Geometry of State Spaces

Armin Uhlmann and Bernd Crell

Institut für Theoretische Physik, Universität Leipzig
armin.uhlmann@itp.uni-leipzig.de
bernd.crell@itp.uni-leipzig.de

In the Hilbert space description of quantum theory one has two major inputs: Firstly its linearity, expressing the superposition principle, and, secondly, the scalar product, allowing to compute transition probabilities. The scalar product defines an Euclidean geometry. One may ask for the physical meaning in quantum physics of geometric constructs in this setting. As an important example we consider the length of a curve in Hilbert space and the "velocity", i. e. the length of the tangents, with which the vector runs through the Hilbert space.

The Hilbert spaces are generically complex in quantum physics: There is a multiplication with complex numbers: Two linear dependent vectors represent the same state. By restricting to unit vectors one can diminish this arbitrariness to phase factors.

As a consequence, two curves of unit vectors represent the same curve of states if they differ only in phase. They are physically equivalent. Thus, considering a given curve — for instance a piece of a solution of a Schrödinger equation – one can ask for an equivalent curve of minimal length. This minimal length is the "Fubini-Study length". The geometry, induced by the minimal length requirement in the set of vector states, is the "Fubini-Study metric".

There is a simple condition from which one can read off whether all pieces of a curve in Hilbert space fulfill the minimal length condition, so that their Euclidean and their Study-Fubini length coincide piecewise: It is the parallel transport condition, defining the geometric (or Berry) phase of closed curves by the following mechanism: We replace the closed curve by changing its phases to a minimal length curve. Generically, the latter will not close. Its initial and its final point will differ by a phase factor, called the geometric phase (factor). We only touch these aspects in our essay and advice the reading of [6] instead. We discuss, as quite another application, the Tam-Mandelstamm inequalities.

The set of vector states associated to a Hilbert space can also be described as the set of its 1-dimensional subspaces or, equivalently, as the set of all rank one projection operators. Geometrically it is the "projective space" given by

the Hilbert space in question. In finite dimension it is a well studied manifold[1]. Again, we advice the reader to a more comprehensive monograph, say [23], to become acquainted with projective spaces in quantum theory. We just like to point to one aspect: Projective spaces are rigid. A map, transforming our projective space one–to–one onto itself and preserving its Fubini-Study distances, is a Wigner symmetry. On the the Hilbert space level this is a theorem of Wigner asserting that such a map can be induced by a unitary or by an anti-unitary transformation.

After these lengthy introduction to our first chapter, we have not much to comment to our third one. It is just devoted to the (partial) extension of what has been said above to general state spaces. It will be done mainly, but not only, by purification methods.

The central objects are the generalized transition probability ("fidelity"), the Bures distance, and its Riemann metric. These concepts can be defined, and show similar features, in all quantum state spaces. They are "universal" in quantum physics.

However, at the beginning of quantum theory people were not sure whether density operators describe states of a quantum system or not. In our days, we think, the question is completely settled. There are genuine quantum states described by density operators. But not only that, the affirmative answer opened new insight into the basic structure of quantum theory. The second chapter is dedicated to these structural questions.

To a Hilbert spaces $\mathcal{H}$ one associates the algebra $\mathcal{B}(\mathcal{H})$ of all bounded operators which map $\mathcal{H}$ into itself. With a density operator ω and any operator $A \in \mathcal{B}(\mathcal{H})$, the number

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

is the expectation value of A, provided the "system is in the state given by ω". Now $\underline{\omega}$ is linear in A, it attains positive values or zero for positive operators, and it returns 1 if we compute the expectation value of the identity operator $\mathbf{1}$. These properties are subsumed by saying "$\underline{\omega}$ is a normalized positive linear functional on the algebra $\mathcal{B}(\mathcal{H})$". Exactly such functionals are also called "states of $\mathcal{B}(\mathcal{H})$ ", asserting that every one of these objects can possibly be a state of our quantum system.

If the Hilbert space is of finite dimension then every state of $\mathcal{B}(\mathcal{H})$ can be characterized by a density operator. But in the other, infinite cases, there are in addition states of quite different nature, the so-called singular ones. They can be ignored in theories with finitely many degrees of freedom, for instance in Schrödinger theory. But in treating unbounded many degrees of freedom we have to live with them.

One goes an essential step further in considering more general algebras than $\mathcal{B}(\mathcal{H})$. The idea is that a quantum system is defined, ignoring other

[1] It is certainly the most important algebraic variety.

demands, by its observables. States should be identified by their expectation values. However, not any set of observables should be considered as a quantum system. There should be an algebra, say $\mathcal{A}$, associated to a quantum system containing the relevant observables. Besides being an algebra (over the complex numbers), an Hermitian adjoint, $A^\dagger$ should be defined for every $A \in \mathcal{A}$ and, finally, there should be "enough" states of $\mathcal{A}$.

As a matter of fact, these requirement are sufficient if $\mathcal{A}$ is of finite dimension as a linear space. Then the algebra can be embedded as a so-called *-subalgebra in any algebra $\mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H}$ sufficient large or infinite. Relying on Wedderburn's theorem, we describe all these algebras and their state spaces, They all can be gained by performing direct products and direct sums of some algebras $\mathcal{B}(\mathcal{H})$. Intrinsically they are enumerated by their "type", a finite set of positive numbers. We abbreviate this set by $\mathbf{d}$ to shorten the more precise notation $I_{\mathbf{d}}$ for so-called type one algebras.

If the algebras are not finite, things are much more involved. There are von Neumann (i. e. concrete W*-) algebras, C*-algebras, and more general classes of algebras. About them we say (almost) nothing but refer, for a physical motivated introduction, to [10].

Let us stress, however, a further point of view. In a bipartite system, which is the direct product of two other ones, (– say Alice's and Bob's –), both systems are embedded in the larger one as subsystems. Their algebras become subalgebras of another, larger algebra.

There is a more general point of view: It is a good idea to imagine the quantum world as a hierarchy of quantum systems. An "algebra of observables" is attached to each one, together with its state space. The way, an algebra is a subalgebra of another one, is describing how the former one should be understood as a quantum subsystem of a "larger" one.

Let us imagine such a hierarchical structure. A state in one of these systems determines a state in every of its subsystems: We just have to look at the state by using the operators of the subsystem in question only, i. e. we recognize what possibly can be observed by the subsystems observables.

Restricting a state of a quantum system (of an algebra) to a subsystem (to a subalgebra) is equivalent to the "partial trace" in direct products. It extends the notion to more general settings.

On the other hand, starting with a system, every operator remains an operator in all larger systems containing the original one as a subsystem. To a large amount the physical meaning of a quantum system, its operators and its states, is determined by its relations to other quantum systems.

There is an appendix devoted to the geometric mean, the perhaps most important operator mean. It provides a method to handle two positive operators in general position. Only one subsection of the appendix is directly connected with the third chapter: How parallel lifts of Alice's states are seen by Bob.

As a matter of fact, a chapter describing more subtle problems of convexity had been proposed. But we could not finish it in due time. Most of the appendix has been prepared for it. Nevertheless, in the main part there is a short description of the convex structure of states spaces in finite systems (faces, extremal points, rigidity).

1 Geometry of pure states

1.1 Norm and distance in Hilbert space

Let us consider a Hilbert space[2] $\mathcal{H}$. Its elements are called *vectors.* If not explicitly stated otherwise, we consider complex Hilbert spaces, i. e. the multiplication of vectors with complex numbers is allowed. Vectors will be denoted either by their mathematical notation, say ψ, or by Dirac's, say $|\psi\rangle$.

For the scalar product we write accordingly[3]

$$\langle\varphi, \psi\rangle \text{ or } \langle\varphi|\psi\rangle \ .$$

The *norm* or *vector norm* of $\psi \in \mathcal{H}$ reads

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

It is the Euclidian length of the vector ψ. One defines

$$\| \psi - \psi' \| = \text{ distance between } \psi \text{ and } \psi'$$

which is an *Euclidean distance:* If in any basis, $|j\rangle$, $j = 1, 2, \ldots$, one gets

$$|\psi\rangle = \sum z_j |j\rangle, \quad z_j = x_j + iy_j \ ,$$

and, with coefficients z'_j, the similar expansion for ψ', then

$$\| \psi - \psi' \| = (\sum [x_j - x'_j]^2 + [y_j - y'_j]^2)^{1/2} \ ,$$

and this justifies the name "Euclidean distance".

The scalar product defines the norm and the norm the Euclidean geometry of $\mathcal{H}$. In turn one can obtain the scalar product from the vector norm:

$$4\langle\psi, \psi'\rangle = \| \psi + \psi' \|^2 - \| \psi - \psi' \|^2 - i \| \psi + i\psi' \|^2 + i \| \psi - i\psi' \|^2$$

The scalar product allows for the calculation of quantum probabilities. Now we see that, due to the complex structure of $\mathcal{H}$, these probabilities are also encoded in its Euclidean geometry.

[2] We only consider Hilbert spaces with countable bases.

[3] We use the "physicist's convention" that the scalar product is anti-linear in φ and linear in ψ.

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

We ask for the *length* of a curve in Hilbert space. The curve may be given by

$$t \to \psi_t, \quad 0 \le t \le 1 \tag{1}$$

where t is a parameter, not necessarily the time. We assume that for all $\varphi \in \mathcal{H}$ the function $t \to \langle \varphi, \psi_t \rangle$ of t is continuous.

To get the length of (1) we have to take all subdivisions

$$0 \le t_0 < t_1 < \ldots < t_n \le 1$$

in performing the sup,

$$\text{length of the curve } = \sup \sum_{j=1}^{n} \parallel \psi_{t_{j-1}} - \psi_{t_j} \parallel \ . \tag{2}$$

The length is independent of the parameter choice.

If we can guaranty the existence of

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

then one knows

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

The vector $\dot{\psi}_t$ is the (contra-variant) tangent along (1). Its lengths is the *velocity* with which[4] $\psi = \psi_t$ travels through $\mathcal{H}$, i. e.

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

Interesting examples are solutions $t \to \psi_t$ of a Schrödinger equation,

$$H\psi = i\hbar\dot{\psi} \ . \tag{6}$$

In this case the tangent vector is time independent and we get

$$\frac{ds}{dt} = \hbar^{-1}\sqrt{\langle \psi, H^2\psi \rangle}\, . \tag{7}$$

The length of the solution between the times t_0 and t_1 is

$$\text{length } = \hbar^{-1}(t_1 - t_0)\ \sqrt{\langle \psi, H^2\psi \rangle} \tag{8}$$

Anandan,[28], has put forward the idea to consider the Euclidean length (5) as an intrinsic and universal parameter in Hilbert space. For example, consider

[4] We often write just ψ instead of ψ_t.

$$\frac{dt}{ds} = \langle \dot{\psi}, \dot{\psi} \rangle^{-1/2} = \hbar \left(\langle \psi, H^2 \psi \rangle \right)^{-1/2}$$

instead of ds/dt and interpret it as the *quantum velocity with which time is elapsing during a Schrödinger evolution.* Also other metrical structures, to which we come later on, allow for similar interpretations.
Remark: Though we shall be interested mostly in finite dimensional Hilbert spaces or, in the infinite case, in bounded operators, let us have a short look at the general case.

In the Schrödinger theory H, the Hamilton operator, is usually "unbounded" and there are vectors not in the domain of definition of H. However, there is always an integrated version: A unitary group

$$t \to U(t) = \exp(\frac{tH}{i\hbar})$$

which can be defined rigorously for self-adjoint H.
Then ψ_0 belongs to the domain of definition of H exactly if the tangents (3) of the curve $\psi_t = U(t)\psi_0$ exist. If the tangents exist then the Hamiltonian can be gained by

$$i\hbar \lim_{\epsilon \to 0} \frac{U(t+\epsilon) - U(t)}{\epsilon} \psi = H\psi$$

and (7) and (8) apply.
If, however, ψ_0 does not belong to the domain of definition of H, then (2) returns ∞ for the length of every piece of the curve $t \to U(t)\psi_0$. In this case the vector runs, during a finite time interval, through an infinitely long piece of $t \to \psi_t$. The velocity ds/dt must be infinite.

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) $\mathrm{dist}(\xi_1, \xi_2) = \mathrm{dist}(\xi_2, \xi_1)$
b) $\mathrm{dist}(\xi_1, \xi_2) + \mathrm{dist}(\xi_2, \xi_3) \geq \mathrm{dist}(\xi_1, \xi_3)$,
c) $\mathrm{dist}(\xi_1, \xi_2) = 0 \Leftrightarrow \xi_1 = \xi_2$.
A set with a distance is a *metric space.*

Given the distance, dist(.,.), of a metric space and two different points, say ξ_0 and ξ_1, one may ask for the length of a continuous curve connecting these two points[5]. The inf of the lengths over all these curves is again a distance, the *inner distance.* The inner distance, $\mathrm{dist}_{\mathrm{inner}}(\xi_0, \xi_1)$ is never smaller than the original one,

$$\mathrm{dist}_{\mathrm{inner}}(\xi_0, \xi_1) \geq \mathrm{dist}(\xi_0, \xi_1)$$

If equality holds, the distance (and the metric space) is called *inner.* A curve, connecting ξ_0 and ξ_1, the length of which equals the distance between the

[5] We assume that for every pair of points such curves exist.

two point, is called a "short geodesic" arc. A curve, which is short geodesic between all sufficiently near points, is a *geodesic.*

The Euclidian distance is an inner distance. It is easy to present the shortest curves between to vectors, ψ_1 and ψ_0, in Hilbert space:

$$t \to \psi_t = (1-t)\psi_0 + t\psi_1, \quad \dot{\psi} = \psi_1 - \psi_0 \ . \tag{9}$$

is a short geodesic arc between both vectors.

In Euclidean spaces the shortest connection between two points is a piece of a straight line, and this *geodesic arc* is unique. Indeed, from (9) we conclude

$$\| \psi_t - \psi_r \| = |t-r| \ \| \psi_1 - \psi_0 \| \ . \tag{10}$$

With this relation we can immediately compute (2) and we see that the length of (9) is equal to the distance between starting and end point.
We have seen something more general: *If in a linear space the distance is defined by a norm, the metric is inner and the geodesics are of the form* (9).

1.4 Curves on the unit sphere

Restricting the 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 in an Euclidean 3-space. In computing a length by (2) only curves on the sphere are allowed.

The geodesics on a unit sphere are *great circles.* These are sections of the sphere with a plane that contains the center of the sphere. The spherical distance of two points, say ψ_0 and ψ_1, is the angle, α, between the rays from the center of the sphere to the two points in question.

$$\mathrm{dist}_{\mathrm{spherical}}(\psi_1, \psi_0) = \text{angle between the radii pointing to } \psi_0 \text{ and } \psi_1 \tag{11}$$

with the restriction $0 \le \alpha \le \pi$ of the angle α. By the additivity modulo 2π of the angle one can compute (2) along a great circle to see that (11) is an inner metric.

If the two points are not antipodes, $\psi_0 + \psi_1 \neq 0$, then the great circle crossing through them and short geodesic arc between the two vectors is unique.
For antipodes the great circle crossing through them is not unique and there are many short geodesic arcs of length π connecting them.

By elementary geometry

$$\| \psi_1 - \psi_0 \| = \sqrt{2 - 2\cos\alpha} = 2 \sin \frac{\alpha}{2} \tag{12}$$

and $\cos\alpha$ can be computed by

$$\cos\alpha = \frac{\langle \psi_0, \psi_1 \rangle + \langle \psi_1, \psi_0 \rangle}{2} \tag{13}$$

We see from (13) that

$$\cos\alpha \leq |\langle\psi_0, \psi_1\rangle| \tag{14}$$

Therefore, we have the following statement:
The length of a curve on the unit sphere connecting ψ_0 with ψ_1 is at least

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

Applying this observation to the solution of a Schrödinger equation one gets the *Mandelstam-Tamm inequalities,* [57]. To get them one simply combines (14) with (8): If a solution ψ_t of the Schrödinger equation (6) goes at time $t = t_0$ through the unit vector ψ_0 and at time $t = t_1$ through ψ_1, then

$$(t_1 - t_0)\sqrt{\langle\psi, H^2\psi\rangle} \geq \hbar \arccos|\langle\psi_0, \psi_1\rangle| \tag{15}$$

must be valid. (Remember that H is conserved along solutions of (6) and we can use any $\psi = \psi_t$ from the assumed solution.)
However, a sharper inequality holds,

$$(t_1 - t_0)\sqrt{\langle\psi, H^2\psi\rangle - \langle\psi, H\psi\rangle^2} \geq \hbar \arccos|\langle\psi_0, \psi_1\rangle| \ . \tag{16}$$

Namely, the right-hand side is invariant against "gauge transformations"

$$\psi_t \mapsto \psi'_t = (\exp i\gamma t)\ \psi_t \ .$$

The left side of (16) does not change in substituting H by $H' = H - \gamma\mathbf{1}$ and ψ'_t is a solution of

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

Hence we can "gauge away" the extra term in (16) to get the inequality (15).

Remarks:
a) The reader will certainly identify the square root expression in (16) as the "uncertainty" $\triangle_\psi(H)$ of H in the state given by the unit vector ψ. More specific, (16) provides the strict lower bound $T\triangle_\psi(H) \geq h/4$ for the time T to convert ψ to a vector orthogonal to ψ by a Schrödinger evolution.
b) If $U(r)$ is a one-parameter unitary group with generator A, then

$$|r|\triangle_\psi(A) \geq \arccos|\langle\psi, U(r)\psi\rangle|.$$

Interesting candidates are the position and the momentum operators, the angular momentum along an axis, occupation number operators, and so on.
c) The tangent space consists of pairs $\{\psi, \dot{\psi}\}$ with a tangent or velocity vector $\dot{\psi}$, reminiscent from a curve crossing through ψ. The fiber of all tangents based at ψ carries the positive quadratic form

$$\psi_1, \psi_2 \to \langle\psi, \psi\rangle\ \langle\psi_1, \psi_2\rangle - \langle\psi_1, \psi\rangle\ \langle\psi, \psi_2\rangle$$

gained by polarization.
d) More general than proposed in (16), one can say something about time dependent Hamiltonians, $H(t)$, and the Schödinger equation

$$H(t)\psi = i\hbar\dot{\psi} \; . \tag{17}$$

If a solution of (17) is crossing the unit vectors ψ_j at times t_j, then

$$\int_{t_0}^{t_1} \sqrt{\langle\psi, H^2\psi\rangle - \langle\psi, H\psi\rangle^2}dt \geq \hbar \arccos|\langle\psi_0, \psi_1\rangle| \; . \tag{18}$$

For further application to the speed of quantum evolutions see [43].

1.5 Phases

If the vectors ψ and ψ' are linearly dependent, they describe the same state.

$$\psi \mapsto \pi_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi, \psi\rangle} \tag{19}$$

maps the vectors of $\mathcal{H}$ onto the pure states, with the exception of the zero vector. Multiplying a vector by a complex number different from zero is the *natural gauge transformation* offered by $\mathcal{H}$.
From this freedom in choosing a state vector for a pure state,

$$\psi \to \epsilon\psi, \quad |\epsilon| = 1, \tag{20}$$

the *phase change,* is of primary physical interest.

In the following we consider parameterized curve as in (1) on the unit sphere of $\mathcal{H}$. At first we see

$$\langle\psi_t, \dot{\psi}_t\rangle \text{ is purely imaginary.} \tag{21}$$

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.
The curves

$$t \to \psi_t \text{ and } t \to \psi'_t := \epsilon_t\psi_t, \quad \epsilon_t = \exp(i\gamma_t), \tag{22}$$

are *gauge equivalent.* The states themselves,

$$t \to \pi_t = |\psi_t\rangle\langle\psi_t| \tag{23}$$

are gauge invariant.

From the transformation (22) we deduce for the tangents

$$\dot{\psi}' = \dot{\epsilon}\psi + \epsilon\dot{\psi}, \quad \epsilon^{-1}\dot{\epsilon} = i\dot{\gamma} \tag{24}$$

with real γ. Thus, by an appropriate choice of the gauge, one gets

$$\langle \psi', \dot{\psi}' \rangle = 0 \ , \text{ the geometric phase transport condition} \tag{25}$$

(Fock, [41], from adiabatic reasoning). Indeed, (25) is the equation

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

Because of $\langle \psi, \psi \rangle = 1$ and $\langle \psi, \dot{\psi} \rangle = -\langle \dot{\psi}, \psi \rangle$ we get

$$\epsilon_t = \exp \int_{t_0}^{t} \langle \dot{\psi}, \psi \rangle \, dt. \tag{26}$$

For a curve $t \to \psi_t$, $0 \leq t \leq 1$, with $\psi_1 = \psi_0$ the integral is the *geometric or Berry phase,* [33]. For more about phases see [6].
Remark: This is true on the unit sphere. If the vectors are not normalized one has to replace (25) by the vanishing of the "gauge potential"

$$\frac{\langle \psi, \dot{\psi} \rangle - \langle \dot{\psi}, \psi \rangle}{2i} \ \text{ or } \ \frac{\langle \psi, \dot{\psi} \rangle - \langle \dot{\psi}, \psi \rangle}{2i\langle \psi, \psi \rangle} \ . \tag{27}$$

In doing so, we conclude: The phase transport condition and the Berry phase do not depend on the normalization.

1.6 Fubini-Study distance

With the Fubini-Study distance, [42], [56], the set of pure states becomes an inner metric space. But at first we introduce a slight deviation from its original form which is defined on the positive operators of rank one. To this end we look at (19) in two steps. First we skip normalization and replace (19) by

$$\psi \mapsto |\psi\rangle\langle\psi| \ ,$$

and only after that we shall normalize.

Let ψ_0 and ψ_1 be two vectors from $\mathcal{H}$. We start with the first form of the Fubini-Study distance:

$$\mathrm{dist}_{\mathrm{FS}}(|\psi_1\rangle\langle\psi_1|, |\psi_0\rangle\langle\psi_0|) = \min_{\epsilon} \parallel \psi_1 - \epsilon\psi_0 \parallel \tag{28}$$

where the minimum is over the complex numbers ϵ, $|\epsilon| = 1$. One easily finds

$$\mathrm{dist}_{\mathrm{FS}}(|\psi_1\rangle\langle\psi_1|, |\psi_0\rangle\langle\psi_0|) = \sqrt{\langle \psi_0, \psi_0 \rangle + \langle \psi_1, \psi_1 \rangle - 2|\langle \psi_1, \psi_0 \rangle|} \ . \tag{29}$$

Therefore, (28) coincides with $\parallel \psi_1 - \psi_0 \parallel$ after choosing the relative phase appropriately, i. e. after choosing $\langle \psi_1, \psi_0 \rangle$ real and not negative.

(28) is a distance in the set of positive rank one operators: Choosing the phases between ψ_2, ψ_1 and between ψ_1, ψ_0 appropriately,

$$\mathrm{dist}_{\mathrm{FS}}(|\psi_2\rangle\langle\psi_2|, |\psi_1\rangle\langle\psi_1|) + \mathrm{dist}_{\mathrm{FS}}(|\psi_1\rangle\langle\psi_1|, |\psi_0\rangle\langle\psi_0|)$$

becomes equal to

$$\| \psi_2 - \psi_1 \| + \| \psi_1 - \psi_0 \| \geq \| \psi_2 - \psi_0 \|$$

and, therefore,

$$\| \psi_2 - \psi_0 \| \geq \mathrm{dist}_{\mathrm{FS}}(|\psi_2\rangle\langle\psi_2|, |\psi_0\rangle\langle\psi_0|) \; .$$

Now we can describe the geodesics belonging to the distance $\mathrm{dist}_{\mathrm{FS}}$ and see that (28) is an inner distance: If the scalar product between ψ_0 and ψ_1 is real and not negative, then this is true for the scalar products between any pair of the vectors

$$t \to \psi_t := (1-t)\psi_0 + t\psi_1, \quad \langle\psi_1, \psi_0\rangle \geq 0. \tag{30}$$

Then we can conclude

$$\mathrm{dist}_{\mathrm{FS}}(|\psi_r\rangle\langle\psi_r|, |\psi_t\rangle\langle\psi_t|) = \| \psi_r - \psi_t \| \tag{31}$$

and (30) is geodesic in $\mathcal{H}$. Furthermore,

$$t \to |\psi_t\rangle\langle\psi_t|, \quad 0 \leq t \leq 1, \tag{32}$$

is the shortest arc between $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$. Explicitly,

$$t \to (1-t)^2|\psi_0\rangle\langle\psi_0| + t^2|\psi_1\rangle\langle\psi_1| + t(1-t)\left(|\psi_0\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_0|\right) . \tag{33}$$

If ψ_0 and ψ_1 are unit vectors, $\pi_j = |\psi_j\rangle\langle\psi_j|$ are (density operators of) pure states. Then (31) simplifies to

$$\mathrm{dist}_{\mathrm{FS}}(\pi_1, \pi_0) = \sqrt{2 - 2|\langle\psi_1, \psi_0\rangle|} = \sqrt{2 - 2\sqrt{\mathrm{Pr}(\pi_0, \pi_1)}} \tag{34}$$

where we have used the notation $\mathrm{Pr}(\pi_1, \pi_0)$ for the *transition probability*

$$\mathrm{Pr}(\pi_1, \pi_0) = \mathrm{tr}\, \pi_0\pi_1 \; . \tag{35}$$

The transition probability is the probability to get an affirmative answer in testing wether the system is in the state π_1 if it was actually in state π_0.

However, the geodesic arc (33) cuts the set of pure states only at π_0 and at π_1. Therefore, the distance (28) is not an inner one for the space of pure states. To obtain the appropriate distance, which we call

$$\mathrm{Dist}_{\mathrm{FS}} \; ,$$

we have to minimize the length with respect to curves consisting of pure states only.
This problem is quite similar to the change from Euclidean to spherical geometry in $\mathcal{H}$ (and, of course, in ordinary 3-space). We can use a great circle on the unit sphere of our Hilbert space which obeys the condition (25), $\langle\dot{\psi},\psi\rangle = 0$. Then the map $\psi \to |\psi\rangle\langle\psi| = \pi$ is one to one within small intervals of the parameter: The map identifies antipodes in the unit sphere of the Hilbert space. Thus "in the small" the map is one to one. Using this normalization we get

$$\mathrm{Dist}_{\mathrm{FS}}(\pi_0,\pi_1) = \arccos\sqrt{\mathrm{Pr}(\pi_0,\pi_1)}\ . \tag{36}$$

The distance of two pure states become maximal if π_0 and π_1 orthogonal. This is at the angle $\pi/2$. As in the unit sphere the geodesics are closed, but now have length π.
Remark: If "Dist" is multiplied by a positive real number, we get again a distance. (This is obviously so for any distance.) Therefore, another normalization is possible. Fubini and Study, who considered these geodesics at first, "stretched" them to become metrical isomorph to the unit circle:

$$\mathrm{Dist}_{\mathrm{Study}}(\pi_0,\pi_1) = 2\arccos\sqrt{\mathrm{Pr}(\pi_0,\pi_1)}.\ \text{(Study, 1904)}$$

1.7 Fubini-Study metric

As we have seen, with $\mathrm{dist}_{\mathrm{FS}}$ the set of positive operators of rank one becomes an inner metric space. We now convince ourselves that it is a Riemannian manifold. Its Riemannian metric, called *Fubini-Study metric,* reads

$$ds_{\mathrm{FS}} = \sqrt{\langle\psi,\psi\rangle\langle\dot{\psi},\dot{\psi}\rangle - \langle\dot{\psi},\psi\rangle\langle\psi,\dot{\psi}\rangle}\ dt \tag{37}$$

for curves

$$t \mapsto \psi_t \mapsto |\psi_t\rangle\langle\psi_t|\ , \tag{38}$$

where in (37) the index t in ψ_t is suppressed.

To prove it, we consider firstly normalized curves ψ_t remaining on the unit sphere of $\mathcal{H}$. Imposing the geometric phase transport condition (25), the map (38) becomes an isometry for small parameter intervals. Simultaneously (37) reduces to the Euclidean line element along curves fulfilling (25). Hence, for curves on the unit sphere, (37) has been proved.
To handle arbitrary normalization, we scale by

$$\psi'_t = z_t\psi_t, \quad z_t \neq 0\ , \tag{39}$$

and obtain

$$\langle\psi',\psi'\rangle\langle\dot{\psi}',\dot{\psi}'\rangle - \langle\dot{\psi}',\psi'\rangle\langle\psi',\dot{\psi}'\rangle = (z^*z)^2\ [\langle\psi,\psi\rangle\langle\dot{\psi},\dot{\psi}\rangle - \langle\dot{\psi},\psi\rangle\langle\psi,\dot{\psi}\rangle] \tag{40}$$

Therefore, (37) shows the correct scaling as required by $\mathrm{dist}_{\mathrm{FS}}$, and it is valid on the unit sphere. Thus, (37) is valid generally, i. e. for curves of positive operators of rank one.

We now express (37) in terms of states. If

$$\pi_t = |\psi_t\rangle\langle\psi_t|, \quad \langle\psi_t, \psi_t\rangle = 1, \qquad \langle\dot{\psi}_t, \psi_t\rangle = 0, \tag{41}$$

then one almost immediately see $\mathrm{tr}\,\dot{\pi}\dot{\pi} = 2\langle\dot{\psi}, \dot{\psi}\rangle$ or

$$ds_{\mathrm{FS}} = \sqrt{\frac{1}{2}\mathrm{tr}\;\dot{\pi}^2}\; dt \tag{42}$$

for all (regular enough) curves $t \to \pi_t$ of pure states.
These curves satisfy

$$\mathrm{tr}\,\pi = 1, \quad \mathrm{tr}\,\dot{\pi} = 0, \quad \mathrm{tr}\;\dot{\pi}\pi = 0\;. \tag{43}$$

The latter assertion follows from $\pi^2 = \pi$ by differentiation, $\dot{\pi} = \dot{\pi}\pi + \pi\dot{\pi}$, and by taking the trace of both sides.

Let now $\rho = |\psi\rangle\langle\psi|$ with $\psi = \psi_t$ a curve somewhere in the Hilbert space. We lost normalization of ψ, but we are allowed to require the vanishing of the gauge potential (27). Then

$$\mathrm{tr}\dot{\rho} = 2\langle\psi, \dot{\psi}\rangle, \quad \mathrm{tr}\dot{\rho}^2 = 2\langle\psi, \psi\rangle\;\langle\dot{\psi}, \dot{\psi}\rangle + \frac{1}{2}(\mathrm{tr}\dot{\rho})^2$$

and we conclude

$$ds^2_{\mathrm{FS}} = \frac{1}{2}[\mathrm{tr}\;\dot{\rho}^2 - (\mathrm{tr}\;\dot{\rho})^2]\,dt^2 \tag{44}$$

for curves $t \to \rho_t$ of positive operators of rank one.

There is a further expression for the Fubini-Study metric. It is $\rho^2 = (\mathrm{tr}\rho)\rho$ for a positive operator of rank one. From this, by differentiating and some algebraic manipulations, one arrives at

$$ds^2_{\mathrm{FS}} = [(\mathrm{tr}\rho)^{-1}\mathrm{tr}\;\rho\dot{\rho}^2 - (\mathrm{tr}\;\dot{\rho})^2]\,dt^2\;. \tag{45}$$

1.8 Symmetries

It is a famous idea of Wigner, [22], to use the transition probability to define the concept of symmetry in the set of pure states. If $\pi \to T(\pi)$ maps the set of pure states onto itself, T is a *symmetry* if it satisfies

$$\mathrm{Pr}(T(\pi_1), T(\pi_2)) = \mathrm{Pr}(\pi_1, \pi_2)\;. \tag{46}$$

Looking at (34) or (30) it becomes evident that (46) is valid if and only if T is an isometry with respect to $\mathrm{dist}_{\mathrm{FS}}$ and also to $\mathrm{Dist}_{\mathrm{FS}}$.

Before stating the two main results, we discuss the case $\dim \mathcal{H} = 2$, the "one-qubit-space". Here, the pure states are uniquely parameterized by the point of a 2-sphere, the Bloch sphere. Indeed, π is a pure state exactly if

$$\pi = \frac{1}{2}\left(\mathbf{1} + \sum_{j=1}^{3} x_j \sigma_j\right), \quad x_1^2 + x_2^2 + x_3^2 = 1 \tag{47}$$

with a "Bloch vector" $\{x_1, x_2, x_3\}$. Clearly,

$$\operatorname{tr} \dot{\pi}^2 = \frac{1}{2} \sum \dot{x}_j^2$$

and, by (42),

$$ds_{\mathrm{FS}} = \frac{1}{2} \sqrt{\sum \dot{x}_j^2}\, dt\ . \tag{48}$$

It follows already, that a symmetry, T, in the sense of Wigner, is a map of the 2-sphere into itself conserving the metric induced on the sphere by the Euclidean one. Hence there is an orthogonal matrix with entries O_{jk} changing the Bloch vector as

$$\pi \to T(\pi) \ \Leftrightarrow\ x_j \to \sum_k O_{jk} x_k\ . \tag{49}$$

As is well known, proper orthogonal transformations can be implemented by a unitary transformation, i. e., with a suitable unitary U,

$$U(\sum_j x_j \sigma_j) U^{-1} = \sum_j x'_j \sigma_j\ .$$

An anti-unitary, say V, can be written $V = U\theta_f$ with a spin-flip

$$\theta_f(c_0|0\rangle + c_1|1\rangle) = (c_1^*|0\rangle - c_0^*|1\rangle)$$

producing the inversion $x_j \to -x_j$ of the Bloch sphere. This says, in short, that

$$T(\pi) = V\pi V^{-1}, \quad V \text{ either unitary or anti-unitary.} \tag{50}$$

Wigner has proposed the validity of (50) for all Hilbert spaces. And he was right.

There is a stronger result[6] for $\dim \mathcal{H} > 2$ saying that it suffices that T preserves orthogonality. To put it together:
Theorem: In order that (50) holds for all pure states π it is necessary and sufficient that one of the following conditions take place:
a) It is a symmetry in the sense (46) of Wigner.
b) It is an isometry of the Study-Fubini distance.
c) It is $\dim \mathcal{H} \geq 3$ and

[6] The 1-qubit case is too poor in structure compared with higher dimensional ones.

$$\Pr(\pi_1, \pi_2) = 0 \Leftrightarrow \Pr(T(\pi_1), T(\pi_2)) = 0\,. \tag{51}$$

If $\dim \mathcal{H} > 2$, the condition c) is an obviously more advanced statement than a) or b). An elementary proof is due to Uhlhorn, [58]. Indeed, the theorem is also a corollary to deeper rooted results of Dye, [38].

1.9 Comparison with other norms

While for the vectors of a Hilbert space one has naturally only one norm, the vector norm, there are many norms to estimate an Operator, say A. For instance one defines

$$\| A \|_2 = \sqrt{\operatorname{tr} A^\dagger A}, \quad \| A \|_1 = \operatorname{tr} \sqrt{A^\dagger A}\,. \tag{52}$$

The first one is called *Frobenius* or *von Neumann norm.* The second is the *functional-* or *1-norm.* If $\mathcal{H} = \infty$, these norms can be easily infinite and their finiteness is a strong restriction to the operator. If A is of finite rank, then

$$\| A \|_2 \leq \| A \|_1 \leq \sqrt{\operatorname{rank}(A)}\ \| A \|_2\ . \tag{53}$$

The rank of A is at most as large as the dimension of the Hilbert space.

For $r \geq 1$ one also defines the *Schatten norms*

$$\| A \|_r = (\operatorname{tr}\,(A^\dagger A)^{r/2})^{1/r}\,. \tag{54}$$

If π a pure state's density operator then $\| \pi \|_r = 1$ always. The Schatten norms of the difference $\nu = \pi_2 - \pi_1$ is also easily computed. One may assume $\dim \mathcal{H} = 2$ as all calculations are done in the space spanned by the vectors ψ_j with $\pi_j = |\psi_j\rangle\langle\psi_j|$. Now ν is Hermitian and with trace 0, its square is a multiple of $\mathbf{1}$. We get

$$\lambda^2 \mathbf{1} = \nu^2 = \pi_1 + \pi_2 - \pi_1\pi_2 - \pi_2\pi_1$$

and, taking the trace, $\lambda^2 = 1 - \Pr(\pi_1, \pi_2)$, by (35). Thus

$$\| \pi_2 - \pi_1 \|_r = 2^{1/r} \sqrt{1 - \Pr(\pi_1, \pi_2)}\,. \tag{55}$$

Comparing with

$$\operatorname{dist}_{FS}(\pi_1, \pi_2) = \sqrt{2}\sqrt{1 - |\langle\psi_1, \psi_2\rangle|} = \sqrt{2}\sqrt{1 - \sqrt{\Pr(\pi_1, \pi_2)}}$$

results in

$$\| \pi_2 - \pi_1 \|_r = \frac{2^{1/r}}{\sqrt{2}}\, \operatorname{dist}_{\mathrm{FS}}(\pi_1, \pi_2)\, \sqrt{1 + \sqrt{\Pr(\pi_1, \pi_2)}}\,. \tag{56}$$

As the value of transition probability is between 0 and 1, the identity provides tight inequalities between Schatten distances and the Fubini-Study distance for two pure states.

One important difference between the Schatten distances (55) and the Fubini-Study one concerns the geodesics. We know the geodesics with respect of a norm read

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

and, therefore, they consist of mixed density operators for $0 < t < 1$. The Study-Fubini geodesics, however, do not leave the set of pure states, an important aspect.

2 Operators, observables, and states

Let us fix some notions. We denote the algebra of all bounded linear operators, acting on an Hilbert space $\mathcal{H}$, by $\mathcal{B}(\mathcal{H})$.

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

$$\| A \|_\infty = \sup_\psi \| A\psi \|, \quad \| \psi \| = 1 \tag{57}$$

and calls A bounded if this sup over all unit vectors is finite. To be bounded means that the operator cannot stretch unit vectors to arbitrary length. One has

$$\lim_{r\to\infty} \| A \|_r = \| A \|_\infty \tag{58}$$

if the Schatten norms are finite for large enough r. The "∞-norm" (57) of every unitary operator and of every projection operator (different from the operator $\mathbf{0}$) is one.

(57) is an "operator norm" because one has

$$\| AB \|_\infty \leq \| A \|_\infty \, \| B \|_\infty$$

in addition to the usual norm properties. For $1 < r < \infty$ no Schatten norm is an operator norm. On the other hand, there are many operator norms. However, among them the ∞-norm has a privileged position. It satisfies

$$\| A^\dagger A \| = \| A \|^2, \quad \| A^\dagger \| = \| A \| \ . \tag{59}$$

An operator norm, satisfying (59), is called a C*-norm. *There is only one C^*-norm in $\mathcal{B}(\mathcal{H})$, the ∞-norm.*

Remark: In mathematics and in mathematical physics the operation $A \to A^\dagger$ is called "the star operation": In these branches of science the Hermitian adjoint of an operator A is called A^*. The notion $A^\dagger$ has been used by Dirac in his famous book "The Principles of Quantum Mechanics", [7].

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 trace one:

$$\omega \geq \mathbf{0}, \quad \mathrm{tr}\,\omega = 1\ . \tag{60}$$

One can prove

$$\| \rho \|_1 = \mathrm{tr}\ \rho = 1 \Leftrightarrow \rho \text{ is a density operator.} \tag{61}$$

A bounded operator on an infinite dimensional Hilbert space is said to be of "trace class" if its 1-norm is finite. The trace class operators constitute a tiny portion of $\mathcal{B}(\mathcal{H})$ in the infinite case.

2.1 States and expectation values

Let ω be a density operator and $A \in \mathcal{B}(\mathcal{H})$ an operator. The value $\mathrm{tr}\,A\omega$ is called "expectation value of A in state ω". There are always operators with different expectation values for two different density operators. In this sense one may say: "Observables distinguish states".

Remark: Not every operator in $\mathcal{B}(\mathcal{H})$ represents an observable in the strict sense: An observable should have a spectral decomposition. Therefore, observables are represented by normal operators, i. e. $A^\dagger A = AA^\dagger$ must be valid. (For historical but not physical reasons, often hermiticity or, if $\dim \mathcal{H} = \infty$, self-adjointness is required in textbooks. A critical overview is in [45].) On the other hand, to distinguish states, the expectation values of projection operators are sufficient.

As already said, 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 to describe a state in a different way, namely as the set of its expectation values. To do so, one consider a state as a function defined for all operators. In particular, if ω is a density operator, one considers the function (or "functional", or "linear form")

$$A \to \underline{\omega}(A) := \mathrm{tr}\,A\omega\ . \tag{62}$$

Let us stress the following properties of (62)
1) Linearity: $\underline{\omega}(c_1A_1 + c_2A_2) = c_1\underline{\omega}(A_1) + c_2\underline{\omega}(A_2)$
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) the essential conditions and calls *every* functional on $\mathcal{B}(\mathcal{H})$ which fulfils these three conditions a *state of the algebra* $\mathcal{B}(\mathcal{H})$.

In 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.

Let us see, how it works. If $\dim \mathcal{H} < \infty$, every functional obeying 1), 2), and 3) can be written

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

as in (62). Here the definition just reproduces the density operators.

Indeed, every linear form can be written $\underline{\omega}(A) = \operatorname{tr} BA$ with an operator $B \in \mathcal{B}(\mathcal{H})$. However, if $\operatorname{tr} BA$ is a real and non-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 identifies $\omega := B$.

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 unit element,* i.e. decompositions

$$\mathbf{1} = \sum_j P_j \tag{63}$$

with projection operators P_j. These decompositions are necessarily orthogonal, $P_k P_l = \mathbf{0}$ if $k \neq l$, and in one-to-one relation to decompositions of the Hilbert space into orthogonal sums,

$$\mathcal{H} = \bigoplus_j \mathcal{H}_j, \quad \mathcal{H}_j = P_j \mathcal{H} \, . \tag{64}$$

(If a sum of projections is a projection, it must be an orthogonal sum. To see it, square the equation and take the trace. The trace of a product of two positive operators is not negative and can be zero only if the product of the operators is zero.)

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

$$\sum_j \underline{\omega}(P_j) = \underline{\omega}(\mathbf{1}) = 1 \tag{65}$$

is valid. $\underline{\omega}$ is normal exactly if its expectation values are given as in (62) with the help of a density operator ω.

There is a further class of states, the *singular states.* A state $\underline{\omega}$ of $\mathcal{B}(\mathcal{H})$ is called "singular", if $\underline{\omega}(P) = 0$ 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

$$\underline{\omega} = (1-p)\underline{\omega}_{\text{normal}} + p\underline{\omega}_{\text{singular}}, \quad 0 \leq p \leq 1 \, . \tag{66}$$

In mathematical measure theory a general $\underline{\omega}$ corresponds to an "additive measure", in contrast to the genuine measures which are countably additive. Accordingly we are invited to consider a normal state of $\mathcal{B}(\mathcal{H})$ to be a "countably additive non-commutative probability measure", and any other state to be an "additive non-commutative probability measure".

I cannot but at this point of my lecture to mention the 1957 contribution of Gleason, [44]. He asked whether it will be possible to define states already by their expectation values at projections.

Assume $P \to f(P) \geq 0$, $f(\mathbf{1}) = 1$, is a function which is defined only on the projection operators $P \in \mathcal{B}(\mathcal{H})$ and which satisfies

$$\sum_j f(P_j) = 1 \tag{67}$$

for all orthogonal partitions (63) of the unity $\mathbf{1}$. Gleason has proved: If[7] $\dim \mathcal{H} > 2$, there is a density operator ω with $\operatorname{tr} P\omega = f(P)$ for all $P \in \mathcal{B}(\mathcal{H})$, i.e. $\underline{\omega}(P) = f(P)$.

The particular merit of Gleason's theorem consists in relating directly quantum probabilities to the concept of "state" as defined above: Suppose our quantum system is in state ω, and we test whether P is valid, the answer is YES with probability $\underline{\omega}(P) = \operatorname{tr} P\omega$.

It lasts about 30 years to find out what is with general states. There is a lengthy proof by Maeda, Christensen, Yeadon, and others, see [51], with a lot of (mostly not particular difficult) steps and with a rich architecture. Indeed, they examined the problem for general von Neumann algebras, but in the case at hand they assert the following extension of Gleason's finding.

Theorem: Assume $\dim \mathcal{H} \geq 3$. Given a function $f \geq 0$ on the projection operators satisfying (65) for all *finite* partitions of $\mathbf{1}$. Then there is a state $\underline{\omega}$ fulfilling $\underline{\omega}(P) = f(P)$ for all projection operators of $\mathcal{B}(\mathcal{H})$.

2.2 Subalgebras and subsystems

There is a consistent solution to the question: What is a subsystem of a quantum system with Hilbert space $\mathcal{H}$ and algebra $\mathcal{B}(\mathcal{H})$? The solution is unique in the finite dimensional case. Below we list some necessary requirements which become sufficient if $\dim \mathcal{H} < \infty$. As already indicated, a subsystem of a quantum system should consist of less observables (operators) than the larger one. For the larger one we start with $\mathcal{B}(\mathcal{H})$ to be on (more or less) known grounds.

Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a subset.
a) If $\mathcal{A}$ is a linear space and
b) if $A, B \in \mathcal{A}$ then $AB \in \mathcal{A}$,
$\mathcal{A}$ is called a *subalgebra* of $\mathcal{B}(\mathcal{H})$ or, equivalently, an *operator algebra* on $\mathcal{H}$. Essential is also the condition:

[7] As already said, in two dimensions the set of projections is too poor in relations.

c) If $A \in \mathcal{A}$ then $A^\dagger \in \mathcal{A}$.

A subset $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ satisfying a), b), and c) is called an *operator *-algebra.* In an operator algebra the scalar product of the Hilbert space is reflected by the star operation, $A \to A^\dagger$. A further point to mention concerns positivity of operators: $B \in \mathcal{B}(\mathcal{H})$ is positive if and only if it can be written $B = A^\dagger A$.

Finally, an algebra $\mathcal{A}$ is called *unital* if it contains an *identity* or *unit element*, say $\mathbf{1}_\mathcal{A}$. The unit element, if it exists, is uniquely characterized by

$$\mathbf{1}_\mathcal{A}\ A = A\ \mathbf{1}_\mathcal{A} = A \ \text{ for all } \ A \in \mathcal{A} \tag{68}$$

and we refer to its existence as
d) $\mathcal{A}$ is unital.

Assume $\mathcal{A}$ fulfils all four conditions a) to d). Then one can introduce the concept of "state". We just mimic what has been said to be a state of $\mathcal{B}(\mathcal{H})$ and obtain a core definition.

→ A *state of* $\mathcal{A}$ is a function $A \to \underline{\omega}(A) \in \mathbb{C}$ of the elements of $\mathcal{A}$ satisfying for all elements of $\mathcal{A}$
1') $\underline{\omega}(c_1A_1 + c_2A_2) = c_1\underline{\omega}(A_1) + c_2\underline{\omega}(A_2)$, "linearity",
2') $\underline{\omega}(A^\dagger A) \geq 0$, "positivity",
3') $\underline{\omega}(\mathbf{1}_\mathcal{A}) = 1$, "normalization".

Let us stop for a moment to ask what has changed. The change is in 2) to 2'). In 2') no reference is made to the Hilbert space. It is a purely algebraic definition which only refers to operations defined in $\mathcal{A}$. It circumvents the way, A is acting on $\mathcal{H}$. That implies: The concept of state does *not* depend how $\mathcal{A}$ is embedded in $\mathcal{B}(\mathcal{H})$, or "at what place $\mathcal{A}$ is sitting within a larger *-algebra". Indeed, to understand the abstract skeleton of the quantum world, one is confronted with (at least!) two questions:
→ What is a quantum system, what is its structure?
→ How is a quantum system embedded in other ones as a subsystem?
Now let us proceed more prosaic.

$A, B \to \underline{\omega}(A^\dagger B)$ is a positive Hermitian form. Therefore,

$$\underline{\omega}(A^\dagger A)\ \underline{\omega}(B^\dagger B) \geq \underline{\omega}(A^\dagger B), \tag{69}$$

which is the important Schwarz inequality.

The set of all states of $\mathcal{A}$ is the *state space of* $\mathcal{A}$. It will be denoted by $\underline{\Omega}(\mathcal{A})$. The state space is naturally convex[8]:

$$\underline{\omega} := \sum p_j \underline{\omega}_j \in \underline{\Omega}(\mathcal{A}) \tag{70}$$

for any convex combination of the $\underline{\omega}_j$, i.e. for all these sums with

[8] For more about convexity see [18, 3].

$$\sum p_j = 1 \text{ and } p_j > 0 \text{ for all } j \; . \tag{71}$$

A *face* of $\underline{\Omega}(\mathcal{A})$ is a subset with the following property: If $\underline{\omega}$ is contained in this subset then for every convex decomposition (70), (71), of $\underline{\omega}$ also all states $\underline{\omega}_j$ belong to this subset.

Main example: Let $P \in \mathcal{A}$ be a projection. Define $\underline{\Omega}(\mathcal{A})_P$ to be the set of all $\underline{\omega} \in \underline{\Omega}(\mathcal{A})$ such that $\underline{\omega}(P) = 1$. This set is a face of $\underline{\Omega}(\mathcal{A})$.
To see it, one looks at the definition of states and concludes from (70) and (71) that $\underline{\omega}_j(P) = 1$ necessarily.

Statement: If $\mathcal{A}$ is a *-subalgebra of $\mathcal{B}(\mathcal{H})$ and $\dim \mathcal{H} < \infty$, then every face of $\underline{\Omega}(\mathcal{A})$ is of the form $\underline{\Omega}(\mathcal{A})_P$ with a projection $P \in \mathcal{A}$.

$\rightarrow$ If a face consists of just one state $\underline{\pi}$, then $\underline{\pi}$ is called *extremal in* $\underline{\Omega}(\mathcal{A})$. This is the mathematical definition. In quantum physics a state $\underline{\pi}$ of $\mathcal{A}$ is called *pure* if and only if $\underline{\pi}$ is extremal in $\underline{\Omega}(\mathcal{A})$.

These are rigorous and fundamental definitions. We do not assert that every $\mathcal{A}$ satisfying the requirements a) to d) above represents or "is" a quantum system. But we claim that every quantum system, which can be represented by bounded operators, can be based on such an algebra. Its structure gives simultaneously meaning to the concepts of "observable", "state", and "pure state". It does so in a clear and mathematical clean way.

Subsystems

Now we consider some relations between operator algebras, in particular between quantum systems. We start by asking for the concept of "subsystems" of a given quantum system. Let $\mathcal{A}_j$ be *-subalgebras of $\mathcal{B}(\mathcal{H})$ with unit element $\mathbf{1}_j$ respectively. From $\mathcal{A}_1 \subset \mathcal{A}_2$ it follows $\mathbf{1}_1\mathbf{1}_2 = \mathbf{1}_1$ and $\mathbf{1}_1$ is a projection in $\mathcal{A}_2$. To be a *subsystem* of the quantum system $\mathcal{A}_2$ we require

$$\mathcal{A}_1 \subset \mathcal{A}_2, \quad \mathbf{1}_1 = \mathbf{1}_2 \; . \tag{72}$$

In mathematical terms, $\mathcal{A}_1$ is a *unital subalgebra* of $\mathcal{A}_2$. Thus, if two quantum systems are represented by two unital *-algebras $\mathcal{A}_j$ satisfying (72), then $\mathcal{A}_1$ is said to be a *subsystem* of $\mathcal{A}_2$.

In particular, $\mathcal{A}$ is a subsystem of $\mathcal{B}(\mathcal{H})$ if it contains the identity operator, $\mathbf{1}_\mathcal{H}$ or simply $\mathbf{1}$ of $\mathcal{H}$ because $\mathbf{1}$ is the unit element of $\mathcal{B}(\mathcal{H})$.

The case $\mathbf{1}_1 \neq \mathbf{1}_2$ will be paraphrased by calling $\mathcal{A}_1$ an *incomplete subsystem* of $\mathcal{A}_2$.

Let $\mathcal{A}_1$ be a subsystem of $\mathcal{A}_2$ and let us ask for relations between their states. At first we see: A state $\underline{\omega}_2 \in \underline{\Omega}(\mathcal{A}_2)$ gives to us automatically a state $\underline{\omega}_1$ on $\mathcal{A}_1$ by just defining $\underline{\omega}_1(A) := \underline{\omega}_2(A)$ for all operators of $\mathcal{A}_1$. $\underline{\omega}_1$ is called the *restriction of* $\underline{\omega}_2$ *to* $\mathcal{A}_1$. Clearly, the conditions 1') to 3') remain valid in restricting a state to a subsystem.

Of course, it may be that there are many states in $\mathcal{A}_2$ with the same restriction to $\mathcal{A}_1$. Two (and more) different states of $\mathcal{A}_2$ may "fall down" to one and the same state of the subalgebra $\mathcal{A}_1$. From the point of view of a subsystem two or more different states of a larger system can become identical.

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

As a consequence, we associate to the words "a quantum system is a subsystem of another quantum system" a precise meaning. Or, to be more cautious, we have a necessary condition for the validity of such a relation. Imaging that every system might be a subsystem of many other ones, one get a faint impression how rich the architecture of that hierarchy may be.

Notice: The restriction of a state to an incomplete subsystem will conserve the linearity and the positivity conditions 1') and 2'). The normalization 3') cannot be guaranteed in general.

$\dim \mathcal{H} = \infty$. Some comments

As a matter of fact, the conditions a) to c) for a *-subalgebra of $\mathcal{B}(\mathcal{H})$ are not strong enough for infinite dimensional Hilbert spaces. There are two classes of algebras in the focus of numerous investigations, the C*- and the von Neumann algebras. We begin by defining[9] C*-algebras and then we turn to von Neumann ones. Much more is in [10].

Every subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ is equipped with the ∞-norm, $\| \, . \, \|_\infty$. One requires the algebra to be closed[10] with respect to that norm: For every sequence $A_j \in \mathcal{A}$ which converges to $A \in \mathcal{B}(\mathcal{H})$ in norm, $\| A - A_j \|_\infty \to 0$, the operator A must be in $\mathcal{A}$ also. In particular, a *-subalgebra is said to be a *C*-algebra* if it is closed with respect to the operator norm. The ∞-norm is a C*-norm in these algebras, see (59). One can prove that in a C*-algebra there exists just one operator norm which is a C*-norm.

In the same spirit there is an 1-norm (or "functional norm") $\| \, . \, \|_1$, estimating the linear functionals of $\mathcal{A}$. $\| \underline{\nu} \|_1$ is the smallest number λ for which $|\underline{\nu}(A)| \leq \lambda \, \| A \|_\infty$ is valid for all $A \in \mathcal{A}$.

With respect to a unital C*-algebra we can speak of its states and its normal operators are its observables. However, a C*-algebra does not necessarily provide sufficiently many projection operators: There are C*-algebras containing no projection different from the trivial ones, $\mathbf{0}$ and $\mathbf{1}_{\mathcal{A}}$.

In contrast, von Neumann algebras contain sufficient many projections. $\mathcal{A}$ is called a *von Neumann algebra,* if it is closed with respect to the so-called weak topology.

[9] We define the so-called "concrete" C*-algebras.
[10] Then the algebra becomes a "Banach algebra".

To explain it, let $\mathcal{F}$ be a set of operators and $B \in \mathcal{B}(\mathcal{H})$. B is a "weak limit point" of $\mathcal{F}$ if for every n, every $\epsilon > 0$, and for every finite set $\psi_1, \ldots, \psi_n$ of vectors from $\mathcal{H}$ there is an operator $A \in \mathcal{F}$ fulfilling the inequality

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

The set of all weak limit points of $\mathcal{F}$ is the "weak closure" of $\mathcal{F}$.

A von Neumann algebra $\mathcal{A}$ is a *-subalgebra of $\mathcal{B}(\mathcal{H})$ which contains all its weak limit points. In addition one requires to every unit vector $\psi \in \mathcal{H}$ an operator $A \in \mathcal{A}$ with $A\psi \neq 0$.
Because of the last requirement, the notion of a von Neumann algebra is defined relative to $\mathcal{H}$. (If $\mathcal{A}$ is just weakly closed, then there is a subspace, $\mathcal{H}_0 \subset \mathcal{H}$, relative to which $\mathcal{A}$ is von Neumann.)

J. von Neumann could give a purely algebraic definition of the algebras carrying his name. It is done with the help of commutants. For a subset $\mathcal{F} \subset \mathcal{B}(\mathcal{H})$ the *commutant*, $\mathcal{F}'$, of $\mathcal{F}$ is the set of all $B \in \mathcal{B}(\mathcal{H})$ commuting with all $A \in \mathcal{F}$. The commutant of a set of operators is always a unital and weakly closed subalgebra of $\mathcal{B}(\mathcal{H})$.

If $\mathcal{F}^\dagger = \mathcal{F}$, i. e. $\mathcal{F}$ contains with A always also $A^\dagger$, its commutant $\mathcal{F}'$ becomes a unital *-algebra which, indeed, is a von Neumann algebra.

But then also the double commutant $\mathcal{F}''$, the commutant of the commutant, is a von Neumann algebra. Even more, von Neumann could show:
$\mathcal{A}$ is a von Neumann algebra if and only if $\mathcal{A}'' = \mathcal{A}$.

We need one more definition. The *center* of an algebra consists of those of its elements which commute with every element of the algebra. The center of $\mathcal{A}$ is in $\mathcal{A}'$ and vice versa. We conclude

$$\mathcal{A} \cap \mathcal{A}' = \text{ center of } \mathcal{A} \; . \tag{73}$$

If $\mathcal{A}$ is a von Neumann algebra, $\mathcal{A} \cap \mathcal{A}'$ is the center of both, $\mathcal{A}$ and $\mathcal{A}'$.

A von Neumann algebra is called a *factor* if its center consists of the multiples of $\mathbf{1}$ only. Thus, a factor may be characterized by

$$\mathcal{A} \cap \mathcal{A}' = \mathbb{C}\mathbf{1} \; . \tag{74}$$

2.3 Classification of finite quantum systems

There are two major branches in group theory, the groups themselves and their representations. We have a similar situation with quantum systems if they are seen as operator algebras: There is a certain *-algebra and its concrete realizations as operators on a Hilbert space. However, at least in finite dimensions, our task is much easier than in group theory.

Wedderburn [21] has classified all finite dimensional matrix algebras[11], or, what is equivalent, all subalgebras of $\mathcal{B}(\mathcal{H})$ if $\dim \mathcal{H} < \infty$. Here we report and comment his results for *-subalgebras only. (These results could also be read of from the classification of factors by Murray and von Neumann[12], see [10], section III.2 .)

One calls two *-algebras,

$$\mathcal{A} \subset \mathcal{B}(\mathcal{H}) \text{ and } \tilde{\mathcal{A}} \subset \mathcal{B}(\tilde{\mathcal{H}}) \tag{75}$$

*-*isomorph* if there is a map Ψ from $\mathcal{A}$ onto $\tilde{\mathcal{A}}$,

$$A \mapsto \Psi(A) = \tilde{A} \in \tilde{\mathcal{A}}, \quad A \in \mathcal{A} \,,$$

satisfying
A) $\Psi(c_1 A_1 + c_2 A_2) = c_1 \Psi(A_1) + c_2 \Psi(A_2)$,
B) $\Psi(AB) = \Psi(A)\,\Psi(B)$,
C) $\Psi(A^\dagger) = \Psi(A)^\dagger$,
D) $A \neq B \;\Rightarrow\; \Psi(A) \neq \Psi(B)$,
E) $\Psi(\mathcal{A}) = \tilde{\mathcal{A}}$.

The first three conditions guarantee the conservation of all algebraic relations under the map Ψ. From them it follows the positivity of the map Ψ because an element of the form $A^\dagger A$ is mapped to $\tilde{A}^\dagger \tilde{A}$.
Condition E) says that $\mathcal{A}$ is mapped *onto* $\tilde{\mathcal{A}}$, i. e. every $\tilde{A}$ can be gained as $\Psi(A)$. It follows that the unit element of $\mathcal{A}$ is transformed into that of $\tilde{\mathcal{A}}$.
Condition D) now shows that Ψ is invertible because to every $A \in \mathcal{A}$ there is exactly one $\tilde{A}$ with $\Psi(A) = \tilde{A}$.

If only A) to D) is valid, Ψ maps $\mathcal{A}$ *into* $\tilde{\mathcal{A}}$. Replacing E) by
E'): $\Psi(\mathcal{A}) \subset \tilde{\mathcal{A}}$
and requiring A) to D) defines an *embedding of* $\mathcal{A}$ *into* $\tilde{\mathcal{A}}$.

If $\mathcal{A} \to \Psi(\mathcal{A}) \subseteq \mathcal{B}(\mathcal{H})$ is an embedding of $\mathcal{A}$, the embedding is also said to be a *-*representation* of $\mathcal{A}$ as an operator algebra.
A "unital *-representation" of $\mathcal{A}$ maps $\mathbf{1}_{\mathcal{A}}$ to the identity operator of $\mathcal{H}$.

Important examples of unital *-representations and *-isomorphisms of $\mathcal{B}(\mathcal{H})$ are given by "matrix representations". Every ortho-normal basis $\psi_1, \psi_2, \ldots \psi_n$ of the Hilbert space $\mathcal{H}$, $\dim \mathcal{H} = n$, induces via the map

$$A \to \text{matrix}\, \mathbf{A} = \{A_{ij}\} \text{ with matrix elements } A_{ij} = \langle \psi_i, A\psi_j \rangle, \; A \in \mathcal{B}(\mathcal{H}),$$

a unital *-isomorphism between $\mathcal{B}(\mathcal{H})$ and the algebra $M_n(\mathbb{C})$, of complex $n \times n$ matrices. If $\dim \mathcal{H} = \infty$, however, matrix representations are a difficult matter.

---

[11] He extends the Jordan form from matrices to matrix algebras.
[12] Von Neumann and Murray introduced and investigated von-Neumann algebras in a famous series of papers on "Rings of operators" [52].

Direct product and the direct sum constructions

Let us review some features of direct products. We start with

$$\mathcal{H} = \mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{B}} \,. \tag{76}$$

The algebra $\mathcal{B}(\mathcal{H}^{\mathrm{A}})$ is not a subalgebra of $\mathcal{B}(\mathcal{H})$, but it becomes one by

$$\mathcal{B}(\mathcal{H}^{\mathrm{A}}) \mapsto \mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathbf{1}^{\mathrm{B}} \subset \mathcal{B}(\mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{B}}) \,: \tag{77}$$

Here "$\mapsto$" points to the unital embedding

$$A \in \mathcal{B}(\mathcal{H}^{\mathrm{A}}) \mapsto A \otimes \mathbf{1}^{\mathrm{B}} \in \mathcal{B}(\mathcal{H}) \tag{78}$$

of $\mathcal{B}(\mathcal{H}^{\mathrm{A}})$ into $\mathcal{B}(\mathcal{H})$. It is a *-isomorphism from the algebra $\mathcal{B}(\mathcal{H}^{\mathrm{A}})$ onto $\mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathbf{1}^{\mathrm{B}}$. Similarly, $\mathcal{B}(\mathcal{H}^{\mathrm{B}})$ is *-isomorph to $\mathbf{1}^{\mathrm{A}} \otimes \mathcal{B}(\mathcal{H}^{\mathrm{B}})$ and embedded into $\mathcal{B}(\mathcal{H})$ as a *-subalgebra. $\mathbf{1}^{\mathrm{A}} \otimes \mathcal{B}(\mathcal{H}^{\mathrm{B}})$ is the commutant of $\mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathbf{1}^{\mathrm{B}}$ and vice vera. Based on $A \otimes B = (A \otimes \mathbf{1}^{\mathrm{B}})\,(\mathbf{1}^{\mathrm{A}} \otimes B)$ there is the identity

$$\mathcal{B}(\mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{B}}) = \mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathcal{B}(\mathcal{H}^{\mathrm{B}}) = (\mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathbf{1}^{\mathrm{B}})\,(\mathbf{1}^{\mathrm{A}} \otimes \mathcal{B}(\mathcal{H}^{\mathrm{B}}))\,. \tag{79}$$

The algebras of $\mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathbf{1}^{\mathrm{B}}$ and $\mathbf{1}^{\mathrm{A}} \otimes \mathcal{B}(\mathcal{H}^{\mathrm{B}})$ are not only subalgebras, but also factors. In finite dimensions every von Neumann factor on $\mathcal{H}$ is of that structure:
If $\mathcal{A}$ is a sub-factor of $\mathcal{B}(\mathcal{H})$ and $\dim \mathcal{H} < \infty$ then there is a decomposition (76) such that $\mathcal{A} = \mathcal{B}(\mathcal{H}^{\mathrm{A}}) \otimes \mathbf{1}^{\mathrm{B}}$.

It is worthwhile to notice the *information* contained in an embedding of $\mathcal{B}(\mathcal{H}^{\mathrm{A}})$ into $\mathcal{B}(\mathcal{H})$: We need a definite decomposition (76) of $\mathcal{H}$ into a direct product of Hilbert spaces with correct dimensions of the factors. Most unitary transformations of $\mathcal{H}$ would give another possible decomposition of the form (77) resulting in another embedding (77). Generally speaking, distinguishing a subsystem of a quantum system enhance our knowledge and can be well compared with the information gain by a measurement.

One knows how to perform direct sums of linear spaces. To apply it to algebras one has to say how the multiplication between direct summands is working. Indeed, it works in the most simple way:

$\mathcal{A}$ is the *direct sum* of its subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_m$ if every $A \in \mathcal{A}$ can be written as a sum

$$A = A_1 + \ldots + A_m, \quad A_j \in \mathcal{A}_j \tag{80}$$

and the multiplication obeys

$$A_j A_k = \mathbf{0} \;\text{ whenever }\; j \neq k \,. \tag{81}$$

One can rewrite the direct sum construction in block matrix notation. Let us illustrate it for the case $m = 3$.

$$A = A_1 + A_2 + A_3 = \begin{pmatrix} A_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_3 \end{pmatrix} , \quad A_j \in \mathcal{A}_j , \tag{82}$$

is the "block matrix representation" of the direct sum. If one considers, say $\mathcal{A}_2$, as an algebra in its own right, its embedding into $\mathcal{A}$ is given by

$$A_2 \hookrightarrow \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} . \tag{83}$$

In contrast to the direct product construction the embedding (83) is not a unital one. (82) illustrate the two ways to direct sums: Either an algebra $\mathcal{A}$ can be decomposed as in (80), (81), or there are algebras $\mathcal{A}_j$ and we build up $\mathcal{A}$ by a direct sum construction out of them. In the latter case one writes

$$\mathcal{A} = \mathcal{A}_1 \oplus \ldots \oplus \mathcal{A}_m .$$

We shall use both possibilities below.

Types

Our aim is to characterize invariantly the set of *-isomorphic finite von Neumann algebras and to choose in it distinguished ones. The restriction to finite dimensions make the task quite simple:
Any *-subalgebra of $\mathcal{B}(\mathcal{H})$ is *-isomorph to a direct sum of factors.

Let $\mathbf{d}$ be a set of natural numbers,

$$\mathbf{d} = \{d_1, \ldots, d_m\}, \quad |\mathbf{d}| = \sum d_j . \tag{84}$$

The number m is called the *length of* $\mathbf{d}$.

We say that $\mathbf{d}' = \{d'_1, \ldots, d'_m\}$ is *equivalent* to $\mathbf{d}$ and we write $\mathbf{d} \sim \mathbf{d}'$ if the numbers d'_j are a permutation of the d_j. Exactly if this takes place, i. e. if $\mathbf{d} \sim \mathbf{d}'$, we say that $\mathbf{d}$ is *is of the same type* as $\mathbf{d}'$.

Given $\mathbf{d}$ as in (84) and Hilbert spaces $\mathcal{H}_j$ of dimensions $\dim \mathcal{H}_j = d_j$, we consider

$$\mathcal{B}_{\mathbf{d}} = \mathcal{B}_{d_1,\ldots,d_m} := \mathcal{B}(\mathcal{H}_1) \oplus \mathcal{B}(\mathcal{H}_2) \oplus \ldots \oplus \mathcal{B}(\mathcal{H}_m) . \tag{85}$$

Similar we can proceed with $\mathbf{d}'$ and Hilbert spaces $\mathcal{H}'_j$ of dimensions d'_j. We assert

$$\mathbf{d} \sim \mathbf{d}' \Leftrightarrow \mathcal{B}_{\mathbf{d}} \text{ is } ^*\text{-isomorph to } \mathcal{B}_{\mathbf{d}'} \tag{86}$$

To see the claim we use the permutation $d_j \to d_{i_j}$. In $\mathcal{H}_j$ we choose a basis $|k\rangle_j$, $k = 1, \ldots, d_j$ and a basis $|k\rangle_{i_j}$ in $\mathcal{H}_{i_j}$. Obviously there is a unitary U with $U|k\rangle_j = |k\rangle_{i_j}$ for all j, k. We see that both algebras become *-isomorphic by $A' = UAU^{-1}$ for any operator A out of (85).

We are now allowed to define: A *-algebra is *of type* $\mathbf{d}$, if it is *-isomorph to the algebra (85).

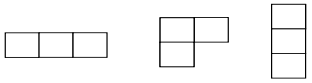
Remark: The number $|\mathbf{d}|$ is occasionally called the *algebraic dimension of* $\mathcal{A}$. Its logarithm (in bits ore nats) is called *entropy of* $\mathcal{A}$.

It is often convenient to choose within a type a standard one. This can be done by convention. A usual way is to require $d_1 \geq \ldots \geq d_m$. One then calls $\mathbf{d}$ *standardly* or *decreasingly ordered.* It opens the possibility to visualize the types with Young tableaux, see [8].

The following example is with $|\mathbf{d}| = 3$. The standard representations are

$$\{3\}, \quad \{2,1\}, \quad \{1,1,1\} \; .$$

The first one is the "full" algebra $\mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} = 3$, the last one is a maximally commutative subalgebra, while the middle one is $\mathcal{B}(\mathcal{H}_2) \oplus \mathbb{C}$. ($\mathbb{C}$ stands for the algebra over an 1-dimensional Hilbert space.) Their Young diagrams are



One may put (part of) Wedderburn's theorem in the form:
Every finite dimensional *-subalgebra of an algebra $\mathcal{B}(\mathcal{H})$ is *-isomorph to an algebra (85), i. e. it is of a certain type $\mathbf{d}$.

The algebra (85) can be identified with a subalgebra of $\mathcal{B}(\mathcal{H})$ where

$$\mathcal{H} = \mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_m, \quad \mathcal{H}_j = Q_j \mathcal{H}, \quad \mathbf{1} = \sum Q_j \; , \tag{87}$$

with projections Q_j. The Q_j sum up to $\mathbf{1}$, the identity operator of $\mathcal{H}$. In the course of constructing $\mathcal{B}_\mathbf{d}$, the unit element $\mathbf{1}_j \in \mathcal{B}(\mathcal{H}_j)$ is mapped onto the projection $Q_j \in \mathcal{B}(\mathcal{H})$. (We may use alternatively both notations. $\mathbf{1}_j$ can indicate a use "inside" the algebra, while Q_j indicates a definite embedding in a larger algebra.)

Let us restrict the trace over $\mathcal{H}$ to operators $A = A_1 + \ldots + A_m$, $A_j \in \mathcal{B}(\mathcal{H}_j)$. We get

$$\mathrm{tr}\, A = \sum \mathrm{tr}_j \, A_j, \quad \mathrm{tr}_j \text{ is the trace over } \mathcal{H}_j \; .$$

Notice $\dim \mathcal{H} = \mathrm{tr}\, \mathbf{1} = |\mathbf{d}|$.
The restriction of the trace of $\mathcal{H}$ to $\mathcal{B}_\mathbf{d}$ is called *the canonical trace* of $\mathcal{B}_\mathbf{d}$.

Let us denote the canonical trace of $\mathcal{B}_\mathbf{d}$ by $\mathrm{tr}^{\mathrm{can}}$ and let us try to explain the word "canonical". The point of this extra notation is its "intrinsic" nature: Let us think of $\mathrm{tr}^{\mathrm{can}}$ as a linear functional over $\mathcal{B}_\mathbf{d}$. It can be characterized by two properties: $\mathrm{tr}^{\mathrm{can}}$ is positive integer valued at the projections $P \neq \mathbf{0}$ of $\mathcal{B}_\mathbf{d}$, and it is the smallest with that property. It means:
The canonical trace is a type invariant. We can recover the canonical trace in every algebra *-isomorph to $\mathcal{B}_\mathbf{d}$.

There is yet another aspect to consider. The Lüders-von Neumann, or projective measurements, [14], [50], [5], are in one-to-one correspondence with the partition of the identity (87) of $\mathcal{H}$. We can associate

$$\mathbf{d} = \{d_1, \ldots, d_m\}, \quad d_j = \mathrm{rank}(Q_j) \tag{88}$$

with the measurement. The average measurement result is given by a trace preserving and unital completely positive map[13],

$$A \to \Phi(A) := \sum Q_j A Q_j, \quad A \in \mathcal{B}(\mathcal{H}) \; . \tag{89}$$

In the direct sum (85), the term $\mathcal{B}(\mathcal{H}_j)$ can be identified with $Q_j \mathcal{B}(\mathcal{H}) Q_j$, the algebra of all operators which can be written $Q_j A Q_j$. Hence,

$$\mathcal{B}_{d_1,\ldots,d_m} := \bigoplus Q_j \mathcal{B}(\mathcal{H}) Q_j, \quad d_j = \mathrm{rank}(Q_j) \; . \tag{90}$$

$\to$ Φ maps $\mathcal{B}(\mathcal{H})$ onto $\mathcal{B}_{\mathbf{d}}$.

Remark: Φ is a completely positive unital map which maps the algebra onto a subalgebra, though it does not preserve multiplication: Generally $QABQ$ is not equal to $QAQBQ$ with a projection Q. Several interesting questions appear. For instance, which channels result after several applications of projective ones? The problem belongs to the theory of conditional expectations.

The state space of $\mathcal{B}_{\mathbf{d}}$

To shorten notation we shall write $\Omega(\mathcal{H})$ instead of $\Omega(\mathcal{B}(\mathcal{H}))$.
Let us now examine the state space $\Omega(\mathcal{B}_{\mathbf{d}})$ which is a subset of $\Omega(\mathcal{H})$. Indeed, a state $\underline{\omega}$ of $\mathcal{B}_{\mathbf{d}}$ can be written $\underline{\omega}(A) = \mathrm{tr}\, \omega A$ and we conclude

$$\mathrm{tr}\, \omega A = \mathrm{tr}\, \omega \sum Q_j A Q_j = \mathrm{tr}\, (\sum Q_j \omega Q_j) A$$

by (87). Hence we can choose $\omega \in \mathcal{B}_{\mathbf{d}}$ and, after doing so, ω becomes unique. In conclusion, $\Omega(\mathcal{B}_{\mathbf{d}}) \subset \Omega(\mathcal{H})$ and

$$\omega \in \Omega(\mathcal{B}_{\mathbf{d}}) \; \Leftrightarrow \; \sum Q_j \omega Q_j = \omega \tag{91}$$

for density operators $\omega \in \Omega(\mathcal{H})$.

A density operators ω_j of $\mathcal{B}(\mathcal{H}_j)$ can be identified with a density operators on $\mathcal{H}$ supported by $\mathcal{H}_j = Q_j \mathcal{H}$. Equivalently we have $\underline{\omega}_j(Q_j) = 1$ for the corresponding states. These states form a face of $\Omega(\mathcal{H})$, and these faces are orthogonal one to another. We get the convex combination

$$\omega \in \Omega(\mathcal{B}_{\mathbf{d}}) \; \Leftrightarrow \; \omega = \sum_{j=1}^{m} p_j \omega_j, \quad \mathrm{tr}\, Q_j \omega_j = \underline{\omega}_j(Q_j) = 1 \; . \tag{92}$$

[13] Complete positive maps respect the superposition principle in tensor products, [15], [5].

The convex combination (92) is uniquely determined by ω, a consequence of the orthogonality $\omega_j\omega_k = \mathbf{0}$ if $j \neq k$.
$\rightarrow$ The state space of $\mathcal{B}_\mathbf{d}$, embedded in $\Omega(\mathcal{H})$, $\dim\mathcal{H} = |\mathbf{d}|$, is the direct convex sum of the state spaces $\Omega(\mathcal{H}_j)$ with $\dim\mathcal{H}_j = d_j$ and $d_j \in \mathbf{d}$.
We see further: Φ defined in (89) maps $\Omega(\mathcal{H})$ onto $\Omega(\mathcal{B}_\mathbf{d})$.

A picturesque description is in saying we have a simplex with m corners and we "blow up", for all j, the j-th corner to the convex set $\Omega(\mathcal{H}_j)$. Then we perform the convex hull.

From (87), (92), and the structure of $\mathcal{B}_\mathbf{d}$ we find the pure (i. e. extremal) density operators by selecting j and a unit vector $|\psi\rangle \in \mathcal{H}_j$ to be $P = |\psi\rangle\langle\psi|$. (We may also write $\pi = P$, but presently we like to see the density operator of a pure state as a member of the projections. This double role of rank one projectors is a feature of discrete type I von Neumann algebras.) Let $\underline{\pi}$ be the state of $\mathcal{B}_\mathbf{d}$ with density operator P. Just by insertion we see

$$PAP = \underline{\pi}(A)P \text{ for all } A \in \mathcal{B}_\mathbf{d} . \tag{93}$$

On the other hand, if for any projector P there is a linear form $\underline{\pi}$ such that (93) is valid, $\underline{\pi}$ must be a state and P its density operator. (Inserting $A = P$ we find $\underline{\pi}(P) = 1$. Because $PA^\dagger AP$ is positive, $\underline{\pi}(A^\dagger A)$ must be positive. Hence it follows from (93), if P is a projection P, $\underline{\pi}$ is a state.) It is also not difficult to see that (93) requires P to be of rank one and $\underline{\pi}$ is pure. We now have another criterium for pure states which refers to the algebra only.

Let $\mathcal{A}$ be *-isomorph to an algebra $\mathcal{B}_\mathbf{d}$. A state $\underline{\pi}$ of $\mathcal{A}$ is pure if and only if there is a projection P such that (93) is valid for all $A \in \mathcal{A}$. Then P is the density operator of the pure state, or, in other terms, $\pi = P$.

The projections, which are density operators of pure states, enjoy a special property, they are *minimal.* A projection P is minimal in an algebra, if from $P = P_1 + P_2$ with P_j projections, it follows either $P_1 = P$ or $P_1 = \mathbf{0}$.

It is quite simple to see $P = |\psi\rangle\langle\psi|$ for a minimal projection operator of $\mathcal{B}_\mathbf{d}$ and, hence, it is a density operator of a pure state. Therefore in algebras *-isomorph to an algebra $\mathcal{B}_\mathbf{d}$ we can assert:
A projection P of $\mathcal{A}$ is minimal if and only if it is the density operator of a pure state of $\mathcal{A}$.

A further observation: Let $\mathcal{A}$ be of type $\mathbf{d}$. There is a linear functional over $\mathcal{A}$ which attains the value 1 for all minimal projections. This linear form is the canonical trace of $\mathcal{A}$.

By slightly reformulating some concepts from Hilbert space we have obtained purely algebraic ones. This way of thinking will also dominate our next issue.

Transition probabilities for pure states

We start again with $\mathcal{B}_{\mathbf{d}}$ as a subalgebra of $\mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = |\mathbf{d}|$. Let us consider some pure states $\underline{\pi}_j$ of $\mathcal{B}_{\mathbf{d}}$. They can be represented by unit vectors,

$$\underline{\pi}_j(A) = \langle \psi_j, A\psi_j \rangle, \quad \pi_j \equiv P_j = |\psi_j\rangle\langle\psi_j| \; . \tag{94}$$

Let us agree "as usual" that

$$\Pr(\underline{\pi}_1, \underline{\pi}_2) = \Pr(\pi_1, \pi_2) = |\langle \psi_1, \psi_2 \rangle|^2 \tag{95}$$

is the transition probability. To obtain the same value for two pure states of an algebra $\mathcal{A}$ *-isomorph to $\mathcal{B}_{\mathbf{d}}$, we reformulate (95) in an invariant way: The right-hand side of (95) is the trace of $\pi_1\pi_2$. In $\mathcal{B}_{\mathbf{d}}$ the canonical trace coincides with the trace over $\mathcal{H}$. Hence, for a general algebra $\mathcal{A}$, we have to use the canonical trace. We get

$$\Pr(\underline{\pi}_1, \underline{\pi}_2) = \Pr(\pi_1, \pi_2) = \mathrm{tr}^{\mathrm{can}} \pi_1\pi_2 \; . \tag{96}$$

Switching, for convenience, to the notation $P_j = \pi_j$, we get $P_1P_2P_1 = \underline{\pi}_1(P_2)P_1$ by inserting $A = P_2$ in the appropriate equation (93) for $\underline{\pi}_1$. By taking the trace we get the expression (96) for the transition probability. Interchanging the indices we finally get

$$\Pr(\underline{\pi}_1, \underline{\pi}_2) = \Pr(\pi_1, \pi_2) = \underline{\pi}_1(P_2) = \underline{\pi}_2(P_1) \; . \tag{97}$$

This and (96) express the transition probability for any two pure states of an algebra $\mathcal{A}$, *-isomorph to a finite dimensional von Neumann algebra.

Our next aim is to prove

$$\Pr(\underline{\pi}_1, \underline{\pi}_2) = \inf_{A > \mathbf{0}} \underline{\pi}_1(A)\underline{\pi}_2(A^{-1}) \; , \tag{98}$$

A is running through all invertible positive elements of $\mathcal{A}$.

It suffices to prove the assertion for $\mathcal{B}_{\mathbf{d}}$. Relying on (94) we observe

$$|\langle \psi_1, \psi_2 \rangle|^2 \leq \langle A^{1/2}\psi_1, A^{1/2}\psi_1 \rangle \, \langle A^{-1/2}\psi_2, A^{-1/2}\psi_2 \rangle \; .$$

Therefore, the left-hand side of (98) cannot be larger than the right one. It remains to ask, whether the asserted infimum can be reached. For this purpose we set

$$A_s = s\mathbf{1} + P_2, \quad A_s^{-1} = \frac{1}{s}\mathbf{1} - \frac{1}{s(1+s)} P_2 \; .$$

A_s is positive for $s > 0$. We find

$$\underline{\pi}_1(A_s) = s + \Pr(\underline{\pi}_1, \underline{\pi}_2), \quad \underline{\pi}_2(A_s^{-1}) = \frac{1}{s} - \frac{1}{s(1+s)} = (1+s)^{-1}$$

and it follows

$$\lim_{s\to+0} \underline{\pi}_1(A_s)\underline{\pi}_2(A_s^{-1}) = \Pr(\underline{\pi}_1, \underline{\pi}_2)$$

and (98) has been proven.

In [1] a similar inequality is reported:

$$2|\langle\psi_1, \psi_2\rangle| = \inf_{A>\mathbf{0}} \langle\psi_1, A\psi_1\rangle + \langle\psi_2, A^{-1}\psi_2\rangle$$

with A varying over all invertible positive operators on a Hilbert space. The equation remains valid for pairs of pure states in a finite *-subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$. The slight extensions of the inequality reads

$$2\sqrt{\Pr(\underline{\pi}_1, \underline{\pi}_2)} = \inf_{\mathbf{0}<A\in\mathcal{A}} \underline{\pi}_1(A) + \underline{\pi}_2(A^{-1}) \; . \tag{99}$$

To prove it we write out the inequality

$$0 \leq (t\sqrt{\underline{\pi}_1(A)} - t^{-1}\sqrt{\underline{\pi}_2(A^{-1})})^2 \; ,$$

t a positive number. We get

$$2\sqrt{\underline{\pi}_1(A)\underline{\pi}_2(A^{-1})} \leq t^2\underline{\pi}_1(A) + t^{-2}\underline{\pi}_2(A^{-1})$$

and, by (98), the right-hand side of (99) is not less than the left one. Adjusting the operators A_s above to $B_s = t_s^2 A_s$ in such a way that $\underline{\pi}_1(B_s) = \underline{\pi}_2(A^{-1})$, then

$$2\sqrt{\underline{\pi}_1(B_s)\underline{\pi}_2(B_s^{-1})} = \underline{\pi}_1(B_s) + \underline{\pi}_2(B_s^{-1}) \; .$$

Performing the limes $s \to 0$ as in the proof of (98) shows that the asserted infimum can be approached arbitrarily well.

Last not least we convince ourselves that the transition probability between pure states *is already fixed by the convex structure* of $\Omega(\mathcal{A})$ respectively of $\underline{\Omega}(\mathcal{A})$.

We prove it for $\Omega(\mathcal{A})$. Let l be a real linear form over the Hermitian operators of $\mathcal{A}$ such that for all density operators ω one has $0 \leq l(\omega) \leq 1$. Then $l(A) \geq 0$ for all positive operators A. Now assume $l(P) = 1$ for a minimal projection. Combining both assumptions we find $l(\mathbf{1}_{\mathcal{A}}) = 1$. Hence l is a pure state $\underline{\pi}$ of $\mathcal{A}$. If P' is another minimal projection, i. e. an extremal element of $\Omega(\mathcal{A})$, we can calculate the transition probability $l(P') = \underline{\pi}(P')$.

The result implies: Our state spaces are *rigid:* If a linear map Φ,

$$\Phi \; : \; \mathcal{A} \mapsto \mathcal{A} \; ,$$

maps $\Omega(\mathcal{A})$ one–to–one onto itself, it must preserve the transition probabilities between pure density operators.

In the particular case $\mathcal{A} = \mathcal{B}(\mathcal{H})$ the map Φ must be a Wigner symmetry. A useful reformulation of this statement reads:
Let Φ_1, Φ_2 denote invertible linear maps from $\mathcal{B}(\mathcal{H})$ onto $\mathcal{B}(\mathcal{H})$. Assume $\Omega(\mathcal{H})$

is mapped by both maps onto the same set of operators. Then there is a unitary or an anti-unitary V such that

$$\Phi_2(X) = \Phi_1(VXV^*) \text{ for all } X \in \mathcal{B}(\mathcal{H}) .$$

Indeed, $\Phi_1^{-1}\Phi_2$ must be a Wigner symmetry.
Remark: Mielnik has defined a "transition probability" between extremal states of a compact convex set K in this way. Let P and P' be two extremal points of K. The "probability" of the transition $P \to P'$ is defined to be $\inf l(P')$ where l runs through all real affine functionals on K with values between 0 and 1 and with $l(P) = 1$. Indeed, for $\Omega(\mathcal{A})$ the procedure gives the correct transition probability as shown above.

2.4 All subsystems for $\dim \mathcal{H} < \infty$

Here we are interested in Wedderburn's description, of the *-subalgebras of $\mathcal{B}(\mathcal{H})$, $\dim \mathcal{H} < \infty$, [21, 37]. In short, such a subalgebra is *-isomorph to a certain algebra $\mathcal{B}_{\mathbf{d}}$.

We change our notations towards its use in quantum information. We think of a quantum system with algebra $\mathcal{B}^{\mathrm{A}}$, owned by some person, say Alice. We may assume the algebra $\mathcal{B}^{\mathrm{A}}$ to be a unital *-subalgebra of a larger algebra $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$. The type of $\mathcal{B}^{\mathrm{A}}$ is the not ordered list $\mathbf{d}^{\mathrm{A}} = \{d_1^{\mathrm{A}}, \ldots, d_m^{\mathrm{A}}\}$. Alice is allowed to operate freely within her subsystem, which is also called "the A-system".
Theorem: Let $\mathcal{B}^{\mathrm{A}}$ be a unital *-subalgebra of $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$ of type $\mathbf{d}^{\mathrm{A}}$. Then There is a decomposition

$$\mathcal{H}^{\mathrm{AB}} = \mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_m, \quad \mathcal{H}_j = \mathcal{H}_j^{\mathrm{A}} \otimes \mathcal{H}_j^{\mathrm{B}}, \tag{100}$$

$$d_j^{\mathrm{A}} = \dim \mathcal{H}_j^{\mathrm{A}}, \quad d_j^{\mathrm{B}} := \dim \mathcal{H}_j^{\mathrm{B}} ,$$

such that

$$\mathcal{B}^{\mathrm{A}} = (\mathcal{B}(\mathcal{H}_1^{\mathrm{A}}) \otimes \mathbf{1}_1^{\mathrm{B}}) \oplus \ldots \oplus (\mathcal{B}(\mathcal{H}_m^{\mathrm{A}}) \otimes \mathbf{1}_m^{\mathrm{B}}) . \tag{101}$$

Equally well we may represent $\mathcal{B}^{\mathrm{A}}$ as a diagonal block matrix with diagonal blocks $\mathcal{B}(\mathcal{H}_j^{\mathrm{A}}) \otimes \mathbf{1}_j^{\mathrm{B}}$.

In the theorem we denote by $\mathbf{1}_j^{\mathrm{A}}$ the identity operator of $\mathcal{H}_j^{\mathrm{A}}$ and by $\mathbf{1}_j^{\mathrm{B}}$ the one of $\mathcal{H}_j^{\mathrm{B}}$. Therefore, $\mathbf{1}_j^{\mathrm{A}} \otimes \mathbf{1}_j^{\mathrm{B}}$ is equal to $\mathbf{1}_j$, the identity operator of $\mathcal{H}_j$. The latter can be identified with the projection Q_j projecting $\mathcal{H}$ onto $\mathcal{H}_j$, i. e. $\mathbf{1}_j = Q_j$. (100) and (101) describe how $\mathcal{B}_{\mathbf{d}^{\mathrm{A}}}$ is embedded into $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$ to become $\mathcal{B}^{\mathrm{A}}$ by the embedding *-isomorphism

$$A_1 + \ldots + A_m \leftrightarrow A_1 \otimes \mathbf{1}_1 + \ldots + A_m \otimes \mathbf{1}_m, \quad A_j \in \mathcal{B}(\mathcal{H}_j^{\mathrm{A}}) . \tag{102}$$

Now we can see, why, by identifying $\mathcal{B}^{\mathrm{A}}$ as a subsystem of $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$, a second subsystem, called "Bob's system", appears quite naturally. It consists

of those operators of $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$ which can be executed independently of Alice's actions. These operators must commute with those of the A-system. Hence, all of them[14] constitute Bob's algebra, $\mathcal{B}^{\mathrm{B}}$. Therefore, Bob's algebra is the commutant of $\mathcal{B}^{\mathrm{A}}$ in $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$. By (101) we see

$$\mathcal{B}^{\mathrm{B}} := (\mathcal{B}^{\mathrm{A}})' = (\mathbf{1}_1^{\mathrm{A}} \otimes \mathcal{B}(\mathcal{H}_1^{\mathrm{B}})) \oplus \ldots \oplus (\mathbf{1}_m^{\mathrm{A}} \otimes \mathcal{B}(\mathcal{H}_m^{\mathrm{B}})) \,. \tag{103}$$

We further can find the center of $\mathcal{B}^{\mathrm{A}}$, respectively of $\mathcal{B}^{\mathrm{B}}$. The center describes the actions which are allowed to both, Alice and Bob. These operators behave "classical" for them. We get

$$\mathcal{B}^{\mathrm{A}} \cap \mathcal{B}^{\mathrm{B}} = \mathbb{C}Q_1 + \ldots + \mathbb{C}Q_m, \quad Q_j = \mathbf{1}_j^{\mathrm{A}} \otimes \mathbf{1}_j^{\mathrm{B}} = \mathbf{1}_j \,. \tag{104}$$

The type of the commutant consists of m-times the number one.

The types of $\mathcal{B}^{\mathrm{A}}$ and of $\mathcal{B}^{\mathrm{B}}$ are $\mathbf{d}^{\mathrm{A}} = \{d_1^{\mathrm{A}}, \ldots, d_m^{\mathrm{A}}\}$ and $\mathbf{d}^{\mathrm{B}} = \{d_1^{\mathrm{B}}, \ldots, d_m^{\mathrm{B}}\}$ respectively. In general, neither one can be assumed decreasingly ordered. Notice

$$\dim \mathcal{H}^{\mathrm{AB}} = \sum d_j^{\mathrm{A}} d_j^{\mathrm{B}} \,.$$

Let us denote by $\mathcal{B}^{\mathrm{AB}}$ the subalgebra generated by $\mathcal{B}^{\mathrm{A}}$ and $\mathcal{B}^{\mathrm{B}}$. Equivalently, $\mathcal{B}^{\mathrm{AB}}$ is the smallest subalgebra of $\mathcal{B}(\mathcal{H})$ containing $\mathcal{B}^{\mathrm{A}}$ and $\mathcal{B}^{\mathrm{B}}$,

$$\mathcal{B}^{\mathrm{AB}} = \mathcal{B}(\mathcal{H}_1) \oplus \ldots \oplus \mathcal{B}(\mathcal{H}_m) = Q_1 \mathcal{B}(\mathcal{H}) Q_1 + \ldots + Q_m \mathcal{B}(\mathcal{H}) Q_m \,. \tag{105}$$

The fact that $\mathcal{B}^{\mathrm{AB}}$ is generated in a larger algebra by the algebras $\mathcal{B}^{\mathrm{A}}$ and $\mathcal{B}^{\mathrm{B}}$ can be expressed also by $\mathcal{B}^{\mathrm{AB}} = \mathcal{B}^{\mathrm{A}} \vee \mathcal{B}^{\mathrm{B}}$. The type of $\mathcal{B}^{\mathrm{AB}}$ is

$$\mathbf{d}^{\mathrm{AB}} := \{d_1^{\mathrm{A}} d_1^{\mathrm{B}}, \ldots d_m^{\mathrm{A}} d_m^{\mathrm{B}}\} \,.$$

As long as $\mathcal{B}^{\mathrm{AB}}$ is not considered itself as a subsysstem of a larger one, and we are allowed to write $\mathcal{B}^{\mathrm{AB}} = \mathcal{B}_{\mathbf{d}^{\mathrm{AB}}}$.

Embedding and partial trace

Let us stick to the just introduced subalgebras of $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$, namely $\mathcal{B}^{\mathrm{A}}$, $\mathcal{B}^{\mathrm{B}}$, $\mathcal{B}^{\mathrm{AB}}$, and $\mathcal{C} = \mathcal{B}^{\mathrm{A}} \cap \mathcal{B}^{\mathrm{B}}$.

If $\underline{\omega}^{\mathrm{AB}}$ is a state of $\mathcal{B}^{\mathrm{AB}}$, its restriction to $\mathcal{B}^{\mathrm{A}}$ is a state $\underline{\omega}^{\mathrm{A}}$ of $\mathcal{B}^{\mathrm{A}}$. The restriction map lets fall down any functional of $\mathcal{B}^{\mathrm{AB}}$ to $\mathcal{B}^{\mathrm{A}}$. After its application, we have obtained $\underline{\omega}^{\mathrm{A}}$ from $\underline{\omega}^{\mathrm{AB}}$ and all what has changed is: Only arguments from $\mathcal{B}^{\mathrm{A}}$ will be allowed for $\underline{\omega}^{\mathrm{A}}$.

The *partial trace*[15], $\omega^{\mathrm{AB}} \to \omega^{\mathrm{A}}$, concerns the involved density operators. It is a map from $\mathcal{B}^{\mathrm{AB}}$ to $\mathcal{B}^{\mathrm{A}}$. For its definition and for later use we need the canonical traces of $\mathcal{B}^{\mathrm{A}}$ and $\mathcal{B}^{\mathrm{B}}$ which we now denote by tr^{A} and tr^{B} respectively. It is

[14] We ignore that there may be further restrictions to Bob.

[15] The partial trace is a particular "conditional expectation".

$$\mathrm{tr}^{\mathrm{A}}\omega^{\mathrm{A}}X = \underline{\omega}^{\mathrm{AB}}(X) \equiv \mathrm{tr}\,\omega^{\mathrm{AB}}X, \quad X \in \mathcal{B}^{\mathrm{A}}\ . \tag{106}$$

Remark: The algebra $\mathcal{B}^{\mathrm{AB}}$ is of the form (90), (85). Therefore, its canonical trace, $\mathrm{tr}^{\mathrm{AB}}$ is the canonical trace over $\mathcal{B}(\mathcal{H})$, i. e. it is just the trace over $\mathcal{H}$.

We read (106) as follows: The right-hand side becomes a linear form over $\mathcal{B}^{\mathrm{A}}$. Every linear functional over $\mathcal{B}^{\mathrm{A}}$ can be uniquely written by the help of the canonical trace as done at the left-hand side. This defines the partial trace

$$\omega^{\mathrm{AB}} \to \omega^{\mathrm{A}} := \mathrm{tr}_{\mathrm{B}}\omega^{\mathrm{AB}} \tag{107}$$

from $\mathcal{B}^{\mathrm{AB}}$ to $\mathcal{B}^{\mathrm{A}}$. The partial trace is "dual" to the restriction map.

The algebra $\mathcal{B}^{\mathrm{AB}}$ consists of all operators

$$Z = \sum_{j=1}^{m} X_j Y_j = \sum_{j=1}^{m} (A_j \otimes \mathbf{1}_j^{\mathrm{B}})\,(\mathbf{1}_j^{\mathrm{A}} \otimes B_j) \tag{108}$$

with

$$A_j \in \mathcal{B}(\mathcal{H}_j^{\mathrm{A}}), \quad B_j \in \mathcal{B}(\mathcal{H}_j^{\mathrm{B}})\ .$$

This follows from (100) and (101). Now

$$\mathrm{tr}\,Y_j = \mathrm{tr}\,(\mathbf{1}_j^{\mathrm{A}} \otimes B_j) = d_j^{\mathrm{A}}\ \mathrm{tr}\,B_j = d_j^{\mathrm{A}}\ \mathrm{tr}^{\mathrm{B}}Y_j\ . \tag{109}$$

The dimensional factors point to the main difference between the canonical trace of $\mathcal{B}^{\mathrm{A}}$ and of the *induced trace,* which is the trace of $\mathcal{H}$ applied to the operators of the subalgebra $\mathcal{B}^{\mathrm{A}}$. All together we get the partial trace of the operator (108),

$$\mathrm{tr}_{\mathrm{B}}\,Z = \sum (d_j^{\mathrm{A}})^{-1}(\mathrm{tr}\,Y_j)X_j = \sum X_j\,\mathrm{tr}^{\mathrm{B}}Y_j\ . \tag{110}$$

An important conclusion is

$$\mathrm{tr}_{\mathrm{B}}\,XZ = X\mathrm{tr}_{\mathrm{B}}Z, \quad \mathrm{tr}_{\mathrm{B}}\,ZX = (\mathrm{tr}_{\mathrm{B}}Z)X, \quad X \in \mathcal{B}^{\mathrm{A}}. \tag{111}$$

Similar to tr_{B} one treats the partial trace tr_{A}. One can check

$$\mathrm{tr}_{\mathrm{B}}\,\mathrm{tr}_{\mathrm{A}} = \mathrm{tr}_{\mathrm{A}}\,\mathrm{tr}_{\mathrm{B}} = \mathrm{tr}_{\mathrm{AB}}\ . \tag{112}$$

Because $\mathrm{tr}_{\mathrm{AB}}$ projects an operator of $\mathcal{B}^{\mathrm{AB}}$ into both, $\mathcal{B}^{\mathrm{A}}$ and $\mathcal{B}^{\mathrm{B}}$, it projects onto the center, $\mathcal{C} = \mathcal{B}^{\mathrm{A}} \cap \mathcal{B}^{\mathrm{B}}$, of $\mathcal{B}^{\mathrm{AB}}$. By inspection we identify (112) with the partial trace of $\mathcal{B}^{\mathrm{AB}}$ onto its center.

The ansatz (106) applies also to the partial trace from $\mathcal{B}(\mathcal{H}^{\mathrm{AB}})$ to $\mathcal{B}^{\mathrm{AB}}$. Because the latter is the commutant of the center $\mathcal{C} = \mathcal{B}^{\mathrm{A}} \cap \mathcal{B}^{\mathrm{B}}$, we have

$$\mathrm{tr}_{\mathrm{A}\cap\mathrm{B}}(Z) = \sum Q_j Z Q_j, \quad \mathcal{B}^{\mathrm{A}} \cap \mathcal{B}^{\mathrm{B}} = \sum Q_j \mathbb{C}\ , \tag{113}$$

see (87) and (89), where the map has been called Φ because at this occasion the partial trace was not yet defined.

3 Transition probability, fidelity, and Bures distance

The aim is to define transition probabilities, [59], [46], between two states of a quantum system, say $\mathcal{A}$, by operating in larger quantum systems. We call it $\Pr(\underline{\rho}, \underline{\omega})$ or, with density operators, $\Pr(\rho, \omega)$.

The notation for the fidelity, $\mathrm{F}(\rho, \omega)$, used here is that of Nielsen and Chuang[16], [15], i. e. it is the square root of the transition probability,

$$\mathrm{F}(\rho, \omega) := \sqrt{\Pr(\rho, \omega)}\,. \tag{114}$$

This quantity is also denoted by "square root fidelity" or by "overlap". An analogous quantity between two probability measures is known as "Kakutani mean", [47], and, for probability vectors, as "Bhattacharyya coefficient". Occasionally the latter name is also used in the quantum case.

There is a related extension of the Study-Fubini distance to the Bures one, [34]. The Bures distance, $\mathrm{dist}_{\mathrm{B}}(\underline{\rho}, \underline{\omega})$, is an inner distance in the set of positive linear functionals, or, in finite dimensions equivalently, in the set of positive operators. The Bures distance is a quantum version of the Fisher distance, [40].

There is a Riemannian metric, the Bures metric, belonging to the Bures distance, [61]. It extends the Fubini-Study metric to general (i.e. mixed) states. It also extends the Fisher metric, originally defined for spaces of probability measures, to quantum theory. (However, there is a large class of reasonable quantum versions of the Fisher metric, discovered by Petz, [53].)

Below we shall define transition probability and related quantities "operationally". Later we shall discuss several possibilities to get them "intrinsically", without leaving a given quantum system, [59], [46].

From the mathematical point of view, there are some quite useful tricks in handling two positive operators in general position.

3.1 Purification

Purification is a tool to extend properties of pure states to general ones. It lives from the fact that, given a state, say $\underline{\omega}^{\mathrm{A}}$, of a quantum system A, there are pure states in sufficiently larger systems the restriction of which to the A-system coincides with $\underline{\omega}^{\mathrm{A}}$. The same terminology is used for the corresponding density operators. Of special interest is the case of a larger system which purifies all states of the A-system.

We can lift any state of a quantum system to every larger system. We can require that a pure state is lifted to a pure state:
Let $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{B}(\mathcal{H})$ and $\underline{\pi}_1$ a pure state of $\mathcal{A}_1$ with density operator P_1. Being a minimal projection in $\mathcal{A}_1$, P_1 may be not minimal in $\mathcal{A}_2$. But then we can write P_1 as a sum of minimal projections of $\mathcal{A}_2$. If P_2 is one of them and $\underline{\pi}_2$ the corresponding pure state of $\mathcal{A}_2$, then $\underline{\pi}_2$ is a pure lift of $\underline{\pi}_1$.

[16] There are also quite different expressions called "fidelity".

As a matter of fact, every state $\underline{\omega}_2$ satisfying $\underline{\omega}_2(P_1) = 1$ is a lift of $\underline{\pi}_1$ to $\mathcal{A}_2$. These states exhausts all lifts of $\underline{\pi}_1$ to $\mathcal{A}_2$. They constitute a face of the state space of $\mathcal{A}_2$.
Assume the state $\underline{\omega}_1$ of $\mathcal{A}_1$ is written as a convex combination of pure states. After lifting them to pure states of $\mathcal{A}_2$ we get a convex combination which extends $\underline{\omega}_1$ to $\mathcal{A}_2$.
Generally, there is a great freedom in extending states of a quantum system to a larger quantum system.

The most important case is the purification of the states of $\mathcal{B}(\mathcal{H})$ or, equivalently, of $\Omega(\mathcal{H})$, well described in [9], [15], [4], [19], and in other text books on quantum information theory. It works by embedding $\mathcal{B}(\mathcal{H})$ as the subalgebra $\mathcal{B}(\mathcal{H}) \otimes \mathbf{1}'$ into a bipartite system $\mathcal{B}(\mathcal{H} \otimes \mathcal{H}')$, provided $d = \dim \mathcal{H} \leq \dim \mathcal{H}'$. Given $\omega \in \Omega(\mathcal{H})$, a unit vector $\psi \in \mathcal{H} \otimes \mathcal{H}'$ is *purifying* ω, and $\pi = |\psi\rangle\langle\psi|$ is a purification of ω, if

$$\langle \psi, (X \otimes \mathbf{1}')\psi \rangle = \operatorname{tr} X\omega \ \text{ for all } X \in \mathcal{B}(\mathcal{H}) \tag{115}$$

or, equivalently,

$$\underline{\omega}(X) = \underline{\pi}(X \otimes \mathbf{1}') \equiv \operatorname{tr} \pi(X \otimes \mathbf{1}') \ . \tag{116}$$

To get a suitable ψ, one chooses d ortho-normal vectors $|j\rangle'$ in $\mathcal{H}'$ and a basis $|j\rangle$ of eigenvectors of ω. Now

$$|\psi\rangle = \sum \lambda^{1/2} |j\rangle \otimes |j\rangle' \ \text{ with } \omega|j\rangle = \lambda_j |j\rangle \tag{117}$$

purifies ω. Indeed,

$$\langle \psi, (X \otimes \mathbf{1}')\psi \rangle = \sum \lambda_j \langle j|X|j\rangle = \operatorname{tr} X\omega \ .$$

Now let $\mathcal{A}$ be a unital *-subalgebra of $\mathcal{B}(\mathcal{H})$ and $\underline{\omega}^{\mathrm{A}}$ one of its states. We have already seen that we can lift $\underline{\omega}^{\mathrm{A}}$ to a state $\underline{\omega}$ of $\mathcal{B}(\mathcal{H})$. With the density operator ω of $\underline{\omega}$ we now proceed as above.

3.2 Transition probability, fidelity, ...

Let $\mathcal{A}$ be a unital *-subalgebra of an algebra $\mathcal{B}(\mathcal{H})$ with finite dimensional Hilbert space $\mathcal{H}$. Denote by $\underline{\omega}_1^{\mathrm{A}}$ and $\underline{\omega}_2^{\mathrm{A}}$ two states of $\mathcal{A}$ and by ω_1^{A} and ω_2^{A} their density operators.

The task is, to prepare $\underline{\omega}_2$ if the state of our system is $\underline{\omega}_1$.
To do so one thinks of purifications $\underline{\pi}_j$ of our $\underline{\omega}_j^{\mathrm{A}}$ in a larger quantum system in which $\mathcal{A}$ is embedded.
One then tests, in the larger system, whether π_2 is true. If the answer of the test is "yes", then π_2 and, hence, ω_2^{A} is prepared.
The probability of success is $\Pr(\underline{\pi}_1, \underline{\pi}_2)$ as defined in (95), (96), and (97).

One now asks for optimality of the described procedure, i.e. one looks for a projective measurement in a larger system which prepares a purification of ω_2^{A} with maximal probability.

This maximal possible probability for preparing ω_2^{A} with given ω_1^{A} is called the *transition probability* from ω_1^{A} to ω_2^{A} or, as this quantity is symmetric in its entries, the transition probability of the pair $\{\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}\}$. The definition applies to any unital C*-algebra and, formally, to any unital *-algebra, [59, 27].

The definition can be rephrased

$$\Pr(\underline{\omega}_1^{\mathrm{A}}, \underline{\omega}_2^{\mathrm{A}}) := \sup \Pr(\underline{\pi}_1, \underline{\pi}_2) \,, \tag{118}$$

where $\underline{\pi}_1, \underline{\pi}_2$ is running through all simultaneous purifications of $\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}$. We also use the density operator notation

$$\Pr(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) \equiv \Pr(\underline{\omega}_1^{\mathrm{A}}, \underline{\omega}_2^{\mathrm{A}}) \,.$$

In almost the same way we define the fidelity by

$$\mathrm{F}(\omega_1, \omega_2) = \sup |\langle \psi_1, \psi_2 \rangle| \tag{119}$$

where ψ_1, ψ_2 run through all simultaneous purifications of ω_1, ω_2 in some $\mathcal{B}(\mathcal{H})$. Though, we do not include all possible purifications, (by using only "full" algebras,) the relation (114) remains valid.

Remark: Let $\underline{\omega}_1$, $\underline{\omega}_2$ two states of a unital C*-algebra $\mathcal{A}$ and $\underline{\nu}$ one of its a linear functionals. If and only if

$$|\underline{\nu}(A^\dagger B)|^2 \leq \underline{\omega}_1(A^\dagger A)\, \underline{\omega}_2(B^\dagger B) \tag{120}$$

for all $A, B \in \mathcal{A}$ there is an embedding Ψ in an algebra $\mathcal{B}(\mathcal{H})$ such that there are purifying vectors ψ_1, ψ_2 satisfying

$$\underline{\nu}(A) = \langle \psi_1, \Psi(A)\psi_2 \rangle, \quad A \in \mathcal{A} \,. \tag{121}$$

This relation implies

$$|\underline{\nu}(\mathbf{1})|^2 \leq \Pr(\underline{\omega}_1, \underline{\omega}_2) \,. \tag{122}$$

Now the definition above can be rephrased: The transition probability is the sup of $|\underline{\nu}(\mathbf{1})|^2$ with $\underline{\nu}$ running through all linear forms satisfying (121). There exist linear functionals $\underline{\nu}$ satisfying (120) with equality in (122). Their structure and eventual uniqueness has been investigated by Alberti, [26].

The Bures distance

For the next term, the Bures distance, [34], it is necessary, not to insist in normalization of the vectors and not to require the trace one condition for the density operators in (119).

Remembering (28) and (29), one defines the *Bures distance* by

$$\mathrm{dist}_B(\omega_1,\omega_2) = \sup \mathrm{dist}_{FS}(\pi_1,\pi_2) = \sup \parallel \psi_2 - \psi_1 \parallel \tag{123}$$

where the sup is running through all simultaneous purifications of ω_1 and ω_2. Because of (119) this comes down to

$$\mathrm{dist}_B(\omega_1,\omega_2) = \sqrt{\mathrm{tr}\,\omega_1 + \mathrm{tr}\,\omega_2 - 2\mathrm{F}(\omega_1,\omega_2)}\,. \tag{124}$$

Rewritten for two density operators it becomes

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

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

$$\mathrm{Dist}_B(\omega_1,\omega_2) = \arccos\sqrt{\mathrm{Pr}(\omega_1,\omega_2)} \tag{125}$$

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

What remains is to express of (118) or (119) in a more explicit way. The dangerous thing in these definitions is the word "all". How to control all possible purifications of every embedding in suitable larger quantum systems? The answer is in a "saturation" property: One cannot do better in (118) than by the squared algebraic dimension of $\mathcal{A}$ for the purifying system.

3.3 Optimization

Let $\mathcal{A} = \mathcal{B}_{\mathbf{d}}$ with $d = |\mathbf{d}|$ as in (85) and (87). Hence, up to a slight change in notation we have

$$\mathcal{A} = \mathcal{B}(\mathcal{H}_1^{\mathrm{A}}) \oplus \mathcal{B}(\mathcal{H}_2^{\mathrm{A}}) \oplus \ldots \oplus \mathcal{B}(\mathcal{H}_m^{\mathrm{A}})\,, \tag{126}$$

$$\mathcal{H}^{\mathrm{A}} = \mathcal{H}_1^{\mathrm{A}} \oplus \ldots \oplus \mathcal{H}_m^{\mathrm{A}}\,.$$

$\mathcal{A} \subset \mathcal{B}(\mathcal{H}^{\mathrm{A}})$ is an embedding with the least possible Hilbert space dimension. (In contrast to (101), the general case.) Our working space will be

$$\mathcal{H}^{\mathrm{AB}} = \mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{B}}, \quad \dim \mathcal{H}^{\mathrm{B}} = \dim \mathcal{H}^{\mathrm{A}} = d\,. \tag{127}$$

The production of purifying vectors is simplified by first selecting a maximally entangled vector

$$|\varphi\rangle = \sum_{j=1}^{d} |jj\rangle \equiv \sum |j\rangle^{\mathrm{A}} \otimes |j\rangle^{\mathrm{B}} \tag{128}$$

of length d. $\{|j\rangle^{\mathrm{A}}\}$ and $\{|j\rangle^{\mathrm{B}}\}$ are bases of $\mathcal{H}^{\mathrm{A}}$ and of $\mathcal{H}^{\mathrm{A}}$ respectively. For any $X \in \mathcal{B}(\mathcal{H}^{\mathrm{A}})$ we get

$$(X \otimes \mathbf{1}^{\mathrm{A}})|\varphi\rangle = \sum X|j\rangle^{\mathrm{A}} \otimes |j\rangle^{\mathrm{B}}\,.$$

Bases are linearly independent. Hence if the right-hand side is zero, then $X = \mathbf{0}$. Because the dimension of $\mathcal{B}(\mathcal{H}^{\mathrm{a}})$ as a linear space is equal to the dimension of the Hilbert space (127), *every vector ψ in $\mathcal{H}^{\mathrm{AB}}$ has a unique representation* $(X \otimes \mathbf{1})|\varphi\rangle$.
One computes for $X_1, X_2 \in \mathcal{B}(\mathcal{H}^{\mathrm{A}})$ the partial trace

$$\psi_i = (X_i \otimes \mathbf{1})|\varphi\rangle \;\Rightarrow\; \mathrm{tr}_{\mathrm{B}}|\psi_1\rangle\langle\psi_2| = X_1 X_2^\dagger \,, \tag{129}$$

because we have to trace out the B-system in

$$\sum X_1(|j\rangle^{\mathrm{A}}\langle k|)X_2^\dagger \otimes (|j\rangle^{\mathrm{B}}\langle k|) \,.$$

Our choice of $\mathcal{A}$ implies $\Omega(\mathcal{A}) \subset \Omega(\mathcal{H}^{\mathrm{A}})$, see (91) and (92). Therefore, we can apply (129) above to the density operators of the A-system. Now let ω^{A} be a density or just a positive operator from $\mathcal{A}$. It is convenient to call an operator $W \in \mathcal{A}$ an *amplitude of* ω^{A} if $\omega^{\mathrm{A}} = WW^*$. $(W \otimes \mathbf{1})|\varphi\rangle$ is a purifying vector for ω^{A} if W is an amplitude of ω^{A} and vice versa.

There are many amplitudes of ω^{A} and the change from one to another one can be described[17] by *gauge transformations* $W \to W' = WU$ with unitary $U \in \mathcal{A}$. The gauge transformations respect ω^{A} as a gauge invariant.

Let us return to our problem with two density operators, ω_1^{A} and ω_2^{A}, and two purifying vectors, ψ_1 and ψ_2. There are two operators W_1, W_2 in our $\mathcal{A}$ satisfying

$$\psi_j = (W_j \otimes \mathbf{1})\varphi, \quad \omega_j^{\mathrm{A}} = W_j W_j^\dagger \,. \tag{130}$$

With these amplitudes we have

$$\langle\psi_1, \psi_2\rangle = \langle(W_1 \otimes \mathbf{1})\varphi, (W_2 \otimes \mathbf{1})\varphi\rangle = \mathrm{tr}\, W_1^\dagger W_2 \,. \tag{131}$$

Gauging $\psi_2 \to \psi_2'$ by $W_2 \to W_2' = W_2 U$, we see

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

Let us stress that we fix W_1 and vary only W_2 in this relation. Hence

$$F(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \sup_{\psi'} |\langle\psi_1, \psi_2'\rangle| = \sup_{U \in \mathcal{A}} |\mathrm{tr}\, W_1^\dagger W_2 U| \,,$$

provided one cannot get better results in higher dimensional purifications. But this is not the case, as one can prove. (Essentially, this is because the largest dimension of a cyclic representation of $\mathcal{B}(\mathcal{H}^{\mathrm{A}})$ is of dimension d^2.)

It is $|\mathrm{tr}\, BU| \le \mathrm{tr}\, B$ in case $B \ge \mathbf{0}$ and U is unitary. Hence we are done if $W_1^\dagger W_2 \ge \mathbf{0}$ can be reached. This is possible because the polar decomposition

[17] due to our finiteness assumptions

theorem is valid in $\mathcal{A}$ (and, indeed, in every von Neumann algebra). In other words, we can choose a pair of amplitudes such that

$$F(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{B}}) = \mathrm{tr}\, W_1^\dagger W_2, \quad W_1^\dagger W_2 \geq \mathbf{0}\,. \tag{132}$$

Let us restate (132) to respect *-isomorphisms. If $\mathcal{A}$ is any (finite dimensional) *-subalgebra of any $\mathcal{B}(\mathcal{H})$, we have to understand the trace in (132) as the canonical trace. (Remember: Only for the algebras $\mathcal{B}_{\mathbf{d}}$ with $\dim \mathcal{H} = \mathbf{d}$ the canonical trace coincides with the trace of $\mathcal{H}$.)
Whenever for two density operators of $\mathcal{A}$

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

we call the pair of amplitudes W_1, W_2 *parallel*. Parallelity implies

$$\mathrm{F}(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \mathrm{tr}^{\mathrm{can}}\, W_1^\dagger W_2, \quad \mathrm{Pr}(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = (\mathrm{tr}^{\mathrm{can}}\, W_1^\dagger W_2)^2\,. \tag{133}$$

3.4 Why the Bures distance is a distance

Before proceeding along the main line the triangle inequality should be proved. Inserting (132) into (124) yields:

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

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

$$W_1^\dagger W_2 = W_2^\dagger W_1\,. \tag{134}$$

Altogether we proved: If W_1, W_2 are parallel amplitudes then

$$\mathrm{dist}_B(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \sqrt{\mathrm{tr}^{\mathrm{can}}\,(W_1 - W_2)^\dagger (W_1 - W_2)}\,, \tag{135}$$

and for two arbitrary amplitudes the left-hand side cannot be larger than the right one. The latter can also be rewritten $\parallel W_2 - W_1 \parallel_2$.

Consider now three positive operators, ω_1^{A}, ω_2^{A}, and ω_3^{A}. Starting with W_2 we can choose W_1 and W_3 such that the pairs W_2, W_1 and W_2, W_3 are parallel amplitudes. This allows to convert the triangle inequality

$$\parallel W_1 - W_2 \parallel_2 + \parallel W_2 - W_3 \parallel_2 \geq \parallel W_1 - W_3 \parallel_2$$

into

$$\mathrm{dist}_B(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) + \mathrm{dist}_B(\omega_2^{\mathrm{A}}, \omega_3^{\mathrm{A}}) \geq \parallel W_1 - W_3 \parallel_2$$

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

$$\mathrm{dist}_B(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) + \mathrm{dist}_B(\omega_2^{\mathrm{A}}, \omega_3^{\mathrm{A}}) \geq \mathrm{dist}_B(\omega_1^{\mathrm{A}}, \omega_3^{\mathrm{A}})\,. \tag{136}$$

It is instructive to rewrite our finding with purifying vectors. We extend our notation and call a pair of purifying vectors *parallel* if the amplitudes in (130), that is in $\psi_j = (W_j \otimes \mathbf{1})\varphi$, are parallel ones. We can express (135) by

$$\mathrm{dist}_B(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) \leq \| \psi_2 - \psi_1 \| \tag{137}$$

for all pairs of purifying vectors of $\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}$. Equality holds for pairs of parallel purifying vectors.

Some geometric properties of the Bures distance

The Bures distance is an inner one: There are short geodesic arcs with length equal to the Bures distance of their end points. Given $\omega_0^{\mathrm{A}}, \omega_1^{\mathrm{A}}$ we choose parallel amplitudes W_0, W_1. Then any pair of amplitudes belonging to the arc

$$t \mapsto W_t = (1-t)W_0 + tW_1, \quad 0 \leq t \leq 1, \tag{138}$$

is a parallel pair. Exactly as in (10) we get

$$\mathrm{dist}_B(\omega_s^{\mathrm{A}}, \omega_t^{\mathrm{A}}) = \| W_t - W_s \|_2, \quad \omega_s^{\mathrm{A}} = W_s W_s^\dagger \, . \tag{139}$$

Bures did not ask wether his distance is based on a Riemannian 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 it is tempting to ask for.

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

$$\Big(\frac{ds_B}{dt}\Big)^2 = \mathrm{tr}^{\mathrm{can}} G^2 \omega^{\mathrm{A}} = \frac{1}{2} \mathrm{tr}^{\mathrm{can}} \dot{\omega}^{\mathrm{A}} G \, , \tag{140}$$

whenever there is a solution of

$$\dot{\omega}^{\mathrm{A}} = \omega^{\mathrm{A}} G + G \omega^{\mathrm{A}}, \quad G = G^\dagger \, . \tag{141}$$

For invertible positive operators ω^{A} there is a unique solution of (141). At the boundary, where the rank is smaller than the Hilbert space dimension, the existence of G depends on the direction of the tangent $\dot{\omega}^{\mathrm{A}}$. For $\dim \mathcal{H} \geq 3$ there are directions for which the metric becomes singular [36].

However, for invertible ω^{A} the metric behaves regularly. Let

$$t \to \omega_t^{\mathrm{A}} = W_t W_t^\dagger, \quad |\psi_t\rangle = (W_t \otimes \mathbf{1}^{\mathrm{B}})|\varphi\rangle, \tag{142}$$

be an arc of invertible density operators. The the curve $t \to W_t$ of the amplitudes are called *parallel,* if

$$\dot{W}_t^\dagger W_t = W_t^\dagger \dot{W}_t \, . \tag{143}$$

By straightforward computation one proves the equivalence of (141) with the condition

$$\dot{W}_t = G_t W_t, \quad G_t^\dagger = G_t \ . \tag{144}$$

Now one easily gets

$$t \to W_t \ \text{ parallel, } \ \Rightarrow \big(\frac{ds_B}{dt}\big)^2 = \mathrm{tr}^{\mathrm{can}}\, \dot{W}\dot{W}^\dagger \ . \tag{145}$$

It is an easy nice exercise to compute the Bures distance (139) by (145) to establish that the Bures distance can be gained from the metric (140).

After switching to the purifying arc $|\psi_t\rangle = (W_t \otimes \mathbf{1}^{\mathrm{B}})|\varphi$ another form of the results above appears: The Hilbert space length of a purifying lift $t \to |\psi_t\rangle$ of $t \to \omega_t^{\mathrm{A}}$ is never less than its Bures length. Equality is reached exactly with parallel amplitudes (143) in (142).

The extended Mandelstam-Tamm inequality

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

$$t \to \omega_t, \quad 0 \le t \le 1 \tag{146}$$

a solution of time dependent von Neumann - Schrödinger equation

$$i\hbar\dot{\omega} = [H, \omega], \quad H = H_t \ . \tag{147}$$

Then one can prove

$$\int_0^1 dt \sqrt{\mathrm{tr}(\omega H^2) - (\mathrm{tr}\,\omega H)^2} \ge \hbar \arccos \mathrm{F}(\omega_1, \omega_0) \ , \tag{148}$$

see [60]. (One has to look for a lift $t \to W_t$ satisfying the differential parallel condition $\dot{W}^\dagger W = W^\dagger \dot{W}$ and a Schrödinger equation with an Hamiltonian $W \to HW + W\tilde{H}$, where $t \to \tilde{H}_t$ has to be chosen suitably.)

Using this one can get a differential form of (148) :

$$\mathrm{tr}(\omega H^2) - (\mathrm{tr}\,\omega H)^2 \ge \frac{\hbar}{2} \mathrm{tr}\, G\dot{\omega} \ . \tag{149}$$

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 [32]. Another question has been discussed by A. Ericsson, [39].

3.5 Expressions for fidelity and transition probability

Now we return to (132) and (133) to benefit from the positivity of $W_1^\dagger W_2$ for parallel amplitudes. It holds

$$(W_1^\dagger W_2)^2 = W_1^\dagger W_2 W_2^\dagger W_1 = W_1^\dagger \omega_2^{\mathrm{A}} W_1 \,.$$

There is a polar decomposition

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

with a unitary U_1. Putting things together yields

$$(W_1^\dagger W_2)^2 = U_1^{-1} (\omega_1^{\mathrm{A}})^{1/2} \omega_2^{\mathrm{A}} (\omega_1^{\mathrm{A}})^{1/2} U_1 \tag{150}$$

we can take the positive root and obtain

$$W_1^\dagger W_2 = U_1^{-1} \sqrt{(\omega_1^{\mathrm{A}})^{1/2} \omega_2^{\mathrm{A}} (\omega_1^{\mathrm{A}})^{1/2}}\, U_1 \tag{151}$$

The canonical trace of (151) yields the fidelity,

$$\mathrm{F}(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \mathrm{tr}^{\mathrm{can}} \sqrt{(\omega_1^{\mathrm{A}})^{1/2} \omega_2^{\mathrm{A}} (\omega_1^{\mathrm{A}})^{1/2}}\,, \tag{152}$$

its square is the transition probability.

As an application we consider direct products. With two pairs, ω_1, ω_2 and ρ_1, ρ_2 of density operators in two different Hilbert spaces, one can perform their direct products $\omega_j \otimes \rho_j$. The structure of the expression (152) allows to conclude

$$\Pr(\omega_1 \otimes \rho_1, \omega_2 \otimes \rho_2) = \Pr(\omega_1, \omega_2)\, \Pr(\rho_1, \rho_2)\,. \tag{153}$$

In what follows we assume invertible positive operators though the results do not depend on that assumption. As above W_1, W_2 are parallel amplitudes of $\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}$. We define a positive gauge invariant, K,

$$W_1^\dagger W_2 > \mathbf{0} \;\Leftrightarrow\; W_2 = K W_1, \quad K > \mathbf{0}\,. \tag{154}$$

Indeed, $W_1^\dagger W_2 = W_1^\dagger K W_1$ proves $K > \mathbf{0}$ equivalent to parallelity. Now

$$K = W_2 W_1^{-1} > \mathbf{0}, \quad K^{-1} = W_1 W_2^{-1} > \mathbf{0}, \tag{155}$$

and we conclude the existence of $K \in \mathcal{A}$ such that

$$\mathrm{tr}^{\mathrm{can}} \omega_1^{\mathrm{A}} K = \mathrm{tr}^{\mathrm{can}} W_1^\dagger W_2 = \mathrm{tr}^{\mathrm{can}} \omega_2^{\mathrm{A}} K^{-1}\,. \tag{156}$$

But $\mathrm{tr}^{\mathrm{can}} W_1^\dagger W_2$ is the fidelity and with our K we have

$$\mathrm{F}(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \mathrm{tr}^{\mathrm{can}} \omega_1^{\mathrm{A}} K = \mathrm{tr}^{\mathrm{can}} \omega_2^{\mathrm{A}} K^{-1} \tag{157}$$

For a pair ψ_1, ψ_2 of parallel purifications and for every positive $C \in \mathcal{B}(\mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{B}})$ we know from (99)

$$\mathrm{F}(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = |\langle \psi_1, \psi_2 \rangle| \leq (1/2)(\langle \psi_1, C\psi_1 \rangle + \langle \psi_2, C^{-1}\psi_2 \rangle)$$

Inserting $C = X \otimes \mathbf{1}^{\mathrm{B}}$ it becomes clear, that the right-hand side cannot become smaller than

$$\frac{1}{2} \inf_{X > \mathbf{0}} (\mathrm{tr}\, \omega_1 X + \mathrm{tr}\, \omega_2 X^{-1}), \quad X \in \mathcal{A}\,.$$

The particular case $X = K$ proves

$$\mathrm{F}(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \frac{1}{2} \inf_{X > \mathbf{0}} (\mathrm{tr}\, \omega_1^{\mathrm{A}} X + \mathrm{tr}\, \omega_2^{\mathrm{A}} X^{-1})\,. \tag{158}$$

Let us reformulate (158) to change from density operators to states. Finally there is no reference on any bipartite structure.
Let $\mathcal{A}$ be a *-subalgebra of $\mathcal{B}(\mathcal{H})$ and $\underline{\omega}$ and $\underline{\rho}$ two of its states or positive linear forms. Then

$$\mathrm{F}(\underline{\omega}, \underline{\rho}) = \frac{1}{2} \inf_{\mathbf{0} < X \in \mathcal{A}} \underline{\omega}(X) + \underline{\rho}(X^{-1})\,, \tag{159}$$

$$\mathrm{Pr}(\underline{\omega}, \underline{\rho}) = \inf_{\mathbf{0} < X \in \mathcal{A}} \underline{\omega}(X)\, \underline{\rho}(X^{-1})\,. \tag{160}$$

Thanks to the work of Araki and Raggio, [25], and Alberti, [24], the two last assertions are known to be true for any pair of states of any unital C*-algebra.

Super-additivity

For all decompositions

$$\omega = \sum \omega_j, \quad \rho = \sum \rho_j \tag{161}$$

of positive operators the inequality

$$\mathrm{F}(\omega, \rho) \geq \sum_j \mathrm{F}(\omega_j, \rho_j) \tag{162}$$

is valid. The inequality expresses "super-additivity" of the fidelity.

For simplicity we prove super-additivity assuming ω and ρ invertible and choose $K \in \mathcal{A}$ satisfying

$$\mathrm{F}(\omega, \rho) = \mathrm{tr}^{\mathrm{can}} \omega K = \mathrm{tr}^{\mathrm{can}} \rho K^{-1}$$

as in (156) and (157). We now have

$$2\mathrm{F}(\omega,\rho) = \sum \mathrm{tr}^{\mathrm{can}}\omega_j K + \sum \mathrm{tr}^{\mathrm{can}}\rho_j K^{-1} \; .$$

The proof terminates by estimating the right part by (158). This is the finite dimensional case. (For von Neumann and C^*-algebras one returns to positive linear forms for which super-additivity of the fidelity can be proved equally well.)

Let us mention how (161) implies joint concavity. Because of

$$\mathrm{F}(a\omega, b\rho) = \sqrt{ab}\,\mathrm{F}(\omega,\rho), \quad a,b \in \mathbb{R}^+ \tag{163}$$

it follows from (161) for convex sums of equal length

$$\mathrm{F}(\sum_j p_j\omega_j, \sum_k q_k\rho_k) \geq \sum_j \sqrt{p_j q_j}\,\mathrm{F}(\omega_j,\rho_j) \; . \tag{164}$$

From (152) one can conclude: *Equality holds in (164) if for $j \neq k$ it holds $\omega_j\rho_k = \mathbf{0}$*. Similar (indeed equivalent) statements are true for (159) and (160).

Monotonicity

Choi, [35], proved for positive unital maps

$$\Psi(A^{-1}) \geq \Psi(A)^{-1} \text{ if } A \geq \mathbf{0} \; . \tag{165}$$

In the case of a 2-positive and unital Ψ the conclusion

$$\begin{pmatrix} A & C \\ C^\dagger & B \end{pmatrix} \geq \mathbf{0} \Rightarrow \begin{pmatrix} \Psi(A) & \Psi(C) \\ \Psi(C^\dagger) & \Psi(B) \end{pmatrix} \geq \mathbf{0} \tag{166}$$

comes simply from the very definition of 2-positivity. Then (165) follows with $B = A^{-1}$, $C = \mathbf{1}$ and unitality, $\Psi(\mathbf{1}) = \mathbf{1}$. However, according to Choi, in the particular case $C \geq \mathbf{0}$ just positivity and unitality are sufficient for the validity of (166). Therefore, (165) is valid for positive unital maps.

Let us apply (165) to the fidelity. To this end we denote by Φ the map dual to Ψ,

$$\mathrm{tr}\, X\Psi(Y) = \mathrm{tr}\,\Phi(X)Y \; . \tag{167}$$

Ψ is positive if Φ is positive. Φ is trace preserving if Ψ is unital. Not every positive operator might be of the form $\Psi(X)$ with positive X. Therefore, by (158) or (159),

$$\mathrm{F}(\omega,\rho) \leq \frac{1}{2} \inf_{\mathbf{0}<X\in\mathcal{A}} \mathrm{tr}^{\mathrm{can}}\Psi(X)\omega \, + \, \mathrm{tr}^{\mathrm{can}}\Psi(X)^{-1}\rho \; .$$

We can replace $\Psi(X)^{-1}$ by the smaller $\Psi(X^{-1})$ in virtue of (165) to get an even larger right hand side:

$$\mathrm{F}(\omega,\rho) \leq \frac{1}{2} \inf_{\mathbf{0}<X\in\mathcal{A}} \mathrm{tr}^{\mathrm{can}}\Psi(X)\omega \,+\, \mathrm{tr}^{\mathrm{can}}\Psi(X^{-1})\rho \,.$$

Now we apply duality, (167), and obtain

$$\mathrm{F}(\omega,\rho) \leq \frac{1}{2} \inf_{\mathbf{0}<X\in\mathcal{A}} \mathrm{tr}^{\mathrm{can}} X\Phi(\omega) \,+\, \mathrm{tr}^{\mathrm{can}} X^{-1}\Phi(\rho) \,.$$

The right-hand side is an expression for the fidelity of the pair $\Phi(\omega), \Phi(\rho)$ and the proof of the *monotonicity property* is done:
Let Φ be positive and trace preserving. Then

$$\mathrm{F}(\omega_1,\omega_2) \leq \mathrm{F}(\Phi(\omega_1),\Phi(\omega_2)) \,. \tag{168}$$

As a consequence, trace preserving positive maps are "Bures-contracting",

$$\mathrm{dist}_B(\omega_1,\omega_2) \geq \mathrm{dist}_B(\Phi(\omega_1),\Phi(\omega_2)) \,. \tag{169}$$

Density operators (and states) become closer one to another under the action of these maps.

Remark: It is well known that there are many Riemannian metrics in a state space $\Omega(\mathcal{H})$ which are monotone decreasing with respect to channels, i. e. with respect to completely positive and trace preserving maps[18]. Thanks to the work of Petz, [53], they can be constructed by the help of certain operator means. Kubo and Ando,[49], could enumerate all operator means by operator monotone functions. Another, but much related story is the question for functions $\mathrm{Pr}'(.,.)$, depending on two states, which are
a) monotone increasing with respect to channels and which
b) coincide with the transition probability for pure states.
Some of them are related to distances, i. e. inserting in dist_B the square root of Pr' for F returns a distance. Most of them, however, are not related to any distance. But what our alternative choice may be, only the transition probability (118) is "operational" defined. Just by this very definition one finds, for pairs of density operators,

$$\mathrm{Pr}(\omega,\rho) \geq \mathrm{Pr}'(\omega,\rho) \tag{170}$$

for all Pr' satisfying the two condition a) and b) above. A nice example is

$$\mathrm{Pr}(\omega,\rho) \geq \mathrm{tr}\,\omega^{1-s}\rho^s, \quad 0<s<1 \,. \tag{171}$$

Indeed, the right hand side fulfills a) and b) above, see [30], where one can also find a more "direct" proof of (171).

[18] Though their geodesics and distances are mostly unknown.

3.6 Estimates and a "hidden symmetry"

I use the notation

$$\mathrm{char}(A) = \text{ all roots of the characteristic equation of } A\,. \tag{172}$$

Clearly, this are the eigenvalues, counted with the appropriate multiplicity, if A is diagonalisable. 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

$$\mathrm{char}(\omega_1^{1/2}\omega_2\omega_1^{1/2}) = \mathrm{char}(\omega_1\omega_2)\,. \tag{173}$$

An estimate

Denoting the characteristic values of (173) by $\lambda_1, \lambda_2, \ldots$, we get

$$\Pr(\omega_1, \omega_2) = (\sum \sqrt{\lambda_j})^2\,. \tag{174}$$

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

$$\Pr(\omega_1, \omega_2) = \mathrm{tr}\,\omega_1\omega_2 + 2\sum_{j<k}\sqrt{\lambda_j\lambda_k}\,.$$

We write $2r$ for the last term and use

$$\sqrt{r^2} = \sqrt{\sum_{j<k}\lambda_j\lambda_k + \ldots}\,.$$

The dots abbreviate some non-negative terms. The other term in the sum is the second elementary symmetric function of the characteristic values λ_k of $\omega_1\omega_2$. Expressing the latter by traces yields

$$\Pr(\omega_1, \omega_2) \geq \mathrm{tr}\,\omega_1\omega_2 + \sqrt{2}\sqrt{(\mathrm{tr}\,\omega_1\omega_2)^2 - \mathrm{tr}(\omega_1\omega_2)^2} \tag{175}$$

with equality for $\mathrm{rank}(\omega_1\omega_2) \leq 2$. For $\dim\mathcal{H} = 3$ closer inspection produces

$$\Pr(\rho, \omega) = \mathrm{tr}\,\rho\omega + \sqrt{2}\sqrt{(\mathrm{tr}\,\rho\omega)^2 - \mathrm{tr}(\rho\omega\rho\omega) + 4F(\rho, \omega)\sqrt{\det(\rho\omega)}}\,.$$

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

In the one qubit case (175) becomes an equality. $\lambda_1\lambda_2$ is the determinant of $\omega_1\omega_2$. Thus

$$\Pr(\omega_1, \omega_2) = \mathrm{tr}\,\omega_1\omega_2 + 2\sqrt{\det\omega_1\,\det\omega_2}\,. \tag{176}$$

Let us represent our density matrices by

$$\omega_1 = \frac{1}{2}(\mathbf{1} + \sum x_n \sigma_n), \quad \omega_2 = \frac{1}{2}(\mathbf{1} + \sum y_n \sigma_n) \tag{177}$$

and let us define a new coordinate by

$$x_4 := 2\sqrt{\det \omega_1}, \quad y_4 := 2\sqrt{\det \omega_2}\ . \tag{178}$$

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

$$x_1^2 + \ldots + x_4^2 = y_1^2 + \ldots + y_4^2 = 1 \tag{179}$$

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

$$\Pr(\omega_1, \omega_2) = \frac{1}{2}(1 + \sum_{j=1}^{4} x_j y_j)\ . \tag{180}$$

A "hidden symmetry"

Remember first the equality (173) for the characteristic numbers. Let Z be invertible and consider the change

$$\omega_1' = Z^{-1}\omega_1 (Z^{-1})^\dagger, \quad \omega_2' = Z^\dagger \omega_2 Z\ . \tag{181}$$

One immediately sees

$$\omega_1'\omega_2' = Z^{-1}(\omega_1\omega_2)Z \tag{182}$$

and

$$\mathrm{char}(\omega_1'\omega_2') = \mathrm{char}(\omega_1\omega_2)\ . \tag{183}$$

Now (173) implies:
The eigenvalues of $\sqrt{\omega_1}\omega_2\sqrt{\omega_1}$ *do not change if* ω_1, ω_2 *are transformed according to (181).* In particular

$$F(\omega_1, \omega_2) = F(Z^{-1}\omega_1(Z^{-1})^\dagger, Z^\dagger\omega_2 Z)\ . \tag{184}$$

Indeed, the argument is valid for every symmetric function of the characteristic numbers in question. We can even refrain from the invertibility of Z by substituting

$$\omega_1 \to Z\omega_1 Z^\dagger$$

in (184):

$$F(Z\omega_1 Z^\dagger, \omega_2) = F(\omega_1, Z^\dagger\omega_2 Z)\ . \tag{185}$$

Relaying on continuity we can state (185) for all operators Z.

3.7 "Operational fidelity"

The question is, whether the fidelity concept can be extended to pairs of quantum channels. It seems, the first relevant studies are done by Raginski, [55]. More recent developments can be seen from Belavkin et al., [31], and Kretschmann et al., [48]. Essentially, our aim is to define what is called *operational fidelity* and to arrive at it via the Bures distance. We restrict ourselves to maps Φ from $\mathcal{B}(\mathcal{H})$ into itself, and we assume $\dim \mathcal{H} = d$ finite.

We denote by I the identity map $I(X) = X$, $X \in \mathcal{B}(\mathcal{H})$. Later on we need the identity maps I_k of auxiliary algebras $\mathcal{B}(\mathcal{H}_k)$ with $\dim \mathcal{H}_k = k$.

With two positive maps, Φ_1 and Φ_2, and a density operator $\omega \in \Omega(\mathcal{H})$ we observe that

$$\Phi_1, \Phi_2 \to \mathrm{dist}_B(\Phi_1(\omega), \Phi_2(\omega))$$

is symmetric in the maps and fulfils, for three positive maps, the triangle inequality. This is because the Bures distance does so. As $\Phi_1(\omega) = \Phi_2(\omega)$ may happen, we do not necessarily get a metrical distance, but only a "semi-distance". As one can check, the sup of arbitrary many semi-distances is again a semi-distance. Therefore,

$$\mathrm{dist}_1(\Phi_1, \Phi_2) := \sup_{\omega \in \Omega} \mathrm{dist}_B(\Phi_1(\omega), \Phi_2(\omega)) \tag{186}$$

is a distance in the space of positive maps.
Indeed, as said above, it is a semi-distance. But if two maps are not equal one to another, there must be a density operator at which they take different values. The index "1" in (186) reflects our assumption that the maps are just positive, i. e. 1-positive. Obviously, the distance $\mathrm{dist}_1(\Phi, I)$ estimates how strongly Φ deviates from the identity map.

If a map Φ is k-positive, then the map $\Phi \otimes I_k$ is still positive. For pairs of k-positive maps the expression

$$\mathrm{dist}_k(\Phi_1, \Phi_2) := \mathrm{dist}_1(\Phi_1 \otimes I_k, \Phi_2 \otimes I_k) \tag{187}$$

is well defined. More explicitly, (187) is a sup,

$$\sup_{\rho} \mathrm{dist}_B([\Phi_1 \otimes I_k](\rho), [\Phi_2 \otimes I_k](\rho)) \ , \tag{188}$$

over all density operators $\rho \in \Omega(\mathcal{H} \otimes \mathcal{H}_k)$.

One can unitarily embed $\mathcal{H} \otimes \mathcal{H}_k$ into $\mathcal{H} \otimes \mathcal{H}_{k+1}$. This implies that the sup in (188) is running over less states as in the case dist_{k+1}, resulting in

$$\mathrm{dist}_k(\Phi_1, \Phi_2) \leq \mathrm{dist}_{k+1}(\Phi_1, \Phi_2) \ . \tag{189}$$

This enables the introduction of an *operational Bures distance*

$$\mathrm{dist}_\infty(\Phi_1, \Phi_2) := \lim_{k \to \infty} \mathrm{dist}_k(\Phi_1, \Phi_2) \tag{190}$$

which takes into account possible entanglement in the input. Possible entanglement based on $\mathcal{H} \otimes \mathcal{H}_k$ saturates with $\dim \mathcal{H} = k$,

$$\mathrm{dist}_\infty(\Phi_1, \Phi_2) = \mathrm{dist}_d(\Phi_1, \Phi_2), \quad d = \dim \mathcal{H} \,. \tag{191}$$

To obtain what has been called *operational fidelity* in the literature, we have to go back to the relation (124)

$$\mathrm{dist}_B(\omega_1, \omega_2) = \sqrt{\mathrm{tr}\,\omega_1 + \mathrm{tr}\,\omega_2 - 2\mathrm{F}(\omega_1, \omega_2)}$$

and try to replace accordingly states by maps. The task can be done quite naturally for completely positive, trace preserving maps: For all these maps $\mathrm{tr}\,\Phi(\omega) = 1$ for density operators. It suggests

$$\mathrm{dist}_\infty(\Phi_1, \Phi_2) = \sqrt{2 - 2\mathrm{F}_{(\infty)}(\Phi_1, \Phi_2)} \,. \tag{192}$$

The quantity $\mathrm{F}_{(\infty)}(\Phi_1, \Phi_2)$ is called "operational fidelity". The index (∞) is not standard and stands here only to respect the possibility of the same procedure with k-positive and trace preserving maps. For such maps one can consider an "operational fidelity for k-positivity", $\mathrm{F}_{(k)}$, as well.

As a matter of fact, one can transmit several properties of the fidelity and the Bures distance for states to completely or k-positive and trace preserving maps. The joint concavity for instance allows to perform the sup in (186) or in (187) over pure states only.

4 Appendix: The geometrical mean

Let A, B, and C be positive operators in a finite dimensional Hilbert space. A remarkable observation due to Pusz and Woronowicz, [54], can be rephrased in the following form:

Given $A \geq \mathbf{0}$ and $B \geq \mathbf{0}$, there is a largest operator in the set of all C satisfying

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq \mathbf{0}, \quad C \geq \mathbf{0} \,. \tag{193}$$

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

$$A \# B \,. \tag{194}$$

In other words: (193) is valid if and only if

$$\mathbf{0} \leq C \leq A \# B \,. \tag{195}$$

From the definition we get the relation

$$A \# B = B \# A, \quad A^{-1} \# B^{-1} = (A \# B)^{-1} \,, \tag{196}$$

the latter is true for invertible positive operators. If just A is invertible then the block matrix (193) is positive if and only if

$$B \geq CA^{-1}C \tag{197}$$

and one concludes: $A\#B$ is the unique positive solution X of the equation

$$B = XA^{-1}X, \quad X \geq \mathbf{0}\,. \tag{198}$$

The equation can be solved and one gets

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}\,. \tag{199}$$

To prove it, one rewrites (198) as

$$A^{-1/2}BA^{-1/2} = \left(A^{-1/2}XA^{-1/2}\right)^2$$

and takes the root.

In case B^{-1} exists too, we may rewrite (198) as

$$\mathbf{1} = B^{-1/2}XA^{-1}XB^{-1/2} = (B^{-1/2}XA^{-1/2})\,(B^{-1/2}XA^{-1/2})^\dagger\,.$$

Therefore, if A und B are strictly positive, the following three statements are equivalent: a) $X = A\#B$, b) $(B^{-1/2}XA^{-1/2})$ is unitary, c) $(A^{-1/2}XB^{-1/2})$ is unitary.

In turn, as shown in [29], an operator $X = (A^{1/2}UB^{1/2})$ with unitary U is positive exactly if $X = A\#B$.

If A and B commute one can see from (193)

$$AB = BA \;\Rightarrow\; A\#B = (AB)^{1/2} \tag{200}$$

To get it one uses a common eigenbasis which reduces (193) to

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

for three positive numbers a, b, and c.

4.1 Geometric mean and fidelity

Let us return to the operator K, defined in (154) by

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

for parallel and invertible amplitudes. One can calculate

$$K = \omega_1^{-1/2}(\omega_1^{1/2}\omega_2\omega_1^{1/2})^{1/2}\omega_1^{-1/2} \tag{201}$$

and conclude: At first we can rewrite (201)

$$K = \omega_2 \# \omega_1^{-1} \,, \tag{202}$$

and, secondly, the expression remains meaningful for not invertible ω_j. In this sense we understand the right-hand side of (201) and (202) to be valid for all pairs of positive operators in our finite dimensional setting. One further concludes that for general pairs of positive operators we have to substitute

$$K \to \omega_2 \# \omega_1^{-1} \text{ and } K^{-1} \to \omega_1 \# \omega_2^{-1}$$

in order that (156) and (157) can be applied to not necessarily invertible positive operators. With this convention one arrives at

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

For parallel purifications in $\mathcal{H}^{\mathrm{AB}} = \mathcal{H}^{\mathrm{A}} \otimes \mathcal{H}^{\mathrm{A}}$ of states ω_1^{A} and ω_2^{A} one may ask, how these purifications appear in Bob's system, i. e. after tracing out the A-system. To this end let us first review the general situation, not requiring parallelity. We choose amplitudes, W_1, W_2, for ω_1^{A} and ω_2^{A}, so that, for $m = 1, 2$,

$$\omega_j^{\mathrm{A}} = W_m W_m^\dagger = \mathrm{tr}_{\mathrm{B}} |\psi_m\rangle\langle\psi_m|, \quad |\psi_m\rangle = (W_m \otimes \mathbf{1}^{\mathrm{B}})|\varphi\rangle \tag{203}$$

with $|\varphi\rangle$ chosen maximally entangled and with norm $d = \dim \mathcal{H}^{\mathrm{A}}$ as in (128). For any $Y \in \mathcal{B}(\mathcal{H}^{\mathrm{B}})$ we get

$$\mathrm{tr}\,(|\psi_n\rangle\langle\psi_m|)^{\mathrm{B}} Y = \langle\psi_m|(\mathbf{1}^{\mathrm{A}} \otimes Y)|\psi_n\rangle = \langle\varphi|(W_m^\dagger W_n \otimes Y)|\varphi\rangle \,. \tag{204}$$

The last expression is equal to

$$\sum \langle j|W_m^\dagger W_n|k\rangle \, \langle j|Y|k\rangle \,.$$

Let us define the transposition $X \to X^\top$ by

$$\langle j|X^\top|k\rangle_{\mathrm{B}} = \langle k|X|j\rangle_{\mathrm{A}} \,. \tag{205}$$

In this equation we used the two basis of the subsystems with which φ is represented in (128). It now follows

$$\mathrm{tr}\,(|\psi_n\rangle\langle\psi_m|)^{\mathrm{B}} Y = \mathrm{tr}\,(W_m^\dagger W_n)^\top Y$$

and, therefore,

$$(|\psi_n\rangle\langle\psi_m|)^{\mathrm{B}} = (W_m^\dagger W_n)^\top \,. \tag{206}$$

Next we conclude, for $m = 1, 2$,

$$\omega_m^{\mathrm{B}} := (|\psi_m\rangle\langle\psi_m|)^{\mathrm{B}} = (W_m^\dagger W_m)^\top \ . \tag{207}$$

We are now prepared to use parallel amplitudes and proceed in showing

$$W_m^\dagger W_n \geq \mathbf{0} \ \Rightarrow \ W_m^\dagger W_n = (W_m^\dagger W_m)\#(W_n^\dagger W_n) \ . \tag{208}$$

At first we assume invertibility of W_m and use the identity

$$(W_n^\dagger W_n) = (W_n^\dagger W_m)(W_m^\dagger W_m)^{-1}(W_m^\dagger W_n)$$

which is true for parallel amplitudes. According to (198) our assertion must be true. In the general case we mention that (193) forces the geometric mean $A\#B$ to give zero if applied to any null-vector of either A or B. Therefore, the support of $A\#B$ is the intersection of the supports of A and of B. Thus A and B become invertible if restricted onto the support of their geometric mean and the reasoning above applies.

We can now return to (208) and state:
If W_m are parallel amplitudes for ω_m^{A} then

$$W_1^\dagger W_2 = (\omega_1^{\mathrm{B}}\#\omega_2^{\mathrm{B}})^\top \tag{209}$$

and, in particular,

$$F(\omega_1^{\mathrm{A}}, \omega_2^{\mathrm{A}}) = \mathrm{tr}\,(\omega_1^{\mathrm{B}}\#\omega_2^{\mathrm{B}}) \ . \tag{210}$$

4.2 The transformer identity

To get further insight one may use the fact that $A\#B$ is an *operator mean.* In particular it satisfies the so-called *transformer identity,* i.e. for invertible Z it enjoys

$$Z(A\#B)Z^\dagger = (ZAZ^\dagger)\#(ZBZ^\dagger) \ . \tag{211}$$

For the proof one relays on

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

for invertible Z.

Now we can combine (200) and (211) with $Z = A + B$. To check the positivity of (193) it is sufficient to do so on the support space of $A+B$. Thus we may assume that this operator is invertible. Then

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

commute. Indeed, it follows

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

and we can apply (211). Therefore we can express $A\#B$ by

$$(A+B)^{1/2}\big((A+B)^{-1/2}A(A+B)^{-1}B(A+B)^{-1/2}\big)^{1/2}(A+B)^{1/2}\ . \quad (212)$$

Resume: *An expression is equal to $A\#B$ if it do so for commuting positive operators and if it satisfy the transformer identity.*

A further application arises from the integral

$$\mathbf{1}\#A = \sqrt{A} = \frac{1}{\pi}\int_0^1 (xA^{-1} + (1-x)\mathbf{1})^{-1}\frac{dx}{\sqrt{x(1-x)}}\ .$$

Substituting $A \to B^{-1/2}AB^{-1/2}$ the transformer identity

$$B\#A = B^{1/2}\ (\mathbf{1}\#(B^{-1/2}AB^{-1/2}))\ B^{1/2}$$

allows to infer algebraically

$$A\#B = \frac{1}{\pi}\int_0^1 (xA^{-1} + (1-x)B^{-1})^{-1}\frac{dx}{\sqrt{x(1-x)}}\ . \quad (213)$$

Super-additivity

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

$$A\#B \geq \sum A_j\#B_j\ . \quad (214)$$

This nice inequality is the key to further estimates. We enumerate m positive operators, $A_1, \ldots A_m$, modulo m by $A_{k+m} = A_m$. Then we set $B_j = A_{j+1}$. Then the sum A of m consecutive A_j is equal to that of the B_j. Then, on the left side of (214), we have $A\#B = A$. (214) yields

$$\sum_1^m A_j \geq \sum_j^m A_j\#A_{j+1}\,. \quad (215)$$

Take $m = 2$ as a particular case and respect (196).We get

$$(A+B)/2 \geq A\#B\ .$$

Replacing A and B by A^{-1} and B^{-1} we get

$$(A^{-1} + B^{-1})/2 \geq (A\#B)^{-1}$$

by (196). Taking the inverse of that inequality:

$$A\#B \geq 2(A^{-1} + B^{-1})^{-1} \ .$$

The right-hand side is the *harmonic mean* of A and B.

A hint to further developments: It is not obvious how to define a "geometrical mean" of more than two operators. One of the proposals is by Ando et al., [29]. It fits to the equality (215) and we describe it for just three positive operators, A, B, C. We define recursively

$$A_{j+1} = B_j\#C_j, \quad B_{j+1} = C_j\#A_j, \quad C_{j+1} = A_j\#B_j \ , \tag{216}$$

starting with $A_0 = A, B_0 = B, C_0 = C$. (215) proves (216) to be a decreasing sequence of positive operators. There is, therefore, a limiting operator which is, up to a factor, the geometric mean $G(A, B, C)$ favored by Ando, Li, and Mathias,

$$G(A, B, C) = \frac{1}{3} \lim_{j\to\infty} A_j + B_j + C_j \ . \tag{217}$$

For three commuting operators one gets $(ABC)^{1/3}$. However, for three positive Operators in general position no explicit expression is known for (217) — even if the operators live on a 2-dimensional Hilbert space.

Monotonicity

Our next task is to prove a monotonicity theorem. Let Ψ be a positive superoperator and $\Psi(\mathbf{1}) > \mathbf{0}$. According to Choi, [35], Ψ is "almost" 2-positive: A 2×2 positive block matrix with Hermitian off-diagonal remains positive by applying Ψ. That is

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geq \mathbf{0}, \ \ C = C^\dagger \Rightarrow \begin{pmatrix} \Psi(A) & \Psi(C) \\ \Psi(C) & \Psi(B) \end{pmatrix} \geq \mathbf{0} \ . \tag{218}$$

Therefore, applied to $C = A\#B$, with

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

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

$$\Psi(A\#B) \leq \Psi(A)\#\Phi(B) \tag{219}$$

is valid by the very definition of the geometric mean.

A rank criterion

Here we like to prove the following:
Let A, B, and C be positive operators in a Hilbert space of dimension d. We further assume that A^{-1} and B^{-1} exist. With these data we consider the matrix

$$X = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \tag{220}$$

which is an operator in $\mathcal{B}(\mathcal{H} \oplus \mathcal{H})$. Then

$$\operatorname{rank} X \geq d = \dim \mathcal{H} \,, \tag{221}$$

and equality holds if and only if $C = A\#B$.

Proof: At first we mention that the invertibility of A and B is essential. It allows to introduce the matrix

$$Y = \begin{pmatrix} \mathbf{1} & D \\ D^\dagger & \mathbf{1} \end{pmatrix} := \begin{pmatrix} A^{-1/2} & \mathbf{0} \\ \mathbf{0} & B^{-1/2} \end{pmatrix} X \begin{pmatrix} A^{-1/2} & \mathbf{0} \\ \mathbf{0} & B^{-1/2} \end{pmatrix} \,, \tag{222}$$

which is of the same rank as X. The set of eigenvectors of Y with eigenvalue 0 is a subspace $\mathcal{H}_0$ of $\mathcal{H} \oplus \mathcal{H}$. The unit vector $\psi \oplus \varphi$ belongs to $\mathcal{H}_0$ if

$$\psi + D\varphi = 0, \quad D^\dagger \psi + \varphi = 0 \,.$$

At first we see that neither ψ nor φ can be the zero-vector of $\mathcal{H}$. (Otherwise both, ψ and φ must be zero.) Hence, the dimension of $\mathcal{H}_0$ cannot exceed d, confirming (221) (This is so because otherwise one must have a non-zero vector in $\mathcal{H}_0$ with either ψ or φ equal to the zero-vector. This we had already excluded.) Secondly, we deduce, by eliminating ψ respectively φ,

$$(\mathbf{1} - DD^\dagger)\psi = 0, \quad (\mathbf{1} - D^\dagger D)\varphi = 0 \,.$$

If $\dim \mathcal{H}_0 = d$, then these equations are valid for all $\psi \in \mathcal{H}$ and all $\varphi \in \mathcal{H}$. Therefore, D is unitary if the ranks of Y and X are equal to d. Then $D = A^{-1/2} C B^{-1/2}$ is unitary. However, we already know that this can be true if and only if $C = A\#B$. Finally, if D is unitary, we easily can find d linear independent vectors of Y with eigenvalue zero.

Geometrical mean in dimension two

Let A, B be two positive operators acting on a 2-dimensional Hilbert space. Explicit expressions for the geometric mean are known:

$$A\#B = \frac{\sqrt{st}}{\sqrt{\det(A/s + B/t)}}(A/s + B/t)$$

with

$$s = \sqrt{\det A}, \quad t = \sqrt{\det B} \,.$$

4.3 #-Convexity

To handle more than two positive operators in general position is a very hard task. One of these problems we like to pose. We think it is important and, perhaps, not completely hopeless.

A set $\mathcal{K}$ of positive operators is called #-convex , if
a) K contains all its limiting operators and
b) K contains with A, B also $A\#B$.

Let us denote by $\#[A, B]$ the smallest #-convex set containing A and B. Assuming $AB = BA$ and both invertible, then

$$\#[A, B] = \{A^s B^{1-s} \,|\, 0 \leq s \leq 1\} \tag{223}$$

If the operators A and B are not invertible, then A^0 and B^0 must be interpreted as the projections P_A and P_B onto the support of A and B respectively. The general case of two non-commuting positive operators can be settled by the transformer inequality

$$Z\,\#[A, B]\,Z^\dagger = \#[ZAZ^\dagger, ZBZ^\dagger]\,. \tag{224}$$

Denote by $\#[A_1, A_2, \ldots, A_m]$ the smallest #-convex set containing $A_1, \ldots, A_m$. If the operators A_j are invertible and pairwise commuting, it is not hard to show that $\#[A_1, A_2, \ldots, A_m]$ consists of all operators

$$A_1^{s_1} A_2^{s_2} \cdots A_m^{s_m}, \quad \sum s_j = 1, \tag{225}$$

with all $s_j \geq 0$. What happens without commutativity is unknown.

References

1. E. F. Beckenbach und R. Bellman: *Inequalities* (Springer Berlin Heidelberg 1961).
2. R. Bhatia: *Matrix Analysis*, Graduate Texts in Mathematics 169 (Springer Berlin New York 1996).
3. S. Boyd and L. Vandenberghe: *Convex Optimization* (Cambridge University Press Cambridge 2004).
4. J. L. Brylinski and G. Chen: *Mathematics of Quantum Computation* (Chapman & Hall Boca Raton London New York 2002).
5. P. Busch, P. Lahti, and P. Mittelstaedt: *The Quantum Theory of Measurement*, 2nd edn (Springer Berlin New York 1996).
6. D. Chruściński and A. Jamiolkowski: *Geometric Phases in Classical and Quantum Mechanics* (Birkhäuser Boston 2004).
7. P. A. M. Dirac: *The Principles of Quantum Mechanics* (Clarendon Press Oxford 1930).
8. W. Fulton: *Young Tableaux* (Cambridge University Press Cambridge 1997).
9. J. Gruska: *Quantum Computing* (McGraw-Hill London 1999).

10. R. Haag: *Local Quantum Physics: Fields, Particles, Algebras* (Springer Berlin New York 1993).
11. R. Horn and C. R. Johnson: *Topics in Matrix Analysis*(Cambridge University Press Cambridge 1994).
12. R. A. Horn and C. R. Johnson: *Matrix Analysis* (Cambridge University Press Cambridge 1990).
13. O. Johnson: *Information Theory and the Central Limit Theorem* (World Scientific Singapore 2004).
14. J. v. Neumann: *Mathematische Grundlagen der Quantenmechanik* (Springer Berlin 1932).
15. M. Nielsen and I. Chuang: *Quantum Computation and Quantum Information* (Cambridge University Press Cambridge 2000).
16. M. Ohya and D. Petz: *Quantum Entropy and Its Use* (Springer Berlin New York 1993).
17. V. Paulsen: *Completely Bounded Maps and Operator Algebras*, Cambridge Studies in Advanced Mathematics 78 (Cambridge University Press Cambridge 2002).
18. R. T. Rockafellar: *Convex Analysis* (Princeton University Press Princeton 1970).
19. W.-H. Steeb and Y. Hardy: *Problems and Solutions in Quantum Computing and Quantum Information* (World Scientific Singapore 2004).
20. W. Thirring: *Lehrbuch der mathematischen Physik IV: Quantenmechanik großer Systeme* (Springer Wien New York, 1980) and
A Course in Mathematical Physics IV: Quantum Mechanics of Large Systems (Springer Wien New York 2002).
21. J. H. M. Wedderburn: *Lectures on Matrices* (American Mathematical Society New York 1934). Online version: www.ams.org.
22. E. Wigner: *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Vieweg Braunschweig 1931).
23. I. Bengtsson and K. Życzkowski: *Geometry of Quantum States.* (Cambridge University Press, to appear 2006).
24. P. M. Alberti: *A note on the transition probability over C^*-algebras.* Lett. Math. Phys. **7**, 25–32 (1983).
25. H. Araki and G. Raggio: *A remark on transition probability.* Lett. Math. Phys. **6**, 237–240 (1982).
26. P. M. Alberti: *On the geometry of pairs of positive linear forms (I).* Wiss. Z. KMU Leipzig, MNR **3**9, 579–597 (1990); P. M. Alberti: *A study on the geometry of pairs of positive linear forms, algebraic transition probablities and geometric phase over non-commutative operator algebras (I).* Z. Anal. Anwendungen **1**1, 293–334 (1992); V. Heinemann: *Geometrie und Transformation von Paaren positiver Linearformen über C^*–Algebren.* Dissertation, Universität Leipzig (1991).
27. P. M. Alberti and A. Uhlmann: *On Bures-distance and *–algebraic transition probability between inner derived positive linear forms over W^*–algebras.* Acta Applicandae Mathematicae **6**0, 1–37 (2000).
see also: P. M. Alberti and A. Uhlmann: *Transition probabilities on C^*- and W^*-algebras.* In: *Proceedings of the Second International Conference on Operator Algebras, Ideals, and Their Applications in Theoretical Physics, Leipzig 1983*, ed by H. Baumgärtel, G. Laßner, A. Pietsch, and A. Uhlmann, Teubner-Texte zur Mathematik 67 (B.G.Teubner Leipzig 1984) pp 5–11.

28. J. Anandan: *A geometric view of quantum mechanics.* In: *Quantum Coherence*, ed. by J. S. Anandan (World Scientific Singapore 1990)
J. Anandan: *A geometric approach to quantum mechanics.* Found. Phys. **21**, 1265-1284 (1991).
29. T. Ando, Chi-Kwang Li, and R. Mathias: *Geometric means.* Lin. Algebra and Appl. **3**85, 305–334 (2004).
30. K. M. R. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, Ll. Masanes, A. Acin. and F. Verstraete: *The Quantum-Chernoff Bound.* quant-ph/0610027.
31. V. P. Belavkin, M. D'Ariano, and M. Raginski: *Operational distance and fidelity for quantum channels.* J. Math. Phys. **4**6, 062106 (2005).
32. I. Bengtsson: *Geometrical statistics - classical and quantum.* quant-ph/0509017.
33. M. Berry: *Quantal phase factors accompanying adiabatic changes.* Proc.Royal Soc. London A **3**92, 45–57 (1984).
34. D. J. C. Bures: *An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite W^*-algebras.* Trans. Amer. Math. Soc. **1**35, 199 (1969).
35. M.-D. Choi: *Some assorted inequalities for positive linear maps on C^*-algebras.* J. Operator Theory **4**, 271–285 (1980).
36. J. Dittmann, private communication.
37. R. Dubisch: *The Wedderburn structure theorems.* Am. Math. Monthly **5**4, 253–259 (1947).
38. H. A. Dye: *On the geometry of projections in certain operator algebras.* Ann. of Math. **6**1, 73–89 (1955).
39. A. Ericsson: *Geodesics and the best measurement for distinguishing quantum states.* J. Phys. A **3**8, L725–L730 (2005).
40. R. A. Fisher: *Theory of statistical estimation.* Proc. Camb. Philos. Soc. **2**2, 700–725 (1925).
41. V. Fock: *Über die Beziehung zwischen den Integralen der quantenmechanischen Bewegungsgleichungen und der Schrödingerschen Wellengleichung.* Zs.Physik **4**9, 323–338 (1928).
42. G. Fubini: *Sulle metriche definite da una forma Hermitiana.* Istituto Veneto **L**XIII, 2 (1904).
43. V. Giovannetti, S. Lloyd, L. Maccone: *Quantum limits to dynamical evolution.* Phys. Rev. A **6**7, 052109 (2003).
J. Batle, M. Casas, A. Plastino, A. R. Plastino: *On the connection between entanglement and the speed of quantum evolution.* Phys. Rev. A **7**2, 032337 (2005).
A. Boras, M. Casas, A. R. Plastino, A. Plastino: *Entanglement and the lower bound on the speed of quantum evolution.* Phys. Rev. A **7**4, 0222326 (2006).
44. A. M. Gleason: *Measures on the closed subspaces of a Hilbert space.* J. of Math. and Mechanics **6**, 885–893 (1957).
45. J. M. Jauch: *On bras and kets.* In: *Aspects of Quantum Theory*, ed by A. Salam and E. Wigner (Cambridge University Press Cambridge 1972) pp 137–167.
46. R. Jozsa: *Fidelity for mixed quantum states.* J. Mod. Opt. **4**1, 2315–2323 (1994); Ch. A. Fuchs and C. M. Caves: *Mathematical techniques for quantum communication theory.* Open Sys. & Inf. Dyn. **3**, 345–356 (1995); Ch. A. Fuchs: *Distinguishability and Accessible Information in Quantum Theory.* Dissertation, University of New Mexico (1995); H. N. Barnum: *Quantum Information Theory.* Dissertation, University of New Mexico (1998).

47. S. Kakutani: *On equivalence of infinite product measures.* Ann. of Math. **4**9, 214–224 (1948).
48. D. Kretschmann, D. Schlingemann, and R. F. Werner: *The information-disturbance tradeoff and the continuity of Stinespring's representation.* quant-ph/0605009.
49. F. Kubo and T. Ando: *Means of positive linear operators.* Math. Ann. **2**46, 205–224 (1980).
50. G. Lüders: *Über die Zustandsänderung durch den Meßprozeß.* Ann.d.Physik **8**, 322–328 (1951).
51. S. Maeda: *Probability measures on projections in von Neumann algebras.* Rev. Math. Phys. **1**, 235–290 (1990).
52. J. v. Neumann: *On a certain topology for rings of operators.* Ann. Math. **3**7, 111–115 (1936); *On rings of operators.* Ann. Math. **3**7, 116–229 (1936) (with F. J. Murray); *On rings of operators II.* Trans. Amer. Math. Soc. **4**1, 208–248 (1937) (with F. J. Murray); *On rings of operators III.* Ann. Math. **4**1, 94–61 (1940); *On rings of operators IV.* Ann. Math. **4**4, 716–808 (1943) (with F. J. Murray); *On rings of operators: Reduction theory.* Ann. Math. **5**0, 401–485 (1949).
    A general reference is
    R. V. Kadison and J. R. Ringrose: *Fundamentals of the Theory of Operator Algebras - Volume I and II*, 2nd edn, Graduate Studies in Mathematics Volume 15 and 16 (American Mathematical Society 1997).
53. D. Petz: *Monotone metrics on matrix spaces.* Lin.Algebra and Appl. **244** 81-96 (1996); D. Petz and Cs. Sudár: *Geometry of quantum states.* J. Math. Phys. **37**, 2662-2673 (1996).
54. W. Pusz and L. Woronowicz: *Functional calculus for sesquilinear forms and the purification map.* Rep. Math. Phys. **8**, 159–170 (1975).
55. M. Raginski: *A fidelity measure for quantum channels.* Phys. Lett. A **290**, 11-18 (2001).
56. E. Study: *Kürzeste Wege im komplexen Gebiet.* Math. Annalen **L**X, 321–378 (1905).
57. L. Mandelstam and I. Tamm: *The uncertainty relation between energy and time in non-relativistic quantum mechanics.* J. Phys. USSR **9**, 249–254 (1945). Reprinted in: Igor Tamm: *Selected Papers* (Springer Berlin New York 1991) pp 116–123.
58. U. Uhlhorn: *Representation of transformations in quantum mechanics.* Arkiv f. Fysik **2**3, 307–340 (1962).
59. A. Uhlmann: *The "transition probability" in the state space of a *-algebra.* Rep. Math. Phys. **9**, 273–279 (1976).
60. A. Uhlmann: *An energy dispersion estimate.* Phys. Lett. A **1**61 329-331 (1992).
61. A. Uhlmann: *The Metric of Bures and the Geometric Phase.* In: *Quantum Groups and Related Topics*, ed by R. Gielerak et al (Kluwer Academic Publishers 1992) pp 267–264.
62. K. Życzkowski and I. Bengtsson: *An introduction to quantum entanglement: A geometric approach.* quant-ph/0606228.