

Armin Uhlmann: Geometry of State Spaces. (incomplete manuscript version)

## 1 Geometry of pure states

### 1.1 Distance in $\mathcal{H}$

Hilbert space  $\mathcal{H}$ , vectors  $\psi \equiv |\psi\rangle$ , scalar product  $\langle \cdot, \cdot \rangle$ .

$$\text{vector norm} = \|\psi\| := \sqrt{\langle \psi, \psi \rangle}$$

= Euclidian length of the vector  $0 \rightarrow \psi$ . In any basis

$$\psi = \sum z_j \psi_j, \quad z_j = x_j + iy_j$$

we get

$$\|\psi\| = (\sum x_j^2 + y_j^2)^{1/2}.$$

Therefore,

$$\|\psi_1 - \psi_0\| = \text{distance between } \psi_0 \text{ and } \psi_1$$

is the Euclidean distance in  $\mathcal{H}$ .

### 1.2 Length of curves in $\mathcal{H}$

*Length* of a connected curve

$$t \rightarrow \psi_t, \quad 0 \leq t \leq 1, \tag{1}$$

( $t$  is a parameter, not necessarily the time.)

To get the length we have to take all subdivisions

$$0 \leq t_0 < t_1 < \dots < t_n \leq 1$$

in performing the supremum

$$\text{length of the curve} = \sup \sum_{j=1}^n \|\psi_{j-1} - \psi_j\| \tag{2}$$

*The length is independent of the parameter choice.*

Iff we can guaranty

$$\dot{\psi}_t = \frac{d}{dt} \psi_t \in \mathcal{H} \tag{3}$$

then

$$\text{length of the curve} = \int_0^1 \sqrt{\langle \dot{\psi}, \dot{\psi} \rangle} dt \tag{4}$$

and the velocity with which  $\psi$  travels through  $\mathcal{H}$  reads

$$\frac{ds}{dt} = \sqrt{\langle \dot{\psi}, \dot{\psi} \rangle}. \tag{5}$$

### 1.3 Distance and length

Generally, a distance “dist” in a space attaches a real and not negative number to any pair of points satisfying

- a)  $\text{dist}(\xi_1, \xi_2) = \text{dist}(\xi_2, \xi_1)$
- b)  $\text{dist}(\xi_1, \xi_2) + \text{dist}(\xi_2, \xi_3) \geq \text{dist}(\xi_1, \xi_3)$ ,
- c)  $\text{dist}(\xi_1, \xi_2) = 0 \Rightarrow \xi_1 = \xi_2$ .

A set with a distance is a *metric space*.

Given the distance,  $\text{dist}(\cdot, \cdot)$ , of a metric space and two different points, say  $\xi_0$  and  $\xi_1$ , one may ask for the length of a continuous curve connecting these two points. The infimum of the lengths over all these curves is again a distance, the *inner distance*. The inner distance,  $\text{dist}_i(\xi_0, \xi_1)$  is never smaller than the original one,

$$\text{dist}_i(\xi_0, \xi_1) \geq \text{dist}(\xi_0, \xi_1)$$

If equality holds, the distance (and the metric space) is called *inner*. The Euclidian distance is inner. It is easy to present the shortest curves between to vectors in Hilbert space:

$$t \rightarrow \psi_t = (1 - t)\psi_0 + t\psi_1 \tag{6}$$

is a short geodesic arc between both vectors. (In Euclidean spaces the shortest connection between two points is a straight line.) The geodesic arc is unique. Note also

$$\dot{\psi} = \psi_1 - \psi_0 .$$

### 1.4 Curves on the unit sphere

Restricting geometry of  $\mathcal{H}$  to the unit sphere  $\{\psi \in \mathcal{H}, \|\psi\| = 1\}$  can be compared with the change from Euclidean geometry to spherical geometry.

Assume  $\psi_0$  and  $\psi_1$  on the unit sphere. The shortest curve on the sphere between them is a piece of a great circle.

This circle of radius 1 is unique if  $\psi_0 + \psi_1 \neq 0$ . Its length is given by the angle  $\alpha$  between the two radii terminating at the two vectors, restricted by  $0 \leq \alpha \leq \pi$ . By elementary geometry

$$\|\psi_1 - \psi_0\| = \sqrt{2 - 2\cos\alpha} = 2\sin\frac{\alpha}{2} \tag{7}$$

and the  $\cos\alpha$  can be computed by

$$\cos\alpha = \frac{\langle\psi_0, \psi_1\rangle + \langle\psi_1, \psi_0\rangle}{2} \tag{8}$$

For a general curve  $t \rightarrow \psi_t$  on the unit sphere one gets

$$\langle\psi_t, \dot{\psi}_t\rangle \text{ is purely imaginary.} \tag{9}$$

To see this one differentiates

$$0 = \frac{d}{dt} \langle \psi, \psi \rangle = \langle \dot{\psi}, \psi \rangle + \langle \psi, \dot{\psi} \rangle$$

and this is equivalent with the assertion.

## 1.5 Mandelstam-Tam-inequality

We see from (8) that

$$\cos \alpha \leq |\langle \psi_0, \psi_1 \rangle| \quad (10)$$

Therefore, we have the following statement:

*The length of A curve on the unit sphere connecting  $\psi_0$  and  $\psi_1$  is not less than*

$$\arccos |\langle \psi_0, \psi_1 \rangle| .$$

If a solution of a Schrödinger equation

$$H\psi = i\hbar\dot{\psi}, \quad \text{length} = \hbar^{-1} \sqrt{\langle \psi, H^2\psi \rangle} (t_1 - t_0) . \quad (11)$$

with Hamiltonian  $H$  connects two vectors, we get the Mandelstam-Tam equality (1946) from this observation:

$$\hbar^{-1} \sqrt{\langle \psi, H^2\psi \rangle} (t_1 - t_0) \geq \arccos |\langle \psi_0, \psi_1 \rangle| \quad (12)$$

By the arguments below one shows a little bit more:

$$\sqrt{\langle \psi, H^2\psi \rangle - \langle \psi, H\psi \rangle^2} (t_1 - t_0) \geq \hbar \arccos |\langle \psi_0, \psi_1 \rangle| \quad (13)$$

Remark that the length and the expectation values of  $H$ ,  $H^2$ , are constants of motion. We can, therefore, use  $\psi = \psi_t$  at any time to compute these expectation values.

## 1.6 Phases

If the vectors  $\psi$  and  $\psi'$  are linearly dependent, they describe the same state. From this freedom in choosing a state vector the *phase change*

$$\psi \rightarrow \epsilon\psi, \quad |\epsilon| = 1, \quad (14)$$

is of primary interest. (14) is the *natural gauge transformation* offered by  $\mathcal{H}$ . The curves

$$t \rightarrow \psi_t \quad \text{and} \quad t \rightarrow \psi'_t := \epsilon_t \psi_t \quad (15)$$

are *gauge equivalent*. The states themselves,

$$t \rightarrow \pi_t = |\psi_t\rangle \langle \psi_t| \quad (16)$$

are gauge invariant. Thus, we have

$$\begin{array}{ccc} \psi_t & \xrightarrow{\text{bundle structure}} & \pi_t \xrightarrow{\text{lifts}} \begin{array}{l} \psi_t \\ \psi'_t \\ \dots \end{array} \end{array} \quad (17)$$

From the transformation (15) we deduce for the tangents

$$\dot{\psi}' = \dot{\epsilon}\psi + \epsilon\dot{\psi}, \quad \epsilon^{-1}\dot{\epsilon} = i\gamma \quad (18)$$

with real  $\gamma$ . By an appropriate choice of the gauge one gets

$$\langle \psi', \dot{\psi}' \rangle = 0, \quad \text{the geometric phase transport condition} \quad (19)$$

(Fock, 1928, from adiabatic reasoning). Indeed, (19) is the equation

$$\langle \psi', \dot{\psi}' \rangle = i\gamma \langle \psi, \dot{\psi} \rangle + \langle \psi, \dot{\psi} \rangle = 0$$

Thus

$$\epsilon_t = \exp - \int_{t_0}^t \langle \psi, \dot{\psi} \rangle dt. \quad (20)$$

For a closed curve  $t \rightarrow \psi_t$  with  $\psi_1 = \psi_0$  the integral is the *geometric or Berry phase* if it extends from  $t_0$  to  $t_1$ .

Remark: This is true on the unit sphere. Generally one requires the vanishing of the “gauge potential”

$$\frac{\langle \psi', \dot{\psi}' \rangle - \langle \dot{\psi}', \psi' \rangle}{2i} \quad \text{or} \quad \frac{\langle \psi', \dot{\psi}' \rangle - \langle \dot{\psi}', \psi' \rangle}{2i \langle \psi, \dot{\psi} \rangle} \quad (21)$$

Appropriately formulated, the phase transport and the Berry phase do not depend on the normalization.

## 1.7 Fubini-Study-distance

Let us start with the Fubini-Study distance between two pure states

$$\text{dist}_{FS}(\pi_1, \pi_0) = \min_{\epsilon} \| \psi_1 - \epsilon\psi_0 \| \quad (22)$$

where

$$\pi_j = |\psi_j\rangle\langle\psi_j|. \quad (23)$$

One may use (22) not only with normalized vectors and positive rank one operators. Having this in mind one easily finds

$$\text{dist}_{FS}(\pi_1, \pi_0) = \sqrt{\langle \psi_0, \psi_0 \rangle + \langle \psi_1, \psi_1 \rangle - 2|\langle \psi_1, \psi_0 \rangle|}. \quad (24)$$

For the modulus of the scalar product involved we choose the old notation *overlap* (remind “overlapping integral”), shortened to “olap” for convenience:

$$\text{olap}(|\psi_1\rangle\langle\psi_1|, |\psi_0\rangle\langle\psi_0|) \equiv \text{olap}(\psi_1, \psi_0) = |\langle \psi_1, \psi_0 \rangle|,$$

while the *transition probability* reads

$$\Pr\{\psi_0 \rightarrow \psi_1\} \equiv \Pr(\pi_0, \pi_1) = \frac{\langle \psi_1, \psi_0 \rangle \langle \psi_0, \psi_1 \rangle}{\langle \psi_0, \psi_0 \rangle \langle \psi_1, \psi_1 \rangle}$$

(It seems, there is no general accepted notation.)

In particular, on the unit sphere, one gets

$$\text{dist}_{FS}(\pi_1, \pi_0) = \sqrt{2 - 2|\langle \psi_1, \psi_0 \rangle|} = \sqrt{2 - 2\sqrt{\Pr(\pi_0, \pi_1)}}. \quad (25)$$

By taking attention to (22), short geodesics between two  $\pi_0$  and  $\pi_1$  can be described as follows. We choose in (23) the phases of the vectors  $\psi_j$  in such a way that their scalar product becomes real and not negative. Then we proceed as in the Euclidean case:

$$t \rightarrow \psi_t := (1 - t)\psi_0 + t\psi_1, \quad \langle \psi_1, \psi_0 \rangle \geq 0. \quad (26)$$

Then it is obvious that

$$\text{dist}_{FS}(\pi_r, \pi_s) = \|\psi_r - \psi_t\| \quad (27)$$

holds. That is, (27) describes a geodesic

$$t \rightarrow \pi_t = (1 - t)^2\pi_0 + t^2\pi_1 + t(1 - t)(|\psi_0\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_0|). \quad (28)$$

There is no normalization here:  $\text{tr} \pi_t$  varies with  $t$ . The scalar product is real and not negative between two vectors of the curve (27). Therefore, the gauge potential (21) is zero along the curve.

## 1.8 Comparison with other norms

In this subsection

$$\pi_j = |\psi_j\rangle\langle\psi_j|, \quad \langle\psi_j, \psi_j\rangle = 1.$$

Then we get

$$\text{dist}_{FS}(\pi_1, \pi_2) = \sqrt{2}\sqrt{1 - |\langle\psi_1, \psi_2\rangle|}. \quad (29)$$

For operators the von Neumann norm reads

$$\|A\|_{vN} = \sqrt{\text{tr} A^\dagger A} \quad (30)$$

while the 1-norm (or functional norm) is defined by

$$\|A\|_1 = \text{tr} \sqrt{A^\dagger A} \quad (31)$$

$\|B - A\|$  is a distance for every norm. In case of the distance of two pure states we easily conclude

$$\|\pi_2 - \pi_1\|_{vN} = \text{dist}_{FS}(\pi_1, \pi_2) \sqrt{1 + |\langle\psi_1, \psi_2\rangle|}. \quad (32)$$

A further difference concerns the shortest curves: The geodesics with respect of a norm are always

$$t \rightarrow \pi_t = (1 - t)\pi_0 + t\pi_1$$

because then

$$\| \pi_t - \pi_r \| = |t - r| \| \pi_1 - \pi_0 \|$$

In particular, the geodesic between two pure states runs through the mixed states. In contrast, the Fubini-Study geodesic between two pure states remain completely within the pure states.

Another observation: There is a Riemann metric belonging to the von Neumann norm, It is an Euclidean metric as seen by writing (30) in terms of the matrix entries with respect to any basis. The 1-norm, in contrast, does not belong to a Riemann metric.

We shall now see, that the Fubini-Study distance does.

## 1.9 Fubini-Study metric

We normalize the curve

$$t \rightarrow \psi_t := (1 - t)\psi_0 + t\psi_1, \quad \langle \psi_1, \psi_0 \rangle \geq 0.$$

to get

$$t \rightarrow \varphi_t = \| \psi_t \|^{-1} \psi_t, \quad (33)$$

$$t \rightarrow \rho_t = |\varphi_t\rangle\langle\varphi_t|.$$

The metric is obtained by calculating

$$\frac{1}{2} \text{tr} \left( \frac{d}{dt} \rho_t \right)^2 = \frac{\langle \dot{\psi}, \dot{\psi} \rangle \langle \psi, \psi \rangle - \langle \psi, \dot{\psi} \rangle \langle \dot{\psi}, \psi \rangle}{\langle \psi, \psi \rangle^2} \quad (34)$$

The Fubini-Study metric becomes simpler by going to normalized vectors (33)

$$\frac{1}{2} \text{tr} (\dot{\rho})^2 = \langle \dot{\varphi}, \dot{\varphi} \rangle - \langle \varphi, \dot{\varphi} \rangle \langle \dot{\varphi}, \varphi \rangle \quad (35)$$

and particular simple by requiring the transport condition for the geometric (Berry) phase:

$$\langle \dot{\varphi}, \varphi \rangle = 0 \Rightarrow \frac{1}{2} \text{tr} (\dot{\rho})^2 = \langle \dot{\varphi}, \dot{\varphi} \rangle. \quad (36)$$

Because  $\rho^2 = \rho$  it is  $\dot{\rho} = \dot{\rho}\rho + \rho\dot{\rho}$ . Hence

$$\frac{1}{2} \text{tr} (\dot{\rho})^2 = \text{tr} \rho \dot{\rho}^2. \quad (37)$$

## 1.10 Once more: Tam-Mandelstam

Along the unit sphere  $\|\psi\|=1$  we have

$$ds_{FS} = \sqrt{\langle \dot{\psi}, \dot{\psi} \rangle - \langle \psi, \dot{\psi} \rangle \langle \dot{\psi}, \psi \rangle} dt \quad (38)$$

Inserting a solution of a Schrödinger equation

$$H\psi = i\hbar\dot{\psi}$$

we get

$$ds_{FS} = \Delta H dt, \quad \Delta H = \sqrt{\langle \psi, H^2\psi \rangle - \langle \psi, H\psi \rangle^2} \quad (39)$$

One may consider  $\hbar^{-1}\Delta H$  the velocity with which the states run through the space of pure states. Comparing again with the shortest possible line within this space one gets

$$(t_1 - t_0)\Delta H \geq \hbar \arccos |\langle \psi_1, \psi_0 \rangle| \quad (40)$$

## 2 Density operators, states, partial traces

### 2.1 Density operators

Let us fix some notions. Let  $\mathcal{H}$  be an Hilbert space. With  $\mathcal{B}(\mathcal{H})$  we denote the set of all bounded operators.

If  $\dim \mathcal{H} < \infty$ , every linear operator  $A$  is bounded. To control it in general one introduces the *operator norm*

$$\|A\| = \sup_{\psi} \|A\psi\|, \quad \|\psi\|=1 \quad (41)$$

and calls  $A$  bounded if this supremum over all unit vectors is finite. Thus boundedness means that the operator cannot stretch unit vectors to arbitrary length. The operator norm of every unitary operator and of every projection operator (different from the operator  $\mathbf{0}$ ) is one.

The operator norm just introduced is also denoted by  $\|A\|_\infty$  and also called the “infinity norm”.

Because of a special property the operator norm is also said to be a  $C^*$ -norm. This refers to the relations

$$\|A^\dagger A\| = \|A\|^2, \quad \|A^\dagger\| = \|A\| \quad (42)$$

In mathematics and in mathematical physics the operation  $A \rightarrow A^\dagger$  is called “the star operation”: In these branches of science the Hermitian adjoint of an operator  $A$  is called  $A^*$ , following in that Hermite. The notion  $A^\dagger$  has been used by Dirac in his famous book “The principles of Quantum Mechanics”.

Let us come now to the density operators. Density operators describe states. We shall indicate that by using small Greek letters for them. Density operators are positive operators with norm one:

$$\omega \geq \mathbf{0}, \quad \text{tr } \omega = 1 . \quad (43)$$

*Heuristics:* “ $\omega$  is a non-commutative probability measure.”

## 2.2 Describing states by expectation values

Let  $\omega$  be a density operator and  $A \in \mathcal{B}(\mathcal{H})$  an operator. The value  $\text{tr } A\omega$  is called “expectation value of  $A$  in state  $\omega$ ”. In this sense one may say: “Observables distinguish states”.

((Observables should have spectral decompositions. An Observable is, therefore, represented by a normal Operator, i.e.  $A^\dagger A = AA^\dagger$  should be valid. In textbooks often the stronger hermiticity is required for historical but not physical reasons. On the other hand, the expectation values of projection operators are sufficient to distinguish states.))

As observables (or operators) distinguish states, more observables allow for a finer description, i.e. they allow to discriminate between more states. To use less observables is like “coarse graining”: Some states cannot be distinguished any more.

These dumb rules will be condensed in a precise scheme later on. The first step in this direction is describing states in a different way, namely, as indicated above, as the set of expectation values.

The function

$$A \mapsto \underline{\omega}(A) := \text{tr } A\omega \quad (44)$$

enjoys the following properties:

- 1) Linearity:  $\underline{\omega}(c_1 A_1 + c_2 A_2) = c_1 \underline{\omega}(A_1) + c_2 \underline{\omega}(A_2)$
- 2) Positivity:  $\underline{\omega}(A) \geq 0$  if  $A \geq \mathbf{0}$
- 3) It is normalized:  $\underline{\omega}(\mathbf{1}) = 1$ .

At this point one inverts the reasoning. One considers 1) to 3) as conditions and calls *every* functional on  $\mathcal{B}(\mathcal{H})$  which fulfils these three conditions a *state of the algebra*  $\mathcal{B}(\mathcal{H})$ . With other words, 1) to 3) is the definition of the term “state of  $\mathcal{B}(\mathcal{H})$ ”. The definition does not discriminate between pure and mixed states from the beginning!

If  $\dim \mathcal{H} < \infty$ , every functional obeying 1), 2), and 3) can be written

$$\underline{\omega}(A) = \text{tr } A\omega, \quad \omega \geq \mathbf{0}, \quad \text{tr } \omega = 1$$

as in (44). To see this one at first remarks, that every linear form can be written  $\underline{\omega}(A) = \text{tr } BA$  with an operator  $B \in \mathcal{B}(\mathcal{H})$ . However, if  $\text{tr } BA$  is a real and not negative number for every  $A \geq \mathbf{0}$ , one infers  $B \geq \mathbf{0}$ . (Take the trace with a basis of eigenvectors of  $B$  to see it.) Finally, condition 3) forces  $B$  to have trace one. Now one can rewrite  $\omega := B$ .  $\diamond$

The case  $\dim \mathcal{H} = \infty$  is more intriguing. A measure in “classical” mathematical measure theory has to respect the condition of countable additivity. The translation to the non-commutative case needs the so-called *partitions of the unity*, i.e. decompositions

$$\mathbf{1} = \sum_j P_j, \quad P_k P_l = \mathbf{0} \text{ if } k \neq l \quad (45)$$

with projection operators  $P_j$ . These decompositions are in one-to-one relation to decompositions of the Hilbert space into an orthogonal sum,

$$\mathcal{H} = \bigoplus_j \mathcal{H}_j, \quad \mathcal{H}_j = P_j \mathcal{H}. \quad (46)$$

A state  $\underline{\omega}$  is called *normal* if for all decompositions of  $\mathbf{1}$

$$\underline{\omega}(\mathbf{1}) = \sum_j \underline{\omega}(P_j) \quad (47)$$

is valid. ((For a state  $\underline{\omega}(\mathbf{1}) = 1$ ). But to define normality the condition “3” is not really important.))

Exactly if  $\underline{\omega}$  is normal its expectation values are given as in (44) by the help of a density operator  $\omega$ .

There is a further class of states, the *singular states*. A state  $\underline{\omega}$  of  $\mathcal{B}(\mathcal{H})$  is called “singular”, iff  $\underline{\omega}(P) = 0$  for all projection operators of finite rank. Thus, if  $\dim P\mathcal{H} < \infty$ , one gets  $\underline{\omega}(P) = 0$  for singular states.

There is a theorem asserting that every state  $\underline{\omega}$  of  $\mathcal{B}(\mathcal{H})$  has a unique decomposition

$$\underline{\omega} = (1 - p)\underline{\omega}_{\text{normal}} + p\underline{\omega}_{\text{singular}} \quad (48)$$

Remark: In mathematical measure theory this correspond to an “additive measure”, in contrast to the genuine measures which are countably additive.

I cannot but at this point of my lecture to mention the theorem of Gleason, published in J.Math. and Mechanics **6**, 885-893 (1957) entitled “Measures on the closed subspaces of a Hilbert-space. (Though I will not make use of it later on.)

Assume  $f = f(P) \geq 0$  is a function which is defined only on the projection operators  $P \in \mathcal{B}(\mathcal{H})$  and which satisfies

$$f(\mathbf{1}) = \sum_j f(P_j) \quad (49)$$

for all partitions (45) of the unity  $\mathbf{1}$ . Then Gleason proved that there is a normal state  $\underline{\omega}$  coinciding on the projections with  $f$ , i.e.  $\underline{\omega}(P) = f(P)$  for all  $P \in \mathcal{B}(\mathcal{H})$ .

The only restriction of his theorem concerns the Hilbert spaces of dimension two. If  $\dim \mathcal{H} = 2$ , the set of projection is too poor of relations and

Gleason's theorem does not apply to the states of  $\mathcal{B}(\mathcal{H})$  in this particular case. One calls the operator algebra acting on an one-qubit Hilbert space a "spin factor". (( "spin" for an obvious reason, "factor" because  $\mathcal{B}(\mathcal{H})$  acts irreducibly on  $\mathcal{H}$ .)

It lasts about 30 years to find what is with general states. It is a lengthy proof with a lot of not particular difficult steps but with a rich architecture. The proof is in Rev. Math. Phys., **1**, 235-290 (1990), written by Maeda and is entitled "Probability measures on Projections in von Neumann algebras".

In the case at hand it asserts the following: Assume  $\dim \mathcal{H} = \infty$ . Given a function  $f \geq 0$  on the projection operators fulfilling (49) for all *finite* decompositions of  $\mathbf{1}$ . Then there is a state  $\underline{\omega}$  satisfying  $\underline{\omega}(P) = f(P)$  for all projection operators of  $\mathcal{B}(\mathcal{H})$ .

### 2.3 Subalgebras

There is a consistent solution to the question: What is a subsystem of a quantum system with Hilbert space  $\mathcal{H}$  and algebra  $\mathcal{B}(\mathcal{H})$ . As already said, a subsystem should consist of less operators (observables) than the larger system  $\mathcal{B}(\mathcal{H})$ .

Let  $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$  a subset. We require

- a)  $\mathcal{A}$  is a linear space.
- b) If  $A \in \mathcal{A}$  then  $A^\dagger \in \mathcal{A}$ .
- c)  $\mathbf{1} \in \mathcal{A}$ . ( $\mathbf{1}$  is the identity map of  $\mathcal{H}$ .)
- d) If  $A, B \in \mathcal{A}$  then  $AB \in \mathcal{A}$
- e) Something more if  $\dim \mathcal{H} = \infty$ .

If these requirements are fulfilled, we say that  $\mathcal{A}$  represents (or "is") a *subsystem* of  $\mathcal{B}(\mathcal{H})$ . As a mathematical object a) to d) is called a *unital \*-subalgebra of  $\mathcal{B}(\mathcal{H})$* .

The next two remarks concern point e) above.

Remark 1:  $\mathcal{A}$  is a  $C^*$ -algebra, if it is closed with respect to the operator norm: For every sequence  $A_j \in \mathcal{A}$  which converges to  $A \in \mathcal{B}(\mathcal{H})$  in norm,  $\|A - A_j\|_\infty \rightarrow 0$ , the operator  $A$  must be in  $\mathcal{A}$  also.

Remark 2:  $\mathcal{A}$  is called a *von Neumann algebra*, if it is closed with respect to the so-called weak topology: Let be  $\mathcal{F}$  a set of operators from  $\mathcal{A}$  and  $A \in \mathcal{B}(\mathcal{H})$ . Then  $A$  should be in  $\mathcal{A}$  if for every  $n$  and for every finite set  $\psi_1, \dots, \psi_n$  of vectors from  $\mathcal{H}$  there is a  $B \in \mathcal{F}$  fulfilling the inequality

$$\sum_{j=1}^n |\langle \psi_j, (A - B)\psi_j \rangle| \leq 1/n$$

Then  $A$  is in the "weak closure" of  $\mathcal{F}$ .

The *commutant* of a  $*$ -subalgebra  $\mathcal{A}$  of  $\mathcal{B}(\mathcal{H})$  is the set of all  $B \in \mathcal{B}(\mathcal{H})$  commuting with all  $A$ , i.e.  $AB = BA$  for every  $A \in \mathcal{A}$ . The commutant of  $\mathcal{A}$  is denoted by  $\mathcal{A}'$ . The commutant of a  $*$ -subalgebra is always a von

Neumann algebra. Von Neumann has shown that  $\mathcal{A}$  is a von Neumann algebra if and only if  $\mathcal{A}'' = \mathcal{A}$ . ( $\mathcal{A}''$  is the commutant of the commutant  $\mathcal{A}'$ .)

Assume the algebras  $\mathcal{A}_1$  and  $\mathcal{A}_2$  represent subsystems of  $\mathcal{B}(\mathcal{H})$ . Then  $\mathcal{A}_2$  is considered a subsystem of  $\mathcal{A}_1$  if  $\mathcal{A}_2 \subset \mathcal{A}_1$ .

If a von Neumann algebra  $\mathcal{A}$  acts irreducibly on  $\mathcal{H}$ , it is called a *factor*. Equivalently, a factor has a center consisting of the multiples of  $\mathbf{1}$  only. Therefore, a factor may be characterized by

$$\mathcal{A} \cap \mathcal{A}' = \mathbb{C}\mathbf{1}.$$

What is a state of  $\mathcal{A}$ ? We just mimic what has been said to be a state of  $\mathcal{B}(\mathcal{H})$ . A state of  $\mathcal{A}$  is a function

$$A \rightarrow \underline{\omega}(A), \quad A \in \mathcal{A} \tag{50}$$

such that we have

- 1) Linearity:  $\underline{\omega}(c_1 A_1 + c_2 A_2) = c_1 \underline{\omega}(A_1) + c_2 \underline{\omega}(A_2)$
- 2) Positivity:  $\underline{\omega}(A) \geq 0$  if  $A \geq \mathbf{0}$
- 3) It is normalized:  $\underline{\omega}(\mathbf{1}) = 1$ .

We do not require 1), 2), and 3) for all operators of  $\mathcal{B}(\mathcal{H})$ , but only for those in  $\mathcal{A}$ . This is the only difference. As a consequence two different states of  $\mathcal{B}(\mathcal{H})$  may “fall down” to one and the same state of the subalgebra  $\mathcal{A}$ .

Slightly more general: Let be  $\underline{\omega}_1$  a state of  $\mathcal{A}_1$  and  $\underline{\omega}_2$  a state of  $\mathcal{A}_2$  and  $\mathcal{A}_2 \subset \mathcal{A}_1$ . Then  $\underline{\omega}_2$  is the *restriction of  $\underline{\omega}_1$  onto  $\mathcal{A}_2$*  if  $\underline{\omega}_2(A_2) = \underline{\omega}_1(A_2)$  for all  $A_2 \in \mathcal{A}_2$ .

Conversely,  $\underline{\omega}_1$  is an *extension or lift* of  $\underline{\omega}_2$ . The task of extending  $\underline{\omega}_2$  to a state of a larger system is always possible, but not unique: Seen from the subsystem  $\mathcal{A}_2$ , (almost) nothing can be said about expectation values  $\underline{\omega}_1(A_1)$  if  $A_1$  is in  $\mathcal{A}_1$  but not in  $\mathcal{A}_2$ .

Remark: Wedderburn obtained a classification of all subalgebras of  $\mathcal{B}(\mathcal{H})$ ,  $\dim \mathcal{H} < \infty$ . See the hand written part for an overview about all subalgebras describing quantum subsystems.

## 2.4 Back to density operators

Here we assume  $\dim \mathcal{H} < \infty$  throughout. If  $\underline{\omega}$  is a state of  $\mathcal{B}(\mathcal{H})$  there is a density operator  $\omega$  such that

$$\underline{\omega}(A) = \text{tr } \omega A \quad \text{for all } A \in \mathcal{B}(\mathcal{H}) \tag{51}$$

In a subsystem, given by an subalgebra  $\mathcal{A}$  we can find an operator  $\omega_{\mathcal{A}}$  such that

$$\underline{\omega}(A) = \text{tr } \omega_{\mathcal{A}} A \quad \text{for all } A \in \mathcal{A} \tag{52}$$

and this procedure becomes unique by requiring

$$\omega_A \in \mathcal{A}. \quad (53)$$

With some caution (see (59) below)  $\omega_A$  may be called a *partial trace* of  $\omega$ . The map

$$\omega \rightarrow \omega_A \quad (54)$$

is a trace preserving and completely positive map, whatsoever the unital  $*$ -subalgebra is.

Remark: There are more than one trace in  $\mathcal{A}$  if  $\mathcal{A}$  is not a factor. Thus some care is needed in defining partial traces in general. But even if there is essential only one trace, there are different normalizations possible. Above we used the *induced trace* defined by  $\mathcal{H}$  (i.e. by the representation of our algebras). Another possibility is the so-called *canonical trace* which is representation independent. The canonical trace gives the value 1 to each of its minimal projection operators  $P$ . A minimal projection projects onto minimal  $\mathcal{A}$ -invariant subspaces of  $\mathcal{H}$ . One knows:  $P \in \mathcal{A}$  is minimal, if there is a state  $\underline{\pi}$  of  $\mathcal{A}$  such that

$$PAP = \underline{\pi}(A)P \text{ for all } A \in \mathcal{A} \quad (55)$$

In the same spirit a state  $\underline{\pi}$  of  $\mathcal{A}$  is called *pure* exactly if there is a projection operator in  $\mathcal{A}$  with which (55) is satisfied.

Main examples of subsystems are direct product constructions. Starting with

$$\mathcal{H} = \mathcal{H}^a \otimes \mathcal{H}^b \quad (56)$$

The algebra  $\mathcal{B}(\mathcal{H}^a)$  becomes a subalgebra of  $\mathcal{B}(\mathcal{H})$  by embedding

$$A \in \mathcal{B}(\mathcal{H}^a) \mapsto A \otimes \mathbf{1} \in \mathcal{B}(\mathcal{H}) \quad (57)$$

so that

$$\mathcal{A} := \mathcal{B}(\mathcal{H}^a) \otimes \mathbf{1} \subset \mathcal{B}(\mathcal{H}^a \otimes \mathcal{H}^b)$$

shows how the embedding is working. (This subalgebra is again a factor, hence a “subfactor”.)

Now let be  $\underline{\omega}$  a state of  $\mathcal{B}(\mathcal{H})$ ,

$$\underline{\omega}(C) = \text{tr } \omega C, \quad C \in \mathcal{B}(\mathcal{H}) \quad (58)$$

Then we get for all  $A \in \mathcal{A}$

$$\underline{\omega}(A \otimes \mathbf{1}) = \text{tr } \omega^a A, \quad \omega_A = \omega^a \otimes \frac{1}{d} \mathbf{1} \quad (59)$$

with

$$\omega^a = \text{tr}_b \omega$$

by definition of  $\text{tr}_b$ . More explicitly

$$\text{tr}_b \left( \sum A_j \otimes B_j \right) = \sum (\text{tr } B_j) A_j. \quad (60)$$

### 3 Transition probabilities

The aim is to define transition probabilities by operating in larger quantum systems. The same idea provides the extension of the Study-Fubini distance (and metric) to the Bures one. There are some quite useful tricks in handling two positive operators in general position behind.

One meets different notations in textbooks and papers. In brackets I collect the words used for the same entity: (transition probability, fidelity) and (overlap, fidelity, square root fidelity). Here I use always the first in the row for definiteness. (Though “fidelity” is much more in use!)

#### 3.1 Purification

I restrict myself to the mostly treated case (56)

$$\mathcal{H} = \mathcal{H}^a \otimes \mathcal{H}^b, \quad \omega^a \text{ a density operator of } \mathcal{B}(\mathcal{H}^a) \quad (61)$$

$\pi = |\psi\rangle\langle\psi|$  and, to shorten language, also  $\pi$  is called a *purification* of  $\omega^a$  if

$$\pi(A \otimes \mathbf{1}) = \underline{\omega}^a(A) \text{ for all } A \in \mathcal{B}(\mathcal{H}^a) \quad (62)$$

Equivalently, again for  $A \in \mathcal{B}(\mathcal{H}^a)$  it is

$$\text{tr } \omega^a A = \text{tr } \pi(A \otimes \mathbf{1}) = \langle \psi, (A \otimes \mathbf{1}) \psi \rangle \quad (63)$$

#### 3.2 Transition probability, overlap, and so on

Denote by  $\omega_1^a$  and  $\omega_2^a$  two density operators of  $\mathcal{B}(\mathcal{H}^a)$ .

The task is, to prepare  $\omega_2^a$  if the state of our system is  $\omega_1^a$ . To do so one think of purifications  $\pi_j$  of our  $\omega_j^a$  in a larger system  $\mathcal{B}(\mathcal{H}^a \otimes \mathcal{H}^b)$ . One then tests whether  $\pi_2$  is true. If the answer of the test is “yes”, then  $\pi_2$  and, hence,  $\omega_2^a$  is prepared.

The probability of success is  $|\langle \psi_1, \psi_2 \rangle|^2$ .

One now asks for optimality of the described procedure, i.e. one looks for a projective measurement in the larger system which prepares a purification of  $\omega_2^a$  with maximal probability.

This maximal possible probability for preparing  $\omega_2^a$  with given  $\omega_1^a$  is called the *transition probability* from  $\omega_1^a$  to  $\omega_2^a$  or, as this quantity is symmetric in its entries, the transition probability of the pair  $\{\omega_1^a, \omega_2^a\}$ .

The definition can be rephrased, with shortened notation, to

$$\text{Pr}(\omega_1, \omega_2) := \sup |\langle \psi_1, \psi_2 \rangle|^2, \quad \psi_1, \psi_2 \text{ all purifications of } \omega_1, \omega_2 \quad (64)$$

In the same manner we define the overlap by

$$\text{olap}(\omega_1, \omega_2) = \sup |\langle \psi_1, \psi_2 \rangle| \quad (65)$$

where  $\psi_1, \psi_2$  run through all simultaneous purifications of  $\omega_1, \omega_2$ . In (65) we do assume positivity but not the trace one condition.

One defines the *Bures distance* by

$$\text{dist}_B(\omega_1, \omega_2) = \sup \text{dist}_{FS}(\pi_1, \pi_2) = \sup \| \psi_2 - \psi_1 \| \quad (66)$$

and by (65) this comes down to

$$\text{dist}_B(\omega_1, \omega_2) = \sqrt{\text{tr}\omega_1 + \text{tr}\omega_2 - 2\text{olap}(\omega_1, \omega_2)} \quad (67)$$

which can be rewritten for two density operators

$$\text{dist}_B(\omega_1, \omega_2) = \sqrt{2 - 2\sqrt{\text{Pr}(\omega_1, \omega_2)}}, \quad \text{tr}\omega_j = 1$$

If only curves entirely within the density operators are allowed in optimizing for the shortest path, we get a further variant of the Bures distance, namely

$$\text{Dist}_B(\omega_1, \omega_2) = \arccos \sqrt{\text{Pr}(\omega_1, \omega_2)} \quad (68)$$

in complete analogy to the discussion of the Study-Fubini case.

What remains is the computation of the overlap (66) to have access to the other quantities defined in the subsection at hand.

### 3.3 Optimization

It is not obvious from the beginning, but nevertheless true, that we can restrict ourselves to the case

$$\mathcal{H} = \mathcal{H}^a \otimes \mathcal{H}^b, \quad \dim \mathcal{H}^a = \dim \mathcal{H}^b = d \quad (69)$$

((This is because the largest dimension for a cyclic representation of  $\mathcal{B}(\mathcal{H}^a)$  is of dimension  $d^2$ . That is implicit in what follows.))

We shall use a maximally entangled vector, and we choose

$$\varphi = \sum_{j=1}^d |jj\rangle \equiv \sum |j\rangle^a \otimes |j\rangle^b \quad (70)$$

Then for all  $A \in \mathcal{B}(\mathcal{H}^a)$

$$\langle \varphi, (A \otimes \mathbf{1})\varphi \rangle = \text{tr } A \quad (71)$$

is valid. From

$$\psi = (W \otimes \mathbf{1})\varphi, \quad \omega^a = \text{tr}_b |\psi\rangle\langle\psi| \quad (72)$$

it follows

$$\omega^a = WW^\dagger \quad (73)$$

and, vice versa, (73) implies (72).

(73) points to the gauge transformation  $W \rightarrow WU$ ,  $U$  unitary, respecting  $\omega^a$  as a gauge invariant. I can, however, only mention the interesting challenge to construct a complete gauge theory governing these gauge transformations.

Let us return to our problem with two density operators  $\omega_1^a$  and  $\omega_2^a$  and two purifying vectors  $\psi_1$  and  $\psi_2$ . There are just two operators  $W_1, W_2$  in  $\mathcal{B}(\mathcal{H}^a)$  satisfying

$$\psi_j = (W_j \otimes \mathbf{1})\varphi, \quad \omega_j^a = W_j W_j^\dagger \quad (74)$$

and with them we have

$$\langle \psi_1, \psi_2 \rangle = \langle (W_1 \otimes \mathbf{1})\varphi, (W_2 \otimes \mathbf{1})\varphi \rangle = \text{tr } W_1^\dagger W_2 \quad (75)$$

Gauging  $\psi_2 \rightarrow \psi_2'$  by  $W_2 \rightarrow W_2 U$ , we see

$$\langle \psi_1, \psi_2' \rangle = \text{tr } W_1^\dagger W_2 U$$

Let us stress that we fix  $W_1$  and vary only  $W_2$  in this relation. We thus arrive at

$$\text{olap}(\omega_1^a, \omega_2^a) = \sup_{\psi'} |\langle \psi_1, \psi_2' \rangle| = \sup_U \text{tr } W_1^\dagger W_2 U$$

It is  $|\text{tr } BU| \geq \text{tr } B$  in case  $B \geq \mathbf{0}$  and  $U$  is unitary. Thus

$$\text{olap}(\omega_1^a, \omega_2^a) = \text{tr } W_1^\dagger W_2 \text{ if } W_1^\dagger W_2 \geq \mathbf{0}. \quad (76)$$

This condition can always be fulfilled due to the polar decomposition theorem.

(76) can be restated as follows: Whenever

$$\omega_1^a = W_1 W_1^\dagger, \quad \omega_2^a = W_1 W_2^\dagger, \quad W_1^\dagger W_2 \geq \mathbf{0}$$

we conclude

$$\text{Pr}(\omega_1^a, \omega_2^a) = (\text{tr } W_1^\dagger W_2)^2.$$

### 3.4 Why the Bures distance is a distance

Before proceeding along the main line the triangle inequality should be proved. Inserting (76) into (67) yields

$$\text{dist}_B(\omega_1^a, \omega_2^a) = \sqrt{\text{tr } W_1 W_1^\dagger + \text{tr } W_2 W_2^\dagger - 2 \text{tr } W_1^\dagger W_2}$$

Now observe that the traces of  $WW^\dagger$  and  $W^\dagger W$  are equal. Further remind that  $W_1^\dagger W_2$  is assumed positive and, therefore, hermitian:

$$W_1^\dagger W_2 = W_2^\dagger W_1. \quad (77)$$

Altogether we proved

$$\text{dist}_B(\omega_1^a, \omega_2^a) = \sqrt{\text{tr}(W_1 - W_2)^\dagger(W_1 - W_2)} \text{ if } W_1^\dagger W_2 \geq \mathbf{0} \quad (78)$$

Consider now three density (or only positive) operators,  $\omega_1^a$ ,  $\omega_2^a$ , and  $\omega_3^a$ . Starting with  $W_1$  we can choose  $W_2$  and  $W_3$  such that

$$\omega_j^a = W_j W_j^\dagger, \quad W_1^\dagger W_2 \geq \mathbf{0}, \quad W_1^\dagger W_3 \geq \mathbf{0}$$

is valid. We can translate (78) to get

$$\text{dist}_B(\omega_1^a, \omega_k^a) = \|W_k - W_1\|_{vN}, \quad k = 2, 3$$

We apply the triangle inequality for the von Neumann norm to (78).

$$\text{dist}_B(\omega_1^a, \omega_2^a) + \text{dist}_B(\omega_1^a, \omega_3^a) = \|W_2 - W_1\|_{vN} + \|W_3 - W_1\|_{vN} \geq \|W_2 - W_3\|_{vN}$$

and the last term cannot be smaller than the Bures distance. Hence

$$\text{dist}_B(\omega_1^a, \omega_2^a) + \text{dist}_B(\omega_1^a, \omega_3^a) \geq \text{dist}_B(\omega_2^a, \omega_3^a) \quad (79)$$

An application is the extended Tam-Mandelstam inequality. Let be

$$t \rightarrow \omega_t, \quad t' \leq t \leq t'' \quad (80)$$

a solution of

$$i\hbar\dot{\omega} = [H, \omega], \quad H = H_t \quad (81)$$

Then one can prove

$$\int_{t'}^{t''} \sqrt{\text{tr}(\omega H^2) - (\text{tr}\omega H)^2} \geq \hbar \arccos \text{overlap}(\omega_1, \omega_0) \quad (82)$$

(see Phys. Lett. A161 (1992) 329-331. One has to look for a lift  $t \rightarrow W_t$  satisfying a Schrödinger equation with an Hamiltonian  $W \rightarrow HW + W\tilde{H}$ , where  $t \rightarrow \tilde{H}_t$  has to be chosen suitably.)

Bures did not ask whether his distance is based on a Riemann metric. He was interested in cases with infinite tensor products of von Neumann algebras and the theory of infinite dimensional manifolds had not been developed. But for finite dimension one can ask for.

There is, indeed, a Riemann metric reproducing the Bures distance. Its line element is given by

$$\left(\frac{ds}{dt}\right)_B = \text{tr} G^2 \omega = \frac{1}{2} \text{tr} \dot{\omega} G, \quad (83)$$

valid for invertible  $\omega$ , with  $G$  the unique solution of

$$\dot{\omega} = \omega G + G \omega . \quad (84)$$

Using this one can get a differential form of (82) :

$$\text{tr}(\omega H^2) - (\text{tr}\omega H)^2 \geq \frac{\hbar}{2} \text{tr} G \dot{\omega} \quad (85)$$

One may compare this inequality with the “quantum Rao-Cramers inequality”, which, however, plays its role in a quite different context (hypothesis testing and other questions of mathematical statistics). A recent overview, discussing these relationships, is in I. Bengtsson’s paper quant-ph/0509017. Further interesting things you have heard in the talk of A. Ericsson, a written version is quant-ph/0508133 .

### 3.5 An expression for $\text{Pr}(\omega_1, \omega_2)$

We make use of the positivity of  $W_1^\dagger W_2$  which is sufficient for the validity of (76), and which yields also (??). It holds

$$(W_1^\dagger W_2)^2 = W_1^\dagger W_2 W_2^\dagger W_1 = W_1^\dagger \omega_2^a W_1$$

There is a polar decomposition

$$W_1 W_1^\dagger = \omega_1^a, \quad W_1 = (\omega_1^a)^{1/2} U_1$$

with a unitary  $U_1$ . Putting things together yields

$$(W_1^\dagger W_2)^2 = U_1^{-1} \sqrt{\omega_1^a} \omega_2^a \sqrt{\omega_1^a} U_1 \quad (86)$$

we can take the positive root and obtain

$$W_1^\dagger W_2 = U_1^{-1} (\sqrt{\omega_1^a} \omega_2^a \sqrt{\omega_1^a})^{1/2} U_1 . \quad (87)$$

The trace of (87) yields the overlap, its square the transition probability

$$\text{Pr}(\omega_1, \omega_2) = (\text{tr} (\sqrt{\omega_1^a} \omega_2^a \sqrt{\omega_1^a})^{1/2})^2 \quad (88)$$

### 3.6 An estimate

I use the notation

$$\text{char}(A) = \text{all roots of the characteristic equation of } A \quad (89)$$

Clearly, this are the eigenvalues, counted with the appropriate multiplicity, if  $A$  is diagonalisable. If  $A$  is hermitian or normal, (89) is often called the “spectrum” of  $A$ .

Because of

$$\omega_1^{1/2}(\omega_1^{1/2}\omega_2\omega_1^{1/2})\omega_1^{-1/2} = \omega_1\omega_2$$

one concludes

$$\text{char}(\omega_1^{1/2}\omega_2\omega_1^{1/2}) = \text{char}(\omega_1\omega_2) \quad (90)$$

Denoting the characteristic values of (90) by  $\lambda_1, \lambda_2, \dots$ ,

$$\text{Pr}(\omega_1, \omega_2) = \left(\sum \sqrt{\lambda_j}\right)^2 \quad (91)$$

The sum of the  $\lambda_j$  is the trace of  $\omega_1\omega_2$ . Hence

$$\text{Pr}(\omega_1, \omega_2) = \text{tr} \omega_1\omega_2 + \sum_{j \neq k} \sqrt{\lambda_j\lambda_k}$$

We use the inequality

$$\sum_{j \neq k} \sqrt{\lambda_j\lambda_k} \geq \sqrt{\sum_{j \neq k} \lambda_j\lambda_k}$$

which is an equality for  $\dim \mathcal{H} = 2$ . For higher dimensions the estimate becomes weak. Better estimates or a recursive procedure should be possible. Anyway, replacing the left hand side of the estimate, which is the second symmetric function of the  $\lambda_j$ , by traces,

$$\text{Pr}(\omega_1, \omega_2) \geq \text{tr} \omega_1\omega_2 + \sqrt{2}\sqrt{(\text{tr}\omega_1\omega_2)^2 - \text{tr}(\omega_1\omega_2)^2} \quad (92)$$

is valid.

### 3.7 One qubit, $\dim \mathcal{H} = 2$

In the one qubit case (92) becomes an inequality. The trace expression in the root can be much simplified also, being essentially the determinant of  $\omega_1\omega_2$ . Thus

$$\text{Pr}(\omega_1, \omega_2) \geq \text{tr} \omega_1\omega_2 + 2\sqrt{\det \omega_1 \det \omega_2} \quad (93)$$

Let us represent our density matrices by

$$\omega_1 = \frac{1}{2}(\mathbf{1} + \sum x_n \sigma_n), \quad \omega_2 = \frac{1}{2}(\mathbf{1} + \sum y_n \sigma_n) \quad (94)$$

and let us define a new coordinate by

$$x_4 := 2\sqrt{\omega_1}, \quad y_4 := 2\sqrt{\omega_2} \quad (95)$$

We have now placed the density operators on the upper 3-hemisphere,

$$x_1^2 + \dots + x_4^2 = y_1^2 + \dots + y_4^2 = 1 \quad (96)$$

with  $x_4 \geq 0, y_4 \geq 0$  The transition probability becomes

$$\text{Pr}(\omega_1, \omega_2) = \frac{1}{2}\left(1 + \sum_{j=1}^4 x_j y_j\right) \quad (97)$$

### 3.8 A “hidden” symmetry

Remember first the equality (90)

$$\text{char}(\omega_1^{1/2}\omega_2\omega_1^{1/2}) = \text{char}(\omega_1\omega_2)$$

of the characteristic numbers. Let  $Z$  be invertible and consider the change

$$\omega'_1 = Z^{-1}\omega_1(Z^{-1})^\dagger, \quad \omega'_2 = Z^\dagger\omega_2Z \quad (98)$$

One immediately see

$$\omega'_1\omega'_2 = Z^{-1}(\omega_1\omega_2)Z \quad (99)$$

and

$$\text{char}(\omega'_1\omega'_2) = \text{char}(\omega_1\omega_2). \quad (100)$$

Now (90) implies:

The eigenvalues of  $\sqrt{\omega_1}\omega_2\sqrt{\omega_1}$  *do not change* if  $\omega_1, \omega_2$  are transformed according to (98). In particular

$$\text{olap}(\omega_1, \omega_2) = \text{olap}(Z^{-1}\omega_1(Z^{-1})^\dagger, Z^\dagger\omega_2Z) \quad (101)$$

Indeed, the argument is valid for every symmetric function of the characteristic numbers in question.

We even can relax from the invertibility of  $Z$  by substituting

$$\omega_1 \rightarrow Z\omega_1Z^\dagger$$

in (101):

$$\text{olap}(Z\omega_1Z^\dagger, \omega_2) = \text{olap}(\omega_1, Z^\dagger\omega_2Z) \quad (102)$$

Relying on continuity we can state (102) for all operators  $Z$ .

### 3.9 Super-additivity of the overlap

At first let  $\omega_1, \omega_2$  invertible. Then  $W_1$  and  $W_2$  (used to purify) are invertible also. Now

$$W_1^\dagger W_2 \geq \mathbf{0} \Leftrightarrow W_2 = KW_1, \quad K > \mathbf{0} \quad (103)$$

To see the positivity of  $K$ , multiply from the left with  $W_1^\dagger$ . We now observe

$$\text{tr}\omega_1 K = \text{tr}KW_1W_1^\dagger = \text{tr}W_1^\dagger W_2, \quad \text{tr}\omega_2 K^{-1} = \text{tr}W_1^\dagger W_2 \quad (104)$$

But  $\text{tr}W_1^\dagger W_2$  is the overlap and with this particular  $K$  we have

$$\text{olap}(\omega_1, \omega_2) = \frac{1}{2}(\text{tr}\omega_1 K + \text{tr}\omega_2 K^{-1}) \quad (105)$$

However, for every positive  $C \in \mathcal{B}(\mathcal{H}^a \otimes \mathcal{H}^b)$

$$2|\langle\psi_1, \psi_2\rangle| \leq \langle\psi_1, C\psi_1\rangle + \langle\psi_2, C^{-1}\psi_2\rangle$$

by the Schwartz inequality. If we restrict ourself to  $C = A \otimes \mathbf{1}$  it becomes clear, that the right hand side of (105) cannot become smaller in substituting another invertible positive operator for  $K$ . This proves

$$\mathrm{tr} \sqrt{\sqrt{\omega_1} \omega_2 \sqrt{\omega_1}} = \frac{1}{2} \inf_{A > \mathbf{0}} (\mathrm{tr} \omega_1 A + \mathrm{tr} \omega_2 A^{-1}) \quad (106)$$

Remind that the left hand side is nothing but  $\mathrm{olap}(\omega_1, \omega_2)$ .

For all decompositions

$$\omega = \sum \omega_j, \quad \rho = \sum \rho_j \quad (107)$$

the inequality

$$\mathrm{olap}(\omega, \rho) \geq \sum_j \mathrm{olap}(\omega_j, \rho_j) \quad (108)$$

is valid. The inequality shows what is called “super-additivity” for the overlap.

The proof uses (106). We use  $A = K$  from (104) to get the left hand side of (108). But with this choice every term on the right of (108) cannot be smaller than the corresponding overlaps because of (106). Thus we are done.

### 3.10 Monotonicity

Choi proved for positive unital super-operators

$$\Psi(A^{-1}) \geq \Psi(A)^{-1} \text{ if } A \geq \mathbf{0} \quad (109)$$

A straightforward proof is possible if  $\Psi$  is 2-positive. One considers

$$\begin{pmatrix} A & \mathbf{1} \\ \mathbf{1} & A^{-1} \end{pmatrix} \Rightarrow \begin{pmatrix} \Psi(A) & \Psi(\mathbf{1}) \\ \Psi(\mathbf{1}) & \Psi(A^{-1}) \end{pmatrix} \quad (110)$$

and concludes, by 2-positivity, the positivity of the right hand side if the matrix at the left is positive. But this takes place for  $A > \mathbf{0}$ .

The matrix at the right hand side is positive if

$$\Psi(A^{-1}) \geq \Psi(\mathbf{1}) \Psi(A)^{-1} \Psi(\mathbf{1}) \quad (111)$$

Now (109) follows with  $\Psi(\mathbf{1}) = \mathbf{1}$ , i.e. by the assumed unitality of  $\Psi$ .

Let us apply (109) to the overlap. To this end we denote by  $\Phi$  the super-operator dual to  $\Psi$ ,

$$\mathrm{tr} X \Psi(Y) = \mathrm{tr} \Phi(X) Y \quad (112)$$

$\Psi$  is positive if and only if  $\Phi$  is positive.  $\Psi$  is unital iff  $\Phi$  is trace preserving. (106) provides us with

$$\mathrm{olap}(\omega_1, \omega_2) \leq \frac{1}{2} \inf_{A > \mathbf{0}} (\mathrm{tr} \omega_1 \Psi(A) + \mathrm{tr} \omega_2 \Psi(A)^{-1})$$

Indeed, the set of all  $\Psi(A)$  with  $A > \mathbf{0}$  is contained in the set of all  $A > \mathbf{0}$  because of positivity of  $\Psi$ . Now Choi's inequality (109) provides us with

$$\text{olap}(\omega_1, \omega_2) \leq \frac{1}{2} \inf_{A > \mathbf{0}} (\text{tr} \omega_1 \Psi(A) + \text{tr} \omega_2 \Psi(A^{-1}))$$

At this point we switch to the dual map  $\Phi$  to obtain

$$\text{olap}(\omega_1, \omega_2) \leq \frac{1}{2} \inf_{A > \mathbf{0}} (\text{tr} \Phi(\omega_1) A + \text{tr} \Phi(\omega_2) A^{-1})$$

To finish the proof we use (106) again to get the *monotonicity property*

$$\text{olap}(\omega_1, \omega_2) \leq \text{olap}(\Phi(\omega_1), \Phi(\omega_2)) \quad (113)$$

for all trace-preserving positive super-operators  $\Phi$ .

((We proved it for 2-positive  $\Psi$ . Following Choi, refined arguments show that positivity is enough.))

### 3.11 Two additional remarks

I like to add two remarks, pointing in different directions. The first concerns direct products and is as well important as simple. The second concerns the geometric mean, referred to in an appendix, which allows a nice reformulation of what has already been said. Its importance, however, is in allowing to do almost all, what I have described in the finite case, for von Neumann (and even for unital  $C^*$ -) algebras.

Direct products. With two pairs,  $\omega_1, \omega_2$  and  $\rho_1, \rho_2$  in two different Hilbert spaces, one can perform their direct products  $\omega_j \otimes \rho_j$ . The structure of the expression (88) allows to conclude

$$\text{Pr}(\omega_1 \otimes \rho_1, \omega_2 \otimes \rho_2) = \text{Pr}(\omega_1, \omega_2) \text{Pr}(\rho_1, \rho_2) \quad (114)$$

Using the geometric mean. The geometric mean is defined and discussed in the appendix to which I refer to. The operator  $K$ , defined in (103) by

$$W_2 = KW_1 \text{ if } W_1^\dagger W_2 > \mathbf{0}$$

(and which was called  $M$  in the talk of A. Ericsson,) kann be written

$$K = \omega_1 \# \omega_2^{-1} = \omega_2^{-1} \# \omega_1 \quad (115)$$

Remembering the notations

$$\underline{\omega}(A) = \text{tr} \omega A$$

we may rewrite (104) as

$$\Pr(\omega_1, \omega_2) = \underline{\omega}_2(\omega_2^{-1} \# \omega_1) = \underline{\omega}_1(\omega_1^{-1} \# \omega_2) \quad (116)$$

and (106) reads

$$\text{olap}(\omega_1, \omega_2) = \frac{1}{2} \inf_{A > \mathbf{0}} (\omega_1(A) + \omega_2(A^{-1})) \quad (117)$$

To go further in the direction that defines states as positive and normalized linear form of an algebra as indicated in subsections 2.3 and 2.4, we consider special linear functionals,

$$\underline{\nu}(A) = \text{tr } \nu A, \quad (118)$$

which may be eventually called “transition forms”. To do so, we require

$$|\underline{\nu}(A^\dagger B)|^2 \leq \omega_1(A^\dagger A) \omega_2(B^\dagger B) \quad (119)$$

for all  $A, B$  in, say,  $\mathcal{B}(\mathcal{H}^a)$  ((or in any other suitable algebra)). Let  $\psi'_1, \psi'_2$  be a purifying pair of the states and  $\psi_j = (W'_j \otimes \mathbf{1})\varphi$ . Then

$$\underline{\nu}'(A) = \langle \psi_1, (A \otimes \mathbf{1})\psi_2 \rangle = \text{tr} (W'_1)^\dagger W'_2 \quad (120)$$

satisfies (119) (use Cauchy’s inequality).

One can show that (119) is sufficient for representing  $\underline{\nu}$  in the manner (120) by a pair of purifications.

Therefore, setting  $A = \mathbf{1}$ , we can assert

$$\Pr(\omega_1, \omega_2) = \sup_{\underline{\nu}} |\underline{\nu}(\mathbf{1})|^2 \quad (121)$$

such that  $\underline{\nu}$  runs over all linear forms fulfilling (119). Indeed, to take this as a definition has been my starting point for all that.

## 4 Appendices

### 4.1 The geometrical mean

Let  $A, B$ , and  $C$  positive operators in a finite dimensional Hilbert space. A remarkable observation due to Pusz and Woronowicz can be rephrased in the following form:

Fixing  $A$  and  $B$ , there is a largest operator in the set of all  $C$  satisfying

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq \mathbf{0}. \quad (122)$$

This unique element is called *the geometrical mean of  $A$  and  $B$*  and it will be denoted, following Ando, by

$$A \# B. \quad (123)$$

In other words: (122) is valid if and only if

$$C \leq A\#B . \quad (124)$$

From the definition we get the relation

$$A\#B = B\#A, \quad A^{-1}\#B^{-1} = (A\#B)^{-1} , \quad (125)$$

the latter being true if the operators are invertible. If just  $A$  is invertible then the block matrix (122) is positive if and only if

$$B \geq CA^{-1}C \quad (126)$$

and one can conclude that  $A\#B$  is the unique positive solution  $X$  of the equation

$$B = XA^{-1}X, \quad X \geq \mathbf{0} . \quad (127)$$

The equation can be solved and one gets

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \quad (128)$$

If  $A$  and  $B$  commute one can see from (122)

$$AB = BA \Rightarrow A\#B = (AB)^{1/2} \quad (129)$$

To see it one uses a common eigenbasis to reduce (122) to the positivity of

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \geq \mathbf{0} \Leftrightarrow ab \geq c^2$$

for three positive numbers  $a$ ,  $b$ , and  $c$ .

To get a further description of  $A\#B$  we shall use the fact that it is an operator mean, i.e. for invertible  $Z$  it enjoys

$$Z(A\#B)Z^* = (ZAZ^*)\#(ZBZ^*) . \quad (130)$$

For the proof one relays on

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq \mathbf{0} \Leftrightarrow \begin{pmatrix} ZAZ^* & ZCZ^* \\ ZCZ^* & ZBZ^* \end{pmatrix} \geq \mathbf{0}$$

for invertible  $Z$ .

Now we can combine (129) and (130) with  $Z = A + B$ . To check the positivity of (122) it is sufficient to so so an the support space of  $A + B$ . Thus we may assume that this operator is invertible. Then

$$A' = (A + B)^{-1/2}A(A + B)^{-1/2} \quad \text{and} \quad B' = (A + B)^{-1/2}B(A + B)^{-1/2}$$

commute. Indeed, it follows

$$A' + B' = \mathbf{1}, \quad A'\#B' = (A'B')^{1/2}$$

and we can apply (130). Thus

$$A\#B = (A+B)^{1/2}((A+B)^{-1/2}A(A+B)^{-1}B(A+B)^{-1/2})^{1/2}(A+B)^{1/2} \quad (131)$$

Next we prove super-additivity. Let

$$A = \sum A_j, \quad B = \sum B_j, \text{ and}$$

$$C_j = A_j\#B_j, \quad C = \sum C_j$$

Then

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} = \sum \begin{pmatrix} A_j & C_j \\ C_j & B_j \end{pmatrix}$$

is a positive block matrix. Thus  $C$  is smaller than  $A\#B$  and that proves

$$A\#B \geq \sum A_j\#B_j. \quad (132)$$

Our next task is monotonicity.

Let  $\Phi$  be a 2-positive super-operator. Then with

$$\begin{pmatrix} A & A\#B \\ A\#B & B \end{pmatrix} \text{ also } \begin{pmatrix} \Phi(A) & \Phi(A\#B) \\ \Phi(A\#B) & \Phi(B) \end{pmatrix}$$

must be a positive block operator with positive entries. Hence,

$$\Phi(A\#B) \leq \Phi(A)\#\Phi(B) \quad (133)$$

is valid by the very definition of the geometric mean.