Physics 125c
Course Notes
Density Matrix Formalism
Solutions to Problems
040520 Frank Porter

1 Exercises

1. Show that any linear operator in an n-dimensional Euclidean space may be expressed as an n-term dyad. Show that this may be extended to an infinite-dimensional Euclidean space.

   Solution: Consider operator A in n-dimensional Euclidean space, which may be expressed as a matrix in a given basis:

$$\begin{aligned} A &= \sum_{i,j=1}^{n} a_{ij}|e_i\rangle\langle e_j| \\ &= \sum_i |e_i\rangle \left(\sum_j a_{ij}\langle e_j|\right). \end{aligned} \tag{1}$$

   This is in the form of an n-term dyad $A = \sum_{i=1}^{n} |\alpha_i\rangle\langle\beta_i|$, with $|\alpha_i\rangle = |e_i\rangle$ and $\langle\beta_i| = \sum_{j=1}^{n} a_{ij}\langle e_j|$.

   An arbitrary vector in an infinite dimensional Euclidean space may be expanded in a countable basis according to:

$$|\alpha\rangle = \sum_{i=1}^{\infty} \alpha_i |e_i\rangle. \tag{2}$$

   Another way to say this is that the basis is "complete", with "completeness relation":

$$I = \sum_i |e_i\rangle\langle e_i|. \tag{3}$$

   An arbitrary linear operator can thus be defined in terms of its actions on the basis vectors:

$$A = IAI = \sum_{i,j} |e_i\rangle\langle e_i|A|e_j\rangle\langle e_j|. \tag{4}$$

   The remainder proceeds as for the finite-dimensional case.

2. Suppose we have a system with total angular momentum 1. Pick a basis corresponding to the three eigenvectors of the z-component of angular momentum, J_z, with eigenvalues $+1, 0, -1$, respectively. We are given an ensemble described by density matrix:

$$\rho = \frac{1}{4}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

(a) Is ρ a permissible density matrix? Give your reasoning. For the remainder of this problem, assume that it is permissible. Does it describe a pure or mixed state? Give your reasoning.

Solution: Clearly ρ is hermitian. It is also trace one. This is almost sufficient for ρ to be a valid density matrix. We can see this by noting that, given a hermitian matrix, we can make a transformation of basis to one in which ρ is diagonal. Such a transformation preserves the trace. In this diagonal basis, ρ is of the form:

$$\rho = a|e_1\rangle\langle e_1| + b|e_2\rangle\langle e_2| + c|e_3\rangle\langle e_3|,$$

where a, b, c are real numbers such that $a + b + c = 1$. This is clearly in the form of a density operator. Another way of arguing this is to consider the n-term dyad representation for a hermitian matrix.

However, we must also have that ρ is positive, in the sense that a, b, c cannot be negative. Otherwise, we would interpret some probabilities as negative. There are various ways to check this. For example, we can check that the expectation value of ρ with respect to any state is not negative. Thus, let an arbitrary state be: $|\psi\rangle = (\alpha, \beta, \gamma)$. Then

$$\langle\psi|\rho|\psi\rangle = 2|\alpha|^2 + |\beta|^2 + |\gamma|^2 + 2\Re(\alpha^*\beta) + 2\Re(\alpha^*\gamma). \qquad (5)$$

This quantity can never be negative, by virtue of the relation:

$$|x|^2 + |y|^2 + 2\Re(x^*y) = |x + y|^2 \geq 0. \qquad (6)$$

Therefore ρ is a valid density operator.

To determine whether ρ is a pure or mixed state, we consider:

$$\mathrm{Tr}(\rho^2) = \frac{1}{16}(6 + 2 + 2) = \frac{5}{8}.$$

This is not equal to one, so ρ is a mixed state. Alternatively, one can show explicitly that $\rho^2 \neq \rho$.

(b) Given the ensemble described by ρ, what is the average value of J_z?

Solution: We are working in a diagonal basis for J_z:

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The average value of J_z is:

$$\langle J_z \rangle = \mathrm{Tr}(\rho J_z) = \frac{1}{4}(2 + 0 - 1) = \frac{1}{4}.$$

(c) What is the spread (standard deviation) in measured values of J_z?

Answer: We'll need the average value of J_z^2 for this:

$$\langle J_z^2 \rangle = \mathrm{Tr}(\rho J_z^2) = \frac{1}{4}(2 + 0 + 1) = \frac{3}{4}.$$

Then:

$$\Delta J_z = \sqrt{\langle J_z^2 \rangle - \langle J_z \rangle^2} = \frac{\sqrt{11}}{4}.$$

3. Prove the first theorem in section **??**.

Solution: The theorem we wish to prove is:

Theorem: Let P_1, P_2 be two primitive Hermitian idempotents (*i.e.*, rays, or pure states, with $P^\dagger = P$, $P^2 = P$, and $\mathrm{Tr}P = 1$). Then:

$$1 \geq \mathrm{Tr}(P_1 P_2) \geq 0. \tag{7}$$

If $\mathrm{Tr}(P_1 P_2) = 1$, then $P_2 = P_1$. If $\mathrm{Tr}(P_1 P_2) = 0$, then $P_1 P_2 = 0$ (vectors in ray 1 are orthogonal to vectors in ray 2).

First, suppose $P_1 = P_2$. Then $\mathrm{Tr}(P_1 P_2) = \mathrm{Tr}(P_1^2) = \mathrm{Tr}(P_1) = 1$. If $P_1 P_2 = 0$, then $\mathrm{Tr}(P_1 P_2) = \mathrm{Tr}(0) = 0$.

More generally, expand P_1 and P_2 with respect to an orthonormal basis $\{|e_i\rangle\}$:

$$P_1 = \sum_{i,j} a_{ij} |e_i\rangle\langle e_j| \quad (8)$$

$$P_2 = \sum_{i,j} b_{ij} |e_i\rangle\langle e_j| \quad (9)$$

$$P_1 P_2 = \sum_{i,j,k} a_{ik} b_{kj} |e_i\rangle\langle e_j|. \quad (10)$$

We know from the discussion on pages 11,12 in the Density Matrix note, that we can work in a basis in which $a_{ij} = \delta_{ij}\delta i1$. In this basis,

$$P_1 P_2 = |e_1\rangle \sum_i b_{1i} \langle e_i|. \quad (11)$$

The trace, which is invariant under choice of basis, is

$$\mathrm{Tr}(P_1 P_2) = b_{11} \quad (12)$$

We are almost there, but we need to show that $b_{11} > 0$, if $P_1 P_2 \neq 0$. A simple way to see this is to notice that P_2 is the outer product of a vector with itself, hence $b_{11} \geq 0$, with $b_{11} = 0$ if and only if $P_1 P_2 = 0$ (since $b_{1i} = b_{i1} = 0$ for all i if $b_{11} = 0$). Finally, $b_{11} < 1$ if $P_1 \neq P_2$.

4. Prove the von Neumann mixing theorem.

Solution: The mixing theorem states that, given two distinct ensembles $\rho_1 \neq \rho_2$, a number $0 < \theta < 1$, and a mixed ensemble $\rho = \theta\rho_1 + (1-\theta)\rho_2$, then

$$s(\rho) > \theta s(\rho_1) + (1-\theta) s(\rho_2). \quad (13)$$

Let us begin by proving the following:

Lemma: Let

$$x = \sum_{i=1}^{n} \lambda_i x_i, \quad (14)$$

where $1 > x_i > 0$, $0 < \lambda_i < 1$ for all i, and $\sum_{i=1}^{n} \lambda_i = 1$. Then

$$-x \ln x > -\sum_{i=1}^{n} \lambda_i x_i \ln x_i. \quad (15)$$

Proof: This follows because $-x \ln x$ is a **concave** function of x. Its second derivative is

$$\frac{d^2}{dx^2}(-x \ln x) = \frac{d}{dx}(-\ln x - 1) = -1/x. \tag{16}$$

For $1 > x > 0$ this is always negative; $-x \ln x$ is a concave function for $1 > x > 0$. Hence, any point on a straight line between two points on the curve of this function lies below the curve. The theorem is for a linear combination of n points on the curve of $-x \ln x$. Here, x is a weighted average of points x_i. The function $-x \ln x$ evaluated at this weighted average point is to be compared with the weighted average of the values of the function at the n points $x_1, x_2, \ldots, x_n$. Again, the function evaluated at the linear combination is a point on the curve, and the weighted average of the function over the n points must lie below that point on the curve. The region of possible values of the weighted average of the function is the polygon joining neighboring points on the curve, and the first and last points. See Fig. 1.

Now we must see how our present problem can be put in the form where this lemma may be applied. Consider the spectral decompositions of ρ_1, ρ_2:

$$\begin{aligned} \rho_1 &= \sum_i a_i P_i = \sum_i a_i |e_i\rangle\langle e_i| & (17) \\ \rho_2 &= \sum_i b_i Q_i = \sum_i b_i |f_i\rangle\langle f_i|, & (18) \end{aligned}$$

where the decompositions have been "padded" with complete sets of one-dimensional projections. That is, some of the a_i's and b_i's may be zero. The idea is that the sets $\{|e_i\rangle\}$ and $\{|f_i\rangle\}$ form complete orthonormal bases. Note that we cannot have $P_i = Q_i$ in general.

Then we have:

$$\begin{aligned} \rho &= \theta\rho_1 + (1-\theta)\rho_2 & (19) \\ &= \sum_i \left[\theta a_i |e_i\rangle\langle e_i| + (1-\theta) b_i |f_i\rangle\langle f_i|\right] & (20) \\ &= \sum_i c_i |g_i\rangle\langle g_i|, & (21) \end{aligned}$$

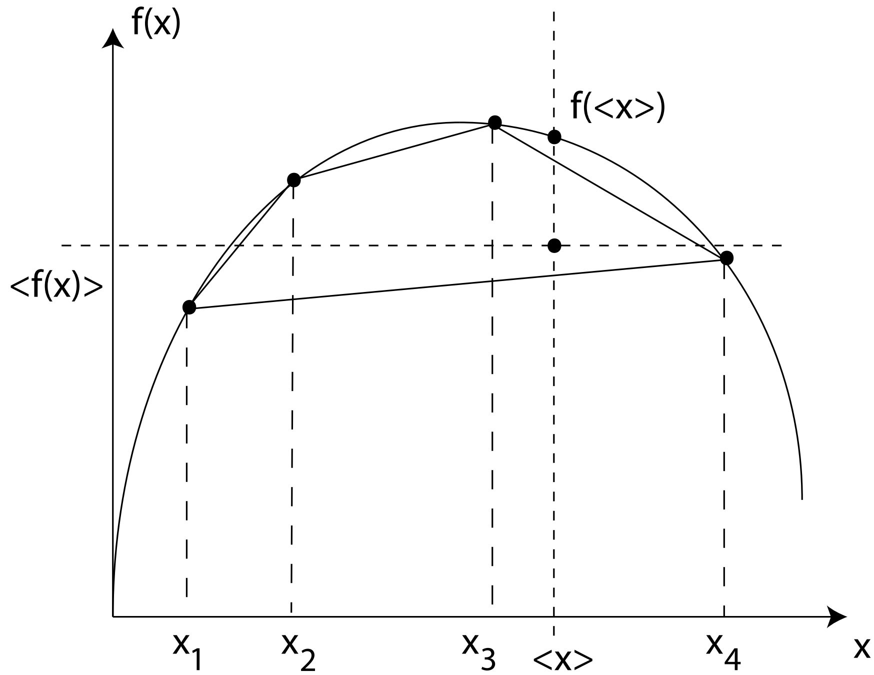

Figure 1: Illustration showing that the weighted average of a concave function is smaller than the function evaluated at the weighted average point. The allowed values of the ordered pairs $\langle x\rangle, \langle f(x)\rangle$ lie in the polygon.

where we have defined another complete orthonormal basis, $\{|g_i\rangle\}$, corresponding to a basis in which ρ is diagonal.

We may expand the $\{|e_i\rangle\}$ and $\{|f_i\rangle\}$ bases in terms of the $\{|g_i\rangle\}$ basis. For example, let

$$|e_i\rangle = \sum_j A_{ij}|g_j\rangle, \tag{22}$$

where $A = \{A_{ij}\}$ is a unitary matrix. The inverse transformation is

$$|g_i\rangle = \sum_j A^*_{ji}|e_j\rangle. \tag{23}$$

Also,

$$\langle e_i| = \left[\sum_j A_{ij}|g_j\rangle\right]^\dagger = \sum_j A^*_{ij}\langle g_j|, \tag{24}$$

and hence,

$$|e_i\rangle\langle e_i| = \sum_j \sum_k A_{ij}|g_j\rangle\langle g_k|A^*_{ik}. \tag{25}$$

Similarly, we define matrix B such that

$$|f_i\rangle\langle f_i| = \sum_j \sum_k B_{ij}|g_j\rangle\langle g_k|B^*_{ik}. \tag{26}$$

Substituting Eqns. 25 and 26 into Eqn: 20:

$$\rho = \sum_i \left[\theta a_i \sum_{j,k} A^*_{ik}A_{ij} + (1-\theta)b_i \sum_{j,k} B^*_{ik}B_{ij}\right] |g_j\rangle\langle g_k|. \tag{27}$$

Thus, the numbers c_ℓ are:

$$\begin{aligned} c_\ell &= \langle g_\ell|\rho|g_\ell\rangle \\ &= \sum_i \left[\theta a_i \sum_{j,k} A^*_{ik}A_{ij} + (1-\theta)b_i \sum_{j,k} B^*_{ik}B_{ij}\right] \delta_{\ell j}\delta_{\ell k} \qquad (28) \\ &= \sum_i \left[\theta|A_{i\ell}|^2 a_i + (1-\theta)|B_{i\ell}|^2 b_i\right]. \qquad (29) \end{aligned}$$

The entropy for density matrix ρ is:

$$\begin{aligned} s(\rho) &= -\sum_i c_i \ln c_i \qquad (30) \\ &= -\sum_i \left\{\sum_j \left[\theta|A_{ji}|^2 a_j + (1-\theta)|B_{ji}|^2 b_j\right]\right\} \ln \left\{\sum_j \left[\theta|A_{ji}|^2 a_j + (1-\theta)|B_{ji}|^2 b_j\right]\right\}. \end{aligned}$$

Note that c_i is of the form

$$c_i = \sum_j (\lambda^{(a)}_{ij} a_j + \lambda^{(b)}_{ij} b_j), \tag{31}$$

where

$$\begin{aligned} \lambda^{(a)}_{ij} &\equiv \theta|A_{ji}|^2 \qquad (32) \\ \lambda^{(b)}_{ij} &\equiv (1-\theta)|B_{ji}|^2. \qquad (33) \end{aligned}$$

Furthermore,
$$\sum_j (\lambda_{ij}^{(a)} + \lambda_{ij}^{(b)}) = 1. \tag{34}$$

Thus, according to the lemma (some of the c_i's might be zero; there is an equality, $0 = 0$ in such cases),

$$-c_i \ln c_i > -\sum_j \left[\lambda_{ij}^{(a)} a_j \ln a_j + \lambda_{ij}^{(b)} b_j \ln b_j\right]. \tag{35}$$

Finally, we sum the above inequality over i:

$$\begin{aligned} s(\rho) &= -\sum_i c_i \ln c_i \\ &> -\sum_i \sum_j \left[\lambda_{ij}^{(a)} a_j \ln a_j + \lambda_{ij}^{(b)} b_j \ln b_j\right] \\ &> -\sum_j \left[\theta a_j \ln a_j + (1-\theta) b_j \ln b_j\right] \qquad (36) \\ &> \theta s(\rho_1) + (1-\theta) s(\rho_2) \qquad (37) \end{aligned}$$

This completes the proof.

5. Show that an arbitrary linear operator on a product space $H = H_1 \otimes H_2$ may be expressed as a linear combination of operators of the form $Z = X \otimes Y$.

   Solution: We are given an arbitrary linear operator A on $H = H_1 \otimes H_2$. We wish to show that there exists a decomposition of the form:

   $$A = \sum_i A_i Z_i = \sum_i A_i X_i \otimes Y_i, \tag{38}$$

   where X_i are operators on H_1 and Y_i are operators on H_2.

   Let $\{f_i : i = 1, 2, \ldots\}$ be an orthonormal basis in H_1 and $\{g_i : i = 1, 2, \ldots\}$ be an orthonormal basis in H_2. Then we may obtain an orhtonormal basis for H composed of vectors of the form:

   $$e_{ij} = f_i \otimes g_j, \quad i = 1, 2, \ldots; \; j = 1, 2, \ldots. \tag{39}$$

   It is readily checked that $\{e_{ij}\}$ is, in fact, an orthonormal basis for H.

Expand A with respect to basis $\{e_{ij}\}$:

$$\begin{aligned} A &= \sum_{i,j,m,n} A_{ij,mn} |e_{ij}\rangle\langle e_{mn}| && (40)\\ &= \sum_{i,j,m,n} A_{ij,mn} |f_i\rangle \otimes |g_j\rangle\langle f_m| \otimes \langle g_n| && (41)\\ &= \sum_{i,j,m,n} A_{ij,mn} |f_i\rangle\langle f_m| \otimes |g_j\rangle\langle g_n| && (42)\\ &= \sum_k A_k X_k \otimes Y_k, && (43)\end{aligned}$$

where k is a relabeling for i, j, m, n.

The only step above which requires further comment is setting:

$$|f_i\rangle \otimes |g_j\rangle\langle f_m| \otimes \langle g_n| = |f_i\rangle\langle f_m| \otimes |g_j\rangle\langle g_n|. \tag{44}$$

One way to check this is as follows. Pick our bases to be in the form:

$$\begin{aligned} (f_i)_k &= \delta_{ik} && (45)\\ (g_j)_\ell &= \delta_{j\ell}. && (46)\end{aligned}$$

Then

$$(|f_i\rangle \otimes |g_j\rangle\langle f_m| \otimes \langle g_n|)_{k\ell,pq} = \delta_{ik}\delta_{j\ell}\delta_{mp}\delta_{nq}. \tag{47}$$

and

$$(|f_i\rangle\langle f_m| \otimes |g_j\rangle\langle g_n|)_{k\ell,pq} = \delta_{ik}\delta_{j\ell}\delta_{mp}\delta_{nq}. \tag{48}$$

6. Let us try to improve our understanding of the discussions on the density matrix formalism, and the connections with "information" or "entropy" that we have made. Thus, we consider a simple "two-state" system. Let ρ be any general density matrix operating on the two-dimensional Hilbert space of this system.

   (a) Calculate the entropy, $s = -\mathrm{Tr}(\rho \ln \rho)$ corresponding to this density matrix. Express your result in terms of a single real parameter. Make sure the interpretation of this parameter is clear, as well as its range.

   Solution: Density matrix ρ is Hermitian, hence diagonal in some basis. Work in such a basis. In this basis, ρ has the form:

$$\rho = \begin{pmatrix} \theta & 0 \\ 0 & 1-\theta \end{pmatrix}, \tag{49}$$

where $0 \leq \theta \leq 1$ is the probability that the system is in state 1. We have a pure state if and only if either $\theta = 1$ or $\theta = 0$.

The entropy is

$$s = -\theta \ln \theta - (1-\theta)\ln(1-\theta). \tag{50}$$

(b) Make a graph of the entropy as a function of the parameter. What is the entropy for a pure state? Interpret your graph in terms of knowledge about a system taken from an ensemble with density matrix ρ.

Solution:

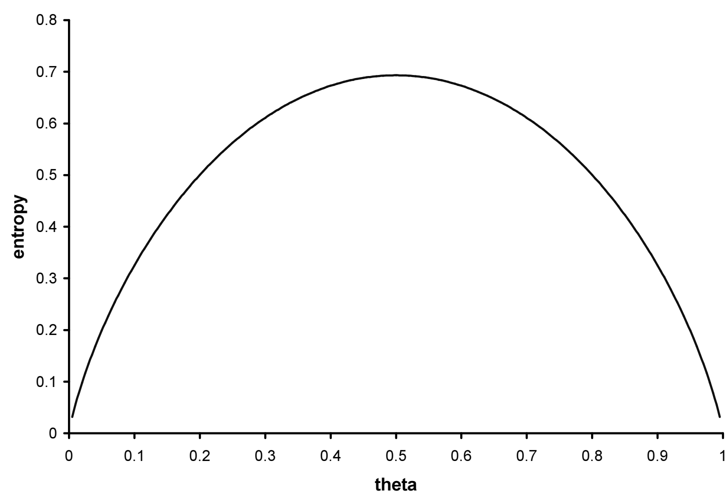

Figure 2: The entropy as a function of θ.

The entropy for a pure state, with $\theta = 1$ or $\theta = 0$, is zero. The entropy increases as the state becomes "less pure", reaching maximum when the probability of being in either state is 1/2, reflecting minimal "knowledge" about the state.

(c) Consider a system with ensemble ρ a mixture of two ensembles ρ_1, ρ_2:

$$\rho = \theta\rho_1 + (1-\theta)\rho_2, \qquad 0 \leq \theta \leq 1 \tag{51}$$

As an example, suppose

$$\rho_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad \rho_2 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{52}$$

in some basis. Prove that VonNeuman's mixing theorem holds for this example:

$$s(\rho) \geq \theta s(\rho_1) + (1-\theta)s(\rho_2), \tag{53}$$

with equality iff $\theta = 0$ or $\theta = 1$.

Solution: The entropy of ensemble 1 is:

$$s(\rho_1) = -\text{Tr}\rho_1 \ln \rho_1 = -\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\ln\frac{1}{2} = \ln 2 = 0.6931 \tag{54}$$

It may be noticed that $\rho_2^2 = \rho_2$, hence ensemble 2 is a pure state, with entropy $s(\rho_2) = 0$. Next, we need the entropy of the combined ensemble:

$$\rho = \theta\rho_1 + (1-\theta)\rho_2 = \frac{1}{2}\begin{pmatrix} 1 & 1-\theta \\ 1-\theta & 1 \end{pmatrix}. \tag{55}$$

To compute the entropy, it is convenient to determine the eigenvalues; they are $1-\theta/2$ and $\theta/2$. Note that they are in the range from zero to one, as they must be. The entropy is

$$s(\rho) = -\left(1-\frac{\theta}{2}\right)\ln\left(1-\frac{\theta}{2}\right) - \left(-\frac{\theta}{2}\right)\ln\frac{\theta}{2}. \tag{56}$$

We must compare $s(\rho)$ with

$$\theta s(\rho_1) + (1-\theta)s(\rho_2) = \theta \ln 2. \tag{57}$$

It is readily checked that equality holds for $\theta = 1$ or $\theta = 0$. For the case $0 < \theta < 1$, take the difference of the two expressions:

$$\begin{aligned} s(\rho) - [\theta s(\rho_1) + (1-\theta)s(\rho_2)] &= -\left(1-\frac{\theta}{2}\right)\ln\left(1-\frac{\theta}{2}\right) - \left(-\frac{\theta}{2}\right)\ln\frac{\theta}{2} - \theta\ln 2 \\ &= -\ln\left[\left(1-\frac{\theta}{2}\right)^{1-\theta/2}\left(\frac{\theta}{2}\right)^{\theta/2} 2^{\theta}\right]. \end{aligned} \tag{58}$$

This must be larger than zero if the mixing theorem is correct. This is equivalent to asking whether

$$\left(1-\frac{\theta}{2}\right)^{1-\theta/2}\left(\frac{\theta}{2}\right)^{\theta/2} 2^{\theta} \tag{59}$$

is less than 1. This expression may be rewritten as

$$\left(1-\frac{\theta}{2}\right)^{1-\theta/2}(2\theta)^{\theta/2}. \tag{60}$$

It must be less than one. To check, let's find its maximum value, by setting its derivative with respect to θ equal to 0:

$$\begin{aligned} 0 &= \frac{d}{d\theta}\left(1-\frac{\theta}{2}\right)^{1-\theta/2}(2\theta)^{\theta/2} \\ &= \frac{d}{d\theta}\exp\left[\left(1-\frac{\theta}{2}\right)\ln\left(1-\theta/2\right)+(\theta/2)\ln(2\theta)\right] \\ &= -\frac{1}{2}\ln(1-\theta/2)+\frac{1}{2}\ln(2\theta)-\frac{1}{2}+\frac{2}{4} \\ &= \ln(2\theta)-\ln(1-\theta/2). \end{aligned} \tag{61}$$

Thus, the maximum occurs at $\theta = 2/5$. At this value of θ, $s(\rho) = 0.500$, and $\theta s(\rho_1) + (1-\theta)s(\rho_2) = (2/5)\ln 2 = 0.277$. The theorem holds.

7. Consider an N-dimensional Hilbert space. We define the real vector space, $\mathcal{O}$ of Hermitian operators on this Hilbert space. We define a scalar product on this vector space according to:

$$(x, y) = \mathrm{Tr}(xy), \quad \forall x, y \in \mathcal{O}. \tag{62}$$

Consider a basis $\{B\}$ of orthonormal operators in $\mathcal{O}$. The set of density operators is a subset of this vector space, and we may expand an arbitrary density matrix as:

$$\rho = \sum_i B_i \mathrm{Tr}(B_i\rho) = \sum_i B_i \langle B_i\rangle_\rho. \tag{63}$$

By measuring the average values for the basis operators, we can thus determine the expansion coefficients for ρ.

(a) How many such measurements are required to completely determine ρ?

Solution: The question is, how many independent basis operators are there in $\mathcal{O}$? An arbitrary $N \times N$ complex matrix is described by $2N^2$ real parameters. The requirement of Hermiticity provides the independent constraint equations:

$$\begin{aligned} \Re(H_{ij}) &= \Re(H_{ji}), \quad i < j & (64) \\ \Im(H_{ij}) &= -\Im(H_{ji}), \quad i \leq j. & (65) \end{aligned}$$

This is $N + 2[N(N-1)/2] = N^2$ equations. Thus, $\mathcal{O}$ is an N^2-dimensional vector space. But to completely determine the density matrix, we have one further constraint, that $\mathrm{Tr}\rho = 1$. Thus, it takes $N^2 - 1$ measurements to completely determine ρ.

(b) If ρ is known to be a pure state, how many measurements are required?

Solution: We note that a complex vector in N dimensions is completely specified by $2N$ real parameters. However, one parameter is an arbitrary phase, and another parameter is eaten by the normalization constraint. Thus, it takes $2(N-1)$ parameters to completely specify a pure state.

If ρ is a pure state, then $\rho^2 = \rho$. How many additional constraints over the result in part (a) does this imply? Let's try to get a more intuitive understanding by attacking this issue from a slightly different perspective. Ask, instead, how many parameters it takes to build an arbitrary density matrix as a mixture of pure states. Our response will be to add pure states into the mixture one at a time, counting parameters as we go, until we cannot add any more.

It takes $2(N-1)$ parameters to define the first pure state in our mixture. The second pure state must be a distinct state. That is, it must be drawn from an $N-1$-dimensional subspace. Thus the second pure state requires $2(N-2)$ parameters to define. There will also be another parameter required to specify the relative probabilities of the first and second state, but we'll count up these probablilities later. The third pure state requires $2(N-3)$ parameters, and so forth, stopping at $2 \cdot 1$ paramter for the $(N-1)$st pure state. Thus, it takes

$$2 \sum_{k=1}^{N-1} k = N(N-1) \tag{66}$$

parameters to define all the pure states in the arbitrary mixture. There can be a total of N pure states making up a mixture (the Nth one required no additional parameters in the count we just made). It takes $N-1$ parameters to specify the relative probabiliteis of these N components in the mixture. Thus, the total number of parameters required is:

$$N(N-1)+(N-1)=N^2-1. \tag{67}$$

Notice that this is just the result we obtained in part (a).

8. Two scientists (they happen to be twins, named "Oivil" and "Livio", but never mind...) decide to do the following experiment: They set up a light source, which emits two photons at a time, back-to-back in the laboratory frame. The ensemble is given by:

$$\rho=\frac{1}{2}(|LL\rangle\langle LL|+|RR\rangle\langle RR|), \tag{68}$$

where "L" refers to left-handed polarization, and "R" refers to right-handed polarization. Thus, $|LR\rangle$ would refer to a state in which photon number 1 (defined as the photon which is aimed at scientist Oivil, say) is left-handed, and photon number 2 (the photon aimed at scientist Livio) is right-handed.

These scientists (one of whom is of a diabolical bent) decide to play a game with Nature: Oivil (of course) stays in the lab, while Livio treks to a point a light-year away. The light source is turned on and emits two photons, one directed toward each scientist. Oivil soon measures the polarization of his photon; it is left-handed. He quickly makes a note that his brother is going to see a left-handed photon, sometime after next Christmas.

Christmas has come and gone, and finally Livio sees his photon, and measures its polarization. He sends a message back to his brother Oivil, who learns in yet another year what he knew all along: Livio's photon was left-handed.

Oivil then has a sneaky idea. He secretly changes the apparatus, without telling his forlorn brother. Now the ensemble is:

$$\rho=\frac{1}{2}(|LL\rangle+|RR\rangle)(\langle LL|+\langle RR|). \tag{69}$$

He causes another pair of photons to be emitted with this new apparatus, and repeats the experiment. The result is identical to the first experiment.

(a) Was Oivil just lucky, or will he get the right answer every time, for each apparatus? Demonstrate your answer explicitly, in the density matrix formalism.

Solution: Yup, he'll get it right, every time, in either case. Let's first define a basis so that we can see how it all works with explicit matrices:

$$|LL\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |LR\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |RL\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |RR\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{70}$$

In this basis the density matrix for the first apparatus is:

$$\begin{aligned} \rho &= \frac{1}{2}(|LL\rangle\langle LL| + |RR\rangle\langle RR|) \\ &= \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{71}$$

Since $Tr(\rho^2) = 1/2$, we know that this is a mixed state.

Now, Oivil observes that his photon is left-handed. His left-handed projection operator is

$$P_L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{72}$$

so once he has made his measurement, the state has "collapsed" to:

$$P_L\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{73}$$

This corresponds to a pure $|LL\rangle$ state, hence Livio will observe left-handed polarization.

For the second apparatus, the density matrix is

$$\begin{aligned} \rho &= \frac{1}{2}(|LL\rangle + |RR\rangle)(\langle LL| + \langle RR|) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{74}$$

Since $Tr(\rho^2) = 1$, we know that this is a pure state. Applying the left-handed projection for Oivil's photon, we again obtain:

$$P_L\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{75}$$

Again, Livio will observe left-handed polarization.

(b) What is the probability that Livio will observe a left-handed photon, or a right-handed photon, for each apparatus? Is there a problem with causality here? How can Oivil know what Livio is going to see, long before he sees it? Discuss! Feel free to modify the experiment to illustrate any points you wish to make.

Solution: Livio's left-handed projection operator is

$$P_L(\text{Livio}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{76}$$

The probability that Livio will observe a left-handed photon for the first apparatus is:

$$\langle P_L(\text{Livio})\rangle = \text{Tr}\left[P_L(\text{Livio})\rho\right] = 1/2. \tag{77}$$

The same result is obtained for the second apparatus.

Here is my take on the philosophical issue (beware!):

If causality means propagation of information faster than the speed of light, then the answer is "no", causality is not violated. Oivil has not propagated any information to Livio at superluminal velocities. Livio made his own observation on the state of the system. Notice that the statistics of Livio's observations are unaltered; independent of what Oivil does, he will still see left-handed photons 50% of the time. If this were not the case, then there *would* be a problem, since Oivil could exploit this to propagate a message to Livio long after the photons are emitted.

However, people widely interpret (and write flashy headlines) this sort of effect as a kind of "action at a distance": By measuring the state of his photon, Oivil instantly "kicks" Livio's far off photon into a particular state (without being usable for the propagation of information, since Oivil can't tell Livio about it any faster than the speed of light). Note that this philosophical dilemma is not silly: The wave function for Livio's photon has both left- and right-handed components; how could a measurement of Oivil's photon "pick" which component Livio will see? Because of this, quantum mechanics is often labelled "non-local".

On the other hand, this philosophical perspective may be avoided (ignored): It may be suggested that it doesn't make sense to talk this way about the "wave function" of Livio's photon, since the specification of the wave function involves also Oivil's photon. Oivil is merely making a measurement of the state of the two-photon system, by observing the polarization of one of the photons, and knowing the coherence of the system. He doesn't need to make two measurements to know both polarizations, they are completely correlated. Nothing is causing anything else to happen at faster than light speed. We might take the (deterministic?) point of view that it was already determined at production which polarization Livio would see for a particular photon – we just don't know what it will be unless Oivil makes his measurement. There appears to be no way of falsifying this point of view, as stated. However, taking this point of view leads to to the further philosophical question of how the pre-determined information is encoded – is the photon propagating towards Livio somehow carrying the information that it is going to be measured as left-handed? This conclusion seems hard to avoid. It leads to the notion of "hidden variables", and there are theories of this sort, which are testable.

We know that our quantum mechanical foundations are compatible with special relativity, hence with the notion of causality that implies.

As Feynman remarked several years ago in a seminar I arranged concerning EPR, the substantive question to be asking is, "Do the predictions of quantum mechanics agree with experiment?". So far the answer is a resounding "yes". Indeed, we often rely heavily on this quantum coherence in carrying out other research activities. Current experiments to measure CP violation in B^0 decays crucially depend on it, for example.

9. Let us consider the application of the density matrix formalism to the problem of a spin-1/2 particle (such as an electron) in a static external magnetic field. In general, a particle with spin may carry a magnetic moment, oriented along the spin direction (by symmetry). For spin-1/2, we have that the magnetic moment (operator) is thus of the form:

$$\boldsymbol{\mu} = \frac{1}{2}\gamma\boldsymbol{\sigma}, \tag{78}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices, the $\frac{1}{2}$ is by convention, and γ is a constant, giving the strength of the moment, called the gyromagnetic ratio. The term in the Hamiltonian for such a magnetic moment in an external magnetic field, $\boldsymbol{B}$ is just:

$$H = -\boldsymbol{\mu}\cdot\boldsymbol{B}. \tag{79}$$

Our spin-1/2 particle may have some spin-orientation, or "polarization vector", given by:

$$\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle. \tag{80}$$

Drawing from our classical intuition, we might expect that in the external magnetic field the polarization vector will exhibit a precession about the field direction. Let us investigate this.

Recall that the expectation value of an operator may be computed from the density matrix according to:

$$\langle A\rangle = \mathrm{Tr}(\rho A). \tag{81}$$

Furthermore, recall that the time evolution of the density matrix is given by:

$$i\frac{\partial\rho}{\partial t} = [H(t),\rho(t)]. \tag{82}$$

What is the time evolution, $d\boldsymbol{P}/dt$, of the polarization vector? Express your answer as simply as you can (more credit will be given for right answers that are more physically transparent than for right answers which are not). Note that we make no assumption concerning the purity of the state.

Solution: Let us consider the ith-component of the polarization:

$$\begin{aligned}
i\frac{dP_i}{dt} &= i\frac{d\langle\sigma_i\rangle}{dt} && (83)\\
&= i\frac{\partial}{\partial t}\mathrm{Tr}(\rho\sigma_i) && (84)\\
&= i\mathrm{Tr}\left(\frac{\partial\rho}{\partial t}\sigma_i\right) && (85)\\
&= \mathrm{Tr}\left([H,\rho]\sigma_i\right) && (86)\\
&= \mathrm{Tr}\left([\sigma_i,H]\rho\right) && (87)\\
&= -\frac{1}{2}\gamma\sum_{j=1}^{3}B_j\mathrm{Tr}\left([\sigma_i,\sigma_j]\rho\right). && (88)
\end{aligned}$$

To proceed further, we need the density matrix for a state with polarization $\boldsymbol{P}$. Since ρ is hermitian, it must be of the form:

$$\rho = a(1+\boldsymbol{b}\cdot\boldsymbol{\sigma}). \tag{89}$$

But its trace must be one, so $a = 1/2$. Finally, to get the right polarization vector, we must have $\boldsymbol{b}$=$\boldsymbol{P}$.

Thus, we have

$$i\frac{dP_i}{dt} = -\frac{1}{4}\gamma\sum_{j=1}^{3}B_j\left\{\mathrm{Tr}[\sigma_i,\sigma_j]+\sum_{k=1}^{3}P_k\mathrm{Tr}\left([\sigma_i,\sigma_j]\sigma_k\right)\right\}. \tag{90}$$

Now $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$, which is traceless. Further, $\mathrm{Tr}\left([\sigma_i, \sigma_j]\sigma_k\right) = 4i\epsilon_{ijk}$. This gives the result:

$$\frac{dP_i}{dt} = -\gamma \sum_{j=1}^{3}\sum_{k=1}^{3} \epsilon_{ijk} B_j P_k. \tag{91}$$

This may be re-expressed in the vector form:

$$\frac{d\boldsymbol{P}}{dt} = \gamma \boldsymbol{P} \times \boldsymbol{B}. \tag{92}$$

10. Let us consider a system of N spin-1/2 particles (see the previous problem) per unit volume in thermal equilibrium, in our external magnetic field $\boldsymbol{B}$. Recall that the canonical distribution is:

$$\rho = \frac{e^{-H/T}}{Z}, \tag{93}$$

with partition function:

$$Z = \mathrm{Tr}\left(e^{-H/T}\right). \tag{94}$$

Such a system of particles will tend to orient along the magnetic field, resulting in a bulk magnetization (having units of magnetic moment per unit volume), $\boldsymbol{M}$.

(a) Give an expression for this magnetization (don't work too hard to evaluate).

Solution: Let us orient our coordinate system so that the z-axis is along the magnetic field direction. Then $M_x = 0$, $M_y = 0$, and:

$$M_z = N\frac{1}{2}\gamma\langle\sigma_z\rangle \tag{95}$$

$$= N\gamma\frac{1}{2Z}\mathrm{Tr}\left[e^{-H/T}\sigma_z\right], \tag{96}$$

where $H = -\gamma B_z \sigma_z/2$.

(b) What is the magnetization in the high-temperature limit, to lowest non-trivial order (this I want you to evaluate as completely as you can!)?

Solution: In the high temperature limit, we'll discard terms of order higher than $1/T$ in the expansion of the exponential: $e^{-H/T} \approx 1 - H/T = 1 + \gamma B_z \sigma_z / 2T$. Thus,

$$\begin{aligned} M_z &= N\gamma \frac{1}{2Z} \mathrm{Tr}\left[(1 + \gamma B_z \sigma_z/2T)\sigma_z\right] && (97) \\ &= N\gamma^2 B_z \frac{1}{2ZT}. && (98) \end{aligned}$$

Furthermore,

$$\begin{aligned} Z &= \mathrm{Tr} e^{-H/T} && (99) \\ &= 2 + O(1/T^2). && (100) \end{aligned}$$

And we have the result:

$$M_z = N\gamma^2 B_z / 4T. \qquad (101)$$

This is referred to as the "Curie Law" (for magnetization of a system of spin-1/2 particles).