Entanglement of Formation of an Arbitrary State of Two Qubits

William K. Wootters

Department of Physics, Williams College, Williamstown, Massachusetts 01267

(Received 12 September 1997)

The entanglement of a pure state of a pair of quantum systems is defined as the entropy of either member of the pair. The entanglement of formation of a mixed state ρ is the minimum average entanglement of an ensemble of pure states that represents ρ. An earlier paper conjectured an explicit formula for the entanglement of formation of a pair of *binary* quantum objects (qubits) as a function of their density matrix, and proved the formula for special states. The present paper extends the proof to arbitrary states of this system and shows how to construct entanglement-minimizing decompositions. [S0031-9007(98)05470-2]

PACS numbers: 03.67.–a, 03.65.Bz, 89.70.+c

Entanglement is the quantum mechanical property that Schrödinger singled out many decades ago as "*the* characteristic trait of quantum mechanics" [1] and that has been studied extensively in connection with Bell's inequality [2]. A pure state of a pair of quantum systems is called entangled if it is unfactorizable, as is the case, for example, for the singlet state of two spin-$\frac{1}{2}$ particles, $(1/\sqrt{2})\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. A *mixed* state is entangled if it cannot be represented as a mixture of factorizable pure states. In the last couple of years a good deal of work has been devoted to finding physically motivated quantitative measures of entanglement, particularly for mixed states of a bipartite system [3–5]. Perhaps the most basic of these measures is the *entanglement of formation*, which is intended to quantify the resources needed to create a given entangled state [5].

Having a well justified and mathematically tractable measure of entanglement is likely to be of value in a number of areas of research, including the study of decoherence in quantum computers [6] and the evaluation of quantum cryptographic schemes [7]. Unfortunately, most proposed measures of entanglement involve extremizations which are difficult to handle analytically, the entanglement of formation being no exception to this rule. However, in the special case of entanglement between two *binary* quantum systems such as the spin of a spin-$\frac{1}{2}$ particle or the polarization of a photon—systems that are generically called "qubits"—an explicit formula for the entanglement of formation has recently been conjectured and has been proved for a special class of density matrices [8]. In this Letter we prove the formula for arbitrary states of two qubits.

The entanglement of formation is defined as follows [5]. Given a density matrix ρ of a pair of quantum systems A and B, consider all possible pure-state decompositions of ρ, that is, all ensembles of states $|\psi_i\rangle$ with probabilities p_i, such that

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \,. \tag{1}$$

For each pure state, the entanglement E is defined as the entropy of either of the two subsystems A and B [3]:

$$E(\psi) = -\mathrm{Tr}(\rho_A \log_2 \rho_A) = -\mathrm{Tr}(\rho_B \log_2 \rho_B) \,. \tag{2}$$

Here ρ_A is the partial trace of $|\psi\rangle\langle\psi|$ over subsystem B, and ρ_B has a similar meaning. The entanglement of formation of the mixed state ρ is then defined as the average entanglement of the pure states of the decomposition, minimized over all decompositions of ρ:

$$E(\rho) = \min \sum_i p_i E(\psi_i) \,. \tag{3}$$

The basic equation (2) is justified by the physical interconvertibility of a collection of pairs in an arbitrary pure state $|\psi\rangle$ and a collection of pairs in the standard singlet state, the asymptotic conversion ratio being given by $E(\psi)$ [3]. The central claim of this Letter is that for a pair of qubits, the minimum value specified in Eq. (3) can be expressed as an explicit function of ρ, which we develop in the next few paragraphs. For ease of expression we will usually refer to the entanglement of formation simply as "the entanglement."

Our formula for entanglement makes use of what can be called the "spin flip" transformation, which is a function applicable to states of an arbitrary number of qubits. For a pure state of a single qubit, the spin flip, which we denote by a tilde, is defined by

$$|\tilde{\psi}\rangle = \sigma_y |\psi^*\rangle \,, \tag{4}$$

where $|\psi^*\rangle$ is the complex conjugate of $|\psi\rangle$ when it is expressed in a fixed basis such as $\{|\uparrow\rangle, |\downarrow\rangle\}$, and σ_y expressed in that same basis is the matrix $\left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$. For a spin-$\frac{1}{2}$ particle this is the standard time reversal operation and indeed reverses the direction of the spin [9]. To perform a spin flip on n qubits, one applies the above transformation to each individual qubit. For example, for a general state ρ of two qubits—the object of interest in this Letter—the spin-flipped state is

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y) \,, \tag{5}$$

where again the complex conjugate is taken in the standard basis, which for a pair of spin-$\frac{1}{2}$ particles is $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$. In this case the spin flip is equivalent [10] to

"complex conjugation in the magic basis," which appears in Ref. [8].

Though we have introduced the spin flip primarily to deal with mixed states, the concept is also convenient for expressing the entanglement of a *pure* state of two qubits. One can show that this entanglement, defined in Eq. (2), can be written as [8]

$$E(\psi) = \mathcal{E}\big(C(\psi)\big), \tag{6}$$

where the "concurrence" C is defined as

$$C(\psi) = |\langle\psi \,|\, \tilde{\psi}\rangle|, \tag{7}$$

and the function $\mathcal{E}$ is given by

$$\mathcal{E}(C) = h\left(\frac{1 + \sqrt{1 - C^2}}{2}\right);$$
$$h(x) = -x \log_2 x - (1 - x)\log_2(1 - x). \tag{8}$$

$\mathcal{E}(C)$ is monotonically increasing and ranges from 0 to 1 as C goes from 0 to 1, so that one can take the concurrence as a measure of entanglement in its own right. For example, the singlet state $|\psi\rangle = (1/\sqrt{2})\,(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ is left unchanged by a spin flip (except for an overall negative sign), so that its concurrence $|\langle\psi \,|\, \tilde{\psi}\rangle|$ is equal to 1. At the other extreme, an unentangled pure state such as $|\uparrow\downarrow\rangle$ is always mapped by the spin flip transformation into an orthogonal state, so that its concurrence is zero. Later we will use another fact about $\mathcal{E}(C)$, namely, that it is a convex function (that is, curving upward).

Having defined the spin flip and the function $\mathcal{E}(C)$, we can now present the promised formula for the entanglement of formation of a mixed state ρ of two qubits:

$$E(\rho) = \mathcal{E}\big(C(\rho)\big), \tag{9}$$

where

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \tag{10}$$

and the λ_is are the eigenvalues, in decreasing order, of the Hermitian matrix $R \equiv \sqrt{\sqrt{\rho}\,\tilde{\rho}\,\sqrt{\rho}}$. Alternatively, one can say that the λ_is are the square roots of the eigenvalues of the non-Hermitian matrix $\rho\tilde{\rho}$. Note that each λ_i is a non-negative real number.

The formula (9) was shown in Ref. [8] to be correct for all density matrices of two qubits having no more than two nonzero eigenvalues. More recently, Smolin has tested the formula numerically on several thousand randomly chosen two-qubit density matrices and has found complete agreement [11]. We now prove the formula for all states of this system by constructing an entanglement-minimizing decomposition of any given density matrix.

We begin with the fact that *every* decomposition of a density matrix ρ can be obtained via the following prescription [12]. First, find a complete set of orthogonal eigenvectors $|v_i\rangle$ corresponding to the nonzero eigenvalues of ρ, and "subnormalize" these vectors so that $\langle v_i \,|\, v_i\rangle$ is equal to the ith eigenvalue. Then a general decomposition $\{|w_i\rangle\}$ of ρ is given by

$$|w_i\rangle = \sum_{j=1}^{n} U_{ij}^{*} |v_j\rangle, \qquad i = 1, \ldots, m. \tag{11}$$

Here n is the rank of ρ, that is, the number of vectors $|v_i\rangle$, and U is an $m \times m$ unitary matrix, m being greater than or equal to n. (The complex conjugation is included only for later convenience.) Alternatively, since only the first n columns of U are used, we can take U to be an $m \times n$ matrix whose columns are orthonormal vectors. The states $|w_i\rangle$ in Eq. (11) are automatically subnormalized so that $\langle w_i \,|\, w_i\rangle$ is equal to the probability of $|w_i\rangle$ in the decomposition. We can thus write $\rho = \sum_i |w_i\rangle\langle w_i|$. In what follows, we express all decompositions of ρ in terms of such subnormalized vectors.

It is helpful to consider separately two classes of density matrix: (i) Those for which $\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4$ is positive or zero, and (ii) those for which the same combination is negative. We consider class (i) first.

For any density matrix ρ in this class, we will define successively three specific decompositions of ρ, the last of which is the optimal decomposition that we seek. Each of these decompositions consists of exactly n pure states, n being the rank of ρ as above. For the system we are considering, n is always less than or equal to 4.

The first decomposition consists of states $|x_i\rangle$, $i = 1, \ldots, n$, satisfying

$$\langle x_i \,|\, \tilde{x}_j\rangle = \lambda_i \delta_{ij}. \tag{12}$$

We obtain such a decomposition as follows. First note that if the set $\{|x_i\rangle\}$ is defined via an $n \times n$ unitary matrix U as in Eq. (11), then the "tilde inner products" $\langle x_i \,|\, \tilde{x}_j\rangle$ can be written as

$$\langle x_i \,|\, \tilde{x}_j\rangle = (U\tau U^{T})_{ij}, \tag{13}$$

where $\tau_{ij} \equiv \langle v_i \,|\, \tilde{v}_j\rangle$ is a symmetric but not necessarily Hermitian matrix. (The states $|v_i\rangle$ are the eigenstates of ρ defined earlier.) In order that condition (12) be satisfied, we want $U\tau U^{T}$ to be diagonal. It happens that for any symmetric matrix τ, one can always find a unitary U that diagonalizes τ in this way [13]. Moreover, the diagonal elements of $U\tau U^{T}$ can always be made real and nonnegative, in which case they are the square roots of the eigenvalues of $\tau\tau^{*}$. (To see how this works, note that U must diagonalize $\tau\tau^{*}$ in the usual sense; that is, $U\tau\tau^{*}U^{\dagger}$ is diagonal.) The square roots of the eigenvalues of $\tau\tau^{*}$ are the same as the eigenvalues of R, so that condition (12) is fulfilled as long as the diagonalizing matrix U is chosen in such a way that the numbers λ_i appear in their proper order. Thus one can always find a decomposition with the desired property.

Our second decomposition of ρ, which we label $\{|\,y_i\rangle\}$, $i = 1, \ldots, n$, is hardly different from the first:

$$|\,y_1\rangle = |x_1\rangle;$$
$$|\,y_j\rangle = i|x_j\rangle, \quad \text{for } j \neq 1. \tag{14}$$

It is indeed physically equivalent to the first decomposition, but the phase factors will become important shortly when we take linear combinations of these vectors.

The decomposition $\{|y_i\rangle\}$ has the following special property. Define the "preconcurrence" c of a pure state $|\psi\rangle$ (possibly subnormalized) to be

$$c(\psi) = \frac{\langle\psi\,|\,\tilde{\psi}\rangle}{\langle\psi\,|\,\psi\rangle}; \tag{15}$$

that is, the preconcurrence is the same as the concurrence of Eq. (7) but without the absolute value sign. Then the average preconcurrence of $\{|y_i\rangle\}$ has the value $C(\rho)$ specified in Eq. (10):

$$\begin{aligned}\langle c\rangle &= \sum_i \langle y_i\,|\,y_i\rangle \frac{\langle y_i\,|\,\tilde{y}_i\rangle}{\langle y_i\,|\,y_i\rangle} \\ &= \sum_i \langle y_i\,|\,\tilde{y}_i\rangle = \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 = C(\rho)\,.\end{aligned} \tag{16}$$

Here we have used the fact that if $n < 4$, the numbers λ_i with $i > n$ are all zero.

We would like to find a decomposition that, like $\{|y_i\rangle\}$, has $\langle c\rangle = C(\rho)$, but which also has the property that the preconcurrence (and hence the concurrence) of each individual *state* is equal to $C(\rho)$. It would then follow immediately that the average entanglement is $\mathcal{E}(C(\rho))$, since this would be the entanglement of each state in the decomposition.

To accomplish this, note first that any decomposition with n elements can be written in terms of the states $|y_i\rangle$ via the equation

$$|z_i\rangle = \sum_{j=1}^{n} V_{ij}^{*}|y_j\rangle, \tag{17}$$

where V is an $n \times n$ unitary matrix. The average preconcurrence of the ensemble $\{|z_i\rangle\}$ is

$$\langle c\rangle = \sum_i \langle z_i\,|\,\tilde{z}_i\rangle = \sum_i (VYV^T)_{ii} = \mathrm{Tr}(VYV^T)\,, \tag{18}$$

where Y is the real diagonal matrix defined by $Y_{ij} = \langle y_i\,|\,\tilde{y}_j\rangle$. Thus the average preconcurrence is unchanged by any *real* unitary matrix V (that is, any orthogonal matrix), since in that case $V^T = V^{-1}$ and the trace in Eq. (18) is preserved.

Let us now restrict ourselves to such orthogonal matrices, so as to preserve the average, and use them to try to equalize the preconcurrences. One way to do this is as follows. First, select the two states $|y_i\rangle$ with the largest and smallest values of the preconcurrence. Unless all the preconcurrences are already equal, the largest one must be too large and the smallest one too small (typically negative). In the latter case, consider the set of positive-determinant orthogonal transformations that act only on these two extreme states as in Eq. (17), changing them into new states that we call $|z_a\rangle$ and $|z_b\rangle$ and leaving the other states $|y_i\rangle$ unchanged. Among this set of transformations is one that simply interchanges the two extreme states and thus interchanges their preconcurrences. Therefore, by continuity there must exist an intermediate transformation that makes the preconcurrence of $|z_a\rangle$ equal to $C(\rho)$. Perform this transformation, thereby fixing one element of the ensemble to have the correct concurrence. By repeating this procedure on the remaining elements of the ensemble, one finally arrives at a set of states all having concurrence equal to $C(\rho)$. This we take to be our final decomposition $\{|z_i\rangle\}$, which, as we have argued above, achieves the claimed minimum average entanglement $\mathcal{E}(C(\rho))$. Thus the value of entanglement given in our formula (9) can always be attained, at least for the case in which $\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 \geq 0$.

We now show that no decomposition of ρ has a *smaller* average entanglement. For this it is enough to show that no decomposition has a smaller average *concurrence:* The average entanglement cannot be less than $\mathcal{E}(\langle C\rangle)$ because of the convexity of the function $\mathcal{E}$. Now, the average concurrence of a general decomposition is given by an equation similar to Eq. (18) but with an absolute value sign:

$$\langle C\rangle = \sum_i |(VYV^T)_{ii}|\,. \tag{19}$$

Here V is an $m \times n$ matrix whose n columns are orthonormal vectors. The dimension m of these vectors can be arbitrarily large, since the decomposition may consist of an arbitrarily large number of pure states (though prior results guarantee that one need not consider values of m larger than sixteen [14]). In terms of the components of V and Y, we can write the average concurrence as

$$\langle C\rangle = \sum_i \left| \sum_j (V_{ij})^2 Y_{jj} \right|\,. \tag{20}$$

To obtain the desired lower bound on this sum, we need use only the fact that $\sum_i |(V_{ij})^2| = 1$. That is, we can show that for any complex numbers α_{ij} such that $\sum_i |\alpha_{ij}| = 1$, we have

$$\sum_i \left| \sum_j \alpha_{ij} Y_{jj} \right| \geq \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\,. \tag{21}$$

The proof is straightforward: first note that there is no loss of generality in taking each α_{i1} to be real and positive. (The phases of the other α_{ij}s can be changed to compensate.) Then we can say

$$\begin{aligned}\sum_i \left| \sum_j \alpha_{ij} Y_{jj} \right| &\geq \left| \sum_{ij} \alpha_{ij} Y_{jj} \right| \\ &= \left| \lambda_1 - \sum_{j=2}^{n} \left(\sum_i \alpha_{ij} \right) \lambda_j \right| \\ &\geq \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 = C(\rho)\,.\end{aligned} \tag{22}$$

Thus no decomposition of ρ can achieve an average concurrence lower than $C(\rho)$ or an average entanglement lower than $\mathcal{E}(C(\rho))$.

There remains one case to consider, namely, density matrices for which $\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 < 0$. For these density matrices our formula predicts that the entanglement should be zero; that is, that there should be a decomposition of ρ into unentangled pure states. To show that this is indeed the case, we again start with the decomposition $\{|x_i\rangle\}$, $i = 1, \ldots, n$, of Eq. (12). If n is equal to 3—the values $n = 1$ and $n = 2$ are not possible for the case we are now considering—it is convenient to supplement this set with a dummy state $|x_4\rangle$ equal to the zero vector. From the complete set we directly form our final decomposition $\{|z_i\rangle\}$:

$$
\begin{aligned}
|z_1\rangle &= \tfrac{1}{2}(e^{i\theta_1}|x_1\rangle + e^{i\theta_2}|x_2\rangle + e^{i\theta_3}|x_3\rangle + e^{i\theta_4}|x_4\rangle)\,,\\
|z_2\rangle &= \tfrac{1}{2}(e^{i\theta_1}|x_1\rangle + e^{i\theta_2}|x_2\rangle - e^{i\theta_3}|x_3\rangle - e^{i\theta_4}|x_4\rangle)\,,\\
|z_3\rangle &= \tfrac{1}{2}(e^{i\theta_1}|x_1\rangle - e^{i\theta_2}|x_2\rangle + e^{i\theta_3}|x_3\rangle - e^{i\theta_4}|x_4\rangle)\,,\\
|z_4\rangle &= \tfrac{1}{2}(e^{i\theta_1}|x_1\rangle - e^{i\theta_2}|x_2\rangle - e^{i\theta_3}|x_3\rangle + e^{i\theta_4}|x_4\rangle)\,,
\end{aligned}
\tag{23}
$$

where the phase factors are chosen so that

$$
\sum_j e^{2i\theta_j}\lambda_j = 0\,. \tag{24}
$$

Such phase factors can always be found when $\lambda_1 < \lambda_2 + \lambda_3 + \lambda_4$ (λ_1 being the largest of the four numbers as always). The condition (24) together with the property (12) of the set $\{|x_i\rangle\}$ guarantee that each state $|z_i\rangle$ has zero concurrence and hence zero entanglement. This completes the proof of the formula (9).

Our formula makes possible the easy evaluation of entanglement of formation for a pair of qubits, and should thus facilitate the investigation of any number of questions concerning entanglement. However, there remains a basic question concerning the *interpretation* of the entanglement of formation that has not yet been resolved. For any pure state $|\psi\rangle$ of a bipartite system, the entanglement $E(\psi)$ defined in Eq. (2) can be interpreted roughly as the number of qubits that must have been exchanged between two observers in order for them to share the state $|\psi\rangle$ [15]. It seems likely that this interpretation applies also to the entanglement of formation of a *mixed* state [5], but this conclusion depends on a property of $E(\rho)$ that has not yet been demonstrated [16]. The question is whether $E(\rho)$ is *additive*, that is, whether the entanglement of formation of two pairs of quantum systems is the sum of the entanglements of formation of the individual pairs and not less. If it is determined that $E(\rho)$ is indeed additive, then this finding will considerably strengthen the physical interpretation of our formula.

I thank a number of colleagues whose comments and suggestions have been of great help in this work: Valerie Coffman, Scott Hill, Joydip Kundu, Hideo Mabuchi, Michael Nielsen, David Park, Eric Rains, John Smolin, Ashish Thapliyal, and especially Chris Fuchs.

[1] E. Schrödinger, Proc. Cambridge Philos. Soc. **31**, 555 (1935).
[2] J. S. Bell, Physics **1**, 195 (1964).
[3] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996); S. Popescu and D. Rohrlich, quant-ph/9610044.
[4] A. Shimony, in *Fundamental Problems in Quantum Theory*, edited by D. M. Greenberger and A. Zeilinger (New York Academy of Sciences, New York, 1995); C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996); V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997); N. J. Cerf and C. Adami, quant-ph/9605039; V. Vedral and M. B. Plenio, quant-ph/9707035; M. Lewenstein and A. Sanpera, quant-ph/9707043.
[5] C. H. Bennett, D. P. DiVincenzo, J. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[6] See, for example, D. P. DiVincenzo, Science **270**, 255 (1995).
[7] See, for example, C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, Phys. Rev. A **56**, 1163 (1997), and references therein.
[8] S. Hill and W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997).
[9] J. J. Sakurai, *Modern Quantum Mechanics*, edited by San Fu Tuan (Benjamin/Cummings, Menlo Park, CA, 1985), p. 277.
[10] The equivalence was pointed out to the author by V. Coffman and J. Kundu.
[11] J. Smolin (private communication).
[12] E. Schrödinger, Proc. Cambridge Philos. Soc. **32**, 446 (1936); N. Hadjisavvas, Lett. Math. Phys. **5**, 327 (1981); L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).
[13] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1985), p. 205.
[14] A. Uhlmann, quant-ph/9701014. Evidence that four states are sufficient can be found in F. Benatti, H. Narnhofer, and A. Uhlmann, Rep. Math. Phys. **38**, 123 (1996).
[15] This interpretation follows directly from the results in Ref. [3].
[16] S. Popescu (private communication).