Entanglement Invariant Theory

Jonas Green Enander

Bachelor's Thesis at the Department of Physics at Stockholm University

September 2008

Abstract

In the last decades Quantum Information Theory has become an intense field of research. In this work we investigate the method of operating on quantum states, namely Local Operations and Classical Communication. This method gives rise to a partitioning on the set of states, where two states are said to be equivalent if they can be transformed into one another by unit probability. We show that this holds if the transformation is local unitary. A similar partitioning is based on the demand that the transformation can be carried out with non-zero probability. For this scenario we give a new proof that this corresponds to general invertible transformations.

We then investigate the equivalence problem, namely to find conditions when two states can be transformed into one another. By reviewing the classical Invariant Theory we see how the concepts of invariants of homogenous polynomials can be used to solve, or at least reiterate, the equivalence problem in a classical framework.

Using these concepts we give a description of the case of bipartite and tripartite qubit states.

Acknowledgment

I would like to thank my supervisor Professor Ingemar Bengtsson for the helpful discussions, comments and proofreading done during the course of this work.

Contents

1 Introduction

Quantum Information Theory has undergone a rapid development during the last two decades thanks to the insight that the phenomena of entanglement can be used to produce new protocols capable of tasks beyond the reach of classical methods. Indeed, already Schrödinger realized that entanglement was the defining feature of quantum mechanics, but it is only in recent years that the implications for information processing have been investigated. Entanglement produces statistical correlations between subsystems that can not be reproduced by classical schemes. This property could therefore be used to device new communication protocols and information processing algorithms.

As a part of the development of Quantum Information Theory comes the classification of information carrying units, *qubits*. Since qubits can be entangled between different systems in many different ways there is a need to fully understand the types and structures that the entanglement gives rise to. One would also like to know when a given quantum state, or a set of qubits having some overall entanglement, can be transformed to some other states with unit or non-zero probability. The classification one needs thus involves the entanglement structure as well as understanding how physical properties of a state change under different transformations.

Interestingly enough, the techniques for dealing with properties that are invariant under different transformations was one of the important themes of late 19:th century mathematics. The so-called Invariant Theory was developed in order to understand how functions in the coefficients of homogenous polynomials changed under general linear transformations. Specifically one wanted to find a set of polynomial functions in the coefficients (and possibly variables) that remain unchanged under a transformation and therefore can be used to decide whether two homogenous polynomials can be transformed into one another. This idea carries over immediately to the case of classifying quantum states. The goal is to find a finite set of functions of the coefficients of a quantum state such that two states can be transformed into one another if and only if they are equal on the functions.

In the following paper we shall investigate how this works for the case of bipartite and tripartite qubit systems. But we also need to know how specific transformations correspond to different types of protocols applicable in the laboratory. More specifically, we are interested in Local Operations and Classical Communication (LOCC), which means that parties perform local actions on their shared state and communicate the results classically. Our question is thus when two states can be transformed into one another with unit probability and what types of transformation this corresponds to. But we are also interested when there is some non-zero probability of successful transformation since this may also be applied in our protocols. This scenario is called stochastic LOCC, and we will see that this also corresponds to specific types of transformations.

This paper is organized as follows. In Chapter 2 we look at the basics, namely the quantum states and how we operate on them. In Chapter 3 we see how these operations are used for specific protocols and how to decide when two states are equivalent (that is, can be transformed into one another) in these protocols. The notion of equivalence of two quantum states is put in a mathematical framework in Chapter 4 where we also look at the problems encountered when trying to classify different states. In Chapter 5 we shift gear and look at the classical Invariant Theory and how it might help us with these problems. In Chapter 6 two specific examples are studied in depth, namely the bipartite and tripartite quantum states. Chapter 7 ends the paper with some conclusions and a look ahead.

2 General framework: States and Operations

When we go into the laboratory we basically want to produce quantum states and operate on them. We therefore need to have a clear picture of how these experimental procedures look in a mathematical setting. The states that we are working with are represented as normalized vectors in some Hilbert space

$$\mathcal{H} = V_1 \otimes V_2 \otimes \cdots \otimes V_n,$$

where V_i is the i:th factor of the Hilbert space. For qubits we have that dim $V_i = 2$. Generally we will have a mixture of different states that we operate on and such a mixture is described by a *density matrix.*

We will now describe some mathematical properties of the density operator and see how we manipulate it in the laboratory.

2.1 States and density operators

We begin with a formal definition:

Definition *A non-negative, hermitian operator with unit trace is called a density operator.*

The density operator ρ contains all the statistical information that can be available to us when we do measurements on a system. It can be decomposed in a non-unique way as a weighted sum of projection operators:

$$\rho = \sum_i w_i \left| f_i \right\rangle \left\langle f_i \right|$$

with the normalization condition $\sum w_i = 1$. A density operator is called *pure* if the decomposition has rank one, i.e. $\rho = \left| f \right\rangle \left\langle f \right|$, for some unit vector $\left| f \right\rangle \in \mathcal{H}$.

Suppose now that we have a composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ with a density operator ρ_{12} defined on this system. We say that ρ_{12} is uncorrelated if it can be decomposed as

$$\rho_{12} = \sum_i w_i \rho_1^i \otimes \rho_2^i, \tag{1}$$

where ρ_1^i and ρ_2^i can be taken to be pure without loss of generality. If ρ_{12} has further correlations, that is does not admit a decomposition of the form (1), we say that the two subsystems $\mathcal{H}_1$ and $\mathcal{H}_2$ are entangled. Entanglement will produce statistical correlation between the two subsystems that are not reproducible in a classical framework.

Given a density operator ρ_{12} on $\mathcal{H}_1 \otimes \mathcal{H}_2$ we might be interested in the statistical distribution of one of the subsystems without caring about the other. We would thus ask if there exists a density operator ρ_1 defined on $\mathcal{H}_1$ that encompasses all the statistical information about $\mathcal{H}_1$ that ρ_{12} does. We would also like to known if this operator is unique. The answer is given in the affirmative in both cases and we are lead to the following definition:

Definition. *If ρ_{12} is a density operator defined on $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $\{f_j\}$ is an orthonormal basis for $\mathcal{H}_2$ then $\rho_1 \equiv \mathrm{tr}_2\, \rho_{12}$ is given by*

$$(x,\, \rho_1 y) = \sum_j \left(x \otimes f_j,\, \rho_{12}\left(x \otimes f_j\right)\right),$$

for all $x, y \in \mathcal{H}_1$. The partial trace $\rho_2 \equiv \mathrm{tr}_1\, \rho_{12}$ is defined analogously.

As an equivalent description we could take an arbitrary operator T_1 defined on $\mathcal{H}_1$ and demand that

$$\mathrm{tr}\left(\left(T_1 \otimes 1_2\right) \rho_{12}\right) = \mathrm{tr}\left(T_1 \rho_1\right)$$

This would mean that ρ_1 reproduces the same statistical distribution in the sense that we could either apply T_1 on the whole system but leaving $\mathcal{H}_2$ unaffected or apply T_1 directly on system $\mathcal{H}_1$. Note that if the density operator of the composite system is just the uncorrelated tensor product $\rho_{12} = \sigma \otimes \phi$, then performing the partial trace gives us back the density operators of the individual subsystems:

$$\sigma = \mathrm{tr}_2\, \rho_{12}$$

$$\phi = \mathrm{tr}_1\, \rho_{12}.$$

The operator $\rho_1 \equiv \mathrm{tr}_2\, \rho_{12}$ is sometimes called a *reduced density operator.*

2.2 Effects and operations

Given a state ρ we would like to know the evolution of this system when it comes in contact with an environment (or a measuring apparatus). In the following we develop the theory put forward by Kraus in [Kr1, Kr2, Kr3] . We follow the exposition in [Kr1] closely but omit most of the mathematical details. We recommend [Ca] and [Sch] for nice introductions.

We start off with an operational definition of an *effect.* Given a system in some state we interact with the system with an apparatus and read off a macroscopic change in the apparatus. This macroscopic change is called an effect. Notice that while the density operator ρ may contain information about microscopic properties, the effect is always related to a macroscopically observable change. Associated with an effect F is some apparatus, or more precisely the equivalence class of apparatuses producing the same effect.

Given a state and some effect F we might be interested in knowing the probability $p\,(F, \rho)$ that the effect will occur. We can take it as a postulate (although an axiomatic deduction is possible) that this is given by

$$p\,(F, \rho) = \mathrm{tr}\,(F\rho)\,.$$

Notice that nothing has been said about the formal mathematical properties of these effects. Originally they were taken to be projection operators, but we will see that in reality a larger set of operators has to be used to describe the effects.

We now introduce the notion of an operation. If we perform an effect F on an ensemble ρ of $N \gg 1$ microsystems and assume that the effect does not annihilate any of the microsystems we will get a new state $\tilde{\rho}$ which also consists of N microsystems. Keeping F fixed and varying ρ we thus get a mapping

$$\tilde{\mathcal{L}} : K(\mathcal{H}) \to K(\mathcal{H})$$

from the set of density operators $K(\mathcal{H})$ to itself. This mapping defines an *operation* and sometimes a operator that maps operators to operators is called a *superoperator*. We call such an operation *non-selective* since we have not made any selection on the microsystems based on the outcome of applying F. On the other hand we could perform a *selective* operation as follows. The apparatus F will be triggered in $N_+ = p(F, \rho) N$ cases. If we after measurement only select those N_+ microsystems for which the apparatus was triggered and disregard the remaining $N_- = N - N_+$ systems we get a new ensemble with only N_+ systems. We call this ensemble $\hat{\rho}$. By varying ρ while keeping F fixed we could define a mapping by $\rho \mapsto \hat{\rho}$. But with such a mapping we would loose the information concerning the transition probabilities and it would not be defined for $p(F, \rho) = 0$. Instead we define the following mapping:

$$\mathcal{L}(\rho) = \begin{cases} p(F,\rho)\,\hat{\rho} & \text{if } \operatorname{tr}(F\rho) \neq 0 \\ 0 & \text{if } \operatorname{tr}(F\rho) = 0 \end{cases} \tag{2}$$

Since $\operatorname{tr}(\hat{\rho}) = 1$ and the trace satisfies $\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B)$ we have that

$$p(F, \rho) = \operatorname{tr}(\mathcal{L}(\rho))$$

so the final state is thus

$$\hat{\rho} = \frac{\mathcal{L}(\rho)}{\operatorname{tr}(\mathcal{L}(\rho))}.$$

Before proceeding, let us look at an example with a projection operator P. A general postulate states that after application of a projection operator on a state ρ we get a new state

$$\hat{\rho} = \frac{P\rho P}{\operatorname{tr}(P\rho)}.$$

Comparing this to our definition (2) we get that

$$\mathcal{L}(\rho) = P\rho P.$$

Notice that we have done a measurement disregarding those states that are orthogonal to the space spanned by the support of P. If we would like to do a non-selective operation we would have to include this space in our final state:

$$\tilde{\mathcal{L}}(\rho) = P\rho P + (I - P)\,\rho\,(I - P)$$

Here $I - P$ could be viewed as the "non-occurence of P". So in a non-selective operation our final state is in a mixture of the two states $P\rho P$ and $(I - P)\,\rho\,(I - P)$ with weights $\operatorname{tr}(P\rho)$ and $\operatorname{tr}((1 - P)\,\rho)$.

Having thus defined our selective and non-selective operations we would like to know their mathematical properties. Our operations are characterized by three important properties which are stated here without proof:

i) The operations are *convex-linear.* The set of all density operators $K(\mathcal{H})$ is convex and given λ with $0 \leq \lambda \leq 1$ and two operators ρ_1 and ρ_2 we have

$$\mathcal{L}\left(\lambda\rho_1 + (1-\lambda)\rho_2\right) = \lambda\mathcal{L}(\rho_1) + (1-\lambda)\mathcal{L}(\rho_2)$$

ii) A non-selective operation is *trace-preserving*, i.e. $\operatorname{tr}\left(\tilde{\mathcal{L}}(\rho)\right) = \operatorname{tr}(\rho)$. A selective operation is *trace-decreasing*, i.e. $\operatorname{tr}(\mathcal{L}(\rho)) \leq \operatorname{tr}(\rho)$. This can be seen directly from the definitions.

iii) $\mathcal{L}(\rho)$ is *completely positive.* Positivity is the requirement that positive operators map to positive operators, that is if $\rho \geq 0$ then $\mathcal{L}(\rho) \geq 0$. Complete positivity is a more subtle but necessary condition. If we extend our Hilbert space to $\mathcal{H}_R \otimes \mathcal{H}$, where $\mathcal{H}_R$ describes some reference system and ρ is the reduced density operator when we trace out the reference system, we demand that the mapping $1_R \otimes \mathcal{L}$ on the composite density operator is also positive. Complete positivity is thus a requirement that positivity is also respected if we add a system that does not participate in the dynamics. In classical probability theory this condition is trivial, but in the quantum context it puts a restriction on the possible operators.

For every selective operator $\mathcal{L}$ there is an associated effect F. We can use this to define a complementary operator to $\mathcal{L}$ labeled $\mathcal{L}'$. This is done by instead of selecting the microsystems that F triggered we select those that F did not trigger. We call this F'. F and F' are related by

$$F + F' = I$$

so

$$\begin{aligned}\operatorname{tr}(F\rho) + \operatorname{tr}(F'\rho) &= \operatorname{tr}\left((F+F')\rho\right) \\ &= \operatorname{tr}(I\rho) = 1.\end{aligned}$$

We can use this to understand the non-selective operator. Applying F we keep both the $N_+ = \operatorname{tr}(F\rho)N$ microsystems that are in a state

$$\hat{\rho} = \frac{\mathcal{L}(\rho)}{\operatorname{tr}(\mathcal{L}(\rho))}$$

and the remaining $N_- = \operatorname{tr}(F'\rho)N$ systems that are in the state

$$\hat{\rho}' = \frac{\mathcal{L}'(\rho)}{\operatorname{tr}(\mathcal{L}'(\rho))}.$$

So the final state $\tilde{\rho}$ is a mixture of $\hat{\rho}$ and $\hat{\rho}'$ in the form

$$\begin{aligned}\tilde{\rho} &= \operatorname{tr}(\mathcal{L}(\rho))\hat{\rho} + \operatorname{tr}(\mathcal{L}'(\rho))\hat{\rho}' \\ &= \mathcal{L}(\rho) + \mathcal{L}'(\rho).\end{aligned}$$

So the mapping $\tilde{\mathcal{L}} : K(\mathcal{H}) \to K(\mathcal{H})$ is simply given by

$$\tilde{\mathcal{L}}(\rho) = \mathcal{L}(\rho) + \mathcal{L}'(\rho). \tag{3}$$

The transition probability for non-selective operators is immediately seen to be unity:

$$\operatorname{tr}\left(\tilde{\mathcal{L}}(\rho)\right) = \operatorname{tr}(\mathcal{L}(\rho)) + \operatorname{tr}(\mathcal{L}'(\rho)) = 1.$$

From (3) we see that non-selective operators are a special case of selective operators.

2.3 First Representation Theorem

So far we have only described $\mathcal{L}$ and $\tilde{\mathcal{L}}$ operationally. We now shift gear and state the important First Representation Theorem which provides us with a mathematical framework for our superoperators. The full proof of the theorem can be found in [Kr1].

First Representation Theorem. *For an arbitrary operation $\mathcal{L}$ there exists operators A_k, $k \in K$ (a finite or countably infinite index set) on the state space $\mathcal{H}$ satisfying*

$$\sum_{k \in K_o} A_k^\dagger A_k \leq I$$

for all finite subsets $K_o \subseteq K$, such that for arbitrary $\rho \in K(\mathcal{H})$ the mapping $\mathcal{L}$ is given by

$$\mathcal{L}(\rho) = \sum_{k \in K} A_k \rho A_k^\dagger. \tag{4}$$

In particular, the effect F corresponding to $\mathcal{L}$ is given by

$$F = \sum_{k \in K} A_k^\dagger A_k.$$

In the theorem we have excluded some facts related to the adjoint of $\mathcal{L}$ since we do not need it in the following discussion. The theorem tells us that we can always perform a so-called operator sum decomposition of our superoperator. There also exists a second representation theorem which shows how a given operator can be represented through a reduced trace on a unitary transformation of the composite system of the state and the apparatus operating on the state.

Given the operator sum decomposition we can then state the measurement postulate which sums up our previous discussion (quoted from [BZ]):

Measurement postulate. *Let the space of possible measurement outcomes consist of n elements related to n measurement operators A_k satisfying the completeness relation*

$$\sum_{k=1}^{n} A_k^\dagger A_k = I. \tag{5}$$

The quantum measurement performed on the initial state ρ produces the k:th outcome with probability p_k and transforms ρ into ρ_k according to

$$\rho \mapsto \rho_k = \frac{A_k \rho A_k^\dagger}{\operatorname{tr}\left(A_k \rho A_k^\dagger\right)} \tag{6}$$

with $p_k = \operatorname{tr}\left(A_k \rho A_k^\dagger\right)$.

3 LOCC and stochastic LOCC

Having thus defined how we operate on quantum states we can now describe the protocols we make use of in the laboratory. As mentioned in the introduction we are interested in Local Operations and Classical Communications (LOCC). In order to understand what this implies, let us look at how the two generic characters Alice and Bob work with a common state.

3.1 Alice and Bob in the laboratory

Suppose that Alice and Bob share a common state $|\psi\rangle$. They only have access to their part of the common state, so $|\psi\rangle$ could for example be the entangled state of a decayed pion. So we have an entangled electron and positron going in opposite direction where Alice and Bob each have access to either the electron or positron. Since Alice and Bob can only operate on their part we say that they can only perform local operations. This means that if the entangled state lives in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, then Alice can only operate on $\mathcal{H}_A$ and Bob on $\mathcal{H}_B$. Furthermore, they can add a so-called ancilla to the system, which is an external reference system. Alice or Bob could let the electron be seen as a part of a larger system that does not participate in the dynamics, for example some other experimental equipment in the room next door. They could also communicate their results classically to each other. Alice could call Bob up and say: "Hey, I measured that my electron had a spin component along the direction of the z-axis." This would in turn make Bob carry out a special kind of measurement, perhaps a measurement of the spin component along the x-axis. This type of protocol, where the participants only perform local operations and possibly communicate their result classically is a LOCC protocol. We now turn to the formal properties of LOCC.

3.2 Formal definitions of LOCC

Before proceeding with the definitions we need to introduce the notion of multilocality in relation to our superoperators acting on the density operator. The condition of multilocality means that in the decomposition (4) the operators can be written $A_i = A_1^i \otimes A_2^i \otimes \cdots \otimes A_n^i$. This notion will be important since we consider the operations performed on the density operators as carried out by different parties in the laboratory who only have access to the subsystems.

Definition. *A LOCC is represented mathematically as a multilocally implementable superoperator, i.e. a completely positive linear map that does not increase the trace, which can be applied with the aid of classical coordination.*

With this definition in hand we can proceed to discuss the problem under what conditions a given state can be transformed into another state. Suppose that we want to implement some algorithm which requires a specific state ρ but we can only produce a state ρ'. The question is then whether there exists a protocol implementable under LOCC so that we can transform ρ into ρ'. We would like to know what types of transformations this corresponds to mathematically and also if we can decide whether this is possible just by knowing ρ and ρ'. The problem of deciding whether ρ and ρ' can be transformed into one another is known as the *equivalence problem.*

3.3 Transformations and equivalence

Because of the statistical nature of our quantum operations we have to distinguish between two procedures. When ρ is transformed into ρ' this can be done with unit probability or with non-zero, non-unit probability. We say that when a given state can be transformed into another with some non-zero probability it is reducible, and exact reducability means that the probability for this transformation is unity. The exact definitions for our two procedures was given for the first time in [BPRST]. It can be stated as:

Definition. *A state ρ' is exactly reducible to a state ρ under LOCC if and only if there exists a multilocally implementable trace preserving superoperator $\mathcal{L}$ such that $\rho' = \mathcal{L}(\rho)$.*

Definition. *A state ρ' is stochastically reducible to a state ρ under LOCC with yield p if and only if there exists a multilocally implementable superoperator $\mathcal{L}$ such that*

$$\rho' = \frac{\mathcal{L}(\rho)}{\operatorname{tr}(\mathcal{L}(\rho))}$$

and $p = \operatorname{tr}(\mathcal{L}(\rho))$.

We thus see that exact reducibility correspond to the special case of stochastic reducibility where $p = 1$.

Given these definition we also see that they establish equivalence relations on the set of density operators (that is, a symmetric, transitative and reflexive relation). Two states ρ and ρ' are said to be exactly (or LOCC) equivalent if they can be converted to one another under LOCC. Similarly, two states are said to be stochastically equivalent if they can be converted with some non-zero probability to one another under stochastic LOCC. These two relations yield two partitions of the set of density operators into a set of equivalence classes under either LOCC or stochastic LOCC.

Given such a partition, a plethora of questions naturally arises. Are the different entanglement measures constant on equivalence classes? What types of operators do LOCC and stochastic LOCC correspond to? How do we classify (both in principle and by computationally feasible means) the equivalence classes? In what sense are they relevant to the different protocols that are constructed for the production of concrete tasks?

Many of these questions are ongoing fields of intense research. Two of them will be discussed in the following, namely the types of operators corresponding to LOCC and stochastic LOCC and the classification of the equivalence classes.

If two operators ρ and ρ' are related under stochastic LOCC it means that instead of using ρ for our concrete task we could use ρ', albeit with a non-zero probability. Starting with a given state ρ we have access to a larger set of states ρ' under stochastic LOCC, but we can not be certain that our transformation protocol will always be successful.

We now move on to consider which transformations correspond to LOCC and stochastic LOCC.

3.3.1 Equivalence under LOCC

We have already seen what a multilocally implementable superoperator is. The question of locality also applies to unitary transformations, so it is natural to state that a unitary operator, that is an operator such that $U^{\dagger}U = 1$, is called a local unitary transformation (LUT) if it can be written as

$$U = U_1 \otimes U_2 \otimes \cdots \otimes U_n. \tag{7}$$

In [BPRST] the following important result was proved:

Theorem. *Two states are LOCC equivalent if and only if they are equivalent under local unitary transformations.*

That is if there exists a local unitary transformation U such that

$$\rho' = U\rho U^{\dagger} \tag{8}$$

then the two states ρ' and ρ are LOCC equivalent. And conversely, if they are LOCC equivalent then they are also related in the form of (8) for some local unitary operator. This theorem actually came out as a corollary from a theorem relating entropic properties to equivalence under LUT. But by observing the properties giving for the superoperator $\mathcal{L}$ and how it is related to LOCC the theorem follows almost directly from the definitions. For suppose that ρ' and ρ are LOCC equivalent. Then there exists a multilocally implementable superoperator such that $\rho' = \mathcal{L}(\rho)$. But by (6) this can be written as $\rho' = A\rho A^{\dagger}$, where A satisfies $AA^{\dagger} = I$ according to (5). So A is an unitary operator. Since we also demand multilocality under LOCC A has to be of the form (7). Suppose instead that there exists a unitary operator U on the form (7) such that (8) holds. But then this holds as our transformation protocol and the two operators are then obviously reducible under LOCC. The unitary operator guarantees the existence of an inverse operator and so the two states must be equivalent.

3.3.2 Equivalence under stochastic LOCC

For stochastic LOCC the case is not as clear. But in [DVC] the following theorem was shown:

Theorem. *Two states are stochastically equivalent if and only if there exists an invertible operator relating them.*

The proof relied chiefly on the use of the Schmidt decomposition. But once again the properties of the superoperator relating states under stochastic LOCC can be used to give an alternative, direct proof. Suppose first that there exists a multilocally implementable superoperator such that

$$\hat{\rho} = \frac{\mathcal{L}\left(\rho\right)}{\mathrm{tr}\left(\mathcal{L}\left(\rho\right)\right)}.$$

According to the mathematical properties described previously the mapping $\mathcal{L}$ can be written as

$$\sum_k A_k \rho A_k^\dagger = A\rho A^\dagger$$

for pure states. Since we have equivalence under stochastic LOCC the operators A_k are of the form

$$A_k = A_1^k \otimes A_2^k \otimes \cdots \otimes A_n^k$$

and so the operator A must also be of this form. Furthermore the operator A satisfies $0 \leq A^\dagger A \leq I$ (see [Kr2]). But

$$\mathrm{tr}\left(A\rho A^\dagger\right) > 0$$

so $A^\dagger A > 0$. We also have the reverse implementation, namely

$$\rho = \frac{\mathcal{L}^{'}\left(\hat{\rho}\right)}{\mathrm{tr}\left(\mathcal{L}^{'}\left(\hat{\rho}\right)\right)}$$

with

$$\sum_k A_k^{'} \hat{\rho} A_k^{'\dagger} = A^{'}\hat{\rho}A^{'\dagger}.$$

Again the operator $A^{'}$ satisfies $0 < A^{'\dagger}A^{'} \leq I$. With $p = \mathrm{tr}\left(\mathcal{L}\left(\rho\right)\right)$ and $p' = \mathrm{tr}\left(\mathcal{L}^{'}\left(\hat{\rho}\right)\right)$ we thus have

$$p'\rho = A^{'}\hat{\rho}A^{'\dagger} \tag{9}$$

$$p\hat{\rho} = A\rho A^\dagger. \tag{10}$$

We know that $A^\dagger A$ and $A^{'\dagger}A^{'}$ are non-singular since they have non-zero eigenvalues (which follows from their positive definite form). So

$$\det A^\dagger A = \det A^\dagger \det A \neq 0$$

and therefore $\det A^\dagger \neq 0$ and $\det A \neq 0$. So A and $A^\dagger$ are also non-singular. By rewritting (9) and (10) as

$$\rho = \frac{A^{'}}{\sqrt{p'}}\hat{\rho}\frac{A^{'\dagger}}{\sqrt{p'}}$$

$$\hat{\rho} = \frac{A}{\sqrt{p}}\rho\frac{A^{\dagger}}{\sqrt{p}}$$

we thus see that stochastic LOCC implies that the two pure states ρ and $\hat{\rho}$ are connected by an invertible local operator.

Suppose instead that ρ and $\hat{\rho}$ are two pure states related through an transformation with probability p given by an invertible local operator A such that

$$\hat{\rho} = A\rho A^{\dagger}.$$

We then have that $\mathrm{tr}\left(A\rho A^{\dagger}\right) = 1$. Define a new operator $A^{'}$ by

$$A^{'} = \sqrt{p}A.$$

We then have that $\mathrm{tr}\left(A^{'}\rho A^{'\dagger}\right) = p$ and $A^{'}$ is of course invertible and multilocally implementable. We also have that $A^{'}$ satisfies $0 < A^{'\dagger}A \leq I$, so by defining $\mathcal{L}\left(\rho\right) = A^{'}\rho A^{'\dagger}$ we see that ρ and $\hat{\rho}$ are stochastically equivalent.

3.3.3 Entanglement properties under LOCC and stochastic LOCC

We have seen that equivalence under LOCC yields a partitioning of the set of states where two states are in the same equivalence class if and only if they are related by a local unitary transformation. But if one of the parties sharing a composite state applies a local unitary transformation which is not a unitary evolution it only means that a local change of basis is done. So the physical properties of the composite state should remain invariant under the application of a local unitary transformation, and in particular the entanglement properties should not change. We can therefore identify the partitioning the LOCC equivalence gives rise to as a partitioning of the set of entanglement values. Every entanglement measure should therefore remain constant on the equivalence classes.

For stochastic equivalence the situation is somewhat different. We have seen that the stochastic equivalence amounts to equivalence under general invertible transformations. This means that we have a coarser partitioning of the set of states, where the equivalence classes are related to structural features of the entanglement rather than the actual value given by some entanglement measure. This coarser partitioning is valuable, however, since it shows which states are accessible to different parties with non-zero probability.

The transformations under stochastic equivalence live in

$$GL\left(n_1\right) \otimes GL\left(n_2\right) \otimes \cdots \otimes GL\left(n_m\right), \tag{11}$$

where n_i is the dimension of the i:th subsystem. When we deal with unnormalized state we may fix the determinant of all the local operators to one, since the two local operators A and kA, where $k \in \mathbb{C}$, will only differ through the introduction of a global constant to the transformed state. We thus identify equivalence under stochastic LOCC by the multilocal unimodular transformations in

$$SL\left(n_1\right) \otimes SL\left(n_2\right) \otimes \cdots \otimes SL\left(n_m\right). \tag{12}$$

One must be careful, however, to note that transformations in (11) will not give rise to the same partitioning as those in (12) when one deals with normalized states. We will address this issue in more detail in section 6.5.2 when dealing with tripartite qubit states. To avoid confusion we will use the transformations in (11).

Let us now move on to study how we can classify states according to their equivalence under either LOCC or stochastic LOCC.

4 Classifications: Orbits, entanglement types and invariants

So far we have discussed two kinds of protocols we might apply in our laboratory, LOCC and stochastic LOCC. We have seen that these protocols are essential when it comes to manipulating our states so that they fit into the scheme we want to deploy in the production of some concrete task. We have also seen that at the heart lies the phenomenon of entanglement. The equivalence conditions that we impose on our states under either LOCC or stochastic LOCC are intimately connected to the entanglement properties of the states. We would thus like to impose some kind of classification of the states so that we would know what physical properties they correspond to. This is where Carl von Linné enters the laboratory.

4.1 Carl von Linné in the laboratory

In his great classification of the natural world (although mainly confined to plants) Linné used a static taxonomy where each individual organism would belong to some species according to a description of its constituent parts. This stands in contrast to the modern approach where individuals belong to a species due to a shared evolutionary origin. But the Linnéan approach is very similar to the classification schemes that we would like to employ on our quantum states. Given an organism, Linné would investigate a finite set of traits of the organism to see which species it belonged to. In the same sense, we would, for a given quantum state, like to find a finite set of properties (in our case, as we will see later on, polynomial functions) that completely determines which class the state belongs to. In terms of multipartite qubit systems we would thus like to produce a great catalogue of classes where the polynomial functions are used to order all classes and their physical properties are clearly distinguished through these polynomial functions. This catalogue would then be used in the construction of more elaborate information processing technology.

It turns out that the creation of such a catalogue is exceedingly difficult. To compute all necessary polynomial functions is very hard for multipartite systems and the complexity grows quickly with the number of subsystems involved. It is possible to do algebraic construction for small systems, but one has to resort to the aid of computers for more complex systems. Even so, without the advent of some improved algorithm, the Linnean ambition comes to a halt quite early in the classification endeavour.

But nevertheless we can put the project of classifying all quantum states up to equivalence in a clear mathematical framework since we now know what types of transformations the two partitions under LOCC and stochastic LOCC correspond to. Since the unitary and general (or unimodular) transformations have a group structure, we may use this to relate our ambitions to some classical mathematical problems.

4.2 Orbit spaces and entanglement types

We have seen that exact reducibility implies LUT equivalence and stochastic equivalence implies equivalence under general invertible local operators. This means that we can shift perspective and work in a group theoretical framework. In order to do this we need to introduce some relevant definitions.

Definition. *An action of a group G on a set X is a mapping*

$$\pi : G \times X \to X$$

denoted by $gx \equiv \pi(g, x)$ satisfying

$$ex = x \qquad h(kx) = (hk)x$$

for all $h, k \in G$ and $x \in X$.

We refer to X as a G-set when we let G act on X. Now let R be a relation such that xRy if and only if there exists a $g \in G$ with $gx = y$. This is an equivalence relation.

Definition. *The equivalence classes under R are called orbits. The orbit of a given element x is formed by the elements gx with $g \in G$ and is denoted Gx.*

The set X is thus partitioned into orbits and the set of all orbits is denoted ${}^{X}\!/_{G}$. These two definitions immediately give us the big picture. The groups

$$U(n_1) \otimes U(n_2) \otimes \cdots \otimes U(n_m) \tag{13}$$

and

$$GL(n_1) \otimes GL(n_2) \otimes \cdots \otimes GL(n_m) \tag{14}$$

act on the set of states and thus induce a partitioning where two elements are exactly or stochastically equivalent if they are on the same orbit. Our space of interest will thus be the orbit space. Orbits under (13) thus correspond to states with the same entanglement measure and orbits under (14) correspond to states with the same entanglement structure. To get a better grasp of the partitions and the entanglement properties we want to know the dimension of the orbit space and the dimension of the orbits. Furthermore we would like to know how we can decide whether two states are on the same orbit. Using the coefficients of our states we would thus form some function that should be invariant for all states in the orbit. If the function differs for two states we know that they are inequivalent. If the function is the same, however, can we infer that the two states are equivalent? The answer is no; there is no reason why the function could not have the same value on two orbits. So how many functions do we need to compute before we can know that two states are equivalent? How can we even be certain that there exists a finite set of functions that allows us to distinguish the orbits? Staring at the complicated topology of our orbit space our initial enthusiasm is quickly buried in the pond of despair that our questions gives rise to. If we can not even be sure that two states are equivalent, what use is our entire construction? But then we suddenly remember something that the math department talked about a long time ago, and at once we are fit to proceed thanks to Hilbert's Basis Theorem.

5 Classical Invariant Theory: Invariant properties of Forms

From the high-tech quantum laboratory of the 21:st century we now move to the provincial German town of Göttingen at the end of the 19:th century. It was here that Hilbert and others grappled with the mathematics underlying our basic questions. The body of knowledge created through the study of transformation properties of homogenous polynomials came to be known as Invariant Theory.

Even though the reader might never have heard of Invariant Theory, she has undoubtedly encountered invariants already in high-school while solving quadratic polynomial equations. In the following all coefficients are complex. The general quadratic polynomial equation

$$ax^2 + 2bx + c = 0$$

has solutions

$$x = \frac{-b \pm \sqrt{\Delta}}{a}$$

where

$$\Delta = b^2 - ac$$

is known as the *discriminant*. We immediately see that if $\Delta = 0$ we have a double root and if $\Delta < 0$ the roots are complex conjugate. What will happen to the discriminant if we do an affine transformation

$$\bar{x} = \alpha x + \beta \qquad\qquad \alpha \neq 0$$

that corresponds to a scaling and a translation? A straightforward calculation yields that our new polynomial is

$$\bar{a}\bar{x}^2 + 2\bar{b}\bar{x} + \bar{c}$$

where

$$\bar{a} = \frac{a}{\alpha^2}$$
$$\bar{b} = \frac{b\alpha - a\beta}{\alpha^2}$$
$$\bar{c} = \frac{c\alpha^2 + a\beta^2 - 2b\beta\alpha}{\alpha^2}$$

Upon calculating the determinant for the polynomial expressed in our new variable $\bar{x}$ we see that

$$\bar{\Delta} = \frac{1}{\alpha^2}\left(b^2 - ca\right) = \frac{1}{\alpha^2}\Delta$$

Interestingly enough, up to a multiplicative factor depending only on the transformation, the discriminant has remained unchanged under our affine transformation. This is not a coincidence; as we saw the discriminant gave us information about the types of roots our polynomial has, and these properties of the roots should remain unchanged as long as we only rescale and translate our polynomial.

This simple example of a property that is invariant under a transformation and thus gives us some purely geometric information about our polynomial really describes the subject matter of Invariant Theory. The reader might think that the study of invariant properties of polynomials is a rather special topic without bifurcations into more general spheres of mathematics or physics. But nothing could be more wrong. Classical Invariant Theory is to algebra what Klein's Erlangen Program is to geometry. Both helped to classify invariant properties under algebraic or geometric manipulations (and not to say the least about the deep connections between them!) which in turn was an important underpinning of one of the great themes of 20:th century mathematical physics: The question of symmetry and invariance.

With this background we now introduce the basic definitions and results. If the reader wishes to read a more comprehensive introduction [Olv] is highly recommended. It appeals to mathematicians as much as to physicists since everything is done in characteristic zero. For a more classical approach involving the somewhat enigmatic symbolic method the reader should consult [Tu]. Finally, Hilbert's original lectures [Hi] on the topic are also recommended. They are clear, accessible and give a firm grasp of the kind of problems that occupied the mind of late 19:th century mathematicians.

5.1 Linear Transformations of Forms

In the preceding example we looked at a quadratic polynomial in one single variable. In Invariant Theory the interest is chiefly on homogenous polynomials in several variables x_i. Homogeneity means that

$$Q\left(tx\right)=t^{d}Q\left(x\right)$$

where $x=\left(x_1,\ldots,x_n\right)$. Here d is called the degree of the polynomial and n the order. A homogenous polynomial is classically referred to as a *form.* A general homogenous polynomial of order d in n variables is written as

$$Q\left(x\right)=\sum_{I}\left(\begin{array}{c} d \\ I \end{array}\right)a_I x^I$$

where

$$x^I=x_1^{i_1}\cdots x_n^{i_n}$$

and the summation is over all multi-indices $I=\left(i_1,\ldots,i_n\right)$ with $i_k\geq 0$ and $i_1+\cdots+i_n=d$. Thus, for example, a quadratic form in two variables (also known as a binary form) is written as

$$Q\left(x,y\right)=a_2x^2+2a_1xy+a_0y^2$$

and a general binary form is written as

$$Q\left(x,y\right)=\sum_{i=0}^{n}\left(\begin{array}{c}n\\ i\end{array}\right)a_{i}x^{i}y^{n-i}.$$

The binomial coefficients are introduced for convenience. In the following we restrict the attention to binary forms since this will simplify the notation. There is no loss of generality in doing this since the theory in principle looks the same for higher orders. Notice that we have two fundamental sets: the coefficients a_i and the variables x_i. A general linear invertible transformation of two variables takes the form

$$\bar{x}=\alpha x+\beta y$$
$$\bar{y}=\gamma x+\delta y$$

where $\alpha\delta-\beta\gamma\neq 0$. When the variables are transformed a given polynomial is also transformed as

$$\bar{Q}\left(\bar{x},\bar{y}\right)=\bar{Q}\left(\alpha x+\beta y,\gamma x+\delta y\right)=Q\left(x,y\right)$$

and we thus get an induced transformation of the coefficients. There is an explicit formula for the new coefficients, but in reality it does not carry much importance in the determination of invariants and therefore we will not state it here. But it is important to note that the general linear invertible transformation transforms homogenous polynomials to homogenous polynomials of the same degree. This is the reason why attention is almost exclusively paid to homogenous polynomials: We can always decompose a general inhomogenous polynomial as a sum of homogenous polynomials.

5.2 Polynomial Invariants

Having seen that a general linear invertible transformation transforms a form to another form of the same degree and order we can thus state the general definition of a invariant.

Definition. *An invariant of a binary form $Q(x,y)$ is a function $I\left(a\right)=I\left(a_0,\ldots,a_n\right)$ depending on the coefficients $a=\left(a_0,\ldots,a_n\right)$ of the form, which, up to a determinantal factor, does not change under the general linear transformation:*

$$I\left(a\right)=\left(\alpha\delta-\beta\gamma\right)^{k}I\left(\bar{a}\right),$$

where $\bar{a}=\left(\bar{a}_0,\ldots,\bar{a}_n\right)$ are the coefficients of the transformed polynomial.

The integer k is called the *weight* of the invariant. In particular we will pay attention to when $I\left(a\right)$ is a polynomial. It is then called a polynomial invariant. The definition raises several questions. How many independent polynomial invariants are there of a given degree? What do they tell us about the form? Can we find a basis for the polynomial invariants?

Before we present some of the results relating to these questions we have to introduce another important definition. The function $I(a)$ depended only on the coefficients $a = (a_0, \ldots, a_n)$ of the form. But we previously made the rather trivial remark that a form is characterized by two sets, the coefficients and the variables. We would thus like to identify a new invariant function that also includes the variables. Such a function is called a *covariant* and is given by the following definition:

Definition. *A covariant of weight k of a binary form Q of degree n is a function $J(a,x) = J(a_0, \ldots, a_n, x, y)$ depending both on the coefficients a_i and on the independent variables $x = (x,y)$ which, up to a determinantal factor, is unchanged under linear transformations:*

$$J(a,x) = (\alpha\delta - \beta\gamma)^k \bar{J}(\bar{a}, \bar{x})$$

From this definition follows that an invariant is just a covariant that does not depend explicitly on the variables. In the analysis of quantum states we will not make use of the concept of a covariant. But it is such a central concept in Invariant Theory that it would be shameful not to mention it.

Let us look at the invariants and covariants of a cubic binary form

$$Q(x) = a_3 x^3 + 3a_2 x^2 y + 3a_1 x y^2 + a_0 y^3.$$

There is one fundamental invariant which is the discriminant of the cubic:

$$\Delta = a_0^2 a_3^2 - 6a_0 a_1 a_2 a_3 + 4a_0 a_2^3 - 3a_1^2 a_2^2 + 4a_1^3 a_3.$$

The obvious covariant for all forms is the form itself, Q. Another important covariant is the *Hessian*

$$H = \det \begin{pmatrix} Q_{xx} & Q_{xy} \\ Q_{xy} & Q_{yy} \end{pmatrix} = Q_{xx} Q_{yy} - Q_{xy}^2$$

which for the cubic binary form is given by

$$\frac{1}{36} H = \left(a_1 a_3 - a_2^2\right) x^2 + \left(a_0 a_3 - a_1 a_2\right) xy + \left(a_0 a_2 - a_1^2\right) y^2.$$

The final covariant (all other covariants can be given by polynomial combinations of these three and the discriminant) is the *Jacobian* between the Hessian and the form itself:

$$T = Q_x H_y - Q_y H_x$$

We will not give the complete expression here since it is rather lengthy and not especially enlighting. So given these three covariants and the discriminant we can use them to completely classify the cubic binary form in terms of the character of its roots. The interested reader may consult [Olv] for this and related classifications.

Let us now move on to the fundamental results concerning invariants and covariants that have direct bearing on our original goal: How to characterize the entanglement types of quantum states. The reader may not be entirely convinced that the study of polynomials is suitable to this end but the following theorems will hopefully show how algebraic generalizations can be made which are applicable in other frameworks.

5.3 Fundamental Results

There are two Fundamental Theorems related to Invariant Theory. To introduce them would take us too far afield since they deal with concepts relating to different methods of creating invariants and covariant. Suffice to say, however, they basically claim that there is a process, called transvection, for creating all the possible independent invariants and covariants and that there is a finite list of non-polynomial relations (so-called syzygies) between the invariants and covariants.

But what about the finiteness of the invariants and covariants? Is there an infinite supply of invariants and covariants for a given form, or is there a finite basis so that all other invariants and covariants can be expressed in this basis? Let us start with the relevant definition:

Definition. *Suppose $Q_1, \ldots, Q_l$ are a collection of binary forms. A finite collection of invariants $I_1, \ldots, I_m$ forms a Hilbert basis if every other invariant can be written as a polynomial function of the basis invariants: $I = P(I_1, \ldots, I_m)$. Similarly, a finite collection of covariants $J_1, \ldots, J_k$ forms a Hilbert basis if every other covariant J can be written as a polynomial in the basis covariants: $J = P(J_1, \ldots, J_k)$.*

It was postulated that for binary forms of degree seven or more there does not exist a finite basis. But in 1868 Gordan proved his Finiteness Theorem (also known as Gordan's Theorem) for binary forms. So, irrespective of degree, all binary forms have a Hilbert basis. The theorem was constructive since it actually gave an explicit procedure for the construction of the basis. But what about higher order forms? Once again, a conjecture was made that there does not exist a finite basis in general. So the mathematical community was subsequently shocked when Hilbert proved that this actually holds for all forms:

Theorem. *Any finite system of homogeneous polynomials admits a Hilbert Basis for its invariants, as well as for its covariants.*

His proof was not constructive, however, which made Gordan remark: “Das ist Theologie und nicht Mathematik.” Hilbert created a second, more difficult but constructive proof. It is an open question how to actually implement this constructive proof computationally.

Our fundamental result that we carry over into our investigation of quantum states is thus the existance of a finite set of generating invariants which may be used to distinguish between inequivalent states.

Furthermore, we might wonder if this finiteness also holds when we restrict our transformations to a subgroup of the general invertible transformations, wich was the case for the LUT equivalence of LOCC. Hilbert and Weyl showed that for every reductive group (that is a group for which all linear representations are direct sums of irreducible representations) this holds. In particular, it will hold for our local unitary transformations. Whether it holds for all subgroups of the general linear invertible transformations became Hilbert's 14:th problem and it was not until 1959 that Nagata gave a counterexample.

We now turn to the question of equivalence and canonical forms.

5.4 Equivalence and Canonical Forms

If we let $P(n, d)$ be the set of all homogenous polynomials of order n and degree d, then $G = GL(n, \mathbb{C})$ defines an action on $P(n, d)$ through the mapping $x' = Gx$ where $x = (x_1, x_2, \ldots, x_n)$ are the variables of the polynomial. The equivalence problem is then to find conditions when two polynomials p and p' in $P(n, d)$ lie on the same orbit, that is $p' = gp$ for some $g \in GL(n, \mathbb{C})$. For each such orbit we can pick a polynomial that has some particularly simple form and let that polynomial be a representative of the orbit. Such a polynomial is called a canonical form and a complete list of canonical forms can thus be identified with a complete list of the orbits. So by observing whether two polynomials can be transformed to the same canonical form one can solve the equivalence problem.

In Invariant Theory we saw that there exist invariants and covariants that, up to a determinantal factor, do not change under general linear transformations. We would therefore like to have a complete list of invariants and covariants so that a given polynomial can be transformed into a canonical form if and only if they both attain the same value on the invariants and covariants of our list. As we saw through Hilbert's Basis Theorem, such a list does indeed exist.

We now move on to see how we may translate this procedure into the framework of quantum states.

5.5 From Forms to States

We have seen that Invariant Theory describes invariant properties of homogenous polynomials under general linear transformations. How may we adapt this knowledge to our problem of finding conditions of equivalence under LOCC and stochastic LOCC? If we write a qubit state in multilinear form

$$f\left(x\right)=f\left(x^{(1)},\ldots,x^{(k)}\right)=\sum_{i_1,\ldots,i_k=0}^{1}a_{i_1i_2\cdots i_k}x_{i_1}^{(1)}\cdots x_{i_k}^{(k)}, \tag{15}$$

where $x^{(j)}=\left(x_0^{(j)},x_1^{(j)}\right)$ are pairs of variables, we get an action of the operator $g=\left(g^{(1)},\ldots,g^{(k)}\right)$ by $gx=x'$ where $x'^{(j)}=g^{(j)}x^{(j)}$. The coefficients $a'_{i_1\cdots i_k}$ are defined by the equation

$$\sum_{i_1,\ldots,i_k=0}^{1}a_{i_1i_2\cdots i_k}x_{i_1}^{(1)}\cdots x_{i_k}^{(k)}=\sum_{i_1,\ldots,i_k=0}^{1}a'_{i_1i_2\cdots i_k}x_{i_1}'^{(1)}\cdots x_{i_k}'^{(k)}.$$

We thus want to find invariants of the form

$$I\left(a,\bar{a}\right)=I\left(a',\bar{a}'\right)$$

where a is shorthand for the coefficients and $\bar{a}$ is the complex conjugate. In this case we can apply results from Invariant Theory in order to construct and investigate these invariants.

6 Case study: Bipartite and tripartite qubit states

Having thus discussed Invariant Theory and its procedures for the classification of forms we now return to see how this could help us when analyzing quantum states. We will look at the two fully understood states: the bipartite and tripartite qubit systems.

6.1 The bit, the qubit and the tensor product

In classical information procedures the standard information unit is the bit. It is characterized by two possible states: 0 and 1. The simplicity of these two discrete states makes it possible to implement the bit in a wide variety of physical medium. In quantum information theory we no longer have a discrete distribution of states, but rather a linear superposition of basic states. A qubit is thus a superposition

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$$

subject to the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. Upon measurement the state will be found either as $|0\rangle$ with probability $|\alpha|^2$ or as $|1\rangle$ with probability $|\beta|^2$. This interplay of superposition and probability changes some of the fundamental laws in classical information theory. For example, the total entropy of a system might be smaller than the entropy of one of the subsystems. Such a situation would never be possible in a classical context.

The outstanding feature of quantum information theory is the phenomenon of entanglement. This is a statistical effect that is not reproducible by any classical scheme. In quantum mechanics systems are joined under a tensor product which means that we have the possibility of global superpositions of joint states. The tensor product is written as $\mathcal{H}_A \otimes \mathcal{H}_B$ and if $|e_i\rangle$ and $|f_j\rangle$ is a basis for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, then $|e_i\rangle \otimes |f_j\rangle$ will be a basis of $\mathcal{H}_A \otimes \mathcal{H}_B$. If n and m are the dimensions of $\mathcal{H}_A$ and $\mathcal{H}_B$, the dimension of $\mathcal{H}_A \otimes \mathcal{H}_B$ will be nm. So a system is defined by the number of tensor products and the dimension of the subsystems. If all subsystems have dimension two they are called qubit systems. If we have a tensor product between two qubit systems it is called bipartite and if we have a tensor product between three qubit systems it is called tripartite.

We now move on to study the bipartite qubit system.

6.2 Bipartite qubit states

The bipartite qubit system is written as $\mathbb{C}^2 \otimes \mathbb{C}^2$ and a generic state has the form

$$|\psi\rangle = \sum_{i,j=0}^{1} a_{ij}\,|e_i\rangle \otimes |f_j\rangle\,. \tag{16}$$

For a bipartite system (irrespective of the dimensions of the two subsystems) we have a very powerful tool called the *Schmidt decomposition.* It allows us to change the basis of a given state so that we can write it as a single instead of a double sum.

Theorem. *Every pure state in the Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ can be expressed in the form*

$$|\psi\rangle = \sum_{i=1}^{n} \sqrt{\lambda_i}\, |e_i\rangle \otimes |f_i\rangle$$

where $\{|e_i\rangle\}_{i=0}^{n_1}$ and $\{|f_j\rangle\}_{j=0}^{n_2}$ is an orthonormal basis for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, and $n \leq \min\{n_1, n_2\}$.

A proof of the theorem may be given as follows. Starting from the generic form

$$|\psi\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} a_{ij}\, |\hat{e}_i\rangle \otimes \left|\hat{f}_j\right\rangle \tag{17}$$

we can form the density operator $\rho_\psi = |\psi\rangle\langle\psi|$. Without loss of generality we can assume that $n_1 \geq n_2$. We can perform a partial trace on ρ_ψ to get the reduced density operator for system A:

$$\rho_A = \mathrm{tr}_B\left(\rho_\psi\right) = \sum_{i,j=1}^{n_1}\sum_{n=1}^{n_2} a_{in} a_{jn}^{\dagger}\, |\hat{e}_i\rangle\langle\hat{e}_j| . \tag{18}$$

But this operator can always be written in diagonal form using a unitary transformation:

$$\rho_A = \sum_{i=1}^{n_1} \lambda_i\, |e_i\rangle\langle e_i| . \tag{19}$$

We now reexpress (17) in terms of this new basis:

$$|\psi\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} c_{ij}\, |e_i\rangle \otimes \left|\hat{f}_j\right\rangle , \tag{20}$$

where $c_{ij} = \left\langle e_i \otimes \hat{f}_j \mid \psi \right\rangle$. We can now define

$$|f_i\rangle = \sum_{j=1}^{n_2} \frac{c_{ij}}{\sqrt{\lambda_i}} \left|\hat{f}_j\right\rangle . \tag{21}$$

The new basis for system B will satisfy the orthonormality relation $\langle f_i \mid f_j\rangle = \delta_{ij}$. Using this together with (20) we get the desired result:

$$|\psi\rangle = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \sqrt{\lambda_i}\, |e_i\rangle \otimes \frac{c_{ij}}{\sqrt{\lambda_i}} \left|\hat{f}_j\right\rangle$$

$$= \sum_{i=1}^{n_1} \sqrt{\lambda_i} |e_i\rangle \otimes |f_i\rangle .$$

The coefficients $\sqrt{\lambda_i}$ are called *Schmidt coefficients* and n is the *rank* of $|\psi\rangle$. If we write $|\psi\rangle$ in terms of a pure density operator, $\rho = |\psi\rangle \langle\psi|$, the rank of the operator is defined as the dimension of the range of the operator. This concept is important for density operators since if the rank of a reduced density operator is greater than one we have the presence of entanglement. In the case of bipartite qubit systems this means that every state can be written as

$$|\psi\rangle = x_{00} |00\rangle + x_{11} |11\rangle \tag{22}$$

Here $|00\rangle$ and $|11\rangle$ is shorthand for $|0\rangle \otimes |0\rangle$ and $|1\rangle \otimes |1\rangle$. x_{00} and x_{11} are the Schmidt coefficients. Notice that if either x_{00} or x_{11} equals zero the state is unentangled.

In order to change an arbitrary bipartite qubit state to the form (22) one applies only unitary transformations. We thus have an immediate answer to the question when two states are equivalent under LOCC:

Proposition. *Two bipartite qubit states are equivalent under LOCC when they have equal Schmidt coefficients under a Schmidt decomposition.*

Notice, however, that the general form of a bipartite qubit state is

$$|\psi\rangle = a_{00} |00\rangle + a_{01} |01\rangle + a_{10} |10\rangle + a_{11} |11\rangle , \tag{23}$$

whereas in (22) we had transformed (23) according to the Schmidt decomposition. But we want to express our equivalence in terms of invariant polynomals of the cofficients of (23). This is possible if we include the complex conjugate of the coefficients. If we notice that the square of the Schmidt coefficients are eigenvalues of the reduced density operator ρ_A (or ρ_B) we conclude that the coefficients of the characteristic polynomial of ρ_A must remain invariant. The characteristic polynomial for a 2×2-matrix A can be written:

$$p(A) = \lambda^2 - \operatorname{tr}(A)\lambda + \frac{1}{2}\left(\operatorname{tr}(A)^2 - \operatorname{tr}\left(A^2\right)\right). \tag{24}$$

So in particular $\operatorname{tr}(\rho_A)$ and $\operatorname{tr}\left(\rho_A^2\right)$ must be invariant. For clarity we compute these for the bipartite qubit state. The density operator is given as a matrix

$$\rho = \sum_{ij} \sum_{kl} \rho_{ij,kl} |ij\rangle \langle kl|$$

and the reduced density operator is

$$\rho_A = \operatorname{tr}_B(\rho) = \sum_i \sum_k \sigma_{ik} |i\rangle \langle k|$$

where σ_{ik} is given by

$$\sigma_{ik} = \sum_j \rho_{ij,kj}.$$

Starting from (23) we thus write the pure density operator as the matrix

$$\rho = |\psi\rangle \langle\psi| = \begin{pmatrix} a_{00}\bar{a}_{00} & a_{00}\bar{a}_{01} & a_{00}\bar{a}_{10} & a_{00}\bar{a}_{11} \\ a_{01}\bar{a}_{00} & a_{01}\bar{a}_{01} & a_{01}\bar{a}_{10} & a_{01}\bar{a}_{11} \\ a_{10}\bar{a}_{00} & a_{10}\bar{a}_{01} & a_{10}\bar{a}_{10} & a_{10}\bar{a}_{11} \\ a_{11}\bar{a}_{00} & a_{11}\bar{a}_{01} & a_{11}\bar{a}_{10} & a_{11}\bar{a}_{11} \end{pmatrix}.$$

The reduced density operator is then

$$\rho_A = \mathrm{tr}_B\left(\rho\right) = \sum_{ik} \sigma_{ik} |i\rangle \langle k|$$

where

$$\sigma_{ik} = \sum_j a_{ij}\bar{a}_{kj}$$

Writing this out in detail gives us

$$\rho_A = \begin{pmatrix} a_{00}\bar{a}_{00} + a_{01}\bar{a}_{01} & a_{00}\bar{a}_{10} + a_{01}\bar{a}_{11} \\ a_{10}\bar{a}_{00} + a_{11}\bar{a}_{01} & a_{10}\bar{a}_{10} + a_{11}\bar{a}_{11} \end{pmatrix}$$

and

$$\rho_A^2 = \begin{pmatrix} c_{00} & c_{01} \\ c_{10} & c_{11} \end{pmatrix}$$

where

$$c_{00} = \left(a_{00}\bar{a}_{00} + a_{01}\bar{a}_{01}\right)^2 + \left(a_{00}\bar{a}_{10} + a_{01}\bar{a}_{11}\right)\left(a_{10}\bar{a}_{00} + a_{11}\bar{a}_{01}\right)$$

$$c_{01} = \left(a_{00}\bar{a}_{00} + a_{01}\bar{a}_{01}\right)\left(a_{00}\bar{a}_{10} + a_{01}\bar{a}_{11}\right) + \left(a_{00}\bar{a}_{10} + a_{01}\bar{a}_{11}\right)\left(a_{10}\bar{a}_{10} + a_{11}\bar{a}_{11}\right)$$

$$c_{10} = \left(a_{10}\bar{a}_{00} + a_{11}\bar{a}_{01}\right)\left(a_{00}\bar{a}_{00} + a_{01}\bar{a}_{01}\right) + \left(a_{10}\bar{a}_{10} + a_{11}\bar{a}_{11}\right)\left(a_{10}\bar{a}_{00} + a_{11}\bar{a}_{01}\right)$$

$$c_{11} = \left(a_{10}\bar{a}_{00} + a_{11}\bar{a}_{01}\right)\left(a_{00}\bar{a}_{10} + a_{01}\bar{a}_{11}\right) + \left(a_{10}\bar{a}_{10} + a_{11}\bar{a}_{11}\right)^2.$$

From this we thus see that the invariant polynomials are

$$I_1 = \mathrm{tr}\left(\rho_A\right) = a_{00}\bar{a}_{00} + a_{01}\bar{a}_{01} + a_{10}\bar{a}_{10} + a_{11}\bar{a}_{11}$$

and

$$\begin{aligned} I_2 = \mathrm{tr}\left(\rho_A^2\right) &= a_{00}^2\bar{a}_{00}^2 + a_{01}^2\bar{a}_{01}^2 + a_{10}^2\bar{a}_{10}^2 + a_{11}^2\bar{a}_{11}^2 \\ &+2a_{00}a_{01}\bar{a}_{00}\bar{a}_{01} + 2a_{10}a_{11}\bar{a}_{10}\bar{a}_{11} + 2a_{00}a_{10}\bar{a}_{00}\bar{a}_{10} \\ &+2a_{00}a_{11}\bar{a}_{10}\bar{a}_{01} + 2a_{01}a_{10}\bar{a}_{11}\bar{a}_{00} + 2a_{01}a_{11}\bar{a}_{11}\bar{a}_{01}. \end{aligned}$$

Furthermore, it is a standard result in linear algebra that the coefficients of the characteristic polynomial can be written as elementary symmetric polynomials of the eigenvalues. This follows from the fact that the expansion of a monic polynomial (that is, a polynomial in which the leading coefficient is 1) is

$$\prod_{i=1}^{n}\left(\lambda - X_j\right) = \lambda^n - e_1\left(X_1,\ldots,X_n\right)\lambda^{n-1}$$

$$+e_2\left(X_1,\ldots,X_n\right)\lambda^{n-2} - \cdots + \left(-1\right)^n e_n\left(X_1,\ldots,X_n\right)$$

where $e_i\left(X_1,\ldots,X_n\right)$ is the i:th elementary symmetric polynomial in n variables. The polynomials are given by the formulae

$$\begin{aligned}
e_0\left(X_1,\ldots,X_n\right) &= 1\\
e_1\left(X_1,\ldots,X_n\right) &= \sum_{1\leq i\leq n} X_i\\
e_2\left(X_1,\ldots,X_n\right) &= \sum_{1\leq i<j\leq n} X_iX_j\\
e_m\left(X_1,\ldots,X_n\right) &= \sum_{1\leq i<j<k\leq n} X_iX_jX_k\\
&\vdots\\
e_n\left(X_1,\ldots,X_n\right) &= X_1X_2\cdots X_n.
\end{aligned}$$

In the case of bipartite qubit states we have

$$\left(\lambda - x_{00}^2\right)\left(\lambda - x_{11}^2\right) = \lambda^2 - \left(x_{00}^2 + x_{11}^2\right)\lambda + x_{00}^2x_{11}^2.$$

and comparing this with (24) we thus have

$$\lambda^2 - \left(x_{00}^2 + x_{11}^2\right)\lambda + x_{00}^2x_{11}^2 = \lambda^2 - \mathrm{tr}\left(\rho_A\right)\lambda + \frac{1}{2}\left(\mathrm{tr}\left(\rho_A\right)^2 - \mathrm{tr}\left(\rho_A^2\right)\right).$$

So we can relate our invariants to symmetric elementary polynomials of the Schmidt coefficients:

$$\begin{aligned}
I_1 &= x_{00}^2 + x_{11}^2\\
\frac{1}{2}\left(I_1^2 - I_2\right) &= x_{00}^2x_{11}^2
\end{aligned}$$

It also makes sense that we have to check two invariant polynomials, since we have two Schmidt coefficients that have to be equal for LOCC equivalence.

The previous results can be generalized to the fact that the polynomial invariants of bipartite states are the polynomials $I_n = \mathrm{tr}\left(\rho^n\right)$ where n goes from 1 to the minimum of the dimension of the two systems involved. The polynomials I_n can also be related to elementary symmetric polynomials of the eigenvalues of ρ. See [Sud] for a more thorough discussion of the permutation structure behind bipartite states.

What then characterizes equivalence under stochastic LOCC? We will get two discrete orbits under SLOCC corresponding to entanglement and no entanglement.

Proposition. *Two bipartite qubit states are equivalent under stochastic LOCC when they have equal rank.*

Two states have bipartite states have equal rank if they have the same number of non-zero Schmidt coefficients. In the qubit scenario this simply means that a state is entangled if both Schmidt coefficients are non-zero and unentangled if one is zero.

We have arrived at the type of classification that LOCC and stochastic LOCC should correspond to. LOCC should single out the orbits where the entanglement is the same, while stochastic LOCC should have a classification according to structural properties of entanglement.

6.3 Tripartite qubit states

We now move on to look at tripartie qubit states, $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$. In the bipartite scenario the Schmidt decomposition was a valuable tool in analyzing the orbits. But the decomposition does not generalize to multiple tensor products. We have to resort to other tools to analyze the orbits. In the case of LOCC equivalence, Invariant Theory is used to analyze the oribts. For stochastic LOCC the orbits are characterized by the rank of the reduced density operators.

In the bipartite scenario the situation was clear-cut regarding the structural properties of entanglement: Either there was the presence of entanglement because of superposition or not. For tripartite qubit states we can have different kinds of entanglement. For example, the first system could be entangled with the second, but not to the third. Or all systems are entangled with each other. This means that we should have different orbits under stochastic LOCC that correspond to these different scenarios.

We will see that there is a finite set of polynomial invariants that can be used to describe the orbits under both LOCC and stochastic LOCC.

6.4 Orbits under LOCC

The generators of the polynomial invariants under LOCC have been computed in several places, for example [Gr], [Ke], [Sud] and [LTT]. The fundamental invariants are interesting in themselves since they tell us something about the physical properties about the states (so they are not only mathematical devices used to separate the orbits). In the following we will look at the invariants as calculated by Sudbery in [Sud] since they can be used rather neatly to distinguish the orbits under stochastic LOCC. Sudbery only computed six polynomial invariants, but [Gr] and [LTT] gave the complete list of the polynomial invariant generators with seven invariants. But it is the invariants that Sudbery computed that will be our chief attention since they can be used in the classification of the orbits under stochastic LOCC.

6.4.1 Computing the polynomial invariants

Let a generic tripartite qubit state be written as

$$|\psi\rangle = \sum_{i,j,k=0}^{1} a_{ijk} \, |ijk\rangle$$

In order to calculate the invariants Sudbery used the fact that any polynomial in a_{ijk} which is invariant under the action of $U(2) \otimes U(2) \otimes U(2)$ is a sum of homogeneous polynomials of even degree of the form

$$P_{\sigma\tau}(a) = a^{i_1 j_1 k_1} \cdots a^{i_r j_r k_r} \bar{a}_{i_1 j_{\sigma(1)} k_{\tau(1)}} \cdots \bar{a}_{i_r j_{\sigma(r)} k_{\tau(r)}}$$

where σ and τ are permutations of $(1, \ldots, r)$ and $2r$ is the degree of the polynomial. We use the summation convention of repeated indices with one index in the upper position and one in the lower. Using this procedure together with an analysis of the relevant permutation structure the following invariants were calculated.

The only independent polynomial invariant of degree 2 is the norm of $|\psi\rangle$:

$$I_1 = a^{ijk} \bar{a}_{ijk} = \langle \psi | \psi \rangle \, . \tag{25}$$

For degree 4 we have the following linearly independent quartic invariants:

$$I_2 = a^{i_1 j_1 k_1} \bar{a}_{i_1 j_1 k_2} a^{i_2 j_2 k_2} \bar{a}_{i_2 j_2 k_1} = \operatorname{tr}\left(\rho_C^2\right) \tag{26}$$

$$I_3 = a^{i_1 j_1 k_1} \bar{a}_{i_1 j_2 k_1} a^{i_2 j_2 k_2} \bar{a}_{i_2 j_1 k_2} = \operatorname{tr}\left(\rho_B^2\right) \tag{27}$$

$$I_4 = a^{i_1 j_1 k_1} \bar{a}_{i_1 j_2 k_2} a^{i_2 j_2 k_2} \bar{a}_{i_2 j_1 k_1} = \operatorname{tr}\left(\rho_A^2\right) . \tag{28}$$

The polynomial invariant of degree 6 is

$$\begin{aligned} I_5 &= a^{i_1 j_1 k_1} a^{i_2 j_2 k_2} a^{i_3 j_3 k_3} \bar{a}_{i_1 j_2 k_3} \bar{a}_{i_2 j_3 k_1} \bar{a}_{i_3 j_1 k_2} \\ &= \left(\rho_{BC}\right)^{j_1 k_1}_{j_2 k_3} \left(\rho_{BC}\right)^{j_2 k_2}_{j_3 k_1} \left(\rho_{BC}\right)^{j_3 k_3}_{j_1 k_2} . \end{aligned} \tag{29}$$

This expression would be the same if, instead for ρ_{BC}, we used ρ_{AC} or ρ_{AB}. The invariant of degree 8 is given as

$$I_6 = 2 \left| \epsilon_{i_1 i_2} \epsilon_{i_3 i_4} \epsilon_{j_1 j_2} \epsilon_{j_3 j_4} \epsilon_{k_1 k_3} \epsilon_{k_2 k_4} a^{i_1 j_1 k_1} a^{i_2 j_2 k_2} a^{i_3 j_3 k_3} a^{i_4 j_4 k} \right| \tag{30}$$

where $\epsilon_{01} = -\epsilon_{10} = 1$ and $\epsilon_{00} = \epsilon_{11} = 0$ is the antisymmetric tensor. All of these six polynomial invariants carry a physical significance related to the entanglement properties of the tripartite qubit state. But there also exists an invariant of degree 12, whose physical significance has not been explored, with 5988 terms [Gr]. We will not state it here.

Before we investigate the properties of our six invariants of degree up to 8, we need to introduce the concept of the *2-tangle* and *3-tangle*.

6.4.2 Distributed Entanglement

In [CKW] Coffman, Kundu and Wootters introduced the entanglement measure between tripartite states called the *3-tangle*. It is a measure of the global entanglement between the three systems that can not be expressed as the sum of local entanglement between the subsystems.

To begin with, let ρ_{AB} be a density matrix of a pure or mixed bipartite state and define the "spin-flipped" density matrix as

$$\bar{\rho}_{AB} = (\sigma_y \otimes \sigma_y)\, \rho^*_{AB}\, (\sigma_y \otimes \sigma_y)\,.$$

Here ρ^*_{AB} is complex conjugation of ρ_{AB} and σ_y is one of the Pauli matrices and can be written as $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. Since both ρ_{AB} and $\bar{\rho}_{AB}$ are positive operators, their product $\rho_{AB}\bar{\rho}_{AB}$ will also have real and non-negative eigenvalues, according to [CKW]. Let the square roots of these eigenvalues, in decreasing order, be λ_1, λ_2, λ_3 and λ_4. The tangle of ρ_{AB} is then defined as

$$\tau_{AB} = \left[\max\left\{\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4, 0\right\}\right]^2.$$

This is a measure of the entanglement for mixed states, where $\tau = 0$ corresponds to an unentangled state and $\tau = 1$ to a competely entangled state. The question Coffman et al. asked is whether there exists, for a pure tripartite qubit state, an inequality relating the tangles τ_{AB} and τ_{AC} with the tangle between the systems B and C taken together, $\tau_{A(BC)}$. Notice that the composite system BC has dimension four but is still treated in this context as a qubit of dimension two. This is possible since only two dimensions are required for a full description of the pure state ρ_{ABC}, namely those that are spanned by the eigenstates of ρ_{BC} that have non-zero eigenvalues. With this in background Coffman et al. showed that a relevant inequality exists, namely

$$\tau_{AB} + \tau_{AC} \leq \tau_{A(BC)}.$$

This inequality is also as sharp as it can be, since that for the state

$$|W\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right)$$

it reduces to an equality. The 3-tangle, or residual tangle, is defined as

$$\tau_{ABC} = \tau_{A(BC)} - \tau_{AB} - \tau_{AC}.$$

Mathematically it is unchanged under permutation of the systems:

$$\tau_{ABC} = \tau_{B(CA)} - \tau_{BC} - \tau_{BA}$$

and so on. It therefore measures an entanglement that is shared by all systems and is not restricted to the entanglement of one system with the other systems. Interestingly enough, the 3-tangle is directly expressible in terms of the modulus of the hyperdeterminant.

6.4.3 Enter the hyperdeterminant

The hyperdeterminant was originally introduced by Cayley in 1845 in one of the first investigations of Invariant Theory. We can arrange the 8 coefficients as a $2 \times 2 \times 2$ *hypermatrix.* The hyperdeterminant is then a quartic polynomial in the coefficients of the hypermatrix. More precisely, if we label the coefficients as a_{ijk} where $i, j, k = \{0, 1\}$ the hyperdeterminant of the hypermatrix A is written as

$$\text{Hdet}\,(A) = d_1 - 2d_2 + 4d_3 \tag{31}$$

where

$$d_1 = a_{000}^2 a_{111}^2 + a_{001}^2 a_{110}^2 + a_{010}^2 a_{101}^2 + a_{100}^2 a_{011}^2$$

$$d_2 = a_{000}a_{111}a_{011}a_{100} + a_{000}a_{111}a_{101}a_{010} + a_{000}a_{111}a_{110}a_{001}$$

$$+a_{011}a_{100}a_{101}a_{010} + a_{011}a_{100}a_{110}a_{001} + a_{101}a_{010}a_{110}a_{001}$$

$$d_3 = a_{000}a_{110}a_{101}a_{011} + a_{111}a_{001}a_{010}a_{100}.$$

We can get a nice pictorial interpretation of d_1, d_2 and d_3 if we arrange the coefficients a_{ijk} as corners of a cube where $i = \{0, 1\}$ corresponds to vertical displacement, $j = \{0, 1\}$ to horizontal discplacement and $k = \{0, 1\}$ to depth displacement (see Figure 1). In this scenario the terms in d_1 corresponds to diagonal lines, in d_2 to diagonal planes and in d_3 to a tetrahedron. For all the different terms in d_1, d_2 and d_3 the "center of mass" in the geometrical arrangement will always lie in the center of the cube. We will observe in section 6.5.1 that this arrangement makes sense for tripartite qubit systems and can also be used to identify canonical forms geometrically.

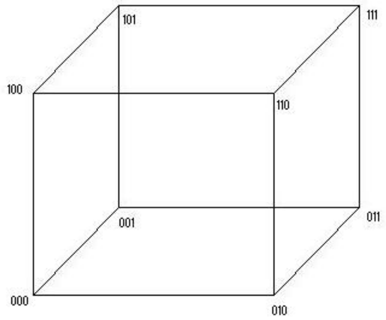

Figure 1: Distribution of eight coefficients on a cube to give a geometrical interpretation of the hyperdeterminant.

Since a generic state in $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ can be arranged in terms of a hypermatrix there is no unambiguity in writting

$$\text{Hdet}\,(|\psi\rangle) = d_1 - 2d_2 + 4d_3.$$

It now turns out that the 3-tangle of a state $|\psi\rangle$ can be written (see [CKW]):

$$\tau_{ABC} = 4\left|d_1 - 2d_2 + d_3\right| = 4\left|\mathrm{Hdet}\left(|\psi\rangle\right)\right|$$

At this point the reader should be impressed by the fact that an abstract mathematical device invented in 1845 resurfaces in the beginning of the 21:th century as a tool for classifying entanglement properties.

6.4.4 Physical properties in terms of polynomial invariants

Looking back at I_2, I_3 and I_4 and remembering that $\mathrm{tr}\left(\rho_A^2\right) = 1$ for unentangled states we see that these quartic invariants express whether system C, B and A are entangled or not. But we also saw that the 2-tangle τ_{AB} is a measure of the entanglement of the bipartite system. It should therefore not come as a surprise that the invariants and the 2-tangle are related. Sudbery computed the exact relations[1] as

$$\tau_{AB} = 1 + I_2 - I_3 - I_4 - \frac{1}{2}I_6 \tag{32}$$

$$\tau_{AC} = 1 - I_2 + I_3 - I_4 - \frac{1}{2}I_6 \tag{33}$$

$$\tau_{BC} = 1 - I_2 - I_3 + I_4 - \frac{1}{2}I_6. \tag{34}$$

Furthermore the 3-tangle is directly expressible in terms of our invariant of degree 8:

$$I_6 = \tau_{ABC}. \tag{35}$$

Since the tangles are expressed in terms of the polynomial invariants, we see that they are constant under local unitary transformations. Furthermore, they will be important when we analyze the orbits under SLOCC.

6.5 The discrete space of stochastic LOCC

We have seen the we have six basic polynomial invariants under LOCC. We thus have a continuous parametrization of the orbits. For stochastic LOCC we get a different situation in the tripartite qubit case. We will only have a discrete set of orbits characterized mainly by the rank of the reduced density operators. It is through the study of the rank of the reduced density operators that Dür, Vidal and Cirac [DVC] classified the orbits under stochastic LOCC. We will now review and comment upon their results.

6.5.1 Rank distribution

We basically have three possible scenarios: 1) All reduced density operators have rank one. 2) Two reduced density operators have rank two and the remaining rank one. 3) All three reduced density operators have rank two. The first case

[1] Here we state the relations in [Sud] but with I_2 and I_4 interchanged. In [Sud] the relation $4\det(\rho_A) = 2\left(I_1^2 - I_2\right)$ was used in the computation whereas $4\det(\rho_A) = 2\left(I_1^2 - I_4\right)$ is the correct one.

is rather trivial and corresponds to a completely unentangled situation. The second case corresponds to the presence of entanglement between two systems whereas the third is not entangled. In the third case all three systems share some kind of entanglement, but interestingly enough this genuine tripartite entanglement can be constructed in two inequivalent ways. The canonical form of these two inequivalent states are

$$|GHZ\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right)$$

$$|W\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right)$$

Since the number of product terms are inequal (namely two and three), these two states can not be transformed into one another by any invertible linear transformation. Furthermore, it was shown in [DVC] that any state with maximal local rank of two can be converted to either $|GHZ\rangle$ or $|W\rangle$ through a invertible linear transformation. We thus have the following canonical forms for the orbits under stochastic LOCC:

$$|\psi_{A-B-C}\rangle = |000\rangle \tag{36}$$

$$|\psi_{A-BC}\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |011\rangle\right) \tag{37}$$

$$|\psi_{B-AC}\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |101\rangle\right) \tag{38}$$

$$|\psi_{C-AB}\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |110\rangle\right) \tag{39}$$

$$|W\rangle = \frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right) \tag{40}$$

$$|GHZ\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right) \tag{41}$$

How do we distinguish between these states in terms of the coefficients? In this scenario the tangle comes in handy. All six states can be distinguished by observing whether the tangles vanish or not. To see this, we first look at the values of the polynomial invariants (25)-(30). They are presented in Table 1. Notice that all the states are normalized, so I_1 is therefore equal to unity.

Class	I_1	I_2	I_3	I_4	I_5	I_6
$\|\psi_{A-B-C}\rangle$	1	1	1	1	1	0
$\|\psi_{A-BC}\rangle$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{8}$	0
$\|\psi_{B-AC}\rangle$	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{8}$	0
$\|\psi_{C-AB}\rangle$	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$	0
$\|W\rangle$	1	$\frac{5}{9}$	$\frac{5}{9}$	$\frac{5}{9}$	$\frac{2}{9}$	0
$\|GHZ\rangle$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	1

Table 1: Values of polynomial invariants for the canonical forms of tripartite qubit system.

Using the values of the polynomial invariants we can calculate the different tangles through formulae (32)-(35). We have the following table:

Class	τ_{AB}	τ_{AC}	τ_{BC}	τ_{ABC}
$\|\psi_{A-B-C}\rangle$	0	0	0	0
$\|\psi_{A-BC}\rangle$	0	0	x	0
$\|\psi_{B-AC}\rangle$	0	x	0	0
$\|\psi_{C-AB}\rangle$	x	0	0	0
$\|W\rangle$	x	x	x	0
$\|GHZ\rangle$	0	0	0	x

Table 2: The tangle is used to distinguish the orbits under stochastic LOCC. An x means that the value is non-vanishing.

We thus see that the hyperdeterminant enters as arbiter between the $|GHZ\rangle$ state and the other states through the 3-tangle. The tangle, which is a polynomial expression in the coefficients, can thus be used to clearly distinguish between the different canonical forms. It singles out the relevant physical properties of each form, namely the type of entanglement the form has both locally between different subsystems and globally.

If we return to our arrangement of the coefficients a_{ijk} as a cube we can get a geometrical interpretation of the canonical forms on the cube. When we have a tensor product $A \otimes B$ it is isomorphic to $B \otimes A$. For our tripartite qubit system this means that $\mathbb{C}_1^2 \otimes \mathbb{C}_2^2 \otimes \mathbb{C}_3^2$ is isomorphic to $\mathbb{C}_{\sigma(1)}^2 \otimes \mathbb{C}_{\sigma(2)}^2 \otimes \mathbb{C}_{\sigma(3)}^2$, where $\sigma \in S(3)$ (the group of permutations on 3 letters). So when we place our coefficients on the cube we want the terms in (31) to retain their geometrical interpretation as lines, planes and tetrahedrons under the permutations $a_{ijk} \mapsto a_{\sigma(ijk)}$. But the labeling we used in 6.4.3 preserves this aspect. Geometrically, then, we have the following interpretation:

The completely unentangled state (36) corresponds to corners of the cube.

The partially entangled states (37), (38) and (39) correspond to diagonal lines in the same plane.

The state (41) corresponds to diagonal lines across the cube (with center of mass in the center of the cube).

Finally the state (40) corresponds to a triangle inscribed into the cube.

6.5.2 The case of special linear transformations and stochastic LOCC

We now wish to comment on an issue raised by a statement in [DVC]. When operating on the unnormalized states in

$$\mathcal{H}^{(n)} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2,$$

where the tensor product enters n times, Dür et. al. want to identify vectors related through an invertible linear transformation in $GL(2) \otimes GL(2) \otimes \cdots \otimes GL(2)$. They argue that the determinant of each local party operator of this transformation can be fixed to one, $\det(A) = 1$, since the only difference between A and kA is that the latter will introduce an extra global constant to our transformed state. Since we use an unnormalized state this should not make any difference and we can therefore restrict our attention to $SL(2) \otimes SL(2) \otimes \cdots \otimes SL(2)$. The equivalence classes that we are interested in are thus the quotient space

$$\frac{\mathcal{H}^{(n)}}{SL(2) \otimes SL(2) \otimes \cdots \otimes SL(2)}.$$

But it is a well known result that the hyperdeterminant remains invariant under the action of $SL(2) \otimes SL(2) \otimes SL(2)$. If we look at the state

$$|\psi\rangle = \frac{1}{\sqrt{3}}|000\rangle + \sqrt{\frac{2}{3}}|111\rangle$$

we see by (31) that

$$\mathrm{Hdet}(|\psi\rangle) = \frac{2}{9}$$

whereas

$$\mathrm{Hdet}(|GHZ\rangle) = \frac{1}{4}.$$

So there is no operator in $SL(2) \otimes SL(2) \otimes \cdots \otimes SL(2)$ that can transform $|\psi\rangle$ into $|GHZ\rangle$ directly. However, if we say that two states are equivalent up to a global complex constant, which tacitly was the assumption when operating on $\mathcal{H}^{(n)}$, we can introduce the extra global constant $\left(\frac{9}{8}\right)^{\frac{1}{4}}$ to $|\psi\rangle$ which will alter the determinant to $\frac{1}{4}$. In this case we may say that $|\psi\rangle$ and $|GHZ\rangle$ are equivalent under the unimodular transformation.

7 Conclusions

We established a mathematical framework for the operation on quantum states. We saw that we could use this framework to understand mathematically what we do in the laboratory in terms of Local Operations and Classical Communications (LOCC). We found out what it means mathematically that two states can be transformed into one another with either unit or non-zero probability, namely that these two cases correspond to local unitary or general transformations. We then asked under what conditions such transformations exist given two states. This made us shift focus and look at a similar problem, namely when two polynomial functions can be transformed into one another. Reviewing classical Invariant Theory we found that we could its results and concepts. Returning back with these new weapons after a short raid into the mathematical armory we laid siege to the bipartite and tripartite qubit states to fully explore their equivalence properties.

After having seen that the bipartite and tripartite qubit states are indeed accessible to a full description, one might ponder if this is possible for more complex systems? Not surprisingly the answer is no. The polynomial invariants for four qubits (that is, for $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$) were constructed by Luque, Thibon and Toumazet but they also report that "in practice, obtaining a description of the algebra in terms of generators and syzygies seems to be out of reach for more than four qubits". The computational complexity increases quickly as the systems become bigger.

But an even more important issue than performing actual computation might be to find a relevant entanglement measure. Many different measures have been proposed. They all have to be constant on the orbits of LOCC (that is be unaffected by local unitary transformations), but they should single out important properties of stochastic LOCC. So far no agreement has been found on a canonical measure. But investigations into the invariant structures might provide insights on how such a measure should be constructed.

8 Bibliography

[BPRST] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin and A. V. Thapliyal, *Exact and Asymptotic Measures of Multipartite Pure State Entanglement*, Phys. Rev. A 63, 012307 (2000).

[BZ] I. Bengtsson, K. Zyczkowski, *Geometry of Quantum States*, Cambridge University Press, 2006.

[Ca] C. M. Caves, *Quantum Error Correction and Reversible Operations*, quant-ph/9811082v1.

[CKW] V. Coffman, J. Kundu, and W. K. Wootters, *Distributed Entanglement,* Phys. Rev. A, 61 (2306), 2000.

[DVC] W. Dür, G. Vidal, and J.I. Cirac, *Three qubits can be entangled in two inequivalent ways*, Phys. Rev. A 62 062314 (2001).

[Gr] M. Grassl, *Entanglement and invariant theory*, talk available at `iaks-www.ira.uka.de/home/grassl/paper/MSRI_InvarTheory.pdf`.

[Hi] D. Hilbert, *The Theory of Algebraic Invariants*, Cambridge Univ. Press, New York, 1993.

[LTT] J.-G. Luque, J.-Y. Thibon and F. Toumazet, *Unitary invariants of qubit systems*, Mathematical Structures in Computer Science (2007), 17:1133-1151.

[Ke] J. Kempe, *Multiparticle entanglement and its applications to cryptography*, Phys. Rev. A 60, 910–916 (1999).

[Kr1] K. Kraus, *States, Effects, and Operations*, Springer-Verlag, Berlin, 1983.

[Kr2] K. E. Hellwig and K.Kraus, *Pure Operations and Measurements*, Commun. math. Phys. 11, 214-220 (1969)

[Kr3] K. E. Hellwig and K. Kraus, *Operations and Measurements II*, Commun. math. Phys. 16, 142-147 (1970)

[Olv] P. J. Olver, *Classical Invariant Theory*, Cambridge Univ. Press, 1999.

[Sch] B. Schumacher, *Sending entanglement through noisy quantum channels*, Phys. Rev. A, 54 4 (1996).

[Sud] A. Sudbery, *On local invariants of pure three-qubit states*, J. Phys. A, Math. Gen. 34 643-652.

[Tu] Turnbull, *The Theory of Determinants, Matrices and Invariants*, Blackie & Son, Glasgow, 1945.