1 Problems: Quantum mechanics essentials

1.1. Show that if a wavefunction $\psi(x,t)$ satisfies the one-dimensional Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\hbar^2\frac{\partial^2\psi}{\partial x^2} + V(x)\psi$$

for a real potential $V(x)$, and assuming that ψ and $\frac{\partial\psi}{\partial x}$ vanish as $x \to \pm\infty$, then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} |\psi(x,t)|^2 \,\mathrm{d}x = 0.$$

What does this equation mean?

Solution:

First note that taking the complex conjugate of the Schrödinger equation gives

$$i\hbar\frac{\partial\psi^*}{\partial t} = \hbar^2\frac{\partial^2\psi^*}{\partial x^2} - V(x)\psi^*.$$

The asymptotic conditions on ψ allow us to integrate by parts without any boundary terms in the following:

$$\begin{aligned} i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} |\psi(x,t)|^2 \,\mathrm{d}x &= \int_{-\infty}^{\infty}\left[\left(\hbar^2\frac{\partial^2\psi^*}{\partial x^2} - V(x)\psi^*\right)\psi + \psi^*\left(-\hbar^2\frac{\partial^2\psi}{\partial x^2} + V(x)\psi\right)\right]\mathrm{d}x \\ &= \hbar^2\int_{-\infty}^{\infty}\left[\frac{\partial^2\psi^*}{\partial x^2}\psi - \psi^*\frac{\partial^2\psi}{\partial x^2}\right]\mathrm{d}x \\ &= \hbar^2\int_{-\infty}^{\infty}\left[-\frac{\partial\psi^*}{\partial x}\frac{\partial\psi}{\partial x} + \frac{\partial\psi^*}{\partial x}\frac{\partial\psi}{\partial x}\right]\mathrm{d}x = 0 \end{aligned}$$

The physical content of this equation is that the total probability of finding a particle somewhere on the real line is constant in time (no probability "leaks out").

1.2. Suppose $|\alpha\rangle$ and $|\beta\rangle$ are eigenstates of a self-adjoint operator $\hat{A}$, with eigenvalues α and β respectively. Using the definition of the adjoint of $\hat{A}$ and inner products such as $\left\langle\alpha\middle|\hat{A}\middle|\alpha\right\rangle$ and $\left\langle\alpha\middle|\hat{A}\middle|\beta\right\rangle$, show that:

(a) $\alpha \in \mathbb{R}$ (hence all eigenvalues of a self-adjoint operator are real.)

(b) If $\alpha \neq \beta$ then $\langle\alpha|\beta\rangle = 0$ (hence eigenstates of a self-adjoint operator with *different* eigenvalues are orthogonal.)

Solution:

(a) We have $\hat{A}|\alpha\rangle = \alpha|\alpha\rangle$ so $\left\langle\alpha\middle|\hat{A}\middle|\alpha\right\rangle = \alpha\langle\alpha|\alpha\rangle$. But taking the complex conjugate we have:

$$\alpha^*\langle\alpha|\alpha\rangle = \left\langle\alpha\middle|\hat{A}\middle|\alpha\right\rangle^* = \left\langle\alpha\middle|\hat{A}^\dagger\middle|\alpha\right\rangle = \left\langle\alpha\middle|\hat{A}\middle|\alpha\right\rangle = \alpha\langle\alpha|\alpha\rangle\,.$$

So, unless $|\alpha\rangle = 0$ in which case it is not really an eigenstate of $\hat{A}$, we must have $\alpha^* = \alpha$.

(b) First calculate

$$\left\langle\alpha\middle|\hat{A}\middle|\beta\right\rangle = \beta\langle\alpha|\beta\rangle$$

and then (noting from the previous part β must be real) take the complex conjugate to find

$$\beta \langle \beta | \alpha \rangle = \left\langle \alpha \middle| \hat{A} \middle| \beta \right\rangle^* = \left\langle \beta \middle| \hat{A}^\dagger \middle| \alpha \right\rangle = \left\langle \beta \middle| \hat{A} \middle| \alpha \right\rangle = \alpha \langle \beta | \alpha \rangle .$$

So, if $\alpha \neq \beta$ the only way this can be true is if $\langle \beta | \alpha \rangle = 0$, as so also its complex conjugate $\langle \alpha | \beta \rangle = 0$.

1.3. Suppose a self-adjoint operator $\hat{A}$ has normalised eigenstates $|1\rangle$, $|2\rangle$ and $|3\rangle$ with eigenvalues 1, 2, 3 respectively. **PROBLEMS CLASS 1**

(a) What is the probability of measuring $A = 2$ if the system is described by the state:

(i) $|\psi\rangle = \frac{1}{2}|1\rangle + \frac{1}{4}|2\rangle + \frac{1}{4}|3\rangle$?

(ii) $|\phi\rangle = \frac{3}{4}|1\rangle - \frac{1}{2}|2\rangle + \frac{1}{4}|3\rangle$?

(b) Calculate $\langle A \rangle$ for the state $|\psi\rangle$ and for the state $|\phi\rangle$.

Solution:

(a) Note from the previous question we can deduce that the three eigenstates are orthogonal. We then notice that the states $|\psi\rangle$ and $|\phi\rangle$ are not normalised, so first find the norms:

$$\begin{aligned} || \, |\psi\rangle \, ||^2 &= \langle \psi | \psi \rangle = \frac{1}{4} + \frac{1}{16} + \frac{1}{16} = \frac{3}{8} \\ || \, |\phi\rangle \, ||^2 &= \langle \phi | \phi \rangle = \frac{9}{16} + \frac{1}{4} + \frac{1}{16} = \frac{7}{8} \end{aligned}$$

Then the probability of measuring $A = 2$ is given by the magnitude squared of the coefficient of $|2\rangle$ divided by the norm squared. I.e.

(i) For $|\psi\rangle$ the probability is $\frac{8}{3} \times |\frac{1}{4}|^2 = \frac{1}{6}$.

(ii) For $|\phi\rangle$ the probability is $\frac{8}{7} \times |-\frac{1}{2}|^2 = \frac{2}{7}$.

(b) The expectation value of A is just the sum of each value of A weighted by the probability of getting that value. The probabilities for $A = 1$ and $A = 3$ can easily be calculated as for $A = 2$ above. The results are:

(i) For $|\psi\rangle$ the expectation value is $\langle A \rangle = \frac{2}{3} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 = \frac{3}{2}$.

(ii) For $|\phi\rangle$ the expectation value is $\langle A \rangle = \frac{9}{14} \times 1 + \frac{2}{7} \times 2 + \frac{1}{14} \times 3 = \frac{10}{7}$.

1.4. If we have a two-dimensional Hilbert space and represent states by two-component column vectors such as u and v, find the necessary and sufficient conditions on the 2×2 matrix M so that $u^\dagger M v$ defines an inner product on the Hilbert space.

Hint: Recall the three conditions for the inner product on physical states, but you can just state without proof any properties which follow automatically from matrix multiplication.

Solution:

The linear property follows immediately from the linearity of matrix multiplication so there is no need to check anything.

The 'symmetry' of the inner product is the statement that the complex conjugate of the inner product of v with u is the inner product of u with v. I.e.

$$u^\dagger M v = (v^\dagger M u)^* = u^\dagger M^\dagger v.$$

For this to hold for all vectors u and v we must have $M^\dagger = M$, i.e. M must be a Hermitian matrix.

The 'physical state' property requires that for any non-zero vector u, its norm squared (given by the inner product of u with itself) is positive, i.e. $u^\dagger M u > 0$. This means that M must be a positive definite matrix. Since we know that M is Hermitian, this is equivalent to the requirement that the eigenvalues of M are all positive.

1.5. Recall that functions such as the exponential of operators are defined through their Taylor series.

(a) Show that if $\left[\hat{A}, \hat{B}\right] = 0$ then

$$\exp\left(\hat{A}\right)\exp\left(\hat{B}\right) = \exp\left(\hat{A}+\hat{B}\right).$$

(b) Show that if $\left[\hat{A}, \hat{B}\right] \neq 0$ then

$$\exp\left(\alpha\hat{A}\right)\exp\left(\beta\hat{B}\right) \neq \exp\left(\alpha\hat{A}+\beta\hat{B}\right)$$

for arbitrary $\alpha, \beta \in \mathbb{C}$. (The expressions can be equal for specific values of α, β.)

(c) Show that if $\hat{J}^2 = -\hat{I}$ where $\hat{I}$ is the identity operator then

$$\exp\left(\theta\hat{J}\right) = (\cos\theta)\hat{I} + (\sin\theta)\hat{J}$$

for any $\theta \in \mathbb{C}$.

Solution:

(a) Since $\left[\hat{A}, \hat{B}\right] = 0$ we can write any product of these operators as $\hat{A}^m\hat{B}^n$ with some coefficient and we don't have to distinguish terms which differ by the order of the products of the operators. E.g. we have $\hat{A}\hat{B}\hat{A}\hat{B}^3 = \hat{A}^2\hat{B}^4$. So, to prove the identity we just have to show that the coefficient of $\hat{A}^m\hat{B}^n$ is the same on both sides for all (non-negative) integers m and n.
On the left hand side we have coefficient $\frac{1}{m!}\frac{1}{n!}$ from the definition of the exponential function. On the right hand side this term will appear when expanding the term $(\hat{A}+\hat{B})^{m+n}$. That term has a coefficient $\frac{1}{(m+n)!}$ from the exponential, and then we have a binomial coefficient $\frac{(m+n)!}{m!\,n!}$ from expanding the polynomial. Therefore we see that the coefficient matches on both sides, so we have proven the identity.

(b) Here we can think of the expressions on each side as 'function' of two variables, α and β. Then expanding the exponentials will give a two-variable Taylor series, and for the two functions to be equal their Taylor series must match. This means that each term with any given powers of α and β must match.
It is easy to see that the terms with $\alpha^0\beta^0 = 1$, $\alpha^1\beta^0 = \alpha$ and $\alpha^0\beta^1 = \beta$ all match as they don't care about the commutation relation between $\hat{A}$ and $\hat{B}$. In fact the same is true for any term $\alpha^m\beta^0 = \alpha^m$ or $\alpha^0\beta^n = \beta^n$. The first non-trivial check comes for the term with $\alpha\beta$. On the LHS we get $\alpha\beta\hat{A}\hat{B}$ while on the RHS the term comes from the exponent squared so we have

$$\frac{1}{2!}\alpha\beta\left(\hat{A}\hat{B}+\hat{B}\hat{A}\right).$$

For non-zero α and β these expression are the same only if $\hat{A}\hat{B} = \hat{B}\hat{A}$, i.e. only if $\left[\hat{A},\hat{B}\right]=0$.

(c) Use the definition of the exponential of an operator in terms of the Taylor series. Since $\hat{J}^2=-\hat{I}$ all even powers of $\hat{J}$ will be proportional to the identity, specifically $\hat{J}^{2n}=(-1)^n\hat{I}$. For odd powers there will be an extra factor of $\hat{J}$ so $\hat{J}^{2n+1}=(-1)^n\hat{J}$. Then we simply calculate

$$\begin{aligned}\exp\left(\theta\hat{J}\right) &= \sum_{m=0}^{\infty}\frac{1}{m!}\left(\theta\hat{J}\right)^m \\ &= \sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(\theta\hat{J}\right)^{2n}+\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(\theta\hat{J}\right)^{2n+1} \\ &= \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\theta^{2n}\hat{I}+\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\theta^{2n+1}\hat{J} \\ &= \cos(\theta)\hat{I}+\sin(\theta)\hat{J}\end{aligned}$$

2 Problems: Measurement and uncertainty

2.1. Consider a two-dimensional Hilbert space with states represented by two-component column vectors, and with the standard inner product. **PROBLEMS CLASS 1**

(a) Find the (normalised) density matrix for each of the following states:

$$\begin{pmatrix}1\\0\end{pmatrix},\ \begin{pmatrix}0\\1\end{pmatrix},\ \begin{pmatrix}1\\1\end{pmatrix},\ \begin{pmatrix}1\\-1\end{pmatrix}.$$

(b) Find states corresponding to the following density matrices, if possible:

$$\begin{pmatrix}\frac{3}{4} & \frac{\sqrt{3}}{4}i\\ -\frac{\sqrt{3}}{4}i & \frac{1}{4}\end{pmatrix},\ \begin{pmatrix}\frac{3}{4} & \frac{\sqrt{3}}{4}i\\ \frac{\sqrt{3}}{4}i & \frac{1}{4}\end{pmatrix},\ \begin{pmatrix}\frac{3}{4} & \frac{\sqrt{3}}{4}\\ \frac{\sqrt{3}}{4} & \frac{1}{4}\end{pmatrix},\ \begin{pmatrix}\frac{3}{4} & 0\\ 0 & \frac{1}{4}\end{pmatrix}.$$

(c) Calculate $\mathrm{Tr}(\rho^2)$ for each of the density matrices ρ in parts (a) and (b).

Solution:

(a) Recall that for a pure state $|\psi\rangle$ the density operator is $\hat{\rho}=|\psi\rangle\langle\psi|$. For the usual vector representation of states this means that for a vector u we have the density matrix $\rho=uu^\dagger$. If the state/vector is not normalised we need to either first normalise it, or divide by the norm squared when calculating the density operator/matrix. For the four given vectors we have the following density matrices ρ:

$$\begin{aligned}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} &= \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\\ \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} &= \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\\ \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\\ \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1 & -1\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}\end{aligned}$$

Note that in all cases the density matrices are properly normalised, i.e. $\mathrm{Tr}\rho=1$ and the matrices are Hermitian.

(b) You can just consider an arbitrary normalised pure state which is represented by a vector $\begin{pmatrix} a \\ b \end{pmatrix}$ for some $a, b \in \mathbb{C}$ with $|a|^2 + |b|^2 = 1$. This will have density matrix

$$\rho = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} a^* & b^* \end{pmatrix} = \begin{pmatrix} |a|^2 & ab^* \\ ba^* & |b|^2 \end{pmatrix}.$$

It is then easy to see that the first and the third matrices are given by the vectors $\begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{-i}{2} \end{pmatrix}$ and $\begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}$ (and these vectors are unique up to multiplication by a phase $\exp(i\phi)$ for any $\phi \in \mathbb{R}$.

The second matrix is not Hermitian so it cannot be a density matrix (for a pure or a mixed state.)

The final matrix clearly cannot be the density matrix for a pure state, but it is Hermitian and has trace one. In fact it is also clearly a positive matrix so it is a mixed state density matrix. For mixed states the ensemble is not unique, but an obvious example here is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ with probability $\frac{3}{4}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ with probability $\frac{1}{4}$.

(c) For the density matrices from part (a), and for the first and third matrices in (b) you should find $\mathrm{Tr}(\rho^2) = 1$ since they are pure states. For the final matrix in (b) $\mathrm{Tr}(\rho^2) = \frac{5}{8} < 1$ as expected for a mixed state.

3 Problems: Qubits and the Bloch sphere

3.1. An arbitrary qubit density matrix (for a mixed or pure state) can be written

$$\hat{\rho} = \frac{1}{2}\left(I + \mathbf{r} \cdot \sigma\right)$$

where I is the 2×2 identity matrix, σ_i are the three Pauli sigma-matrices and the real 'Bloch' vector $\mathbf{r}$ has length $|\mathbf{r}| \leq 1$.

(a) Suppose $\hat{\rho}_1$ and $\hat{\rho}_2$ are density matrices for the pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively. What condition on $\hat{\rho}_1\hat{\rho}_2$ is equivalent to the statement that $|\psi_1\rangle$ is orthogonal to $|\psi_2\rangle$?

(b) Express the condition for $\hat{\rho}_1\hat{\rho}_2$ in part (a) (and now allowing pure or mixed states) in terms of conditions on the Bloch vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ defining the two density matrices. How are two orthogonal qubit states represented on the Bloch sphere?

Solution:

(a) Note that if $|\psi_1\rangle$ is orthogonal to $|\psi_2\rangle$ then $\langle\psi_1|\psi_2\rangle = 0$ so

$$\hat{\rho}_1\hat{\rho}_2 = |\psi_1\rangle\langle\psi_1||\psi_2\rangle\langle\psi_2| = (\langle\psi_1|\psi_2\rangle)|\psi_1\rangle\langle\psi_2| = 0.$$

Conversely if $\hat{\rho}_1\hat{\rho}_2 = 0$ then

$$0 = \mathrm{Tr}\left(\hat{\rho}_1\hat{\rho}_2\right) = (\langle\psi_1|\psi_2\rangle)\mathrm{Tr}\left(|\psi_1\rangle\langle\psi_2|\right) = (\langle\psi_1|\psi_2\rangle)(\langle\psi_2|\psi_1\rangle) = |\langle\psi_1|\psi_2\rangle|^2.$$

However, this means that $\langle\psi_1|\psi_2\rangle = 0$.

Therefore we have the following:

$$\hat{\rho}_1\hat{\rho}_2 = 0 \iff \langle \psi_1 | \psi_2 \rangle = 0 \iff \langle \psi_2 | \psi_1 \rangle = 0 \iff \hat{\rho}_2\hat{\rho}_1 = 0.$$

(b) Using the relation between the density matrix and the Bloch vector we have

$$4\hat{\rho}_1\hat{\rho}_2 = (I + \mathbf{r}_1 \cdot \sigma)(I + \mathbf{r}_2 \cdot \sigma) = I + (\mathbf{r}_1 + \mathbf{r}_2) \cdot \sigma + (\mathbf{r}_1 \cdot \mathbf{r}_2)I + i(\mathbf{r}_1 \times \mathbf{r}_2) \cdot \sigma$$

where the last two terms come from using an identity for the σ-matrices in terms of their commutation ($[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$) and anti-commutation ($\{\hat{\sigma_i}, \hat{\sigma_j}\} \equiv \hat{\sigma_i}\hat{\sigma_j} + \hat{\sigma_j}\hat{\sigma_i} = 2\delta_{ij}I$) relations.
Since the identity matrix and the three σ-matrices are linearly independent, the product of density matrices can only vanish if

$$\begin{aligned} 1 + \mathbf{r}_1 \cdot \mathbf{r}_2 &= 0 \\ \mathbf{r}_1 + \mathbf{r}_2 + i\mathbf{r}_1 \times \mathbf{r}_2 &= \mathbf{0} \end{aligned}$$

Recall that the Bloch vectors cannot have magnitude greater than one, i.e. they must give points on (for pure states) or inside (for mixed states) the Bloch sphere which has radius one. With this restriction the first condition is equivalent to $\mathbf{r}_2 = -\mathbf{r}_1$ and $|\mathbf{r}_1| = 1$. I.e. the two states must be pure states and must sit on antipodal points of the Bloch sphere. Clearly the second condition is then satisfied. For a single qubit we see that a mixed state cannot be orthogonal to any other state.
We could alternatively take the orthogonality condition to be $\mathrm{Tr}(\hat{\rho}_1\hat{\rho}_2) = 0$ and since the σ-matrices are traceless we would only get the first constraint. However, note that when $\mathbf{r}_2 = -\mathbf{r}_1$ the second condition above is automatically satisfied, so these conditions for orthogonality are equivalent.

3.2. Consider a qubit system with standard basis $\{|0\rangle, |1\rangle\}$. Define the following states:

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad , \quad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ |L\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \quad , \quad |R\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \end{aligned}$$

(a) Find the density matrices for each of the pure states $|0\rangle, |1\rangle, |+\rangle, |-\rangle, |L\rangle, |R\rangle$.

(b) Find the density matrices for each of the following mixed states:

(i) $|0\rangle$ with probability $\frac{1}{2}$, $|1\rangle$ with probability $\frac{1}{4}$, $|+\rangle$ with probability $\frac{1}{4}$.

(ii) $|0\rangle$ with probability $\frac{1}{2}$, $|1\rangle$ with probability $\frac{1}{2}$.

(iii) $|+\rangle$ with probability $\frac{1}{2}$, $|-\rangle$ with probability $\frac{1}{2}$.

(iv) $|L\rangle$ with probability $\frac{1}{2}$, $|R\rangle$ with probability $\frac{1}{2}$.

(c) Using the Bloch sphere, sketch the regions which can be described as an ensemble (with arbitrary probabilities whose total is 1) of the following pure states:

(i) $|0\rangle$ and $|1\rangle$

(ii) $|+\rangle$ and $|-\rangle$

(iii) $|L\rangle$ and $|R\rangle$

(iv) $|1\rangle$ and $|L\rangle$

(v) $|0\rangle$, $|+\rangle$ and $|R\rangle$

Solution:

(a) For any pure state $|\psi\rangle$ the density operator is $\hat{\rho} = |\psi\rangle\langle\psi|$. For a qubit system we can always write the operator $\hat{\rho}$ using the standard basis as

$$\hat{\rho} = \rho_{00}|0\rangle\langle 0| + \rho_{01}|0\rangle\langle 1| + \rho_{10}|1\rangle\langle 0| + \rho_{11}|1\rangle\langle 1| .$$

Obviously this can be represented in matrix form as

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}$$

You can either calculate in Dirac notation or immediately write the states as vectors u and then calculate $\rho = uu^\dagger$. E.g. for $|L\rangle$ either note that $\langle L| = \frac{1}{\sqrt{2}}(\langle 0| - i\langle 1|)$ or just calculate

$$\rho_L = \left(\frac{1}{\sqrt{2}}\right)^2 \begin{pmatrix} 1 \\ i \end{pmatrix} \begin{pmatrix} 1 & -i \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

Similarly for the other cases you should find

$$\begin{aligned}
\rho_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\
\rho_1 &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\
\rho_+ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\
\rho_- &= \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \\
\rho_R &= \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.
\end{aligned}$$

(b) For a mixed state the density matrix is just the linear combination of the density matrices for each component of the ensemble, weighted by the probability. For the four cases in this question we have density matrices

$$\begin{aligned}
\frac{1}{2}\rho_0 + \frac{1}{4}\rho_1 + \frac{1}{4}\rho_+ &= \frac{1}{8}\begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix} \\
\frac{1}{2}\rho_0 + \frac{1}{2}\rho_1 &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\frac{1}{2}\rho_+ + \frac{1}{2}\rho_- &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\frac{1}{2}\rho_L + \frac{1}{2}\rho_R &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}$$

(c) Either recall for lectures, or check by extracting the Bloch vector from the density matrix, where each of these pure states sits on the Bloch sphere. You should find that $|0\rangle$ and $|1\rangle$ sit on the positive and negative z-axis respectively. Similarly for $|+\rangle$ and $|-\rangle$ on the x-axis and $|L\rangle$ and $|R\rangle$ on the y-axis. Of course, being pure states, all are distance one from the origin.

Now a mixed state of $|0\rangle$ and $|1\rangle$ will have density matrix given by a linear combination of ρ_0 and ρ_1, but the coefficients are probabilities adding up to one so the density matrix must be of the form

$$\rho = (1-\lambda)\rho_0 + \lambda\rho_1 = \rho_0 + \lambda(\rho_1 - \rho_0)$$

where $0 \leq \lambda \leq 1$. But this is just the parametric equation for the straight line between the points on the Bloch sphere given by ρ_0 and ρ_1, i.e. in this case the points on the z-axis with $z = 1$ and $z = -1$. The limited range of λ means that we only consider the line segment with endpoints given by ρ_0 and ρ_1.

Similarly we can describe the other mixed states of two pure states. Note that as for $|0\rangle$ and $|1\rangle$, the lines for $|+\rangle$ and $|-\rangle$ and for $|L\rangle$ and $|R\rangle$ go through the origin. Indeed the origin is the point corresponding to equal weighting in each case, and this explains the corresponding results in the previous part.

For the final case we have a mixture of 3 pure states. Then the linear combination gives the parametric equation for the plane containing the three points on the Bloch sphere. Again the coefficients are probabilities totally one, and this constraint defines the triangle with the three points giving the vertices.

In general a mixed state ensemble with arbitrary probabilities will be represented by the polyhedron, and its interior, defined by the vertices given by the pure states in the ensemble. In special cases where the vertices are co-planar or co-linear the shape will degenerate to a polygon or a line.

3.3. Suppose a single qubit system is in the state $|0\rangle$.

(a) What are the possible outcomes, and associated probabilities, of a measurement of the observable σ_3?

(b) If instead σ_1 is measured, what are the possible outcomes and probabilities?

(c) Now suppose σ_1 is measured first, then σ_3 is measured and then σ_1 is measured again. Describe the possible outcomes and probabilities at each stage. Is there any relation between the results of the two measurements of σ_1?

Solution:

(a) The outcome would be 1 with probability 1 since $|0\rangle$ is an eigenstate of σ_3 with eigenvalue 1.

(b) The eigenstates of σ_1 are $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ with eigenvalues ± 1. Since $|0\rangle = (|+\rangle + |-\rangle)/\sqrt{2}$, either outcome ± 1 is possible with probability $1/2$.

(c) The first measurement of σ_1 will produce result ± 1 and change the state to $|\pm\rangle$ with equal probability. Since the states $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, the measurement of σ_3 will give result ± 1 and change the state to $|0\rangle$ or $|1\rangle$ with equal probability, and without any correlation to the earlier measurement of σ_1. Since these states are $(|+\rangle \pm |-\rangle)/\sqrt{2}$ the second measurement of σ_1 will give result ± 1 and produce the state $|\pm\rangle$ with equal probability. The final result does not depend in any way on the previous measurements, in particular it is independent of the result of the first measurement of σ_1.

4 Problems: Bipartite systems

4.1. For a single classical bit the NOT gate implements the logical operation

$$m \to \overline{m} \equiv NOT(m)$$

defined by

$$\overline{0} = 1 \ , \quad \overline{1} = 0.$$

(a) Show that a quantum NOT gate can be implemented as a unitary operation on a single qubit, i.e. sending $|m\rangle \to |\overline{m}\rangle$, and write the unitary operator as a 2×2 matrix in the standard basis.

(b) We can erase a classical bit by sending

$$0 \to 0 \ \text{ and } \ 1 \to 0.$$

Explain why this operation cannot be implemented by a unitary transformation of a single qubit system.

(c) Can we construct a unitary transformation to erase a single qubit in a larger system formed by taking a tensor product of the qubit Hilbert space with another (finite dimensional) Hilbert space?

Solution:

1. Note that any linear operator is fully determined by its action on the basis states. Also, for a qubit system we can write any linear operator in the form

$$\hat{A} = |\psi_0\rangle \langle 0| + |\psi_1\rangle \langle 1|$$

and it will map $|0\rangle \to |\psi_0\rangle$ and $|1\rangle \to |\psi_1\rangle$.
So, in this case the operator must be

$$\hat{U}_N = |1\rangle \langle 0| + |0\rangle \langle 1| \ .$$

Written in the standard basis this is

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

To show that the operator is unitary, it is sufficient to note that the matrix is unitary. Alternatively, since the adjoint of $|\psi\rangle \langle \phi|$ is $|\phi\rangle \langle \psi|$, we see that $\hat{U}_N^\dagger = \hat{U}_N$ and using the orthonormality of the basis it is easy to see that

$$\hat{U}_N^\dagger \hat{U}_N = |0\rangle \langle 0| + |1\rangle \langle 1| = \hat{I}$$

as required to show that $\hat{U}_N$ is a unitary operator.

2. By definition, a unitary operator must be invertible. Since the linearly independent states $|0\rangle$ and $|1\rangle$ are both mapped to $|0\rangle$, this is not an invertible map.

3. Yes, this is possible and there is not a unique way to do it. To check this, let's embed the single qubit states into a 2-qubit system (this is the simplest possibility, so we try it first) as

$$|0\rangle \to |0\rangle \otimes |0\rangle \ , \quad |1\rangle \to |1\rangle \otimes |0\rangle \ .$$

(Another natural choice is $|1\rangle \to |1\rangle \otimes |1\rangle$ which will give a slightly different solution.)

Now the erase operation should set the first qubit to $|0\rangle$ but we can choose what happens to the second qubit. From part (b) we know that we cannot set it to the same state for both cases. So, let's choose a linear operator which acts as

$$|0\rangle \otimes |0\rangle \to |0\rangle \otimes |0\rangle \ , \ \ |1\rangle \otimes |0\rangle \to |0\rangle \otimes |1\rangle \ .$$

Note that we are still free to decide how the operator acts of the other 2-qubit basis states $|0\rangle \otimes |1\rangle$ and $|1\rangle \otimes |1\rangle$. It is straightforward to check that we get a unitary transformation if we choose

$$|0\rangle \otimes |1\rangle \to |1\rangle \otimes |0\rangle \ , \ \ |1\rangle \otimes |1\rangle \to |1\rangle \otimes |1\rangle \ .$$

If you work in the standard vector representation, the required mapping is

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \to \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

so the matrix must be of the form

$$\begin{pmatrix} 1 & * & 0 & * \\ 0 & * & 1 & * \\ 0 & * & 0 & * \\ 0 & * & 0 & * \end{pmatrix}$$

and then it is easy to check that

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is a unitary matrix of that form.

4.2. Consider a two-qubit bipartite system and use the standard orthonormal basis states for each qubit subsystem.

(a) Explain why any pure state can be written in the form

$$|\Psi\rangle = a\,|0\rangle \otimes |\phi_0\rangle + b\,|1\rangle \otimes |\phi_1\rangle$$

where $|\phi_0\rangle$ and $|\phi_1\rangle$ are normalised states in system B.

(b) What, if any, constraints must $a, b \in \mathbb{C}$, $|\phi_0\rangle$ and $|\phi_1\rangle$ satisfy so that $|\Psi\rangle$ is normalised.

(c) Write the density operator $\hat{\rho}$ for the system, and calculate the reduced density operators $\hat{\rho}_A$ and $\hat{\rho}_B$.

(d) Show that $\mathrm{Tr}(\hat{\rho}_A^2) = \mathrm{Tr}(\hat{\rho}_B^2)$ and find the range of possible values for this quantity.

(e) What conditions must be satisfied in order to maximise the value of $\mathrm{Tr}(\hat{\rho}_A^2)$ and what property does the state $|\Psi\rangle$ have in this case.

(f) What conditions must be satisfied in order to minimise the value of $\mathrm{Tr}(\hat{\rho}_A^2)$.

(g) Suppose now that system B has a Hilbert space with dimension larger than two, but system A is still an single qubit. Does that change any of the results above?

Solution:

1. We can take the basis $\{|0\rangle \otimes |i\rangle, |1\rangle \otimes |i\rangle\}$ for the system where $i \in \{0,1\}$ but in fact we could easily generalise to the case where system B was a larger system by simply allowing more values for i. Then an arbitrary pure state must be a linear combination of these basis states, so for some constants c_{xi},

$$|\Psi\rangle = \sum_{x,i} c_{xi} |x\rangle \otimes |i\rangle$$

where $x \in \{0,1\}$. However, for each x, $\sum_i c_{xi} |i\rangle$ is a linear combination of the basis states for system B, hence is a state (not necessarily normalised) in system B. Then we just choose normalisation constants so that

$$|\phi_0\rangle = \frac{1}{a} \sum_i c_{0i} |i\rangle \text{ and } |\phi_1\rangle = \frac{1}{b} \sum_i c_{1i} |i\rangle$$

are normalised. (In the exceptional case where $\sum_i c_{0i} |i\rangle = 0$ we can take any normalised state $|\phi_0\rangle$ with $a = 0$, and similarly $b = 0$ if $\sum_i c_{1i} |i\rangle = 0$.)

2. Since $\{|0\rangle, |1\rangle\}$ are orthonormal, and $\langle\phi_0|\phi_0\rangle = \langle\phi_1|\phi_1\rangle = 1$ (but without assuming anything about $\langle\phi_0|\phi_1\rangle$),

$$1 = \langle\Psi|\Psi\rangle = |a|^2 + |b|^2 .$$

3. By definition

$$\begin{aligned}\hat{\rho} &= |\Psi\rangle\langle\Psi| = \Big(a|0\rangle \otimes |\phi_0\rangle + b|1\rangle \otimes |\phi_1\rangle\Big) \otimes \Big(a^*\langle 0| \otimes \langle\phi_0| + b^*\langle 1| \otimes \langle\phi_1|\Big) \\ &= |a|^2 |0\rangle\langle 0| \otimes |\phi_0\rangle\langle\phi_0| + ab^* |0\rangle\langle 1| \otimes |\phi_0\rangle\langle\phi_1| + \\ &+ ba^* |1\rangle\langle 0| \otimes |\phi_1\rangle\langle\phi_0| + |b|^2 |1\rangle\langle 1| \otimes |\phi_1\rangle\langle\phi_1|\end{aligned}$$

We can then easily calculate $\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho})$ and $\hat{\rho}_B = \mathrm{Tr}_A(\hat{\rho})$.

$$\begin{aligned}\hat{\rho}_A &= |a|^2 |0\rangle\langle 0| + ab^* \langle\phi_1|\phi_0\rangle |0\rangle\langle 1| + ba^* \langle\phi_0|\phi_1\rangle |1\rangle\langle 0| + |b|^2 |1\rangle\langle 1| \\ \hat{\rho}_B &= |a|^2 |\phi_0\rangle\langle\phi_0| + |b|^2 |\phi_1\rangle\langle\phi_1|\end{aligned}$$

4. Using the properties of the states we find

$$\begin{aligned}\mathrm{Tr}(\hat{\rho}_A^2) &= \mathrm{Tr}(\hat{\rho}_B^2) = |a|^4 + 2|a|^2|b|^2 |\langle\phi_0|\phi_1\rangle|^2 + |b|^4 \\ &= 2|a|^2(|a|^2 - 1)\Big(1 - |\langle\phi_0|\phi_1\rangle|^2\Big) + 1\end{aligned}$$

Note that in the first line the 3 quantities are non-negative, while the first term in the second line is non-positive.

5. This means that the maximum value is 1 which is attained precisely when either $|a| = 0$, $|a| = 1$, or $|\langle\phi_0|\phi_1\rangle| = 1$. These conditions are equivalent to $a = 0$, $b = 0$, or $|\phi_0\rangle = e^{i\phi}|\phi_1\rangle$. In all cases this means that $|\Psi\rangle$ is separable (and any separable state is of that form.)
6. For any (non-zero) values of a and b the minimum value is clearly obtained when $|\langle\phi_0|\phi_1\rangle| = 0$ so that the middle term in the first line vanishes. Then using the constraint $|a|^2 = |b|^2 = 1$ it is simple algebra to check that the minimum value is $1/2$, attained when $|a|^2 = |b|^2 = 1/2$.
7. No, since we did not assume that system B was a single qubit system above.

4.3. Repeat question 4.2 for a bipartite system where system A is a two-qubit system and system B is a two-qubit system or a larger system. Write a suitable generalisation of the form of a pure state $|\Psi\rangle$ in part (a).

Solution:

This is similar to the solution to question 4.2 but letting x run over 4 values corresponding to 4 orthonormal basis states for the 2-qubit system. E.g. by similar arguments

$$|\Psi\rangle = \sum_x a_x |x\rangle \otimes |\phi_x\rangle \ .$$

5 Problems: Entanglement applications

5.1. Consider the operators

PROBLEMS CLASS 2

$$\hat{N}_1 \equiv \sigma_1 \otimes \hat{I} \quad \text{and} \quad \hat{N}_2 \equiv \sigma_3 \otimes \sigma_1$$

acting on a 2-qubit system.

(a) Write $\hat{N}_1$ and $\hat{N}_2$ as 4×4 matrices in the representation where the standard basis states $|m\rangle \otimes |n\rangle$ are written as 4-component column vectors with all components zero except a 1 in the row counted by the 2-digit binary number $(mn)_2$, e.g. $|1\rangle \otimes |0\rangle \to (0\ 0\ 1\ 0)^T$.

(b) Write the operators $\hat{N}_+ \equiv \hat{N}_1 + \hat{N}_2$ and $\hat{N}_\times \equiv \hat{N}_1\hat{N}_2$ in matrix form. Explain why the structure of the 4×4 matrices shows that $\hat{N}_\times$ can be written in the form $\hat{A} \otimes \hat{B}$ but $\hat{N}_+$ cannot.

(c) Show that $\hat{U} \equiv \frac{1}{\sqrt{2}}\hat{N}_+$ and $\hat{N}_\times$ are unitary operators.

(d) Show that $\hat{U}$ acting on the 4 basis states $|m\rangle \otimes |n\rangle$ produces the 4 Bell states

$$|\beta_{xy}\rangle \equiv \frac{1}{\sqrt{2}} \left(|0\rangle \otimes |y\rangle + (-1)^x |1\rangle \otimes |\overline{y}\rangle \right)$$

and show that none of the Bell states is a separable state.

(e) Find the 4 states produced by $\hat{U}$ acting on the 4 states $|\pm\rangle \otimes |\pm\rangle$ where $|\pm\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Are any of the resulting states separable?

If any are separable, can you explain why? (Hint: look at $\hat{N}_1$ and $\hat{N}_2$ acting on these states.)

Solution:

1. In this representation standard matrix multiplication shows that the columns of the matrices are (left to right) the results of the action of the basis vectors corresponding to the binary numbers $00, 01, 10, 11$ in that order. Also recall that σ_1 acts as $|0\rangle \to |1\rangle$ and $|1\rangle \to |0\rangle$ while σ_3 acts as $|0\rangle \to |0\rangle$ and $|1\rangle \to -|1\rangle$. E.g. for $\hat{N}_1$ we map $|0\rangle \otimes |0\rangle \to |1\rangle \otimes |0\rangle$ so the first column of $\hat{N}_1$ is given by the vector for $|1\rangle \otimes |0\rangle$, i.e. $(0\ 0\ 1\ 0)^T$. Repeating for the other basis states, and doing the same for $\hat{N}_2$ we find

$$N_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} , \quad N_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} .$$

Note the structure is of 2×2 blocks given by the form of the first operator in the tensor product (i.e. σ_1 for $\hat{N}_1$) and the structure of each individual block (itself a 2×2 matrix) is given by the second operator in the tensor product i.e. I for $\hat{N}_1$).

2. Just add and multiply the matrices to find

$$N_+ = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \; , \; N_\times = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \; .$$

We see that the blocks in N_+ are not of the same form (some are proportional to σ_1, others to I) so it cannot be written as a single tensor product. For $N_\times$ all the blocks are proportional to σ_1 so it can be written as $-i\sigma_2 \otimes \sigma_1$.

3. Just check that the corresponding matrices are unitary.

4. Just write the 4 Bell states in vector form and note that these are the 4 columns of the matrix N_+. Note that the values of x and y are $x = m$ and $y = \overline{n}$ for $m = 0$ while $y = n$ for $m = 1$.

   Any separable state can be written in the form $|\psi\rangle \otimes |\phi\rangle$. Since we can always write $|\psi\rangle = |0\rangle + b\,|1\rangle$ for a qubit system, we see that any separable state must be of the form $a\,|0\rangle \otimes |\phi\rangle + b\,|1\rangle \otimes |\phi\rangle$ but in the Bell states, since $|y\rangle$ and $|\overline{y}\rangle$ are linearly independent, we clearly cannot have both $a\,|\phi\rangle = |y\rangle$ and $b\,|\phi\rangle = (-1)^x\,|\overline{y}\rangle$.

5. Either write these states as a linear combination of the standard basis states (and we know how $\hat{U}$ acts on these states from the above parts) or just calculate the action of $\hat{N}_1$ and $\hat{N}_2$ directly. In particular note that $\sigma_3\,|\pm\rangle = |\mp\rangle$ while $\sigma_1\,|\pm\rangle = \pm\,|\pm\rangle$. This means that if we denote the $\pm$ signs by $s_1 = \pm 1$ and $s_2 = \pm 1$ we have

$$\hat{N}_1\,|s_1\rangle \otimes |s_2\rangle = s_1\,|s_1\rangle \otimes |s_2\rangle \;\; , \;\; \hat{N}_2\,|s_1\rangle \otimes |s_2\rangle = s_2\,|-s_1\rangle \otimes |s_2\rangle \;\; .$$

   Obviously the second factor in the tensor product is the same in both cases so the linear combination corresponding to the action of $\hat{U}$ is the separable state

$$\frac{1}{\sqrt{2}}\Big(s_1\,|s_1\rangle + s_2\,|-s_1\rangle\Big) \otimes |s_2\rangle \;\; .$$

   Note that if $s_1 = s_2$ the first factor is proportional to $|0\rangle$ while if $s_1 = -s_2$ it is proportional to $|1\rangle$.

5.2. For a 2-qubit system we define the 4 Bell states in terms of the standard basis states as: **PROBLEMS CLASS 3**

$$|\beta_{xy}\rangle \equiv \frac{1}{\sqrt{2}}\left(|0\rangle \otimes |y\rangle + (-1)^x\,|1\rangle \otimes |\overline{y}\rangle\right).$$

(a) Write the Bell states in the basis $\{\,|{+}{+}\rangle\,, |{+}{-}\rangle\,, |{-}{+}\rangle\,, |{-}{-}\rangle\,\}$.

(b) Write the Bell states in the basis $\{\,|LL\rangle\,, |LR\rangle\,, |RL\rangle\,, |RR\rangle\,\}$.

(c) If Alice and Bob share the Bell state $|\beta_{00}\rangle$ (one qubit each) what are the possible outcomes and the associated probabilities, and what are Alice and Bob's final states in each case:

   (i) Alice measures σ_3.

   (ii) Bob measures σ_3.

   (iii) Alice measures σ_1.

   (iv) Alice measures σ_2.

   (v) Alice measures σ_1 and then Bob measures σ_3.

   (vi) Alice measures σ_1 after Bob measures σ_3.

   (vii) Alice measures σ_1 and then Bob measures σ_1.

Solution:

1. Just replace the basis states using

$$|0\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle + |-\rangle \Big) \ , \ \ |1\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle - |-\rangle \Big) \ .$$

For example, using notation $|01\rangle = |0\rangle \otimes |1\rangle$ etc.

$$|00\rangle = \frac{1}{2}\Big(|+\rangle + |-\rangle \Big) \otimes \Big(|+\rangle + |-\rangle \Big) = \frac{1}{2}\Big(|{+}{+}\rangle + |{+}{-}\rangle + |{-}{+}\rangle + |{-}{-}\rangle \Big) \ .$$

similarly we find

$$\begin{aligned} |01\rangle &= \frac{1}{2}\Big(|{+}{+}\rangle - |{+}{-}\rangle + |{-}{+}\rangle - |{-}{-}\rangle \Big) \\ |10\rangle &= \frac{1}{2}\Big(|{+}{+}\rangle + |{+}{-}\rangle - |{-}{+}\rangle - |{-}{-}\rangle \Big) \\ |11\rangle &= \frac{1}{2}\Big(|{+}{+}\rangle - |{+}{-}\rangle - |{-}{+}\rangle + |{-}{-}\rangle \Big) \ . \end{aligned}$$

Then we just take the appropriate linear combinations to find

$$\begin{aligned} |\beta_{00}\rangle &= \frac{1}{\sqrt{2}}\Big(|{+}{+}\rangle + |{-}{-}\rangle \Big) \\ |\beta_{01}\rangle &= \frac{1}{\sqrt{2}}\Big(|{+}{+}\rangle - |{-}{-}\rangle \Big) \\ |\beta_{10}\rangle &= \frac{1}{\sqrt{2}}\Big(|{+}{-}\rangle + |{-}{+}\rangle \Big) \\ |\beta_{11}\rangle &= \frac{1}{\sqrt{2}}\Big(-|{+}{-}\rangle + |{-}{+}\rangle \Big) \ . \end{aligned}$$

2. This part is essentially the same as above, using

$$|0\rangle = \frac{1}{\sqrt{2}}\Big(|L\rangle + |R\rangle \Big) \ , \ \ |1\rangle = \frac{-i}{\sqrt{2}}\Big(|L\rangle - |R\rangle \Big) \ .$$

The result is

$$\begin{aligned} |\beta_{00}\rangle &= \frac{1}{\sqrt{2}}\Big(|LR\rangle + |RL\rangle \Big) \\ |\beta_{01}\rangle &= \frac{-i}{\sqrt{2}}\Big(|LL\rangle - |RR\rangle \Big) \\ |\beta_{10}\rangle &= \frac{1}{\sqrt{2}}\Big(|LL\rangle + |RR\rangle \Big) \\ |\beta_{11}\rangle &= \frac{-i}{\sqrt{2}}\Big(-|LR\rangle + |RL\rangle \Big) \ . \end{aligned}$$

3. Noting that the eigenstates of σ_1, σ_2, σ_3 are respectively $|+\rangle$, $|L\rangle$, $|0\rangle$ with eigenvalue 1 and $|-\rangle$, $|R\rangle$, $|1\rangle$ with eigenvalue -1, the above results can be used:
   (i) With probability $1/2$ in each case, Alice measures 1 and the state becomes $|00\rangle$ or -1 and the state becomes $|11\rangle$.
   (ii) Exactly as above, with probability $1/2$ in each case, Bob measures 1 and the state becomes $|00\rangle$ or -1 and the state becomes $|11\rangle$.

(iii) Now noting that

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}\Big(|{+}{+}\rangle + |{-}{-}\rangle\Big)$$

again with probability $1/2$ in each case, Alice measures 1 and the state becomes $|{+}{+}\rangle$ or -1 and the state becomes $|{-}{-}\rangle$.

(iv) This time noting that

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}\Big(|LR\rangle + |RL\rangle\Big)$$

again with probability $1/2$ in each case, Alice measures 1 and the state becomes $|LR\rangle$ or -1 and the state becomes $|RL\rangle$.

For parts (v) and (vi) use the results from parts (iii) and (ii) respectively. Note that in both cases, after the first measurement the state is a separable state, so for (v) Bob's measurement does not change Alice's state (which is either $|+\rangle$ or $|-\rangle$), and for (vi) Alice's measurement does not change Bob's state ($|0\rangle$ or $|1\rangle$). Also, if Bob has the state $|+\rangle$ or $|-\rangle$ and measures σ_3, in either case he will get results 1 or -1 with probability $1/2$ and his state will become $|0\rangle$ or $|1\rangle$ respectively. Similarly, Alice gets 1 or -1, producing state $|+\rangle$ or $|-\rangle$ respectively, with probability $1/2$ when measuring σ_1 on state $|0\rangle$ or $|1\rangle$. The final result is that for (v) or (vi) the results are exactly the same (as should be expected since the measurements correspond to the commuting operators $\sigma_1 \otimes I$ and $I \otimes \sigma_3$). With probability $1/4$ we have the following results and final states:

$$(1,1)\ ,\ |{+}0\rangle\ ;\ \ (1,-1)\ ,\ |{+}1\rangle\ ;\ \ (-1,1)\ ,\ |{-}0\rangle\ ;\ \ (-1,-1)\ ,\ -|{-}1\rangle\ .$$

Part (vii) follows immediately from part (iii). Bob will always get the same result as Alice so with probability $1/2$ the combined results and final states are :

$$(1,1)\ ,\ |{+}{+}\rangle\ ;\ \ (-1,-1)\ ,\ |{-}{-}\rangle\ .$$

5.3. Charlie knows that Alice and Bob are meeting for lunch on Saturday. He wants to send them a message with a surprise announcement that they can read together at that time, and not before. The problem is, he will be travelling and unable to communicate with them that day. However, he will see Alice on Thursday and Bob on Friday before he leaves, so he decides to give them each half the message with instructions to combine the information over lunch on Saturday. Suppose Charlie needs 100 bits for his message.

(a) At first Charlie decides to give Alice half the bits, say the odd ones (i.e. the first, the third, the fifth, etc.) and to give Bob the even ones. However, he soon realises that this won't work since Alice and Bob will probably be able to guess some of the message from their 50 bits. After some thought he realises that he can avoid this problem by giving them each 100 bits. How might this work?

(b) Just before Alice arrives on Thursday, Charlie realises the flaw in his plan. Alice and Bob are both impatient so will just send each other the information before Saturday! Then (what luck!) he remembers he has 50 Bell states. He checks his 100-bit message and performs some unitary transformations of the Bell states. When Alice leaves he gives her 50 qubits and some instructions, and the next day gives the remaining 50 qubits to Bob.

Assuming Alice and Bob cannot meet before Saturday, and they do not have a quantum communication channel, explain how Charlie's plan might work.

Solution:

1. One way is that the message can be found by combining Alice and Bob's 100 bits using bitwise addition modulo 2. I.e. each bit of the message will be a 0 if Alice and Bob's corresponding bits are the same, and a 1 if they are different. Obviously Alice or Bob alone cannot deduce anything about the message since for each bit they have it is equally likely that the other will have the same or different value for that bit.
2. He can use the 50 Bell states to encode 50 pairs of bits. As in superdense coding, if Alice and Bob have one qubit of each Bell pair, they cannot determine anything about which of the four Bell states it is. Of course, when they meet they can combine the qubits and measure which Bell states they have, generating the $100 = 50 \times 2$ bit message.

   Note that this is not necessarily a flawless mechanism to keep the message secret. E.g. if Alice and Bob each measure σ_3 and discuss their results (using classical communication), they can determine half the message. This is because they will get the same result for $|\beta_{00}\rangle$ or $|\beta_{10}\rangle$, but different results for $|\beta_{01}\rangle$ or $|\beta_{11}\rangle$. Of course, once they have done this they will never be able to recover the full message, so it is not quite the same as the case where they each have 50 classical bits of the message.

5.4. Alice needs to urgently send a secret message to Bob. For simplicity, assume she can do this with just 2 bits. They would like to do this using superdense coding, but they don't share any entangled states and to make matters worse someone ordered the wrong equipment – Alice's quantum entangler is broken, as is her quantum receiver and Bob's quantum transmitter. Fortunately, they each share a Bell state with Charlie whom they trust. So, their plan is for Alice to send the message to Charlie using superdense coding, and for Charlie to read it and transmit it to Bob, again using superdense coding. Then, they get news that Eve is spying on Charlie (it is vital that she does not see the message) and also blocking all his quantum communications. All seems lost, but actually there is a way for Alice to send the message to Bob using superdense coding with Charlie's help. Explain how this can be done in a way that does not require Charlie to use quantum communication, and so that Eve gains no information about the message from eavesdropping on any classical communications, or watching what Charlie does.

Solution:

You can picture this in terms of a line connecting Alice to Charlie, and another line connecting Bob to Charlie. Each line represents a shared Bell state. Effectively, since there is a line from Alice to Charlie to Bob, it is possible for Charlie to manipulate the two qubits he has so that Alice and Bob share a Bell state. This can be done using teleportation. E.g. Alice can transform the Bell pair she shares according to her 2-bit message (the first part of superdense coding), Charlie can teleport the qubit he shares with Alice to Bob (using classical communication which does not reveal any details of the state being teleported), Alice can send her qubit to Bob (quantum communication). Finally Bob will have the Bell state chosen by Alice so can measure to read the 2-bit message. Of course Alice must transform the Bell state before she sends her qubit to Bob, but otherwise the order of the three steps does not matter.

6 Problems: Information theory

6.1. (a) Show that the operator $\hat{S} = \mathbf{n} \cdot \sigma$, where $\mathbf{n}$ is a three-dimensional unit vector, has eigenvalues ± 1. PROBLEMS CLASS 4

(b) Recall that any qubit density matrix can be written in terms of a Bloch vector $\mathbf{r}$. Calculate the expectation value of $\hat{S}$ in terms of $\mathbf{n}$ and $\mathbf{r}$.

(c) By considering the range of values possible in part (b), state what the eigenstates of $\hat{S}$ are in terms of a relation between $\mathbf{r}$ and $\mathbf{n}$.

(d) For each Bell state, how is the expectation value of $\hat{S} \otimes \hat{I}$ related to that of $\hat{I} \otimes \hat{S}$?

(e) For each Bell state, calculate the expectation value of $\hat{S} \otimes \sigma_3$ and $\hat{S} \otimes \sigma_1$.

Solution:

(a) You could write out $\hat{S}$ as a 2×2 matrix and calculate the eigenvalues. Alternatively, note that since $\mathrm{Tr}\sigma_i = 0$, $\mathrm{Tr}\hat{S} = 0$ which means that the sum of the (two) eigenvalues is 0. Then note that $\hat{S}^2 = \hat{I}$ since

$$(\mathbf{n}\cdot\sigma)^2 = n_i n_j \sigma_i \sigma_j = \frac{1}{2}(n_i n_j \sigma_i \sigma_j + n_j n_i \sigma_j \sigma_i) = \frac{1}{2} n_i n_j (\sigma_i \sigma_j + \sigma_j \sigma_i) = \frac{1}{2} n_i n_j 2\delta_{ij} I = (\mathbf{n}\cdot\mathbf{n}) I = I \ .$$

Therefore each eigenvalue must square to 1, hence the two eigenvalues are 1 and -1.

(b) Since the trace is cyclic

$$\mathrm{Tr}(\sigma_i \sigma_j) = \mathrm{Tr}(\sigma_j \sigma_i) = \frac{1}{2}\mathrm{Tr}(\sigma_i \sigma_j + \sigma_j \sigma_i) = \frac{1}{2}\mathrm{Tr}(2\delta_{ij} I) = \delta_{ij}\mathrm{Tr}I = 2\delta_{ij} \ .$$

Therefore we have

$$\left\langle \hat{S} \right\rangle = \mathrm{Tr}(\hat{\rho}\hat{S}) = \frac{1}{2}\mathrm{Tr}\Big((I + r_i \sigma_i) n_j \sigma_j\Big) = \frac{1}{2} r_i n_j 2\delta_{ij} = \mathbf{r}\cdot\mathbf{n} \ .$$

(c) Since $|\mathbf{r}| \leq 1$ and $\mathbf{n}$ is a unit vector, $-1 \leq \mathbf{r}\cdot\mathbf{n} \leq 1$. On the other hand, for an eigenstate of $\hat{S}$, the expectation value is equal to the eigenvalue. Since the expectation value is an average over the possible outcomes (given by the eigenvalues), in this case any state which is not an eigenstate must have expectation value strictly between -1 and 1. Hence the eigenstates are exactly the states with $\mathbf{r} = \pm\mathbf{n}$.

(d) One approach is to directly calculate the expectation value by considering the action of these operators on the Bell states. Another approach is to note that we can calculate the trace in a bipartite system by first taking the partial trace in one subsystem, then taking the trace in the other. So

$$\left\langle \hat{S} \otimes \hat{I} \right\rangle = \mathrm{Tr}(\hat{\rho}(\hat{S} \otimes \hat{I})) = \mathrm{Tr}_A \mathrm{Tr}_B(\hat{\rho}(\hat{S} \otimes \hat{I})) = \mathrm{Tr}_A(\hat{\rho}_A \hat{S}) \ .$$

Similarly

$$\left\langle \hat{I} \otimes \hat{S} \right\rangle = \mathrm{Tr}_B(\hat{\rho}_B \hat{S}) \ .$$

Also, recall that for any Bell state both reduced density matrices are equal to $\frac{1}{2}\hat{I}$. This means that is all cases we have the same result

$$\left\langle \hat{S}\otimes\hat{I}\right\rangle = \left\langle \hat{I}\otimes\hat{S}\right\rangle = \frac{1}{2}\mathrm{Tr}(\hat{S}) \ .$$

The above is true for any operator $\hat{S}$, but in this specific case we also know that $\mathrm{Tr}(\hat{S}) = 0$.

(e) Consider

$$\begin{aligned}\sqrt{2}\hat{S}\otimes\sigma_3\left|\beta_{xy}\right\rangle &= \hat{S}\left|0\right\rangle\otimes\sigma_3\left|y\right\rangle + (-1)^x\hat{S}\left|1\right\rangle\otimes\sigma_3\left|\overline{y}\right\rangle \\ &= (-1)^y\left(\hat{S}\left|0\right\rangle\otimes\left|y\right\rangle - (-1)^x\hat{S}\left|1\right\rangle\otimes\left|\overline{y}\right\rangle\right)\end{aligned}$$

So we see that

$$2\left\langle \hat{S}\otimes\sigma_3\right\rangle = \left\langle\beta_{xy}\right|\hat{S}\otimes\sigma_3\left|\beta_{xy}\right\rangle = (-1)^y\left(\left\langle 0\right|\hat{S}\left|0\right\rangle - \left\langle 1\right|\hat{S}\left|1\right\rangle\right)$$

The above didn't use any properties of $\hat{S}$. It is easy to now evaluate explicitly, noting that only σ_3 contributes since the other σ-matrices have vanishing diagonal components. Alternatively, note that we can also use $\sigma_3\left|0\right\rangle = \left|0\right\rangle$ and $\sigma_3\left|1\right\rangle = -\left|1\right\rangle$ to write

$$2\left\langle \hat{S}\otimes\sigma_3\right\rangle = (-1)^y\left(\left\langle 0\right|\hat{S}\sigma_3\left|0\right\rangle + \left\langle 1\right|\hat{S}\sigma_3\left|1\right\rangle\right) = (-1)^y\mathrm{Tr}(\hat{S}\sigma_3) \ .$$

Then, since $\mathrm{Tr}(\sigma_i\sigma_j) = 2\delta_{ij}$ we see that

$$\left\langle \hat{S}\otimes\sigma_3\right\rangle = (-1)^y n_3 \ .$$

Similarly, note that since $\sigma_1\left|0\right\rangle = \left|1\right\rangle$ and $\sigma_1\left|1\right\rangle = \left|0\right\rangle$ we can write $\sigma_1\left|y\right\rangle = \left|\overline{y}\right\rangle$ so

$$\begin{aligned}\sqrt{2}\hat{S}\otimes\sigma_1\left|\beta_{xy}\right\rangle &= \hat{S}\left|0\right\rangle\otimes\sigma_1\left|y\right\rangle + (-1)^x\hat{S}\left|1\right\rangle\otimes\sigma_1\left|\overline{y}\right\rangle \\ &= \hat{S}\left|0\right\rangle\otimes\left|\overline{y}\right\rangle + (-1)^x\hat{S}\left|1\right\rangle\otimes\left|y\right\rangle \\ &= \hat{S}\sigma_1\left|1\right\rangle\otimes\left|\overline{y}\right\rangle + (-1)^x\hat{S}\sigma_1\left|0\right\rangle\otimes\left|y\right\rangle \\ &= (-1)^x(\hat{S}\sigma_1\otimes\hat{I})\left|\beta_{xy}\right\rangle \ .\end{aligned}$$

As above we can then easily calculate

$$\left\langle \hat{S}\otimes\sigma_1\right\rangle = (-1)^x\frac{1}{2}\mathrm{Tr}(\hat{S}\sigma_1) = (-1)^x n_1 \ .$$

For completeness, note that we can calculate the similar expression with σ_2. Using $\sigma_2\left|0\right\rangle = i\left|1\right\rangle$ and $\sigma_2\left|1\right\rangle = -i\left|0\right\rangle$ we can write $\sigma_2\left|y\right\rangle = i(-1)^y\left|\overline{y}\right\rangle$ so

$$\begin{aligned}\sqrt{2}\hat{S}\otimes\sigma_2\left|\beta_{xy}\right\rangle &= \hat{S}\left|0\right\rangle\otimes\sigma_2\left|y\right\rangle + (-1)^x\hat{S}\left|1\right\rangle\otimes\sigma_2\left|\overline{y}\right\rangle \\ &= i(-1)^y\left(\hat{S}\left|0\right\rangle\otimes\left|\overline{y}\right\rangle - (-1)^x\hat{S}\left|1\right\rangle\otimes\left|y\right\rangle\right) \\ &= -(-1)^y\left(\hat{S}\sigma_2\left|1\right\rangle\otimes\left|\overline{y}\right\rangle + (-1)^x\hat{S}\sigma_2\left|0\right\rangle\otimes\left|y\right\rangle\right) \\ &= -(-1)^{x+y}(\hat{S}\sigma_2\otimes\hat{I})\left|\beta_{xy}\right\rangle \ .\end{aligned}$$

As above we can then easily calculate

$$\left\langle \hat{S}\otimes\sigma_2\right\rangle = -(-1)^{x+y}\frac{1}{2}\mathrm{Tr}(\hat{S}\sigma_2) = -(-1)^{x+y}n_2 \ .$$

6.2. Suppose a device performs a unitary transformation on the tensor product of any **PROBLEMS CLASS 4**

input state with a fixed state $|\Omega\rangle$. For two different input states $|\psi\rangle$ and $|\phi\rangle$ the device seems to be a quantum copier, i.e. it maps

$$|\psi\rangle \otimes |\Omega\rangle \to |\psi\rangle \otimes |\psi\rangle \quad \text{and} \quad |\phi\rangle \otimes |\Omega\rangle \to |\phi\rangle \otimes |\phi\rangle .$$

Show that this is only possible if the two states $|\psi\rangle$ and $|\phi\rangle$ are either the same or orthogonal.

Solution:

Unitary transformations preserve inner products. Therefore the norm of $|\psi\rangle \otimes |\Omega\rangle$ must equal the norm of $|\psi\rangle \otimes |\psi\rangle$, which means that $|\psi\rangle$ and $|\Omega\rangle$ have the same norm. Similarly we see that $|\phi\rangle$ also. has the same norm.

Now compare the inner product of the two states before and after the transformation, i.e.

$$\langle \psi | \phi \rangle \langle \Omega | \Omega \rangle = (\langle \psi | \phi \rangle)^2$$

so either $\langle \psi | \phi \rangle = 0$, in which case $|\psi\rangle$ and $|\phi\rangle$ are orthogonal, or $\langle \psi | \phi \rangle = \langle \Omega | \Omega \rangle$. Since the states all have the same norm, the final relation is only possible if $|\psi\rangle = |\phi\rangle$.

6.3. Show that the Shannon entropy of a random variable which can take N different values is maximised if and only if the probability distribution is uniform. Calculate the maximum Shannon entropy (for a given fixed N).

Solution:

Recall that the Shannon entropy is

$$H = -\sum_{i=1}^{N} p_i \log p_i$$

where the p_i are the probabilities of each possible outcome. Note that these are not N independent variables since we have the constraint $\sum_{i=1}^{N} p_i = 1$. There are two obvious ways to look for the maximum value. We could substitute $p_N = 1 - \sum_{i=1}^{N-1} p_i$ and then set $\frac{\partial H}{\partial p_i} = 0$ for $i \in \{1, 2, \ldots, N-1\}$ or we could instead introduce a Lagrange multiplier λ and set $\frac{\partial \tilde{H}}{\partial \lambda} = 0$ and $\frac{\partial \tilde{H}}{\partial p_i} = 0$ for $i \in \{1, 2, \ldots, N\}$ where

$$\tilde{H} = H + \lambda \left(\sum_{i=1}^{N} p_i - 1 \right) .$$

Using the second method we find

$$\begin{aligned} 0 = \frac{\partial \tilde{H}}{\partial \lambda} &= \sum_{i=1}^{N} p_i - 1 \\ 0 = \frac{\partial \tilde{H}}{\partial p_i} &= -\log p_i - \frac{1}{\ln 2} + \lambda \end{aligned}$$

The second equation shows that all p_i have the same value, and so from the first equation (imposing the constraint on the total probability) we see that $p_i = 1/N$. In this case we find

$$H = -\sum_{i=1}^{N} \frac{1}{N} \log \left(\frac{1}{N} \right) = \log N .$$

You could check that this is really a maximum and not some other turning point. A simple argument is that since we only found one turning point, the only other possibility for a maximum is at the boundary of the region on parameter space, i.e.

where some of the $p_i = 0$. (This includes the case where one $p_i = 1$ since then all other p_i must vanish.) However, these give lower values for H, $\log(N - M)$ for $M \in \{1, 2, \ldots, N - 1\}$.

6.4. Suppose we have messages encoded as strings of N bits, where each bit can be considered a random variable having value 1 with probability p.

(a) Calculate the Shannon entropy of a single bit.

(b) What is the Shannon entropy of a string of N bits? (Recall that if X and Y are independent random variables, $H(X, Y) = H(X) + H(Y)$.)

(c) Suppose $N = 1000$ and $p = 3/4$. What is the minimum average length of message which could contain the same information? (You do not have to give a specific method to encode the message.)

Solution:

(a) We have two possible outcomes, 1 with probability p and 0 with probability $1 - p$. Therefore the Shannon entropy is

$$H = -p \log p - (1 - p) \log(1 - p) \ .$$

(b) Let's consider a string of N bits, X_N, as a string of $N - 1$ bits, X_{N-1}, plus a single bit, X_1. Since the bits are independent

$$H(X_N) \equiv H(X_{N-1}, X_1) = H(X_{N-1}) + H(X_1); .$$

So, by induction we find

$$H(X_N) = NH(X_1) = -Np \log p - N(1 - p) \log(1 - p)$$

using the result in part (a) for $H(X_1)$.

(c) For $p = 3/4$ we find

$$H(X_{1000}) = -750 \log(3/4) - 250 \log(1/4) = 750 \times 2 - 750 \log 3 + 250 \times 2 = 2000 - 750 \log 3 \simeq 811.3$$

so such a message could on average be encoded using only about 811.3 bits.

6.5. Suppose two bits are described by random variables X and Y.

(a) Calculate the joint Shannon entropy, the conditional entropy of X given Y, the conditional entropy of Y given X, and the mutual information in each of the following cases, where (X, Y) can take (with equal probability) only the values:

(i) $(0, 0), (0, 1), (1, 0), (1, 1)$

(ii) $(0, 1), (1, 0), (1, 1)$

(iii) $(0, 1), (1, 0)$

(iv) $(0, 1), (1, 1)$

(b) Calculate the relative entropy for the above probability distributions of

- Case (iii) to case (ii)
- Case (iii) to case (i)
- Case (ii) to case (i)
- Case (iv) to case (ii)
- Case (iv) to case (i)

Solution:

1. Let's use notation $p_x = P(X=x)$, $p_y = P(Y=y)$ and $p_{x,y} = P(X=x \,\&\, Y=y)$, where for the different cases we have the following probabilities

Case	$P(X=0)$	$P(X=1)$	$P(Y=1)$	$P(Y=1)$	$p_{0,0}$	$p_{0,1}$	$p_{1,0}$	$p_{1,1}$
(i)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
(ii)	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
(iii)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	0
(iv)	$\frac{1}{2}$	$\frac{1}{2}$	0	1	0	$\frac{1}{2}$	0	$\frac{1}{2}$

   1. First let's calculate the Shannon entropy for X and for Y, then the joint entropy. Remember that for these calculations we take logarithms in base 2.

$$\begin{aligned}
H(X) &\equiv -\sum_x p_x \log(p_x) = -2 \times \frac{1}{2}\log\left(\frac{1}{2}\right) = 1 \\
H(Y) &\equiv -\sum_y p_y \log(p_y) = -2 \times \frac{1}{2}\log\left(\frac{1}{2}\right) = 1 \\
H(X,Y) &\equiv -\sum_{x,y} p_{x,y} \log(p_{x,y}) = -4 \times \frac{1}{4}\log\left(\frac{1}{4}\right) = 2
\end{aligned}$$

   We can now calculate the conditional entropy of X given Y, $H(X|Y)$ and similarly of Y given X, $H(Y|X)$

$$\begin{aligned}
H(X|Y) &\equiv H(X,Y) - H(Y) = 2 - 1 = 1 \\
H(Y|X) &\equiv H(X,Y) - H(X) = 2 - 1 = 1
\end{aligned}$$

   noting that $H(X,Y) = H(Y,X)$. Finally, the mutual information is

$$H(X:Y) \equiv H(X) + H(Y) - H(X,Y) = 1 + 1 - 2 = 0\,.$$

   2.

$$\begin{aligned}
H(X) &= -\frac{1}{3}\log\left(\frac{1}{3}\right) - \frac{2}{3}\log\left(\frac{2}{3}\right) = \log(3) - \frac{2}{3} \\
H(Y) &= -\frac{1}{3}\log\left(\frac{1}{3}\right) - \frac{2}{3}\log\left(\frac{2}{3}\right) = \log(3) - \frac{2}{3} \\
H(X,Y) &= -3 \times \frac{1}{3}\log\left(\frac{1}{3}\right) = \log(3) \\
H(X|Y) &= \log(3) - \left(\log(3) - \frac{2}{3}\right) = \frac{2}{3} \\
H(Y|X) &= \log(3) - \left(\log(3) - \frac{2}{3}\right) = \frac{2}{3} \\
H(X:Y) &= \log(3) - \frac{4}{3}
\end{aligned}$$

3.

$$\begin{aligned} H(X) &= -2 \times \frac{1}{2} \log\left(\frac{1}{2}\right) = 1 \\ H(Y) &= -2 \times \frac{1}{2} \log\left(\frac{1}{2}\right) = 1 \\ H(X,Y) &= -2 \times \frac{1}{2} \log\left(\frac{1}{2}\right) = 1 \\ H(X|Y) &= 1 - 1 = 0 \\ H(Y|X) &= 1 - 1 = 0 \\ H(X:Y) &= 1 + 1 - 1 = 1 \end{aligned}$$

4.

$$\begin{aligned} H(X) &= -2 \times \frac{1}{2} \log\left(\frac{1}{2}\right) = 1 \\ H(Y) &= -1 \log(1) = 0 \\ H(X,Y) &= -2 \times \frac{1}{2} \log\left(\frac{1}{2}\right) = 1 \\ H(X|Y) &= 1 - 0 = 1 \\ H(Y|X) &= 1 - 1 = 0 \\ H(X:Y) &= 1 + 0 - 1 = 0 \end{aligned}$$

2. For the relative entropy let's label the probability distribution of case A, $p_{x,y}$ and the probability distribution of case B, $q_{x,y}$. Then the relative entropy of case A to case B is

$$H(A||B) \equiv \sum_{x,y} p_{x,y} \left(\log(p_{x,y}) - \log(q_{x,y})\right) \ .$$

Note that it is not symmetric. We then find

$$\begin{aligned} H(iii||ii) &= 2 \times \frac{1}{2} \left(\log\left(\frac{1}{2}\right) - \log\left(\frac{1}{3}\right)\right) = \log(3) - 1 \\ H(iii||i) &= 2 \times \frac{1}{2} \left(\log\left(\frac{1}{2}\right) - \log\left(\frac{1}{4}\right)\right) = -1 + 2 = 1 \\ H(ii||i) &= 3 \times \frac{1}{3} \left(\log\left(\frac{1}{3}\right) - \log\left(\frac{1}{4}\right)\right) = 2 - \log(3) \\ H(iv||ii) &= 2 \times \frac{1}{2} \left(\log\left(\frac{1}{2}\right) - \log\left(\frac{1}{3}\right)\right) = \log(3) - 1 \\ H(iv||i) &= 2 \times \frac{1}{2} \left(\log\left(\frac{1}{2}\right) - \log\left(\frac{1}{4}\right)\right) = 1 \end{aligned}$$

6.6. Consider a qubit with density matrix **PROBLEMS CLASS 4**

$$\rho = \frac{1}{2} \begin{pmatrix} 1+z & 0 \\ 0 & 1-z \end{pmatrix}$$

(a) Calculate the von Neumann entropy $S(\rho)$ as a function of z.

(b) Show that $S(\rho)$ is a monotonic function for $z \in [0,1]$ and find the minimum and maximum values of the entropy.

(c) Calculate the entanglement entropy, as a function of θ, of the state

$$|\Psi\rangle = \cos\theta\,|0\rangle \otimes |0\rangle + \sin\theta\,|1\rangle \otimes |1\rangle \ .$$

Solution:

(a) Since ρ is diagonal we can easily calculate

$$\begin{aligned} S(\rho) &= -\mathrm{Tr}\left(\rho\log\rho\right) = -\frac{1+z}{2}\log\left(\frac{1+z}{2}\right) - \frac{1-z}{2}\log\left(\frac{1-z}{2}\right) \\ &= 1 - \frac{1}{2}(1+z)\log(1+z) - \frac{1}{2}(1-z)\log(1-z) \end{aligned}$$

(b) We can then easily find

$$\frac{dS}{dz} = -\frac{1}{2}\log(1+z) - \frac{1}{2\ln 2} + \frac{1}{2}\log(1-z) + \frac{1}{2\ln 2} = \frac{1}{2}\log\left(\frac{1-z}{1+z}\right) .$$

For the given range of z, this quantity vanishes at the endpoint $z = 0$ and is negative for all other values of z, so the function is monotonic with maximum $S = 1$ at $z = 0$ and minimum

$$S = 1 - \frac{1}{2}2\log 2 - \frac{1}{2}0\log 0 = 1 - 1 - 0 = 0$$

with our usual understanding that $0\log 0 = 0$ by considering it as the limit $\lim_{p\to 0} p\log p$.

(c) The entanglement entropy is the von Neumann entropy of the reduced density matrix (in either subsystem). Here we have

$$\hat{\rho}_A = \mathrm{Tr}_B\left(\cos^2\theta\,|0\rangle\langle 0|\otimes|0\rangle\langle 0| + \sin^2\theta\,|1\rangle\langle 1|\otimes|1\rangle\langle 1|\right) = \cos^2\theta\,|0\rangle\langle 0| + \sin^2\theta\,|1\rangle\langle 1|$$

Therefore the entanglement entropy is

$$S(A) = S(\hat{\rho}_A) = -\cos^2\theta\log(\cos^2\theta) - \sin^2\theta\log(\sin^2\theta) .$$

6.7. Calculate the von Neumann entropy of

(i) $\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$

(ii) The ensemble of $|0\rangle$ and $|1\rangle$ with equal probabilities.

and calculate the relative entropy of state (i) to state (ii).

Solution:

The state (i) is a pure state so its von Neumann entropy is 0. If you want to see this explicitly, work in the orthonormal basis $\{|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}\}$ and represent $|+\rangle$ by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|-\rangle$ by $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then the state (i) is just $|+\rangle$ so its density matrix is

$$\rho = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1\ \ 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} .$$

Therefore we have

$$S(\rho) = -1\log 1 - 0\log 0 = 0 .$$

For the mixed state (ii) we can work in the usual basis, representing $|0\rangle$ by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

and $|1\rangle$ by $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then we easily find

$$\rho = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}(1 \;\; 0) + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}(0 \;\; 1) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

giving the von Neumann entropy

$$S(\rho) = -\frac{1}{2}\log\left(\frac{1}{2}\right) - \frac{1}{2}\log\left(\frac{1}{2}\right) = \log 2 = 1 \; .$$

Now, to calculate the relative entropy we use the definition

$$S(\rho_1||\rho_2) = \mathrm{Tr}(\rho_1 \log \rho_1) - \mathrm{Tr}(\rho_1 \log \rho_2) \; .$$

Here we take ρ_1 to be the density matrix for state (i), and ρ_2 for state (ii). The first term is just minus the von Neumann entropy of state (i) so we know from above that this is 0. To calculate the second term we must write the two density matrices in the same representation. There are two natural ways to do this. We could note that since ρ_2 is proportional to the identity matrix in one orthonormal basis, it is the same in any orthonormal basis. However, to illustrate the method more generally, let's work in the basis representing $|0\rangle$ by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle$ by $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. note that when choosing the basis, it is important that ρ_2 is diagonal so that we can calculate $\log \rho_2$ be just taking the logarithm of each diagonal element. In this representation state (i) is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ so

$$\rho_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(1 \;\; 1) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \; .$$

Then we can calculate

$$\begin{aligned} \mathrm{Tr}(\rho_1 \log \rho_2) &= \mathrm{Tr}\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \log(1/2) & 0 \\ 0 & \log(1/2) \end{pmatrix}\right] \\ &= -\frac{1}{2}\mathrm{Tr}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = -1 \end{aligned}$$

which gives

$$S(\rho_1||\rho_2) = 0 - (-1) = 1 \; .$$

6.8. Consider the state

$$|\Psi\rangle = \cos\theta\,|0\rangle \otimes |0\rangle + \sin\theta\,|1\rangle \otimes |1\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B.$$

Calculate the relative entropy of this state to the state with density matrix

$$\hat{\rho} = \hat{\rho}_A \otimes \hat{\rho}_B$$

where $\hat{\rho}_A$ and $\hat{\rho}_B$ are the reduced density matrices in each system.

Solution:
Here we want to calculate

$$S(\hat{\rho}||\hat{\rho}_A \otimes \hat{\rho}_B) = \mathrm{Tr}(\hat{\rho}\log\hat{\rho}) - \mathrm{Tr}(\hat{\rho}\log(\hat{\rho}_A \otimes \hat{\rho}_B)) \; .$$

Since $\hat{\rho}$ is the density matrix for the pure state $|\Psi\rangle$ we know that the first term, which is $-S(\hat{\rho})$, is 0.

For the second term we need to calculate the reduced density matrices. This is easy and by the symmetry both have the same form

$$\cos^2\theta\,|0\rangle\,\langle 0| + \sin^2\theta\,|1\rangle\,\langle 1|$$

so we find

$$\begin{aligned}\hat{\rho}_A\otimes\hat{\rho}_B &= \cos^4\theta(|0\rangle\otimes|0\rangle)(\langle 0|\otimes\langle 0|) + \cos^2\theta\sin^2\theta(|0\rangle\otimes|1\rangle)(\langle 0|\otimes\langle 1|) + \\ &\quad + \cos^2\theta\sin^2\theta(|1\rangle\otimes|0\rangle)(\langle 1|\otimes\langle 0|) + \sin^4\theta(|1\rangle\otimes|1\rangle)(\langle 1|\otimes\langle 1|)\ .\end{aligned}$$

This means that in the standard basis for the 2-qubit system

$$\rho_A\otimes\rho_B = \mathrm{diag}(\cos^4\theta, \cos^2\theta\sin^2\theta, \cos^2\theta\sin^2\theta, \sin^4\theta)\ .$$

On the other hand

$$\begin{aligned}\hat{\rho} &= |\Psi\rangle\,\langle\Psi| = \cos^2\theta(|0\rangle\otimes|0\rangle)(\langle 0|\otimes\langle 0|) + \cos\theta\sin\theta(|0\rangle\otimes|0\rangle)(\langle 1|\otimes\langle 1|) + \\ &\quad + \cos\theta\sin\theta(|1\rangle\otimes|1\rangle)(\langle 0|\otimes\langle 0|) + \sin^2\theta(|1\rangle\otimes|1\rangle)(\langle 1|\otimes\langle 1|)\ .\end{aligned}$$

So, in 4×4 matrix form we have

$$\begin{aligned}S(\hat{\rho}||\hat{\rho}_A\otimes\hat{\rho}_B) &= -\mathrm{Tr}(\hat{\rho}\log(\hat{\rho}_A\otimes\hat{\rho}_B)) \\ &= -\mathrm{Tr}\left[\begin{pmatrix}\cos^2\theta & 0 & 0 & \cos\theta\sin\theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \cos\theta\sin\theta & 0 & 0 & \sin^2\theta\end{pmatrix}\times\right. \\ &\quad\left.\times\begin{pmatrix}\log(\cos^4\theta) & 0 & 0 & 0 \\ 0 & \log(\cos^2\theta\sin^2\theta) & 0 & 0 \\ 0 & 0 & \log(\cos^2\theta\sin^2\theta) & 0 \\ 0 & 0 & 0 & \log(\sin^4\theta)\end{pmatrix}\right] \\ &= -2\left(\cos^2\theta\log(\cos^2\theta) + \sin^2\theta\log(\sin^2\theta)\right)\ .\end{aligned}$$