Quantum Computation and Quantum Information

-Michael A. Nielsen and Isaac L. Chuang

Chapter 2. Introduction to Quantum Mechanics

Selected Problems

Set-I

Exercises: 2.16, 2.60, 2.67, 2.73, 2.78, 2.80, 2.81, 2.82

Problems: 1, 2 and 3

2.16) Show that any linear operator P defined on a vector space V is a projector if and only if $P^2 = P$.

Solution:

$\boxed{P \text{ is a projector} \Rightarrow P^2 = P}$:

Suppose W is the k-dimensional vector subspace of the d-dimensional vector space V. Using the Gram-Schmidt procedure it is possible to construct an orthonormal basis $|1\rangle, |2\rangle, \ldots, |d\rangle$ for V such that $|1\rangle, |2\rangle, \ldots, |k\rangle$ is an orthonormal basis for W. By definition,

$$P \equiv \sum_{i=1}^{k} |i\rangle\langle i| \tag{1}$$

is a projector onto the subspace W.

$$P^2 = \left(\sum_{i=1}^{k} |i\rangle\langle i|\right)\left(\sum_{j=1}^{k} |i\rangle\langle i|\right)$$

$$\Rightarrow \quad P^2 = \sum_{i,j=1}^{k} |i\rangle \underbrace{\langle i|j\rangle}_{0 \text{ for } i \neq j} \langle j|$$

We see that $|i\rangle$ and $|j\rangle$ being orthonormal results to $\langle i|j\rangle = 0$ for $i \neq j$ and $\langle i|j\rangle = 1$ for $i = j$. Hence,

$$\Rightarrow \quad P^2 = \sum_{i,j=1}^{k} |i\rangle\langle i|j\rangle\langle j| = \sum_{i,j=1}^{k} |i\rangle\langle j|\delta_{ij} = \sum_{i=1}^{k} |i\rangle\langle i| = P \quad \Rightarrow \quad P^2 = P$$

$\boxed{P^2 = P \Rightarrow P \text{ is a projector}}$:

It is given that $P : V \to V$ is a linear operator defined on the vector space V such that $P = P^2$. We show that P is a projector by showing:

- P is the identity operator on the subspace $U = \mathrm{im}(P)$, i.e. $\forall\, |u\rangle \in U$, $P|u\rangle = |u\rangle$.
- $V = U \oplus W$, where the subspace $W = \ker(P)$, i.e. $\forall\, |v\rangle \in V$, $|v\rangle = |u\rangle + |w\rangle$, where $|u\rangle = P|v\rangle$ and $|w\rangle = (I - P)|v\rangle$ and $|u\rangle \in U$, $|w\rangle \in W$.

a) First, we show that the operator P is an identity matrix on the subspace U.
U is the image of P which is defined as

$$\mathrm{im}(P) = U = \left\{ |u\rangle \in V \,\middle|\, P|v\rangle = |u\rangle \text{ for some } |v\rangle \in V \right\}$$

Suppose the vector space V is spanned by the orthonormal basis $\left\{|v_1\rangle, |v_2\rangle, \ldots, |v_d\rangle\right\}$. Using the Gram-Schmidt procedure we can produce the orthonormal basis $\left\{|x_1\rangle, |x_2\rangle, \ldots, |x_k\rangle\right\}$ that

spans the subspace U. It is given that $P=P^2$. So, for some $|u\rangle \in U$ (and some $|v\rangle \in V$) we have $P|u\rangle = P(P|v\rangle) = P^2|v\rangle = P|v\rangle = |u\rangle$.
Hence

$$P|u\rangle = |u\rangle \quad \Rightarrow \quad P|u\rangle = \sum_{i=1}^{k} \langle x_i|u\rangle |x_i\rangle = \sum_{i=1}^{k} |x_i\rangle \langle x_i|u\rangle$$
$$P|u\rangle = \left(\sum_{i=1}^{k} |x_i\rangle\langle x_i|\right)|u\rangle \quad \Rightarrow \quad P \equiv \sum_{i=1}^{k} |x_i\rangle\langle x_i| \underbrace{=}_{\substack{\text{Completeness}\\ \text{Relation}}} I$$

Hence, P is an identity matrix on the subspace U.

b) Next, we show that any vector $|v\rangle \in V$ can be uniquely decomposed into a sum of the vector $|u\rangle \in U$ and a vector $|w\rangle \in W$. We also show that the set of vectors in the intersection $U \cap W$ is $\{0\}$. W is the kernel of P which is defined as

$$\ker(P) = W = \left\{ |w\rangle \in V \middle| P|w\rangle = 0 \right\}$$

Consider a vector $|x\rangle \in V$. Since it is given that $P^2 = P$, we note that $P|x\rangle = P^2|x\rangle \Rightarrow P(I|x\rangle - P|x\rangle) = |0\rangle \Rightarrow P(|x\rangle - P|x\rangle) = |0\rangle$. Hence, we can define a vector $|w\rangle \in V$ as $|w\rangle = |x\rangle - P|x\rangle$ where $|x\rangle \in V$. Clearly, $|w\rangle$ satisfies $P|w\rangle = 0$. And so we can redefine the subspace $W = \ker(P)$ as

$$\ker(P) = W = \left\{ |w\rangle \in V \middle| |w\rangle = |x\rangle - P|x\rangle \text{ for some } |x\rangle \in V \right\} \tag{2}$$

Therefore, every vector $|x\rangle \in V$ can be decomposed as $|x\rangle = P|x\rangle + |x\rangle - P|x\rangle$. By definition of U and W, $P|x\rangle \in U$ and $(|x\rangle - P|x\rangle) \in W$. Next, we show that such a decomposition is unique by showing that $U \cap V = \{0\}$. Consider some $|z\rangle \in U \cap V$. Because $|z\rangle \in U$, $|z\rangle = P|v\rangle$ for some $|v\rangle \in V$. Applying P on both sides we get $P|z\rangle = P^2|z\rangle$. Also, $|z\rangle \in W$. So, $0 = P|z\rangle = P^2|v\rangle = P|v\rangle = |z\rangle$. Hence for some $|z\rangle \in U \cap W$ we have $|z\rangle = 0$. So, $U \cap W = \{\emptyset\}$. Hence,

$$V = U \oplus W \quad \text{i.e.} \quad V = \text{im}(P) \oplus \ker(P) \tag{3}$$

Therefore, $P : V \to V$ with $P = P^2$ is a projector.

2.60) Show that $\vec{v}\cdot\vec{\sigma}$ has eigenvalues ±1, and that the projectors on the corresponding eigenspaces are given by $P_\pm = (I \pm \vec{v}\cdot\vec{\sigma})/2$.

Solution:
We use the definition of the Pauli matrices, $\sigma_1, \sigma_2, \sigma_3$. Also, $\vec{v}$ is a unit vector which means $\sqrt{v_1^2 + v_2^2 + v_3^2} = 1$. Let λ be the eigenvalue of $\vec{v}\cdot\vec{\sigma}$.

$$\vec{v}\cdot\vec{\sigma} = v_1\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + v_2\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} + v_3\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}v_3 & v_1 - iv_2\\ v_1 + iv_2 & -v_3\end{bmatrix}$$

$$|\,\lambda - \vec{v}\cdot\vec{\sigma}\,| = 0$$

$$\Rightarrow \quad (\lambda - v_3)(\lambda + v_3) + (iv_2 - v_1)(iv_2 + v_1) = 0$$

$$\Rightarrow \quad \lambda^2 = v_3^2 + v_2^2 + v_1^2 = 1$$

$$\Rightarrow \quad \lambda = \pm 1$$

Let $\lambda_+ = 1$ and $\lambda_- = -1$. And, P_+ be the projector onto the eigenspace of $\vec{v}\cdot\vec{\sigma}$ corresponding to $\lambda_+ = 1$ and P_- be the projector onto the eigenspace of $\vec{v}\cdot\vec{\sigma}$ corresponding to $\lambda_- = -1$. We can see that the observable $\vec{v}\cdot\vec{\sigma}$ is Hermitian.

$$(\vec{v}\cdot\vec{\sigma})^\dagger = \left(\begin{bmatrix}v_3 & v_1 - iv_2\\ v_1 + iv_2 & -v_3\end{bmatrix}'\right)^\star = \begin{bmatrix}v_3 & v_1 + iv_2\\ v_1 - iv_2 & -v_3\end{bmatrix}^\star = \begin{bmatrix}v_3 & v_1 - iv_2\\ v_1 + iv_2 & -v_3\end{bmatrix} = \vec{v}\cdot\vec{\sigma}$$

Then, we have

$$\vec{v}\cdot\vec{\sigma} = \lambda_+ P_+ + \lambda_- P_- = P_+ - P_- \tag{4}$$

Also, projectors are Hermitian operators, which means they are self adjoint, i.e. $P^\dagger = P$. Also, $P^2 = P$. Hence from the completeness relation $\sum P^\dagger P = \sum PP = \sum P^2 = \sum P = I$, we have the following.

$$P_+ + P_- = I \tag{5}$$

Solving (2) and (3) we get $P_+ = (I + \vec{v}\cdot\vec{\sigma})/2$ and $P_- = (I - \vec{v}\cdot\vec{\sigma})/2$.

2.67) Suppose V is a Hilbert space with a subspace W. Suppose $U : W \to V$ is a linear operator which preserves inner products, i.e. for any $|w_1\rangle$ and $|w_2\rangle$ in W we have

$$\langle w_1|U^\dagger U|w_2\rangle = \langle w_1|w_2\rangle$$

Prove that there exists a Unitary operator $U' : V \to V$ which extends U, i.e. $U'|w\rangle = U|w\rangle$ for all $|w\rangle \in W$, where U' is defined in the entire space V.

Solution:

It is given that $W \subset V$ is a subspace of V. Then $V = W \bigoplus W^\perp$, where $W^\perp \subset V$ is the orthogonal complement of W. Any vector in $W \subset V$ is orthogonal to a vector in $W^\perp \subset V$. Let, the dimension of W, $\dim W = k$, i.e. the subspace W has basis vectors, say $\left\{|w_1\rangle, |w_2\rangle, \ldots, |w_k\rangle\right\}$. Then by Gram-Schmidt procedure we can add basis vectors $\left\{\left|w'_{k+1}\right\rangle, \left|w'_{k+2}\right\rangle, \ldots, \left|w'_{k+m}\right\rangle\right\} \in W^\perp$, such that the basis vectors in the entire space V are $\left\{|w_1\rangle, |w_2\rangle, \ldots, |w_k\rangle, \left|w'_{k+1}\right\rangle, \left|w'_{k+2}\right\rangle, \ldots, \left|w'_{k+m}\right\rangle\right\}$. (Note that $\dim V = \dim W + \dim W^\perp$).

Now, by definition a vector in W is orthogonal to a vector in $W^{\perp}$. So, we have

$$\left\langle w_i \middle| w'_j \right\rangle = 0 \quad \text{for} \quad \begin{cases} i = 1,2,\ldots,k \\ j = k+1, k+2, \ldots, k+m \end{cases}$$

Next, we see that

$$\text{image}(U) = \left\{ U\left|w\right\rangle = \left|u\right\rangle \in V \;\middle|\; \left|w\right\rangle \in W \right\}$$

Hence, $\left|u_i\right\rangle = U\left|w_i\right\rangle \in \text{image}(U), i = 1,\ldots,k$. Now, the problem tells us that such a U is a linear operator that preserves inner products, i.e. for any $\left|w_1\right\rangle$ and $\left|w_2\right\rangle \in W$, we have $\left\langle u_1|u_2\right\rangle = \left\langle w_1\right|U^{\dagger}U\left|w_2\right\rangle = \left\langle w_1|w_2\right\rangle$. Hence, U preserves orthogonality.
Therefore, using Gram-Schmidt we can extend this for $i = k+1,\ldots,k+m$. We can add vectors $\left|u'_{k+1}\right\rangle, \left|u'_{k+2}\right\rangle, \ldots, \left|u'_{k+m}\right\rangle$ such that $\left|u'_i\right\rangle \in (\text{image}(U))^{\perp} \subset V, i = k+1,\ldots,k+m$. Also, $\left\langle u_i \middle| u'_j \right\rangle = 0$ for $i = 1,\ldots,k$ and $j = k+1,\ldots,k+m$. There is a linear map defined in $W^{\perp}$ that maps vectors $\left|w'_i\right\rangle \in W^{\perp}$ to vectors $\left|u'_i\right\rangle \in V$.

Next, using $\left|w_i\right\rangle, \left|u_i\right\rangle, \left|w'_i\right\rangle, \left|u'_i\right\rangle$ we construct the following matrix.

$$\sum_{i=1}^{k} \left|u_i\right\rangle\left\langle w_i\right| + \sum_{i=k+1}^{k+m} \left|u'_i\right\rangle\left\langle w'_i\right| \tag{6}$$

We operate the matrix constructed above on a vector $\left|v\right\rangle \in V$. We want to see if the matrix defined in (4) sends vectors defined in V to V itself. Since we have considered $V = W \bigoplus W^{\perp}$, any vector $\left|v\right\rangle \in V$ can be uniquely represented as $\left|v\right\rangle = \sum_{i=1}^{k} \alpha_i \left|w_i\right\rangle + \sum_{i=k+1}^{k+m} \beta_i \left|w'_i\right\rangle$.

$$\begin{aligned}
&\left(\sum_{i=1}^{k} \left|u_i\right\rangle\left\langle w_i\right| + \sum_{i=k+1}^{k+m} \left|u'_i\right\rangle\left\langle w'_i\right|\right)\left|v\right\rangle \\
&= \left(\sum_{i=1}^{k} \left|u_i\right\rangle\left\langle w_i\right| + \sum_{i=k+1}^{k+m} \left|u'_i\right\rangle\left\langle w'_i\right|\right)\left(\sum_{i=1}^{k} \alpha_i \left|w_i\right\rangle + \sum_{i=k+1}^{k+m} \beta_i \left|w'_i\right\rangle\right) \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k} \alpha_j \left|u_i\right\rangle\left\langle w_i|w_j\right\rangle + \sum_{i=1}^{k}\sum_{j=k+1}^{k+m} \beta_j \left|u_i\right\rangle\left\langle w_i \middle| w'_j\right\rangle + \sum_{i=k+1}^{k+m}\sum_{j=1}^{k} \alpha_j \left|u'_i\right\rangle\left\langle w'_i|w_j\right\rangle + \sum_{i=k+1}^{k+m}\sum_{j=k+1}^{k+m} \beta_j \left|u'_i\right\rangle\left\langle w'_i \middle| w'_j\right\rangle \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k} \alpha_j \left|u_i\right\rangle \delta_{ij} + \sum_{i=k+1}^{k+m}\sum_{j=k+1}^{k+m} \beta_j \left|u'_i\right\rangle \delta_{ij} \\
&= \sum_{i=1}^{k} \alpha_i \left|u_i\right\rangle + \sum_{i=k+1}^{k+m} \beta_i \left|u'_i\right\rangle \in V
\end{aligned}$$

We see that in the four double summation terms above, the second and the third double summation terms vanish because of the fact that $\left|w_i\right\rangle \in W$ is orthogonal to $\left|w'_j\right\rangle \in W^{\perp}$. So, the matrix $\sum_{i=1}^{k} \left|u_i\right\rangle\left\langle w_i\right| + \sum_{i=k+1}^{k+m} \left|u'_i\right\rangle\left\langle w'_i\right|$ maps a $\left|v\right\rangle \in V$ to V. Next, we define this matrix as

$$U' = \sum_{i=1}^{k} \left|u_i\right\rangle\left\langle w_i\right| + \sum_{i=k+1}^{k+m} \left|u'_i\right\rangle\left\langle w'_i\right|. \tag{7}$$

So, there exists a linear map U' that sends $\left|v\right\rangle \in V$ to V.

Next, we want to show that such a linear map defined in the entire space V agrees with

U defined in W on any $|w\rangle \in W$, i.e., $U|w\rangle = U'|w\rangle$.

$$\begin{aligned} U'|w\rangle &= \left(\sum_{i=1}^{k} |u_i\rangle\langle w_i| + \sum_{i=k+1}^{k+m} |u_i'\rangle\langle w_i'|\right)|w\rangle \\ &= \sum_{i=1}^{k} |u_i\rangle\langle w_i|w\rangle + \sum_{i=k+1}^{k+m} |u_i'\rangle\langle w_i'|w\rangle \end{aligned}$$

Now, any $|w\rangle \in W$ can be written in terms of its basis vectors $|w_i\rangle$. And because $|w_i'\rangle \in W^{\perp}$ is orthogonal to $|w_i\rangle \in W$, the terms in the second summation evaluate to zero. Hence, we have

$$U'|w\rangle = \sum_{i=1}^{k} |u_i\rangle\langle w_i|w\rangle = \left(\sum_{i=1}^{k} |u_i\rangle\langle w_i|\right)|w\rangle = U|w\rangle \tag{8}$$

where $\sum_{i=1}^{k} |u_i\rangle\langle w_i|$ is the outer product expression for the matrix U defined in the subspace W.

Next, we check that such a U' as defined above is unitary. We want to show that $U'^{\dagger}U' = I$.

$$U'^{\dagger}U' = \left(\sum_{i=1}^{k} |w_i\rangle\langle u_i| + \sum_{i=k+1}^{k+m} |w_i'\rangle\langle u_i'|\right)\left(\sum_{i=1}^{k} |u_i\rangle\langle w_i| + \sum_{i=k+1}^{k+m} |u_i'\rangle\langle w_i'|\right)$$

There are double summation terms that involve inner products $\langle u_i|u_i'\rangle$ and $\langle w_i|w_i'\rangle$ which are zero by virtue of orthogonality. The reduced expression reads

$$\begin{aligned} U'^{\dagger}U' &= \sum_{i=1}^{k}\sum_{j=1}^{k} |w_i\rangle\langle u_i|u_j\rangle\langle w_j| + \sum_{i=k+1}^{k+m}\sum_{j=k+1}^{k+m} |w_i'\rangle\langle u_i'|u_j'\rangle\langle w_j'| \\ &= \sum_{i=1}^{k} |w_i\rangle\langle u_i|u_i\rangle\langle w_i| + \sum_{i=k+1}^{k+m} |w_i'\rangle\langle u_i'|u_i'\rangle\langle w_i'| \\ &= \sum_{i=1}^{k} |w_i\rangle\langle w_i| + \sum_{i=k+1}^{k+m} |w_i'\rangle\langle w_i'| \\ &= I \end{aligned}$$

Hence, $U'^{\dagger}U' = I$. Similarly, we can show that $U'U'^{\dagger} = I$. Hence, U' is unitary.

2.73) Let ρ be a density operator. A minimum ensemble is an ensemble $\{p_i, |\psi_i\rangle\}$ containing a number of elements equal to the rank of ρ. Let $|\psi\rangle$ be any vector in the support of ρ. A support for a Hermitian operator is the space spanned by the eigenvectors with non-zero eigenvalues. Show that there is a minimum ensemble for ρ containing $|\psi\rangle$. In any such ensemble $|\psi\rangle$ may appear with a probability

$$p_i = \frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}$$

Solution:

Let the dimension for the state space of the entire system be d. Since the density opera-

tor, ρ is positive, it has a spectral decomposition. So, $\rho = \sum_{j=1}^{d} \lambda_j |x_j\rangle\langle x_j|$, where λ_j's are the eigenvalues and $|x_j\rangle$'s are the corresponding eigenvectors. The support of ρ is the space spanned by the eigenvectors corresponding to the non-zero eigenvalues. Let the dimension of the support be k ($\le d$), i.e. the number of $|x_j\rangle$'s spanning the support is k. Hence, the spectral decomposition can be rewritten as $\rho = \sum_{j=1}^{k} \lambda_j |x_j\rangle\langle x_j|$.
Next, consider a state vector $|\psi_i\rangle$ in the support of ρ. We can write $|\psi_i\rangle$ as a linear combination of $|x_j\rangle$'s since the $|x_j\rangle$'s span the support of ρ. So, $|\psi_i\rangle = \sum_{j=1}^{k} c_{ij} |x_j\rangle$.
In order to show that there exists a minimal ensemble containing any such $|\psi_i\rangle$ we create an ensemble of states $\{|\psi_i\rangle\}_{i=1}^{k}$ and show that the density matrix obtained from such a set is equal to the density $\rho = \sum_{j=1}^{k} \lambda_j |x_j\rangle\langle x_j|$.
Next, ρ when considered as an operator acting on its own support is full rank. This can be inferred from the diagonal representation of $\rho = \sum_{i=1}^{k} \lambda_i |x_i\rangle\langle x_i|$, where all the $|x_i\rangle$'s are the eigenvectors corresponding to non-zero λ_i's. Consider a density matrix $\sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i|$ where p_i is the probability of the states $|\psi_i\rangle$ in the support of ρ. Define a unitary relationship between $|\psi_i\rangle$ and $|\lambda_j\rangle$ as follows:

$$\sqrt{p_i}\,|\psi_i\rangle = \sum_{j=1}^{k} \sqrt{\lambda_j} u_{ij} |x_j\rangle$$

where u_{ij} is a unitary matrix of complex numbers with indices i and j. Plugging in the expression for $|\psi_i\rangle$, i.e. $|\psi_i\rangle = \sum_{j=1}^{k} c_{ij} |x_j\rangle$ in the above expression and comparing term-by-term we have

$$\sqrt{p_i}\sum_{j=1}^{k} c_{ij} |x_j\rangle = \sum_{j=1}^{k} \sqrt{\lambda_j} u_{ij} |x_j\rangle \Rightarrow \sqrt{p_i} c_{ij} = \sqrt{\lambda_j} u_{ij} \Rightarrow u_{ij} = \frac{\sqrt{p_i} c_{ij}}{\sqrt{\lambda_j}}$$

Since u_{ij} are the elements of a unitary matrix, it must satisfy $\sum_{j=1}^{k} |u_{ij}|^2 = 1$. Plugging in the expression for u_{ij} obtained above we have

$$\sum_{j=1}^{k} \frac{p_i |c_{ij}|^2}{\lambda_j} = 1 \Rightarrow p_i \sum_{j=1}^{k} \frac{|c_{ij}|^2}{\lambda_j} = 1 \Rightarrow p_i = \frac{1}{\sum_{j=1}^{k} \frac{|c_{ij}|^2}{\lambda_j}} = \left(\sum_{j=1}^{k} \frac{|c_{ij}|^2}{\lambda_j}\right)^{-1} \qquad (\star)$$

Now, we show that the density matrix $\sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i|$ refer to the same density matrix $\rho = \sum_{j=1}^{k} \lambda_j |x_j\rangle\langle x_j|$.

$$\sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{k} \sqrt{p_i}|\psi_i\rangle \sqrt{p_i}\langle\psi_i| = \sum_{i=1}^{k}\sum_{j=1}^{k} \sqrt{\lambda_j} u_{ij} |x_j\rangle \sum_{j=1}^{k} \sqrt{\lambda_j} u_{ji}^{\star} \langle x_j|$$
$$= \sum_{i=1}^{k}\sum_{jl} \sqrt{\lambda_j \lambda_l} u_{ij} u_{li}^{\star} |x_i\rangle\langle x_l| = \sum_{jl} \sqrt{\lambda_j\lambda_l} \underbrace{\sum_{i=1}^{k} u_{ij} u_{li}^{\star}}_{\delta_{jl}} |x_i\rangle\langle x_l| = \sum_{j=1}^{k} \lambda_j |x_j\rangle\langle x_j| = \rho$$

where $\sum_{i=1}^{k} u_{ij} u_{li}^{\star} = \delta_{jl}$ since u_{ij} is a unitary matrix.
Hence we have shown that $\sum_{i=1}^{k} p_i |\psi_i\rangle\langle\psi_i| = \rho = \sum_{j=1}^{k} \lambda_j |x_j\rangle\langle x_j|$. So the sets $\{p_i, |\psi_i\rangle\}_{i=1}^{k}$ and $\{\lambda_j, |x_j\rangle\}_{j=1}^{k}$ generate the same density matrix. Therefore, we have shown that we can construct a minimal ensemble $\{p_i, |\psi_i\rangle\}_{i=1}^{k}$ containing the state $|\psi_i\rangle$ in the support of ρ.

Now, we want to obtain the expression of the probability p_i of the state $|\psi_i\rangle$ in the minimal ensemble we just constructed.
Since ρ is full rank (when acting on its own support), it is invertible, i.e. ρ^{-1} exists. We can also see this from the eigen-decomposition of ρ which is $\sum_{j=1}^{k}\lambda_j|x_j\rangle\langle x_j|$. Since $\det(\rho)=\lambda_1\lambda_2\cdots\lambda_k$ and all the λ_j's are non-zero, we can see that ρ^{-1} exists. In fact, ρ^{-1} is also Hermitian. This is due to one of the properties of Hermitian matrices which states that the inverse of an invertible Hermitian matrix is Hermitian. Indeed, suppose if A is Hermitian and invertible then $A^{-1}A=AA^{-1}=I$.
Then $I^\dagger=I\Rightarrow(AA^{-1})^\dagger=AA^{-1}\Rightarrow(A^{-1})^\dagger A^\dagger=AA^{-1}\Rightarrow(A^{-1})^\dagger A^\dagger=A^{-1}A\Rightarrow(A^{-1})^\dagger A=A^{-1}A$ and hence $(A^{-1})^\dagger=A^{-1}$.
Next, the eigenvalues of ρ^{-1} are λ_j^{-1} for $i=1,\ldots,k$ and the eigen-decomposition is $\rho^{-1}=\sum_{j=1}^{k}\lambda_j^{-1}|x_j\rangle\langle x_j|$. So, ρ^{-1} can be considered as an observable with outcomes λ_j^{-1} in the support of ρ.
Therefore, the probability of $|\psi_j\rangle$ in the minimal ensemble can be obtained as shown below.

$$\langle\psi_i|\rho^{-1}|\psi_i\rangle=\langle\psi_i|\left(\sum_{j=1}^{k}\lambda_j^{-1}|x_j\rangle\langle x_j|\right)\langle\psi_i|=\sum_{j=1}^{k}\lambda_j^{-1}\langle\psi_i|x_j\rangle\langle x_j|\psi_i\rangle=\sum_{j=1}^{k}\lambda_j^{-1}|\langle\psi_i|x_j\rangle|^2$$

Now, from $\psi_i=\sum_{j=1}^{k}c_{ij}|x_j\rangle$ we have $\langle\psi_i|x_j\rangle=c_{ij}$ and hence

$$\langle\psi_i|\rho^{-1}|\psi_i\rangle=\sum_{j=1}^{k}\lambda_j^{-1}|\langle\psi_i|x_j\rangle|^2=\sum_{j=1}^{k}\lambda_j^{-1}|c_{ij}|^2=\sum_{j=1}^{k}\frac{|c_{ij}|^2}{\lambda_j}\underbrace{=}_{\text{From }(\star)}p_i^{-1}$$
$$\Rightarrow p_i=\left(\langle\psi_i|\rho^{-1}|\psi_i\rangle\right)^{-1}=\frac{1}{\langle\psi_i|\rho^{-1}|\psi_i\rangle}$$

2.78) a) Prove that a state $|\psi\rangle$ of a composite system AB is a product state if and only if it has a Schmidt number 1. b) Prove that $|\psi\rangle$ is a product state if and only if ρ_A (and thus ρ_B) are pure states.

Solution:

a) Schmidt Number $=1\Rightarrow|\psi\rangle$ is a product state:

It is given that the Schmidt number is 1. So, there exists only one non-zero λ_i such that $|\psi\rangle=\sum_i\lambda_i|i_A\rangle|i_B\rangle=\lambda|i_A\rangle|i_B\rangle$. The condition $\sum_i\lambda_i^2=1$ and the fact that there is only one such λ_i implies that $\lambda_i=\lambda=1$. Hence, $|\psi\rangle=\lambda|i_A\rangle|i_B\rangle=|i_A\rangle|i_B\rangle=|i_A\rangle\otimes|i_B\rangle$.
In addition, we see that

$$\rho_{AB}=|\psi\rangle\langle\psi|=|i_A\rangle|i_B\rangle\langle i_A|\langle i_B|. \tag{9}$$

Now, $\rho_A=\mathrm{tr}_B(\rho_{AB})=\mathrm{tr}_B|i_A\rangle|i_B\rangle\langle i_A|\langle i_B|=|i_A\rangle\langle i_A|\langle i_B|i_B\rangle=|i_A\rangle\langle i_A|$.
Similarly, $\rho_B=\mathrm{tr}_A(\rho_{AB})=|i_B\rangle\langle i_B|$. We take the (tensor) product of the two density matrices ρ_A and ρ_B.

$$\rho_A \otimes \rho_B = |i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B| = |i_A\rangle \otimes |i_B\rangle\langle i_A| \otimes \langle i_B| = |i_A\rangle|i_B\rangle\langle i_A|\langle i_B| = \rho_{AB} \tag{10}$$

where we have used the property of tensor products: $(S_1T_1)\otimes(S_2T_2) = (S_1 \otimes S_2)(T_2 \otimes T_2)$.

$|\psi\rangle$ is a product state $\Rightarrow$ Schmidt number = 1:

The Schmidt number is always greater than or equal to 1.
Now, we are given that $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. We assume that the states $|\psi_A\rangle$ and $|\psi_B\rangle$ are normalized, i.e. they are of unit length. We claim that $|\psi_{AB}\rangle$ is already in the Schmidt decomposed form. This is because $|\psi_A\rangle$ and $|\psi_B\rangle$ are unit vectors (by assumption). And we can find normalized states $\{|\psi'_A\rangle\}$ and $\{|\psi'_B\rangle\}$ that are orthogonal to $\{|\psi_A\rangle\}$ and $\{|\psi_B\rangle\}$ respectively and write the decomposition as $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle + \sum 0 \cdot |\psi'_A\rangle \otimes |\psi'_B\rangle$.
This shows that the Schmidt number of $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ is 1.

Next, we show that $|\psi\rangle$ is a product state if and only if ρ_A and ρ_B are pure states.

b) $|\psi\rangle$ is a product state $\Rightarrow$ ρ_A and ρ_B are pure states:

$|\psi\rangle$ is a product state. As shown in part a) of the problem the Schmidt number is equal to 1. Hence, the Schmidt decomposition of $|\psi\rangle$ is $|\psi\rangle = \lambda |i_A\rangle|i_B\rangle$ and because λ must satisfy the condition $\lambda^2 = 1$ and $\lambda \geq 0$, we have $\lambda = 1$. So, $|\psi\rangle = |i_A\rangle|i_B\rangle$. Therefore, the density matrix for $|\psi\rangle$ is $\rho_{AB} = |\psi\rangle\langle\psi| = |i_A i_B\rangle\langle i_A i_B|$. Taking the partial trace with respect to A and B respectively gives us ρ_A and ρ_B.

$$\rho_A = \mathrm{tr}_B \rho_{AB} = \mathrm{tr}_B(|i_A i_B\rangle\langle i_A i_B|) = |i_A\rangle\langle i_A| \mathrm{tr}(|i_B\rangle\langle i_B|) = |i_A\rangle\langle i_A|\langle i_B|i_B\rangle = |i_A\rangle\langle i_A|$$

Taking the trace of ρ_A^2 gives us the following:

$$\mathrm{tr}\left(\rho_A^2\right) = \mathrm{tr}(|i_A\rangle\langle i_A|i_A\rangle\langle i_A|) = \mathrm{tr}(|i_A\rangle\langle i_A|) = \mathrm{tr}\rho_A = 1$$

Similarly, obtaining ρ_B using $\mathrm{tr}_A \rho_{AB}$ and taking the trace $\mathrm{tr}\left(\rho_B^2\right)$ we get $\mathrm{tr}\left(\rho_B^2\right) = 1$. Hence, ρ_A and ρ_B are pure states.

ρ_A and ρ_B are pure states $\Rightarrow$ $|\psi\rangle$ is a product state:

Consider the Schmidt decomposition: $|\psi\rangle = \sum_i \lambda_i |i_A\rangle|i_B\rangle$.
Given that ρ_A and ρ_B are pure states, we show that $|\psi\rangle$ is a product state, i.e. the Schmidt decomposition turns out to be $|\psi\rangle = |i_A\rangle|i_B\rangle$. It suffices to show that the Schmidt number of $|\psi\rangle$ is equal to 1. Now, we have $\rho_A = \sum_i \lambda_i^2 |i_A\rangle\langle i_A|$ and $\rho_B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B|$ by taking the partial trace of ρ_{AB} with respect to B and A respectively.

Also,

$$\rho_A^2 = \sum_i \lambda_i^2 |i_A\rangle\langle i_A| \sum_i \lambda_i^2 |i_A\rangle\langle i_A| = \sum_{ij} \lambda_i^2 \lambda_j^2 |i_A\rangle \underbrace{\langle i_A | j_A\rangle}_{=\delta_{ij}} \langle j_A| = \sum_i \lambda_i^4 |i_A\rangle\langle i_A|$$

Similarly, $\rho_B^2 = \sum_i \lambda_i^4 |i_B\rangle\langle i_B|$. Since ρ_A and ρ_B are pure states,

$$\mathrm{tr}\left(\rho_A^2\right) = \mathrm{tr}\left(\sum_i \lambda_i^4 |i_A\rangle\langle i_A|\right) = 1$$
$$\mathrm{tr}\left(\rho_B^2\right) = \mathrm{tr}\left(\sum_i \lambda_i^4 |i_B\rangle\langle i_B|\right) = 1$$

Note that $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ are respectively sets of orthonormal states. By Gram-Schmidt procedure we can extend them to orthonormal basis sets. And hence $\sum_i \lambda_i^4 |i_A\rangle\langle i_A|$ and $\sum_i \lambda_i^4 |i_B\rangle\langle i_B|$ are the diagonal representations of ρ_A^2 and ρ_B^2 respectively. The trace conditions (obtained above) for both of them show that $\sum_i \lambda_i^4 = 1$. Also, since ρ_A and ρ_B are densities, by their expressions obtained from the partial trace of ρ_{AB} we must have $\mathrm{tr}\rho_A = \mathrm{tr}\rho_B = \sum_i \lambda_i^2 = 1$. Hence, λ_i's must satisfy: $\sum_i \lambda_i^4 = \sum_i \lambda_i^2 = 1$ with the constraint $\lambda_i \geq 0$ (with at least one $\lambda_i > 0$). The only permissible solution is: $\lambda_i = 1$ for some index i and the rest $\lambda_{\neq i} = 0$.
Hence, we have shown that the Schmidt number is 1 and therefore, the Schmidt decomposition of $|\psi\rangle$ turns out to be $|i_A\rangle|i_B\rangle$ which shows that it is a product state.

2.80) Suppose $|\psi\rangle$ and $|\phi\rangle$ are two pure states of a composite quantum system with components A and B, with identical Schmidt coefficients. Show that there are unitary transformations U on system A and V on system B such that $|\psi\rangle = (U \otimes V)|\phi\rangle$.

Solution:

We represent the pure states of the composite quantum system AB in their Schmidt decomposition form. The states $|\psi\rangle$ and $|\phi\rangle$ can be written as

$$|\psi\rangle = \sum_i \lambda_i \left|x_i^A\right\rangle \left|x_i^B\right\rangle$$
$$|\phi\rangle = \sum_i \lambda_i \left|y_i^A\right\rangle \left|y_i^B\right\rangle$$

The vectors $|x_i^A\rangle$, $|x_i^B\rangle$, $|y_i^A\rangle$, $|y_i^B\rangle$ are orthonormal bases. λ_i is the Schmidt coefficient.

Note: In *Nielsen and Chuang* $|x_i^A\rangle$, $|x_i^B\rangle$, $|y_i^A\rangle$, $|y_i^B\rangle$ have been referred to as orthonormal states. However, it depends on how we choose to see it. If we allow $\lambda_i = 0$ in the expression for Schmidt decomposition, we can refer $|x_i^A\rangle$, $|x_i^B\rangle$, $|y_i^A\rangle$, $|y_i^B\rangle$ as orthonormal bases. The following example explains it further.
For the state $|\psi\rangle = \frac{1}{2}(|00\rangle - |01\rangle - |10\rangle + |11\rangle)$ in the composite system AB we can write $|\psi\rangle$ in the Schmidt decomposition formula as $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle|-\rangle$ with a Schmidt coefficient of $\lambda_1 = 1$. However, we can also see it as $|\psi\rangle = 1 \times |-\rangle|-\rangle + 0 \times |+\rangle|+\rangle$. And $|-\rangle \left(= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right)$ and $|+\rangle = \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right)$ are orthonormal bases for the state space of both A and B.

Next, we define linear maps $U: A \to A$ and $V: B \to B$ as follows.

$$U = \sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right| \quad V = \sum_i \left|x_i^B\right\rangle \left\langle y_i^B\right|$$

We can check that they are unitary, i.e. they satisfy $U^\dagger U = UU^\dagger = I$ and $V^\dagger V = VV^\dagger = I$.

$$U^\dagger U = \left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right)^\dagger \left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right) = \left(\sum_i \left|y_i^A\right\rangle \left\langle x_i^A\right|\right)\left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right) = \sum_{ij} \left|y_i^A\right\rangle \underbrace{\left\langle x_i^A \middle| x_j^A\right\rangle}_{=\delta_{ij}} \left\langle y_j^A\right| = \underbrace{\sum_i \left|y_i^A\right\rangle \left\langle y_i^A\right|}_{=I}$$

Since $|y_i^A\rangle$'s are orthonormal $|y_i^A\rangle\langle y_i^A|$'s are projective measurements and they follow the completeness relation, i.e. $\sum_i |y_i^A\rangle\langle y_i^A| = I$. The same is true for $|x_i^A\rangle$'s.

$$UU^\dagger = \left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right)\left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right)^\dagger = \left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right)\left(\sum_i \left|y_i^A\right\rangle \left\langle x_i^A\right|\right) = \sum_{ij} \left|x_i^A\right\rangle \underbrace{\left\langle y_i^A \middle| y_j^A\right\rangle}_{=\delta_{ij}} \left\langle x_j^A\right| = \underbrace{\sum_i \left|x_i^A\right\rangle \left\langle x_i^A\right|}_{=I}$$

Similarly, we can show that $V = \sum_i |x_i^B\rangle\langle y_i^B|$ satisfies $V^\dagger V = VV^\dagger = I$. Hence, both U and V are unitary.
Next, we do the following.

$$(U \otimes V)|\phi\rangle = (U \otimes V)\sum_i \lambda_i \left|y_i^A\right\rangle \left|y_i^B\right\rangle = \sum_i \lambda_i U \left|y_i^A\right\rangle \otimes V \left|y_i^B\right\rangle$$
$$= \sum_i \lambda_i \underbrace{\left(\sum_i \left|x_i^A\right\rangle \left\langle y_i^A\right|\right)\left|y_i^A\right\rangle}_{=|x_i^A\rangle} \otimes \underbrace{\left(\sum_i \left|x_i^B\right\rangle \left\langle y_i^B\right|\right)\left|y_i^B\right\rangle}_{=|x_i^B\rangle} = \sum_i \lambda_i \left|x_i^A\right\rangle \left|x_i^B\right\rangle = |\psi\rangle$$

Note that $\left(\sum_i |x_i^A\rangle\langle y_i^A|\right)|y_i^A\rangle = \sum_k |x_k^A\rangle\langle y_k^A|y_i^A\rangle\delta_{ik} = |x_i^A\rangle$ and $\left(\sum_i |x_i^B\rangle\langle y_i^B|\right)|y_i^B\rangle = \sum_k |x_k^B\rangle\langle y_k^B|y_i^B\rangle\delta_{ik} = |x_i^B\rangle$. Hence, we have shown that there exists unitary transformations U on system A and V on system B such that $|\psi\rangle = (U \otimes V)|\phi\rangle$.

> **2.81) (Freedom in purifications)** Let $|AR_1\rangle$ and $|AR_2\rangle$ be two purifications of a state ρ^A to a composite system AR. Prove that there exists a unitary transformation U_R acting on system R such that $|AR_1\rangle = (I_A \otimes U_R)|AR_2\rangle$.

Solution:

It is given that both pure states $|AR_1\rangle$ and $|AR_2\rangle$ are purifications of the impure state ρ^A (in system A). Hence, it must be that

$$\rho^A = \mathrm{tr}_R\, |AR_1\rangle\langle AR_1| = \mathrm{tr}_R\, |AR_2\rangle\langle AR_2| \tag{11}$$

Suppose the pure states $|AR_1\rangle$ and $|AR_2\rangle$ have Schmidt decomposition: $|AR_1\rangle = \sum_i \lambda_i \left|u_i^A\right\rangle\left|u_i^R\right\rangle$ and $|AR_2\rangle = \sum_j \mu_j \left|v_j^A\right\rangle\left|v_j^R\right\rangle$. We consider the states $\left|u_i^A\right\rangle, \left|u_i^R\right\rangle, \left|v_j^A\right\rangle, \left|v_j^R\right\rangle$ to be orthonormal basis in systems A and R. Then the Schmidt number or Schmidt rank is the number of non-zero λ_i (and μ_j) in the Schmidt decomposition of $|AR_1\rangle$ (and $|AR_2\rangle$). From (11) we have

$$\rho^A = \sum_i \lambda_i^2 \left|u_i^A\right\rangle \left\langle u_i^A\right| = \sum_j \mu_j^2 \left|v_j^A\right\rangle \left\langle v_j^A\right| \tag{12}$$

The Schmidt coefficients λ_i (and μ_j) are the square root of the eigenvalues of the reduced density matrix (impure state) ρ^A. The eigenvalues (and hence the coefficients) are unique, so for (2) we must have $\lambda_i = \mu_j$. With the assumption that the eigenvalues are non-degenerate, the eigenvectors $|u_i^A\rangle$ and $\left|v_j^A\right\rangle$ associated with λ_i and μ_j respectively are also uniquely determined. If $\lambda_i = \mu_j$, then it must be that $|u_i^A\rangle = \left|v_j^A\right\rangle$.

Next, we define $U_R = \sum_{i=1}^k |u_i^R\rangle\langle v_i^R|$ where $k = \mathrm{Sch}(AR_1) = \mathrm{Sch}(AR_2)$.

$$(I_A \otimes U_R)|AR_2\rangle = (I_A \otimes U_R)\sum_{j=1}^k \mu_j \left|v_j^A\right\rangle \otimes \left|v_j^R\right\rangle = \sum_{j=1}^k \mu_j I \left|v_j^A\right\rangle \otimes U_R \left|v_j^R\right\rangle = \sum_{j=1}^k \lambda_j \left|u_j^A\right\rangle \left|u_j^R\right\rangle = |AR_1\rangle$$

Hence, $|AR_1\rangle = \left(I_A \otimes U_R\right)|AR_2\rangle$.

Note: $U_R \left|v_j^R\right\rangle = \left(\sum_{i=1}^k |u_i^R\rangle\langle v_i^R|\right)\left|v_j^R\right\rangle = \sum_{i=1}^k \left|u_i^R\right\rangle \left\langle v_i^R \middle| v_j^R\right\rangle \delta_{ij} = \left|u_j^R\right\rangle$.

2.82) Suppose $\left\{p_i, |\psi_i\rangle\right\}$ is an ensemble of states generating a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ for a quantum system A. Introduce a system R with an orthonormal basis $|i\rangle$.
(1) Show that $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ is a purification of ρ.
(2) Suppose we measure R in the basis $|i\rangle$, obtaining outcome i. With what probability do we obtain the result i, and what is the corresponding state of system A ?
(3) Let $|AR\rangle$ be any purification of ρ to the system AR. Show that there exists an orthonormal basis $|i\rangle$ in which R can be measured such that the corresponding post-measurement state for system A is $|\psi_i\rangle$ with probability p_i.

Solution:

(1) We denote $\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ as $|AR\rangle$. $|\psi_i\rangle$ is any state in the system A (not necessarily orthonormal). $|AR\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$

$$|AR\rangle\langle AR| = \left(\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle\right)\left(\sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle\right)$$

$$\Rightarrow |AR\rangle\langle AR| = \sum_{ij}\sqrt{p_i p_j}|\psi_i\rangle|i\rangle\langle\psi_j|\langle j|$$

$$\mathrm{tr}_R(|AR\rangle\langle AR|) = \sum_{ij}\sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j|\mathrm{tr}_R(|i\rangle\langle j|) = \sum_{ij}\sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j|\langle j|i\rangle = \sum_{ij}\sqrt{p_i p_j}|\psi_i\rangle\langle\psi_j|\delta_{ij}$$

$$\Rightarrow \mathrm{tr}_R(|AR\rangle\langle AR|) = \sum_i p_i|\psi_i\rangle\langle\psi_i| = \rho$$

Hence $\sum_i p_i|\psi_i\rangle|i\rangle$ is a purification of ρ since the partial trace $\mathrm{tr}_R(|AR\rangle\langle AR|)$ gives the density matrix ρ generated by the ensemble of states $\{p_i, |\psi_i\rangle\}$.

(2) We want to obtain the outcome i by measuring system R in the basis $|i\rangle$. *To measure in basis* $|i\rangle$ means to perform projective measurement using the projectors $|i\rangle\langle i|$. We use the projective operator $I\otimes|i\rangle\langle i|$ in the space of system AR on the state $|AR\rangle=\sum_j\sqrt{p_j}\,|\psi_j\rangle|j\rangle$ to measure R. The probability of obtaining the outcome i is

$$\begin{aligned}
&\Big(\sum_j\sqrt{p_j}\,|\psi_j\rangle|j\rangle\Big)^\dagger\Big(I\otimes|i\rangle\langle i|\Big)\Big(\sum_j\sqrt{p_j}\,|\psi_j\rangle|j\rangle\Big)\\
&=\Big(\sum_j\sqrt{p_j}\,\langle\psi_j|\langle j|\Big)\Big(I\otimes|i\rangle\langle i|\Big)\Big(\sum_j\sqrt{p_j}\,|\psi_j\rangle|j\rangle\Big)\\
&=\Big(\sum_j\sqrt{p_j}\,\langle\psi_j|\langle j|\Big)\Big(\sum_j\sqrt{p_j}I\,|\psi_j\rangle\otimes|i\rangle\langle i|j\rangle\delta_{ij}\Big)\\
=\Big(\sum_j\sqrt{p_j}\,\langle\psi_j|\langle j|\Big)&\sqrt{p_i}\,|\psi_i\rangle\otimes|i\rangle=\sum_j\sqrt{p_j}\sqrt{p_i}\,\langle\psi_j|\psi_i\rangle\delta_{ij}=p_i\,\langle\psi_i|\psi_i\rangle=p_i
\end{aligned}$$

Hence, the probability of obtaining the outcome i by measuring R in the basis $|i\rangle$ is p_i. The post measurement state of the system AR is

$$\frac{\Big(I\otimes|i\rangle\langle i|\Big)|AR\rangle}{\sqrt{p_i}}=\frac{\Big(I\otimes|i\rangle\langle i|\Big)\sum_j\sqrt{p_j}\,|\psi_j\rangle|j\rangle}{\sqrt{p_i}}=\frac{\sqrt{p_i}\,|\psi_i\rangle|i\rangle}{\sqrt{p_i}}=|\psi_i\rangle|i\rangle$$

We see that the corresponding state of system A is $|\psi_i\rangle$.

(3) The density matrix generated by the ensemble of states $\{p_i,|\psi_i\rangle\}$ is $\sum_i p_i\,|\psi_i\rangle\langle\psi_i|$. *We assume the states $|\psi_i\rangle$ to be normalized.* However, they are not necessarily orthogonal. $|AR\rangle$ is a purification of ρ to the system AR. Let the orthonormal decomposition of ρ be $\sum_k\mu_k\,|u_k^A\rangle\langle u_k^A|$. As a consequence of Theorem 2.6 (unitary freedom in the ensemble of density matrices) we can say that $\rho=\sum_k\mu_k\,|u_k^A\rangle\langle u_k^A|=\sum_i p_i\,|\psi_i\rangle\langle\psi_i|$ for normalized states $|u_k^A\rangle$ and $|\psi_i\rangle$ if and only if

$$\sqrt{\mu_k}\,\big|u_k^A\big\rangle=\sum_i v_{ik}\sqrt{p_i}\,|\psi_i\rangle$$

where v_{ik} are the entries of a unitary matrix and we can pad the smaller ensemble with zeroes to make the two ensembles the same size.

The Schmidt representation for the pure state $|AR\rangle$ is $\sum_k\sqrt{\mu_k}\,|u_k^A\rangle|k^R\rangle$ where $|k^R\rangle$ are a set of orthonormal states in system R. Substituting the expression for $\sqrt{\mu_k}\,|u_k^A\rangle$ we have the following

$$|AR\rangle=\sum_k\sqrt{\mu_k}\,\big|u_k^A\big\rangle\big|k^R\big\rangle=\sum_k\Big(\sum_i\sqrt{p_i}v_{ik}\,|\psi_i\rangle\Big)\big|k^R\big\rangle=\sum_i\sum_k\sqrt{p_i}v_{ik}\,|\psi_i\rangle\otimes\big|k^R\big\rangle=\sum_i\sqrt{p_i}\,|\psi_i\rangle\otimes\sum_k v_{ik}\,\big|k^R\big\rangle$$

Next, we define $|i\rangle=\sum_k v_{ik}\,|k^R\rangle$. We can check if $|i\rangle$ (as defined) is orthonormal.

$$\langle j|i\rangle=\Big(\sum_k v_{jk}^\star\,\big\langle k^R\big|\Big)\sum_k v_{ik}\,\big|k^R\big\rangle=\sum_{lk}v_{jl}^\star v_{ik}^\star\underbrace{\big\langle l^R\big|k^R\big\rangle}_{=\delta_{lk}}=\underbrace{\sum_l v_{jl}^\star v_{il}=\delta_{ij}}_{\text{due to unitarity of } v_{ik}}$$

We see that $|i\rangle = \sum_k v_{ik} \left|k^R\right\rangle$ is orthonormal. But it might not be a basis of system R. We can use Gram Schmidt method to extend it to an orthonormal basis for system R. So, there exists an orthonormal basis to measure R such that the corresponding post-measurement state for system A is $\left|\psi_i\right\rangle$ with probability p_i.

Problem 2.1: (Functions of Pauli Matrices) Let $f(\cdot)$ be any function from complex numbers to complex numbers. Let $\vec{n}$ be a normalized vector in three dimensions, and let θ be real. Show that

$$f(\theta\vec{n}\cdot\vec{\sigma}) = \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma} \tag{13}$$

Solution: $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices. In Problem 2.60 it was already shown that $\vec{n}\cdot\vec{\sigma} = n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3$ has eigenvalues ± 1. Suppose $|v_+\rangle$ and $|v_-\rangle$ are the eigenvectors corresponding to the eigenvalues $+1$ and -1. Then we can write $\vec{n}\cdot\vec{\sigma}$ as

$$\vec{n}\cdot\vec{\sigma} = |v_+\rangle\langle v_+| - |v_-\rangle\langle v_-| \tag{14}$$

Also, because $|v_+\rangle\langle v_+|$ and $|v_-\rangle\langle v_-|$ are measurement operators they add up to I.

$$|v_+\rangle\langle v_+| + |v_-\rangle\langle v_-| = I \tag{15}$$

Given a function from complex numbers to complex numbers it is possible to define a corresponding operator function. If $A = \sum_a a|a\rangle\langle a|$ is the spectral decomposition of the operator A then $f(A) = \sum_a f(a)|a\rangle\langle a|$. Using this we can proceed with the given problem as shown below.

$$\begin{aligned}
\theta\vec{n}\cdot\vec{\sigma} &= \theta|v_+\rangle\langle v_+| + (-\theta)|v_-\rangle\langle v_-| \\
f(\theta\vec{n}\cdot\vec{\sigma}) &= f(\theta)|v_+\rangle\langle v_+| + f(-\theta)|v_-\rangle\langle v_-| \\
f(\theta\vec{n}\cdot\vec{\sigma}) &= \frac{f(\theta)+f(-\theta)-f(-\theta)+f(\theta)}{2}|v_+\rangle\langle v_+| + \frac{f(-\theta)+f(\theta)-f(\theta)+f(-\theta)}{2}|v_-\rangle\langle v_-| \\
&= \frac{f(\theta)+f(-\theta)}{2}|v_+\rangle\langle v_+| + \frac{f(\theta)-f(-\theta)}{2}|v_+\rangle\langle v_+| + \frac{f(\theta)+f(-\theta)}{2}|v_-\rangle\langle v_-| + \frac{f(-\theta)-f(\theta)}{2}|v_-\rangle\langle v_-| \\
&= \frac{f(\theta)+f(-\theta)}{2}\Big(|v_+\rangle\langle v_+| + |v_-\rangle\langle v_-|\Big) + \frac{f(\theta)-f(-\theta)}{2}\Big(|v_+\rangle\langle v_+| - |v_-\rangle\langle v_-|\Big) \\
\Rightarrow f(\theta\vec{n}\cdot\vec{\sigma}) &= \frac{f(\theta)+f(-\theta)}{2}I + \frac{f(\theta)-f(-\theta)}{2}\vec{n}\cdot\vec{\sigma} \quad \text{from (4) and (5)}
\end{aligned}$$

Problem 2.2: (Properties of the Schmidt Number) Suppose $|\psi\rangle$ is a pure state of a composite system with components A and B.
(1) Prove that the Schmidt number of $|\psi\rangle$ is the rank of the reduced density matrix $\rho_A = \mathrm{tr}_B |\psi\rangle\langle\psi|$. (Note that the rank of a Hermitian operator is equal to the dimension of its support).
(2) Suppose $\sum_j |\alpha_j\rangle|\beta_j\rangle$ is a representation for $|\psi\rangle$, where $|\alpha_j\rangle$ and $|\beta_j\rangle$ are (un-normalized) states for systems A and B, respectively. Prove that the number of terms in such a decomposition is greater than or equal to the Schmidt number of $|\psi\rangle$, Sch(ψ).
(3) Suppose $|\psi\rangle = \alpha|\phi\rangle + \beta|\gamma\rangle$. Prove that

$$\mathrm{Sch}(\psi) \geq \left|\mathrm{Sch}(\phi) - \mathrm{Sch}(\gamma)\right| \tag{16}$$

Solution:

1) Let the Schmidt decomposition of $|\psi\rangle$ be $\sum_i \lambda_i |u_i\rangle |v_i\rangle$, where we consider $|u_i\rangle$ and $|v_i\rangle$ as the orthonormal basis of systems A and B respectively. The Schmidt number is the number of non-zero λ_i's in the Schmidt decomposition of $|\psi\rangle$. The reduced density matrix obtained by taking the partial trace of $|\psi\rangle\langle\psi|$ with respect to B is $\mathrm{tr}_B|\psi\rangle\langle\psi| = \rho_A = \sum_i \lambda_i^2 |u_i\rangle\langle u_i|$. This is the diagonal representation of ρ_A since $|u_i\rangle$ is an orthonormal basis set in A.
Next, by the Rank-Nullity theorem we see that the rank of ρ_A is the number of non-zero λ_i's in the expression $\sum_i \lambda_i^2 |u_i\rangle\langle u_i|$. Since this expression is obtained by taking the partial trace of $|\psi\rangle\langle\psi|$ wrt B, it follows that the Schmidt number of $|\psi\rangle$ is the rank of the reduced density density matrix $\rho_A = \mathrm{tr}_B|\psi\rangle\langle\psi|$. And since the support of ρ_A (a Hermitian operator) is the space spanned by its eigenvectors corresponding to non-zero eigenvalues, it follows that the rank of ρ_A is the dimension of its support.

(2) We start with the Schmidt decomposition of $|\psi\rangle$.
$|\psi\rangle = \sum_{i=1}^{I} \lambda_i |i_A\rangle|i_B\rangle$, where $I = \mathrm{Sch}(\psi) \leq \min(d_A, d_B)$ and $d_A = \dim(A), d_B = \dim(B)$. We know the states $|i_A\rangle$ and $|i_B\rangle$ are orthonormal.
Rewriting the Schmidt decomposition as $|\psi\rangle = \sum_{i=1}^{I} \sqrt{\lambda_i}|i_A\rangle \sqrt{\lambda_i}|i_B\rangle = \sum_{i=1}^{I} |\tilde{i}_A\rangle|\tilde{i}_B\rangle$, we see that the states $|\tilde{i}_A\rangle = \sqrt{\lambda_i}|i_A\rangle$ and $|\tilde{i}_B\rangle = \sqrt{\lambda_i}|i_B\rangle$ are un-normalized states.
Next, we write $|\tilde{i}_B\rangle$ in the standard basis $|k\rangle$ of B: $|\tilde{i}_B\rangle = \sum_{k=1}^{d_B} c_{ik}|k\rangle$. Plugging this in the expression of $|\psi\rangle$ we have

$$|\psi\rangle = \sum_{i=1}^{I} |\tilde{i}_A\rangle|\tilde{i}_B\rangle = \sum_{i=1}^{I} |\tilde{i}_A\rangle \left(\sum_{k=1}^{d_B} c_{ik}|k\rangle\right) = \underbrace{\sum_{i=1}^{I}\sum_{k=1}^{d_B}}_{\substack{\text{Re-index} \\ j=1 \text{ to } Id_B}} \underbrace{|\tilde{i}_A\rangle}_{|\alpha_j\rangle} \underbrace{c_{ik}|k\rangle}_{|\beta_j\rangle} = \sum_{j=1}^{Id_B} |\alpha_j\rangle|\beta_j\rangle$$

where $|\alpha_j\rangle$ and $|\beta_j\rangle$ are un-normalized.
In the expression $\sum_j |\alpha_j\rangle|\beta_j\rangle$ the number of terms Id_B is greater than or equal to $I = \mathrm{Sch}(\psi)$.

(3) $|\psi\rangle = \alpha|\phi\rangle + \beta|\gamma\rangle$. Expressing $|\psi\rangle, |\phi\rangle$ and $|\gamma\rangle$ in the standard basis we have

$$\left|\psi\right\rangle = \alpha\left|\phi\right\rangle + \beta\left|\gamma\right\rangle$$
$$\Rightarrow \sum_{jk} a_{jk}\,|j\rangle\,|k\rangle = \alpha\sum_{jk} b_{jk}\,|j\rangle\,|k\rangle + \beta\sum_{jk} c_{jk}\,|j\rangle\,|k\rangle = \sum_{jk}\left(\alpha b_{jk} + \beta c_{jk}\right)|j\rangle\,|k\rangle$$

Comparing term by term we have $a_{jk} = \alpha b_{jk} + \beta c_{jk}$. Each of the terms a_{jk}, b_{jk}, c_{jk} are entries of matrices $\mathbf{A}$,$\mathbf{B}$ and $\mathbf{C}$ respectively. Using singular value decomposition (svd) we have $\mathbf{A} = \mathbf{U_A D_A V_A}$ where $\mathbf{D_A}$ is a diagonal matrix with non-negative entries and $\mathbf{U_A}$ and $\mathbf{V_A}$ are unitary matrices. The rank of $\mathbf{A}$ is the number of non-zero entries of $\mathbf{D_A}$. Likewise, we can use svd on $\mathbf{B}$ and $\mathbf{C}$. From $a_{jk} = \alpha b_{jk} + \beta c_{jk}$ we have $\mathbf{A} = \alpha\mathbf{B} + \beta\mathbf{C}$. One of the classical inequalities in Linear Algebra is the rank-sum inequality that states

$$\left|\text{rank}(\mathbf{B}) - \text{rank}(\mathbf{C})\right| \le \text{rank}(\mathbf{B} + \mathbf{C}) \le \text{rank}(\mathbf{B}) + \text{rank}(\mathbf{C})$$

Hence for $\mathbf{A} = \alpha\mathbf{B} + \beta\mathbf{C}$ we have

$$\text{rank}(\mathbf{A}) \ge \left|\text{rank}(\mathbf{B}) - \text{rank}(\mathbf{C})\right| \tag{17}$$

Now rank($\mathbf{A}$) is equal to the rank of $\mathbf{D_A}$, i.e. the diagonal matrix in its singular value decomposition, which in turn is equal to the number of non-zero entries along the diagonal. Hence, $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{D_A}) = \text{Sch}(\psi)$. Likewise, $\text{rank}(\mathbf{B}) = \text{Sch}(\phi)$ and $\text{rank}(\mathbf{C}) = \text{Sch}(\gamma)$. Therefore, from (7) we have

$$\text{Sch}(\psi) \ge \left|\text{Sch}(\phi) - \text{Sch}(\gamma)\right|$$

Problem 2.3: (Tsirelson's Inequality) Suppose $\vec{Q} = \vec{q}\cdot\vec{\sigma}$, $\vec{R} = \vec{r}\cdot\vec{\sigma}$, $\vec{S} = \vec{s}\cdot\vec{\sigma}$, $\vec{T} = \vec{t}\cdot\vec{\sigma}$ where $\vec{q},\vec{r},\vec{s},\vec{t}$ are real unit vectors in three dimensions. Show that

$$\left(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T\right)^2 = 4I + [Q,R]\otimes[S,T] \tag{18}$$

Use this result to prove that

$$\left\langle Q\otimes S\right\rangle + \left\langle R\otimes S\right\rangle + \left\langle R\otimes T\right\rangle - \left\langle Q\otimes T\right\rangle \le 2\sqrt{2} \tag{19}$$

so the violation of the Bell inequality is the maximum possible in quantum mechanics.

Solution:

In problem 2.60 we showed that the observable $Q = \vec{q}\cdot\vec{\sigma}$ has eigenvalues ±1 and the corresponding eigenspaces have projectors $P_\pm = (I \pm \vec{q}\cdot\vec{\sigma})/2$. Hence, Q can be written as $Q = \vec{q}\cdot\vec{\sigma} = (+1)P_+ + (-1)P_-$. Note that $P_\pm^2 = P_\pm$.

We see that

$$Q^2 = (\vec{q}\cdot\vec{\sigma})^2 = ((+1)P_+ + (-1)P_-)^2$$

$$\Rightarrow Q^2 = P_+P_+ - P_+P_- - P_-P_+ + P_-P_- = P_+ - P_+P_- - P_-P_+ + P_- \underbrace{=}_{\substack{\text{Completeness}\\\text{Relation}}} I - P_+P_- - P_-P_+$$

$$\Rightarrow Q^2 = I - \frac{1}{2}(I+\vec{q}\cdot\vec{\sigma})\frac{1}{2}(I-\vec{q}\cdot\vec{\sigma}) - \frac{1}{2}(I-\vec{q}\cdot\vec{\sigma})\frac{1}{2}(I+\vec{q}\cdot\vec{\sigma}) = I - \frac{1}{4}(I-(\vec{q}\cdot\vec{\sigma})^2) - \frac{1}{4}(I-(\vec{q}\cdot\vec{\sigma})^2)$$

$$\Rightarrow Q^2 = I - \frac{1}{2}(I-(\vec{q}\cdot\vec{\sigma})^2) = I - \frac{1}{2}(I-Q^2) = \frac{1}{2}I + \frac{1}{2}Q^2$$

$$\Rightarrow \frac{1}{2}Q^2 = \frac{1}{2}I \Rightarrow Q^2 = I$$

This can be shown for R,S and T. Hence, $Q^2 = R^2 = S^2 = T^2 = I$.

Next, we show that $\left(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T\right)^2 = 4I + [Q,R]\otimes[S,T]$.

$$\left(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T\right)^2 = \left(Q\otimes(S-T) + R\otimes(S+T)\right)^2$$

$$= \left(Q\otimes(S-T)\right)\left(Q\otimes(S-T)\right) + \left(R\otimes(S+T)\right)\left(R\otimes(S+T)\right) + \left(Q\otimes(S-T)\right)\left(R\otimes(S+T)\right) + \left(R\otimes(S+T)\right)\left(Q\otimes(S-T)\right)$$

$$= Q^2\otimes(S-T)^2 + R^2\otimes(S+T)^2 + QR\otimes(S-T)(S+T) + RQ\otimes(S+T)(S-T)$$

$$= I\otimes(S^2+T^2-ST-TS) + I\otimes(S^2+T^2+ST+TS) + QR\otimes((S^2+ST-TS-T^2) + RQ\otimes(S^2-ST+TS-T^2)$$

$$= I\otimes(2I-ST-TS) + I\otimes(2I+ST+TS) + QR\otimes(I+ST-TS-I) + RQ\otimes(I-ST+TS-I)$$

$$= 2I\otimes I + 2I\otimes I - I\otimes ST - I\otimes TS + I\otimes ST + I\otimes TS + QR\otimes(ST-TS) + RQ\otimes(TS-ST)$$

$$= 4I\otimes I + QR\otimes(ST-TS) + RQ\otimes(TS-ST) = 4I + QR\otimes(ST-TS) - RQ\otimes(ST-TS)$$

$$\Rightarrow \left(Q\otimes S + R\otimes S + R\otimes T - Q\otimes T\right)^2 = 4I + \left(QR-RQ\right)\otimes\left(ST-TS\right) = 4I + [Q,R]\otimes[S,T]$$

The sup-norm of a tensor product satisfies $\|[Q,R]\otimes[S,T]\|_\infty \le \|[Q,R]\|_\infty\|[S,T]\|_\infty$. Next, $\|[Q,R]\|_\infty = \|QR-RQ\|_\infty \le \|QR\|_\infty + \|RQ\|_\infty \le \|Q\|_\infty\|R\|_\infty + \|R\|_\infty\|Q\|_\infty \le 2\|Q\|_\infty\|R\|_\infty$.

Next, we show that $\|Q\|_\infty \le 1$.

$$\|Q\|_\infty = \max_{\|\psi\|=1}\langle\psi|Q|\psi\rangle = \max_{\|\psi\|=1}\mathrm{tr}\left(Q|\psi\rangle\langle\psi|\right) = \max_{\|\psi\|=1}\mathrm{tr}\left(Q\rho\right) = \max\ \mathrm{tr}\left(\vec{q}\cdot\vec{\sigma}\frac{1}{2}(I+\vec{n}\cdot\vec{\sigma})\right)$$

$\mathrm{tr}\left(\vec{q}\cdot\vec{\sigma}\frac{1}{2}(I+\vec{n}\cdot\vec{\sigma})\right) = \frac{1}{2}\mathrm{tr}\left(\vec{q}\cdot\vec{\sigma}+(\vec{q}\cdot\vec{\sigma})(\vec{n}\cdot\vec{\sigma})\right) = \frac{1}{2}\left(\mathrm{tr}(\vec{q}\cdot\vec{\sigma})+\mathrm{tr}(\vec{q}\cdot\vec{\sigma})(\vec{n}\cdot\vec{\sigma})\right) \underbrace{=}_{\mathrm{tr}(\sigma_i=0)} \frac{1}{2}\mathrm{tr}(\vec{q}\cdot\vec{\sigma})(\vec{n}\cdot\vec{\sigma}) = \frac{1}{2}\mathrm{tr}\sum_{ij}q_in_j\sigma_i\sigma_j = \frac{1}{2}\sum_{ij}q_in_j\mathrm{tr}(\sigma_i\sigma_j)$. We know that $\mathrm{tr}(\sigma_i\sigma_j) = 2\delta_{ij}$. So, $\frac{1}{2}\sum_{ij}q_in_j\mathrm{tr}(\sigma_i\sigma_j) = \frac{1}{2}\sum_{ij}q_in_j2\delta_{ij} = \sum_i q_in_i$.
Next, $\sum_{ij}q_in_i = \langle\vec{q}|\vec{n}\rangle \le |\langle\vec{q}|\vec{n}\rangle| \underbrace{\le}_{C-S} \|\vec{q}\|\|\vec{n}\| = 1\cdot 1 = 1$. Therefore, $\mathrm{tr}(Q\rho) \le 1$ for all $|\psi\rangle$ satisfying $\|\psi\| = 1$. Hence, $\max_{\|\psi\|=1}\mathrm{tr}(Q\rho) \le 1$. And so $\|Q\|_\infty \le 1$.
Similarly, we can show that $\|R\|_\infty \le 1$, $\|S\|_\infty \le 1$, $\|T\|_\infty \le 1$.
Previously we had shown $\|[Q,R]\|_\infty \le 2\|Q\|_\infty\|R\|_\infty$. Hence, $\|[Q,R]\|_\infty \le 2\cdot1\cdot1 = 2$.
Also, $\|[S,T]\|_\infty \le 2$.
So, $\|[Q,R]\otimes[S,T]\|_\infty \le \|[Q,R]\|_\infty\|[S,T]\|_\infty \le 2\cdot 2 = 4$. Putting these results together we have

$$\left\|\left(Q\otimes\left(S-T\right) + R\otimes\left(S+T\right)\right)^2\right\|_\infty = \|4I + [Q,R]\otimes[S,T]\|_\infty \le 4 + \|[Q,R]\|_\infty\|[S,T]\|_\infty \le 4+4 = 8$$

Hence,

$$\left\| Q \otimes (S - T) + R \otimes (S + T) \right\|_\infty \le \sqrt{8} = 2\sqrt{2}$$
$$\Rightarrow \left\langle Q \otimes (S - T) + R \otimes (S + T) \right\rangle \le \left\| Q \otimes (S - T) + R \otimes (S + T) \right\|_\infty \le \sqrt{8} = 2\sqrt{2}$$
$$\text{i.e. } \left\langle Q \otimes S \right\rangle + \left\langle R \otimes S \right\rangle + \left\langle R \otimes T \right\rangle - \left\langle Q \otimes T \right\rangle \le 2\sqrt{2}$$

Quantum Computation and Quantum Information

-Michael A. Nielsen and Isaac L. Chuang

Chapter 2. Introduction to Quantum Mechanics

Selected Problems

Set-II

Exercises: 2.20, 2.21, 2.22, 2.23, 2.24, 2.35, 2.39, 2.54, 2.55, 2.56, 2.57

2.20) (Basis changes) Suppose A' and A'' are matrix representations of an operator A on a vector space V with respect to two different orthonormal bases, $|v_i\rangle$ and $|w_i\rangle$. Then the elements of A' and A'' are $A'_{ij} = \langle v_i|A\left|v_j\right\rangle$ and $A''_{ij} = \langle w_i|A''\left|w_j\right\rangle$. Characterize the relationship between A' and A''.

It is given that the vectors $|v_i\rangle$ and $|w_i\rangle$ are each a set of orthonormal bases on the vector space V. We define the matrix $U = \sum_i |v_i\rangle\langle w_i|$. We can check that the matrix U is a unitary matrix. $U^\dagger U = \sum_{ij} |w_i\rangle\left\langle v_i|v_j\right\rangle\left\langle w_j\right| = \sum_{ij} |w_i\rangle\delta_{ij}\left\langle w_j\right| = \sum_i |w_i\rangle\langle w_i|$. Each $|w_i\rangle\langle w_i|$ is a projection operator and hence by the completeness relation $\sum_i |w_i\rangle\langle w_i| = I$. Therefore, $U^\dagger U = I$. Similarly, we can check that $UU^\dagger = I$.
Next, with respect to the orthonormal bases $|v_i\rangle$, each entry for the matrix A' is

$$A'_{ij} = \langle v_i|A\left|v_j\right\rangle = \langle v_i|U^\dagger U A U^\dagger U\left|v_j\right\rangle = \sum_{\alpha\beta\gamma\theta} \langle v_i|w_\alpha\rangle\left\langle v_\alpha|v_\beta\right\rangle\left\langle w_\beta\right|A\left|w_\gamma\right\rangle\left\langle v_\gamma|v_\theta\right\rangle\left\langle w_\theta|v_j\right\rangle$$

$$\Rightarrow A'_{ij} = \sum_{\alpha\beta\gamma\theta} \langle v_i|w_\alpha\rangle\delta_{\alpha\beta}\left\langle w_\beta\right|A\left|w_\gamma\right\rangle\delta_{\gamma\theta}\left\langle w_\theta|v_j\right\rangle = \sum_{\alpha\gamma} \langle v_i|w_\alpha\rangle\underbrace{\langle w_\alpha|A\left|w_\gamma\right\rangle}_{=A''_{\alpha\gamma}}\left\langle w_\gamma|v_j\right\rangle$$

Therefore, we can characterize the relationship between each entry of the matrices A' and A'' as:

$$A'_{ij} = \sum_{\alpha\gamma} A''_{\alpha\gamma}\langle v_i|w_\alpha\rangle\left\langle w_\gamma|v_j\right\rangle \quad \text{for all } i,j$$

where A'_{ij} and $A''_{\alpha\gamma}$ are the entries of the matrices A' and A'' respectively.

2.21) Repeat the proof of spectral decomposition in Box 2.2 for the case when M is Hermitian, simplifying the proof wherever possible.

When M is Hermitian the spectral decomposition theorem reads: *Any Hermitian operator M on a vector space V is diagonal with respect to some orthonormal basis for V.*
$\cdots\cdots\cdots$ (Incomplete).

2.22) Prove that two eigenvectors of a Hermitian operator with different eigenvalues are necessarily orthogonal.

Let $|v_1\rangle$ be the eigenvector with respect to the eigenvalue λ_1 and $|v_2\rangle$ be the eigenvector with respect to the eigenvalue λ_2. And $\lambda_1 \neq \lambda_2$. Then we have $H|v_1\rangle = \lambda_1|v_1\rangle$ and $H|v_2\rangle = \lambda_2|v_2\rangle$. Taking the adjoint of $H|v_1\rangle = \lambda_1|v_1\rangle$ we get $\langle v_1|H^\dagger = \lambda_1\langle v_1|$. Taking the inner product with $|v_2\rangle$ gives $\langle v_1|H^\dagger|v_2\rangle = \lambda_1\langle v_1|v_2\rangle$. Next $H^\dagger = H$, so $\langle v_1|H^\dagger|v_2\rangle = \langle v_1|H|v_2\rangle$ which is equal to $\lambda_2\langle v_1|v_2\rangle$. So we have $\lambda_1\langle v_1|v_2\rangle = \lambda_2\langle v_1|v_2\rangle$ which implies that $\langle v_1|v_2\rangle$ must necessarily be equal to 0 (and hence $|v_1\rangle$ and $|v_2\rangle$ are orthogonal) since $\lambda_1 \neq \lambda_2$.

2.23) Show that the eigenvalues of the projector P are all either 0 or 1.

Let P be a projector and $|v\rangle$ be an eigenvector with respect to the eigenvalue λ. Then we have $P^2|v\rangle = P|v\rangle$ which implies $P(P|v\rangle) = \lambda|v\rangle \Rightarrow P(\lambda|v\rangle) = \lambda|v\rangle \Rightarrow \lambda^2|v\rangle = \lambda|v\rangle$. Hence $\left(\lambda^2-\lambda\right)|v\rangle = 0$. So, $\lambda(\lambda-1)=0$ which means λ is either 0 or 1. This is true for all such λ's. Hence, the eigenvalues of the projector P are all either 0 or 1.

2.24 (Hermiticity of positive operators) Show that a positive operator is necessarily Hermitian. (*Hint*: Show that an arbitrary operator A can be written $A = B + iC$ where B and C are Hermitian).

The book defines a positive operator A as an operator such that for any vector $|v\rangle$, $\langle v|A|v\rangle$ is real and non-negative. Next, we define $A = \frac{A+A^\dagger}{2} + i\frac{A-A^\dagger}{2i}$. So, $A = B + iC$ where $B = \frac{A+A^\dagger}{2}$ and $C = \frac{A-A^\dagger}{2i}$. We can check that B and C are both Hermitian. $B^\dagger = \frac{1}{2}\left(A+A^\dagger\right)^\dagger = \frac{1}{2}\left(A^\dagger + A\right) = B$. And $C^\dagger = -\frac{1}{2i}\left(A-A^\dagger\right)^\dagger = -\frac{1}{2i}\left(A^\dagger - A\right) = \frac{1}{2i}\left(A-A^\dagger\right) = C$.
Next, we see that $\langle v|A|v\rangle = \langle v|B|v\rangle + i\langle v|C|v\rangle$. But it is given that A is a positive operator which means $\langle v|A|v\rangle \geq 0$ and real. This implies that $\langle v|C|v\rangle$ must vanish for all $|v\rangle$. And hence $\langle v|A|v\rangle = \langle v|B|v\rangle$ for all $|v\rangle$. So, $A = B = \frac{A+A^\dagger}{2} \Rightarrow 2A = A + A^\dagger \Rightarrow A = A^\dagger$ and hence A is Hermitian.

2.35) (Exponential of the Pauli matrices) Let $\vec{v}$ be any real, three dimensional unit vector and θ be a real number. Prove that $\exp\{(i\theta\vec{v}\cdot\vec{\sigma})\} = \cos(\theta)I + i\sin(\theta)(\vec{v}\cdot\vec{\sigma})$ where $\vec{v}\cdot\vec{\sigma} = \sum_{i=1}^{3} v_i\sigma_i$.

In the process of solving Problem 2.60 (in Set-I) we found that for a unit vector in $\vec{v} \in \mathbb{R}^3$, $\vec{v}\cdot\vec{\sigma}$ has eigenvalues 1 and -1 and we can write $\vec{v}\cdot\vec{\sigma}$ in the form $\vec{v}\cdot\vec{\sigma} = |v_+\rangle\langle v_+| - |v_-\rangle\langle v_-|$ where $|v_+\rangle$ and $|v_-\rangle$ are the eigenvectors corresponding to the eigenvalues 1 and -1 respectively. Hence we have

$$i\theta\vec{v}\cdot\vec{\sigma} = i\theta(|v_+\rangle\langle v_+| - |v_-\rangle\langle v_-|) \Rightarrow \exp(i\theta\vec{v}\cdot\vec{\sigma}) = e^{i\theta}|v_+\rangle\langle v_+| + e^{-i\theta}|v_-\rangle\langle v_-|$$

Also, $|v_+\rangle\langle v_+|$ and $|v_-\rangle\langle v_-|$ are projection operators that satisfy the completeness relation $|v_+\rangle\langle v_+| + |v_-\rangle\langle v_-| = I$. Solving the two equations $\vec{v}\cdot\vec{\sigma} = |v_+\rangle\langle v_+| - |v_-\rangle\langle v_-|$ and $|v_+\rangle\langle v_+| + |v_-\rangle\langle v_-| = I$ we get $|v_\pm\rangle\langle v_\pm| = \frac{1}{2}(I \pm \vec{v}\cdot\vec{\sigma})$. Hence we have

$$\exp(i\theta\vec{v}\cdot\vec{\sigma}) = e^{i\theta}|v_+\rangle\langle v_+| + e^{-i\theta}|v_-\rangle\langle v_-|$$
$$\Rightarrow \exp(i\theta\vec{v}\cdot\vec{\sigma}) = e^{i\theta}\left(\frac{I+\vec{v}\cdot\vec{\sigma}}{2}\right) + e^{-i\theta}\left(\frac{I-\vec{v}\cdot\vec{\sigma}}{2}\right) = \frac{e^{i\theta}+e^{-i\theta}}{2}I + \frac{e^{i\theta}-e^{-i\theta}}{2}\vec{v}\cdot\vec{\sigma} = \cos(\theta)I + i\sin(\theta)\vec{v}\cdot\vec{\sigma}$$

where $\cos(\theta) = \frac{e^{i\theta}+e^{-i\theta}}{2}$ and $\sin(\theta) = \frac{e^{i\theta}-e^{-i\theta}}{2i}$.

2.39) (Hilbert-Schmidt inner product on operators) The set L_V of linear operators on a Hilbert space V is obviously a vector space – the sum of two linear operators is a linear operator, zA is a linear operator if A is a linear operator and z is a complex number, and there is a zero element 0. An important additional result is that the vector space L_V can be given a natural inner product structure, turning it into a Hilbert space.
(1) Show that the function $(\cdot,\cdot)$ on $L_V \times L_V$ defined by $(A,B) \equiv \text{tr}(A^\dagger B)$ is an inner product function. This inner product is known as the Hilbert-Schmidt inner product or trace inner product.
(2) If V has d dimensions show that L_V has dimension d^2.
(3) Find an orthonormal basis of Hermitian matrices for the Hilbert space L_V.

(1) In order to show that the function $(\cdot,\cdot)$ on $L_V \times L_V$ defined by $(A,B) \equiv \text{tr}(A^\dagger B)$ is an inner product function we need to show a) Symmetry, b) Linearity and c) Positive-Definite.
a) **Symmetry**:
We have $A,B \in L_V$. Let $A = \sum_{ij} a_{ij} |i\rangle\langle j|$ and $B = \sum_{ij} b_{ij} |i\rangle\langle j|$ be the expansion of A and B in the standard basis. We show that $(A,B) = \overline{(B,A)}$. Using the definition of $(\cdot,\cdot)$ we have

$$(A,B) = \text{tr}(A^\dagger B) = \text{tr}\left(\sum_{ij} a^\star_{ij} |j\rangle\langle i| \sum_{ij} b_{ij} |i\rangle\langle j|\right) = \text{tr}\left(\sum_{ijkl} a^\star_{ij} b_{kl} |j\rangle \underbrace{\langle i|k\rangle}_{=\delta_{ik}} \langle l|\right) = \text{tr}\left(\sum_{ijl} a^\star_{ij} b_{il} |j\rangle\langle l|\right) = \sum_{ij} a^\star_{ij} b_{ij} \tag{20}$$

$$(B,A) = \text{tr}(B^\dagger A) = \text{tr}\left(\sum_{ij} b^\star_{ij} |j\rangle\langle i| \sum_{ij} a_{ij} |i\rangle\langle j|\right) = \text{tr}\left(\sum_{ijkl} b^\star_{ij} a_{kl} |j\rangle \underbrace{\langle i|k\rangle}_{=\delta_{ik}} \langle l|\right) = \text{tr}\left(\sum_{ijl} b^\star_{ij} a_{il} |j\rangle\langle l|\right) = \sum_{ij} b^\star_{ij} a_{ij} \tag{21}$$

$$\Rightarrow (B,A) = \sum_{ij} b^\star_{ij} a_{ij} = \left(\sum_{ij} a^\star_{ij} b_{ij}\right)^\star = \overline{(A,B)} \tag{22}$$

b) **Linearity**:
For some $\alpha \in \mathbb{C}$ we have

$$(\alpha A,B) = \text{tr}\left((\alpha A)^\dagger B\right) = \text{tr}\left(\alpha^\star A^\dagger B\right) = \alpha^\star \text{tr}\left(A^\dagger B\right) = \alpha^\star (A,B) \tag{23}$$

$$(A,\alpha B) = \text{tr}\left(A^\dagger (\alpha B)\right) = \text{tr}\left(\alpha A^\dagger B\right) = \alpha \text{tr}\left(A^\dagger B\right) = \alpha (A,B) \tag{24}$$

Next, for some $A,B,C \in L_V$ we have

$$(A+C,B) = \text{tr}\left((A+C)^\dagger B\right) = \text{tr}\left((A^\dagger + C^\dagger)B\right) = \text{tr}\left(A^\dagger B + C^\dagger B\right) = \text{tr}\left(A^\dagger B\right) + \text{tr}\left(C^\dagger B\right) = (A,B) + (C,B) \tag{25}$$

c) **Positive-Definite**:

$$(A,A) = \mathrm{tr}\left(A^\dagger A\right) = \mathrm{tr}\left(\sum_{ij} a^\star_{ij} |j\rangle\langle i| \sum_{ij} a_{ij} |i\rangle\langle j|\right) = \mathrm{tr}\left(\sum_{ijkl} a^\star_{ij} a_{kl} |j\rangle \underbrace{\langle i|k\rangle}_{=\delta_{ik}} \langle l|\right) = \mathrm{tr}\left(\sum_{ijl} a^\star_{ij} a_{il} |j\rangle\langle l|\right) = \sum_{ij} a^\star_{ij} a_{ij}$$

Hence $(A,A) = \sum_{ij} a^\star_{ij} a_{ij} = \sum_{ij} |a_{ij}|^2 \geq 0$ with equality iff $a_{ij} = 0$ i.e. $A = 0$

(2) Suppose $\{|1\rangle, |2\rangle, \ldots, |d\rangle\}$ is the set of orthonormal basis for the vector space V. Then the dimension of V is the cardinality of the basis of V, i.e. d. L_V is the set of linear operators T on V such that $T : V \to V$. We want to show that the cardinality of the basis of L_V is d^2. By the completeness relation we have $I_V = \sum_{i=1}^d |i\rangle\langle i|$. One application of the completeness relation is to give a means for representing an operator in the outer product notation. For $T : V \to V$ and $\{|i\rangle\}_{i=1}^d$, an orthonormal basis for V, using the completeness relation twice gives

$$T = I_V T I_V = \sum_{i=1}^d |i\rangle\langle i| T \sum_{i=1}^d |i\rangle\langle i| = \sum_{i=1}^d \sum_{j=1}^d |i\rangle\langle i| T |j\rangle\langle j| = \sum_{i=1}^d \sum_{j=1}^d \langle i| T |j\rangle |i\rangle\langle j| \tag{26}$$

which is the outer product representation for T. T has matrix element $\langle i|T|j\rangle$ in the ith row and the jth column, with respect to the input basis $|j\rangle$ and the output basis $|i\rangle$ (where $i,j = 1,2,\ldots,d$). In (26) we have shown that the elements $|i\rangle\langle j|$ span the space of L_V since the matrix representation of any linear operator $T : V \to V$ can be obtained by the linear combination of $|i\rangle\langle j|$. In order to show that the elements $|i\rangle\langle j|$ are indeed basis elements of L_V we need to also show that they are mutually linearly independent. From the matrix representation of any linear operator $T : V \to V$ as obtained in (26) we see that $\langle i|T|j\rangle$ are the weights (real/complex) for $|i\rangle\langle j|$ for all $i,j = 1,\ldots,d$. And $\sum_{i=1}^d \sum_{j=1}^d \langle i|T|j\rangle |i\rangle\langle j| = 0$ if and only if $\langle i|T|j\rangle = 0$ for all $i,j = 1,\ldots,d$. This shows that $\{|i\rangle\langle j|\}_{i,j=1}^d$ are linearly independent.
So, (26) shows that there are d^2 such elements in the representation of a linear operator T. Hence the cardinality of the basis $\{|i\rangle\langle j|\}_{i,j=1}^d$ of L_V is d^2, i.e. $\dim(L_V)$ is d^2.

(3) Assuming a dimension of 2 the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$ form an orthonormal basis of Hermitian matrices for the Hilbert space L_V. For higher dimensions we need some generalization of the Pauli matrices. Suppose we have the standard basis $\{|j\rangle\}_{j=1}^n$ (and $n > 2$) in the Hilbert space V. We define the following three types of matrices:

$$\text{(1) Symmetric}\left(\frac{d(d-1)}{2} \text{ matrices}\right) : \Lambda_s^{jk} = |j\rangle\langle k| + |k\rangle\langle j|, \quad \text{for } 1 \leq j < k \leq n \tag{27}$$

$$\text{(2) Asymmetric}\left(\frac{d(d-1)}{2} \text{ matrices}\right) : \Lambda_a^{jk} = -i\,|j\rangle\langle k| + i\,|k\rangle\langle j|, \quad \text{for } 1 \leq j < k \leq n \tag{28}$$

$$\text{(3) Diagonal}(d-1 \text{ matrices}) : \Lambda_d^l = \sqrt{\frac{2}{l(l+1)}}\left(\sum_{j=1}^l |j\rangle\langle j| - l\,|l+1\rangle\langle l+1|\right), \quad \text{for } 1 \leq l \leq n-1 \tag{29}$$

The $\frac{n(n-1)}{2} + \frac{n(n-1)}{2} + n - 1 = n^2 - 1$ matrices in (27), (28) and (29) are the so-called *generalized Gell-Mann matrices*. They are a generalization of the Pauli matrices to dimensions greater than 2. By definition the matrices $\{\{\Lambda_s^{jk}\}, \{\Lambda_a^{jk}\}, \{\Lambda_d^l\}\}$ are Hermitian. Also, each of them have a trace equal to zero. They are orthogonal in the Hilbert-Schmidt sense, i.e. $\mathrm{tr}\left(\Lambda_i^\dagger \Lambda_j\right) = 2\delta_{ij}$.

For $n=2$ and $n=3$ the collection of matrices defined above recover the Pauli and Gell-Mann matrices respectively.

2.54) Suppose A and B are commuting Hermitian operators. Prove that $\exp(A)\exp(B) = \exp(A+B)$. (*Hint*: Use the results of section 2.1.9).

Given that A and B are commuting Hermitian operators, i.e. $A^\dagger = A$, $B^\dagger = B$ and $[A,B] = AB - BA = 0$ we can use the *simultaneous diagonalization theorem* given in section 2.1.9 which says: *If A and B are Hermitian. Then* $[A,B]=0$ *if and only if there exists an orthonormal basis such that both A and B are diagonal with respect to that basis. We say that A and B are simultaneously diagonalizabe in that case.*
Since $[A,B] = 0$, A and B can be simultaneously diagonalized as $A = \sum_i \lambda_i |i\rangle\langle i|$ and $B = \sum_i \mu_i |i\rangle\langle i|$ where λ_i and μ_i are reals since A and B are Hermitian. Hence

$$\exp(A)\exp(B) = \left(\sum_i \exp(\lambda_i)|i\rangle\langle i|\right)\left(\sum_i \exp(\mu_i)|i\rangle\langle i|\right) = \sum_{ij} \exp(\lambda_i)\exp(\mu_j)|i\rangle \underbrace{\langle i|j\rangle}_{=\delta_{ij}} \langle j|$$

$$\Rightarrow \exp(A)\exp(B) = \sum_i \exp(\lambda_i)\exp(\mu_i)|i\rangle\langle i| = \sum_i \exp(\lambda_i+\mu_i)|i\rangle\langle i| = \exp\left(\sum_i (\lambda_i+\mu_i)|i\rangle\langle i|\right)$$

$$\Rightarrow \exp(A)\exp(B) = \exp\left(\sum_i \lambda_i|i\rangle\langle i| + \sum_i \mu_i |i\rangle\langle i|\right) = \exp(A+B)$$

2.55) Prove that $U(t_1,t_2)$ defined in Equation (2.91) is unitary.

$U(t_1,t_2)$ is defined in Equation (2.91) as

$$U(t_1,t_2) = \exp\left(\frac{-iH(t_2-t_1)}{\hbar}\right)$$

Using the spectral decomposition of the Hamiltonian $H = \sum_E E|E\rangle\langle E|$ and $\sum_E |E\rangle\langle E| = I$

$$U^\dagger(t_1,t_2)U(t_1,t_2) = \exp\left(\frac{iH(t_2-t_1)}{\hbar}\right)\exp\left(\frac{-iH(t_2-t_1)}{\hbar}\right)$$
$$= \exp\left(\frac{i(t_2-t_1)}{\hbar}\sum_E E|E\rangle\langle E|\right)\exp\left(-\frac{i(t_2-t_1)}{\hbar}\sum_E E|E\rangle\langle E|\right)$$
$$\Rightarrow U^\dagger(t_1,t_2)U(t_1,t_2) = \left(\sum_E \exp\left(\frac{i(t_2-t_1)}{\hbar}E\right)|E\rangle\langle E|\right)\left(\sum_E \exp\left(-\frac{i(t_2-t_1)}{\hbar}E\right)|E\rangle\langle E|\right)$$
$$\Rightarrow U^\dagger(t_1,t_2)U(t_1,t_2) = \sum_{E,E'} \exp\left(\frac{i(t_2-t_1)}{\hbar}E\right)\exp\left(-\frac{i(t_2-t_1)}{\hbar}E'\right)|E\rangle\underbrace{\langle E|E'\rangle}_{=\delta_{E,E'}}\langle E'|$$
$$\Rightarrow U^\dagger(t_1,t_2)U(t_1,t_2) = \sum_{E,E'} \exp\left(\frac{i(t_2-t_1)}{\hbar}E - \frac{i(t_2-t_1)}{\hbar}E'\right)|E\rangle\underbrace{\langle E|E'\rangle}_{=\delta_{E,E'}}\langle E'| = \sum_E \exp\left(\frac{i(t_2-t_1)}{\hbar}E - \frac{i(t_2-t_1)}{\hbar}E\right)|E\rangle\langle E|$$
$$\Rightarrow U^\dagger(t_1,t_2)U(t_1,t_2) = \sum_E \exp(0)|E\rangle\langle E| = \sum_E |E\rangle\langle E| = I$$

With a similar approach we can show that $U(t_1,t_2)U^\dagger(t_1,t_2) = I$. Hence $U^\dagger(t_1,t_2)U(t_1,t_2) = U(t_1,t_2)U^\dagger(t_1,t_2) = I$. Hence $U(t_1,t_2)$ is unitary.
<u>Note</u>: In the expression $H = \sum_E E|E\rangle\langle E|$ of the Hamiltonian (a Hermitian operator), $|E\rangle$ is the

energy eigenstate or sometimes referred to as *stationary state* and E is the *energy* of the state $|E\rangle$. The lowest energy is known as the *ground state energy* for the system and the corresponding energy eigenstate (or eigenspace) is known as the *ground state*.

2.56) Use the spectral decomposition to show that $K \equiv -i\log(U)$ is Hermitian for any unitary U, and thus $U = \exp(iK)$ for some Hermitian K.

The unitary operator U satisfies $U^\dagger U = UU^\dagger = I$. Hence, U is normal and by the spectral decomposition theorem we have $U = \sum_j \lambda_j |j\rangle\langle j|$, where λ_j is the eigenvalue corresponding to the eigenvector $|j\rangle$. So, $\log(U) = \sum_j \log(\lambda_j)|j\rangle\langle j|$. Also, $|\lambda_j| = 1$. Let $\lambda_j = e^{i\theta_j}$ for $\theta_j \in \mathbb{R}$. Then $\log(U) = \sum_j \log(\lambda_j)|j\rangle\langle j| = \sum_j \log(e^{i\theta_j})|j\rangle\langle j| = \sum_j i\theta_j|j\rangle\langle j|$. Hence, $K = -i\log(U) = -i\sum_j i\theta_j|j\rangle\langle j| = \sum_j \theta_j|j\rangle\langle j|$. We can check that K is Hermitian. $K^\dagger = \left(\sum_j \theta_j|j\rangle\langle j|\right)^\dagger = \sum_j \theta_j^\star|j\rangle\langle j| = \sum_j \theta_j|j\rangle\langle j| = K$ since $\theta_j \in \mathbb{R}$. Hence, $U = \exp(iK)$ for some Hermitian K. This shows that the exponential of i times a Hermitian operator is unitary. And so there is a one-to-one correspondence between the discrete-time description of dynamics using unitary operators, and the continuous time description using Hamiltonians.

2.57) (Cascaded measurements are single measurements) Suppose $\{L_l\}$ and $\{M_m\}$ are two sets of measurement operators. Show that a measurement defined by the measurement operators $\{L_l\}$ followed by a measurement defined by the measurement operators $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{lm}\}$ with the representation $N_{lm} \equiv M_m L_l$.

Suppose the state of the system immediately before the measurement is $|\psi\rangle$. Then the probability of obtaining the result l using the operator L_l is $p(l) = \langle\psi|L_l^\dagger L_l|\psi\rangle$ and the state of the system after the measurement is $|\phi\rangle = \frac{L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}}$. Next, we use the operator M_m on the state $|\phi\rangle$. The state of the system after the measurement is

$$\frac{M_m|\phi\rangle}{\sqrt{\langle\phi|M_m^\dagger M_m|\phi\rangle}} = \frac{\frac{M_m L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}}}{\sqrt{\frac{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle}}} = \frac{M_m L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}}\sqrt{\frac{\langle\psi|L_l^\dagger L_l|\psi\rangle}{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}} = \frac{M_m L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}}$$

$$= \frac{M_m L_l|\psi\rangle}{\sqrt{\langle\psi|(M_m L_l)^\dagger(M_m L_l)|\psi\rangle}} = \frac{N_{lm}|\psi\rangle}{\langle\psi|N_{lm}^\dagger N_{lm}|\psi\rangle}$$

Hence the measurement using the operator L_l followed by the measurement using the operator M_m is physically equivalent to a single measurement using the operator $N_{lm} \equiv M_m L_l$. Next, we show that the probability of obtaining the outcome lm using the operator N_{lm} is the same as that of the cascaded measurement L_l followed by M_m.

The probability of obtaining lm using N_{lm} is $p(lm) = \langle\psi|N_{lm}^\dagger N_{lm}|\psi\rangle$. In the cascaded measurement we first use the operator L_l on the initial state $|\psi\rangle$ and obtain the outcome l with probability $p(l) = \langle\psi|L_l^\dagger L_l|\psi\rangle$ and obtain the post-measurement state $|\phi\rangle = \frac{L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}}$.

Next, we use the operator M_m on the state $|\phi\rangle$ and obtain the outcome m with probability

$$p(m) = \langle\phi| M_m^\dagger M_m |\phi\rangle = \frac{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}{\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}\sqrt{\langle\psi|L_l^\dagger L_l|\psi\rangle}} = \frac{\langle\psi|L_l^\dagger M_m^\dagger M_m L_l|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle} = \frac{\langle\psi|N_{lm}^\dagger N_{lm}|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle}.$$

Hence, the probability of observing the outcome lm in the cascaded scenario is $p(l)p(m) = \langle\psi|L_l^\dagger L_l|\psi\rangle \times \frac{\langle\psi|N_{lm}^\dagger N_{lm}|\psi\rangle}{\langle\psi|L_l^\dagger L_l|\psi\rangle} = \langle\psi|N_{lm}^\dagger N_{lm}|\psi\rangle$ which is equal to the probability $p(lm)$ of obtaining the outcome lm using the operator N_{lm}.

2.76) Extend the proof of Schmidt decomposition to the case where A and B may have state spaces of different dimensionality.

We begin by stating the Schmidt decomposition theorem.
Suppose H_A and H_B are the Hilbert spaces of dimensions d_A and d_B respectively (with $d_A \geq d_B$). For any state $|\psi\rangle$ in the state space of $H_A \otimes H_B$ there exist orthonormal sets $|i_A\rangle$ and $|i_B\rangle$ such that $|\psi\rangle = \sum_{i=1} \lambda_i |i_A\rangle |i_B\rangle$ where λ_i is real, non-negative and satisfies $\sum_i \lambda_i^2 = 1$. Note that the orthonormal sets $|i_A\rangle$ and $|i_B\rangle$ can be extended to orthonormal basis by the Gram-Schmidt procedure.

Proof: Any state in the state space of $H_A \otimes H_B$ can be viewed as $|\psi\rangle = \sum_{j=1}^{d_A} \sum_{k=1}^{d_B} c_{jk} |j\rangle |k\rangle$, where $\{|j\rangle\}_{j=1}^{d_A}$ and $\{|k\rangle\}_{k=1}^{d_B}$ are the standard basis in the state space of H_A and H_B respectively. Then c_{jk} are the elements of a $d_A \times d_B$ matrix C. By the singular value decomposition theorem of a (real/complex) matrix we can find $d_A \times d_A$ unitary U, a $d_B \times d_B$ unitary V and a $d_B \times d_B$ positive semi-definite diagonal matrix D such that $C = UDV^\dagger$. Note that since we have assumed $d_A \geq d_B$ we can write C as $C = U \begin{pmatrix} D \\ 0 \end{pmatrix} V^\dagger$, where we have appended $d_A - d_B$ rows of 0's to the diagonal matrix D.