Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.453 QUANTUM OPTICAL COMMUNICATION

Problem Set 1 Solutions

Fall 2016

Problem 1.1

Here we shall verify the elementary properties of the 1-D Gaussian probability density function (pdf),

$$p_x(X) = \frac{e^{-(X-m)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}, \quad \text{for } -\infty < X < \infty.$$

(a) For a 1-D deterministic function to be a pdf, it must be non-negative and integrate to one. It is clear that $e^{-(X-m)^2/2\sigma^2}$ is non-negative. To demonstrate that it integrates to one—without recourse to integral tables—we proceed as follows. We have that

$$\left(\int_{-\infty}^{\infty} dX\, e^{-(X-m)^2/2\sigma^2}\right)^2 = \int_{-\infty}^{\infty} dX \int_{-\infty}^{\infty} dY\, e^{-(X-m)^2/2\sigma^2-(Y-m)^2/2\sigma^2},$$

by writing out the square of the single integral as the product of single integrals with different dummy variables of integration, and then combining the product of these single integrals into a double integral. Converting the double integral to polar coordinates, via $X - m = R\cos(\Phi)$, $Y - m = R\sin(\Phi)$ and $dX\,dY = R\,dR\,d\Phi$, yields,

$$\begin{aligned}\int_{-\infty}^{\infty} dX \int_{-\infty}^{\infty} dY\, e^{-(X-m)^2/2\sigma^2-(Y-m)^2/2\sigma^2} &= \int_0^{2\pi} d\Phi \int_0^{\infty} dR\, Re^{-R^2/2\sigma^2} \\ &= 2\pi \int_0^{\infty} dR\, Re^{-R^2/2\sigma^2} = 2\pi\sigma^2,\end{aligned}$$

where we have first done the Φ integral and then the R integral. Taking the square root of this result then verifies the normalization constant for the 1-D Gaussian pdf.

(b) We have that

$$\begin{aligned}M_x(jv) &= E(e^{jvx}) = \int_{-\infty}^{\infty} dX\, \frac{e^{jvX-(X-m)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \\ &= \int_{-\infty}^{\infty} dX\, \frac{e^{jvm-v^2\sigma^2/2-[X-(m+jv\sigma^2)]^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} = e^{jvm-v^2\sigma^2/2},\end{aligned}$$

where we have used the fact that

$$\int_{-\infty}^{\infty} dX\, \frac{e^{-(X-a)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} = 1,$$

is valid when $\sigma^2 > 0$, even if a is complex valued.

(c) To get the mean value of x we differentiate $M_x(jv)$ once,

$$E(x) = \left.\frac{dM_x(jv)}{d(jv)}\right|_{jv=j0} = \left.\left[(m + jv\sigma^2)e^{jvm - v^2\sigma^2/2}\right]\right|_{jv=j0} = m.$$

To get the mean-square value of x, we differentiate once more,

$$\begin{aligned} E(x^2) &= \left.\frac{d^2M_x(jv)}{d(jv)^2}\right|_{jv=j0} \\ &= \left.\left\{[(m + jv\sigma^2)^2 + \sigma^2]e^{jvm - v^2\sigma^2/2}\right\}\right|_{jv=j0} = m^2 + \sigma^2. \end{aligned}$$

Now, using $\text{var}(x) = E(x^2) - [E(x)]^2$, we find that $\text{var}(x) = \sigma^2$.

Problem 1.2

Here we shall verify the elementary properties of the Poisson probability mass function (pmf),

$$P_x(n) = \frac{m^n}{n!}e^{-m}, \quad \text{for } n = 0, 1, 2, \ldots, \text{ and } m \geq 0.$$

(a) A probability mass function must be non-negative and sum to one. The Poisson pmf is clearly non-negative. To prove that it is properly normalized we use the power series for e^z to verify that,

$$\sum_{n=0}^{\infty} P_x(n) = e^{-m}\sum_{n=0}^{\infty}\frac{m^n}{n!} = e^{-m}e^m = 1.$$

(b) The characteristic function associated with the Poisson pmf is found via a similar power series calculation:

$$M_x(jv) = E(e^{jvx}) = e^{-m}\sum_{n=0}^{\infty}\frac{e^{jvn}m^n}{n!} = e^{-m}\exp(e^{jv}m) = \exp[m(e^{jv} - 1)].$$

(c) We differentiate $M_x(jv)$ once to get $E(x)$:

$$E(x) = \left.\frac{dM_x(jv)}{d(jv)}\right|_{jv=j0} = \left.\left(me^{jv}\exp[m(e^{jv} - 1)]\right)\right|_{jv=j0} = m.$$

We differentiate a second time to obtain $E(x^2)$:

$$E(x^2) = \left.\frac{d^2M_x(jv)}{d(jv)^2}\right|_{jv=j0} = \left.\left[(me^{jv} + m^2e^{2jv})\exp[m(e^{jv} - 1)]\right]\right|_{jv=j0} = m + m^2.$$

Now, using $\text{var}(x) = E(x^2) - [E(x)]^2$, we find that $\text{var}(x) = m$.

Problem 1.3
Here we perform a simple 1-D random-variable transformation, using the method of events.

(a) The probability distribution function of $y = x^2$ is,

$$\begin{aligned} F_y(Y) &\equiv \Pr(y \le Y) = \Pr(|x| \le \sqrt{Y}) = \int_0^{\sqrt{Y}} dX\, \frac{X}{\sigma^2} e^{-X^2/2\sigma^2} \\ &= 1 - e^{-Y/2\sigma^2}, \quad \text{for } Y \ge 0. \end{aligned}$$

The pdf of y is obtained by differentiating its probability distribution function,

$$p_y(Y) = \frac{dF_y(Y)}{dY} = \begin{cases} \dfrac{e^{-Y/2\sigma^2}}{2\sigma^2}, & \text{for } Y \ge 0, \\ 0, & \text{otherwise,} \end{cases}$$

i.e., y is exponentially distributed.

(b) The moment integrals for the exponential distribution are straightforward. The factorial integral

$$\int_0^\infty dZ\, Z^n e^{-Z} = n!, \quad \text{for } n = 0, 1, 2, \ldots,$$

plus the change of variables $Z = Y/2\sigma^2$ yields:

$$E(y^n) = \int_0^\infty dY\, Y^n e^{-Y/2\sigma^2}/2\sigma^2 = 2^n \sigma^{2n} n!.$$

Hence, we find $E(y) = 2\sigma^2$, $E(y^2) = 8\sigma^4$, and $\text{var}(y) = E(y^2) - [E(y)]^2 = 4\sigma^4$.

Problem 1.4
Here we perform a simple 2-D random variable transformation, using the method of events.

(a) The joint pdf for r, ϕ can be found by differentiating the joint probability distribution function,

$$F_{r,\phi}(R, \Phi) \equiv \Pr(r \le R, \phi \le \Phi),$$

for these random variables. This joint distribution function can, in turn, be calculated from the joint pdf of x, y, as follows,

$$\begin{aligned} F_{r,\phi}(R, \Phi) &= \iint_{\{(X,Y): r \le R, \phi \le \Phi\}} dX\, dY\, p_{x,y}(X, Y) \\ &= \int_0^\Phi d\theta \int_0^R \rho\, d\rho\, p_{x,y}(\rho\cos(\theta), \rho\sin(\theta)) \\ &= \int_0^\Phi d\theta \int_0^R \rho\, d\rho\, \frac{e^{-\rho^2/2\sigma^2}}{2\pi\sigma^2} = \frac{\Phi}{2\pi}(1 - e^{-R^2/2\sigma^2}), \end{aligned}$$

where we have converted to polar coordinates in order to do the integrations. It is now easy to find the joint pdf of r, ϕ:

$$p_{r,\phi}(R,\Phi) = \frac{\partial^2 F_{r,\phi}(R,\Phi)}{\partial R \partial \Phi} = \begin{cases} \dfrac{R}{2\pi\sigma^2} e^{-R^2/2\sigma^2}, & \text{for } 0 \le R,\ 0 \le \Phi < 2\pi, \\ 0, & \text{otherwise.} \end{cases}$$

(b) To find the marginal distributions we can integrate out the unwanted variable from the joint distribution. This procedure yields,

$$p_r(R) = \int_0^{2\pi} d\Phi\, p_{r,\phi}(R,\Phi) = \begin{cases} \dfrac{R}{\sigma^2} e^{-R^2/2\sigma^2}, & \text{for } 0 \le R, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$p_\phi(\Phi) = \int_0^{\infty} dR\, p_{r,\phi}(R,\Phi) = \begin{cases} \dfrac{1}{2\pi}, & \text{for } 0 \le \Phi < 2\pi, \\ 0, & \text{otherwise.} \end{cases}$$

Because $p_{r,\phi}(R,\Phi) = p_r(R)p_\phi(\Phi)$, for all R, Φ, we see that the random variables r and ϕ are statistically independent. Moreover, r is Rayleigh distributed and ϕ is uniformly distributed.

Problem 1.5

Here we will learn about a pmf that will show up in our quantum optics work.

(a) The unconditional pmf of N is found by averaging its conditional pmf over the statistics for x:

$$\begin{aligned} P_N(n) &= \int_{-\infty}^{\infty} dX\, P_{N|x}(\,n \mid x = X\,)p_x(X) \\ &= \int_0^{\infty} dX\, \frac{X^n}{n!} e^{-X} \frac{e^{-X/m}}{m} \\ &= \frac{1}{n!m} \int_0^{\infty} dX\, X^n e^{-X(m+1)/m} \\ &= \frac{m^n}{n!(m+1)^{n+1}} \int_0^{\infty} dZ\, Z^n e^{-Z} = \frac{m^n}{(m+1)^{n+1}}, \quad \text{for } n = 0, 1, 2, \ldots, \end{aligned}$$

where we have used the change of variable $Z = X(m+1)/m$ in the penultimate equality. This pmf is called the Bose-Einstein distribution.

(b) The characteristic function for the Bose-Einstein pmf is found as follows:

$$\begin{aligned}
M_N(jv) &= \sum_{n=0}^{\infty} e^{jvn} \frac{m^n}{(m+1)^{n+1}} \\
&= \frac{1}{m+1} \sum_{n=0}^{\infty} \left(\frac{me^{jv}}{m+1} \right)^n = \frac{1}{(m+1)[1 - me^{jv}/(m+1)]} \\
&= \frac{1}{1 - m(e^{jv} - 1)},
\end{aligned}$$

where the penultimate equality is due to the geometric series formula,

$$\sum_{n=0}^{\infty} Z^n = \frac{1}{1-Z}, \quad \text{for } |Z| < 1.$$

(c) Differentiating $M_N(jv)$ once yields,

$$E(N) = \left. \frac{dM_N(jv)}{d(jv)} \right|_{jv=j0} = \left. \left(\frac{me^{jv}}{[1 - m(e^{jv} - 1)]^2} \right) \right|_{jv=j0} = m.$$

Differentiating a second time we obtain:

$$\begin{aligned}
E(N^2) &= \left. \frac{d^2 M_N(jv)}{d(jv)^2} \right|_{jv=j0} = \left. \left(\frac{me^{jv}}{[1 - m(e^{jv} - 1)]^2} + \frac{2m^2 e^{2jv}}{[1 - m(e^{jv} - 1)]^3} \right) \right|_{jv=j0} \\
&= m + 2m^2.
\end{aligned}$$

Now, using $\text{var}(N) = E(N^2) - [E(N)]^2$, we find that $\text{var}(x) = m + m^2$.

Problem 1.6

Here we will take a first step toward understanding jointly Gaussian random variables.

(a) The transformation

$$\begin{aligned}
w &= x\cos(\theta) + y\sin(\theta) \\
z &= -x\sin(\theta) + y\cos(\theta),
\end{aligned}$$

is equivalent to the picture shown in Fig. 1. Clearly this is rotation by θ.

(b) Because the transformation is linear and x, y are jointly Gaussian, we know that w, z are also going to be jointly Gaussian. Hence all we need to do is to find the

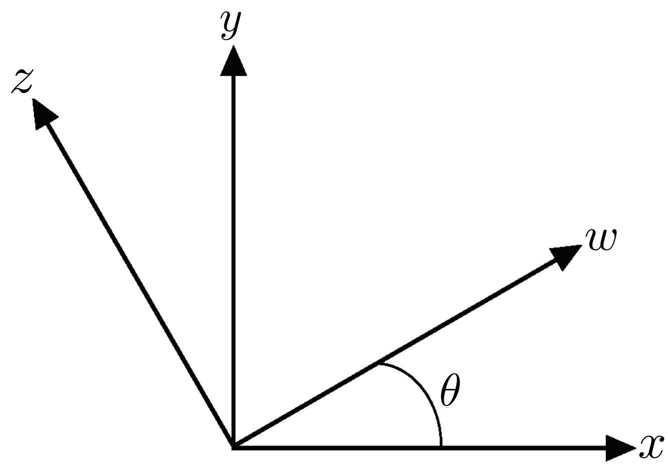

Figure 1: Transformation from (X, Y) to (W, Z)

first and second moments of the new variables and substitute into the standard 2-D Gaussian pdf. We have that,

$$\begin{aligned} E(w) &= E(x)\cos(\theta) + E(y)\sin(\theta) = 0 \\ E(z) &= -E(x)\sin(\theta) + E(y)\cos(\theta) = 0, \end{aligned}$$

because of the linearity of expectation—the average of the sum is the sum of the averages, the average of a constant times a random variable is the constant times the average of the random variable—plus the fact that x and y are both zero mean. So, because w and z are zero mean, their variances—like those for the zero-mean random variables x and y—equal their respective mean-square values. To find these mean-square values we square out the transformation that defines w and z and average:

$$\begin{aligned} \sigma_w^2 &= E(w^2) = E(x^2)\cos^2(\theta) + 2E(xy)\cos(\theta)\sin(\theta) + E(y^2)\sin^2(\theta) \\ &= \sigma^2\cos^2(\theta) + 2\rho\sigma^2\cos(\theta)\sin(\theta) + \sigma^2\sin^2(\theta) \\ \sigma_z^2 &= E(z^2) = E(x^2)\sin^2(\theta) - 2E(xy)\sin(\theta)\cos(\theta) + E(y^2)\cos^2(\theta) \\ &= \sigma^2\sin^2(\theta) - 2\rho\sigma^2\sin(\theta)\cos(\theta) + \sigma^2\cos^2(\theta), \end{aligned}$$

where we have used the zero-mean property to obtain $E(xy) = \text{cov}(x, y)$. Now, with standard trig identities, we can reduce these expressions to,

$$\begin{aligned} \sigma_w^2 &= \sigma^2[1 + \rho\sin(2\theta)], \\ \sigma_z^2 &= \sigma^2[1 - \rho\sin(2\theta)]. \end{aligned}$$

To complete the information needed to pin down the joint pdf of w and z, we must find their covariance. Because they are zero mean, this is found via,

$$\begin{aligned}
\lambda_{wz} &= E(wz) \\
&= -E(x^2)\cos(\theta)\sin(\theta) + E(xy)[\cos^2(\theta) - \sin^2(\theta)] + E(y^2)\sin(\theta)\cos(\theta) \\
&= \rho\sigma^2\cos(2\theta).
\end{aligned}$$

The joint pdf for w and z can now be written down:

$$p_{w,z}(W,Z) = \frac{\exp\left[-\dfrac{\sigma_z^2 W^2 - 2\lambda_{wz}WZ + \sigma_w^2 Z^2}{2(\sigma_w^2\sigma_z^2 - \lambda_{wz}^2)}\right]}{2\pi\sqrt{\sigma_w^2\sigma_z^2 - \lambda_{wz}^2}},$$

where we have used the fact that w and z are zero mean, and the variances and covariance are as derived above.

(c) To make w and z statistically independent, it is sufficent to make them uncorrelated, because they are jointly Gaussian. To be uncorrelated, in turn, means we need to choose θ to make $\lambda_{wz} = \rho\sigma^2\cos(2\theta) = 0$. We have restricted θ to lie between 0 and $\pi/2$, hence the value we need is $\theta = \pi/4$. With this choice, we find that w and z are statistically independent, zero-mean Gaussian random variables, with $\sigma_w^2 = \sigma^2(1+\rho)$ and $\sigma_z^2 = \sigma^2(1-\rho)$. Because $|\rho| \leq 1$, both of these variances will be non-negative.

Problem 1.7

Here we shall examine some of the eigenvalue/eigenvector properties of an Hermitian matrix.

(a) We know that vector-matrix multiplication is associative, so $\mathbf{y}^\dagger(A\mathbf{x}) = (\mathbf{y}^\dagger A)\mathbf{x}$. We also know that the $(A^\dagger\mathbf{y})^\dagger = \mathbf{y}^\dagger A$, and hence we see that $B = A^\dagger$ is the matrix that satisfies $(B\mathbf{y})^\dagger\mathbf{x} = \mathbf{y}^\dagger(A\mathbf{x})$ for all $\mathbf{x},\mathbf{y} \in \mathcal{C}^N$.

(b) From the eigenvalue/eigenvector property we have that

$$\boldsymbol{\phi}_n^\dagger A\boldsymbol{\phi}_n = \boldsymbol{\phi}_n^\dagger(A\boldsymbol{\phi}_n) = \mu_n\boldsymbol{\phi}_n^\dagger\boldsymbol{\phi}_n = \mu_n,$$

and, because $A^\dagger = A$, we also have that

$$\boldsymbol{\phi}_n^\dagger A\boldsymbol{\phi}_n = (\boldsymbol{\phi}_n^\dagger A)\boldsymbol{\phi}_n = \mu_n^*\boldsymbol{\phi}_n^\dagger\boldsymbol{\phi}_n = \mu_n^*.$$

Equating these two results makes it clear that μ_n is real, for $1 \leq n \leq N$.

(c) Using the eigenvalue/eigenvector property we have that,

$$\boldsymbol{\phi}_m^\dagger A \boldsymbol{\phi}_n = \boldsymbol{\phi}_m^\dagger (A \boldsymbol{\phi}_n) = \mu_n \boldsymbol{\phi}_m^\dagger \boldsymbol{\phi}_n.$$

Again using the fact that A is Hermitian—plus the result from (b), that the $\{\mu_n\}$ are real—we have that,

$$\boldsymbol{\phi}_m^\dagger A \boldsymbol{\phi}_n = (\boldsymbol{\phi}_m^\dagger A) \boldsymbol{\phi}_n = \mu_m \boldsymbol{\phi}_m^\dagger \boldsymbol{\phi}_n.$$

Equating these two results makes it clear that if $\mu_m \neq \mu_n$, then $\boldsymbol{\phi}_m^\dagger \boldsymbol{\phi}_n = 0$ must prevail, i.e., the eigenvectors associated with distinct eigenvalues are orthogonal.

(d) If $\boldsymbol{\phi}$ and $\boldsymbol{\phi}'$ are two linearly independent N-D vectors, then, via the Gram-Schmidt process, we can find constants $\{a, b, c, d\}$ such that

$$\boldsymbol{\theta} \equiv a\boldsymbol{\phi} + b\boldsymbol{\phi}',$$

$$\boldsymbol{\theta}' \equiv c\boldsymbol{\phi} + d\boldsymbol{\phi}',$$

are non-zero, orthogonal vectors. It now follows that

$$\begin{aligned} A\boldsymbol{\theta} &= aA\boldsymbol{\phi} + bA\boldsymbol{\phi}' = \mu(a\boldsymbol{\phi} + b\boldsymbol{\phi}') = \mu\boldsymbol{\theta}, \\ A\boldsymbol{\theta}' &= cA\boldsymbol{\phi} + dA\boldsymbol{\phi}' = \mu(c\boldsymbol{\phi} + d\boldsymbol{\phi}') = \mu\boldsymbol{\theta}', \end{aligned}$$

proving that these orthogonal vectors are also eigenvectors of A with the common eigenvalue μ.

(e) Assume that we have orthonormalized our eigenvectors, $\{\,\boldsymbol{\phi}_n : 1 \leq n \leq N\,\}$. These eigenvectors form an orthonormal basis for $\mathcal{C}^N$. Any vector $\mathbf{c} \in \mathcal{C}^N$ can then be written in the form,

$$\mathbf{c} = \sum_{n=1}^{N} \boldsymbol{\phi}_n c_n, \quad \text{where } c_n \equiv \boldsymbol{\phi}_n^\dagger \mathbf{c}.$$

If I_N is the $N \times N$ identity matrix, then $\mathbf{c}$ can also be written as,

$$\mathbf{c} = I_N \mathbf{c}.$$

Subtracting the former equation from the latter we get,

$$\left(I_N - \sum_{n=1}^{N} \boldsymbol{\phi}_n \boldsymbol{\phi}_n^\dagger \right) \mathbf{c} = 0, \quad \text{for all } \mathbf{c} \in \mathcal{C}^N.$$

For this to be true it must be that,

$$I_N = \sum_{n=1}^{N} \boldsymbol{\phi}_n \boldsymbol{\phi}_n^\dagger,$$

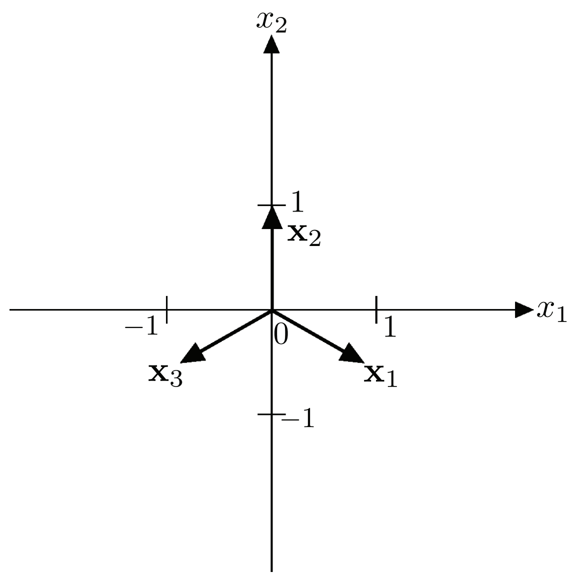

Figure 2: Labeled sketch of $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$.

as advertised. Because I_N is the identity matrix, we now have that

$$A = I_N A = \sum_{n=1}^{N} \boldsymbol{\phi}_n \boldsymbol{\phi}_n^\dagger A = \sum_{n=1}^{N} \mu_n \boldsymbol{\phi}_n \boldsymbol{\phi}_n^\dagger,$$

where the last equality uses the facts that A is Hermitian and the $\{\mu_n\}$ are real.

Problem 1.8

Here we shall introduce the idea of overcompleteness, something that will be of great importance in the quantum optics work to come.

(a) The vectors $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ are sketched in Fig. 3.1. We see that they are unit-length vectors, $\{\, \mathbf{x}_i^T \mathbf{x}_i = 1 : i = 1, 2, 3\,\}$, that are not orthogonal.

(b) With $\mathbf{e}_1^T \equiv [\ 1 \quad 0\ ]$ and $\mathbf{e}_2^T \equiv [\ 0 \quad 1\ ]$ being the standard orthonormal basis for $\mathcal{R}^2$, we know that any $\mathbf{y} \in \mathcal{R}^2$ can be written as

$$\mathbf{y} = y_1 \mathbf{e}_1 + y_2 \mathbf{e}_2 = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Thus to prove that any $\mathbf{y} \in \mathcal{R}^2$ can be expressed as a weighted sum of any two of the $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, it is sufficient to show that $\mathbf{e}_1$ and $\mathbf{e}_2$ can be written as weighted sums of any two of the $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. Figure 3.1 makes it clear that this is so, but let's write out the formulas anyway:

$$\begin{aligned}\mathbf{e}_1 &= \frac{2}{\sqrt{3}}[\mathbf{x}_1 + \mathbf{x}_2/2] = \frac{1}{\sqrt{3}}[\mathbf{x}_1 - \mathbf{x}_3] = -\frac{2}{\sqrt{3}}[\mathbf{x}_2/2 + \mathbf{x}_3]. \\ \mathbf{e}_2 &= \mathbf{x}_2 = -(\mathbf{x}_1 + \mathbf{x}_3) = \mathbf{x}_2.\end{aligned}$$

(c) This part is straight plug-and-chug. We have that,

$$\begin{aligned}\mathbf{x}_1\mathbf{x}_1^T &= \begin{bmatrix} 3/4 & -\sqrt{3}/4 \\ -\sqrt{3}/4 & 1/4 \end{bmatrix} \\ \mathbf{x}_2\mathbf{x}_2^T &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ \mathbf{x}_3\mathbf{x}_3^T &= \begin{bmatrix} 3/4 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/4 \end{bmatrix}.\end{aligned}$$

Summing these terms up yields,

$$\sum_{n=1}^{3} \mathbf{x}_n\mathbf{x}_n^T = \begin{bmatrix} 3/2 & 0 \\ 0 & 3/2 \end{bmatrix}.$$

Multiplying by 2/3 now yields the 2×2 identity matrix, I_2, as desired. Finally, for any $\mathbf{x} \in \mathcal{R}^2$ we have that

$$\mathbf{x} = I_2\mathbf{x} = \frac{2}{3}\left(\sum_{n=1}^{3} \mathbf{x}_n\mathbf{x}_n^T\right)\mathbf{x} = \frac{2}{3}\sum_{n=1}^{3} \mathbf{x}_n(\mathbf{x}_n^T\mathbf{x}).$$

6.453 Quantum Optical Communication
Fall 2016