Definition of entanglement for pure and mixed states

seminar talk given by

Marius Krumm

in the master studies seminar course

„Selected Topics in Mathematical Physics: Quantum Information Theory “

at the University of Heidelberg

Abstract: A keypoint in which quantum mechanics differs from classical mechanics is the possibility of (quantum) entanglement. Entanglement is an important resource in quantum information theory and quantum computation and can be used to realize many surprising effects like quantum teleportation. This document is a written elaboration of the seminar lecture *Definition of entanglement for pure and mixed states* by Marius Krumm as part of the master studies seminar course *Selected Topics in Mathematical Physics: Quantum Information Theory* at the University of Heidelberg. Its purpose is to give an introduction to entanglement. In the beginning, we define entanglement for pure and mixed states. Afterwards, the von Neumann entropy is introduced as an entanglement measure for pure states. At last, there will be a short discussion of entanglement witnesses.

1 Introduction

A fundamental element of quantum information science is quantum entanglement. It can be used to realize surprising effects like quantum teleportation and is considered as an important resource in quantum computation. The aim of this document is to give a first introduction to quantum entanglement and its quantification. This knowledge is collected from the sources found at the end of this document, which are also a good starting point to obtain further knowledge.

2 Quantum Entanglement

Before we define *entanglement*, at first we provide a tool that will prove useful several times.

Theorem 1 (Schmidt decomposition)**.** *Let $\mathcal{H}_A$, $\mathcal{H}_B$ be two Hilbert spaces and $|\psi_{AB}\rangle$ a normalized state in $\mathcal{H}_A \otimes \mathcal{H}_B$. Then there exist orthonormal sets $\{|j_A\rangle\}$, $\{|j'_B\rangle\}$ in $\mathcal{H}_A$, $\mathcal{H}_B$, such that:*

$$|\psi_{AB}\rangle = \sum_j \sqrt{p_j}\,|j_A\rangle \otimes |j'_B\rangle\,. \tag{2.1}$$

Here p_j are the non-zero eigenvalues of $\rho_A := \mathrm{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|)$.

Proof. We define $\rho := |\psi_{AB}\rangle\langle\psi_{AB}|$. Furthermore we expand $|\psi_{AB}\rangle$ in an orthonormal eigenbasis $\{|j_A\rangle\}$ of ρ_A (i.e. $\rho_A = \sum_j p_j |j_A\rangle\langle j_A|$) and an arbitrary orthonormal basis $\{|k''_B\rangle\}$ of $\mathcal{H}_B$: $|\psi_{AB}\rangle = \sum_{j,k} a_{jk}|j_A\rangle \otimes |k''_B\rangle$. Now we define $|\hat{j}_B\rangle := \sum_k a_{jk}|k''_B\rangle$ to obtain $|\psi_{AB}\rangle = \sum_j |j_A\rangle \otimes |\hat{j}_B\rangle$.
We claim, that these vectors are orthogonal: $\langle\hat{k}_B|\hat{j}_B\rangle \propto \delta_{jk}$. To prove this claim, we expand the density matrix: $\rho = \sum_{j,k}\big(|j_A\rangle \otimes |\hat{j}_B\rangle\big)\big(\langle k_A| \otimes \langle\hat{k}_B|\big)$. We take the partial trace:

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B\left(\sum_{j,k}\big(|j_A\rangle \otimes |\hat{j}_B\rangle\big)\big(\langle k_A| \otimes \langle\hat{k}_B|\big)\right) = \sum_m \langle m''_B|\left(\sum_{j,k}\big(|j_A\rangle \otimes |\hat{j}_B\rangle\big)\big(\langle k_A| \otimes \langle\hat{k}_B|\big)\right)|m''_B\rangle \\
&= \sum_{m,j,k}\Big[\langle m''_B|\hat{j}_B\rangle\langle\hat{k}_B|m''_B\rangle\Big]|j_A\rangle\langle k_A| = \sum_{j,k}\left[\langle\hat{k}_B|\left(\sum_m |m''_B\rangle\langle m''_B|\right)|\hat{j}_B\rangle\right]|j_A\rangle\langle k_A| \\
&= \sum_{jk}\langle\hat{k}_B|\hat{j}_B\rangle|j_A\rangle\langle k_A|
\end{aligned}$$

We compare to $\rho_A = \sum_j p_j|j_A\rangle\langle j_A|$ and indeed find $\langle\hat{k}_B|\hat{j}_B\rangle = p_j\delta_{jk}$. For those j with $p_j = 0$ we find $\langle\hat{j}_B|\hat{j}_B\rangle = 0$ and thus $|\hat{j}_B\rangle = 0$. For the j with $p_j \neq 0$ we can define $|j'_B\rangle := \frac{1}{\sqrt{p_j}}|\hat{j}_B\rangle$. Thus we have an orthonormal set ($\langle k'_B|j'_B\rangle = \delta_{jk}$) that allows us to write $|\psi_{AB}\rangle$ like in eq. (2.1). □

Corollary 1. *Let $|\psi_{AB}\rangle$ be a normalized state in $\mathcal{H}_A \otimes \mathcal{H}_B$. Then $\rho_A := \mathrm{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|)$ and $\rho_B := \mathrm{Tr}_A(|\psi_{AB}\rangle\langle\psi_{AB}|)$ have the same non-zero eigenvalues.*

Proof. We apply the Schmidt decomposition: $|\psi_{AB}\rangle = \sum_j \sqrt{p_j}\,|j_A\rangle \otimes |j'_B\rangle$.
Thus we find

$$\rho_B = \mathrm{Tr}_A\left(\sum_{j,k}\sqrt{p_j p_k}\big(\,|j_A\rangle \otimes |j'_B\rangle\,\big)\big(\,\langle k_A| \otimes \langle k'_B|\,\big)\right) = \sum_{m,j,k}\sqrt{p_j p_k}\,\langle m_A|j_A\rangle\,\langle k_A|m_A\rangle\,|j'_B\rangle\,\langle k'_B|$$
$$= \sum_j p_j\,|j'_B\rangle\,\langle j'_B|$$

This is the spectral decomposition of ρ_B, which shows that ρ_B also has the non-zero eigenvalues p_j. □

Definition 1 (Schmidt number)**.** *The **Schmidt number** or **Schmidt rank** of a pure state $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is defined as the number of non-zero eigenvalues of ρ_A or ρ_B:*

$$\mathrm{SR}(|\psi_{AB}\rangle) := \#\{j|p_j \neq 0\} \tag{2.2}$$

Definition 2 (Entanglement of pure states)**.** *Let $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a pure state. We call $|\psi_{AB}\rangle$ **entangled** or **non-separable** if its Schmidt number is larger than 1.*

Theorem 2. *Let $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a pure state. Then:*

$$|\psi_{AB}\rangle \text{ entangled} \Leftrightarrow \nexists\, |a_A\rangle \in \mathcal{H}_A, |b_B\rangle \in \mathcal{H}_B \text{ such that } |\psi_{AB}\rangle = |a_A\rangle \otimes |b_B\rangle$$

Proof. If $|\psi_{AB}\rangle$ is entangled, it cannot be written as $|a_A\rangle \otimes |b_B\rangle$, because that state would have Schmidt number 1. And vice versa, if $|\psi_{AB}\rangle$ cannot be written as $|a_A\rangle \otimes |b_B\rangle$, then its Schmidt decomposition contains at least 2 summands and thus the Schmidt number is at least 2. □

Definition 3. *Let $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a pure state, $dim(\mathcal{H}_A) \leq dim(\mathcal{H}_B)$. Then*

$$|\psi_{AB}\rangle \textbf{ maximally entangled} :\Leftrightarrow \rho_A = \frac{\mathbb{1}}{dim(\mathcal{H}_A)}$$

Example 1. *An example for a maximally entangled state (in a composite system of two spin-$\frac{1}{2}$ systems) is the Bell state $|\psi_{AB}\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle)$. It has Schmidt rank 2. It may describe the full composite quantum system. But only after a measurement we can tell if the spin of system A is up or down. But if have measured the spin of A, we also know the spin of B without any further measurement.*

Definition 4 (General definition of entanglement)**.** *A state ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is called* ***non-entangled/separable*** *if there exist density operators $\rho_A^{(j)}$, $\rho_B^{(j)}$ and $p_j \geq 0$ with $\sum_j p_j = 1$ such that:*

$$\rho = \sum_j p_j \rho_A^{(j)} \otimes \rho_B^{(j)} \tag{2.3}$$

The idea behind that definition is: systems described by $\rho_A^{(j)} \otimes \rho_B^{(j)}$ are completely independent of each other. Now assume two physicists Alice (system A) and Bob (system B) use **classical** communication: Alice rolls a dice, which has the result j with probability p_j, and tells Bob the result; Then Alice produces $\rho_A^{(j)}$ and Bob $\rho_B^{(j)}$. They repeat this procedure many times and forget, which j belongs to the samples. Then the resulting ensemble is described by eq. (2.3).

Theorem 3. *For a pure state $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, definition (4) reduces to definition (2).*

Proof. We first assume $\rho = \sum_j p_j \rho_A^{(j)} \otimes \rho_B^{(j)} = |\psi_{AB}\rangle \langle\psi_{AB}|$. We define $\rho_j := \rho_A^{(j)} \otimes \rho_B^{(j)}$ (again a density operator). Complete $|\psi_{AB}\rangle$ to an orthonormal basis $\{|\psi_{AB}\rangle, |\phi_1\rangle, |\phi_2\rangle ...\}$. Furthermore we perform an eigen-decomposition: $\rho_j = \sum_k a_k^{(j)} |a_k^{(j)}\rangle \langle a_k^{(j)}|$ with $a_k^{(j)} > 0$. Then we find $0 = \langle\phi_x| \rho_{AB} |\phi_x\rangle = \sum_{j,k} p_j a_k^{(j)} \langle\phi_x|a_k^{(j)}\rangle \langle a_k^{(j)}|\phi_x\rangle$ and thus $|\langle\phi_x|a_k^{(j)}\rangle|^2 = 0$. As $|a_k^{(j)}\rangle = \sum_x \langle\phi_x|a_k^j\rangle |\phi_x\rangle + \langle\psi_{AB}|a_k^{(j)}\rangle |\psi_{AB}\rangle$ we find $|a_k^{(j)}\rangle \propto |\psi_{AB}\rangle$. We thus found $\rho_j = \rho_{AB}$ and thus $\rho = \rho_A \otimes \rho_B$ for some ρ_A and ρ_B. Now we notice $\rho_{AB}^2 = \rho_A^2 \otimes \rho_B^2$ and thus $1 = \mathrm{Tr}(\rho_{AB}^2) = \mathrm{Tr}_A(\rho_A^2) \cdot \mathrm{Tr}_B(\rho_B^2)$ and thus $\mathrm{Tr}(\rho_A^2) = 1$, which means that ρ_A is pure and thus Schmidt number = 1.
Now we assume $|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\alpha_B\rangle$. Then $\rho_{AB} = |\phi_A\rangle \langle\phi_A| \otimes |\alpha_B\rangle \langle\alpha_B|$. □

Theorem 4. *ρ_{AB} separable $\Leftrightarrow \exists$ sets of normalized states $\{|a_A^{(j)}\rangle\}, \{|b_B^{(j)}\rangle\}$, $p_j \geq 0$ with $\sum_j p_j = 1$ such that $\rho_{AB} = \sum_j p_j |a_A^{(j)}\rangle \langle a_A^{(j)}| \otimes |b_B^{(j)}\rangle \langle b_B^{(j)}|$*

Proof.
$\Leftarrow$: We simply set $\rho_A^{(j)} := |a_A^{(j)}\rangle \langle a_A^{(j)}|$ and $\rho_B^{(j)} := |b_B^{(j)}\rangle \langle b_B^{(j)}|$.
$\Rightarrow$: We perform an eigenvalue decomposition of all $\rho_A^{(j)}$ and $\rho_B^{(j)}$. By relabeling and change of numeration we achieve the desired form. A full proof is found in Appendix A. □

So far we have only defined entanglement. Now we want to find a notion for how strongly a state is entangled. For a pure state, this is done by the von Neumann entropy.

Definition 5. *Let ρ be an arbitrary density operator. Then the* ***von Neumann entropy*** *is defined as:*

$$S(\rho) = -Tr(\rho \cdot \log_2 \rho) \tag{2.4}$$

If p_j are the eigenvalues of ρ, then we can rewrite the entropy as[1]

$$S(\rho) = -\sum_j p_j \cdot \log_2 p_j \tag{2.5}$$

Example 2.

1. $\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ *has eigenvalues 0 and 1* $\Rightarrow S(\rho) = 0$
2. $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$. *Then* $\rho = |\psi\rangle\langle\psi| \hat{=} \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. *As ρ has eigenvalues* $0, 1$, *again* $S(\rho) = 0$.
3. $\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \Rightarrow S(\rho) = \log_2(2)$

Theorem 5 (General properties of the von Neumann entropy)**.**

1. $S(\rho) \geq 0$ *and* $S(\rho) = 0 \Leftrightarrow \rho$ *pure state*
2. $dim(\mathcal{H}) = d \Rightarrow S(\rho) \leq \log_2(d)$ *and* $S(\rho) = \log(d) \Leftrightarrow \rho = \frac{\mathbb{1}}{d}$
3. ρ_{AB} *on* $\mathcal{H}_A \otimes \mathcal{H}_B$ *pure* $\Rightarrow S(\rho_A) = S(\rho_B)$.
4. *Let* $p_j \geq 0$, $\sum_j p_j = 1$, ρ_j *density operators which have support on orthogonal subspaces*[2]. *Then* $S(\sum_j p_j \rho_j) = H(\{p_j\}) + \sum_j p_j S(\rho_j)$, *where* $H(\{p_j\}) = -\sum_j p_j \log_2 p_j$ *is the* ***Shannon entropy****.*
5. ***Joint Entropy Theorem:*** *Let* $p_j \geq 0$, $\sum_j p_j = 1$, $\{|j_A\rangle\} \subset \mathcal{H}_A$ *orthonormal,* ρ_j *density operators on* $\mathcal{H}_B$. *Then* $S(\sum_j p_j |j_A\rangle\langle j_A| \otimes \rho_j) = H(\{p_j\}) + \sum_j p_j S(\rho_j)$
6. ρ_A, ρ_B *density operators on* $\mathcal{H}_A$, $\mathcal{H}_B$. *Then* $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$.

Proof.

1. As all eigenvalues $p_j \in [0, 1]$, we find $S(\rho) = -\sum_j p_j \log_2 p_j \geq 0$ because $-p_j \log_2 p_j \geq 0$. Furthermore $-p_j \log_2 p_j = 0 \Leftrightarrow p_j \in \{0, 1\}$. As the p_j form a probability distribution, we find $S(\rho) = 0 \Leftrightarrow \exists j : p_j = 1 \Leftrightarrow$ state pure

[1] For a hermitian operator A with eigenvalue-decomposition $A = \sum_j a_j |a_j\rangle\langle a_j|$ (where $\{|a_j\rangle\}$ is an orthonormal basis) and a function $f : \mathbb{C} \to \mathbb{C}$, we define $f(A) := \sum_j f(a_j) |a_j\rangle\langle a_j|$. Especially this implies that $\rho \log_2 \rho$ has the eigenvalues $p_j \log_2 p_j$. Furthermore we define $0 \cdot \log_2 0 := 0$.

[2] For a hermitian operator $A = \sum_j a_j |a_j\rangle\langle a_j|$ (where $\{|a_j\rangle\}$ is an orthonormal basis), the support is defined as $\text{span}\{|a_j\rangle \, | a_j \neq 0\}$. If A and B have orthogonal support, then $B|a_j\rangle = 0$ for $a_j \neq 0$.

2. We prove this claim by induction. At first assume $d = 1$. Then we have $p_1 = 1$ and thus $S(\rho) = 0 = \log_2 1$. Now assume we already know: $-\sum_{j=1}^{d-1} p_j \log_2 p_j \leq \log_2(d-1)$ for all $\{p_j | j \in \{1, 2, ..., d-1\}\}$ with $p_j \geq 0$, $\sum_{j=1}^{d-1} p_j = 1$. We consider now $\{p_j | j \in \{1, 2, ..., d\}\}$ with $p_j \geq 0$, $\sum_{j=1}^{d} p_j = 1$. We impose this last constraint by identifying p_d as function of the other eigenvalues: $p_d = p_d(p_1, ..., p_{d-1}) = 1 - \sum_{j=1}^{d-1} p_j$. This function is defined for $\vec{p} := (p_1, ...p_{d-1}) \in D := [0,1]^{d-1} \cap \{\vec{q} \,|\, \sum_{j=1}^{d-1} q_j \leq 1\}$ (see figure 2.1). Here, we demand $\sum_{j=1}^{d-1} p_j \leq 1$ because we want $p_d \in [0,1]$. As D is compact, there must exist a global maximum. To find the border of this set, we notice that $[0,1]^{d-1}$ is a hypercube, while $\{\vec{q} \,|\, \sum_{j=1}^{d-1} q_j \leq 1\}$ constrains us to one half of that cube, because $\{\vec{q} \,|\, \sum_{j=1}^{d-1} q_j = 1\}$ is a hyperplane (compare Appendix B) which splits $\mathbb{R}^{d-1}$ and $[0,1]^{d-1}$ in two parts. Then the border is given by the sides of the hypercube (with $\exists j : p_j = 0$) and the part of the hyperplane, that lies in the cube (here $p_d = 0$). Thus on ∂D, we always have at least one $p_j = 0$, wlog we assume $p_d = 0$. But we already know $-\sum_{j=1}^{d-1} p_j \log_2 p_j \leq \log_2(d-1)$. However, $S(\frac{\mathbb{1}_d}{d}) = \log_2 d$, so we have to look for our maximum in the interior $D^0 = D \setminus \partial D$. Here, every global maximum is also a local maximum. Such maxima must satisfy $\nabla_{\vec{p}}(-\sum_{j=1}^{d} p_j \log_2 p_j) = 0$. For $k \in \{1, 2, ...d-1\}$ we find $\frac{\partial}{\partial p_k}(-\sum_{j=1}^{d} p_j \log_2(e) \ln(p_j)) = 0 \Leftrightarrow 0 = -\ln p_k - 1 + \ln p_d + 1 \Leftrightarrow p_k = p_d$. Thus our maximum is indeed found for $\rho = \frac{\mathbb{1}_d}{d}$ and $S(\rho) = \log_2 d$.

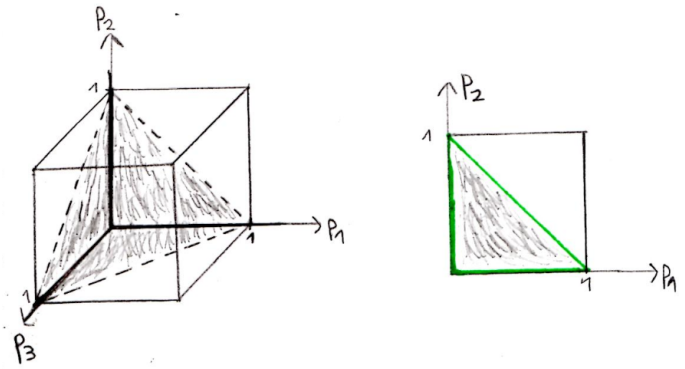

Figure 2.1: The domain of $\vec{p} = (p_1, ...p_{d-1})$ for $d = 4$ dimensions (left) and $d = 3$ dimensions (right).

3. We know that ρ_A and ρ_B have the same non-zero eigenvalues. $S(\rho_A)$ and $S(\rho_B)$ depend only on the eigenvalues.

4. We perform eigenvalue decompositions: $\rho_i = \sum_j p_j^{(i)} |p_j^{(i)}\rangle \langle p_j^{(i)}|$. Because of the orthogonal supports, the $|p_j^{(i)}\rangle$ with $p_j^{(i)} \neq 0$ are also eigenstates of the other ρ_k with $\rho_k |p_j^{(i)}\rangle = 0$. This implies that $\sum_j p_j \rho_j$ also has the eigenstates $|p_j^{(i)}\rangle$, with eigenvalues $p_i p_j^{(i)}$. Thus we find $S(\sum_j p_j \rho_j) = -\sum_{i,j} p_i p_j^{(i)} \log_2(p_i p_j^{(i)}) = -\sum_i p_i \log_2 p_i + \sum_i p_i \big(-\sum_j p_j^{(i)} \log_2 p_j^{(i)}\big) = H(\{p_j\}) + \sum_i p_i S(\rho_i)$.

5. The $|j_A\rangle \langle j_A| \otimes \rho_j$ obviously have orthogonal supports, and $S(|j_A\rangle \langle j_A| \otimes \rho_j) = S(\rho_j)$ because $|j_A\rangle \langle j_A| \otimes \rho_j$ and ρ_j have the same non-zero eigenvalues.

6. Use the joint entropy theorem with $\rho_j := \rho_B \ \forall j$ and $\sum_j p_j \left|j_A\right\rangle \left\langle j_A\right| = \rho_A$. Then $H(\{p_j\}) = S(\rho_A)$ and $\sum_j p_j S(\rho_B) = S(\rho_B)$.

□

Theorem 6 (Von Neumann entropy as measure for pure state entanglement)**.** *Let $\rho = \left|\psi_{AB}\right\rangle \left\langle\psi_{AB}\right|$ be a pure state on $\mathcal{H}_A \otimes \mathcal{H}_B$, $dim(\mathcal{H}_A) \leq dim(\mathcal{H}_B)$. Then:*

1. *$S(\rho_A) = 0 \Leftrightarrow \left|\psi_{AB}\right\rangle$ separable, $\quad S(\rho_A) > 0 \Leftrightarrow \left|\psi_{AB}\right\rangle$ entangled*
2. *$S(\rho_A) = \log(dim\mathcal{H}_A)\big($ = maximal$\big) \ \Leftrightarrow \ \left|\psi_{AB}\right\rangle$ maximally entangled*

Proof.

1. $S(\rho_A) = 0 \Leftrightarrow \rho_A$ pure $\Leftrightarrow$ Schmidt number $= 1 \Leftrightarrow$ separable
2. $S(\rho_A)$ maximal $\Leftrightarrow \rho_A = \frac{\mathbb{1}}{\dim\mathcal{H}_A} \Leftrightarrow$maximally entangled

□

The last theorem shows that the von Neumann entropy is a good measure for entanglement of pure states. For mixed states, it is much harder to find good entanglement measures, because one has additional classical correlations that need to be distinguished from quantum entanglement. There exist many measures, but none of them is perfect. As a simple example to show why the von Neumann entropy is a bad measure for entanglement of mixed states we consider $\rho_{AB} = \rho_A \otimes \rho_B$. This mixed stated is non-entangled. However, we are completely free to choose ρ_A. We could take $\rho_A = \left|\psi_A\right\rangle \left\langle\psi_A\right|$ with $S(\rho_A) = 0$ or $\rho_A = \frac{1}{\dim(\mathcal{H}_A)}\mathbb{1}$ with $S(\rho_A) = \log_2(\dim(\mathcal{H}_A))$. Thus even for non-entangled states, $S(\rho_A)$ can take values between its minimum and maximum and is thus a bad measure.
An idea to detect entanglement of mixed states is given by entanglement witnesses. Before we state the Entanglement Witness Theorem, at first we need to prove a few statements:

Theorem 7. 1. *The density operators form a convex set*[3]*:*

$$\mathcal{D} := \{\rho \mid \ \left\langle\psi\right| \rho \left|\psi\right\rangle \geq 0 \ \forall \left|\psi\right\rangle, \rho^\dagger = \rho, Tr(\rho) = 1\} \tag{2.6}$$

2. *The separable density operators form a convex set:*

$$\mathcal{S} := \{\rho \in \mathcal{D} \mid \rho \text{ separable } \} \tag{2.7}$$

Proof. 1. Let $A, B \in \mathcal{D}, p \in [0,1]$.
Then $\mathrm{Tr}(pA + (1-p)B) = p\mathrm{Tr}(A) + (1-p)\mathrm{Tr}(B) = 1$.
Also $(pA + (1-p)B)^\dagger = (pA + (1-p)B)$
and $\left\langle\psi\right| [pA + (1-p)B] \left|\psi\right\rangle = p \left\langle\psi\right| A \left|\psi\right\rangle + (1-p) \left\langle\psi\right| B \left|\psi\right\rangle \geq 0$

[3]A set $\mathcal{A}$ is called convex, if for all $A, B \in \mathcal{A}$, $p \in [0,1]$ also $p \cdot A + (1-p) \cdot B \in \mathcal{A}$

2. $\rho, \sigma \in \mathcal{D}$ separable $\Rightarrow \exists$ density operators $\rho_A^{(j)}, \rho_B^{(j)}, \sigma_A^{(j)}, \sigma_B^{(j)}$ and $p_j, q_j \geq 0$ with $\sum_j p_j = \sum_j q_j = 1$ such that: $\rho = \sum_{j=1}^{m} p_j \rho_A^{(j)} \otimes \rho_B^{(j)}$ and $\sigma = \sum_{j=1}^{n} q_j \sigma_A^{(j)} \otimes \sigma_B^{(j)}$. For $j \in \{1, ..., m\}$ and $a \in [0, 1]$ we define $r_j := ap_j$ and $\tau_A^{(j)} := \rho_A^{(j)}$, $\tau_B^{(j)} := \rho_B^{(j)}$. For $j \in \{m+1, ..., m+n\}$ we define $r_j := (1-a)q_{j-m}$ and $\tau_A^{(j)} := \sigma_A^{(j-m)}$, $\tau_B^{(j)} := \sigma_B^{(j-m)}$. Then we find $r_j \in [0, 1]$, $\sum_{j=1}^{n+m} r_j = 1$ and $a\rho + (1-a)\sigma = \sum_{j=1}^{n+m} r_j \tau_A^{(j)} \otimes \tau_B^{(j)} \in \mathcal{S}$

□

Theorem 8 (Entanglement witness theorem (EWT))**.**
Let ρ be an arbitrary density operator on a finite-dimensional Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. We define $\mathcal{M} := \{A \mid A$ linear operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ with $A^\dagger = A\}$. Then:
ρ entangled $\Leftrightarrow \exists W \in \mathcal{M}$: $Tr(\rho W) < 0$ and $Tr(\sigma W) \geq 0$ for all separable density operators σ.

Proof(sketch). We assume that $\rho \in \mathcal{D}$ is entangled. $\mathcal{M}$ is a real vector space with scalar product $\langle A, B\rangle := \mathrm{Tr}(A^\dagger B) = \mathrm{Tr}(AB)$ (see Appendix C for more information). As we have learned from theorem 7, we have the situation shown in figure 2.2 (left image). It is intuitive, that there exists a hyperplane which separates ρ and $\mathcal{S}$.[4]

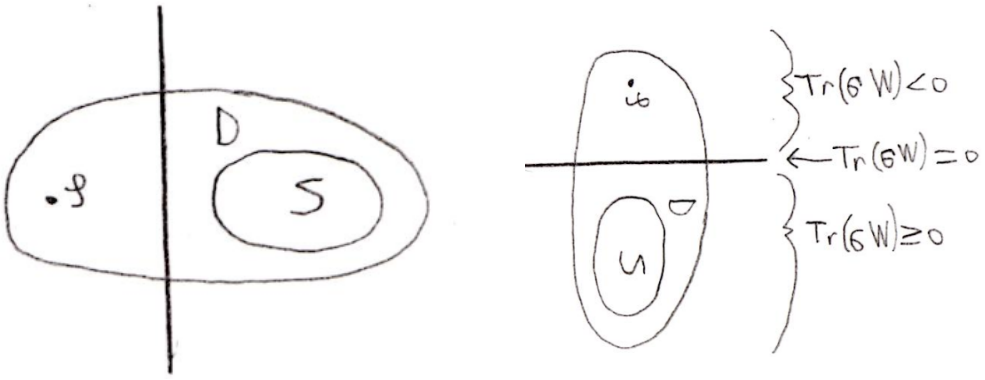

Figure 2.2: The set of density operators $\mathcal{D}$ is convex and contains the convex set of separable density operators $\mathcal{S}$. Also shown is the hyperplane that separates $\mathcal{S}$ and the entangled density operator ρ.

Let this hyperplane be described by $\mathrm{Tr}(\hat{W} \cdot \tau) = \langle \hat{W}, \tau \rangle = c$ (see also Appendix B), where $\hat{W} = \hat{W}^\dagger$ is the normal vector of the hyperplane, chosen such that $\mathrm{Tr}(\hat{W}\rho) = \langle \hat{W}, \rho \rangle < c$ and $\mathrm{Tr}(\hat{W}\sigma) = \langle \hat{W}, \sigma \rangle \geq c \quad \forall \sigma \in \mathcal{S}$. Now instead of $\hat{W}$, we take $W := \hat{W} - c\mathbb{1}$. Then we find ($\tau \in \mathcal{D}$):

$$\langle W, \tau \rangle < 0 \Leftrightarrow \mathrm{Tr}(W\tau) < 0 \Leftrightarrow \mathrm{Tr}(\hat{W}\tau) - c\mathrm{Tr}(\tau) < 0 \Leftrightarrow \mathrm{Tr}(\hat{W}\tau) < c \Leftrightarrow \langle \hat{W}, \tau \rangle < c$$

We also find $\langle \hat{W}, \tau \rangle = c \Leftrightarrow \langle W, \tau \rangle = 0$ and $\langle \hat{W}, \tau \rangle > c \Leftrightarrow \langle W, \tau \rangle > 0$. Thus we are now in the situation shown in figure 2.2(right image) and W fulfills the EWT. □

[4]The existence of this hyperplane follows from the Hahn-Banach theorem of functional analysis.

Comment. *As shown in Appendix C, the density operators can all be found in a hyperplane* $\mathcal{T} := \{A \in \mathcal{M} \mid Tr(A) = 1\} \subset \mathcal{M}$. *This means that* $\mathcal{T}$ *and the other hyperplane orthogonal to* W *intersect in an even lower dimensional line. This gives us the freedom to chose another* W *to effectively tilt the hyperplane such that the origin is also found in that plane. This is what we have done by replacing* $\hat{W}$ *with* W. *The whole situation is visualized in figure 2.3.*

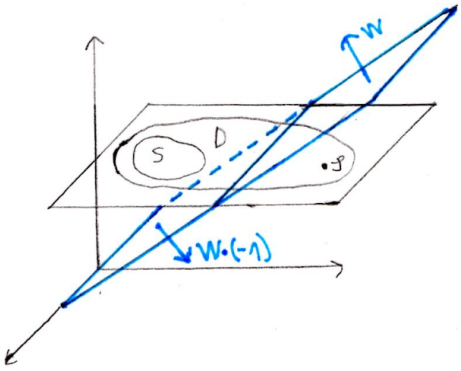

Figure 2.3: The density operators can be found in an affine hyperplane $\mathcal{T}$. The plane that we obtain from the Hahn-Banach theorem can be tilted such that it contains the origin.

3 Conclusion

In this document we have introduced entanglement for pure and mixed states. We have discussed the von Neumann entropy as measure for entanglement of pure states. However, many aspects of entanglement have not been adressed in this introduction, e.g. what entanglement can be used for and what its application for information science is. We also have not discussed the problems with the interpretation of entanglement, e.g. why it does not allow for faster-than-light communication. Also a delicate topic is how to define entanglement measures for mixed states. Here one typically defines some properties that such an measure should satisfy and defines measures constructed such that the constraints are fulfilled. There exist a lot of such measures; but many of them are hard to calculate.

A Appendix: Extended Proof Of Theorem 4

We start with a separable (mixed) state: $\rho = \sum_{j=1}^{a} p_j \rho_A^{(j)} \otimes \rho_B^{(j)}$. We perform a spectral expansion of all density operators and obtain:

$$\rho = \sum_{j=1}^{a} p_j (\sum_{k=1}^{b} q_k \left|q_k\right\rangle \left\langle q_k\right|) \otimes (\sum_{l=1}^{c} r_l \left|r_l\right\rangle \left\langle r_l\right|) = \sum_{j,k,l=1}^{a,b,c} (p_j q_k r_l) \left|q_k\right\rangle \left\langle q_k\right| \otimes \left|r_l\right\rangle \left\langle r_l\right|$$

Now we define a bijection $f : \mathbb{N}^3 \to \mathbb{N}$ and $\mathcal{B} := f(\{1, 2, ...a\} \times \{1, 2, ...b\} \times \{1, 2, ...c\})$ (see e.g. figure A.1). Then $f : \{1, 2, ...a\} \times \{1, 2, ...b\} \times \{1, 2, ...c\} \to \mathcal{B}$ is also a bijection. Now we define $t_{f(j,k,l)} := p_j q_k r_l$, $|\alpha_{f(j,k,l)}\rangle := |q_k\rangle$ and $|\beta_{f(j,k,l)}\rangle := |r_l\rangle$. Thus:

$$\rho = \sum_{j,k,l=1}^{a,b,c} t_{f(j,k,l)} \, |\alpha_{f(j,k,l)}\rangle \langle\alpha_{f(j,k,l)}| \otimes |\beta_{f(j,k,l)}\rangle \langle\beta_{f(j,k,l)}| = \sum_{f\in\mathcal{B}} t_f \, |\alpha_f\rangle \langle\alpha_f| \otimes |\beta_f\rangle \langle\beta_f|$$

$t_{f(j,k,l)} = p_j q_k r_l \geq 0$ and $\sum_{f\in\mathcal{B}} t_f = \sum_{j,k,l=1}^{a,b,c} t_{f(j,k,l)} = (\sum_j p_j)(\sum_k q_k)(\sum_l r_l) = 1$.

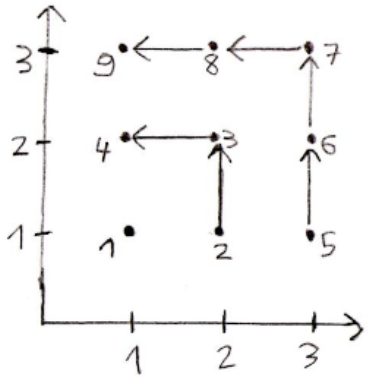

Figure A.1: The figure illustrates how to define a bijection $g : \mathbb{N}^2 \to \mathbb{N}$. With this function a bijection $f : \mathbb{N}^3 \to \mathbb{N}$, $f(j, k, l) = g(j, g(k, l))$ can be defined.

B Appendix: Euclidean space

Here we want to recall some simple facts about planes in Euclidean space (or isomorphic spaces).

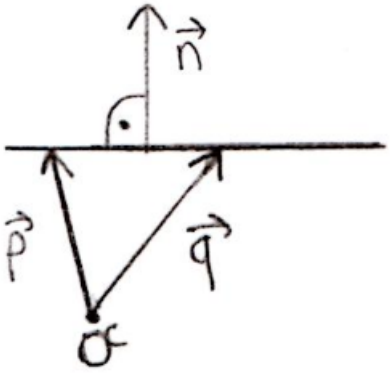

Figure B.1: A hyperplane in 2 dimensions. The difference of two vectors pointing on the hyperplane is orthogonal to the normal vector.

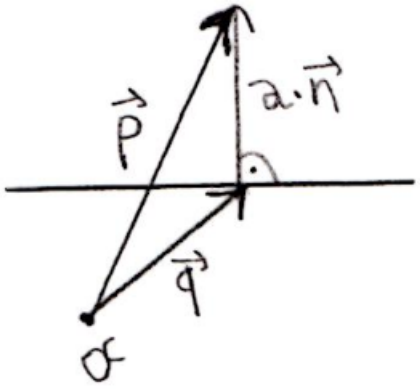

Figure B.2: The vector $\vec{p}$ points above the plane.

Let $\vec{q}$ be an arbitrary vector pointing on the hyperplane with normal vector $\vec{n}$. Then for any vector $\vec{p}$ also pointing on the plane, we find $(\vec{p} - \vec{q}) \cdot \vec{n} = 0$ (see figure B.1). This last equation defines the plane. We note that the exact choice of $\vec{q}$ does not matter, because the different choices only differ in a vector orthogonal to $\vec{n}$. Defining $c := \vec{q} \cdot \vec{n}$, we find the hyperplane to be defined by $\vec{p} \cdot \vec{n} = c$.

Now we consider a vector $\vec{p}$ that does not point on the hyperplane (see figure B.2). We decompose $\vec{p} = \vec{q} + \alpha\vec{n}$, where $\alpha \in \mathbb{R}$ and $\vec{q}$ points on the plane. Then we define $\vec{p}$ to point above the plane, if $\alpha > 0$. This implies $\vec{p} \cdot \vec{n} = \alpha\vec{n}^2 + \vec{q} \cdot \vec{n} = c + \alpha > c$. Thus we say that $\vec{p}$ points above the hyperplane if $\vec{p} \cdot \vec{n} > c$, and equivalently we say that $\vec{p}$ points below the plane if $\vec{p} \cdot \vec{n} < c$. Of course the assignment of *above* and *below* is arbitrary and can be exchanged.

C Hermitian operators form a real vector space

Theorem 9. *The set* $\mathcal{M} := \{A \mid A \text{ linear operator on } \mathcal{H}_A \otimes \mathcal{H}_B \text{ with } A^\dagger = A\}$*, where* $\dim(\mathcal{H}_A \otimes \mathcal{H}_B) < \infty$*, forms a real vector space with a scalar product* $\langle A, B\rangle := Tr(A^\dagger B) = Tr(AB)$.

Proof. Let $A, B \in \mathcal{M}, a, b \in \mathbb{R}$. Addition of matrices of course is associative and $A + B = B + A$, the neutral element 0 is hermitian, and the inverse $-A$ is hermitian if A is. Furthermore $(A+B)^\dagger = A^\dagger + B^\dagger = A + B$, thus $(\mathcal{M}, +)$ is an abelian group. Of course we also have $a \cdot (A+B) = a \cdot A + a \cdot B$, $(a+b) \cdot A = a \cdot A + b \cdot A$, $a \cdot (bA) = (ab) \cdot A$ and $1 \cdot A = A$; $(a \cdot A)^\dagger = a^* A^\dagger = a \cdot A$. Thus $\mathcal{M}$ is a real vector space [5].
$\langle A, B\rangle$ is linear in both arguments because matrix multiplication and Tr() are linear. Because of $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, $\langle A, B\rangle$ is also symmetric. Furthermore $\mathrm{Tr}(AB)^* = \mathrm{Tr}((AB)^\dagger) = \mathrm{Tr}(AB)$, thus $\langle \cdot, \cdot \rangle : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ is well-defined. Furthermore, for $A = \sum_j a_j |a_j\rangle \langle a_j|$ (eigenvalue decomposition) we find $\langle A, A\rangle = \sum_j |a_j|^2$, which means $\langle A, A\rangle \geq 0$ and $= 0$ exactly for $A = 0$. □

Theorem 10. *The set* $\mathcal{T} := \{A \in \mathcal{M} \mid Tr(A) = 1\}$ *is a hyperplane in* $\mathcal{M}$ *given by* $\mathcal{T} = \mathbb{1} + \hat{\mathcal{T}}$*, where* $\hat{\mathcal{T}} := \{A \in \mathcal{M} \mid Tr(A) = 0\}$.

Proof. At first we show $\mathcal{T} = \mathbb{1} + \hat{\mathcal{T}}$:
Every $A \in \mathcal{T}$ we can decompose into $A = M + D$ where $M \in \hat{\mathcal{T}}$ contains all off-diagonal elements of A, and D contains all diagonal elements of A. As also $\mathrm{Tr}(D) = 1$, we find $\mathrm{Tr}(D - \mathbb{1}) = 0$ and thus $A = \mathbb{1} + \big((D - \mathbb{1}) + M\big) \in \mathbb{1} + \hat{\mathcal{T}}$.
Now we argue that $\hat{\mathcal{T}} \subset \mathcal{M}$ is a sub-vectorspace:
$0 \in \hat{\mathcal{T}}$, $A, B \in \hat{\mathcal{T}} \Rightarrow A + B \in \hat{\mathcal{T}}$ because $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B) = 0$ and $A + B$ hermitian too. Furthermore for $A \in \hat{\mathcal{T}}, a \in \mathbb{R}$: $a \cdot A$ also hermitian and $\mathrm{Tr}(a \cdot A) = a \cdot \mathrm{Tr}(A) = 0$. Note that $\hat{\mathcal{T}}$ has one dimension less than $\mathcal{M}$, because the additional condition $\mathrm{Tr}(A) = 0$ puts exactly one constraint on the elements of A. [6] □

[5] For some $a \in \mathbb{C}$, we find $(a \cdot A)^\dagger = a^* A^\dagger \neq a \cdot A$. Thus $\mathcal{M}$ is not a complex vector space.

[6] A basis for $\hat{\mathcal{T}}$ could be given by $E^{(jk)} := (\delta_{ja}\delta_{kb} + \delta_{ka}\delta_{jb})_{ab}$ for $j \neq k$, $\hat{E}^{(jk)} := (i\delta_{ja}\delta_{kb} - i\delta_{jb}\delta_{ka})_{ab}$ for $j \neq k$ and $D^{(j)} := (\delta_{ja}\delta_{jb} - \delta_{(j+1),a}\delta_{(j+1),b})_{ab}$ for $j \in \{1, 2, ..., \dim -1\}$. Here, $E^{(jk)}$ and $\hat{E}^{(jk)}$ are used to obtain the off-diagonal elements. $D^{(j)}$ are used to obtain the diagonal elements: First we use $D^{(1)}$ to choose the first diagonal element; then we use $D^{(2)}$ to choose the second diagonal element,... The last diagonal element is uniquely determined by the condition $\mathrm{Tr}(\cdot) = 0$, which is fulfilled because all basis elements have vanishing trace. To get a basis for $\mathcal{M}$, we can add $(\delta_{a,\dim}\delta_{b,\dim})_{ab}$ such that we can also choose the last diagonal element freely. Thus $\hat{\mathcal{T}}$ is indeed a hyperplane in $\mathcal{M}$.

References

[1] *Quantum Computation and Quantum Information*, M. Nielsen, I. Chuang, Cambridge University Press, tenth printing

[2] lecture notes *Physics 219* by John Preskill

[3] D. Bruß, G. Leuchs *Lectures on Quantum Information*, Wiley-VCH(2007)

[4] V. Vedral, *Introduction to Quantum Information Science*, Oxford University Press

[5] M. Piani, *Lecture Notes on Entanglement Theory*, IQC Waterloo, 2012

[6] bachelor thesis by Lukas Schneiderbauer, *Entanglement or Separability-an introduction*

[7] P. Krammer *Characterizing entanglement with geometric entanglement witnesses*, arXiv: 0807.4830v2[quant-ph]

[8] Diploma thesis by Philipp Krammer, *Quantum Entanglement: Detection, Classification, and Quantification*