{pmatrix} & \lambda_{j1} & \\ \cdots & \vdots & \cdots \\ & \lambda_{jn} & \end{pmatrix}. \tag{A}$$

So $U^\dagger$ is the matrix whose rows are the bra vectors $\left\langle\lambda_i\right|$:

$$U^\dagger = \begin{pmatrix} & \vdots & \\ \lambda_{i1}^* & \cdots & \lambda_{in}^* \\ & \vdots & \end{pmatrix}. \tag{B}$$

U is unitary because

$$U^\dagger U = \begin{pmatrix} & \vdots & \\ \lambda_{i1}^* & \cdots & \lambda_{in}^* \\ & \vdots & \end{pmatrix} \begin{pmatrix} & \lambda_{j1} & \\ \cdots & \vdots & \cdots \\ & \lambda_{jn} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & \left\langle\lambda_i\middle|\lambda_j\right\rangle & \cdots \\ & \vdots & \end{pmatrix} = \left(\delta_{ij}\right) = I. \quad ✔$$

M acting on column j of U gives

$$M\begin{pmatrix}\lambda_{j1}\\ \vdots \\ \lambda_{jn}\end{pmatrix} = M\left|\lambda_j\right\rangle = \lambda_j\left|\lambda_j\right\rangle = \begin{pmatrix}\lambda_j\lambda_{j1}\\ \vdots \\ \lambda_j\lambda_{jn}\end{pmatrix}$$

$$\Rightarrow \quad MU = \begin{pmatrix} & \lambda_j\lambda_{j1} & \\ \cdots & \vdots & \cdots \\ & \lambda_j\lambda_{jn} & \end{pmatrix}. \tag{C}$$

This implies that the *ij*-element of $U^\dagger MU$, row *i* of $U^\dagger$ (expression B) times column *j* of *MU* (expression C), is

$$\left(\lambda_{i1}^* \cdots \lambda_{in}^*\right)\begin{pmatrix}\lambda_j\lambda_{j1}\\ \vdots \\ \lambda_j\lambda_{jn}\end{pmatrix} = \lambda_j\left(\left\langle\lambda_i\right|\right)\left(\left|\lambda_j\right\rangle\right) = \lambda_j\left\langle\lambda_i\middle|\lambda_j\right\rangle = \lambda_j\,\delta_{ij}\,. \tag{D}$$

$$\Rightarrow \qquad U^\dagger MU \overset{\text{(D)}}{=} \left(\lambda_j\,\delta_{ij}\right) = \begin{pmatrix}\lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n\end{pmatrix}. \quad ✔ \qquad\qquad ■$$

Corollary. Let ρ be Alice's density matrix in a pure composite system AB, and denote the eigenvalues of ρ by λ_a. Let *U* be the matrix whose columns are the eigenvectors $\left|\lambda_a\right\rangle$ of ρ. Then $\bar{\rho} = U^\dagger \rho\, U$ is also a density matrix as well as a diagonal matrix whose diagonal entries are $\bar{\rho}_{aa} = \lambda_a = \mathrm{P}_\rho\left(\lambda_a\right)$. Thus the eigenvalues are their own probabilities of occurrence when ρ is the observable.

Proof. Let $\left\{\left|a\right\rangle\right\}$, $\left\{\left|b\right\rangle\right\}$, and $\left\{\left|ab\right\rangle\right\}$ denote orthonormal bases for Alice, Bob, and the composite system, respectively. Since AB is pure, there is a state vector

$$\left|\Psi\right\rangle = \sum_{ab}\psi_{ab}\left|ab\right\rangle, \tag{a}$$

and by definition (7.17) Alice's density matrix is

$$\rho = \begin{pmatrix} & \vdots & \\ \cdots & \rho_{a'a} & \cdots \\ & \vdots & \end{pmatrix} \text{ where } \rho_{a'a} = \sum_b \psi_{ab}^*\psi_{a'b}\,. \tag{b}$$

We can express the eigenvectors of ρ in terms of the basis $\left\{\left|a\right\rangle\right\}$:

$$\left|\lambda_a\right\rangle = \sum_{a'} \lambda_{aa'} \left|a'\right\rangle . \tag{c}$$

The matrix U is

$$U = \begin{pmatrix} \cdots & \left|\lambda_a\right\rangle & \cdots \end{pmatrix} = \begin{pmatrix} & \lambda_{a1} & \\ \cdots & \vdots & \cdots \\ & \lambda_{an} & \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{n1} \\ \vdots & & \vdots \\ \lambda_{1n} & \cdots & \lambda_{nn} \end{pmatrix}, \tag{d}$$

and so

$$U^\dagger = \begin{pmatrix} \lambda_{11}^* & \cdots & \lambda_{1n}^* \\ \vdots & & \vdots \\ \lambda_{n1}^* & \cdots & \lambda_{nn}^* \end{pmatrix}. \tag{e}$$

By Theorem 7.5 U is unitary,

$$U^\dagger = U^{-1}. \tag{f}$$

Also by Theorem 7.5, $\overline{\rho}$ is a diagonal matrix having the eigenvalues of ρ on its diagonal:

$$\overline{\rho} = U^\dagger \rho\, U = \begin{pmatrix} & \vdots & \\ \cdots & \overline{\rho}_{a'a} & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \quad ✔ \tag{g}$$

In order to show that $\overline{\rho}$ is a density matrix we must find a wave function $\left\{ \overline{\psi}_{ab} \right\}$ such that $\overline{\rho}_{a'a} = \sum_b \overline{\psi}^*_{ab} \overline{\psi}_{a'b}$. We will find $\overline{\psi}_{ab}$ by rewriting $\left|\Psi\right\rangle$ in terms of the eigenvector composite basis $\left\{ \left|\lambda_a b\right\rangle \right\}$.

We start by writing Alice's basis elements $\left|a\right\rangle$ in terms of the eigenvector basis elements λ_a.

Claim: $\left|a\right\rangle = \sum_{a'} \lambda^*_{aa'} \left|\lambda_{a'}\right\rangle$ (h)

Define $\vec{A} = \begin{pmatrix} \vdots \\ \left\langle a\right| \\ \vdots \end{pmatrix} = \begin{pmatrix} \left\langle 1\right| \\ \vdots \\ \left\langle n\right| \end{pmatrix}$ and $\vec{\Lambda} = \begin{pmatrix} \left\langle \lambda_1\right| \\ \vdots \\ \left\langle \lambda_n\right| \end{pmatrix}$, vectors composed of row vectors.

$$\langle\lambda_a| \overset{(c)}{=} \sum_{a'} \langle a'|\lambda^*_{aa'} \text{ for } a = 1, \ldots, n$$

$$\Leftrightarrow \quad \begin{cases} \lambda^*_{11}\langle 1| + \cdots + \lambda^*_{1n}\langle n| & = & \langle\lambda_1| \\ \vdots & & \vdots \\ \lambda^*_{n1}\langle 1| + \cdots + \lambda^*_{nn}\langle n| & = & \langle\lambda_n| \end{cases}$$

$$\Leftrightarrow \quad \begin{pmatrix} \lambda^*_{11} & \cdots & \lambda^*_{1n} \\ \vdots & & \vdots \\ \lambda^*_{n1} & \cdots & \lambda^*_{nn} \end{pmatrix} \begin{pmatrix} \langle 1| \\ \vdots \\ \langle n| \end{pmatrix} = \begin{pmatrix} \langle\lambda_1| \\ \vdots \\ \langle\lambda_n| \end{pmatrix}$$

$$\Leftrightarrow \quad U^{-1}\vec{A} \overset{(f)}{=} U^{\dagger}\vec{A} = \vec{\Lambda}$$

$$\Leftrightarrow \quad \vec{A} = U\vec{\Lambda}$$

$$\Leftrightarrow \quad \begin{pmatrix} \langle 1| \\ \vdots \\ \langle n| \end{pmatrix} \overset{(d)}{=} \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{n1} \\ \vdots & & \vdots \\ \lambda_{1n} & \cdots & \lambda_{nn} \end{pmatrix} \begin{pmatrix} \langle\lambda_1| \\ \vdots \\ \langle\lambda_n| \end{pmatrix}$$

$$\Leftrightarrow \quad \begin{cases} \langle 1| & = & \lambda_{11}\langle\lambda_1| + \cdots + \lambda_{n1}\langle\lambda_n| \\ \vdots & & \vdots \\ \langle n| & = & \lambda_{1n}\langle\lambda_1| + \cdots + \lambda_{nn}\langle\lambda_n| \end{cases}.$$

$$\Leftrightarrow \quad \langle a| = \sum_{a'} \lambda_{a'a}\langle\lambda_{a'}| \quad \text{for } a = 1, \ldots, n.$$

$$\Leftrightarrow \quad |a\rangle = \sum_{a'} \lambda^*_{a'a}|\lambda_{a'}\rangle \quad \text{for } a = 1, \ldots, n. \qquad ✔$$

We can use (h) to write the composite state vector $|\Psi\rangle$ in terms of the basis $\{|\lambda_a b\rangle\}$:

$$|\Psi\rangle \overset{(a)}{=} \sum_{ab} \psi_{ab}|ab\rangle \overset{(h)}{=} \sum_{ab} \psi_{ab} \sum_{a'} \lambda^*_{a'a}|\lambda_{a'}b\rangle = \sum_{a'b}\left(\sum_a \psi_{ab}\lambda^*_{a'a}\right)|\lambda_{a'}b\rangle. \tag{i}$$

We have succeeded in writing $|\Psi\rangle$ in terms of the eigenvector composite basis. Next define $\bar{\psi}_{a'b}$ as the coefficient of $|\lambda_{a'}b\rangle$ in equation (i) :

$$\bar{\psi}_{a'b} = \sum_k \psi_{kb}\lambda^*_{a'k}. \tag{j}$$

So

$$\left|\Psi\right\rangle \overset{(i,j)}{=} \sum_{a'b} \bar{\psi}_{a'b} \left|\lambda_{a'}b\right\rangle \tag{k}$$

Replacing a' by a in (j) yields

$$\bar{\psi}_{ab} = \sum_k \psi_{kb} \lambda^*_{ak} \,. \tag{l}$$

Hence

$$\bar{\psi}^*_{ab} \overset{(j)}{=} \sum_j \psi^*_{jb} \lambda_{aj} \,. \tag{m}$$

We can now express $\bar{\rho}$ in terms of a wave function.

Claim $\bar{\rho}_{a'a} = \sum_b \bar{\psi}^*_{ab} \bar{\psi}_{a'b}$: (n)

$$\text{RHS} = \sum_b \bar{\psi}^*_{ab} \bar{\psi}_{a'b} \overset{(m,j)}{=} \sum_b \sum_{j,k} \psi^*_{jb} \psi_{kb} \lambda^*_{a'k} \lambda_{aj} = \sum_{j,k} \lambda^*_{a'k} \lambda_{aj} \sum_b \psi^*_{jb} \psi_{kb}$$

$$\overset{(b)}{=} \sum_{j,k} \rho_{kj} \lambda^*_{a'k} \lambda_{aj}$$

$$\bar{\rho} = U^\dagger \rho\, U \overset{(e,b,d)}{=} \begin{pmatrix} \lambda^*_{11} & \cdots & \lambda^*_{1n} \\ \vdots & & \vdots \\ \lambda^*_{n1} & \cdots & \lambda^*_{nn} \end{pmatrix} \begin{pmatrix} \rho_{11} & \cdots & \rho_{1n} \\ \vdots & & \vdots \\ \rho_{n1} & \cdots & \rho_{nn} \end{pmatrix} \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{n1} \\ \vdots & & \vdots \\ \lambda_{1n} & \cdots & \lambda_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda^*_{11} & \cdots & \lambda^*_{1n} \\ \vdots & & \vdots \\ \lambda^*_{n1} & \cdots & \lambda^*_{nn} \end{pmatrix} \begin{pmatrix} \sum_j \rho_{1j}\lambda_{1j} & \cdots & \sum_j \rho_{1j}\lambda_{nj} \\ \vdots & & \vdots \\ \sum_j \rho_{nj}\lambda_{1j} & \cdots & \sum_j \rho_{nj}\lambda_{nj} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j,k} \lambda^*_{1k}\rho_{kj}\lambda_{1j} & \cdots & \sum_{j,k} \lambda^*_{1k}\rho_{kj}\lambda_{nj} \\ \vdots & & \vdots \\ \sum_{j,k} \lambda^*_{nk}\rho_{kj}\lambda_{1j} & \cdots & \sum_{j,k} \lambda^*_{nk}\rho_{kj}\lambda_{nj} \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & \sum_{j,k} \lambda^*_{a'k}\rho_{kj}\lambda_{aj} & \cdots \\ & \vdots & \end{pmatrix} = \left(\bar{\rho}_{a'a}\right)$$

$\text{LHS} = \bar{\rho}_{a'a} = \sum_{j,k} \rho_{kj} \lambda^*_{a'k} \lambda_{aj} = \text{RHS}$, which proves the claim. ✔

Equation (n) means that $\bar{\rho}$ is a density function. ✔

More specifically, $\bar{\rho}$ is generated from $\left|\Psi\right\rangle \overset{(k)}{=} \sum_{a'b} \bar{\psi}_{a'b} \left|\lambda_{a'}b\right\rangle$ per equation (7.17).

Lastly, we show that $P_\rho(\lambda_a) = \lambda_a$. Our approach is to carefully use Theorem 7.3. First, replace L_A in the theorem by ρ. Since the eigenvectors of ρ are $|\lambda_a\rangle$ with associated eigenvalues λ_a, we must replace the eigenvectors $|a\rangle$ by $|\lambda_a\rangle$ and the eigenvalues *a* by λ_a. Then the composite state vector $|\Psi\rangle = \sum_{ab} \psi_{ab} |ab\rangle$, in terms of the basis $|ab\rangle$, must be replaced by $|\Psi\rangle = \sum_{ab} \bar{\psi}_{ab} |\lambda_a b\rangle$, expressed in terms of the basis $|\lambda_a b\rangle$. (Note that the *subscript a* need not be replaced by *subscript* λ_a). Consequently, RHS of equation (7.17) must be changed from $\sum_b \psi^*_{a'b} \psi_{ab}$ to $\sum_b \bar{\psi}^*_{a'b} \bar{\psi}_{ab}$. The latter sum equals $\bar{\rho}_{aa'}$ by equation (n), so we replace $\rho_{aa'}$ on LHS of (7.17) by $\bar{\rho}_{aa'}$. That is, we replace Alice's density matrix $\rho = (\rho_{aa'})$ in the theorem by the density matrix $\bar{\rho} = (\bar{\rho}_{aa'})$. The conclusion to the theorem, equation (7.22), then becomes $P_\rho(\lambda_a) = \bar{\rho}_{aa'} \overset{(g)}{=} \lambda_a$ ■

Lessons Learned about ρ and $\bar{\rho}$:

1. Diagonal elements of ρ are probabilities $\{ P_{L_A}(a) \}$
2. Diagonal elements of $\bar{\rho} = U^\dagger \rho U$ are $\{\lambda_a\}$, the eigenvalues of ρ
3. $\bar{\rho}_{aa} = \lambda_a$ is a probability but not generally equal to $\rho_{aa} = P_{L_A}(a)$

Theorem 7.6. If *A* and *B* are matrices then Tr *AB* = Tr *BA*.

Proof. $AB = (a_{ik})(b_{kj}) = \left(\sum_k a_{ik} b_{kj}\right)$ and $BA = (b_{ik})(a_{kj}) = \left(\sum_k b_{ik} a_{kj}\right)$. So

$$\text{Tr}\, AB = \sum_{i,k} a_{ik} b_{ki} \overset{(*)}{=} \sum_{k,i} a_{ki} b_{ik} = \sum_{i,k} b_{ik} a_{ki} = \text{Tr}\, BA.$$

(*) Replace *i* by *k* and *k* by *i*. ■

The next theorem applies to $\bar{\rho}$ as well as ρ.

Theorem 7.7. Let $\{\psi_{ab}\}$ be the wave function for a pure composite system AB, and let ρ be the density matrix for system A.

(a) ρ is Hermitian
(b) (Exercise 7.6) Tr (ρ) = 1

(c) If λ is an eigenvalue of ρ, then $0 \le \lambda \le 1$
(d) If any eigenvalue equals 1 then all the others are zero
(e) For a pure state A, $\rho^2 = \rho$ and $\mathrm{Tr}\left(\rho^2\right) = 1$

(f) For an entangled state A, $\rho^2 \ne \rho$ and $\mathrm{Tr}\left(\rho^2\right) < 1$

Proof.

(a) $\rho_{a'a} \overset{(7.17)}{=} \sum_b \psi^*_{a'b}\psi_{ab} \quad \Rightarrow \quad \rho_{aa'} = \sum_b \psi^*_{ab}\psi_{a'b}$ and $\rho^*_{aa'} = \sum_b \psi^*_{a'b}\psi_{ab} = \rho_{a'a}$ ✔

(b) $\mathrm{Tr}\left(\rho\right) = \sum_a \rho_{aa} \overset{(7.17)}{=} \sum_a \sum_b \psi^*_{ab}\psi_{ab} = \left\langle \Psi \middle| \Psi \right\rangle = 1$ ✔

(c) From the Corollary to Theorem 7.5, λ is a probability. ✔

(d) By (b) and (c), if one diagonal element $\lambda = 1$, then the rest are zero. ✔

(e) For a pure system A, by definition there is a state vector $\left|\Psi\right\rangle$, and ρ is the projection operator: $\rho = \left|\Psi\right\rangle\left\langle\Psi\right|$. So

$$\rho^2 \overset{(7.1e)}{=} \rho. \quad ✔$$

Also

$$\mathrm{Tr}\left(\rho^2\right) = \mathrm{Tr}\,\rho \overset{(7.1f)}{=} 1. \quad ✔$$

(f) For a mixed state A, $\rho = \sum_{k=1}^{r} \mathrm{P}\left(k\right)\left|\Psi_k\right\rangle\left\langle\Psi_k\right|$ where $r > 1$ and $0 < \mathrm{P}\left(k\right) < 1 \ \forall k$. So

$$\rho^2 = \sum_k \sum_j \mathrm{P}\left(k\right)\mathrm{P}\left(j\right)\left|\Psi_k\right\rangle\left\langle\Psi_k\right|\left|\Psi_j\right\rangle\left\langle\Psi_j\right| = \sum_k \sum_j \mathrm{P}\left(k\right)\mathrm{P}\left(j\right)\delta_{kj}\left|\Psi_k\right\rangle\left\langle\Psi_j\right|$$

$$= \sum_k \mathrm{P}\left(k\right)^2\left|\Psi_k\right\rangle\left\langle\Psi_k\right| \ne \sum_k \mathrm{P}\left(k\right)\left|\Psi_k\right\rangle\left\langle\Psi_k\right| = \rho. \quad ✔$$

Also, denote $\left|\Psi_k\right\rangle = \sum_j \psi_{kj}\left|j\right\rangle$. Then

$$\mathrm{Tr}\left(\rho^2\right) = \mathrm{Tr}\left(\sum_k \mathrm{P}\left(k\right)^2\left|\Psi_k\right\rangle\left\langle\Psi_k\right|\right) = \sum_k \mathrm{P}\left(k\right)^{2n}\mathrm{Tr}\left(\left|\Psi_k\right\rangle\left\langle\Psi_k\right|\right)$$

$$= \sum_k \mathrm{P}\left(k\right)^{2n}\sum_j \psi^*_{kj}\psi_{kj} = \sum_k \mathrm{P}\left(k\right)^{2n}\left\langle\Psi_k\middle|\Psi_k\right\rangle$$

$$= \sum_k \mathrm{P}\left(k\right)^{2n} < \sum_k \mathrm{P}\left(k\right) = 1 \quad ✔ \qquad ■$$

Theorem 7.8. Suppose Alice's density matrix ρ represents a pure system with state vector $\left|\Psi_A\right\rangle$. Then

(a) $\left|\Psi_A\right\rangle$ is an eigenvector of ρ with eigenvalue +1

(b) If $\left|\vartheta\right\rangle \perp \left|\Psi_A\right\rangle$ then $\left|\vartheta\right\rangle$ is an eigenvector of ρ with eigenvalue 0

(c) ρ has exactly one non-zero eigenvalue λ, and λ = +1

Proof.

(a) Since $\left|\Psi_A\right\rangle$ is pure, by Theorem 7.4 ψ_{ab} factorizes into $\psi_a\phi_b$. So

$$\rho_{a'a} = \sum_b \psi^*_{a'b}\psi_{ab} = \psi_{a'}{}^*\psi_a \sum_b \phi_b{}^*\phi_b = \psi_{a'}{}^*\psi_a\left\langle\Phi_B\middle|\Phi_B\right\rangle = \psi_{a'}{}^*\psi_a$$

$$\rho\left|\Psi_A\right\rangle = \left(\rho_{aa'}\right)\left(\psi_{a'}\right) = \left(\sum_a \rho_{aa'}\psi_{a'}\right) = \left(\sum_a \psi_{a'}{}^*\psi_a\psi_{a'}\right) = \left(\psi_a\left\langle\Psi_A\middle|\Psi_A\right\rangle\right) = \left(\psi_a\right) = \left|\Psi_A\right\rangle$$

(b) $$\rho\left|\vartheta\right\rangle = \left(\rho_{aa'}\right)\left(\varphi_{a'}\right) = \left(\sum_{a'} \rho_{aa'}\varphi_{a'}\right) = \left(\sum_{a'} \psi_{a'}{}^*\psi_a\varphi_{a'}\right) = \left(\psi_a\left\langle\Psi_A\middle|\vartheta\right\rangle\right) = 0 \quad ✔$$

(c) By (a), ρ has one eigenvalue equal to 1. By Theorem 7.7 (d), if any eigenvalue is 1, all the others are 0. ■

Corollary. If the state vector $\left|\Psi\right\rangle$ of a composite system is a product state vector, then (a–c) above hold.

An example of the concepts of the past few pages is helpful: Alice's and Bob's state vectors, wave functions, observables and expected values; density matrix and a change of basis to diagonalize it; product state and wave function; and correlation of the observables.