Husimi coordinates of multipartite separable states

Georges Parfionov and Roman R Zapatrin

Informatics Department, The State Russian Museum, Inženernaya 4, 191186, St Petersburg, Russia
E-mail: Roman.Zapatrin@gmail.com

Received 15 June 2010, in final form 4 October 2010
Published 24 November 2010
Online at stacks.iop.org/JPhysA/43/505304

Abstract
A parametrization of multipartite separable states in a finite-dimensional Hilbert space is suggested. It is proved to be a diffeomorphism between the set of zero-trace operators and the interior of the set of separable density operators. The result is applicable to any tensor product decomposition of the state space. An analytical criterion for separability of density operators is established in terms of the boundedness of a sequence of operators.

PACS numbers: 03.65.Aa, 03.65.Db, 03.67.Mn

1. Introduction
We explore multipartite quantum systems with state space $H^{(1)} \otimes \cdots \otimes H^{(N)}$, assuming the factors are Hilbert spaces of finite dimension:
$$\dim H^{(1)}, \ldots, \dim H^{(N)} < \infty.$$ A pure state of the system is a state whose density operator is a one-dimensional projector. Pure states of special kind:
$$P = P^{(1)} \otimes \cdots \otimes P^{(N)}$$
where $P^{(k)}$ is a one-dimensional projector in $H^{(k)}$ are called factorizable. A density operator $\rho$ is representable as a convex combination
$$\rho = \sum w_j P_j, \quad P_j \in P$$
where $P$ is the set of all pure product states is said to be separable. Such a description of the set of separable density operators is not constructive as their appropriate representation is not unique.

The main result of this paper is a uniform representation of separable density operators $\rho$ in the form
$$\rho = \frac{1}{Z} \int_{P} e^{\text{Tr}(B \cdot P)} P \, d\sigma(P)$$
(1)
where $B$ is a self-adjoint operator in $H$ such that $\text{Tr}B = 0$. This representation turns out to be unique, therefore, it can be treated as coordinatization of the set of separable density operators. Along these lines, a criterion to determine the separability of a given density operator $\rho$ is formulated.

The paper is organized as follows. In section 2, separable density operators are treated as barycenters of continuously distributed unit masses on the set $P$ of pure product states. This gives rise to equation (6) on which the coordinatization of separable density operators is based. In section 3, a scalar function (7) is studied discriminating the separability in a many-particle setting. This makes it possible to formulate a criterion of multipartite separability in terms of boundedness of certain sequence of operators. In section 4, we show that the range of the suggested coordinatization of separable density operators turns out to be the interior of the set of separable density operators, and the coordinatization is proved to be a diffeomorphism. In section 5, we study a degenerate case of 1-separability in order to demonstrate how the suggested approximation techniques can be used to check the separability of density operators.

2. Separable density operators viewed as barycenters

In this section we describe how density operators in a finite-dimensional Hilbert space can be described as barycenters of continuous probability distributions on a set of one-dimensional projectors.

2.1. From finite sums to continuous distributions

We study a multipartite quantum system whose state space $H$ is a finite tensor product of finite-dimensional Hilbert spaces $H = H^{(1)} \otimes \cdots \otimes H^{(N)}$. A density operator $\rho$ in $H$ is called separable whenever it is an element of the convex hull of the set $P$ of pure product states. Since $P$ is the set of extreme points of the set of all separable density operators, the Choquet theorem makes it possible to write down any such $\rho$ as an integral

$$\rho = \int_P P \, d\mu(P)$$

(2)

over certain probability measure $\mu$ on $P$. In other words, we may represent $\rho$ as the barycenter of certain distribution of unit mass on the set $P$. Such representation is known to be essentially non-unique. According to the Carathéodory theorem, this mass may have discrete finite distribution:

$$\rho = \sum w_\alpha P_\alpha$$

with the measure $\mu = \sum w_\alpha \delta_{P_\alpha}$ being a sum of atomic measures $\delta_{P_\alpha}$.

We emphasize that representation (2) also comprises continuous distributions:

$$\rho = \int_P P \, w(P) \, d\sigma(P)$$

(3)

where $w(P)$ is a positive continuous function and the integration is carried out over the probability measure $d\sigma$ invariant with respect to all local unitary transformations in $H$.

So, a density operator $\rho$ in the product space $H = H^{(1)} \otimes \cdots \otimes H^{(N)}$ is separable when it is a barycenter of a continuous distribution $w(P)$ on the set $P$ of pure product vectors. In other words, that means that a function $w(P)$ exists satisfying the equation

$$\int_P P \, w(P) \, d\sigma(P) - \rho = 0.$$
2.2. Main equation

In this section we confine ourselves to distributions \( w(P) \) of specific form and explore the existence of appropriate solutions.

Let us first try to find solutions of (4) of the form

\[
w(P) = e^{\text{Tr}(BP)}
\]  

where \( B \) is an Hermitian operator. Then solving equation (4) reduces to finding an operator \( X \) satisfying what we call the main equation

\[
\int_P e^{\text{Tr}(XP)} d\sigma(P) - \rho = 0.
\]  

In order to verify the existence of the solution of (6), we introduce the function

\[
G(X) = \int_P e^{\text{Tr}(XP)} d\sigma(P) - \text{Tr}(X\rho)
\]

whose gradient is

\[
\nabla G(X) = \int_P e^{\text{Tr}(XP)} d\sigma(P) - \rho.
\]

So, solving (6) reduces to finding an extremum of the function \( G(X) \). Taking into account that \( G(X) \) is convex, the solution of (6) exists only when the minimum of \( G(X) \) exists. In particular, when \( \rho \) is entangled, there is no way for the function \( G(X) \) to have a minimum.

**Theorem 1.** If the density operator \( \rho \) is entangled, then

\[
\inf G(X) = -\infty.
\]

**Proof.** The set of all separable states is closed, therefore, if \( \rho \) is not separable, there exists a hyperplane, defined by a self-adjoint operator \( X \) such that \( \forall P \in P \quad \text{Tr}(PX) < 0 \), while \( \text{Tr}(\rho X) > 0 \). Denote

\[
a = \text{Tr}(\rho X),
\]

\[
b = \max \text{Tr}(PX).
\]

Then \( a > 0, b < 0 \), so \( e^{\text{Tr}(XP)} \leq e^{kb} \) and

\[
G(kX) \leq \int e^{kb} d\sigma(P) - ka \to -\infty
\]

as \( k \to \infty \). \( \Box \)

So far, finding conditions for the existence of the minimum becomes essential to judge if \( \rho \) is entangled or not.

3. Exponential distributions and their approximations

3.1. An interlude on minima of convex functions

**Proposition 2.** Let a convex function \( F(X) \) have a minimum on a finite-dimensional space \( E \) and the minimal point is unique. Then

\[
\lim_{X \to \infty} F(X) = +\infty.
\]
Proof. With no loss of generality, we assume the minimum to be attained at 0, then $F(X) > F(0)$ for any $X \neq 0$. Suppose $\lim_{X \to \infty} F(X) \neq +\infty$; then there exists such $M$ and such sequence $X_k \to \infty$ that $F(X_k) \leq M$. Then

$$\lambda_k = \frac{1}{\|X_k\|} \to 0.$$  

Denoting

$$E_k = \frac{X_k}{\|X_k\|} = \lambda_k X_k$$

we obtain a bounded sequence in the finite-dimensional space $E$, which contains a converging subsequence. With no loss of generality, we denote this subsequence $E_k$ and its limit

$$E = \lim E_k, \quad \|E\| = 1.$$  

Then

$$E_k = \lambda_k X_k = (1 - \lambda_k)0 + \lambda_k X_k$$

so, due to the convexity of $F$

$$F(E_k) \leq (1 - \lambda_k)F(0) + \lambda_k F(X_k);$$

therefore,

$$F(E_k) \leq (1 - \lambda_k)F(0) + \lambda_k M$$

for sufficiently large $k$. Taking the limit and recalling that $\lambda_k \to 0$, and using the continuity of convex functions (see, e.g. [1]), we come to the contradiction: $F(E) \leq F(0)$. □

So far, we obtained a necessary condition for the unique minimum of a convex function to exist. This condition (9) is also sufficient for a minimum (not necessarily strict) to exit. In the following we need a more verifiable sufficient condition.

Proposition 3. Let $F(X)$ be a convex function on a finite-dimensional space $E$. If

$$\forall X \neq 0 \quad \lim_{t \to \infty} F(tX) = +\infty,$$  

(10)

then there exists minimum of $F(X)$ on $E$.

Proof. First prove that

$$\lim_{X \to \infty} F(X) = +\infty.$$  

Suppose that this is not the case, then there exists a number $M$ and a sequence $X_k \to \infty$ such that $F(X_k) \leq M$ for all $k$. Denoting

$$E_k = \frac{X_k}{\|X_k\|}$$

we obtain a bounded sequence in the finite-dimensional space $E$, which contains a converging subsequence. With no loss of generality, we denote this subsequence $E_k$ and its limit

$$E = \lim E_k, \quad \|E\| = 1.$$  

Now take an arbitrary $t \geq 0$ and $N$ such that $\|X_k\| \geq t$ for all $k \geq N$. Then $\|tE_k\| = t \leq \|X_k\|$. Denote

$$\lambda_k = \frac{t}{\|X_k\|};$$

4
then $0 \leq \lambda_k \leq 1$ and $t E_k = \lambda_k X_k$. We may treat $t E_k$ as convex combination

$$t E_k = (1 - \lambda_k)0 + \lambda_k X_k$$

and apply Jensen’s inequality

$$F(t E_k) \leq (1 - \lambda_k)F(0) + \lambda_k F(X_k) \leq F(0) + M$$

for all $k \geq N$. Taking the limit, we get $F(t E) \leq F(0) + M$, which contradicts the statement $F(t E) \to \infty$.

Now let us prove the existence of minimum. For that, consider the set

$$K = \{ X \mid F(X) \leq F(0) \}.$$  

The function $F$ is convex (and hence continuous), so the set $K$ is closed. As proved above, $\lim_{X \to \infty} F(X) = +\infty$, therefore the set $K$ is bounded.

The space $E$ is finite-dimensional and the set $K$ is compact; therefore, the function $F$ always attains minimum on $K$, which is its minimum on the whole space $E$. □

So far, we have obtained a sufficient condition for the existence of the minimum of a convex function.

### 3.2. Multipartite setting

Now return to our initial setting. We are dealing with integrals over the set $P$ of pure product projectors in a tensor product space $H = H^{(1)} \otimes \cdots \otimes H^{(N)}$. We need the following.

**Proposition 4** (Multipolarization identity). *Any quadratic form $Y$ in $H$ is completely defined by its values on $P$:*

$$\langle e^{(1)} \otimes \cdots \otimes e^{(N)} | Y | f^{(1)} \otimes \cdots \otimes f^{(N)} \rangle = \sum_{k^{(1)}, \ldots, k^{(N)}=0}^3 \frac{i^{(k^{(1)}+\cdots+k^{(N)})}}{4^N} \left( \bigotimes_{I=1}^N (e^{(I)} + i k^{(I)} f^{(I)}) \bigg| Y \bigg| \bigotimes_{I=0}^N (e^{(I)} + i k^{(I)} f^{(I)}) \right). \tag{11}$$

**Proof.** Verified by direct calculation. □

**Remark.** This is a multipartite generalization of usual polarization identity in complex space:

$$4\langle e | Y | f \rangle = \langle (e + f) | Y | (e + f) \rangle + i \langle (e + i f) | Y | (e + i f) \rangle - \langle (e - f) | Y | (e - f) \rangle - i \langle (e - i f) | Y | (e - i f) \rangle.$$

In the following we shall use the following consequence of this proposition:

$$\forall P \in P \quad \text{Tr}(XP) = 0 \quad \Rightarrow \quad X = 0. \tag{12}$$

Now return to the convex function (7)

$$G(X) = \int_P e^{\text{Tr}(XP)} \, d\sigma(P) - \text{Tr}(X\rho).$$

**Proposition 5.** *When the function $G(X)$ has minimum, this minimum is strict.*

**Proof.** Suppose there are two minimal points $B_0 \neq B_1$ of $G(X)$. Denote $V = B_1 - B_0$ and consider a family $B_t = B_0 + tV$. Consider the function

$$g(t) = G(B_t) = \int_P e^{\text{Tr}(B_tP)} \, d\sigma(P) - \text{Tr}(B_t\rho)$$

then $0 \leq \lambda_k \leq 1$ and $t E_k = \lambda_k X_k$. We may treat $t E_k$ as convex combination

$$t E_k = (1 - \lambda_k)0 + \lambda_k X_k$$

and apply Jensen’s inequality

$$F(t E_k) \leq (1 - \lambda_k)F(0) + \lambda_k F(X_k) \leq F(0) + M$$

for all $k \geq N$. Taking the limit, we get $F(t E) \leq F(0) + M$, which contradicts the statement $F(t E) \to \infty$.

Now let us prove the existence of minimum. For that, consider the set

$$K = \{ X \mid F(X) \leq F(0) \}.$$  

The function $F$ is convex (and hence continuous), so the set $K$ is closed. As proved above, $\lim_{X \to \infty} F(X) = +\infty$, therefore the set $K$ is bounded.

The space $E$ is finite-dimensional and the set $K$ is compact; therefore, the function $F$ always attains minimum on $K$, which is its minimum on the whole space $E$. □

So far, we have obtained a sufficient condition for the existence of the minimum of a convex function.

### 3.2. Multipartite setting

Now return to our initial setting. We are dealing with integrals over the set $P$ of pure product projectors in a tensor product space $H = H^{(1)} \otimes \cdots \otimes H^{(N)}$. We need the following.

**Proposition 4** (Multipolarization identity). *Any quadratic form $Y$ in $H$ is completely defined by its values on $P$:*

$$\langle e^{(1)} \otimes \cdots \otimes e^{(N)} | Y | f^{(1)} \otimes \cdots \otimes f^{(N)} \rangle = \sum_{k^{(1)}, \ldots, k^{(N)}=0}^3 \frac{i^{(k^{(1)}+\cdots+k^{(N)})}}{4^N} \left( \bigotimes_{I=1}^N (e^{(I)} + i k^{(I)} f^{(I)}) \bigg| Y \bigg| \bigotimes_{I=0}^N (e^{(I)} + i k^{(I)} f^{(I)}) \right). \tag{11}$$

**Proof.** Verified by direct calculation. □

**Remark.** This is a multipartite generalization of usual polarization identity in complex space:

$$4\langle e | Y | f \rangle = \langle (e + f) | Y | (e + f) \rangle + i \langle (e + i f) | Y | (e + i f) \rangle - \langle (e - f) | Y | (e - f) \rangle - i \langle (e - i f) | Y | (e - i f) \rangle.$$  

In the following we shall use the following consequence of this proposition:

$$\forall P \in P \quad \text{Tr}(XP) = 0 \quad \Rightarrow \quad X = 0. \tag{12}$$

Now return to the convex function (7)

$$G(X) = \int_P e^{\text{Tr}(XP)} \, d\sigma(P) - \text{Tr}(X\rho).$$

**Proposition 5.** *When the function $G(X)$ has minimum, this minimum is strict.*

**Proof.** Suppose there are two minimal points $B_0 \neq B_1$ of $G(X)$. Denote $V = B_1 - B_0$ and consider a family $B_t = B_0 + tV$. Consider the function

$$g(t) = G(B_t) = \int_P e^{\text{Tr}(B_tP)} \, d\sigma(P) - \text{Tr}(B_t\rho)$$
for \( t \in [0, 1] \). This function is constant because \( G \) is convex. Therefore, its second derivative vanishes, but

\[
g''(t) = \int_P e^{\text{Tr}(B,P)} (\text{Tr}(V P))^2 d\sigma(P). \tag{13}
\]

In the meantime \( V \neq 0 \); therefore, \( \int_P (\text{Tr}(V P))^2 d\sigma(P) > 0 \) due to (12), so \( g''(t) > 0 \)—contradiction. \( \square \)

However, the existence of the minimum of \( G(X) \) still cannot be directly verified. In the meantime, the approximations of \( G(X) \) by the family of convex functions

\[
G_M(X) = \int_P \left( 1 + \frac{\text{Tr}(X P)}{2M} \right)^{2M} d\sigma(P) - \text{Tr}(X \rho) \tag{14}
\]

possess the following property.

**Proposition 6.** Whatever \( \rho \), for any \( M \) there exists minimum of the function \( G_M(X) \) and this minimum is strict.

**Proof.** First prove that the minimum exists. Fix an arbitrary \( X \neq 0 \) and apply the sufficient condition (10) by showing that \( G_M(t X) \to +\infty \). Let

\[
Y = I + \frac{t X}{2M};
\]

then

\[
G_M(t X) = \int_P \text{Tr}(Y P)^{2M} d\sigma(P) - t \text{Tr}(X \rho).
\]

For all \( Y \neq 0 \), introduce

\[
N(Y) = \int_P \frac{\text{Tr}(Y P)^{2M} d\sigma(P)}{\text{Tr}(Y^2)^M} \geq 0.
\]

Being homogeneous, \( N(Y) \) is completely defined by its values on the compact set defined by \( \text{Tr}(Y^2) = 1 \). Let us prove that \( N(Y) \) is strictly positive. Suppose \( N(Y) = 0 \) for some \( Y \), then \( \text{Tr}(Y P) = 0 \) for all product one-dimensional projectors \( P \in P \). That means, for all vectors \( e^{(1)}, \ldots, e^{(N)} \):

\[
\langle e^{(1)} \otimes \cdots \otimes e^{(N)} \vert Y \vert e^{(1)} \otimes \cdots \otimes e^{(N)} \rangle = 0.
\]

Then, by virtue of (12), \( Y = 0 \).

Being continuous function defined on a compact set, \( N(Y) \) attains its minimal value, denote it \( a \). Then

\[
G_M(t X) \geq a \text{Tr}(Y^2)^M - t \text{Tr}(X \rho)
\]

\[
= a t^{2M} \left( \text{Tr} \left( I - I + \frac{X}{2M} \right)^2 \right)^M - t \text{Tr}(X \rho) \to +\infty
\]

since \( a > 0 \) as \( N(Y) \) was proved to be strictly positive. So the minimum of \( G_M(X) \) exists.

In order to prove that the minimum is strict, we proceed in a way similar to proposition 5 with the only difference that the function \( g(t) \) has the form

\[
g(t) = G_M(B_t) = \int_P \left( 1 + \frac{\text{Tr}(B_t P)}{2M} \right)^{2M} d\sigma(P) - \text{Tr}(B_t \rho)
\]
and we check its $2M$-s derivative (rather than the second one):

$$g^{(2M)}(t) = \int_P \frac{(2M - 1)!}{(2M)^{2M-1}} (\text{Tr}(VP))^{2M} \, d\sigma(P)$$

and obtain the same contradiction. □

**Corollary.** Whatever (separable or entangled) $\rho$, it decomposes into

$$\rho = \int_P P \, \nu(P) \, d\sigma(P) \quad (15)$$

with

$$\nu(P) = \left(1 + \frac{\text{Tr}(BP)}{2M}\right)^{2M-1} \quad (16)$$

where $B$ is the minimal point of the function $G_M$, that is, $\nabla G_M(B) = 0$. We emphasize that the obtained decomposition (15) is in general not barycentric as density (16) may take negative values. Furthermore, for entangled $\rho$ density (16) will always take both positive and negative values. Let us consider it in more detail.

### 3.3. The convergence of the approximations

We start from the main equation (6)

$$\int_P P e^{\text{Tr}(XP)} \, d\sigma(P) - \rho = 0$$

which may not have solutions, and replace it with a sequence of its binomial approximations

$$\int_P P \left(1 + \frac{\text{Tr}(XP)}{2M}\right)^{2M-1} \, d\sigma(P) - \rho = 0 \quad (17)$$

having a unique solution for all $k = 1, 2, \ldots$. So, we have to study the conditions when these approximations (17) turn to the main equation. For that, we explore the convergence $G_M \to G$ of functions associated with the equations in question on the set $E$ of self-adjoint operators in $\mathbb{H}$.

**Theorem 7.** If the function $G(X)$ has strict minimum on $E$ attained in $B$, then the sequence $B_M$ of the minimal points of $G_M(X)$ converge to $B$.

**Proof.** First prove that the convergence $G_M \to G$ is uniform on any compact subset of $E$. For any $0 \leq a \leq n$ direct calculation yields

$$\max_{|x| \leq a} \left| e^x - \left(1 + \frac{x}{n}\right)^n \right| = \max \left\{ e^a - \left(1 + \frac{a}{n}\right)^n, e^{-a} - \left(1 - \frac{a}{n}\right)^n \right\}$$

therefore,

$$\left(1 + \frac{x}{n}\right)^n - e^x \to 0$$

uniformly on any finite interval in $\mathbb{R}$. For $n = 2M$

$$|G_M(X) - G(X)| \leq \int_P \left| \left(1 + \frac{\text{Tr}(XP)}{2M}\right)^{2M} - e^{\text{Tr}(XP)} \right| \, d\sigma(P).$$

In this case the appropriate distribution is a signed measure rather than a probability measure on product states, or, in terms of [7], a continuous local pseudomixture.
Since $X$ ranges over a compact set, $|\text{Tr}(XP)| \leq C$ for some $C$ not depending on $X$. The integration set $\mathbf{P}$ is compact, so the convergence $G_M \to G$ is uniform.

Now let us prove that the sequence $B_M$ tends to $B$. The minimum is strict, so $G(X) > G(B)$ for any $X \neq B$. The sphere $\|X - B\| = \varepsilon$ is compact, then there exists 
$$a = \min_{\|X-B\|=\varepsilon} G(X) - G(B) > 0.$$ 

Since $G_M \to G$ uniformly on compact sets, a number $N_\varepsilon$ exists such that for any $M > N_\varepsilon$
$$G(X) - \frac{a}{3} < G_M(X) < G(X) + \frac{a}{3} < G_M(X) - G(B).$$
for any $X : \|X - B\| \leq \varepsilon$. Then for any such $X$
$$2a = a - \frac{a}{3} \leq G(X) - G(B) - \frac{a}{3} < G_M(X) - G(B).$$

So,
$$\frac{2}{3}a + G(B) < G_M(X) \quad \forall X : \|X - B\| \leq \varepsilon. \quad (19)$$

It remains to check that $\|B_k - B\| \leq \varepsilon$ for all $k > N_\varepsilon$. Suppose a number $k > N_\varepsilon$ exists such that the appropriate minimal point $\|B_k - B\| > \varepsilon$. For $t = \varepsilon/\|B_k - B\|$ consider the convex combination
$$E = (1-t) \cdot B + tB_k$$
then $\|E - B\| = \varepsilon$. Then it follows from (19) and the convexity of $G_M$ that
$$\frac{2}{3}a + G(B) < G_M(E) \leq (1-t) \cdot G_M(B) + t G_M(B_k) \leq$$
$$\leq (1-t) \cdot G_M(B) + t G_M(B) = G_M(B) < \frac{a}{3} + G(B)$$
—contradiction. \(\square\)

**Proposition 8.** If the sequence $B_M$ of the minimal points of $G_M(X)$ converge, then the function $G(X)$ has strict minimum at point $B = \lim B_M$.

**Proof.** For any $X$
$$G_M(X) \geq G_M(B_k).$$

For any fixed $X$ the sequence $G_M(X) \to G(X)$, so
$$G(X) = \lim G_M(X) \geq \lim G_M(B_M). \quad (20)$$

Now check that
$$\lim G_M(B_M) = G(B).$$

The convergence $B_j \to B$ and $G_M \to G$ together with the continuity of each $G_M(X)$ imply
- $\forall M \; G_M(B_j) \to G_M(B)$
- $G_M(B) \to G(B)$.

Applying the Cantor diagonal method, we choose a subsequence $B_{j\mu}$ such that $G_M(B_{j\mu}) \to G(B)$. For arbitrary $X, Y$ we have
$$|G_M(X) - G_M(Y)| \leq$$
$$\leq \int_{\mathbf{P}} \left|\left(1 + \frac{\text{Tr}(XP)}{2M}\right)^{2M} - \left(1 + \frac{\text{Tr}(YP)}{2M}\right)^{2M}\right| d\sigma(P) + \|X - Y\| \cdot \|\rho\|. \quad (19)$$
Observing that for any \(x, y \in \mathbb{R}\)
\[
\left|\left(1 + \frac{x}{n}\right)^n - \left(1 + \frac{y}{n}\right)^n\right| \leq e^{\|x+y\|} \cdot |x - y|
\]
and taking into account that \(\text{Tr}(XP) \leq \|X\|\), we get
\[
|G_M(X) - G_M(Y)| \leq (e^{\|X+Y\|} + \|\rho\|) \cdot \|X - Y\|.
\]
Using this, we have
\[
|G_M(B_M) - G_M(B_jM)| \leq (e^{\|B_M\|} + \|B_jM\| + \|\rho\|) \cdot \|B_M - B_jM\| \to 0.
\]
Therefore, \(G(B) = \lim G_M(B_M)\), so it follows from (20) that \(B\) is the minimal point of \(G(X)\). In accordance with the proposition 5 this minimum is strict. □

The obtained criterion can be strengthened.

**Theorem 9.** If the sequence \(B_M\) of the minimal points of \(G_M(X)\) is bounded, then it converges, and the function \(G(X)\) has strict minimum at \(B = \lim B_M\).

**Proof.** Since the space \(\mathcal{E}\) is finite-dimensional, we can select its converging subsequence \(B_{M_n} \to B\). Let us first show that \(B = \min G(X)\). For any fixed \(X\)
\[
G(X) = \lim G_{M_n}(X) \geq \lim G_{M_n}(B_{M_n}).
\]
Now check that
\[
\lim G_{M_n}(B_{M_n}) = G(B).
\]
Proceeding in a way similar to theorem 6, we select a sub-subsequence \(B_{M_{n_j}}\) such that \(G_{M_n}(B_{M_{n_j}}) \to G(B)\) and get the required: the function \(G(X)\) has minimum. Then, applying theorem 5 we infer that the minimum is strict and using theorem 7 we obtain that \(B_{M_n} \to B\). □

### 4. ‘Temperature’ theorem and its consequences

The mapping
\[
\mathcal{L}(X) = \int_{\mathcal{P}} e^{\text{Tr}(XP)} P \, d\sigma(P)
\]
from all self-adjoint operators to positive operators in \(\mathcal{H}\) was shown (proposition 5) to be injective. However, \(\mathcal{L}\) is not surjective: clearly, no pure product state can be represented this way. We shall show that the image of \(\mathcal{L}\) contains almost all separable density operators.

#### 4.1. Matching theorem

The decompositions of a given density operator \(\rho\), both discrete and continuous, are known to be non-unique. The following theorem shows that any continuous positive decomposition of \(\rho\) can be replaced by an exponential one, which we studied before.

**Theorem 10 (‘Temperature’ theorem).** Let \(\rho\) be a density matrix such that it can be represented in form (3)
\[
\rho = \int_{\mathcal{P}} w(P) \cdot P \, d\sigma(P)
\]
with \(w\) being positive
\[
w(P) > 0.
\]
Then there exists such self-adjoint operator $B$ where

$$\rho = \int_{P} e^{\text{Tr}(BP)} \, P \, d\sigma(P).$$

**Proof.** Taking into account (8), it suffices to prove that the function $G(X) = \int_{P} e^{\text{Tr}(XP)} \, d\sigma(P) - \text{Tr}(X\rho)$ has minimum. Writing down $\rho$ in form (3), we get

$$G(X) = \int_{P} e^{\text{Tr}(XP)} \, d\sigma(P) - \int_{P} w(P) \text{Tr}(XP) \, d\sigma(P) = \int_{P} (e^{\text{Tr}(XP)} - w(P) \text{Tr}(XP)) \, d\sigma(P). \quad (23)$$

For every fixed $P \in P$ introduces a function

$$g(t) = e^{t} - at \quad (24)$$

where

$$a = w(P), \quad t = \text{Tr}(XP).$$

The set $P$ is compact, and the function $w(P)$ is continuous and positive, which is why it attains its extreme values

$$0 < m \leq w(P) \leq M.$$

Since $w(P)$ is a probability density, we have

$$0 < m \leq 1 \leq M. \quad (25)$$

The value of $a$ in (24) lies between $m$ and $M$. The minimal value of $g(t)$ depends on the value of the parameter $a$ as follows:

$$g_{\text{min}} = a - a \ln a.$$

By elementary (for $t \leq 0$) and routine (for $t \geq 0$) calculations we obtain

$$g(t) \geq (m + M)[1 - \ln(m + M)] + m|t| \quad (26)$$

for any real $t$ as illustrated in the following graph:

Since $t = \text{Tr}(XP)$ and $a = w(P)$, this means

$$e^{\text{Tr}(XP)} - w(P)\text{Tr}(XP) \geq (m + M)[1 - \ln(m + M)] + m|\text{Tr}(XP)|.$$

Integrating this inequality over $P$, we obtain the following evaluation:

$$G(X) \geq (m + M)[1 - \ln(m + M)] + m \int_{P} |\text{Tr}(XP)| \, d\sigma(P). \quad (27)$$
Consider the second summand of the following expression:

\[ \nu(X) = \int_{P} |\text{Tr}(XP)| \, d\sigma(P) \]

which is a seminorm. Proceeding as in proposition 6 we see that \( \nu(X) \) is non-degenerate. Since all norms in finite-dimensional space are equivalent, \( \nu(X) \) evaluates as

\[ \nu(X) \geq \text{const} \cdot \|X\|. \]

So far, we obtain

\[ G(X) \geq (m + M)[1 - \ln(m + M)] + \text{const} \cdot \|X\| \rightarrow +\infty. \]

Since \( G(X) \) is convex, it has minimum, which is attained at certain point \( B \in E \). \( \square \)

**Remark.** It can be verified that the distribution \( w(P) = e^{\text{Tr}(BP)} \) yields the maximum for

\[ S = - \int w \ln w \, d\sigma(P). \]

**Theorem 11** (Matching theorem). Any \( \rho \) belonging to the interior \( D \) of the set of separable states can be represented in form (5):

\[ \rho = \int_{P} e^{\text{Tr}(BP)} \, P \, d\sigma(P). \]

**Proof.** Denote by \( \mathcal{M} \) the set of density operators representable as Husimi exponentials. Let \( \rho_1, \rho_2 \in \mathcal{M} \) and form their convex combination \( \rho = (1 - t)\rho_1 + t\rho_2 \). This \( \rho \) can be represented by continuous positive density \( w(P) = (1 - t)\rho_1 + t\rho_2 \). Therefore, it follows from the 'temperature' theorem 10 that such \( B \) exists that \( \rho = \int e^{\text{Tr}(BP)} \, P \, d\sigma(P) \). That means, \( \mathcal{M} \) is convex together with its closure.

Let \( P_0 \in P \) be a product pure state. Any atomic measure on a compact set \( P \) can be approximated by a sequence of strictly positive densities. Take such a sequence \( w_n \) on \( P \) such that for any continuous function \( f \) on \( P \):

\[ \lim_{n \to \infty} \int_{P} f(P)w_n(P) \, d\sigma(P) = f(P_0). \]

Consider the sequence of operators

\[ \rho_n = \int_{P} w_n(P) \, P \, d\sigma(P) \]

each belonging to \( \mathcal{M} \). For any affine function \( h(X) \)

\[ h(\rho_n) = \int_{P} w_n(P)h(P) \, d\sigma(P) \rightarrow h(P_0); \]

therefore, \( \rho_n \to P_0 \). So far, the closure of \( \mathcal{M} \) contains \( P \). The closure is shown to be convex, therefore it contains all separable density operators.

Now consider the 'partition function'

\[ Z(X) = \int_{P} e^{\text{Tr}(XP)} \, d\sigma(P). \quad (28) \]

Note that for any \( V \neq 0 \)

\[ \left. \frac{d^2 Z(V + Vt)}{dt^2} \right|_{t=0} = \int_{P} e^{\text{Tr}(XP)}(\text{Tr}(VP))^2 \, d\sigma(P) > 0 \]
whose positivity was established in (13). However $d^2 Z = d\mathcal{L}$, where $\mathcal{L} = \nabla Z$, see (22). Then the mapping $\mathcal{L}(X)$ is non-degenerate at every point $X \in \mathcal{E}$. According to the inverse image theorem, we see that $\mathcal{L}$ is local diffeomorphism, that is, each point $X$ has a neighborhood whose image under the mapping $\mathcal{L}$ is open in $\mathcal{E}$. So, we conclude that its full image $\mathcal{L}(\mathcal{E})$ is an open subset in $\mathcal{E}$. In the meantime

\[ M = \mathcal{L}(\mathcal{E}) \cap \{ X \in \mathcal{E} \mid \text{Tr}X = 1 \} \]

so, $M$ is an open subset of the set of separable density operators: $M \subseteq \mathcal{D}$.

So far, we have shown that the set $M$ is an open, convex and dense subset of the set of all separable density operators; therefore, $M$ coincides with its interior: $M = \mathcal{D}$. □

4.2. Husimi coordinatization

Starting from the ‘partition function’ $Z$ defined above (28) we introduce

\[ W = \ln Z = \ln \int_{P} \text{e}^{\text{Tr}(XP)} \, d\sigma(P) \]

and first calculate its gradient

\[ \nabla W(X) = \frac{\nabla Z}{Z} = \frac{1}{Z} \int_{P} \text{e}^{\text{Tr}(XP)} P \, d\sigma(P). \]

**Proposition 12.** The quadratic form $d^2 W$ is non-negatively defined and vanishes only on scalar operators.

**Proof.** Calculate the value of $d^2 W$ at point $X$ on element $V$:

\[ d^2 W(X; V) = \left. \frac{d^2 W(X + Vt)}{dt^2} \right|_{t=0} = \frac{d}{dt} \left. \left( \frac{1}{Z} \int_{P} \text{e}^{\text{Tr}(X+Vt)P} \text{Tr}(VP) \, d\sigma(P) \right) \right|_{t=0}. \]

Denoting

\[ w(P) = \frac{\text{e}^{\text{Tr}(XP)}}{Z} \]

it reads

\[ d^2 W(X; V) = \int_{P} (\text{Tr}(VP))^2 w(P) \, d\sigma(P) - \left( \int_{P} \text{Tr}(VP) w(P) \, d\sigma(P) \right)^2 \geq 0. \]

This expression vanishes if and only if $\text{Tr}(VP) = \text{const}$ for every $P \in \mathcal{P}$. All such $V$ are the scalar operators: $V = \lambda I$.

Note that for any scalar operator the function $W$ has the property

\[ W(X + \lambda I) = W(X) + \lambda; \]

therefore, its gradient

\[ \chi(X) = \nabla W(X) \]

is invariant with respect to shifts along scalar operators:

\[ \chi(X + \lambda I) = \chi(X). \]

This shows that the mapping $\chi$ is not injective on $\mathcal{E}$, but its restriction to the set $\mathcal{N}$ of traceless operators from $\mathcal{E}$ becomes injective.
Theorem 13. Mapping (32) establishes a diffeomorphism
\[ \chi : \mathcal{N} \to \mathcal{D} \] (34)
between the set \( \mathcal{N} \) of all traceless self-adjoint operators in \( \mathcal{H} \) and the interior \( \mathcal{D} \) of the set of all separable density operators.

Proof. The mapping \( \chi \) was shown to be invariant with respect to shifts along scalar operators, so \( \chi(\mathcal{N}) = \chi(\mathcal{E}) \). The differential \( d\chi = d^2W \). It follows from proposition 12 that \( d^2W > 0 \) on \( \mathcal{N} \); therefore, \( d\chi \) is non-degenerate on \( \mathcal{N} \). So, \( \chi \) is a diffeomorphism of \( \mathcal{N} \) on the image \( \chi(\mathcal{N}) \). In turn, the image of \( \chi \) coincides with the image of the mapping (22), the density operators representable as (5). Finally, as shown in the matching theorem 11, this set coincides with \( \mathcal{D} \), the interior of the set of all separable density operators.

\[ \square \]

5. The role of approximations: a toy model

In this section we show how Husimi techniques work on a simplest, degenerate example. We prove that any mixed state of 1-partite quantum system is 1-separable. The result itself has no practical value, but it shows a way for implementation of the method of continuous distributions. Namely, verifying the membership of a mixed state to the set of separable states is reduced to finding a common evaluation for a sequence of approximations.

Theorem 14. Any density operator in the state \( \mathcal{H} \) of finite dimension \( n \) can be represented as a barycenter of a positive measure on unit vectors.

Proof. We prove this using theorem 9. It will suffice to show that the sequence of the solutions \( B_M \) of the equation
\[ \int \left( 1 + \frac{\text{Tr}(B_M P)}{2M} \right)^{2M-1} P \, d\sigma(P) - \rho = 0 \] (35)
is bounded (the integration here is carried out over the unit sphere \( \mathcal{S} \) in the state space \( \mathcal{H} \)). Let us evaluate the eigenvalues of \( B_M \), denote them \( b^M_1 \leq \cdots \leq b^M_n \) appropriately ordered. That is, for any \( P \)
\[ b^M_1 \leq \text{Tr}(B_M P) \leq b^M_n \).

Substituting this inequality to the operator equation (35) and taking the trace, we obtain
\[ \left( 1 + \frac{b^M_1}{2M} \right)^{2M-1} \leq \text{Tr}\rho \leq \left( 1 + \frac{b^M_n}{2M} \right)^{2M-1} \].

Since we are dealing with a density operator \( \rho \), its trace equals one, therefore \( b^M_1 \leq 0 \leq b^M_n \).

Now let us evaluate the difference \( b^M_n - b^M_1 \) between the greatest and the least eigenvalues of \( B_M \) from (35). According to (35), the \( j \)th eigenvalue of \( \rho \) reads:
\[ \lambda_j = \int_{t_1 + \cdots + t_n \leq 1} \left( 1 + \frac{b^M_1 t_1 + \cdots + b^M_n t_n - \cdots - t_0}{2M} \right)^{2M-1} \frac{t_j \, dt_1 \cdots d_{t-1}}{2^M} \].

The idea is to replace the integration over the simplex \( t_1 + \cdots + t_n \leq 1 \) by iterated integration over \( t_1 \) first and then over the simplex \( t_2 + \cdots + t_n \leq 1 \). Omitting the details (see appendix in [2]), the difference \( (b^M_n - b^M_1) \lambda_1 \) is shown to satisfy
\[ (b^M_n - b^M_1) \lambda_1 \leq \text{Tr}\rho + \int \left( 1 + \frac{\text{Tr}(B_M P)}{2M} \right)^{2M-1} \frac{\text{Tr}(B_M P)}{2M} \, d\sigma(P). \]
Evaluating the above integral, we obtain
\[(b_n^M - b_1^M)\lambda_1 \leq 1 + \frac{b_n^M - b_1^M}{2M}\]
so \((b_n^M - b_1^M)(\lambda_1 - \frac{1}{2M}) \leq 1\). That is, for any \(M > \lambda_1\) we have \(\lambda_1 - \frac{1}{2M} \geq \frac{\lambda_1}{2}\); therefore,
\[(b_n^M - b_1^M) \leq 2/\lambda_1.\] (36)

This means that for sufficiently big \(M\) the spectra \(\{b_1^M, \ldots, b_n^M\}\) of \(B_M\) remain in the compact set not depending on \(M\):
\[
\begin{align*}
    b_1^M &\leq 0 \leq b_n^M, \\
    (b_n^M - b_1^M) &\leq 2/\lambda_1,
\end{align*}
\]
where \(\lambda_1\) is the smallest eigenvalue of \(\rho\). Then, theorem 9 tells us that the limit \(B_M\) always exists, so any state \(\rho\) is 1-separable. 

We emphasize that theorem 35 only demonstrates how Husimi techniques can be implemented to check if a density matrix is separable or not: for that, a uniform evaluation of the spectrum of all \(B_M\)-s is needed.

6. Concluding remarks

The central result of this paper is Husimi coordinatization — the explicit diffeomorphism \(\rho \leftrightarrow B\)
\[
\rho = \frac{1}{Z} \int \rho P \exp(B \rho P) \rho d\sigma(P)
\]
between the set of all traceless Hermitian operators in the state space \(H\) of a multipartite quantum system and the interior of the set of all separable density operators in \(H\).

The second, more practical result is the approximation techniques for evaluating the parameter \(B\), which were shown to be relevant for checking the separability of a given multipartite mixed state \(\rho\). These techniques are still to be developed, an degenerate example of their application is shown on a toy model in section 5.

In this work, we used common thermodynamical ideas. However, it was not possible to replant directly the standard techniques. The first reason is that the analog of inverse temperature is no longer a scalar \(\beta\), it becomes self-adjoint operator \(B\). That is why the standard thermodynamical ‘temperatur theorem’ had to be reproved in the new setting (section 4).

The overall construction is general enough: we do not dwell on any particular decomposition of the state space \(H\) into tensor product—this notion was shown to be relative to measurements [3]. The results are applicable to any finite multipartite structure.

From a ‘historical perspective’, the idea to treat the density operator as a functional on pure states belongs to Husimi [4]: continuous distributions over pure states were first considered in [5], while the idea to employ uniformly distributed (though, still finite) subsets to decompose separable states is due to [6]. In [7], decompositions of any mixed states into local pseudomixtures were studied, among which minimal ones were found. In this paper, rather, we focus (in terms of [7]) on maximal continuous local pseudomixtures.

Acknowledgments

The authors acknowledge the discussion of the results offered by the participants of the Laboratory of Quantum Information (St Petersburg, Russia), essential remarks made by JPA referees were gratefully taken into account. Support from research grant RFFI 1006-00178 is appreciated.
References

[1] Nikaido H 1968 Convex structures and economic theory (New York: Academic)
[2] Parfionov G and Zapatrin R R 2009 Guessing quantum ensemble using laplace principle arXiv:0901.3771v1 [quant-ph]
[3] Zanardi P, Lidar D and Lloyd S 2004 Quantum tensor product structures are observable-induced Phys. Rev. Lett. 92 060402
[4] Husimi K 1940 Some formal properties of the density matrix Proc. Phys. Math. Soc. Japan 22 264–314
[5] Jozsa R, Robb D and Wootters W K 1994 Lower bound for accessible information in quantum mechanics Phys. Rev. A 49 668
[6] Vidal G and Tarrach R 1999 Robustness of entanglement Phys. Rev. A 59 141
[7] Sanpera A, Tarrach R and Vidal G 1998 Quantum inseparability as local pseudomixture Phys. Rev. A 58 826