Geometrical aspects of entanglement

Jon Magne Leinaas,[1] Jan Myrheim,[2] and Eirik Ovrum[1,3]

[1]*Department of Physics, University of Oslo, P.O. Box 1048 Blindern, 0316 Oslo, Norway*
[2]*Department of Physics, The Norwegian University of Science and Technology, 7491 Trondheim, Norway*
[3]*Centre of Mathematics for Applications, P.O. Box 1053 Blindern, 0316 Oslo, Norway*
(Received 22 May 2006; published 19 July 2006)

We study geometrical aspects of entanglement, with the Hilbert–Schmidt norm defining the metric on the set of density matrices. We focus first on the simplest case of two two-level systems and show that a "relativistic" formulation leads to a complete analysis of the question of separability. Our approach is based on Schmidt decomposition of density matrices for a composite system and nonunitary transformations to a standard form. The positivity of the density matrices is crucial for the method to work. A similar approach works to some extent in higher dimensions, but is a less powerful tool. We further present a numerical method for examining separability and illustrate the method by a numerical study of bound entanglement in a composite system of two three-level systems.

I. INTRODUCTION

Entanglement is considered to be one of the main signatures of quantum mechanics, and in recent years the study of different aspects of entanglement has received much attention. One approach has been to study the formal, geometrical characterization of entangled states as opposed to nonentangled or separable states [1–3]. In such a geometrical approach the Hilbert–Schmidt metric defines in many ways the natural metric on the space of physical states. This metric follows quite naturally from the Hilbert space norm, when the quantum description is extended from pure to mixed states, and it is a Euclidean metric on the set of density matrices. For a composite quantum system the separable states form a convex subset of the full convex set of density matrices, and one of the aims of the geometrical approach is to give a complete specification of this set and thereby of the nonseparable or entangled states.

The purpose of this paper is to examine some questions related to the geometrical description of entanglement. We focus primarily on the simplest composite systems consisting of two two-level systems (2×2 system) or two three-level systems (3×3 system), but examine also some questions relevant for higher dimensions. In the case of two two-level systems the separable states can be fully identified by use of the Peres criterion [4]. This criterion states that every separable density matrix is mapped into a positive semidefinite matrix by partial transposition, i.e., by a transposition relative to one of the subsystems. Since also Hermiticity and trace normalization is preserved under this operation, the partial transpose of a separable density matrix is a new density matrix.

A nonseparable density matrix, on the other hand, may or may not satisfy Peres' condition. This means that, in general, Peres' condition is necessary but not sufficient for separability. However, for the special case of a 2×2 system as well as for a 2×3 system the connection is stronger, and the Peres condition is both necessary and sufficient for a density matrix to be separable [5]. To identify the separable density matrices, and thereby the entangled states, is therefore in

these cases relatively simple. In higher dimensions, in particular for the 3×3 system, that is not the case. Peres' condition is therefore not sufficient for separability, and the convex subset consisting of all the density matrices that satisfy this condition is larger than the set of separable matrices. This is part of the reason that the identification of the set of separable states in higher dimensions is a hard problem [6].

In the present paper we first give, in Sec. II, a brief introduction to the geometry of density matrices and separable states. As a next step, in Sec. III we focus on the geometry of the two-level system and discuss the natural extension of the Euclidean three-dimensional Hilbert–Schmidt metric to a four-dimensional indefinite Lorentz metric [7,8]. This indefinite metric is useful for the discussion of entanglement in the 2×2 system, where Lorentz transformations can be used to transform any density matrix to a diagonal standard form in a way which preserves separability and the Peres condition (Sec. IV). By using this standard form it is straightforward to demonstrate the known fact that any matrix that satisfies the Peres condition is also separable.

The transformation to the diagonal standard form is based on an extension of the Schmidt decomposition to the matrix space with indefinite metric. For general matrices such a decomposition cannot be done, but for density matrices it is possible due to the positivity condition. We show in Sec. V that the Schmidt decomposition in this form can be performed not only for the 2×2 system, but for bipartite systems of arbitrary dimensions. However, only for the 2×2 system the decomposition can be used to bring the matrices to a diagonal form where separability can be easily demonstrated. In higher dimensions the Schmidt decomposition gives only a partial simplification. This indicates that to study separability in higher dimensions one eventually has to rely on the use of numerical methods [9–13].

In Sec. VI we discuss a new numerical method to determine separability [14], and use the method to study entanglement in the 3×3 system. The method is based on a numerical estimation of the Hilbert–Schmidt distance from any chosen density matrix to the closest separable one. We focus particularly on states that are nonseparable but still satisfy the Peres condition. These are states that one usually associ-

ates with bound entanglement. A particular example of such states is the one-parameter set discussed by P. Horodecki [15] and extended by Bruss and Peres [16]. We study numerically the density matrices in the neighborhood of one particular Horodecki state and provide a map of the states in a two-dimensional subspace.

II. THE GEOMETRY OF DENSITY MATRICES

A. The convex set of density matrices

A density matrix ρ of a quantum system has the following properties:

$$\rho^\dagger = \rho \text{ Hermiticity,}$$

$$\rho > 0 \text{ positivity,}$$

$$\text{Tr } \rho = 1 \text{ normalization.} \quad (1)$$

The matrices that satisfy these conditions form a convex set, and can be written in the form

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k| \quad (2)$$

with $|\psi_k\rangle$ as Hilbert space unit vectors and

$$p_k \geq 0, \quad \sum_k p_k = 1. \quad (3)$$

The coefficient p_k, for given k, can be interpreted as the probability that the quantum system is in the pure state $|\psi_k\rangle$. This interpretation depends, however, on the representation (2), which is by no means unique. In particular, the vectors $|\psi_k\rangle$ may be chosen to be orthonormal, then they are eigenvectors of ρ with eigenvalues p_k, and Eq. (2) is called the *spectral representation* of ρ.

The pure states, represented by the one-dimensional projections $|\psi\rangle\langle\psi|$, are the extremal points of the convex set of density matrices. That is, they generate all other density matrices, corresponding to mixed states, by convex combinations of the form (2), but cannot themselves be expressed as nontrivial convex combinations of other density matrices.

The Hermitian matrices form a real vector space with a natural scalar product, $\text{Tr}(AB)$, which is bilinear in the two matrices A and B and is positive definite. From this scalar product a Euclidean metric is derived which is called the Hilbert–Schmidt (or Frobenius) metric on the matrix space,

$$ds^2 = \frac{1}{2} \text{Tr}[(d\rho)^2]. \quad (4)$$

The scalar product between pure state density matrices $\rho_1 = |\psi\rangle\langle\psi|$ and $\rho_2 = |\phi\rangle\langle\phi|$ is

$$\text{Tr}(\rho_1\rho_2) = |\langle\psi|\phi\rangle|^2. \quad (5)$$

For an infinitesimal displacement $|d\psi\rangle$ on the unit sphere in Hilbert space the displacement in the matrix space is

$$d\rho = |d\psi\rangle\langle\psi| + |\psi\rangle\langle d\psi| \quad (6)$$
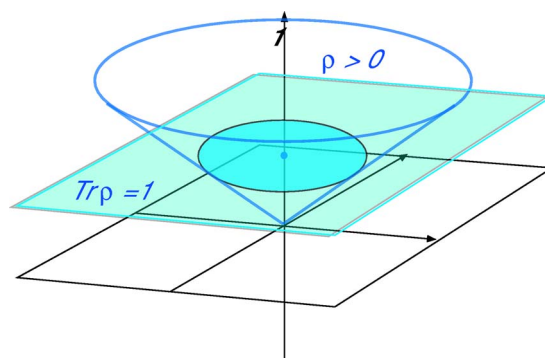
and the Hilbert–Schmidt metric is

FIG. 1. (Color online) A schematic representation of the set of density matrices in the vector space of Hermitian matrices. The positive matrices form a cone about the axis defined by the unit matrix, and the normalization condition restricts the density matrices to a convex subset, here represented by the shaded circle.

$$ds^2 = \langle d\psi|d\psi\rangle - |\langle\psi|d\psi\rangle|^2, \quad (7)$$

where we have used that $\text{Tr}(d\rho) = \langle\psi|d\psi\rangle + \langle d\psi|\psi\rangle = 0$. This may be interpreted as a metric on the complex projective space, called the Fubini–Study metric. It is derived from the Hilbert space metric and is a natural measure of distance between pure quantum states, in fact, ds is the infinitesimal angle between rays (one-dimensional subspaces) in Hilbert space. Since the Hilbert–Schmidt metric on density matrices is a direct extension of the Hilbert space metric, it is a natural metric for all, both pure and mixed, quantum states.

A complete set of basis vectors $\{J_a\}$, in the space of Hermitian matrices, can be chosen to satisfy the normalization condition

$$\text{Tr}(J_a J_b) = \delta_{ab}. \quad (8)$$

For $n \times n$ matrices the dimension of the matrix space, and the number of basis vectors, is n^2. One basis vector, J_0, can be chosen to be proportional to the identity $\mathbb{I}$, then the other basis vectors are traceless matrices. A general density matrix can be expanded in the given basis as

$$\rho = \sum_a \xi_a J_a, \quad (9)$$

where the coefficients ξ_a are real, and the trace normalization of ρ fixes the value of ξ_0. With the chosen normalization of J_0 we have $\xi_0 = 1/\sqrt{n}$.

Due to the normalization, the density matrices are restricted to a hyperplane of dimension $n^2 - 1$, shifted in the direction of J_0 relative to a plane through the origin. The set of density matrices is further restricted by the positivity condition, so it forms a closed, convex set centered around the point $\rho_0 = J_0/\sqrt{n}$. This point corresponds to the *maximally mixed state*, which has the same probability $\langle\psi|\rho_0|\psi\rangle = 1/n$ for any pure state $|\psi\rangle$. The geometry is schematically shown in Fig. 1, where the set of density matrices is pictured as the interior of a circle. One should note that the normalization condition in a sense is trivial and can always be corrected for by a simple scale factor. In the discussion to follow we will find it sometimes convenient to give up this constraint. The

quantum states can then be viewed as rays in the matrix space, and the positivity condition restricts these to a convex sector (the cone in Fig. 1).

B. Unitary transformations

The Hermitian matrices J_a can be viewed as generators of unitary transformations,

$$U = \exp\left(i \sum_a \zeta_a J_a\right) \qquad (10)$$

with real coefficients ζ_a, which act on the density matrices in the following way:

$$\rho \rightarrow \widetilde{\rho} = U\rho U^\dagger. \qquad (11)$$

If ρ is represented as in Eq. (2), then we see that

$$\widetilde{\rho} = \sum_k p_k |\widetilde{\psi}_k\rangle\langle\widetilde{\psi}_k| \qquad (12)$$

where $|\widetilde{\psi}_k\rangle = U|\psi_k\rangle$. Thus the matrix transformation $\rho \rightarrow U\rho U^\dagger$ is induced by the vector transformation $|\psi\rangle \rightarrow U|\psi\rangle$. An immediate consequence of Eq. (12) is that the transformed density matrix $U\rho U^\dagger$ is positive.

Such unitary transformations respect both the trace and positivity conditions and therefore leave the set of density matrices invariant. Also the von Neumann entropy

$$S = -\operatorname{Tr}(\rho \ln \rho) \qquad (13)$$

is unchanged, which means that the degree of mixing, as measured by the entropy, does not change.

Since the dimension of the set of density matrices grows quadratically with the Hilbert space dimension n the geometry rapidly gets difficult to visualize as n increases. However, the high degree of symmetry under unitary transformations simplifies the picture. The unitary transformations define an SU(n) subgroup of the rotations in the n^2-1 dimensional matrix space, and all density matrices can be obtained from the diagonal ones by these transformations. In this sense the geometry of the set of density matrices is determined by the geometry of the set of *diagonal* density matrices. The diagonal matrices form a convex set with a maximal set of n commuting pure states as extremal points. Geometrically, this set is a regular hyperpyramid, a *simplex*, of dimension $n-1$ with the pure states as corners. The geometrical object corresponding to the full set of density matrices is generated from this by the SU(n) transformations.

In Fig. 2 the set of diagonal density matrices is illustrated for $n=2$, 3, and 4, where in the first case the hyperpyramid has collapsed to a line segment, for $n=3$ it is an equilateral triangle and for $n=4$ it is a tetrahedron. For $n=2$, the SU(2) transformations generate from the line segment the three-dimensional *Bloch sphere* of density matrices. This case is special in the sense that the pure states form the complete surface of the set of density matrices. This does not happen in higher dimensions. In fact, the dimension of the set of pure states is $2n-2$, the dimension of SU(n)/[U(1)×SU($n-1$)], because one given pure state has a U(1)×SU($n-1$) invariance group. This dimension grows linearly with n,
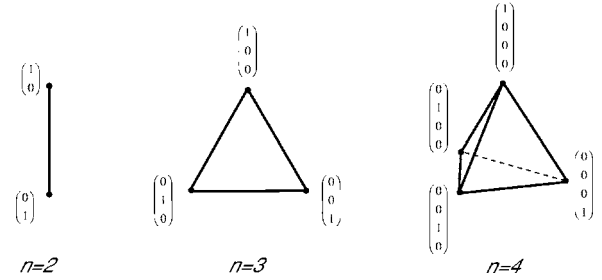

FIG. 2. Geometric representation of the diagonal density matrices for the three cases of Hilbert space dimension 2, 3, and 4. For general dimension n they define a hyperpyramid of dimension $n-1$.

while the dimension of the surface, n^2-2, grows quadratically.

The faces of the hyperpyramid of dimension $n-1$ are hyperpyramids of dimension $n-2$, corresponding to density matrices of the subspace orthogonal to the pure state of the opposite corner. Similarly, the hyperpyramid of dimension $n-2$ is bounded by hyperpyramids of dimension $n-3$, etc. This hierarchy is present also in the full set of density matrices, generated from the diagonal ones by SU(n) transformations. Thus, to each extremal point (pure state) the boundary surface opposite to it is a flat face corresponding to the set of density matrices of one lower dimension. In this way the boundary surface of the set of density matrices contains a hierarchy of sets of density matrices of lower dimensions.

The boundary of the set of density matrices is characterized by at least one of the eigenvalues of the density matrices being zero, since outside the boundary the positivity condition is broken. This means that at the boundary the density matrices satisfy the condition $\det \rho = 0$, which is an algebraic equation for the coordinates of the boundary points. When n is not too large the equation can be solved numerically. This has been done in Fig. 6 where a two-dimensional section of the set of density matrices is shown. One should note that there will be solutions to the equation $\det \rho = 0$ also outside the set of density matrices. The boundary of the set of density matrices can be identified as the *closed* surface, defined by $\det \rho = 0$, that encloses the maximally mixed state and is closest to this point.

C. More general transformations

We shall later make use of the complex extension of the transformations (10) by allowing ζ_a to be complex. This means that the transformation group is extended from SU(n) to SL($n, \mathbb{C}$) (the normalization condition $\det U = 1$, or $\zeta_0 = 0$, is trivial). Transformations of the form $\widetilde{\rho} = V\rho V^\dagger$ do not respect the trace condition $\operatorname{Tr} \rho = 1$ if V is nonunitary, but they do respect the positivity condition because they are still vector transformations of the form (12), with $|\widetilde{\psi}_k\rangle = V|\psi_k\rangle$. This means that they leave the sector of non-normalized density matrices invariant. They no longer keep the entropy unchanged, however. Thus the larger group SL($n, \mathbb{C}$) connects a larger set of density matrices than the restricted group SU(n).

One further generalization is possible. In fact, even if we

allow V to be *antilinear*, the transformation $\tilde{\rho} = V\rho V^\dagger$ still preserves positivity because Eq. (12) still holds with $|\tilde{\psi}_k\rangle = V|\psi_k\rangle$. This point needs some elaboration.

An operator V is antilinear if

$$V(a|\psi\rangle + b|\phi\rangle) = a^* V|\psi\rangle + b^* V|\phi\rangle \qquad (14)$$

for any vectors $|\psi\rangle$, $|\phi\rangle$ and complex numbers a, b. Let $\{|i\rangle\}$ be a set of orthonormal basis vectors, let $\psi_i = \langle i|\psi\rangle$, and write $|V\psi\rangle$ for the vector $V|\psi\rangle$. Then

$$|V\psi\rangle = V\sum_i \psi_i |i\rangle = \sum_{i,j} V_{ji}\psi_i^* |j\rangle \qquad (15)$$

with $V_{ji} = \langle j|V|i\rangle$. The Hermitian conjugate $V^\dagger$ is defined in a basis independent way by the identity

$$\langle\psi|V^\dagger\phi\rangle \equiv \langle\phi|V\psi\rangle = \sum_{i,j} \phi_j^* V_{ji}\psi_i^* \qquad (16)$$

or equivalently,

$$|V^\dagger\phi\rangle = V^\dagger|\phi\rangle = \sum_{i,j} \phi_j^* V_{ji}|i\rangle. \qquad (17)$$

By definition, V is *antiunitary* if $V^\dagger = V^{-1}$.

Familiar relations valid for linear operators are not always valid for antilinear operators. For example, when V is antilinear and $|V\psi\rangle = V|\psi\rangle$, it is no longer true that $\langle V\psi| = \langle\psi|V^\dagger$. This relation cannot hold, simply because $\langle V\psi|$ is a linear functional on the Hilbert space, whereas $\langle\psi|V^\dagger$ is an antilinear functional. What is nevertheless true is that $V|\psi\rangle\langle\psi|V^\dagger = |V\psi\rangle\langle V\psi|$. In fact, both of these operators are linear, and they act on the vector $|\phi\rangle = \Sigma_j \phi_j|j\rangle$ as follows:

$$\begin{aligned} V|\psi\rangle\langle\psi|V^\dagger|\phi\rangle &= V\left(\sum_{i,j} \phi_j^* V_{ji}\psi_i^*|\psi\rangle\right) \\ &= \sum_{i,j} \psi_i V_{ji}^*\phi_j V|\psi\rangle \\ &= |V\psi\rangle\langle V\psi|\phi\rangle. \end{aligned} \qquad (18)$$

As a consequence of this identity the form (12) is valid for the antiunitary transformations, and the positivity is thus preserved.

The transposition of matrices, $\rho \to \rho^T$, obviously preserves positivity, since it preserves the set of eigenvalues. This is not an SU(n) transformation of the form (11), as one can easily check. However, transposition of a Hermitian matrix is the same as complex conjugation of the matrix, and if we introduce the complex conjugation operator K, which is antilinear and antiunitary, we may write

$$\rho^T = K\rho K^\dagger. \qquad (19)$$

Note that transposition is a basis dependent operation. The complex conjugation operator K is also basis dependent, it is defined to be antilinear and to leave the basis vectors invariant, $K|i\rangle = |i\rangle$. We see that $K^\dagger = K = K^{-1}$.

One may ask the general question, which are the transformations that preserve positivity of Hermitian matrices. If we consider an *invertible* linear transformation on the real vector space of matrices, then it has to be a one-to-one mapping of the extremal points of the convex set of positive matrices onto the extremal points. In other words, it is a one-to-one mapping of one-dimensional projections onto one-dimensional projections. In yet other words, it is an invertible vector transformation $|\psi\rangle \to V|\psi\rangle$, defined up to a phase factor, or more generally an arbitrary nonzero complex factor, for each pure state $|\psi\rangle$. One can show that these complex factors can be chosen in such a way that V becomes either linear or antilinear, and that the matrix transformation is $\rho \to V\rho V^\dagger$. However, we will not go into details about this point here.

D. Geometry and separability

We consider next a composite system with two subsystems A and B, of dimensions n_A and n_B, respectively. By definition, the *separable states* of the system are described by density matrices that can be written in the form

$$\rho = \sum_k p_k \rho_k^A \otimes \rho_k^B, \qquad (20)$$

where ρ_k^A and ρ_k^B are density matrices of the two subsystems and p_k is a probability distribution over the set of product density matrices labeled by k. The separable states form a convex subset of the set of all density matrices of the composite system, with the *pure product states* $|\psi\rangle\langle\psi|$, where $|\psi\rangle = |\phi\rangle \otimes |\chi\rangle$, as extremal points. Our interest is to study the geometry of this set, and thereby the geometry of the set of entangled states, defined as the complement of the set of separable states within the full set of density matrices.

The Peres criterion [4] gives a necessary condition for a density matrix to be separable. Let us introduce orthonormal basis vectors $|i\rangle_A$ in $\mathcal{H}_A$ and $|j\rangle_B$ in $\mathcal{H}_B$, as well as the product vectors

$$|ij\rangle = |i\rangle_A \otimes |j\rangle_B. \qquad (21)$$

We write the matrix elements of the density matrix ρ as

$$\rho_{ij;kl} = \langle ij|\rho|kl\rangle. \qquad (22)$$

The partial transposition with respect to the B system is defined as the transformation

$$\rho \to \rho^P: \quad \rho_{ij;kl}^P \equiv \rho_{il;kj}. \qquad (23)$$

This operation preserves the trace, but not necessarily the positivity of ρ. However, for separable states one can see from the expansion (20) that it preserves positivity, because it is just a transposition of the density matrices ρ_k^B of the subsystem B.

Thus the Peres criterion states that preservation of positivity under a partial transposition is a necessary condition for a density matrix ρ to be separable. Conversely, if the partial transpose ρ^P is *not* positive definite, it follows that ρ is nonseparable or entangled. The opposite is not true: if ρ^P is positive, the density matrix ρ is not necessarily separable.

It should be emphasized that the Peres condition, i.e., positivity of both ρ and ρ^P, is independent of the choice of basis vectors $|i\rangle_A$ and $|j\rangle_B$. In fact, a change of basis may result in another definition of the partial transpose ρ^P, which

differs from the first one by a unitary transformation, but this does not change the eigenvalue spectrum of ρ^P. The condition is also the same if transposition is defined with respect to subsystem A rather than B. This is obvious, since partial transposition with respect to the A subsystem is just the combined transformation $\rho \rightarrow (\rho^T)^P$.

Let us consider the Peres condition from a geometrical point of view. We first consider the transposition of matrices, $\rho \rightarrow \rho^T$, and note that it leaves the Hilbert–Schmidt metric invariant. Being its own inverse, transposition is an inversion in the space of density matrices, or a rotation if the number of inverted directions, $n(n-1)/2$ in an $n \times n$ Hermitian matrix, is even. Since ρ and ρ^T have the same set of eigenvalues, transposition preserves positivity and maps the set of density matrices onto itself. Thus the set of density matrices $\mathcal{D}$ is invariant under transposition as well as under unitary transformations.

Similarly, a partial transposition $\rho \rightarrow \rho^P$ preserves the metric and therefore also corresponds to an inversion or rotation in the space of matrices. On the other hand, it does not preserve the eigenvalues and therefore in general does not preserve positivity. This means that the set of density matrices, $\mathcal{D}$, is not invariant under partial transposition, but is mapped into an inverted or rotated copy $\mathcal{D}^P$. These two sets will partly overlap, in particular, they will overlap in a neighborhood around the maximally mixed state, since this particular state is invariant under partial transposition. We note that, even though partial transposition is basis dependent, the set of transposed matrices $\mathcal{D}^P$ does not depend on the chosen basis. Nor does it depend on whether partial transposition is defined with respect to subsystem A or B.

To sum up the situation, we consider the following three convex sets. $\mathcal{D}$ is the full set of density matrices of the composite system, while $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$ is the subset of density matrices that satisfy the Peres condition, and $\mathcal{S}$ is the set of separable density matrices. In general we then have the following inclusions:

$$\mathcal{S} \subset \mathcal{P} \subset \mathcal{D}. \tag{24}$$

The Peres criterion is useful thanks to the remarkable fact that partial transposition does not preserve positivity. This fact is indeed remarkable for the following reason. We have seen that any linear or antilinear vector transformation $|\psi\rangle \rightarrow V|\psi\rangle$ will preserve the positivity of Hermitian matrices by the transformation $\rho \rightarrow V\rho V^\dagger$. It would seem that $V = \mathbb{I}_A \otimes K_B$, a complex conjugation on subsystem B, would be a vector transformation such that $\rho^P = V\rho V^\dagger$, and hence that partial transposition would preserve positivity. What is wrong with this argument is that there exists no such operator as $\mathbb{I}_A \otimes K_B$. To see why, choose a complex number c with $|c|=1$, and consider the transformation of a product vector $|\psi\rangle = |\phi\rangle \otimes |\chi\rangle$,

$$(\mathbb{I}_A \otimes K_B)|\psi\rangle = (\mathbb{I}_A \otimes K_B)[(c|\phi\rangle) \otimes (c^*|\chi\rangle)]$$
$$= c^2[|\phi\rangle \otimes (K|\chi\rangle)]. \tag{25}$$

The arbitrary phase factor c^2 invalidates the attempted definition.

The boundary of the congruent (or reflected) image $\mathcal{D}^P$ of $\mathcal{D}$ is determined by the condition $\det \rho^P = 0$ in the same way as the boundary of the set of density matrices $\mathcal{D}$ is determined by $\det \rho = 0$. As a consequence, to determine whether a density matrix ρ belongs to the set $\mathcal{P}$ is not a hard problem. One simply checks whether the determinants of $\hat{\rho}$ and $\hat{\rho}^P$ are both positive for every $\hat{\rho}$ on the line segment between ρ and the maximally mixed state ρ_0. However, to check whether a density matrix is separable and thus belongs to the subset $\mathcal{S}$ is in general not easy, even though the definition (20) of separability has a simple form. The exceptional cases are the systems of Hilbert space dimensions 2×2, 2×3, or 3×2, where $\mathcal{S} = \mathcal{P}$.

E. Schmidt decomposition and transformation to a standard form

A general density matrix of the composite system can be expanded as

$$\rho = \sum_{a=0}^{n_A^2-1} \sum_{b=0}^{n_B^2-1} \xi_{ab} J_a^A \otimes J_b^B, \tag{26}$$

where the coefficients ξ_{ab} are real, and J_a^A and J_b^B are orthonormal basis vectors of the two subsystems. We may use our convention that $J_0^A = \mathbb{I}_A / \sqrt{n_A}$ and $J_0^B = \mathbb{I}_B / \sqrt{n_B}$, and that J_k^A and J_k^B with $k>0$ are generators of $\mathrm{SU}(n_A)$ and $\mathrm{SU}(n_B)$.

A Schmidt decomposition is a diagonalization of the above expansion. By a suitable choice of basis vectors $\hat{J}_a^A$ and $\hat{J}_b^B$, depending on ρ, we may always write

$$\rho = \sum_{a=0}^{n_A^2-1} \hat{\xi}_a \hat{J}_a^A \otimes \hat{J}_a^B \tag{27}$$

assuming that $n_A \leq n_B$. There exist many different such diagonal representations of a given ρ, in fact it is possible to impose various extra conditions on the new basis vectors. It is usual to impose an orthonormality condition, that the new basis vectors should be orthonormal with respect to some positive definite scalar product. Then the Schmidt decomposition of ρ is the same as the singular value decomposition [17] of the $n_A^2 \times n_B^2$ matrix ξ_{ab}. Below, we will introduce a Schmidt decomposition based on other types of extra conditions.

The usefulness of the representation (27) is limited by the fact that we expand in basis vectors depending on ρ. However, we may make a transformation of the form

$$\rho \rightarrow \tilde{\rho} = V\rho V^\dagger \tag{28}$$

where $V = V_A \otimes V_B$ is composed of transformations $V_A \in \mathrm{SL}(n_A, \mathbb{C})$ and $V_B \in \mathrm{SL}(n_B, \mathbb{C})$ that act independently on the two subsystems and transform the basis vectors $\hat{J}_a^A$ and $\hat{J}_b^B$ into $V_A \hat{J}_a^A V_A^\dagger$ and $V_B \hat{J}_b^B V_B^\dagger$. A transformation of this form obviously preserves the set $\mathcal{S}$ of separable states, since a sum of the form (20) is transformed into a sum of the same form. It also preserves the set $\mathcal{P}$ of density matrices satisfying the Peres condition. In fact, it preserves the positivity not only of ρ, but also of the partial transpose ρ^P, since

$$(\widetilde{\rho})^P = (V_A \otimes V_B^*)\rho^P(V_A \otimes V_B^*)^\dagger. \qquad (29)$$

Here V_B^* is the complex conjugate of V_B. What is not preserved by the transformation is the trace, but this can easily be corrected by introducing a normalization factor. Such transformations have been considered, e.g., by Cen *et al.* [18] and by Osterloh and Siewert [19] in the case of pure states.

As we will later show, it is possible to choose the transformation $V=V_A \otimes V_B$ in such a way that the transformed and normalized density matrix $\widetilde{\rho}$ can be brought into the special form

$$\widetilde{\rho} = \frac{1}{n_A n_B}\left(\mathbb{I} + \sum_{k=1}^{n_A^2-1} \widetilde{\xi}_k \widetilde{J}_k^A \otimes \widetilde{J}_k^B\right), \qquad (30)$$

with $\widetilde{J}_k^A$ and $\widetilde{J}_k^B$ as new sets of traceless orthonormal basis vectors. We have some freedom in choosing V_A and V_B because the form (30) is preserved by unitary transformations of the form $U=U_A \otimes U_B$.

In the case $n_A=n_B=2$ we may choose V_A and V_B so that $\widetilde{J}_k^A$ and $\widetilde{J}_k^B$ are fixed sets of basis vectors, independent of the density matrix ρ. In particular, we may use the standard Pauli matrices as basis vectors. In this way we define a special form of the density matrices $\widetilde{\rho}$, which we refer to as the standard form. Any density matrix can be brought into this form by a transformation that preserves separability and the Peres condition. All matrices of the resulting standard form commute and can be simultaneously diagonalized. This makes it easy to prove the equality $\mathcal{S}=\mathcal{P}$, and thereby solve the separability problem. Although this result is well-known, the proof given here is simpler than the original proof.

The decomposition (30) is generally valid, but when either n_A, n_B, or both are larger than 2, it is impossible to choose both $\widetilde{J}_k^A$ and $\widetilde{J}_k^B$ to be independent of ρ. Simply by counting the number of parameters one easily demonstrates that this cannot be done in higher dimensions. Thus the product transformations $V_A \otimes V_B$ are specified by $2n_A^2+2n_B^2-4$ parameters, while the number of parameters in Eq. (30) is n_A^2-1, when the generators $\widetilde{J}_k^A$ and $\widetilde{J}_k^B$ are fixed. This gives a total number of parameters $3n_A^2+2n_B^2-5$, when $n_A \le n_B$, compared to the number of parameters of the general density matrix, which is $n_A^2 n_B^2-1$. Only for $n_A=n_B=2$ do these numbers match.

The mismatch in the number of parameters shows that the independent transformations V_A and V_B performed on the two subsystems are less efficient in simplifying the form of the density matrices in higher dimensions. In particular, it is impossible to transform all density matrices to a standard form of commuting matrices. Thus the question of separability is no longer trivially solved. Nevertheless, we consider the Schmidt decomposition to be interesting and important, if only because the number of dimensions in the problem is reduced. We expect this to be useful, even if, at the end, separability can only be determined by numerical methods.

III. THE TWO-LEVEL SYSTEM

The density matrices of a two-level system describe the states of a qubit and represent a simple, but important, special case. It is well-known that the normalized density matrices, expressed in terms of the Pauli matrices $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ as

$$\rho = \frac{1}{2}(\mathbb{I} + \mathbf{r} \cdot \boldsymbol{\sigma}), \qquad (31)$$

can geometrically be pictured as the interior of a three-dimensional unit sphere, the *Bloch sphere*, with each point identified by a vector $\mathbf{r}$. The two-level case is special in that the pure states form the complete surface $|\mathbf{r}|=1$ of the set of density matrices. The diagonal 2×2 matrices, in any chosen basis, correspond to a line segment through the origin with the two pure basis states as end points.

The two-level system is special also in the sense that the Euclidean metric of the three dimensional space of density matrices can be extended in a natural way to an indefinite metric in four dimensions. The extension is analogous to the extension from the three-dimensional Euclidean metric to the four-dimensional Lorentz metric in special relativity. Since it is useful for the discussion of entanglement in the two qubit system, we shall briefly discuss it here.

We write the density matrix in relativistic notation as

$$\rho = \frac{1}{2}x^\mu \sigma_\mu, \qquad (32)$$

where σ_0 is the identity matrix $\mathbb{I}$. The trace normalization condition is that $\text{Tr}\,\rho=x^0=1$. We may relax this condition and retain only the positivity condition on ρ, which means that x^0 is positive and dominates the vector part $\mathbf{r}$ of the four-vector x^μ, as expressed by the covariant condition

$$4 \det \rho = (x^0)^2 - |\mathbf{r}|^2 = x^\mu x_\mu = g_{\mu\nu}x^\mu x^\nu \ge 0. \qquad (33)$$

In other words, the four-vector x^μ is restricted by the positivity condition to be either a timelike vector inside the forward light cone, or a lightlike vector on the forward light cone. The lightlike vectors correspond to pure states, and the timelike vectors to mixed states. As already discussed, all points on a given line through the origin represent the same normalized density matrix (see Fig. 3).

Positivity is conserved by matrix transformations of the form

$$\rho \to \widetilde{\rho} = V\rho V^\dagger = \frac{1}{2}L_\mu^\nu x^\mu \sigma_\nu. \qquad (34)$$

If we restrict V to be linear (not antilinear), invertible, and normalized by $\det V=1$, then it belongs to the group $\text{SL}(2,\mathbb{C})$, and L is a continuous Lorentz transformation (continuous in the sense that it can be obtained as a product of small transformations). Thus preservation of positivity by the $\text{SL}(2,\mathbb{C})$ transformations corresponds to preservation of the forward light cone by the continuous Lorentz transformations.
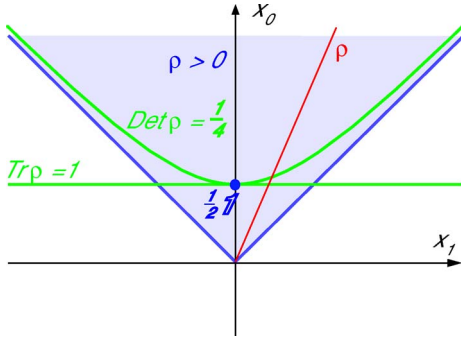
FIG. 3. (Color online) A relativistic view of the density matrices of a two-level system. The positive Hermitian matrices are represented by the forward light cone and a density matrix ρ by a timelike or lightlike ray (red line). The standard normalization $\mathrm{Tr}\,\rho = x^0 = 1$ breaks relativistic invariance (green horizontal line), but may be replaced by the relativistically invariant normalization $\det \rho = x^\mu x_\mu / 4 = 1/4$ (green hyperbola).

In order to compare the indefinite Lorentz metric to the Euclidean scalar product $\mathrm{Tr}(AB)$ introduced earlier, we introduce the operation of space inversion,

$$\sigma_\mu = (\mathbb{I}, \boldsymbol{\sigma}) \to \bar{\sigma}_\mu \equiv (\mathbb{I}, -\boldsymbol{\sigma}). \tag{35}$$

It is obtained as a combination of matrix transposition, or equivalently complex conjugation, which for the standard Pauli matrices inverts the sign of $\sigma_y = \sigma_2$, and a rotation of π about the y axis. Thus for a general Hermitian 2×2 matrix A it acts as

$$A \to \bar{A} = R A^T R^\dagger \tag{36}$$

where

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{37}$$

We may also write $A^T = K A K^\dagger$, where $K = K^\dagger$ is complex conjugation.

The Lorentz metric is now expressed by

$$\mathrm{Tr}(\bar{\sigma}_\mu \sigma_\nu) = 2 g_{\mu\nu} \tag{38}$$

and the Lorentz invariant scalar product between two Hermitian matrices $A = a^\mu \sigma_\mu$ and $B = b^\mu \sigma_\mu$ is

$$\mathrm{Tr}(\bar{A}B) = 2 a^\mu b_\mu. \tag{39}$$

The invariance of the scalar product (38) can be seen directly from the transformation properties under $\mathrm{SL}(2, \mathbb{C})$ transformations,

$$A \to A' = V A V^\dagger \Rightarrow \bar{A} \to \bar{A}' = V^{\dagger -1} \bar{A} V^{-1}. \tag{40}$$

Thus A and $\bar{A}$ transform under contragredient representations of $\mathrm{SL}(2, \mathbb{C})$.

The Lorentz transformed Pauli matrices

$$\tilde{\sigma}_\mu = V \sigma_\mu V^\dagger = L_\mu^{\ \nu} \sigma_\nu, \quad V \in \mathrm{SL}(2, \mathbb{C}) \tag{41}$$

satisfy the same metric condition (38) as the standard Pauli matrices σ_μ. Conversely, any set of matrices $\tilde{\sigma}_\mu$ satisfying

relations of the form (38) are related to the Pauli matrices by a Lorentz transformation, which need not, however, be restricted to the continuous part of the Lorentz group, but may include space inversion or time reversal, or both.

It is clear from the relativistic representation that any density matrix can be reached from the maximally mixed state $\rho_0 = \mathbb{I}/2$ by a transformation corresponding to a boost. The Lorentz boosts generate from ρ_0 a three-dimensional hyperbolic surface (a "mass shell") where $x^\mu x_\mu = 1$. This surface will intersect any timelike line once and only once. Thus any mixed state is obtained from ρ_0 by a unique boost. However, the pure states, corresponding to lightlike vectors, can only be reached asymptotically, when the boost velocity approaches the speed of light, here set equal to 1. The form of the $\mathrm{SL}(2, \mathbb{C})$ transformation corresponding to a pure boost is

$$V(\boldsymbol{\xi}) = \exp\left(\frac{1}{2} \boldsymbol{\xi} \cdot \boldsymbol{\sigma} \right), \tag{42}$$

where the three-dimensional real vector $\boldsymbol{\xi}$ is the boost parameter, called rapidity. Since the boost matrices are Hermitian, a density matrix defined by a boost of the maximally mixed state will have the form

$$\rho = N(\xi) V(\boldsymbol{\xi})^2 = N(\xi)(\cosh \xi \mathbb{I} + \sinh \xi \hat{\boldsymbol{\xi}} \cdot \boldsymbol{\sigma}), \quad \hat{\boldsymbol{\xi}} = \frac{\boldsymbol{\xi}}{\xi}, \tag{43}$$

where $N(\xi)$ is a normalization factor determined by the trace normalization condition. The normalized density matrix is

$$\rho = \frac{1}{2}(\mathbb{I} + \tanh \xi \hat{\boldsymbol{\xi}} \cdot \boldsymbol{\sigma}). \tag{44}$$

Thus the boost parameter $\boldsymbol{\xi}$ gives a representation which is an alternative to the Bloch sphere representation. The relation between the parameters $\mathbf{r}$ and $\boldsymbol{\xi}$ is that

$$\mathbf{r} = \tanh \xi \hat{\boldsymbol{\xi}}, \tag{45}$$

which means that $\mathbf{r}$ can be identified as the velocity of the boost, in the relativistic picture, with $|\mathbf{r}| = 1$ corresponding to the speed of light. We note that the positivity condition gives no restriction on $\boldsymbol{\xi}$, and the extremal points, i.e., the pure states, are points at infinity in the $\boldsymbol{\xi}$ variable.

IV. ENTANGLEMENT IN THE 2×2 SYSTEM

We consider now in some detail entanglement between two two-level systems. We will show that with the use of nonunitary transformations the density matrices can be written in a standardized Schmidt decomposed form. In this form the question of separability is easily determined and the equality of the two sets $\mathcal{P}$ and $\mathcal{S}$ is readily demonstrated.

A. Schmidt decomposition by Lorentz transformations

We consider transformations $V = V_A \otimes V_B$ composed of $\mathrm{SL}(2, \mathbb{C})$ transformations acting independently on the two subsystems, and therefore respecting the product form (20) of the separable matrices. We will show that by means of

such transformations any density matrix of the composite system can be transformed to the form

$$\widetilde{\rho} = \frac{1}{4}\left(\mathbb{I} + \sum_{k=1}^{3} d_k \sigma_k \otimes \sigma_k\right), \tag{46}$$

which we refer to as the standard form. Note that the real coefficients d_k must be allowed to take both positive and negative values.

We start by writing a general density matrix in the form

$$\rho = c^{\mu\nu}\sigma_\mu \otimes \sigma_\nu. \tag{47}$$

The transformation $V=V_A \otimes V_B$ produces independent Lorentz transformation L_A and L_B on the two subsystems,

$$\rho \rightarrow \widetilde{\rho} = V\rho V^\dagger = \widetilde{c}^{\mu\nu}\sigma_\mu \otimes \sigma_\nu \tag{48}$$

with

$$\widetilde{c}^{\mu\nu} = L_{A\rho}^\mu L_{B\sigma}^\nu c^{\rho\sigma}. \tag{49}$$

The Schmidt decomposition consists in choosing L_A and L_B in such a way that $\widetilde{c}^{\mu\nu}$ becomes diagonal. We will show that this is always possible when ρ is strictly positive.

Note that the standard Schmidt decomposition, also called the singular value decomposition, involves a compact group of rotations, leaving invariant a positive definite scalar product. The present case is different because it involves a noncompact group of Lorentz-transformations, leaving invariant an indefinite scalar product.

The positivity condition on ρ plays an essential part in the proof. It states that

$$\langle\psi|\rho|\psi\rangle \geq 0, \tag{50}$$

where $|\psi\rangle$ is an arbitrary state vector. Let us consider a density matrix ρ which is *strictly* positive so that Eq. (50) is satisfied with $>$ and not only with $\geq$, and let us restrict $|\psi\rangle$ to be of product form, $|\psi\rangle=|\phi\rangle \otimes |\chi\rangle$. With ρ expressed by Eq. (47), the positivity condition then implies

$$c^{\mu\nu}m_\mu n_\nu > 0 \tag{51}$$

with

$$m_\mu = \langle\phi|\sigma_\mu|\phi\rangle, \quad n_\nu = \langle\chi|\sigma_\nu|\chi\rangle. \tag{52}$$

These two four-vectors are on the forward light cone, in fact, it is easy to show that

$$m_\mu m^\mu = n_\mu n^\mu = 0, \quad m_0 = n_0 = 1. \tag{53}$$

We note that by varying the state vectors $|\phi\rangle$ and $|\chi\rangle$ all directions on the light cone can be reached. The inequality (51) holds for forward timelike vectors as well, because any such vector may be written as a linear combination of two forward lightlike vectors, with positive coefficients. We may actually write a stronger inequality

$$c^{\mu\nu}m_\mu n_\nu \geq C > 0 \tag{54}$$

valid for all timelike or lightlike vectors m and n with $m_0 = n_0 = 1$. In fact, this inequality holds because the set of such pairs of four-vectors (m,n) is compact.

Now define the function

$$f(m,n) = \frac{c^{\mu\nu}m_\mu n_\nu}{\sqrt{m^\rho m_\rho n^\sigma n_\sigma}}. \tag{55}$$

It is constant for m and n lying on two fixed, one-dimensional rays inside the forward light cone. It goes to infinity when either m or n becomes lightlike, because

$$f(m,n) \geq \frac{Cm_0 n_0}{\sqrt{m^\rho m_\rho n^\sigma n_\sigma}}. \tag{56}$$

Using again a compactness argument, we conclude that there exist four-vectors $\bar{m}$ and $\bar{n}$ such that $f(\bar{m},\bar{n})$ is minimal. We may now choose the Lorentz transformations L_A and L_B such that

$$L_{A\mu}^0 = \bar{m}_\mu, \quad L_{B\mu}^0 = \bar{n}_\mu \tag{57}$$

assuming the normalization conditions $\bar{m}^\mu\bar{m}_\mu = \bar{n}^\mu\bar{n}_\mu = 1$. This defines L_A and L_B uniquely up to arbitrary three-dimensional rotations. Define

$$\widetilde{f}(m,n) = \frac{\widetilde{c}^{\mu\nu}m_\mu n_\nu}{\sqrt{m^\rho m_\rho n^\sigma n_\sigma}}. \tag{58}$$

Since $\widetilde{f}(m,n)=f(\widetilde{m},\widetilde{n})$, with $\widetilde{m}_\mu = L_{A\mu}^\rho m_\rho$ and $\widetilde{n}_\mu = L_{B\mu}^\rho n_\rho$, it follows that $\widetilde{f}(m,n)$ has a minimum at $m=n=(1,0,0,0)$. The condition for an extremum at $m=n=(1,0,0,0)$ is that $\widetilde{c}^{0k} = \widetilde{c}^{k0}=0$ for $k=1,2,3$, so that

$$\widetilde{\rho} = \widetilde{c}^{00}\mathbb{I} + \sum_{k=1}^{3}\sum_{l=1}^{3}\widetilde{c}^{kl}\sigma_k \otimes \sigma_l. \tag{59}$$

The coefficient $\widetilde{c}^{00}$ is the minimum value of $f(m,n)$, and hence positive.

The last term of Eq. (59) can be diagonalized by a standard Schmidt decomposition, and by a further normalization $\widetilde{\rho}$ can be brought into the form (30). Finally, a unitary transformation $\widetilde{\rho} \rightarrow U\widetilde{\rho}U^\dagger$ of the product form $U=U_A \otimes U_B$ may be performed, where the unitary matrices U_A and U_B may be chosen so that $U_A\widetilde{J}_k^A U_A^\dagger = \pm\sigma_k$ and $U_B\widetilde{J}_k^B U_B^\dagger = \pm\sigma_k$. This is always possible because SU(2) transformations generate the full three-dimensional rotation group, excluding inversions. In this way we obtain the standard form (46).

Note that the standard form (46) of a given density matrix ρ is not unique because there exists a discrete subgroup of 24 unitary transformations that transform one matrix of this form into other matrices of the same form. This group includes all permutations of the three basis vectors $\sigma_k \otimes \sigma_k$, as well as simultaneous reversals of any two of the basis vectors. It is the full symmetry group of a regular tetrahedron. If we want to make the standard form unique we may, for example, impose the conditions $d_1 \geq d_2 \geq |d_3|$, allowing both positive and negative values of d_3.

B. Density matrices in the standard form

The density matrices of the standard form (46) define a convex subset of lower dimension than the full set of density matrices. It is a three-dimensional section of the 15-dimensional set of density matrices, consisting of commuting
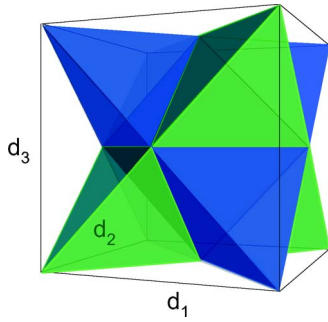
FIG. 4. (Color online) Geometric representation of the diagonal matrices spanned by orthogonal Bell states of the 2×2 system. The density matrices form a tetrahedron (green, its edges are diagonals of the faces of a cube) while the matrices obtained from this by a partial transposition define a mirror image (blue, its edges are the opposite diagonals). The separable states are defined by the intersection between the two tetrahedra and form an octahedron (in the center).

(simultaneously diagonalizable) matrices. The eigenvalues, as functions of the parameters d_k of Eq. (46), are

$$\lambda_{1,2} = \frac{1}{4}[1 \pm (d_1 - d_2) + d_3], \quad \lambda_{3,4} = \frac{1}{4}[1 \pm (d_1 + d_2) - d_3].$$

$$(60)$$

The pure states (with eigenvalues $1,0,0,0$) that are the extremal points of the convex set of commuting matrices are specified by the conditions

$$|d_1| = |d_2| = |d_3| = 1, \quad d_1 d_2 d_3 = -1. \quad (61)$$

There are four such states, corresponding to the four corners of the tetrahedron of diagonal density matrices, and these are readily identified as four orthogonal Bell states (maximally entangled pure states).

We now consider the action of the partial transposition on the tetrahedron of diagonal density matrices. It preserves the standard form and transforms the coefficients as $d_1 \to d_1, d_2 \to -d_2, d_3 \to d_3$. Thus it produces a mirror image of the tetrahedron, by a reflection in the d_2 direction (see Fig. 4). The density matrices of standard form belonging to the set $\mathcal{P}$, i.e., satisfying the Peres condition that they remain positive after the partial transposition, form an octahedron which is the intersection of the two tetrahedra.

We will now show that for the density matrices of standard form the Peres condition is both necessary and sufficient for separability. What we have to show is that all the density matrices of the octahedron are separable. Since, in general, the separable matrices form a convex set, it is sufficient to show that the corners of the octahedron correspond to separable states.

The density matrices of the octahedron satisfy a single inequality

$$|d_1| + |d_2| + |d_3| \le 1 \quad (62)$$

and its six corners are (d_1, d_2, d_3) $= (\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)$, corresponding to the mid-

points of the six edges of each of the two tetrahedra. The corners are separable by the identities

$$\frac{1}{4}(\mathbb{I} \pm \sigma_k \otimes \sigma_k) = \frac{1}{8}[(\mathbb{I} + \sigma_k) \otimes (\mathbb{I} \pm \sigma_k) + (\mathbb{I} - \sigma_k) \otimes (\mathbb{I} \mp \sigma_k)].$$

$$(63)$$

This completes our proof that the Peres condition is both necessary and sufficient for separability of density matrices on the standard form. Furthermore, since any (nonsingular) density matrix can be obtained from a density matrix of standard form by a transformation that preserves both separability and the Peres condition, this reproduces the known result that for the 2×2 system the set of density matrices that remain positive after a partial transposition is identical to the set of separable density matrices.

With this we conclude the discussion of the two-qubit system. The main point has been to show the usefulness of applying the nonunitary Lorentz transformations in the discussion of separability. Also in higher dimensions such transformations can be applied in the form of $SL(n, \mathbb{C})$ transformations, although not in precisely the same form as with two-level systems.

V. HIGHER DIMENSIONS

The relativistic formulation is specific for the two-level system, but some elements can be generalized to higher dimensions. We consider first a single system with Hilbert space dimension n. Again, if the trace condition is relaxed, the symmetry group $SU(n)$ of the set of density matrices is extended to $SL(n, \mathbb{C})$. The Hilbert–Schmidt metric is *not* invariant under this larger group, but the determinant is invariant for any n. However, it is only for $n=2$ that the determinant is quadratic and can be interpreted as defining an invariant indefinite metric.

The generalization to $n > 2$ of the Lorentz boosts are the Hermitian matrices

$$V = V^\dagger, \quad \det V = 1 \quad (64)$$

and expressed in terms of the $SU(n)$ generators $\mathbf{J}$ they have the form

$$V = \exp(\boldsymbol{\xi} \cdot \mathbf{J}) \quad (65)$$

with real group parameters $\boldsymbol{\xi}$. Here $\mathbf{J}$ and $\boldsymbol{\xi}$ are $n^2 - 1$ dimensional vectors. Any *strictly positive* density matrix ρ (with $\det \rho > 0$) can be factorized in terms of Hermitian matrices (65) as

$$\rho = NV^2 = N \exp(2\boldsymbol{\xi} \cdot \mathbf{J}) \quad (66)$$

with the normalization factor

$$N = \frac{1}{\mathrm{Tr} \exp(2\boldsymbol{\xi} \cdot \mathbf{J})} = (\det \rho)^{1/n}. \quad (67)$$

Thus in the same way as for $n=2$, any strictly positive density matrix can be generated from the maximally mixed state by an $SL(n, \mathbb{C})$ transformation of the form (65). The boundary matrices, however, which satisfy $\det \rho = 0$, cannot be ex-

pressed in this way, they can be reached by Hermitian transformations only asymptotically, as $|\boldsymbol{\xi}| \to \infty$. In this limit $N \to 0$, and therefore $\mathrm{Tr}\, V^2 \to \infty$ for the non-normalized density matrix V^2.

Schmidt decomposition

We now consider a composite system, consisting of two subsystems of dimension n_A and n_B, and assume ρ to be a strictly positive density matrix on the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ of the composite system. The general expansion of ρ in terms of the $\mathrm{SU}(n)$ generators is given by Eq. (26). Our objective is to show that by a transformation of the form (28), with $V_A \in \mathrm{SL}(n_A, \mathbb{C})$ and $V_B \in \mathrm{SL}(n_B, \mathbb{C})$, followed by normalization, the density matrix can be transformed to the simpler form (30).

Let $\mathcal{D}_A$ and $\mathcal{D}_B$ be the sets of density matrices of the two subsystems A and B. The Cartesian product $\mathcal{D}_A \times \mathcal{D}_B$, consisting of all product density matrices $\rho_A \otimes \rho_B$ with normalization $\mathrm{Tr}\,\rho_A = \mathrm{Tr}\,\rho_B = 1$, is a compact set of matrices on the full Hilbert space $\mathcal{H}$. For the given density matrix ρ we define the following function of ρ_A and ρ_B, which does not depend on the normalizations of ρ_A and ρ_B,

$$f(\rho_A, \rho_B) = \frac{\mathrm{Tr}[\rho(\rho_A \otimes \rho_B)]}{(\det \rho_A)^{1/n_A} (\det \rho_B)^{1/n_B}}. \tag{68}$$

This function is well-defined on the interior of $\mathcal{D}_A \times \mathcal{D}_B$, where $\det \rho_A > 0$ and $\det \rho_B > 0$. Because ρ is assumed to be strictly positive, we have the strict inequality

$$\mathrm{Tr}[\rho(\rho_A \otimes \rho_B)] > 0 \tag{69}$$

and since $\mathcal{D}_A \times \mathcal{D}_B$ is compact, we have an even stronger inequality on $\mathcal{D}_A \times \mathcal{D}_B$,

$$\mathrm{Tr}[\rho(\rho_A \otimes \rho_B)] \geq C > 0 \tag{70}$$

with a lower bound C depending on ρ. It follows that $f \to \infty$ on the boundary of $\mathcal{D}_A \times \mathcal{D}_B$, where either $\det \rho_A = 0$ or $\det \rho_B = 0$. It follows further that f has a positive minimum on the interior of $\mathcal{D}_A \times \mathcal{D}_B$, with the minimum value attained for at least one product density matrix $\tau_A \otimes \tau_B$ with $\det \tau_A > 0$ and $\det \tau_B > 0$. For τ_A and τ_B we may use the representation (66), written as

$$\tau_A = T_A^\dagger T_A, \quad \tau_B = T_B^\dagger T_B \tag{71}$$

ignoring normalization factors. The matrices $T_A \in \mathrm{SL}(n_A, \mathbb{C})$ and $T_B \in \mathrm{SL}(n_B, \mathbb{C})$ may be chosen to be Hermitian, but they need not be, since they may be multiplied from the left by arbitrary unitary matrices. We further write $T = T_A \otimes T_B$, so that

$$\tau_A \otimes \tau_B = T^\dagger T. \tag{72}$$

Now define a transformed density matrix

$$\tilde{\rho} = T\rho T^\dagger \tag{73}$$

and define

$$\tilde{f}(\rho_A, \rho_B) = \frac{\mathrm{Tr}[\tilde{\rho}(\rho_A \otimes \rho_B)]}{(\det \rho_A)^{1/n_A} (\det \rho_B)^{1/n_B}} = f(T_A^\dagger \rho_A T_A, T_B^\dagger \rho_B T_B). \tag{74}$$

This transformed function $\tilde{f}$ has a minimum for

$$\rho_A \otimes \rho_B = (T^\dagger)^{-1} \tau_A \otimes \tau_B T^{-1} = \mathbb{I}_A \otimes \mathbb{I}_B = \mathbb{I}. \tag{75}$$

Since $\tilde{f}$ is stationary under infinitesimal variations about the minimum, it follows that

$$\mathrm{Tr}[\tilde{\rho}\, \delta(\rho_A \otimes \rho_B)] = 0 \tag{76}$$

for all infinitesimal variations

$$\delta(\rho_A \otimes \rho_B) = \delta\rho_A \otimes \mathbb{I}_B + \mathbb{I}_A \otimes \delta\rho_B \tag{77}$$

subject to the constraints $\det(\mathbb{I}_A + \delta\rho_A) = \det(\mathbb{I}_B + \delta\rho_B) = 1$, or equivalently,

$$\mathrm{Tr}(\delta\rho_A) = \mathrm{Tr}(\delta\rho_B) = 0. \tag{78}$$

The variations satisfying the constraints are the general linear combinations of the SU generators,

$$\delta\rho_A = \sum_i \delta c_i^A J_i^A, \quad \delta\rho_B = \sum_j \delta c_j^B J_j^B. \tag{79}$$

It follows that

$$\mathrm{Tr}(\tilde{\rho} J_i^A \otimes \mathbb{I}_B) = \mathrm{Tr}(\tilde{\rho} \mathbb{I}_A \otimes J_j^B) = 0 \tag{80}$$

for all $\mathrm{SU}(n_A)$ generators J_i^A and all $\mathrm{SU}(n_B)$ generators J_j^B. This means that the terms proportional to $J_i^A \otimes \mathbb{I}_B$ and $\mathbb{I}_A \otimes J_j^B$ vanish in the expansion for $\tilde{\rho}$, which therefore has the form

$$\tilde{\rho} = \frac{1}{n_A n_B}\left(\mathbb{I} + \sum_{k=1}^{n_A^2-1} \sum_{l=1}^{n_B^2-1} \xi_{kl} J_k^A \otimes J_l^B \right). \tag{81}$$

In order to write $\tilde{\rho}$ in the Schmidt decomposed form (30), we have to make a change of basis, from the fixed basis sets J_i^A and J_j^B to other orthonormal SU generators $\tilde{J}_i^A$ and $\tilde{J}_j^B$ depending on $\tilde{\rho}$. This final Schmidt decomposition involves a standard singular value decomposition of the matrix ξ_{kl} by orthogonal transformations. We may make further unitary transformations $U = U_A \otimes U_B$, but as already pointed out, this is in general not sufficient to obtain a standard form independent of ρ.

VI. NUMERICAL APPROACH TO THE STUDY OF SEPARABILITY

In higher dimensions the Peres condition is not sufficient to identify the separable states. In other words, there exist entangled states that remain positive after a partial transposition. This is known not only from general theoretical considerations [5], but also from explicit examples [15]. States of this type have been referred to as having *bound* entanglement. However, whereas it is a fairly simple task to check the Peres condition, it is in general difficult to identify the separable states [6].
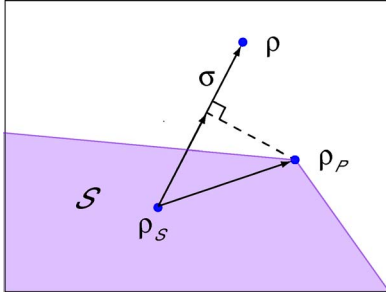
FIG. 5. (Color online) Schematic illustration of the method of finding the separable state closest to a test matrix ρ. The matrix ρ_s represents the best approximation to the closest separable state at a given iteration step, while ρ_p is the product matrix that maximizes the scalar product with the matrix $\sigma = \rho - \rho_s$. This matrix is used to improve ρ_s at the next iteration step. The shaded area $\mathcal{S}$ in the figure illustrates a section through the set of separable density matrices.

In this section we discuss a general numerical method for identifying separability, previously introduced in [14]. It is based on an iterative scheme for calculating the closest separable state and the distance to it, given an arbitrary density matrix (test state). The method can be used to test single density matrices for separability or to make a systematic search to identify the boundary of the set of separable density matrices. After giving an outline of the method we show how to apply the method in a numerical study of bound entanglement in a 3×3 system.

A. Outline of the method

Assume a test state ρ has been chosen. This may typically be close to the boundary of the set of states that satisfy the Peres condition. Let ρ_s be a separable state, an approximation in the iterative scheme to the closest separable state. We may start, for example, with $\rho_s = \mathbb{I}/n$, the maximally mixed state, or with any pure product state. The direction from ρ_s to ρ is denoted $\sigma = \rho - \rho_s$. In order to improve the estimate ρ_s we look for a pure product state ρ_p that maximizes the scalar product

$$s = \text{Tr}[(\rho_p - \rho_s)\sigma] \qquad (82)$$

or equivalently, maximizes $s' = \text{Tr}(\rho_p \sigma)$ (see Fig. 5). If $s > 0$, then it is possible to find a closer separable state ρ_s' by mixing in the product state ρ_p. This search for closer separable states is iterated, either until no pure product state ρ_p can be found such that $s > 0$, which means that ρ_s is already the unique separable state closest to ρ, or until some other convergence criterion is satisfied.

There are two separate mathematical subproblems that have to be solved numerically in this scheme. The first problem is to find the pure product state maximizing the scalar product s'. The second problem is the so-called quadratic programming problem: given a finite number of pure product states, to find the convex combination of these which is closest to the given state ρ. Our approach to these two problems is described briefly below. We refer to Ref. [14] for more details.

To approach the first subproblem, note that a pure product state ρ_p has matrix elements of the form

$$\langle ij|\rho_p|kl\rangle = \phi_i \chi_j \phi_k^* \chi_l^*, \qquad (83)$$

where $\Sigma_i |\phi_i|^2 = \Sigma_j |\chi_j|^2 = 1$. We want to find complex coefficients ϕ_i and χ_j that maximize

$$s' = \text{Tr}(\rho_p \sigma) = \sum_{i,j,k,l} \phi_i^* \chi_j^* \sigma_{ij;kl} \phi_k \chi_l. \qquad (84)$$

The following iteration scheme turns out in practice to be an efficient numerical method. It may not necessarily give a global maximum, but at least it gives a useful local maximum that may depend on a randomly chosen starting point.

The method is based on the observation that the maximum value of s' is actually the maximal eigenvalue μ in the two linked eigenvalue problems

$$\sum_k A_{ik} \phi_k = \mu \phi_i, \quad \sum_l B_{jl} \chi_l = \mu \chi_j, \qquad (85)$$

where

$$A_{ik} = \sum_{j,l} \chi_j^* \sigma_{ij;kl} \chi_l, \quad B_{jl} = \sum_{i,k} \phi_i^* \sigma_{ij;kl} \phi_k. \qquad (86)$$

Thus we may start with any arbitrary unit vector $|\chi\rangle = \Sigma_j \chi_j |j\rangle_B \in \mathcal{H}_B$ and compute the Hermitian matrix A. We compute the unit vector $|\phi\rangle = \Sigma_i \phi_i |i\rangle_A \in \mathcal{H}_A$ as an eigenvector of A with maximal eigenvalue, and we use it to compute the Hermitian matrix B. Next, we compute a new unit vector $|\chi\rangle$ as an eigenvector of B with maximal eigenvalue, and we iterate the whole procedure.

This iteration scheme is guaranteed to produce a nondecreasing sequence of function values s', which must converge to a maximum value μ. This is at least a local maximum, and there corresponds to it at least one product vector $|\phi\rangle \otimes |\chi\rangle$ and product density matrix $\rho_p = (|\phi\rangle\langle\phi|) \otimes (|\chi\rangle\langle\chi|)$.

The above construction of ρ_p implies, if $s > 0$, that there exist separable states

$$\rho_s' = (1 - \lambda)\rho_s + \lambda \rho_p \qquad (87)$$

with $0 < \lambda \leq 1$, closer to ρ than ρ_s is. However, it turns out to be very inefficient to search only along the line segment from ρ_s to ρ_p for a better approximation to ρ. It is much more efficient to append the new ρ_p to a list of product states ρ_{pk} found in previous iterations, and then minimize

$$F = \text{Tr}\left(\rho - \sum_k \lambda_k \rho_{pk}\right)^2 \qquad (88)$$

which is a quadratic function of coefficients $\lambda_k \geq 0$ with $\Sigma_k \lambda_k = 1$. We solve this quadratic programming problem by an adaptation of the conjugate gradient method, and we throw away a given product matrix ρ_{pk} if and only if the corresponding coefficient λ_k becomes zero when F is minimized. In practice, this means that we may construct altogether several hundred or even several thousand product states, but only a limited number of those, typically less than 100 in the cases we have studied, are actually included in the final approximation ρ_s.

B. Bound entanglement in the 3×3 system

For the 3×3 system (composed of two three-level systems) there are explicit examples of entangled states that remain positive under a partial transposition. This was first discussed by Horodecki [15] and then an extended set of states was found by Bruss and Peres [16]. We apply the method outlined above to density matrices limited to a two-dimensional planar section of the full set. The section is chosen to contain one of the Horodecki states and a Bell state in addition to the maximally mixed state, and the method is used to identify the boundary of the separable states in this two-dimensional plane.

Since the separable states $\mathcal{S}$ are contained in the set $\mathcal{P}$ of states that remain positive under partial transposition, we start the search for the boundary of $\mathcal{S}$ with a series of states located at the boundary of $\mathcal{P}$. This boundary is found by solving the algebraic equations det $\rho=0$ and det $\rho^P=0$. For each chosen state we find the distance to the closest separable state and change the test state ρ on a straight line between this point and the maximally mixed state, in a step of length equal to the evaluated distance. In a small number of steps the intersection of the straight line and the boundary of the separable states is found within a small error, typically chosen to be 10^{-6}. (The distance from the maximally mixed state to the pure states in this case is $d=2/3$.)

In Fig. 6 we show a plot of the results of the calculations. The numerically determined points on the border of the set of separable states $\mathcal{S}$ are indicated by black dots, while the border of the set of states $\mathcal{P}$ that satisfy Peres' condition, determined by solving the algebraic equations, is shown as lines which cross at the position of the Horodecki state. One should note that the states with bound entanglement in the corner of the set $\mathcal{P}$ cover a rather small area.

VII. CONCLUSIONS

To summarize, we have in this paper focused on some basic questions concerning the geometry of separability. The simplest case of 2×2 matrices has been used to demonstrate the usefulness of relaxing the normalization requirement Tr $\rho=1$. Thus if this condition is replaced by det $\rho=1$, a relativistic description with a Minkowski metric can be used, where all (nonpure) states can be connected by Lorentz transformations. For a composite system consisting of two two-level systems, independent Lorentz transformations performed on the two subsystems can be used to diagonalize an arbitrary density matrix in a way that respects separability. We have used this diagonalization to demonstrate the known fact that the Peres condition and the separability condition are equivalent in this case.

Although the diagonalization with Lorentz transformations is restricted to the composite 2×2 system, we have shown that the generalized form of the Schmidt decomposition used in this diagonalization can be extended to higher
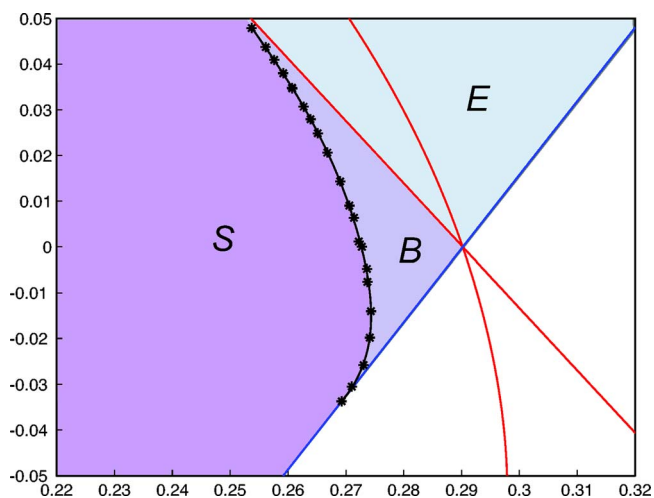

FIG. 6. (Color online) The boundary of the set of separable states in a two-dimensional section, as determined numerically. $\mathcal{S}$ denotes the set of separable states, $\mathcal{B}$ the states with bound entanglement, and $\mathcal{E}$ the entangled states violating the Peres condition. The straight line from lower left to upper right (in blue) is determined by the algebraic equation det $\rho=0$ and gives the boundary of the full set of density matrices. The two curves from upper left to lower right (in red) are determined by the equation det $\rho^P=0$, in particular, the red straight line gives the boundary of the set of states that satisfy the Peres condition. The black dots are the numerically determined points on the boundary of the set of separable states. The coordinate values in the plot are a factor of $\sqrt{2}$ too large according to the definition of distance in the text.

dimensions. The decomposition involves the use of nonunitary SL$(n,\mathbb{C})$ transformations for the two subsystems. Although a full diagonalization is not obtained in this way, we suggest that the Schmidt decomposed form may be of interest in the study of separability and bound entanglement.

A third part of the paper has been focused on the use of a numerical method to study separability. This method exploits the fact that the set of separable states is convex, and is based on an iterative scheme to find the closest separable state for arbitrary density matrices. We have demonstrated the use of this method in a numerical study of bound entanglement in the case of a 3×3 system. A further study of separability with this method is under way.

[1] M. Kus and K. Zyczkowski, Phys. Rev. A **63**, 032307 (2001).

[2] F. Verstraete, J. Dehaene, and B. De Moor, J. Mod. Opt. **49**, 1277 (2002).

[3] A. O. Pittenger and M. H. Rubin, Phys. Rev. A **67**, 012327 (2003).

[4] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[5] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[6] L. Gurvits, in *Proceedings of the Thirty-Fifth ACM Symposium on Theory of Computing* (ACM, New York, 2003), pp. 10-19.

[7] P. Arrighi and C. Patricot, Phys. Rev. A **68**, 042310 (2003).

[8] P. Arrighi and C. Patricot, J. Phys. A **36**, L287 (2003).

[9] L. M. Ioannou, B. C. Travaglione, D. C. Cheung, and A. K. Ekert, Phys. Rev. A **70**, 060303(R) (2004).

[10] F. Hulpke and D. Bruss, J. Phys. A **38**, 5573 (2005).

[11] P. Horodecki, M. Lewenstein, G. Vidal, and I. Cirac, Phys. Rev. A **62**, 032310 (2000).

[12] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002).

[13] P. Badziąg, P. Horodecki, and R. Horodecki, e-print quant-ph/0504041.

[14] G. Dahl, J. M. Leinaas, J. Myrheim, and E. Ovrum, Linear Algebra and its Applications (to be published).

[15] P. Horodecki, Phys. Lett. A **232**, 333 (1997).

[16] D. Bruss and A. Peres, Phys. Rev. A **61**, 030301(R) (2000).

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations* (John Hopkins University Press, Baltimore, MD, 1996).

[18] Li-Xiang Cen, Xin-Qi Li, and Yi Jing Yan, J. Phys. A **36**, 12267 (2003).

[19] A. Osterloh and J. Siewert, Phys. Rev. A **72**, 012337 (2005).