Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.453 Quantum Optical Communication

Problem Set 3 Solutions
Fall 2016

Problem 3.1

Here we shall extend the results of Problem 2.2 to include classically-random polarizations.

(a) We have that

$$\begin{aligned}\mathbf{r}^T\mathbf{r} &= r_1^2 + r_2^2 + r_3^2 \\ &= 4[\text{Re}(\langle\alpha_x^*\alpha_y\rangle)]^2 + 4[\text{Im}(\langle\alpha_x^*\alpha_y\rangle)]^2 + (\langle|\alpha_x|^2\rangle - \langle|\alpha_y|^2\rangle)^2 \\ &= 4|\langle\alpha_x^*\alpha_y\rangle|^2 + (\langle|\alpha_x|^2\rangle - \langle|\alpha_y|^2\rangle)^2 \\ &\leq 4\langle|\alpha_x|^2|\rangle\langle|\alpha_y|^2\rangle + (\langle|\alpha_x|^2\rangle - \langle|\alpha_y|^2\rangle)^2,\end{aligned}$$

via the Schwarz inequality applied to the first term. Squaring out the last term and doing some cancellation then yields,

$$\begin{aligned}\mathbf{r}^T\mathbf{r} &\leq \langle|\alpha_x|^2\rangle^2 + 2\langle|\alpha_x|^2\rangle\langle|\alpha_y|^2\rangle + \langle|\alpha_y|^2\rangle^2 \\ &= (\langle|\alpha_x|^2\rangle + \langle|\alpha_y|^2\rangle)^2 = 1^2 = 1.\end{aligned}$$

(b) Because $\mathbf{r}_a$ is a real-valued, unit-length vector and $\mathbf{r}$ is a real-valued vector whose length is at most one, we have that

$$0 \leq \frac{1 + \mathbf{r}_a^T\mathbf{r}}{2} \leq 1$$

via the Schwarz inequality applied to the second term. This same argument applies to $(1 + \mathbf{r}_b^T\mathbf{r})/2$. Thus to prove that we have a proper probability distribution, we need only show that the probabilities sum to one. We are given that $\mathbf{r}_a$ and $\mathbf{r}_b$ are the Poincaré sphere representations of the orthogonal polarization states $\mathbf{i}_a$ and $\mathbf{i}_b$, respectively. We commented in the solution to Problem 2.2 that these Poincaré sphere representations must then satisfy $\mathbf{r}_b = -\mathbf{r}_a$, hence the proof is trivial once this condition is employed:

$$\begin{aligned}\text{Pr}(\text{polarized along } \mathbf{i}_a) + \text{Pr}(\text{polarized along } \mathbf{i}_b) &= \frac{1 + \mathbf{r}_a^T\mathbf{r}}{2} + \frac{1 + \mathbf{r}_b^T\mathbf{r}}{2} \\ &= \frac{2 + \mathbf{r}_a^T\mathbf{r} - \mathbf{r}_a^T\mathbf{r}}{2} = 1.\end{aligned}$$

Now we need only prove the $\mathbf{r}_a = -\mathbf{r}_b$ assertion.

Defining the component representations,

$$\mathbf{i}_a = \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} \quad \mathbf{i}_b = \begin{bmatrix} \beta_x \\ \beta_y \end{bmatrix},$$

we have that

$$\mathbf{i}_a^\dagger \mathbf{i}_b = \alpha_x^* \beta_x + \alpha_y^* \beta_y = 0, \tag{1}$$

because the $\mathbf{i}_a$ and $\mathbf{i}_b$ polarizations are orthogonal. Thus, without loss of generality we may say that

$$\mathbf{i}_b = \begin{bmatrix} -\alpha_y^* \\ \alpha_x^* \end{bmatrix},$$

as this choice gives a unit-length vector that satisfies the orthogonality condition. Now, by direct calculation we find that

$$\mathbf{r}_b = \begin{bmatrix} 2\text{Re}(\beta_x^* \beta_y) \\ 2\text{Im}(\beta_x^* \beta_y) \\ |\beta_x|^2 - |\beta_y|^2 \end{bmatrix} = \begin{bmatrix} -2\text{Re}(\alpha_y \alpha_x^*) \\ -2\text{Im}(\alpha_y \alpha_x^*) \\ |\alpha_y|^2 - |\alpha_x|^2 \end{bmatrix} = -\mathbf{r}_a,$$

and our proof is done.

(c) If $\mathbf{r} = 0$, then it is obvious (from the measurement probability definitions) that Pr(polarized along $\mathbf{i}_a$) = Pr(polarized along $\mathbf{i}_b$) = 1/2. Note that this is true *regardless* of what pair of orthogonal polarizations are chosen for the measurements.

When

$$\mathbf{r} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{r}_a = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{r}_b = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix},$$

i.e., a right circularly polarized photon measured in the x-y linear polarization basis, we find that Pr(polarized along $\mathbf{i}_a$) = Pr(polarized along $\mathbf{i}_b$) = 1/2. This equiprobable situation does not hold, however, for the right circularly polarized photon when we measure in other bases. In particular, for

$$\mathbf{r}_a = -\mathbf{r}_b = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

i.e., if we measure in the circularly-polarized basis, then we will obtain

$$\text{Pr(polarized along } \mathbf{i}_a) = 1 - \text{Pr(polarized along } \mathbf{i}_b) = 1.$$

Problem 3.2

Here we introduce the notion of a density operator, i.e., a way to account for classical randomness limiting our knowledge of a quantum system's state.

(a) This is standard, simple, classical probability theory. We know that the probability of observing the outcome o_n when we measure $\hat{O}$ on a quantum system in state $|\psi_m\rangle$ is $\Pr(\, o_n \mid |\psi_m\rangle\,) \equiv |\langle o_n|\psi_m\rangle|^2$. If p_m is the probability that the system is in state $|\psi_m\rangle$, for $1 \le m \le M$ with $\sum_{m=1}^M p_m = 1$, then

$$\Pr(o_n) = \sum_{m=1}^{M} p_m \Pr(\, o_n \mid |\psi_m\rangle\,) = \sum_{m=1}^{M} p_m |\langle o_n|\psi_m\rangle|^2, \quad \text{for } 1 \le n < \infty,$$

is the unconditional probability of getting this outcome.

(b) Expanding the squared magnitude that appears in the answer from (a) gives us,

$$\begin{aligned}\Pr(o_n) &= \sum_{m=1}^{M} p_m |\langle o_n|\psi_m\rangle|^2 = \sum_{m=1}^{M} p_m \langle o_n|\psi_m\rangle\langle\psi_m|o_n\rangle \\ &= \langle o_n| \left(\sum_{m=1}^{M} p_m |\psi_m\rangle\langle\psi_m|\right) |o_n\rangle \\ &= \langle o_n|\hat\rho|o_n\rangle, \quad \text{for } 1 \le n < \infty,\end{aligned}$$

QED.

(c) Again, we start with straightforward, classical probability theory:

$$\langle\hat O\rangle \equiv \sum_{n=1}^{\infty} o_n \Pr(o_n),$$

is the expected value of the outcome of the $\hat O$ measurement. Now, using the result of (b) we have that

$$\begin{aligned}\langle\hat O\rangle &= \sum_{n=1}^{\infty} o_n \langle o_n|\hat\rho|o_n\rangle \\ &= \sum_{n=1}^{\infty} \langle o_n| \left(\hat\rho \sum_{k=1}^{\infty} o_k |o_k\rangle\langle o_k|\right) |o_n\rangle = \operatorname{tr}(\hat\rho\hat O),\end{aligned}$$

where the last equality employs the diagonal representation of $\hat O$, viz.,

$$\hat O = \sum_{k=1}^{\infty} o_k |o_k\rangle\langle o_k|,$$

and the penultimate equality relies on the orthonormality of the $\hat O$ eigenkets, i.e.,

$$\langle o_n|o_m\rangle = \delta_{nm}.$$

Problem 3.3
Here we will explore the difference between a pure state and a mixed state, i.e., the difference between knowing that a quantum system is in a definite state $|\psi\rangle$ as opposed to having a classically-random distribution over a set of such states, namely a density operator $\hat{\rho}$.

(a) Suppose we measure an observable $\hat{O}$ with eigenvalues $\{\, o_n : 1 \le n < \infty \,\}$ and complete orthonormal (CON) eigenkets $\{\, |o_n\rangle : 1 \le n < \infty \,\}$. From Problem 3.2 we know that if we measure $\hat{O}$ on the quantum system, when that system has density operator $\hat{\rho}$, the probability of getting the outcome o_n is

$$\Pr(o_n) = \langle o_n|\hat{\rho}|o_n\rangle = \sum_{k=1}^{\infty} \rho_k |\langle o_n|\rho_k\rangle|^2.$$

If the eigenkets of $\hat{O}$ are identical to those of $\hat{\rho}$, i.e., $|o_n\rangle = |\rho_n\rangle$, for $1 \le n < \infty$, then the general result reduces to $\Pr(o_n) = \rho_n$, i.e., the $\{\rho_n\}$ must be a probability distribution. This is what we were asked to show.

(b) This is trivial. We can use any CON basis to evaluate a trace. So, let us choose the eigenkets of $\hat{\rho}$. We then find that,

$$\text{tr}(\hat{\rho}) = \sum_{k=1}^{\infty} \langle \rho_k|\hat{\rho}|\rho_k\rangle = \sum_{k=1}^{\infty} \rho_k \langle \rho_k|\rho_k\rangle = \sum_{k=1}^{\infty} \rho_k = 1.$$

(c) Combining the result of (a) with the setup in Problem 3.2, it should be clear that ρ_k is the probability that the quantum system is in state $|\rho_k\rangle$. If the system is in a pure state $|\psi\rangle$, i.e., there is probability one that the system is in this state, we can represent that situation in density operator form by setting $\rho_1 = 1$ and $|\rho_1\rangle = |\psi\rangle$. This leads to a projector-valued density operator, $\hat{\rho} = |\rho_1\rangle\langle\rho_1| = |\psi\rangle\langle\psi|$. It is now easy to verify that,

$$\hat{\rho}^2 = |\psi\rangle \left(\langle\psi|\psi\rangle\right) \langle\psi| = |\psi\rangle\langle\psi| = \hat{\rho}.$$

Thus, $\text{tr}(\hat{\rho}^2) = \text{tr}(\hat{\rho}) = 1$.

(d) Using the diagonal representation of the density operator, we find that

$$\begin{aligned}
\hat{\rho}^2 &= \left(\sum_{n=1}^{\infty} \rho_n |\rho_n\rangle\langle\rho_n|\right)\left(\sum_{k=1}^{\infty} \rho_k |\rho_k\rangle\langle\rho_k|\right) \\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \rho_n\rho_k |\rho_n\rangle \left(\langle\rho_n|\rho_k\rangle\right) \langle\rho_k| \\
&= \sum_{n=1}^{\infty} \rho_n^2 |\rho_n\rangle\langle\rho_n|,
\end{aligned}$$

via the orthonormality of the density operator's eigenkets. Taking the trace in the $\{|\rho_k\rangle\}$ basis now gives,

$$\mathrm{tr}(\hat{\rho}^2) = \sum_{k=1}^{\infty} \rho_k^2,$$

cf. the derivation in (b). Finally, because $0 \leq \rho_k \leq 1$ implies that $\rho_k^2 \leq \rho_k$ for $1 \leq k < \infty$, we get

$$\mathrm{tr}(\hat{\rho}^2) \leq \sum_{k=1}^{\infty} \rho_k = \mathrm{tr}(\hat{\rho}) = 1.$$

Equality only occurs here if and only if $\rho_k = \delta_{kn}$ for some non-negative integer n, i.e., if $\hat{\rho} = |\rho_n\rangle\langle\rho_n|$, meaning that the system is in the pure state $|\rho_n\rangle$ with probability one. DONE!

Problem 3.4

In this problem we shall explore the density operator for single-photon polarization. Suppose that we are interested in the polarization state of a frequency-ω, $+z$-propagating, single photon. We know that a pure state of such a photon can be written as the 2-D complex-valued ket vector,

$$|\mathbf{i}\rangle \equiv \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix},$$

in the x-y (horizontal-vertical) basis, with $|\alpha_x|^2 + |\alpha_y|^2 = 1$. If we measure the polarization state of this photon using the basis,

$$|\mathbf{i}'\rangle \equiv \begin{bmatrix} \alpha'_x \\ \alpha'_y \end{bmatrix},$$

and

$$|\mathbf{i}'_\perp\rangle \equiv \begin{bmatrix} \alpha'^*_y \\ -\alpha'^*_x \end{bmatrix},$$

where $|\alpha'_x|^2 + |\alpha'_y|^2 = 1$, then we will get outcome $\mathbf{i}'$ with probability

$$\Pr(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,) = |\langle \mathbf{i}'|\mathbf{i}\rangle|^2,$$

and outcome $\mathbf{i}'_\perp$ with probability

$$\Pr(\,\mathbf{i}'_\perp \mid |\mathbf{i}\rangle\,) = |\langle \mathbf{i}'_\perp|\mathbf{i}\rangle|^2 = 1 - \Pr(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,)$$

(a) It is trivial to verify that the density operator for this pure state,

$$\hat{\rho} = |\mathbf{i}\rangle\langle\mathbf{i}|$$

gives these same probabilities via

$$\Pr(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,) = \langle \mathbf{i}'|\hat{\rho}|\mathbf{i}'\rangle,$$

and

$$\Pr(\,\mathbf{i}'_\perp \mid |\mathbf{i}\rangle\,) = \langle \mathbf{i}'_\perp|\hat{\rho}|\mathbf{i}'_\perp\rangle = 1 - \Pr(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,),$$

because

$$\langle \mathbf{i}'|\mathbf{i}\rangle\langle \mathbf{i}|\mathbf{i}'\rangle = |\langle \mathbf{i}'|\mathbf{i}\rangle|^2 = |\alpha_x'^*\alpha_x + \alpha_y'^*\alpha_y|^2,$$

and

$$\langle \mathbf{i}'_\perp|\mathbf{i}\rangle\langle \mathbf{i}|\mathbf{i}'_\perp\rangle = |\langle \mathbf{i}'_\perp|\mathbf{i}\rangle|^2 = |\alpha_y'\alpha_x - \alpha_x'\alpha_y|^2,$$

where the evaluations in terms of the x-y representations will be of use in (b). We also have that

$$\langle \mathbf{i}'|\mathbf{i}\rangle\langle \mathbf{i}|\mathbf{i}'\rangle + \langle \mathbf{i}'_\perp|\mathbf{i}\rangle\langle \mathbf{i}|\mathbf{i}'_\perp\rangle = \langle \mathbf{i}|\left(|\mathbf{i}'\rangle\langle \mathbf{i}'| + |\mathbf{i}'_\perp\rangle\langle \mathbf{i}'_\perp|\right)|\mathbf{i}\rangle = \langle \mathbf{i}|\hat{I}|\mathbf{i}\rangle = 1, \qquad (2)$$

where $\hat{I}$ is the identity operator, and the second equality follows from $\{|\mathbf{i}'\rangle, |\mathbf{i}'_\perp\rangle\}$ being a basis for the polarization state of a $+z$-propagating photon.

(b) Now suppose that the single photon is in a mixed state, i.e., that α_x and α_y are complex-valued random variables whose joint distribution ensures that $|\alpha_x|^2 + |\alpha_y|^2 = 1$ with probability one. To show that the density operator $\hat{\rho}$ can now be written in the form

$$\hat{\rho} = \begin{bmatrix} \langle |\alpha_x|^2\rangle & \langle \alpha_x\alpha_y^*\rangle \\ \langle \alpha_x^*\alpha_y\rangle & \langle |\alpha_y|^2\rangle \end{bmatrix},$$

we will verify that this expression yields the proper formulas for the unconditional measurement probabilities, $\Pr(\mathbf{i}')$ and $\Pr(\mathbf{i}'_\perp)$, i.e.,

$$\langle \mathbf{i}'|\hat{\rho}|\mathbf{i}'\rangle = \Pr(\mathbf{i}') = \int d\alpha_x \int d\alpha_y\, p(\alpha_x,\alpha_y)\Pr(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,),$$

and

$$\langle \mathbf{i}'_\perp|\hat{\rho}|\mathbf{i}'_\perp\rangle = \Pr(\mathbf{i}'_\perp) = \int d\alpha_x \int d\alpha_y\, p(\alpha_x,\alpha_y)\Pr(\,\mathbf{i}'_\perp \mid |\mathbf{i}\rangle\,),$$

where $p(\alpha_x, \alpha_y)$ is the joint probability density for α_x and α_y. This is easy to accomplish by using the x-y representations for $|\mathbf{i}'\rangle$ and $|\mathbf{i}'_\perp\rangle$ in conjunction with the x-y representation for $\hat{\rho}$. We have that

$$\begin{aligned}
\langle \mathbf{i}'|\hat{\rho}|\mathbf{i}'\rangle &= \begin{bmatrix} \alpha_x'^* & \alpha_y'^* \end{bmatrix} \begin{bmatrix} \langle|\alpha_x|^2\rangle & \langle\alpha_x\alpha_y^*\rangle \\ \langle\alpha_x^*\alpha_y\rangle & \langle|\alpha_y|^2\rangle \end{bmatrix} \begin{bmatrix} \alpha_x' \\ \alpha_y' \end{bmatrix} \\
&= |\alpha_x'|^2\langle|\alpha_x|^2\rangle + \alpha_x'^*\alpha_y'\langle\alpha_x\alpha_y^*\rangle + \alpha_x'\alpha_y'^*\langle\alpha_x^*\alpha_y\rangle + |\alpha_y'|^2\langle|\alpha_y|^2\rangle \\
&= \left\langle \begin{bmatrix} \alpha_x'^* & \alpha_y'^* \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} \begin{bmatrix} \alpha_x^* & \alpha_y^* \end{bmatrix} \begin{bmatrix} \alpha_x' \\ \alpha_y' \end{bmatrix} \right\rangle \\
&= \int d\alpha_x \int d\alpha_y \, p(\alpha_x,\alpha_y)|\alpha_x'^*\alpha_x + \alpha_y'^*\alpha_y|^2 \\
&= \int d\alpha_x \int d\alpha_y \, p(\alpha_x,\alpha_y) \,\mathrm{Pr}(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,).
\end{aligned}$$

Likewise we find that

$$\begin{aligned}
\langle \mathbf{i}'_\perp|\hat{\rho}|\mathbf{i}'_\perp\rangle &= \begin{bmatrix} \alpha_y' & -\alpha_x' \end{bmatrix} \begin{bmatrix} \langle|\alpha_x|^2\rangle & \langle\alpha_x\alpha_y^*\rangle \\ \langle\alpha_x^*\alpha_y\rangle & \langle|\alpha_y|^2\rangle \end{bmatrix} \begin{bmatrix} \alpha_y'^* \\ -\alpha_x'^* \end{bmatrix} \\
&= |\alpha_y'|^2\langle|\alpha_x|^2\rangle - \alpha_x'^*\alpha_y'\langle\alpha_x\alpha_y^*\rangle - \alpha_x'\alpha_y'^*\langle\alpha_x^*\alpha_y\rangle + |\alpha_x'|^2\langle|\alpha_y|^2\rangle \\
&= \left\langle \begin{bmatrix} \alpha_y' & -\alpha_x' \end{bmatrix} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} \begin{bmatrix} \alpha_x^* & \alpha_y^* \end{bmatrix} \begin{bmatrix} \alpha_y'^* \\ -\alpha_x'^* \end{bmatrix} \right\rangle \\
&= \int d\alpha_x \int d\alpha_y \, p(\alpha_x,\alpha_y)|\alpha_y'\alpha_x - \alpha_x'\alpha_y|^2 \\
&= \int d\alpha_x \int d\alpha_y \, p(\alpha_x,\alpha_y) \,\mathrm{Pr}(\,\mathbf{i}' \mid |\mathbf{i}\rangle\,).
\end{aligned}$$

Note that we have just shown that the preceding form of the density operator is equivalent to the mixed-state Poincaré vector that we studied in Problem 3.1.

Problem 3.5

Commutators are very important. This problem develops two key points about them.

(a) This part is easy. Because the adjoint of the product of two operators is the reverse-order product of their two adjoints, we have that

$$\left[\hat{A},\hat{B}\right]^\dagger = \hat{B}^\dagger\hat{A}^\dagger - \hat{A}^\dagger\hat{B}^\dagger = \hat{B}\hat{A} - \hat{A}\hat{B} = -\left[\hat{A},\hat{B}\right],$$

where the penultimate equality follows from the fact that $\hat{A}$ and $\hat{B}$ are Hermitian operators. Thus, commutators are anti-Hermitian, viz., they equal minus their adjoints. As a result, if we define an operator $\hat{C}$ via

$$\hat{C} \equiv \frac{1}{j}\left[\hat{A},\hat{B}\right],$$

we find that,

$$\hat{C}^\dagger = \frac{1}{-j}\left[\hat{A},\hat{B}\right]^\dagger = -\frac{1}{-j}\left[\hat{A},\hat{B}\right] = \hat{C},$$

proving that $\hat{C}$ is Hermitian.

(b) First let's employ the $\hat{A}$ eigenvalue/eigenvector properties:

$$\hat{A}|a_n\rangle = a_n|a_n\rangle \quad \text{and} \quad \langle a_n|\hat{A} = a_n\langle a_n|,$$

where we have used the fact that $\hat{A}$ is Hermitian and the $\{a_n\}$ are real, to show that,

$$\langle a_n|\hat{A}\hat{B}|a_m\rangle = a_n\langle a_n|\hat{B}|a_m\rangle,$$

and

$$\langle a_n|\hat{B}\hat{A}|a_m\rangle = a_m\langle a_n|\hat{B}|a_m\rangle,$$

for $1 \leq n, m < \infty$. Because the commutator of $\hat{A}$ and $\hat{B}$ is zero, we know that

$$\langle a_n|\left[\hat{A},\hat{B}\right]|a_m\rangle = 0.$$

The left-hand side of the preceding equation can then be expanded to yield,

$$\langle a_n|\hat{A}\hat{B}|a_m\rangle - \langle a_n|\hat{B}\hat{A}|a_m\rangle = (a_n - a_m)\langle a_n|\hat{B}|a_m\rangle = 0.$$

So, because the eigenvalues of $\hat{A}$ are distinct, we get

$$\langle a_n|\hat{B}|a_m\rangle = 0, \quad \text{for } n \neq m.$$

Because the $\{|a_n\rangle\}$ are complete, this result implies that

$$\hat{B}|a_n\rangle = K_n|a_n\rangle,$$

where K_n is a constant, i.e., $|a_n\rangle$ is an eigenket of $\hat{B}$. This proof works for every eigenket of $\hat{A}$: every $\hat{A}$ eigenket is also a $\hat{B}$ eigenket. To prove that the converse is true, we merely start from

$$\hat{B}|b_n\rangle = b_n|b_n\rangle \quad \text{and} \quad \langle b_n|\hat{B} = b_n\langle b_n|,$$

and then use the zero-commutator to get,

$$(b_m - b_n)\langle b_n|\hat{A}|b_m\rangle = 0, \quad \text{for } n \neq m.$$

Because the eigenvalues of $\hat{B}$ are distinct, this implies that

$$\langle b_n|\hat{A}|b_m\rangle = 0, \quad \text{for } n \neq m,$$

and because the $\{|b_n\rangle\}$ are complete we find that

$$\hat{A}|b_n\rangle = K_n'|b_n\rangle,$$

for some constant K_n', viz., every $|b_n\rangle$ is an eigenket of $\hat{A}$.

Problem 3.6
Here we introduce the notation of tensor products, to permit us to deal with multiple quantum systems.

(a) Basically, this problem is trying to convince you that tensor product stuff is notationally cumbersome, but really easy to work with. Suppose we start with a product state, $|\phi_n\rangle_1 \otimes |\theta_m\rangle_2$ from the basis $\{\, |\phi_n\rangle_1 \otimes |\theta_m\rangle_2 : 1 \leq n, m \leq \infty \,\}$ discussed in the problem statement. The adjoint operator $\hat{C}^\dagger$ must satisfy,

$$\begin{aligned}
(\langle_2\langle\theta_m| \otimes {}_1\langle\phi_n|)[\hat{C}^\dagger(|\phi_k\rangle_1 \otimes |\theta_l\rangle_2)] &= \{({}_2\langle\theta_l| \otimes {}_1\langle\phi_k|)[\hat{C}(|\phi_n\rangle_1 \otimes |\theta_m\rangle_2)]\}^* \\
&= \{({}_2\langle\theta_l| \otimes {}_1\langle\phi_k|)[(\hat{A}_1|\phi_n\rangle_1) \otimes (\hat{B}_2|\theta_m\rangle_2)]\}^* \\
&= ({}_1\langle\phi_k|\hat{A}_1|\phi_n\rangle_1)^*({}_2\langle\theta_l|\hat{B}_2|\theta_m\rangle_2)^* \\
&= ({}_1\langle\phi_n|\hat{A}_1|\phi_k\rangle_1)({}_2\langle\theta_m|\hat{B}_2|\theta_l\rangle_2),
\end{aligned}$$

where the last equality uses the fact that $\hat{A}_1$ and $\hat{B}_2$ are Hermitian. Because this result must hold for all n, m, k, l, it follows that $\hat{C}^\dagger = \hat{A}_1 \otimes \hat{B}_2 = \hat{C}$, i.e., $\hat{C}$ is Hermitian.

Let $\{\, |a_n\rangle_1 : 1 \leq n < \infty \,\}$ and $\{\, |b_m\rangle_2 : 1 \leq m < \infty \,\}$ be the eigenkets of $\hat{A}_1$ and $\hat{B}_2$, respectively. These eigenkets are CON sets on their respective Hilbert spaces, $\mathcal{H}_1$ and $\mathcal{H}_2$. We now have that,

$$\begin{aligned}
\hat{C}(|a_n\rangle_1 \otimes |b_m\rangle_2) &= (\hat{A}_1|a_n\rangle_1) \otimes (\hat{B}_2|b_m\rangle_2) = (a_n|a_n\rangle_1) \otimes (b_m|b_m\rangle_2) \\
&= a_n b_m (|a_n\rangle_1 \otimes |b_m\rangle_2),
\end{aligned}$$

so that $|a_n\rangle_1 \otimes |b_m\rangle_2$ is an eigenket of $\hat{C}$ with associated eignevalue $a_n b_m$, for $1 \leq n, m < \infty$. Because of the CON nature of $\{|a_n\rangle_1\}$ and $\{|b_m\rangle_2\}$ on their respective Hilbert spaces, it follows that $\{|a_n\rangle_1 \otimes |b_m\rangle_2\}$ is CON on $\mathcal{H}$.

(b) It is straightforward to show that

$$\Pr(a_n, b_m) = |\langle\psi|(|a_n\rangle_1 \otimes |b_m\rangle_2)|^2,$$

is a proper probability distribution. Because of the magnitude squared operation, the probability is non-negative. The Schwarz inequality guarantees that

$$\begin{aligned}
\Pr(a_n, b_m) &\leq |\langle\psi|\psi\rangle|^2 |({}_2\langle b_m| \otimes {}_1\langle a_n|)(|a_n\rangle_1 \otimes |b_m\rangle_2)|^2 \\
&= |\langle\psi|\psi\rangle|^2 ||{}_1\langle a_n|a_n\rangle_1|^2 |{}_2\langle b_m|b_m\rangle_2|^2 = 1,
\end{aligned}$$

where the last equality follows because $|\psi\rangle$, $|a_n\rangle_1$, and $|b_m\rangle_2$ are all unit-length kets. To show that the probability distribution sums to one, we argue as follows:

$$\begin{aligned}
\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\Pr(a_n,b_m) &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\langle\psi|(|a_n\rangle_1\otimes|b_m\rangle_2)({}_2\langle b_m|\otimes\,{}_1\langle a_n|)|\psi\rangle \\
&= \langle\psi|\left(\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}(|a_n\rangle_1\otimes|b_m\rangle_2)({}_2\langle b_m|\otimes\,{}_1\langle a_n|)\right)|\psi\rangle \\
&= \langle\psi|\left[\left(\sum_{n=1}^{\infty}|a_n\rangle_{11}\langle a_n|\right)\otimes\left(\sum_{m=1}^{\infty}|b_m\rangle_{22}\langle b_m|\right)\right]|\psi\rangle \\
&= \langle\psi|\left(\hat{I}_1\otimes\hat{I}_2\right)|\psi\rangle,
\end{aligned}$$

where in the next to last equality we have used the tensor form of the outer product

$$(|a_n\rangle_1\otimes|b_m\rangle_2)({}_2\langle b_m|\otimes\,{}_1\langle a_n|) = (|a_n\rangle_{11}\langle a_n|)\otimes(|b_m\rangle_{22}\langle b_m|),$$

and in the last equality we have used the completeness relations for the $\hat{A}_1$ and $\hat{B}_2$ eigenkets. So, because the identity operator for $\mathcal{H}\equiv\mathcal{H}_1\otimes\mathcal{H}_2$ satisfies $\hat{I}=\hat{I}_1\otimes\hat{I}_2$ in terms of the identity operators on $\mathcal{H}_1$ and $\mathcal{H}_2$, we have

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\Pr(a_n,b_m) = \langle\psi|\hat{I}|\psi\rangle = \langle\psi|\psi\rangle = 1,$$

as was to be shown.

Our next task is to find the marginal probability distributions. Using the intermediate steps of the previous derivation we have that,

$$\begin{aligned}
\Pr(a_n) &= \sum_{m=1}^{\infty}\Pr(a_n,b_m) = \langle\psi|\left(\sum_{m=1}^{\infty}(|a_n\rangle_1\otimes|b_m\rangle_2)({}_2\langle b_m|\otimes\,{}_1\langle a_n|)\right)|\psi\rangle \\
&= \langle\psi|\left[|a_n\rangle_{11}\langle a_n|\otimes\left(\sum_{m=1}^{\infty}|b_m\rangle_{22}\langle b_m|\right)\right]|\psi\rangle \\
&= \langle\psi|\left(|a_n\rangle_{11}\langle a_n|\otimes\hat{I}_2\right)|\psi\rangle.
\end{aligned}$$

A similar procedure yields,

$$\Pr(b_m) = \langle\psi|\left(\hat{I}_1\otimes|b_m\rangle_{22}\langle b_m|\right)|\psi\rangle.$$

(c) We now use the results of (b) for $|\psi\rangle = |\psi_1\rangle_1\otimes|\psi_2\rangle_2$, i.e., a product state. Here we find that,

$$\begin{aligned}
\Pr(a_n) &= \langle\psi|\left(|a_n\rangle_{11}\langle a_n|\otimes\hat{I}_2\right)|\psi\rangle \\
&= (\,{}_1\langle\psi_1|a_n\rangle_{11}\langle a_n|\psi_1\rangle_1)\,{}_2\langle\psi_2|\hat{I}_2|\psi_2\rangle_2 \\
&= {}_1\langle\psi_1|a_n\rangle_{11}\langle a_n|\psi_1\rangle_1 = |{}_1\langle\psi_1|a_n\rangle_1|^2.
\end{aligned}$$

A similar derivation yields,

$$\Pr(b_m) = |{}_2\langle\psi_2|b_m\rangle_2|^2.$$

These results are very interesting because they are, respectively, the probability distributions for measuring $\hat{A}_1$ on system 1 when it is in state $|\psi_1\rangle_1$ and for measuring $\hat{B}_2$ on system 2 when it is in state $|\psi_2\rangle_2$.

Turning to the joint statistics we find that the product state $|\psi\rangle$ gives,

$$\begin{aligned}\Pr(a_n, b_m) &= |(\,{}_2\langle\psi_2| \otimes {}_1\langle\psi_1|)(|a_n\rangle_1 \otimes |b_m\rangle_2)|^2 \\ &= (|{}_1\langle\psi_1|a_n\rangle_1|^2)(|{}_2\langle\psi_2|b_m\rangle_2|^2) \\ &= \Pr(a_n)\Pr(b_m).\end{aligned}$$

So, we have shown that measuring a product observable, $\hat{A}_1$ on system 1 and $\hat{B}_2$ on system 2 when the composite system $\mathcal{S}$ is in a product state $|\psi\rangle = |\psi_1\rangle_1 \otimes |\psi_2\rangle_2$ leads to statistically independent outcomes. This is *not* generally the case when $|\psi\rangle$ isn't a product state, as we shall see later this term when we treat entanglement.

Problem 3.7

Here we prove that it is impossible to clone the unknown state of a quantum system by means of a unitary evolution. It is a proof by contradiction. Suppose that we have a quantum system whose Hilbert space of states is $\mathcal{H}_S$, where ${}_S$ indicates that this is the *source* system. Suppose too that we have a *target* system whose Hilbert space of states is $\mathcal{H}_T$. We will assume that these two Hilbert spaces have the same dimensionality, e.g., 2.

We wish to construct a perfect cloner, viz., a unitary operator, $\hat{U}$, on the tensor product space $\mathcal{H} \equiv \mathcal{H}_S \otimes \mathcal{H}_T$ such that

$$\hat{U}(|\psi\rangle_S \otimes |0\rangle_T) = |\psi\rangle_S \otimes |\psi\rangle_T, \tag{3}$$

where $|\psi\rangle_S$ is an *arbitrary* unit-length ket in $\mathcal{H}_S$, and $|0\rangle_T$ is a reference ("blank") unit-length ket in $\mathcal{H}_T$.

Let $|\psi_1\rangle_S$ and $|\psi_2\rangle_S$ be two distinct, unit-length kets in $\mathcal{H}_S$, let α and β be two non-zero complex numbers, and assume that we have found an ideal cloner operator $\hat{U}$ satisfying Eq. (3) for all unit-length source kets.

(a) We define

$$|\psi'\rangle_S = \frac{\alpha|\psi_1\rangle_S + \beta|\psi_2\rangle_S}{\sqrt{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}}.$$

By inspection, we have that $|\psi'\rangle_S$ is a unit-length ket in $\mathcal{H}_S$, i.e., the denominator is what is needed to normalize the numerator to unit length. Thus, because $\hat{U}$ is a unitary operator on $\mathcal{H}$, we have that d_θ, the length of

$$|\theta\rangle \equiv \hat{U}(|\psi'\rangle_S \otimes |0\rangle_T),$$

is given by

$$
\begin{aligned}
d_\theta^2 &= \langle\theta|\theta\rangle = ({}_T\langle 0| \otimes {}_S\langle\psi'|)\hat{U}^\dagger\hat{U}(|\psi'\rangle_S \otimes |0\rangle_T) \\
&= ({}_T\langle 0| \otimes {}_S\langle\psi'|)\hat{I}(|\psi'\rangle_S \otimes |0\rangle_T) = ({}_T\langle 0|0\rangle_T)({}_S\langle\psi'|\psi'\rangle_S) = 1.
\end{aligned}
$$

(b) We now expand out $|\psi'\rangle_S$ appearing in the tensor product $|\psi'\rangle_S \otimes |0\rangle_T$ and obtain

$$
\begin{aligned}
|\psi'\rangle_S \otimes |0\rangle_T &= \left(\frac{\alpha}{\sqrt{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}}\right)(|\psi_1\rangle_S \otimes |0\rangle_T) \\
&+ \left(\frac{\beta}{\sqrt{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}}\right)(|\psi_2\rangle_S \otimes |0\rangle_T). \\
&= \alpha'(|\psi_1\rangle_S \otimes |0\rangle_T) + \beta'(|\psi_2\rangle_S \otimes |0\rangle_T), \qquad (4)
\end{aligned}
$$

with the obvious definitions for α' and β'. From the linearity of $\hat{U}$ we then have that

$$
\begin{aligned}
|\theta\rangle &= \alpha'\hat{U}(|\psi_1\rangle_S \otimes |0\rangle_T) + \beta'\hat{U}(|\psi_2\rangle_S \otimes |0\rangle_T) \\
&= \alpha'(|\psi_1\rangle_S \otimes |\psi_1\rangle_T) + \beta'(|\psi_2\rangle_S \otimes |\psi_2\rangle_T).
\end{aligned}
$$

(c) From (b) we have that the length of $|\theta\rangle$ obeys

$$
\begin{aligned}
d_\theta^2 &= \langle\theta|\theta\rangle \\
&= |\alpha'|^2 + |\beta'^2| + 2\mathrm{Re}[\alpha'^*\beta'({}_S\langle\psi_1|\psi_2\rangle_S)^2] \\
&= \frac{|\alpha|^2 + |\beta^2| + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)^2]}{|\alpha|^2 + |\beta^2| + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}.
\end{aligned}
$$

This expression for d_θ does *not* equal 1 for non-zero α and β unless ${}_S\langle\psi_1|\psi_2\rangle_S = 0$ or 1. Thus, we have a contradiction in that Eq. (3) cannot be satisfied for arbitrary source states. So, there does not exist a unitary operator $\hat{U}$ that is a perfect cloner.

Problem 3.8

Here we prove that it is impossible to erase the unknown state of a quantum system by means of a unitary evolution. It is a proof by contradiction. Suppose that we have a quantum system whose Hilbert space of states is $\mathcal{H}_S$, where ${}_S$ indicates that this is the *source* system. Suppose too that we have an *ancilla* system whose Hilbert space of states is $\mathcal{H}_A$. We will assume that these two Hilbert spaces have the same dimensionality, e.g., 2.

We wish to construct a perfect eraser, viz., a unitary operator, $\hat{U}$, on the tensor product space $\mathcal{H} \equiv \mathcal{H}_S \otimes \mathcal{H}_A$ such that

$$\hat{U}(|\psi\rangle_S \otimes |0\rangle_A) = |0\rangle_S \otimes |0\rangle_A, \tag{5}$$

where $|\psi\rangle_S$ is an *arbitrary* unit-length ket in $\mathcal{H}_S$, and $|0\rangle_A$ is a reference ("blank") unit-length ket in $\mathcal{H}_A$.

Let $|\psi_1\rangle_S$ and $|\psi_2\rangle_S$ be two distinct, unit-length kets in $\mathcal{H}_S$, let α and β be two non-zero complex numbers, and assume that we have found an ideal eraser operator $\hat{U}$ satisfying Eq. (5) for all unit-length source kets.

(a) We define

$$|\psi'\rangle_S = \frac{\alpha|\psi_1\rangle_S + \beta|\psi_2\rangle_S}{\sqrt{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}}.$$

By inspection, we have that $|\psi'\rangle_S$ is a unit-length ket in $\mathcal{H}_S$, i.e., the denominator is what is needed to normalize the numerator to unit length. Thus, because $\hat{U}$ is a unitary operator on $\mathcal{H}$, we have that d_θ, the length of

$$|\theta\rangle \equiv \hat{U}(|\psi'\rangle_S \otimes |0\rangle_A),$$

is given by

$$\begin{aligned} d_\theta^2 &= \langle\theta|\theta\rangle = ({}_A\langle 0| \otimes {}_S\langle\psi'|)\hat{U}^\dagger\hat{U}(|\psi'\rangle_S \otimes |0\rangle_A) \\ &= ({}_A\langle 0| \otimes {}_S\langle\psi'|)\hat{I}(|\psi'\rangle_S \otimes |0\rangle_A) = ({}_A\langle 0|0\rangle_A)({}_S\langle\psi'|\psi'\rangle_S) = 1. \end{aligned}$$

(b) We now expand out $|\psi'\rangle_S$ appearing in the tensor product $|\psi'\rangle_S \otimes |0\rangle_A$ and obtain

$$\begin{aligned} |\psi'\rangle_S \otimes |0\rangle_A &= \left(\frac{\alpha}{\sqrt{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}}\right)(|\psi_1\rangle_S \otimes |0\rangle_A) \\ &+ \left(\frac{\beta}{\sqrt{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}}\right)(|\psi_2\rangle_S \otimes |0\rangle_A). \\ &= \alpha'(|\psi_1\rangle_S \otimes |0\rangle_A) + \beta'(|\psi_2\rangle_S \otimes |0\rangle_A), \end{aligned} \tag{6}$$

with the obvious definitions for α' and β'. From the linearity of $\hat{U}$we then have that

$$\begin{aligned} |\theta\rangle &= \alpha'\hat{U}(|\psi_1\rangle_S \otimes |0\rangle_A) + \beta'\hat{U}(|\psi_2\rangle_S \otimes |0\rangle_A) \\ &= (\alpha' + \beta')(|0\rangle_S \otimes |0\rangle_A). \end{aligned}$$

(c) From (b) we have that the length of $|\theta\rangle$ obeys

$$\begin{aligned} d_\theta^2 &= \langle\theta|\theta\rangle \\ &= |\alpha'|^2 + |\beta'|^2 + 2\mathrm{Re}(\alpha'^*\beta') \\ &= \frac{|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}(\alpha^*\beta)}{|\alpha|^2 + |\beta^2| + 2\mathrm{Re}[\alpha^*\beta({}_S\langle\psi_1|\psi_2\rangle_S)]}. \end{aligned}$$

This expression for d_θ does *not* equal 1 for non-zero α and β unless ${}_S\langle\psi_1|\psi_2\rangle_S = 1$, which would mean that $|\psi_1\rangle_S = |\psi_2\rangle_S$. Thus, we have a contradiction in that Eq. (5) cannot be satisfied for arbitrary source states. So, there does not exist a unitary operator $\hat{U}$ that is a perfect eraser.