SEMINAR PAPER
UNGLEICHUNGEN UND ÄHNLICH VERWIRRENDE KONZEPTE

MAXI BRANDSTETTER, ARNE HEIMENDAHL, FELIX KIRSCHNER

Abstract. Dieses Paper ist eine Ausarbeitung unseres Vortrags im Seminar "Introduction to Quantum Information and Quantum Computing", das zwischen dem 19.9. und 21.9.2018 in Köln stattgefunden hat. Es wird eine kleine Einführung in die mathematischen Methoden gegeben, mit denen die Welt der Quantenfunktion beschrieben werden kann, woraufhin wir "Nonlocal Games" einführen, die Brücke zur (semidefiniten) Optimierung schlagen und hoffentlich noch genug Zeit für die Grothendieck Ungleichungen haben. Die Grothendieck Ungleichungen finden erstaunlicherweise in einer Vielzahl an mathematischen Teilgebieten Verwendung.

Contents

1. **Einleitung**

Einleitung. Es geht um Grothendieck Ungleichung und so weiter.

2. Quantum Grundlagen

2.1. *Basic definitions*

In order to make sure everyone is on the same page we will have the following introduction where the necessary groundwork is done. The aim is to call a few basic definitions back to our memory so one can fluently read through this paper.
A complex matrix $A \in \mathbb{C}^{n\times n}$ is called Hermitian if $A^* = A$, where A^* denotes the conjugate transpose of A. A complex Hermitian matrix A is called positve semidefinite (abbreviated psd.) if one of the following holds:

i, The matrix has only real non-negative eigenvalues
ii, There exist complex $n-$dimensional vectors $z_1, \ldots, z_n$ s.t. $A_{i,j} = \langle z_i, z_j\rangle = \sum_{k=1}^{n} \overline{z_{i_k}} z_{j_k}$
iii, For every $z \in \mathbb{C}^n$ we have $z^* A z \geq 0$
iv, There exists a complex matrix B s.t. $A = B^* B$

It can be shown that i, - iv, are in fact equivalent. The set of positive semidefinite matrices is a cone, meaning for two psd. $n \times n$ matrices A, B and $\alpha, \beta \in \mathbb{R}_+$ we have that $\alpha A + \beta B$ is also positive semidefinite. The set of real $n \times n$ matrices is denoted by $\mathcal{S}_n^+$.

2.2. *Tensor products and Dirac notation*

In case someone is not familiar with the *tensor product* a short introduction with an example or two is given here: Let $\mathcal{X} = \mathbb{C}^{n_1\times m_1}$ and $\mathcal{Y} = \mathbb{C}^{n_2\times m_2}$. Then the tensor product of the vector spaces $\mathcal{X}$ and $\mathcal{Y}$ is defined as $\mathcal{X} \otimes \mathcal{Y} = \mathbb{C}^{n_1 n_2 \times m_1 m_2}$. The tensor product of complex matrices can be obtained as follows: Index the rows and columns of a matrix by $\mathcal{R}$ and $\mathcal{C}$ and think of the matrix as a map from $\mathcal{R} \times \mathcal{C} \to \mathbb{C}$. For two complex matrices $A : \mathcal{R}_1 \times \mathcal{C}_1 \to \mathbb{C}$ and $B : \mathcal{R}_2 \times \mathcal{C}_2 \to \mathbb{C}$ their tensor product is the matrix $A \otimes B : (\mathcal{R}_1 \times \mathcal{R}_2) \times (\mathcal{C}_1 \times \mathcal{C}_2) \to \mathbb{C}$ defined by $(A \otimes B)((r_1, r_2), (c_1, c_2)) = A(r_1, c_1)B(r_2, c_2)$. Considering a lexicographic order understand tensor products in the following way: $A \otimes B = \begin{pmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \ldots & a_{nn}B \end{pmatrix}$, which is the *Kronecker product* and coherent with the definition. For two complex vectors v_1, v_2 when we talk about their tensor products we will mean $v_1 \otimes v_2 = (v_{1_1} v_2, v_{1_2} v_2, \ldots, v_{1_n} v_2)^\top$ from which we can deduce that $\langle x_1 \otimes x_2, y_1 \otimes y_2\rangle = \langle x_1, x_2\rangle\langle y_1, y_2\rangle$. Also for any matrices A, B, C, D (assuming fitting dimensions) we have the following identities:

i, $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
ii, $A \otimes (B + C) = A \otimes B + A \otimes C$
iii, $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$

Also throughout this paper we will stick to the *Dirac notation*, which is the standard notation for describing quantum states. In Dirac notation $|\psi\rangle$ refers to a vector in $\mathbb{C}^n$. The conjugate transpose of this vector is written $\langle\psi|$. The non-negative integers, by

convention, represent the canonical basis vectors, i.e.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, |n-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{1}$$

Usually the tensor product symbol is omitted when taking the tensor product of two vectors in Dirac notation. This means we write $|\psi\rangle|\phi\rangle$ instead of $|\psi\rangle \otimes |\phi\rangle$. We also would like to quickly remind ourselves what a Hilbert space is. Let $\mathcal{H}$ be an inner product space. Endow $\mathcal{H}$ with a norm $||x|| = \sqrt{\langle x, x\rangle}$ and a metric $d(x, y) = ||x - y||$. If every Cauchy sequence in $\mathcal{H}$ converges to an element in $\mathcal{H}$, i.e. $\mathcal{H}$ is complete, then $\mathcal{H}$ is a Hilbert space.

2.3. ***States and measurements***

Now we can define a *state*.
A state is a complex positive semidefinite matrix ρ that satisfies $\mathrm{Tr}(\rho) = 1$. The trace of a psd. matrix is equal to the sum of its eigenvalues. The spectral theorem tells us that any $n \times n$ matrix can be decomposed as $\rho = \sum_{i=1}^{n} \lambda_i |\psi_i\rangle\langle\psi_i|$ with λ_i being its eigenvalues and $|\psi_i\rangle$ the corresponding eigenvectors. We call a state *pure* if it has rank 1, i.e.$\rho = |\psi\rangle\langle\psi|$ for some complex unit vector $|\psi\rangle$. This means every state is a convex combination of pure states. Note that complex unit vectors are often referred to as states even though states are defined as matrices. What is actually meant is the pure state $|\psi\rangle\langle\psi|$. The name state for these mathematical objects is chosen because with them the possible configurations of a quantum system can be modeled. A quantum system X is said to be *in* state ρ and is associated with a positive integer n, referred to as its dimension and a copy of $\mathbb{C}^n$. The states in $\mathbb{C}^{n\times n}$ give the possible configurations of X. A quantum system X may consist subsystems $X_1, \ldots, X_N$, where each subsystem X_i is a quantum system for itself. And X is then associated with $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_N}$ with n_i being the dimensions of the subsystems. A state ρ in which X is then is a matrix of size $n_1 \ldots n_N$.
What physicists usually do is building some mathematical model that is supposed to describe how the universe behaves and then test this model by performing experiments and comparing the results to what the model predicts. We will now define an experiment, a *measurement* of a quantum state, in a mathematical way. It shall be stated that the measurements we are talking about do not compare to a physical measurement like measuring the temperature, atmospheric pressure or any other continuous physical quantity. The outcome of measuring a temperature in Kelvin may be a real number $T \in [0, \infty)$ but we are going to assume the measurements we consider only have a finite set of outcomes $\mathcal{A}$. As before we will have an n-dimensional quantum system X and a measurement on X in state ρ with outcomes in $\mathcal{A}$ is defined by a set of psd. matrices $\{F^a\}_{a\in\mathcal{A}} \subseteq \mathbb{C}^{n\times n}$ that sum up to the identity matrix, i.e. $\sum_{a\in\mathcal{A}} F^a = I$. A *projective* measurement is defined by matrices that satisfy $F^a F^b = \delta_{ab} F^a$ for all $a, b \in \mathcal{A}$. The outcome of a measurement is a random variable χ and its probability distribution is given by $\mathbb{P}[\chi = a] = \mathrm{Tr}(\rho F^a)$. In order to be able to define an expected value for an projective measurement it is convenient to define the outcomes in $\mathcal{A}$ as real numbers.

In that case we have:

$$\mathbb{E}[\chi] = \sum_{a\in\mathcal{A}} a\mathrm{Tr}(\rho F^a) = \mathrm{Tr}(\rho(\sum_{a\in\mathcal{A}} aF^a)) \tag{2}$$

The sum of the matrices times their outcome value is called an *observable* associated to a projective measurement. A very simple case of this would be a $\{-1,1\}$-valued observable, where $\{-1,1\}$ is the set of outcomes. Such an observable is defined as $\sum_{a\in\{-1,1\}} aF^a = (-1)F^- + (1)F^+ = F^+ - F^-$. Since we are considering projective measurements, squaring the difference yields

$$(F^+ - F^-)^2 = \underbrace{F^{+^2}}_{=F^+} - \underbrace{F^+F^-}_{\delta_{+-}=0} + \underbrace{F^{-^2}}_{F^-} = F^+ + F^- = I \tag{3}$$

i.e. a $\{-1,1\}$-valued observable is both unitary and Hermitian.
Now let us take a quantum system X consisting of subsystems $X_1,\ldots,X_N$ an distribute the subsystems among N parties. If X is in state ρ we say that state ρ is shared by the parties. The parties may have an arbitrary distance to each other, i.e. may be located anywhere in the universe. Every party can perform a measurement on their subsystem. This means there are N sets of psd. matrices $\{F^{a_1}\}_{a_1\in\mathcal{A}_1} \in \mathbb{C}^{n_1\times n_1},\ldots,\{F^{a_N}\}_{a_N\in\mathcal{A}_N} \in \mathbb{C}^{n_N\times n_N}$ and the joint probability distribution of the N measurement outcomes $\chi_1,\ldots,\chi_N$ is

$$\mathbb{P}[\chi_1 = a_1, \chi_2 = a_2, \ldots, \chi_N = a_N] = \mathrm{Tr}(\rho F_1^{a_1}\otimes\cdots\otimes F_N^{a_N}) \tag{4}$$

As we defined earlier a pure state ρ has rank 1 and in the case that the system X we have consists of subsystems there is a $|\psi\rangle \in \mathbb{C}^{n_1}\otimes\cdots\otimes\mathbb{C}^{n_N}$ such that $\rho = |\psi\rangle\langle\psi|$. The state is called *product state* if it is of the form $|\psi\rangle = |\psi_1\rangle|\psi_2\rangle\ldots|\psi_N\rangle$. A state that is not a product state is said to be entangled. A mixed state, i.e. a state with rank greater than 1 is said to be separable if it is a convex combination of pure states. The interesting thing and what makes quantum mechanics interesting is that entangled states can give correlated measurement outcomes. What makes this especially mind-boggling is the fact that the parties can be located anywhere in the universe. This means that the information of a measurement can travel at an instant.
In the following example we would like to show that if two players, call them Alice and Bob, share a product state, the result is in fact a product distribution, i.e. the measurement outcome do not correlate. So, let $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$ and let Alice perform a measurement $\{F^a\}_{a\in\mathcal{A}}$ on her $|\psi_A\rangle$ and let Bob perform a measurement $\{G^b\}_{b\in\mathcal{B}}$ on his $|\psi_B\rangle$. The probability of Alice getting measurement outcome $\chi_A = a$ and Bob getting $\chi_B = b$ is equal to:

$$\begin{aligned}
Tr(|\psi\rangle\langle\psi|F^a\otimes G^b) &= \langle\psi|F^a\otimes G^b|\psi\rangle \\
&= (\langle\psi_A|\otimes\langle\psi_B|)(F^a\otimes G^b)(|\psi_A\rangle\otimes|\psi_B\rangle) \\
&= ((\langle\psi_A|F^a)\otimes(\langle\psi_B|G^b))(|\psi_A\rangle\otimes|\psi_B\rangle) \\
&= \langle\psi_A|F^a|\psi_A\rangle\otimes\langle\psi_B|G^b|\psi_B\rangle \\
&= \langle\psi_A|F^a|\psi_A\rangle\langle\psi_B|G^b|\psi_B\rangle
\end{aligned}$$

Where the third and fourth equality follow from the fact that $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$ and that the tensor product of two real numbers is equal to their ordinary product. The result is just the product of the probability of Alice measuring a and Bob measuring b, as desired.

3. Nonlocal Games

3.1. *Local and Quantum correlation matrices*

Definition 3.1.1. Let $(X_i)_{1\leq i\leq m}$ and $(Y_j)_{1\leq j\leq n}$ be families of random variables on a common probability space such that $|X_i|, |Y_j| \leq 1$ almost surely **define the norm**. Then $A = (a_{ij})$ is the corresponding *classical (or local) correlation matrix* if

$$a_{ij} = \mathbb{E}[X_i Y_j]$$

for all $1 \leq i \leq m, 1 \leq j \leq n$.

As we will see in the sequel, the set of $m \times n$ correlation matrices is a polytope, denoted by $\mathrm{LC}_{m,n}$.

Definition 3.1.2. Let $(X_i)_{1\leq i\leq m}$ and $(Y_j)_{1\leq j\leq n}$ be self-adjoint operators on $\mathbb{C}^{d_1}$, respectively $\mathbb{C}^{d_2}$ for some positive integers d_1, d_2, satisfying $|X_i|, |Y_j| \leq 1$. $A = (a_{ij})$ is called *quantum correlation matrix* $\rho \in D(\mathbb{C} \otimes \mathbb{C})$ such that

$$a_{ij} = \mathrm{Tr}\, \rho(X_i \otimes Y_j).$$

We will write $\mathrm{QC}_{m,n}$ for the set of all $m \times n$ quantum cor relation matrices. With regard to quantum information theory it is interesting to analyze the geometry of $\mathrm{LC}_{m,n}$ and $\mathrm{QC}_{m,n}$. As we will see in the following two lemmata, both sets have rather simple descriptions.

Lemma 3.1.3. *An alternative description of* $\mathrm{LC}_{m,n}$ *is given by*

$$\mathrm{LC}_{m,n} = \mathrm{conv}\{\xi\eta^T \,|\, \xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n\}. \tag{5}$$

Proof. Let us denote the right hand side of 5 by M and let $\xi\eta^T \in M$ with $\xi \in \{-1,1\}^m, \eta \in \{-1,1\}^n$. Clearly $\xi_i, \eta_j \in \{-1,1\}$ define constant $\mathbb{R}$-valued random variables satisfying $|X_i|$, $|Y_j| \leq 1$. Hence, it suffices to show that $\mathrm{LC}_{m,n}$ is convex since it contains the vertices of M. Therefore, consider to classical correlation matrices $a_{ij}^{(k)} = \mathbb{E}[X_i^{(k)} Y_j^{(k)}]$ for $k \in \{0,1\}$ which are defined on a common probability space and whose absolute value is smaller than one almost surely. We have to show that there exists random variables $(X_i), (Y_j)$ with $\|X_i\|, \|Y_j\| \leq 1$ almost surely such that

$$\beta a_{ij}^{(0)} + (1-\beta) a_{ij}^{(1)} = \mathbb{E}[X_i Y_j] \tag{6}$$

for all $\beta \in [0,1]$. Let α be a Bernoulli random variable, i.e. $\mathbb{P}(\alpha = 0) = \beta$, $\mathbb{P}(\alpha = 1) = 1 - \beta$ and set $X_i = X_i^{(\alpha)}, Y_j = Y_j^{(\alpha)}$. Then

$$\begin{aligned}\mathbb{E}[X_i Y_j] &= \mathbb{E}[X_i^{(\alpha)} Y_j^{(\alpha)} \mathbb{1}_{\{\alpha=0\}}] + \mathbb{E}[X_i^{(\alpha)} Y_j^{(1)}] \mathbb{1}_{\{\alpha=1\}}] \\ &= \beta \mathbb{E}[X_i^{(0)} Y_j^{(0)}] + (1-\beta)\mathbb{E}[X_i^{(1)} Y_j^{(1)}],\end{aligned}$$

which proofs that $\mathrm{LC}_{m,n}$ is convex.

For the other inclusion, let $(a_{ij}) \in \mathrm{LC}_{m,n}$, i.e. $a_{ij} = \mathbb{E}[X_i Y_j]$ for $\mathbb{R}$-valued random variables $(X_i), (Y_j)$, defined on a common probability space Ω with $|X_i|, |Y_j| \leq 1$ almost surely. We will use the characterization of the $d-$dimensional cube by its vertices, that is $[-1,1]^d = \mathrm{conv}\{\xi \,|\, \xi \in \{-1,1\}^d\}$ (proof by induction). So, for If we define the random variables $X = (X_1, ..., X_m)$ and $Y = (Y_1, ..., Y_m)$ they are essentially functionals

$X : \Omega^m \mapsto [-1,1]^m$ (up to a null set). Using the characterization of the hypercube we can define random variables $\lambda_\xi^{(X)} : \Omega^m \to [0,1]$ for $\Omega \in \{-1,1\}^m$ such that

$$X(\omega) = \sum_{\xi \in \{-1,1\}^m} \lambda_\xi^{(X)}(\omega)\xi$$

and $\sum_{\xi\in\{-1,1\}^m} \lambda_\xi^{(X)}(\omega) = 1$. If we proceed analogously for Y we obtain

$$\begin{aligned} a_{ij} = \mathbb{E}[X_i Y_j] &= \mathbb{E}\big[\big(\sum_{\xi\in\{-1,1\}^m} \lambda_\xi^{(X)}\xi_i\big)\big(\sum_{\eta\in\{-1,1\}^n} \lambda_\eta^{(Y)}\eta_j\big)\big] \\ &= \sum_{\xi\in\{-1,1\}^m,\eta\in\{-1,1\}^n} \mathbb{E}[\lambda_\xi^{(X)}\lambda_\eta^{(Y)}]\xi_i\eta_j \\ &= \Big(\mathbb{E}\big[\sum_{\eta\in\{-1,1\}^m} \lambda_\xi^{(X)}\big]\mathbb{E}\big[\sum_{\eta\in\{-1,1\}^n} \lambda_\eta^{(Y)}\big]\Big)\xi_i\eta_j \\ &= \xi_i\eta_j, \end{aligned}$$

where we used that $\lambda_\xi^{(X)}$ and $\lambda_\eta^{(Y)}$ are independent and sum up to one. Thus, $\{a_{ij}\} \in M$ which finishes the proof. □

Now we can easily count the vertices of $\mathrm{LC}_{m,n}$. Observing that $\xi\eta^T = \tilde{\xi}\tilde{\eta}^T$ if and only if $\xi = \tilde{\xi}$ and $\eta = \tilde{\eta}$ or $\xi = -\tilde{\xi}$ and $\eta = -\tilde{\eta}$ it follows that we have $2^{n+m}/2 = 2^{n+m-1}$ different matrices $\xi\eta$, hence $\mathrm{LC}_{m,n}$ has 2^{n+m-1} vertices. To analyze the facial structure of $\mathrm{LC}_{m,n}$ is rather complicated. However, we will do it later on for $n = m = 2$ and compare it to $\mathrm{QC}_{m,n}$.

In the following, we will proof a similar description for $\mathrm{QC}_{m,n}$ that is:

Lemma 3.1.4.

$$QC_{m,n} = \{(\langle x_i, y_j\rangle)_{1\le 1\le m, 1\le j\le n} \,|\, x_i, y_j \in \mathbb{R}^{\min\{m,n\}}, |x_i| \le 1, |y_j| \le 1\}.$$

In order to proof this we have to review some definitions and introduce a special class of matrices, namely the *Pauli matrices*. For the first inclusion we review the definition of an inner product, ...

4. Grothendieck Inequality

Maxis Teil Lorem Ipsum und so weiter

Appendix

Dinge, die definiert werden sollten.

(1) Injective tensor product
(2) norms
(3) Notation, operators of norms
(4) perhaps what a state is

Appendix alles was vorher keinen Platz findet.