ERRATA, 19 June 2009_V3.0

Additions or modifications to previous Errata versions : **P.64**, **P.574-575**, **Index**

See latest Errata update in
http://www.cambridge.org/9780521881715

ORIGINAL VERSION	CORRECTED VERSION
P.4 **Rolling dice (game 2)** (…) For instance, $p(x=2)=p(x=22)=1/36=0.028$, (…)	**P.4** **Rolling dice (game 2)** (…) For instance, $p(x=2)=p(x=12)=1/36=0.028$, (…)
P.13 **Taking exams** (…) $p(John\ \ pass)\times p(Peter\ \ pass)=0.7\times 0.3=0.21$	**P.13** **Taking exams** (…) $p(John\ \ pass)\times p(Peter\ \ pass)=0.7\times 0.4=0.28$
P.15 **Party meetings** (…) But if Alice is first seen there, it is much more likely to see Bob ($p(Bob\|Alice)=0.7>p(Bob)=0.4$) and (…)	**P.15** **Party meetings** (…) But if Alice is first seen there, it is much more likely to see Bob ($p(Bob\|Alice)=0.5>p(Bob)=0.4$) and (…)

[5] The assumption according to which a system may be accurately or completely defined through a quantum state $\left|\psi\right\rangle$, is equivalent to assume that the system is *closed*, meaning that it is not coupled nor entangled to any other unknown external system.

P.64

$$
\begin{aligned}
H(X) &= -\sum_{x\in X} p(x)\log p(x) \\
&= -x_1 \log p(x_1) - x_2 \log p(x_2) \qquad (4.13)\\
&= -q\log q - (1-q)\log(1-q) \equiv f(q)
\end{aligned}
$$

P.341

Density operator or matrix

Let $V^n = \{|x_i\rangle\} = \{|x_1\rangle, |x_2\rangle, \ldots |x_n\rangle\}$ an orthonormal basis for the space of quantum states $|\psi\rangle$, i.e. verifying $\langle x_i | x_j \rangle = \delta_{ij}$. In this basis, the states have a unique decomposition, which takes the form

$$
|\psi\rangle = x_1|x_1\rangle + x_2|x_2\rangle + \ldots x_n|x_n\rangle = \sum_{i=1}^{n} x_i|x_i\rangle \qquad (17.34)
$$

where x_i $(i = 1\ldots n)$ represent the complex coordinates. If the modulus/length of $|\psi\rangle$ is unity, we have

$$
\langle\psi|\psi\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_i x_j \langle x_i | x_j\rangle = \sum_{i=1}^{n} |x_i|^2 = 1 \qquad (17.35)
$$

As we have seen in Chapter 16, for qubits, the number $p_i = |x_i|^2$, represents the probability to find (or measure) the state $|\psi\rangle$ in the basis state $|x_i\rangle$. Hence the

P.64

$$
\begin{aligned}
H(X) &= -\sum_{x\in X} p(x)\log p(x) \\
&= -p(x_1)\log p(x_1) - p(x_2)\log p(x_2) \qquad (4.13)\\
&= -q\log q - (1-q)\log(1-q) \equiv f(q)
\end{aligned}
$$

P.341

Density operator or matrix

Let $|\psi\rangle$ a pure state, meaning a determinitstic or well-defined quantum state with a probability of unity. Let $V^n = \{|x_i\rangle\} = \{|x_1\rangle, |x_2\rangle, \ldots |x_n\rangle\}$ an orthonormal basis for the space of quantum states $|\psi\rangle$. In this basis, the states have the unique decomposition

$$
|\psi\rangle = x_1|x_1\rangle + x_2|x_2\rangle + \ldots x_n|x_n\rangle = \sum_{i=1}^{n} x_i|x_i\rangle \qquad (17.34)
$$

where x_i $(i = 1\ldots n)$ are complex coordinates. If the modulus/length of $|\psi\rangle$ is unity, we have

$$
\langle\psi|\psi\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_i x_j \langle x_i | x_j\rangle = \sum_{i=1}^{n} |x_i|^2 = 1 \qquad (17.35)
$$

We can now define a Hermitian operator called density operator of density matrix , ρ, associated to the pure state $|\psi\rangle$ as follows :

[4] By application of the property $\lim_{x\to 0}(x\log x) = 0$, which makes the function $x\log x$ analytically defined for any real $x \geq 0$.

[6] The assumption according to which a system may be accurately or completely defined through a quantum state $|\psi\rangle$, is equivalent to assume that the system is *closed*, meaning that it is not coupled nor entangled to any other unknown external system.

coordinates x_i represent complex amplitude probabilities. We can now define the density operator/matrix associated to the state $|\psi\rangle$ as follows:

$$\rho = |x_1|^2 |x_1\rangle\langle x_1| + |x_2|^2 |x_2\rangle\langle x_2| + \ldots |x_n|^2 |x_n\rangle\langle x_n| = \sum_{i=1}^{n} |x_i|^2 |x_i\rangle\langle x_i| \qquad (17.36)$$

The density matrix is diagonal, since its elements ρ_{ij} are given by

$$\rho_{ij} = \langle x_i|\rho|x_j\rangle = = \langle x_i \left| \left(\sum_{k=1}^{n} |x_k|^2 |x_k\rangle\langle x_k| \right) \right| x_j\rangle = \sum_{k=1}^{n} |x_k|^2 \langle x_i|x_k\rangle\langle x_k|x_j\rangle$$

$$= \sum_{k=1}^{n} |x_k|^2 \delta_{ik}\delta_{kj} \equiv |x_i|^2 \delta_{ij} \qquad (17.37)$$

Hence, the density matrix operator takes the diagonal matrix representation:

$$\rho = \begin{pmatrix} |x_1|^2 & 0 & \cdots & 0 \\ 0 & |x_2|^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & |x_n|^2 \end{pmatrix} = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & p_n \end{pmatrix} \qquad (17.38)$$

P.342

It is immediately noted that the trace of the density matrix is unity since

$$\rho = |\psi\rangle\langle\psi| \qquad (17.36)$$

According to the decomposition in eq.(17.34), we obtain

$$\rho = \left(\sum_{i=1}^{n} x_i |x_i\rangle \right) \left(\sum_{j=1}^{n} \bar{x}_j \langle x_j| \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{x}_j |x_i\rangle\langle x_j| \qquad (17.37)$$

The density-matrix elements are thus $\rho_{ij} = \langle i|\rho|j\rangle = x_i \bar{x}_j$, with $\rho_{ij} = \bar{\rho}_{ji}$ since ρ is Hermitian. Given this definition and the property in eq.(17.35), is immediately noted that the trace of the density matrix is unity, since

$$tr(\rho) = \sum_{i=1}^{n} \rho_{ii} = \sum_{i=1}^{n} |x_i|^2 = 1 \qquad (17.38)$$

Consider next the case of a statistical mixture of pure states. Let $\{|\psi_1\rangle, |\psi_2\rangle, \ldots |\psi_N\rangle\}$ an sensemble of N pure states, not necessarily orthogonal to each other. Assume that to each pure state $|\psi_k\rangle$ be associated a probability p_k, meaning that the system state $|\psi\rangle$ is most generally defined by the linear superposition :

$$|\psi\rangle = \alpha_1 |\psi_1\rangle + \alpha_2 |\psi_2\rangle + \ldots \alpha_N |\psi_N\rangle = \sum_{k=1}^{N} \alpha_k |\psi_k\rangle \qquad (17.39a)$$

where α_k are complex numbers verifying $|\alpha_k|^2 = p_k$. The state $|\psi\rangle$ is undefined

---

P.342

since there exist an infinity of sets $\{\alpha_k\}_{k=1\ldots N}$ having this property. However, the density matrix formalism makes it possible to exactly and comprehensively describe any statistical mixture of states, through the definition

$$tr(\rho) = \sum_{i=1}^{n} |x_i|^2 = \sum_{i=1}^{n} p_i = 1 \quad (17.39)$$

We now have the tools to make a short hint to quantum information theory. This may also constitute a nice reward for having gone through the lengthy description of Dirac notations!

Let introduce a new operator called $U \log U$, where U is assumed to be diagonal with non-negative coefficients. To calculate the matrix coefficients of $U \log U$, we must first define $\log U$. Assume then a linear operator V, which verifies $U = \exp(V)$, which defines $V = \log U$. Formally, the exponential operator is determined by the infinite series:

$$U = \exp(V) = \sum_{n}^{\infty} \frac{V^n}{n!} \quad (17.40)$$

Since U is diagonal, any of the powers V^n must be diagonal. The diagonal coefficients of U are thus given by $U_{ii} = \exp(V_{ii})$ or $V_{ii} = \log(U_{ii})$. The matrix $W = UV = U \log U$ is also diagonal. It is clear that its coefficients are given by $W_{ii} = U_{ii}V_{ii} = U_{ii}\log(U_{ii})$. This result shows that the matrix $W = U \log U$ is analytically defined for any diagonal matrix U with coefficients $U_{ii} \geq 0$.[4] We conclude that the density matrix $U = \rho$, for which the coefficients are non-negative, is an eligible candidate for the operator $U \log U$. We thus have $(\rho \log \rho)_{ii} = \rho_{ii} \log \rho_{ii}$ and the matrix definition

$$\rho = \sum_{k=1}^{N} p_k |\psi_k\rangle\langle\psi_k| \equiv \sum_{k=1}^{N} p_k \rho_k \quad (17.39b)$$

where ρ_k is the density matrix associated with the pure state $|\psi_k\rangle$. It is immediately noted that the density-matrix trace is unity, since $tr(\rho) == \sum_k p_k tr(\rho_k) = \sum_k p_k \equiv 1$, according to the previous trace property and the definition of probabilities.
Since ρ is Hermitian, it is possible to find an orthonormal basis $W^n = \{|y_i\rangle\}_{i=1\ldots n}$ where ρ is diagonal. Call this diagonal representation $\tilde{\rho}$. Under the basis transformation $\rho \to \tilde{\rho}$, the trace is preserved, according to the property in eq.(17.33), hence $tr(\tilde{\rho}) = tr(\rho) \equiv 1$. The diagonal density matrix $\tilde{\rho}$ is thus defined by

$$\tilde{\rho} = \sum_{i=1}^{n} y_i \bar{y}_i |y_i\rangle\langle y_i| = \sum_{i=1}^{n} |y_i|^2 |y_i\rangle\langle y_i| \equiv \sum_{i=1}^{n} q_i |y_i\rangle\langle y_i| \quad (17.39c)$$

with coefficients q_i being probabilities. Thus $\tilde{\rho}$ defines a statistical mixture of the pure, orthonormal-basis states $|y_i\rangle$.

The above has shown that the density-matrix representations ρ (of mixture $|\psi_i\rangle$) and $\tilde{\rho}$ (of mixture $|y_i\rangle$) are equivalent, apart from their probability coefficients p_i and q_i. For QIT purposes, it is often easier to use the diagonal representation $\tilde{\rho}$, along with the orthonormal basis $\{|y_i\rangle\}$. For simplicity, we shall use for now on the notations $\rho, p_i, |x_i\rangle$ instead of $\tilde{\rho}, q_i, |y_i\rangle$ to refer to the diagonal representation.

Let introduce next a new operator called $U \log U$, where U is assumed diagonal with non-negative coefficients. To define $\log U$, assume then a linear operator V, which verifies $U = \exp(V)$. The exponential operator is determined by the series:

$$\rho\log\rho=\begin{pmatrix} |x_1|^2\log|x_1|^2 & 0 & \cdots & 0 \\ 0 & |x_2|^2\log|x_2|^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & |x_n|^2\log|x_n|^2 \end{pmatrix} = \begin{pmatrix} p_1\log p_1 & 0 & \cdots & 0 \\ 0 & p_2\log p_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & p_n\log p_n \end{pmatrix} \quad (17.41)$$

Finally, we find that the trace of $\rho\log\rho$ is given by the expression:

$$tr(\rho\log\rho)=\sum_{i=1}^{n}|x_i|^2\log|x_i|^2=\sum_{i=1}^{n}p_i\log p_i \quad (17.42)$$

Based on our background of Shannon's information theory (Chapter 4), we can heuristically define an "entropy" H for the quantum state describe by the density matrix ρ under the form

$$H=-\sum_{i=1}^{n}p_i\log p_i \quad (17.43)$$

P.432

Consider now a quantum system. This system may exist in a quantum state $|\psi\rangle$, which we shall assume here to represent a statistical mixture of pure states $|x_i\rangle$.[5] Such pure

$$U=\exp(V)=\sum_{n}^{\infty}\frac{V^n}{n!} \quad (17.40)$$

Since U is diagonal, any of the powers V^n are diagonal. We have $U_{ii}=\exp(V_{ii})$ or $V_{ii}=\log(U_{ii})$. The matrix $W=UV=U\log U$ is also diagonal. Clearly, $W_{ii}=U_{ii}V_{ii}=U_{ii}\log(U_{ii})$. This shows that $W=U\log U$ is analytically defined for any diagonal matrix U with coefficients $U_{ii}\geq 0$.[4] Therefore, $U=\rho$ is an eligible candidate for the operator $U\log U$. We thus have $(\rho\log\rho)_{ii}=\rho_{ii}\log\rho_{ii}$ and

$$\rho\log\rho=\begin{pmatrix} p_1\log p_1 & 0 & \cdots & 0 \\ 0 & p_2\log p_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & p_n\log p_n \end{pmatrix} \quad (17.41)$$

Finally, we find that the trace of $\rho\log\rho$ is given by the expression:

$$tr(\rho\log\rho)=\sum_{i=1}^{n}|x_i|^2\log|x_i|^2=\sum_{i=1}^{n}p_i\log p_i \quad (17.42)$$

Based on our background of Shannon's information theory (Chapter 4), we can heuristically define an "entropy" H for the quantum state described by ρ under the form

$$H=-\sum_{i=1}^{n}p_i\log p_i \quad (17.43)$$

P.432

Consider now a quantum system. As we have seen in Chapter 17, this system may exist in a quantum state $|\psi\rangle$, which most generally is a statistical mixture of pure states $|x_i\rangle$.[6]

we shall assume here to represent a statistical mixture of pure states $|x_i\rangle$.[5] Such pure states, which cannot be defined by any mixture of the other pure states, are orthogonal to each other and have a unity length, such that $\langle x_i | x_j \rangle = \delta_{ij}$. The set of pure states $\{|x_i\rangle\} = \{|x_1\rangle, |x_2\rangle, \ldots |x_n\rangle\}$ thus defines an orthonormal basis for the n-dimensional space V^n of all possible quantum states $|\psi\rangle$ defining the system. Consistently, any state $|\psi\rangle$ of V^n accepts a unique decomposition of the form

$$|\psi\rangle = x_1|x_1\rangle + x_2|x_2\rangle + \ldots x_n|x_n\rangle = \sum_{i=1}^{n} x_i |x_i\rangle \qquad (21.2)$$

where x_i ($i = 1 \ldots n$) are complex coordinates. We may choose to represent the quantum system with a state $|\psi\rangle$ of unity length, i.e. $\langle\psi|\psi\rangle = \sum_i |x_i|^2 = 1$, in which case the real number $p_i = |x_i|^2$, represents the probability to find the state $|\psi\rangle$ in the pure state $|x_i\rangle$. In the quantum world, the system in the quantum state $|\psi\rangle$ thus plays the role of a "random events" source, and naturally the concepts of "information" and "entropy" may be associated to such a system.

To establish such a connection, we need to use the concept of density operator (or density matrix), which was previously introduced in chapter 17. As we have learnt, the density operator/matrix ρ is an alternative way to define a quantum system in a given state $|\psi\rangle$, by means of an operator. Formally,

$$\rho = \sum_{i=1}^{n} p_i |x_i\rangle\langle x_i| \qquad (21.3)$$

where $|x_i\rangle\langle x_i|$ is the projector (or measurement) operator on the basis state $|x_i\rangle$. It is clear that $\rho|x_i\rangle = p_i|x_i\rangle$, which shows that $|x_i\rangle$ is an eigenstate of ρ with associated

We may assume, without loss of generality for most QIT purposes, that these pure states are orthogonal to each other and have a unity length, such that $\langle x_i | x_j \rangle = \delta_{ij}$. The set of pure states $\{|x_i\rangle\} = \{|x_1\rangle, |x_2\rangle, \ldots |x_n\rangle\}$ thus defines an orthonormal basis for the n-dimensional space V^n of all possible quantum states $|\psi\rangle$ defining the system. Consistently, any state $|\psi\rangle$ of V^n accepts a unique decomposition of the form

$$|\psi\rangle = x_1|x_1\rangle + x_2|x_2\rangle + \ldots x_n|x_n\rangle = \sum_{i=1}^{n} x_i |x_i\rangle \qquad (21.2)$$

where x_i ($i = 1 \ldots n$) are complex coordinates. We may choose to represent the quantum system with a state $|\psi\rangle$ of unity length, i.e. $\langle\psi|\psi\rangle = \sum_i |x_i|^2 = 1$. The definition in eq.(21.2) may also correspond to that of a linear superposition of states, with deterministic (well-known) coefficients x_i. In such a case, $|\psi\rangle$ is a pure state. However, if we attribute to the coefficients x_i the meaning of complex amplitude probabilities, then $|\psi\rangle$ is a statistical mixture of states. In this case, the real number $p_i = |x_i|^2$, represents the probability to find the state $|\psi\rangle$ in the pure state $|x_i\rangle$. In the quantum world, the system in quantum state $|\psi\rangle$ thus plays the role of a "random events" source, and naturally the concepts of "information" and "entropy" may be associated to such a system. As observed in Chapter 17, a statistical mixture of states cannot be defined by eq.(21.2), since there exist an infinity of complex sets $\{x_i\}_{i=1..n}$ verifying $|x_i|^2 = p_i$.

Instead, we need to use density operator (or density matrix) formalism previously introduced in that chapter. As we have learnt, the density operator/matrix ρ (or its diagonal representation $\tilde{\rho}$) is the most comprehensive way to define quantum systems. Formally, eigenvalue p_i. In the case where $|\psi\rangle$ is a pure state, e.g. $|\psi\rangle = |x_k\rangle$ with $p_i = \delta_{ik}$, the density operator is simply given by $\rho = |\psi\rangle\langle\psi|$, and only in such a case. As a general definition, a pure state is a state that has 100% probability to be observed in a quantum system, or which is exactly known. A given basis state (for instance $|0\rangle$ or $|1\rangle$ in a 2D space) may or may not be a pure state, according to this condition be or not be fulfilled (see more on this further down).

As we have also seen in chapter 17, the matrix elements ρ_{ij} of the density operator verify $\rho_{ij} = \langle x_i|\rho|x_j\rangle \equiv |x_i|^2 \delta_{ij} = p_i\delta_{ij}$, showing that the matrix is diagonal in the computational basis $\{|x_i\rangle\}$, as expected from the fact that it is the basis of eigenstates. The matrix representation of ρ is thus:

$$\rho = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & p_n \end{pmatrix} \quad (21.4)$$

P.435

This result shows that quantum information, as defined by the VN entropy, represents an incompressible feature in quantum systems, as is also the case for classical information in random events sources, as defined by Shannon entropy. What is the quantum information contained in a qubit?. The answer is straightforward. Assume a qubit of general definition

$$\begin{cases} \rho = \sum_{i=1}^{n} x_i \bar{x}_j |x_i\rangle\langle x_j| \\ \tilde{\rho} = \sum_{i=1}^{n} q_i |y_i\rangle\langle y_i| \end{cases} \quad (21.3)$$

where $|y_i\rangle\langle y_i|$ are the projectors (or measurement) operators on the basis states $|y_i\rangle$ for which the density matrix $\tilde{\rho}$ is diagonal. As we have also seen in chapter 17, the two representations of ρ and $\tilde{\rho}$ are equivalent, except for the probability coefficients p_i, q_i and the basis representations $|x_i\rangle, |y_i\rangle$. For simplicity, we shall use for the diagonal representation the notations $\rho, p_i, |x_i\rangle$ instead of $\tilde{\rho}, q_i, |y_i\rangle$. In this case, the matrix elements ρ_{ij} thus verify $\rho_{ij} = \langle x_i|\rho|x_j\rangle \equiv |x_i|^2 \delta_{ij} = p_i\delta_{ij}$, and we finally have :

$$\rho = \begin{pmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & p_n \end{pmatrix} \quad (21.4)$$

P.435

This result shows that quantum information, as defined by the VN entropy, represents an incompressible feature in quantum systems, as is also the case for classical information in random events sources, as defined by Shannon entropy. What is the quantum information contained in a qubit? The answer is simple, but not that straightforward. Assume indeed a qubit of general definition

$$|q\rangle = \alpha|0\rangle + \beta|1\rangle \qquad (21.10)$$

where $|0\rangle, |1\rangle$ are two pure states in the 2D quantum space V^2, and $p = |\alpha|^2 = 1 - |\beta|^2$. From the definitions in eq.(21.3) and eq.(21.4), we have

$$\rho = |\alpha|^2 |0\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1| = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| \qquad (21.11)$$

$$\rho = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix} \qquad (21.12)$$

and

$$\begin{aligned} S(\rho) &= -\left(|\alpha|^2 \log|\alpha|^2 + |\beta|^2 \log|\beta|^2\right) \\ &= -p\log p - (1-p)\log(1-p) \equiv f(p) \end{aligned} \qquad (21.13)$$

In the result in eq.(21.13), we recognize the Shannon entropy of a two-events source $X = \{0,1\}$, corresponding to the two possible "states" of a classical bit (see eq.(4.13) and eq.(4.14)). The average qubit information is thus equivalent to that of two classical bits, the amount depending upon the weights p_0, p_1 in the state mixture. As previously described in Chapter 4, the function $f(p)$ has a maximum of unity for $p = 1/2$, corresponding to $|\alpha| = |\beta| = 1/\sqrt{2}$ (phases being arbitrary) or a uniform distribution $p_0 = p_1 = 1/2$, and for the VN entropy, $S_{\max}(\rho) \equiv \log 2 = 1$. In this case, the quantum information amounts to exactly two classical bits. The minimum of $f(p)$ is zero, which is reached either when $p_0 = 1, p_1 = 0$ or when $p_0 = 0, p_1 = 1$, meaning that the qubit is in a pure state, i.e. $|q\rangle = |0\rangle$ or $|q\rangle = |1\rangle$, giving $S_{\max}(\rho) \equiv 0$. In this case, there is no quantum information in the system, as there is no information in a single, deterministic classical bit.

It is clear that the VN entropy of a n-qubit (or <u>qunit</u>) in the quantum space V^n always

$$|q\rangle = \alpha|0\rangle + \beta|1\rangle \qquad (21.10)$$

where $|0\rangle, |1\rangle$ are two pure states in the 2D quantum space V^2, and $p = |\alpha|^2 = 1 - |\beta|^2$. As we know, p and $1-p$ represent the probabilities to measure $|q\rangle$ in the state $|0\rangle$ and $|1\rangle$, respectively, corresponding to classical bits measurements "0" and "1". Such measurement outcomes are those of a classical two-events source $X = \{0,1\}$, with <u>Shannon</u> entropy $H(X) = -p\log p - (1-p)\log(1-p) \equiv f(p)$ (see eq.(4.13)). As previously seen in Chapter 4, the entropy $H(X)$ has a maximum of unity ($H = 1$ bit) for a uniform distribution $(p, 1-p) = (1/2, 1/2)$, or $|\alpha| = |\beta| = 1/\sqrt{2}$ and is identically zero ($H = 0$ bit) for the deterministic distribution $(p, 1-p) = (1,0)$ or $(p, 1-p) = (0,1)$. Thus to any qubit can be associated a <u>Shannon entropy</u> $0 \le H(X) \le 1$. What about the <u>VN entropy</u>? Assume that the system is in the state $|0\rangle$ with probability $p = |\alpha|^2$ and in the state $|1\rangle$ with probability $1 - p = |\beta|^2$. From the definition in eq.(21.3) , the density matrix of this statistical mixture is

$$\rho = |\alpha|^2 |0\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1| = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1| \qquad (21.11)$$

$$\rho = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix} \qquad (21.12)$$

and we obtain the VN entropy

$$S(\rho) = -p\log p - (1-p)\log(1-p) \equiv f(p) \qquad (21.13)$$

The <u>quantum information</u> $S(\rho)$ in the statistical mixture ρ is thus identical to its <u>classical information</u> counterpart $H(X)$ for the single qubit measurement, with $0 \le S(\rho) = H(X) \le 1$. It should be emphasized, however, that the system described by the statistical mixture ρ in eq.(21.11) is <u>not</u> equivalent to that defined by the qubit $|q\rangle$ in has the maximum $S_{\max}(\rho) \equiv \log n$, when the system is in the most homogeneous state superposition with a uniform probability distribution $p_i = 1/n$, hence corresponding to maximum quantum information. In the general case, we have $0 \le S(\rho) \le \log n$.

eq.(21.10)! The two can be connected if we introduce $|q'\rangle = \alpha|0\rangle - \beta|1\rangle$. It is then is easily verified that the statistical mixture $\sigma = (|q\rangle\langle q| + |q'\rangle\langle q'|)/2$ is identical to ρ, or $\sigma = \rho$. As for a system with density matrix $\tau = |q\rangle\langle q|$, its VN entropy is identically zero, since it corresponds to a <u>pure state</u> ($|q\rangle$ is a pure state in the orthonormal basis $\{|q\rangle, |q^{\perp}\rangle\}$ where $\langle q^{\perp}|q\rangle = 0$).

It is clear that the VN entropy of statistical mixture of n qubits reaches the maximum $S_{\max}(\rho) \equiv \log n$, corresponding to the uniform probability distribution $p_i = 1/n$. In the general case, we thus have $0 \le S(\rho) = H(X^n) \le \log n$.

Appendix C

p.574

$$\begin{aligned} s_0 &= 1 - \sum_i p_i \\ &= 1 - \sum_{i=1}^{k} = \exp(\lambda - 1)1 - k\exp(\lambda - 1) \\ &= 0 \end{aligned} \qquad \text{(C7)}$$

p.575

$$\frac{df}{dp_j} = \frac{d}{dp_j}\left[H(X) + \lambda\sum_i p_i + \mu\sum_i x_i\, p_i + \right] = 0 \qquad \text{(C10)}$$

or

Appendix C

p.574

$$\begin{aligned} s_0 &= 1 - \sum_i p_i \\ &= 1 - \sum_{i=1}^{k} \exp(\lambda - 1) \\ &= 1 - k\exp(\lambda - 1) \\ &= 0 \end{aligned} \qquad \text{(C7)}$$

p.575

$$\frac{df}{dp_j} = \frac{d}{dp_j}\left[H(X) + \lambda\sum_i p_i + \mu\sum_i x_i\, p_i \right] = 0 \qquad \text{(C10)}$$

or

$$\frac{df}{dp_j} = \left[-\sum_i p_i \log p_i + \lambda \sum_i p_i + \mu \sum_i x_i p_i + \right]$$
$$= -\log p_j - 1 + \lambda + \mu x_j \qquad \text{(C11)}$$
$$= 0$$

which yields

$$\frac{df}{dp_j} = \left[-\sum_i p_i \log p_i + \lambda \sum_i p_i + \mu \sum_i x_i p_i + \right]$$
$$= -\log p_j - 1 + \lambda + \mu x_j \qquad \text{(C12)}$$
$$= 0$$

$$\frac{df}{dp_j} = \frac{d}{dp_j} \left[-\sum_i p_i \log p_i + \lambda \sum_i p_i + \mu \sum_i x_i p_i \right] \qquad \text{(C11)}$$
$$= 0$$

which yields

$$\frac{df}{dp_j} = -\log p_j - 1 + \lambda + \mu x_j = 0 \qquad \text{(C12)}$$

Index *(include following new entries)*

See latest Errata update in

http://www.cambridge.org/9780521881715