ALGEBRA AND GEOMETRY

IN THE THEORY OF MIXED STATES

Armin Uhlmann

University of Leipzig
Department of Physics
Leipzig, Germany

1. INTRODUCTION

It is my intention to show some aspects of the very the rich theory of mixed states. A more or less subjective selection of the topics for this purpose was necessary. A certain complement to this selection is reviewed in[1,2,3].

My very aim is an introduction to the idea of extending the concept of *topological phases* a la Berry[4] and Simon[5] to the spaces of mixed states. However, a considerable fraction of the material is 'common knowledge'. This is in particular true with chapters 2, 3, and 5. Also the content of the other parts (except of chapter 9 and the last topic in chapter 8) are essentially known though, perhaps, not so widely. The introduction to Berry's phase within chapter 3 touches only those properties the generalization of which to mixed states is known to me.

The problem of the topological phase is explained as a natural problem of transporting phases along curves (not necessarily and loops) within state spaces. The *naturality* is expressed by the fact that the result depends only on the curve along which the transport takes place, and on the general setting of Quantum Theory.

In this paper I generally consider finite dimensional objects (Hilbert spaces, algebras) mainly in order to minimize technicallities, and (hopefully) allowing easier reading. As a matter of fact almost all what is to be said allows for several extensions to much more general situations. One then, however, enters large areas of complicated and partly unsolved mathematical questions.

2. ELEMENTARY DEFINITIONS

The (pure) states of a quantum system can be described by the vectors of an Hilbert space $\mathcal{H}$. In accordance of what has been said previously we assume

$$\dim \mathcal{H} = n < \infty$$

where n is the dimension of $\mathcal{H}$ as a complex linear space, while its geometrical dimension is of course $2n$. The scalar product of $\mathcal{H}$ will be denoted by $< ., . >$.

An operator of $\mathcal{H}$ is nothing than a complex linear map from $\mathcal{H}$ into $\mathcal{H}$. In the finite dimensional case there is no need for further requirements.

Maps can be composed and this is equivalent to perform the product, AB , of two operators, A and B. Hence, in addition of being a complex linear space, the operators constitute an associative algebra (or ring). It is an algebra over the complex numbers, simply because one

Symmetries in Science V, Edited by B. Gruber *et al.*
Plenum Press, New York, 1991

can multiply an operator by a complex number to get another operator. If A is an operator on $\mathcal{H}$, one denotes by A^* its hermitian conjugate. Because of the antilinear operation $A \mapsto A^*$ which satisfies $(AB)^* = B^*A^*$ one may call the algebra of all operators a **-algebra*. (Indeed, it is the most elementary noncommutative *-algebra one can imagine.) There is a standard notation: The algebra of all bounded operators acting on an Hilbert space is denoted by $\mathcal{B}(\mathcal{H})$. In our case of finite dimension every operator is bounded trivially.

An operator A is called *positive* (more accurate *positive semidefinite*) if $< \psi, A\psi >$ is non-negative real for all $\psi \in \mathcal{H}$. One then writes $A \geq 0$. Clearly, A is positive iff A is hermitian and all its eigenvalues are non-negative. A is called *strictly positive* (or *positive definite*) if A is positive and one of the following conditions is fulfilled: a) A^{-1} exists, b) $< \psi, A\psi >= 0$ implies $\psi = 0$, c) all eigenvalues are positive.

For every positive real number $s > 0$ there is one and only one positive s–th power A^s of a positive operator A. As an hermitian operator is uniquely given by its eigenvectors and its eigenvalues, its definition may be given by

$$A^s\psi = a^s\psi \quad \text{if} \quad A\psi = a\psi$$

$$A^s\psi = 0 \quad \text{if} \quad A\psi = 0$$

Remembering $a^s = s \ln a$, with real $\ln a$, this definition could be extented to all complex numbers. This is well done for strictly positive A but comes otherwise in conflict with the notation of the inverse A^{-1} if $\Re s < 0$. In the latter case one speaks of the *operational defined* s–th power if one uses the definition above.

Evidently

$$\lim A^s \to P \quad \text{for} \quad s \to +0$$

is a projection operator. P is called *support* of A and is denoted by $\text{supp}A$, while I shall call *carrier* of A the subspace $P\mathcal{H}$ whereon A acts non-trivially. (This distinction of support and carrier is not generally in use.)

Now consider an arbitrary operator A . Then AA^* and A^*A are positive. The support of AA^* (of A^*A) is the left support, l–supp, (resp. right support, r–supp,) of A . There is the notation

$$\text{carrier}(A) := \text{carrier}(A^*A) \quad \text{and} \quad \text{range}(A) := \text{carrier}(AA^*)$$

The orthogonal complement of carrier(A) is the *kernel*, ker(A), of A . It will be convenient later on to use the *polar decomposition* of an operator in the (unconventional but equivalent) form

$$A = (AA^*)^{1/2}U$$

where U is a partial isometry (i.e. UU^* and U^*U are projection operators) with

$$UU^* = \text{supp}\, AA^* = \text{l–supp}\, A$$

$$U^*U = \text{supp}\, A^*A = \text{r–supp}\, A$$

3. PURE STATES

As is explained in every textbook, the vectors ψ of $\mathcal{H}$ describe pure states — with the exception of the zero vector. The same state is described by ψ as well by $\lambda\psi$, where λ is any non–zero complex number. This means the pure states are in one to one correspondence to the rays $\{\lambda\psi, \lambda \in \mathcal{C}\}$, or, equivalently, to the 1–dimensionsional subspaces of $\mathcal{H}$. These 1–dimensional subspaces form a smooth manifold called complex projective space, $\mathcal{CP}^{(n-1)}$, or, remembering its inventor, the first Grassmann manifold of $\mathcal{H}$. It is, however, somewhat easier to handle operators than subspaces or rays. Therefore one equally well characterizes a pure state by the projection operator projecting $\mathcal{H}$ onto the 1–dimensional subspace in question. Thus the points of $\mathcal{CP}^{(n-1)}$ may and will be in what follows defined as the operators P of the form

$$P = P_\psi = \frac{|\psi >< \psi|}{< \psi, \psi >} \tag{1}$$

It follows from (1) for every operator A

$$PAP = \frac{<\psi, A\psi>}{<\psi, \psi>} P \tag{2}$$

where the coefficient on the right hand side is the *expectation value* of A if the system is in the state P .

A given 1–dimensional projection (1) does not only characterize a pure state uniquely, it is at the same time an observable asking wether the system is in the state (1) or not. Hence given two pure states, P_1, P_2, one considers the expectation value of P_2 in the state P_1 . This expectation value is the well known *transition probability* :

$$\text{tprob}(P_1, P_2) = \frac{<\psi_1, P_2\psi_1>}{<\psi_1, \psi_1>} = \frac{<\psi_1, \psi_2><\psi_2, \psi_1>}{<\psi_1, \psi_1><\psi_2, \psi_2>} \tag{3}$$

which is a symmetrical quantity in its arguments.

Geometrically, the transition probability governs the relative position of two points of the space of pure states: Two pairs of pure states are unitarily equivalent if and only if they have equal transition probability (see below).

At that point it seems we have lost all relative phases. But this is not so. We lost the *superficial* phases, all others remain. This can be seen by solving the following problem: Given two m-tuples

$$P_1, P_2, \ldots, P_m \quad \text{and} \quad Q_1, Q_2, \ldots, Q_m \tag{4}$$

of pure states. The question is wether there exists a unitary operator U with

$$Q_j = UP_jU^{-1}, \qquad j = 1, 2, \ldots, m \tag{5}$$

One may assume that P_j (resp. Q_j) is given by the normalized vector ψ_j (resp. φ_j) . If

$$<\psi_j, \psi_k> = <\varphi_j, \varphi_k> \quad \text{all } j, k$$

then the map $\psi_j \mapsto \varphi_j$ can be extended isometrically to a map of the linear span of the ψ_j onto the linear span of the φ_j . As both spans are of equal dimensions it is elementary to extend the map to an isometry of $\mathcal{H}$ onto $\mathcal{H}$, i.e. to a unitary map U with $U\psi_k = \varphi_k$. But φ_k and $\epsilon_k\varphi_k$ with $|\epsilon_k| = 1$ define the same pure state. Therefore it is sufficient (and necesary) for the validity of (5) to have

$$<\psi_j, \psi_k> = \epsilon_j\epsilon_k <\varphi_j, \varphi_k> \quad \text{all } j, k$$

with some phase factors ϵ_j . This is true iff for all $2 \leq k \leq m$ the numbers

$$<\psi_{i_1}, \psi_{i_2}><\psi_{i_2}, \psi_{i_3}> \cdots <\psi_{i_k}, \psi_{i_1}> \tag{6}$$

remain unchanged after replacing the ψ–vectors by the φ–vectors. A number (6) is either zero or it is the only non-zero eigenvalue of the partial symmetry

$$P_{i_1} P_{i_2} \cdots P_{i_k} \tag{7}$$

One concludes that (5) is valid if and only if

$$spec\{P_{i_1} P_{i_2} \cdots P_{i_k}\} = spec\{Q_{i_1} Q_{i_2} \cdots Q_{i_k}\} \tag{8}$$

for all subsets $i_1, \ldots, i_k$ of $1, \ldots, m$ with $k \geq 2$. (*spec* denotes "spectrum of".) In particular, an essential relative phase appears in describing the relative positions of at least three pure states.

4. TRANSPORT OF PURE STATES

A partial isometry of rank one, V, can be written as

$$V = |\psi_{\text{out}} >< \psi_{\text{in}}|, \quad VV^* = P_{\text{out}} \quad V^*V = P_{\text{in}} \tag{1}$$

with two normed vectors, and it can be interpreted as annihilating $P_{\text{in}} = |\psi_{\text{in}} >< \psi_{\text{in}}|$ and creating $P_{\text{out}} = |\psi_{\text{out}} >< \psi_{\text{out}}|$. Given the in- and out-states this operation is fixed up to a phase factor because there is no other invariant then tprob$(P_{\text{in}}, P_{\text{out}})$. This slight arbitrariness cannot be removed without introducing a new structural element.

This new structural element is a curve, $\mathbf{c}$, connecting smoothly P_{in} and P_{out}:

$$\mathbf{c} : s \mapsto P_s, \quad 0 \leq s \leq 1, \quad P_0 = P_{\text{in}}, \quad P_1 = P_{\text{out}} \tag{2}$$

By the aid of the construction at the end of chapter 3 it is now possible to fix the phase factor in dependence on $\mathbf{c}$. To this end one takes subdivisions

$$1 > s_1 > s_2 > \ldots > s_m > 0 \tag{2}$$

of the parameter s of the curve and perform

$$V = V(\mathbf{c}) := \lim P_{\text{out}} P_{s_1} P_{s_2} \ldots P_{s_m} P_{\text{in}} \tag{4}$$

where the limiting procedure is taken over finer and finer subdivisions of (3). To calculate V one uses a *lifted* path

$$\mathbf{c}^{\text{lift}} : \quad s \mapsto \psi_s, \quad 0 \leq s \leq 1, \quad \text{with} \quad P_s = |\psi_s >< \psi_s| \tag{5}$$

of unit vectors with which (4) is converted into

$$V = |\psi_1 >< \psi_0| \lim < \psi_1, \psi_{s_1} >< \psi_{s_1}, \psi_{s_2} > \ldots < \psi_{s_m}, \psi_0 > \tag{6}$$

It is convenient to require

$$< \psi_s, \dot{\psi}_s >= 0 \tag{7}$$

before performing (6). (The dot indicates s-differentiation.) A lift (5) fulfilling (7) is called a *parallel lift*. Already in the early days of the adiabatic theorems of Born, Fock, and Oppenheimer condition (7) has been in use, and presently it is known as the condition for the *natural parallel transport* in the unit ball of $\mathcal{H}$ after the work of Berry and Simon. For the time being, however, (7) appears as a technical device, and the result of (6) does *not* depend on it.

From (7) and two times differentiability of (5) one estimates by the help of Taylor's theorem

$$|1 - < \psi_s \psi_t > | \leq (s-t)^2 \, \text{const.} \tag{8}$$

where the constant is independent of s and t. But this estimate implies in (6) that the limit goes to one. Hence assuming (7) one gets

$$V(\mathbf{c}) = |\psi_{\text{out}} >< \psi_{\text{in}}|, \quad \psi_{\text{in}} = \psi_0, \quad \psi_{\text{out}} = \psi_1 \tag{9}$$

Now one may relaxe from condition (7) and obtains

$$V(\mathbf{c}) = |\psi_{\text{out}} >< \psi_{\text{in}}| \exp \int < \mathrm{d}\psi, \psi > \tag{10}$$

where the integral can be taken over an arbitrary lift (5).

For the proof one shows the indepenmdence of expession (10) with respect of regaugings $\psi_s \mapsto \epsilon(s)\psi_s$ with unimodular $\epsilon(s)$. Then it remains to see that (10) reduces to (9) for parallel lifts (7).

$V(\mathbf{c})$ does not depend on the way $\mathbf{c}$ is parametrized. It is a function of the curve as a 1-dimensional submanifold of $\mathcal{CP}^{n-1}$ and of its orientation. Reversing the orientation, and hence interchanging the in- and out-state gives the curve $\mathbf{c}^{-1}$. Obviously

$$V(\mathbf{c}^{-1}) = V(\mathbf{c})^* \tag{11}$$

Let $\mathbf{c}_1$ and $\mathbf{c}_2$ be two curves for which the out-state of second coincides with the in-state of the first. Then $\mathbf{c}_1\mathbf{c}_2$ is a well defined curve and

$$V(\mathbf{c}_1\mathbf{c}_2) = V(\mathbf{c}_1)V(\mathbf{c}_2) \tag{12}$$

V as defined by (4) satisfies a differential equation where the variable is the out-state. Considering the curve

$$\mathbf{c}_s : t \mapsto P_s, \quad 0 < t \leq 1 \tag{13}$$

and the corresponding

$$V_s := V(\mathbf{c}_s) \tag{14}$$

one uses (12) and (4) or (10) to arrive at

$$\dot{V}_s = \dot{P}_s V_s \quad \text{or} \quad \mathrm{d}V = (\mathrm{d}P)V \tag{15}$$

Here the total differential notation in the second equation expresses the irrelevance of the choice of the parameter: $\mathrm{d}V$ restricted to a curve $s \mapsto V_s$ results in $\dot{V}_s \mathrm{d}s$.

It is $P_s V_s = V_s$ according to (13) and (4). Thus, surpressing indices,

$$V^*\dot{V} = V^*\dot{P}V = V^*P\dot{P}PV$$

But differentiating $P = P^2$ results in $\dot{P} = \dot{P}P + P\dot{P}$. Multiplying by P gives $P\dot{P}P = 0$. Finally one gets

$$V_s^*\dot{V}_s = 0 \quad \text{or} \quad V^*\mathrm{d}V = 0 \tag{16}$$

There is an interesting geometrical interpretation of (16). The manifold of rank one isometries can be considered as a fibre bundle with the manifold of rank one projection operators as base manifold. The bundle projection, π, is given by

$$\pi : V \mapsto VV^*$$

A curve $\mathbf{c}$ in the base space can be lifted to a path in the rank one isometries. (16) is the condition for being a parallel lift. (16) implies (15) and vice versa, while (10) gives the explicite construction of parallel lifting.

It should be obvious, last not least, that a closed curve (2), a loop with $P_{\text{in}} = P_{\text{out}}$, gives Berry's construction. It is $V(\mathbf{c}) = \epsilon P_{\text{in}}$ with Berry's phase factor

$$\epsilon = \mathrm{Berry}(\mathbf{c}) = \exp \oint < \mathrm{d}\psi, \psi > \tag{17}$$

while the exponent (divided by i) is the Berry phase.

(17) is the trace of $V(\mathbf{c})$ for a closed loop. Otherwise (10) indicates that this trace depends generally only on the curve, wether this curve is closed or not, and takes values within the unit disk.

5. STATES

Pure states are not the only ones which can be attained by a quantum system with Hilbert space $\mathcal{H}$. On the contrary they form only a thiny (i.e. low dimensional) subset within the state space. The latter will be introduced now.

A *state*, ϱ , is a prescription that associates to every "observable", A , an *expectation value* $\varrho(A)$

$$A \mapsto \varrho(A), \qquad A = A^* \in \mathcal{B}(\mathcal{H}) \tag{1}$$

fulfilling three axioms:

1) Linearity.

$\varrho(A)$ is a real linear functional on the linear space of the hermitian operators.

2) Positivity.

$$\varrho(A) \geq 0 \quad \text{if} \quad A \geq 0 \tag{2}$$

3) Normalization.

$$\varrho(\mathbf{1}) = \mathbf{0} \tag{3}$$

Now every operator A can be uniquely decomposed into an hermitiàn and an antihermitian part, i.e. $A = A_1 + iA_2$ with hermitian A_j. One therefore extends the definition of ϱ by

$$\varrho(A_1 + iA_2) = \varrho(A_1) + i\varrho(A_2)$$

to get a complex linear functional on $\mathcal{B}(\mathcal{H})$ and not only on its hermitian part. One then calls $\varrho(A)$ an *expectation value* even if A is not hermitian — though, physically, there is no measurement apparatus for that if the hermitian and the antihermitian part of A do not commute. With this abstraction one usually replace the first requirement by

1') Linearity.

$\varrho(A)$ is a complex linear functional on $\mathcal{B}(\mathcal{H})$

and so I will do in this paper always.

The normalization condition is good for two things. At first it prevends ϱ to be identical zero. Namely, like for state vectors in the Hilbert space the zero is never a state. Secondly, every state is counted one time. If the normalization condition is neglected one calls an objekt with axioms 1 and 2 a *positive linear functional.* (By the by, there is no logical distinction between "functional" und "function" if one not bounds the former to linearity, what is not usual in Physics.)

The set of all states is called *state space* and will be denoted by

$$\Omega \quad \text{or} \quad \Omega(\mathcal{H})$$

This is a closed subspace of the linear space $\mathcal{B}(\mathcal{H})^*$ of all linear functionals on $\mathcal{B}(\mathcal{H})$, and it gains very natural its structure from this imbedding. It is also useful to consider Ω as a subset of the real linear space of real linear functionals on the hermitian part of $\mathcal{B}(\mathcal{H})$.

It results from the axioms that Ω is a *compact convex* set: Let $p_1, \ldots, p_k$ be a set of reals with

$$p_i \geq 0 \qquad \sum p_j = 1 \tag{4}$$

Then for every choice $\varrho_1, \ldots, \varrho_k$ of states the linear form

$$\varrho = \sum p_j \varrho_j \tag{5}$$

is again a state. This explains the *convex* (or *affine*) structure of Ω. One calls (4), (5) a *convex linear combination* or, in more physical terms, a *Gibbsian mixture* of the states $\varrho_1, \ldots, \varrho_k$ with *weights* or *probabilities* $p_1, \ldots, p_k$. If (5) is valid and $p_j > 0$ then one calls ϱ_j *dominated* by ϱ.

An often more convenient representation of states is given by densitiy operators and this description is isomorphic to that given by linear functionals. Here the starting point is the existence of exactly one linear functional τ satisfying

$$\tau(\mathbf{1}) = 1 \quad \text{and} \quad \tau(AB) = \tau(BA) \tag{6}$$

for all linear operators A, B acting on the Hilbert space, and it is necessarily given by the trace

$$\tau(A) = n^{-1} \operatorname{tr} A, \quad n = \dim \mathcal{H} \tag{7}$$

The uniqueness is a consequence of the statement that the trace of an operator A is zero if and only if A is a linear combination of commutators $BC - CB$. That in turn is equivalent to state that the complexified Lie–algebra of SU(n) is simple and consists of all traceless operators. But also explicit verification is not difficult.

τ as defined by (6) and given by (7) is a state, the unique *tracial state* of our state space: The trace of a positive operator different from the zero operator is positive. While the definition of the trace does not depend on the Hilbert space structur, the property to be a positive functional does. There would be no state at all if the *–operation is given by an indefinite, non–degenerate hermitian scalar product.

Next, every $B \in \mathcal{B}$ defines one and only one linear functional

$$A \mapsto \mathrm{tr}(AB) \tag{8}$$

and because $\mathcal{B}$ and $\mathcal{B}^*$ have equal dimensions, every linear functional can be gained by a suitable B. Moreover, tr(AB) will be positive for all positive A iff B itself is positive. Hence to every state ϱ there is just one operator, ϱ_{op}, with

$$A \mapsto \varrho(A) = \mathrm{tr}\,\varrho_{\mathrm{op}} A \tag{9}$$

and

$$\varrho_{\mathrm{op}} \geq 0 \qquad \mathrm{tr}\varrho_{\mathrm{op}} = 1 \tag{10}$$

Every such ϱ_{op} is called *density operator*, and there is a complete correspondence between states and density operators. Having this in mind no explicit notational distiction will be made between state and density operator, and the index *op* will be neglected usually. As the eigenvalues of a density operator are not negative and sum up to one they constitute a probability vector.

The most simple density operator is a 1–dimensional projection operator P characterizing a pure state.

6. SOME CONVEX GEOMETRY

Above I considered states as individual objects which form, as a compact comvex set, the state space Ω. One may reverse this setting and look at a state as a point of the state space, and one may ask to what degree its position in Ω determines its properties. In other words, one wonders how much is encoded in the inner geometry of Ω, forgetting for a moment the nature of its points as linear, normed, and positive functionals or density operators. The symmetry group Aut(Ω) for this task consists of the *affine automorphisms* ϕ , i.e. linear mappings from Ω onto Ω fulfilling

$$\phi(\sum p_j\, \varrho_j) = \sum p_j\, \phi(\varrho_j) \tag{1}$$

for all convex linear combinations. Every unitary operator, U, gives rise to an element of Aut(Ω) by

$$\varrho \to \varrho^U \quad \text{with} \quad \varrho^{\mathrm{U}}(\mathrm{A}) = \varrho(\mathrm{UAU}^{-1}) = (\mathrm{U}^{-1}\varrho\mathrm{U})(\mathrm{A}) \tag{2}$$

These transformations form a subgroup $\mathrm{Aut_U}(\Omega)$ of Aut(Ω). In fact, $\mathrm{Aut_U}$ is a normal subgroup of Aut of index two as will be seen soon.

According to general terminology a state ϱ is called *extremal* iff for every of its decompositions (5–4), (5–5) one necessarily has $\varrho = \varrho_k$ whenever $p_k > 0$. The set of all extremal elements of a convex set is called its *extremal boundary*.

The pure states are exactly the extremal points of Ω.

Indeed, every density operator allows for a decomposition

$$\varrho = \sum \lambda_j P_j \tag{3}$$

as a convex set of 1-dimensional projection operators. Well known examples of (3) provide spectral decompositions whith mutually orthogonal projectors where the λ_j are the corresponding eigenvalues of ϱ. The decompositions (3) show Ω as the convex span of the pure states. But

if P is both, pure and extremal, then so does P^U with unitary U because both properties are stable against the subgroup Aut_U of unitary transformations. But the unitary transformations act transitve on the set of pure states. Hence every pure state has to be extremal.

An extremal point of a convex set is a particular case of what is called a *face*. A convex subset Ω_0 of Ω is called a *face* of Ω if $\varrho \in \Omega$ implies $\tilde{\varrho} \in \Omega$ whenever $\tilde{\varrho}$ is dominated by ϱ.

One knows the following: There is a one-to-one correspondence between faces and subspaces of $\mathcal{H}$. For every face Ω_0 there is a subspace $\mathcal{H} \subset \mathcal{H}_1$ with

$$\Omega_0 = \{\varrho : \mathrm{carrier}\varrho = \mathcal{H}_0\} \tag{4}$$

The infinite-dimensional case is much mor complicate[1]4

Let now be $P_1, P_2 \in \Omega$ two different pure states and let a denote their transition probability:

$$P_1P_2P_1 = aP_1 \quad \text{with} \quad a = \mathrm{tprob}(P1,P_2) \tag{5}$$

Considering a Gibbsian mixture (convex sum)

$$\varrho = \mu_1P_1 + \mu_2P_2 \tag{6}$$

one gets by taking its square and then the trace

$$\lambda_1^2 + \lambda_2^2 = \mu_1^2 + \mu_2^2 + 2a\mu_1\mu_2 \tag{7}$$

where λ_1, λ_2 denote the eigenvalues of ϱ. As the latter sum up to one as is the case with the coeficients μ_j equation (7) is equivalent to

$$\lambda_1\lambda_2 = \mu_1\mu_2(1 - a) \tag{8}$$

From these elementary observations follow some important conclusions. At first, if ϱ is given by (6), one may ask for all possible decompoitons of ϱ as a comvex sum of two pure states. Then the product of the coefficients $\mu_1\mu_2$ takes its minimum if and only if $a = 0$, i.e. iff the coefficients of the Gibbsian mixture coincide with the eigenvalues. Hence the eigenvalues of ϱ are determined by its position in Ω.

Because of this, however, (7) or (8) shows that $\mathrm{trans}(P_1, P_2)$ is uniquely fixed by the relative position of the two pure states within Ω. Thus what is physically relevant of the Hilbert space structure of $\mathcal{H}$ is encoded in the geometry of Ω.

In particular it should be possible to measure transition probabilities of pure states by performing mixtures. (Let $\mu_1 = \mu_2$ then the transition probability equals $(\lambda_1 - \lambda_2)^2$.)

Furthermore, the transition probability between two pure states remains unchanged under the action of an element of $\mathrm{Aut}(\Omega)$. Applying now a well known theorem of Wigner one concludes that every element of $\mathrm{Aut}(\Omega)$ is induced either by a unitary or an antiunitary transformation of $\mathcal{H}$ and the affine automorphisms of the state space reflect the essentials of the Hilbert space transformations.

It is interesting to note that an explicite characterization of the semigroup of affine endomorphisms is not known. Physically it is meaningful to consider restricted classes of them (in particular the duals of comletely positive unital transformations of $\mathcal{B}$). However, a remarkable exception can be found at the end of chapter 8.

Now one may pose the questions: How to extend the concept of transition probability to general states, and how to extend equations (7, (8) to them. I shall explain the answer to the first question later, and give a partial answer to the second.

If ordered in decreasing order, (7) or (8) gives

$$\lambda_1 \geq \mu_1 \geq \mu_2 \geq \lambda_2$$

and $\lambda_1 = \lambda_2$ iff $a = 0$, i.e. iff $P_1P_2 = 0$. Considering the general case

$$\varrho = \sum \mu_j \varrho_j, \qquad \mu_1 \geq \mu_2 \geq \ldots \geq 0 \tag{9}$$

and denoting by $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ the eigenvalues of ϱ in decreasing order, one finds[15,2]

$$\text{for all k}: \quad \sum_{j=1}^{k} \lambda_j \geq \sum_{j=1}^{k} \mu_j \tag{10}$$

and equality holds for all k iff (9) is an orthogonal decomposition with the eigenvalues of ϱ as coefficients. (If the λ- and the μ–vectors are of different length, one adds zeros to compensate this.)

For the proof of (10) one proceeds in the following steps. At first a set of inequalities of the form (10) defines a partial order within the probability vectors appearing in decompositions (9) of a fixed density operator. To every such vector there is at least one maximal element with respect to the partial order. Assumings such a maximal element is not an orthogonal decomposition it follows at least for two pure states, P_i, P_k with $i < k$ a positive transition probability. But this is impossible: Replacing $\mu_i P_i + \mu_k P_k$ by their orthogonal decomposition shows (after some calculations) that a not orthogonal decomposition cannot go with a maximal element of the partial order.

It is more complicated to derive handy relations similar to (7) or (8) for a general decompostion (9). To do so one has to characterize the relative positions of the pure states involved. This demands to know all the numbers

$$\mathrm{tr} P_{i_1} \ldots P_{i_k}$$

constructed from subsets of the $P_1, P_2, \ldots$. This is clearly possible in an implicite form by calculating the trace of the powers of the density operator ϱ as given by (9).

7. Subsystems. Reductions

As before let $\mathcal{B}(\mathcal{H})$ be the algebra of operators acting on $\mathcal{H}$. Within sektion 3 the pure states have been described by normed vectors $\psi \in \mathcal{H}$ or by 1-dimensional projections $P = |\psi><\psi|$. This usage of the word *pure* is bound to an assumption: The algebra $\mathcal{B}(\mathcal{H})$ should be the algebra of observables. To be more accurate, the algebra of observables should contain to every pair of different 1-dimensional projections (or rays) at least one hermitian operator which distinguish them by its expectation value. However, a *-subalgebra of $\mathcal{B}(\mathcal{H})$ with this properties coincides with $\mathcal{B}(\mathcal{H})$.

The conceptual perhaps cleanest way to distinguish a ***subsystem*** is by choosing a *-subalgebra $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$ i.e. one which is generated by its hermitian operators.

Heuristically this can be understood as following. There may be distinct states (pure or not) which, if restricted to a given subsystem, cannot be distinguished any more. Indeed, if this would not be true, every state of the subsystem could be uniquely extended to the bigger system, every state could already be distinguished from all others by the subsystem. But then there is no possibility to explain in what system and subsystem are different, and both had to be considered as identical.

Hence, looking from a genuine subsystem, some states can be no longer distinguished though they are originally as seen from the system. Therefore, an operator with different expectation values for a pair of such states cannot be an observable of the subsystem. Consequently, a subsystem can be characterized by a subset of observables of the full system. This is the key idea in the concept of subsystems.

It is, however, not completely deducable from this considerations that observables of a subsystem can be given by the hermitian elements of a *-subalgebra $\mathcal{A}$ of the system's observables. But here this assumption will be accepted and, moreover, $\mathcal{A}$ should contain the identity map of $\mathcal{H}$, i.e. it should be a unital subalgebra.

Even then the notation of a subsystem is a rather general one ranging from commuting subalgebras describing compatible measurements to *proper* subsystems for which $\mathcal{A}$ is a *factor* for which the centre contains the multiples of the unit element only. A theorem of Wedderburn allows to ennumerate all unital *-subalgebras $\mathcal{A}$ of $\mathcal{B}(\mathcal{H})$. In such an algebra there exists an orthogonal and central set $Q_1, \ldots Q_m$ of projection operators ($Q_j = Q_j^2 = Q_j^*$) with

$$1 = \sum Q_j, \quad Q_j Q_K = 0 \quad \text{if} \quad j \neq k \tag{1}$$

which define subalgebras

$$\mathcal{A}_j = \{A \in \mathcal{A} : Q_j A = A, \quad Q_k A = 0 \text{ otherwise}\} \tag{2}$$

with

$$\mathcal{A} = \mathcal{A}_1 \oplus \mathcal{A}_2 \oplus \ldots \oplus \mathcal{A}_m \tag{3}$$

and with centres consisting of the multiples of Q_j only: They are factors. Let

$$\mathcal{H}_j = Q_j \mathcal{H} \tag{4}$$

denote the carrier of Q_j. (1) implies

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \ldots \oplus \mathcal{H}_m \tag{5}$$

and $\mathcal{A}_j$ can be regardes as a *-subalgebra of $\mathcal{B}(\mathcal{H}_j)$ because the carrier of every element of $\mathcal{A}_j$ is contained in $\mathcal{H}_j$ according to (2). Be!ng a factor of $\mathcal{B}(\mathcal{H}_j)$ there exists a direct product decomposition

$$\mathcal{H}_j = \mathcal{H}_j^{(x)} \otimes \mathcal{H}_j^{(y)} \tag{6}$$

such that

$$\mathcal{A}_j = \mathcal{B}(\mathcal{H}_j^{(x)}) \otimes \mathbf{1}_j^{(y)} \tag{7}$$

In this relation $\mathbf{1}_j^{(y)}$ denotes the identity map of $\mathcal{B}(\mathcal{H}_j^{(y)})$, and the algebra (7) consists just of all operators

$$A_j^{(x)} \otimes \mathbf{1}_j^{(y)} \tag{8}$$

where the first factor runs through all elements of $\mathcal{B}(\mathcal{H}_j^{(x)})$.

It is fairly clear, how a *state* of $\mathcal{A}$ is to define. One literally repeats the definition of section 5: A state of $\mathcal{A}$ is a real linear functional on the real linear space of hermitian operators of $\mathcal{A}$ which takes non-negative values for positive operators, and is equal to one for the unity of $\mathcal{A}$. Of course one can and does extend these functionals uniquely to a complex linear one defined on all $\mathcal{A}$.

The states of $\mathcal{A}$ form a convex set (an affine space) $\Omega_{\mathcal{A}}$, the *state space* of $\mathcal{A}$ or, equivalently, of the *subsystem* defined by $\mathcal{A}$. Now one *defines* the pure state of $\mathcal{A}$ to be the extremal states of its state space.

If ϱ is a state of $\mathcal{B}(\mathcal{H})$ its restriction

$$\omega = \varrho_{\mathcal{A}} : \omega(A) = \varrho(A) \quad \text{all} \quad A \in \mathcal{A} \tag{9}$$

is a state of $\mathcal{A}$. This state, as defined by (9), is called *the restriction* of ϱ on $\mathcal{A}$. If ω is the reduction of ϱ , ϱ is called an *extension* of ω.

One knows by general extension theorems (of Hahn Banach type) for positive functionals that every state of $\mathcal{A}$ is a reduced state and can be gained as the reduction of a state of $\mathcal{B}(\mathcal{H})$. Thus, the reduction of states is a map from Ω, the state space of $\mathcal{B}(\mathcal{H})$, onto $\Omega_{\mathcal{A}}$, the state space of $\mathcal{A}$.

Generally, the reducing map does not directly respect the concept of purity. If ω is pure, its extension ϱ may not be pure. (This can occur also classically.) If ϱ is pure its reduction may not be pure. This is an important fact showing the influence of the superpostion principle of Quantum Physics. Even more, every state can be considered as the reduction of a pure state. This is contrary to all what is known from Classical Statistical Physics.

The next issue is to ask for the concept of density operators for staes of $\mathcal{A}$. It should not depend on the way, $\mathcal{A}$ is embedded within $\mathcal{B}(\mathcal{H})$ or within another algebra. To this end one needs a particular trace, i.e. a linear functional of $\mathcal{A}$ satisfying

$$\mathrm{tr}_{\mathcal{A}}(AB) = \mathrm{tr}_{\mathcal{A}}(BA) \ .$$

and taking value one for all minimal projection operators of $\mathcal{A}$. By the last condition this trace is uniquely defined and called *canonical trace* of $\mathcal{A}$. Using the general form of $\mathcal{A}$ explained by (3) and (7) its existence can be shown by construction.

Given now ω, a state of $\mathcal{A}$, there is an operator $\omega = \omega_{op}$ in $\mathcal{A}$, called *density operator*, such that

$$A \mapsto \omega(A) = \mathrm{tr}_{\mathcal{A}}(A) \quad \text{for all} \quad A \in \mathcal{A} \tag{10}$$

In case the state ω appears as the reduction of a state of a larger system, the density operator ω is called the *reduced* density operator.

8. EXTENSIONS AND PURIFICATIONS

It is evident how to reverse the procedure explained in chapter 7, and how to ask for extensions of the given system, i.e. for embeddings of $\mathcal{B}(\mathcal{H})$ in a larger observable algebra, and extending the states of $\Omega(\mathcal{H})$ appropriately. Though elementary, it is a key ingredient that an observable (operator) of a system remains an observable in every larger system.

In extending $\mathcal{H}$, $\mathcal{B}$, Ω I restrict myself to cases where $\mathcal{B}(\mathcal{H})$ as a subalgebra of the larger system becomes a factor. The general case is obtained by performing direct sums of such extensions.

Thus let

$$\mathcal{H}^{\mathrm{ext}} = \mathcal{H} \otimes \mathcal{H}' \tag{1}$$

describe a system composed of our original one, $\mathcal{H}$, and a supplementary one, $\mathcal{H}'$. Then $\mathcal{B}(\mathcal{H})$ is embedded in $\mathcal{B}(\mathcal{H}^{\mathrm{ext}})$ by

$$A \in \mathcal{B}(\mathcal{H}) \mapsto A \otimes \mathbf{1}' \in \mathcal{B}(\mathcal{H}^{\mathrm{ext}}) \tag{2}$$

where $\mathbf{1}'$ is the unit element of $\mathcal{B}(\mathcal{H}')$. In the same spirit the map

$$A' \in \mathcal{B}(\mathcal{H}') \mapsto \mathbf{1} \otimes A' \in \mathcal{B}(\mathcal{H}^{\mathrm{ext}}) \tag{3}$$

is a unital isomorphism of $\mathcal{B}(\mathcal{H}')$ into $\mathcal{B}(\mathcal{H}^{\mathrm{ext}})$. The decomposition (1) induces in this way uniquely these two embeddings (2) and (3), and as long one sticks to (1) it is possible and convenient to consider $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{H}')$ as subalgebras of $\mathcal{B}(\mathcal{H}^{\mathrm{ext}})$. This convention will be used in the following. Then $\mathcal{B}(\mathcal{H}')$ is the commutant of $\mathcal{B}(\mathcal{H})$, i.e. the set of operators of $\mathcal{B}(\mathcal{H}^{\mathrm{ext}})$ commuting with every element of $\mathcal{B}(\mathcal{H})$, and $\mathcal{B}(\mathcal{H})$ is the commutant of $\mathcal{B}(\mathcal{H}')$.

An element $\tilde{\varrho}$ of $\Omega^{\mathrm{ext}} = \Omega(\mathcal{H}^{\mathrm{ext}})$ is called an *extension* of ϱ iff $\tilde{\varrho}$ gives ϱ if restricted to $\mathcal{B}(\mathcal{H})$.

Given $\varrho \in \Omega(\mathcal{H})$ the set of all extensions of ϱ to Ω^{ext} will be denoted by $\{\varrho\}^{\mathrm{ext}}$.

$\{\varrho\}^{\mathrm{ext}}$ is a convex compact subset of Ω^{ext}. According to Minkowski it is therefore the convex hull of its extremal points. However, no effective procedure is known to ennumerate these points. This is a common difficulty arising mathematically from the positivity condition, and from the point of view of Physics from the superposition principle: In a composed quantum system there are much more states then obviously visible from the subsystems. Here the difficult question is the following. Given density operators ϱ_j and ϱ'_j from $\Omega(\mathcal{H})$ and $\Omega(\mathcal{H}')$. For what coefficients α_{jk} is the operator

$$\sum \alpha_{jk} \varrho_j \otimes \varrho_k$$

a positive one?

Because of these circumstances I will pose a not so ambitious question and ask for the set of pure states within $\{\varrho\}^{\mathrm{ext}}$. Clearly, every such pure state is automatically extremal in $\{\varrho\}^{\mathrm{ext}}$ while the reverse assertion is wrong.

Let $\psi_1, \psi_2, \ldots, \psi_n$ and $\psi'_1, \ldots, \psi'_{n'}$ be a complete orthonormal system of $\mathcal{H}$ and $\mathcal{H}'$ respectively. Every normed vector $\varphi \in \mathcal{H}\mathrm{ext}$ can be represented as

$$\varphi = \sum_{j=1}^{n} \sum_{k=1}^{n'} c_{jk} \psi_j \otimes \psi'_k, \quad \sum |c_j k| = 1 \tag{4}$$

To reduce $|\varphi><\varphi|$ to $\mathcal{B}(\mathcal{H})$ one has to perform with $A \in \mathcal{B}(\mathcal{H})$

$$<\varphi, A\varphi> = \sum \bar{c}_{jk} c_{ir} <\psi_j \otimes \psi'_k, (A\psi_i) \otimes \psi'_r>$$

to obtain

$$<\varphi, A\varphi> = \sum_k \bar{c}_{jk} c_{ik} <\psi_j, A\psi_i>, \tag{5}$$

This should be equivalent to

$$\varrho(A) = \operatorname{tr} \varrho A = \sum_{j,i} <\psi_i, \varrho\psi_j><\psi_j, A\psi_i> \tag{6}$$

If the linear operator C from $\mathcal{H}'$ into $\mathcal{H}$ is now defined by

$$<\psi_i, C\psi'_k> = c_{ik} \tag{7}$$

then one gets

$$\varrho = C C^* \tag{8}$$

The result is: There are pure states in $\{\varrho\}^{\text{ext}}$ if and only if

$$\operatorname{rank} \varrho \leq n' = \dim \mathcal{H} \tag{9}$$

Therefore *every* state $\varrho \in \Omega(\mathcal{H})$ has an extension in Ω^{ext} which is pure if $n \leq n'$. Every pure extension of $\varrho \in \Omega(\mathcal{H})$ is called a *purification* of ϱ.

Looking at (7) one immediately sees: If there is one purification then there are many. With an operator C of (7) also every other one given by

$$C^{\text{new}} = CU, \quad U \in \mathcal{B}(\mathcal{H}') \tag{10}$$

with a unitary U will give the same ϱ. Thus for the purification ambiguity a *gauging* with the group of all unitarities of $\mathcal{B}(\mathcal{H}')$ is responsible. It is possible (see later on) to construct a gauge field theory to handle this situation. While one has

$$U\{\varrho\}^{\text{ext}}U^* = \{\varrho\}^{\text{ext}} \tag{11}$$

with unitary $U \in \mathcal{H}'$, it remains to remark the following. Let $\tilde{\varrho} \in \{\varrho\}^{\text{ext}}$ and let $\in \mathcal{B}(\mathcal{H}')$ be a partial isometry the right support of which contains the carrier of $\tilde{\varrho}$. Then $U\tilde{\varrho}U^*$ is again in $\{\varrho\}^{\text{ext}}$. However, the partial isometries are not equipped with the structure of a group but only with that of a groupoid.

The next aim is to introduce the notation of transition probability for general states. Let ϱ_1, ϱ_2 denote two states of $\Omega = \Omega(\mathcal{H})$. In a suitable extended system (1) one can consider purifications φ_1, φ_2 of them which give rise to operators C_1, C_2 as indicated by (4) and (7). A similar calculation yields

$$<\varphi_1, \varphi_2> = \operatorname{tr} C_1^* C_2 \tag{12}$$

This number depends on the choice of the purifications. Both operators, C_1, C_2, can be gauged independently according to (9). This changes the number (12). This means that one cannot control completely by a system what is going on in a larger system. In particular, the absolute square of (12) gives the a priori probability that ϱ_1 goes over to ϱ_2 if probed in the extended system by the *obsewrvable* P_2 wether the state which is originally P_1 is equal to the *state* P_2.

Though (12) and $\operatorname{tprob}(P_1, P_2)$ is not determined by ϱ_1 and ϱ_2, one can correctly speak of the supremum

$$\operatorname{tprob}(\varrho_1, \varrho_2) = \sup \operatorname{tprob}(P_1, P_2) \tag{13}$$

where P_1, P_2 run through all possible pairs of simultaneous purifications of ϱ_1 and ϱ_2. This supremum is *by definition* called the *transition probability* of ϱ_1, ϱ_2, see[16].

The calculation of (13) can be performed as following. There are polar decompositions

$$C_j = \varrho_j^{\frac{1}{2}} U_j, \quad j = 1,2 \tag{14}$$

where uniquness is achieved with maps U_j from $\mathcal{H}'$ onto carrier(ϱ_j) in $\mathcal{H}$.

Now

$$\operatorname{tr} C_1^* C_2 = \operatorname{tr} \varrho_1^{\frac{1}{2}} \varrho_2^{\frac{1}{2}} U_2 U_1^* = \operatorname{tr} |\varrho_1^{\frac{1}{2}} \varrho_2^{\frac{1}{2}}| U U_2 U_1^* \tag{15}$$

In this relation the polar decomposition

$$\varrho_1^{\frac{1}{2}} \varrho_2^{\frac{1}{2}} = |\varrho_1^{\frac{1}{2}} \varrho_2^{\frac{1}{2}}| U, \quad |\varrho_1^{\frac{1}{2}} \varrho_2^{\frac{1}{2}}| = (\varrho_1^{\frac{1}{2}} \varrho_2 \varrho_1^{\frac{1}{2}})^{\frac{1}{2}} \tag{16}$$

has been used. However, U, U_1, U_2 have operator norm one, and so the operator norm of their product cannot exceed one either. This has the consequence that the supremum will be reached if

$$U U_2 U_1^* \geq \operatorname{supp} |\varrho_1^{\frac{1}{2}} \varrho_2^{\frac{1}{2}}| \tag{17}$$

Hence

$$\operatorname{tprob}(\varrho_1 \varrho_2) = (\operatorname{tr}(\varrho_1^{\frac{1}{2}} \varrho_2 \varrho_1^{\frac{1}{2}})^{\frac{1}{2}})^2 \tag{18}$$

and there are indeed always purifications P_1, P_2 of ϱ_1, ϱ_2 with which the supremum of (13) is attained.

I call *parallel* purifications those pairs P_1, P_2 which purify simultaneously ϱ_2, ϱ_2, and for which (13) becomes an equality. The same notation (i.e. parallel purification) applies for pairs of normed vectors φ_1, φ_2, of the extended system if there pure states purify simultaneously, and if in addition there scalar product is real and not negative.

(17) can also expressed by

$$C_1^* C_2 = C_2^* C_1 \geq 0 \tag{19}$$

(18) indicates the possibility to calculate the transition probability in the original system. Without proofs I describe some properties arising in this connection[17-22].

Let ν be a linear function on $\mathcal{B}(\mathcal{H})$ satisfying

$$|\nu(A^* B)|^2 \leq \varrho_1(A^* A)\, \varrho_2(B^* B) \tag{20}$$

The typical example is

$$\nu(A) = \operatorname{tr} A C_2 C_1^* \tag{21}$$

With C_j given by (14). It can be shown that always

$$|\nu(\mathbf{1})|^2 \leq \operatorname{tprob}(\varrho_1, \varrho_2) \tag{22}$$

and, using an ansatz (21), that equality can be obtained in (22). If ν satisfies (20), if the carriers of φ_1 and φ_2 coincides, and

$$\nu(\mathbf{1}) > 0 \quad \text{with} \quad |\nu(\mathbf{1})|^2 = \operatorname{tprob}(\varrho_1, \varrho_2) \tag{23}$$

is valid then ν is uniquely determined. If the carrier condition is not fulfilled the uniqueness question is more complicated.

It is also possible to get the transition probability by an infimum. Surprising enough one gets

$$\operatorname{tprob}(\varrho_1, \varrho_2) = \inf \varrho_1(A)\, \varrho_2(A^{-1}) \tag{24}$$

where A runs through all the positive operators of $\mathcal{B}(\mathcal{H})$.

The pecular role of the transition probability is further underlined by the statement: Let ϕ be an affine endomorphism of Ω. Then

$$\operatorname{tprob}(\phi \circ \varrho_1, \phi \circ \varrho_2) \geq \operatorname{tprob}(\varrho_1, \varrho_2) \tag{25}$$

9. PARALLEL TRANSPORT

To treat the problem of chapter 4 in the case of general states[24] is now more or less straightforward: Given a curve of states (respectively density operators)

$$s \mapsto \varrho_s \qquad \text{with} \quad 0 \leq s \leq 1 \tag{1}$$

defined on $\mathcal{H}$, one looks for a purified lift

$$s \mapsto P_s = \frac{|\varphi_s><\varphi_s|}{<\varphi_s, \varphi_s>}, \quad 0 \leq s \leq 1, \tag{2}$$

in an extended system with Hilbert space $\mathcal{H}^{\text{ext}}$ as defined in chapter 8.

To restrict as much as possible the arbitrariness of such a lift the simple idea is to require parallelity for 'infinitely neighboured' purifications. The Hilbert space metric

$$\| \varphi_s - \varphi_t \| = \sqrt{<\varphi_s - \varphi_t, \varphi_s - \varphi_t>} \tag{3}$$

will be used to make this attempt correct.

Assuming at first the normed vectors φ_s and φ_t to be purifications of ϱ_s and ϱ_t then, according to the defining equation (8-13), it is

$$\| \varphi_s - \varphi_t \|^2 \leq 2 - 2\text{tprob}(\varrho_s, \varrho_t) \tag{4}$$

and the equality sign holds for parallel lifts[23,24]. The latter case gives the metric of Bures[25] in the state space one is starting from:

$$\| \varrho_s - \varrho_t \| = \sqrt{2 - 2\text{tprob}(\varrho_s, \varrho_t)} \tag{5}$$

Returning to (4) and concerned with 'infinitely neighboured' states it is tempting to use

$$\lim \epsilon^{-1} \| \varphi_{s+\epsilon} - \varphi_s \| = \sqrt{<\dot{\varphi}_s, \dot{\varphi}_s>} \tag{6}$$

and to require a minimal right hand side in order that (2) is a parallel lifting of (1). For normed purifying vectors we then are done. It is, however, convenient to relaxe from norming the purifying vectors. In this slightly more general case the following seems to be very natural:

A lift (2) is called a *parallel* purification of (1) if the expression

$$\frac{<\dot{\varphi}, \dot{\varphi}>}{<\varphi, \varphi>} \tag{7}$$

attains for every value of the parameter s its minimum with respect to all purifications of a given curve of states (1).

For the transport of frames of degenerated eigenstates a similar proposel, though not worked out at all, has been made already by Fock in the appendix of[26]. As a first extension of Berry's idea it is discussed, however from quite another side, by Wilczek and Zee[27]. Their theory concerns the case of a path of density operators (1) which is proportional to a path of projection operators of fixed rank.

In (7) a fixed parametrization of the original curve (1) is assumed. The essentials are parameter independent of course. This can be made obvious by considering the line element

$$\sqrt{\frac{<\dot{\varphi}, \dot{\varphi}>}{<\varphi, \varphi>}}\,\mathrm{d}s \tag{8}$$

There is, by the by, a slight difference between the Bures metric (5) and the metric given by the line element (8) because of the occurence of the denominator in (7) and (8).

The next aim is to get conditions for the curve

$$s \mapsto \varphi_s \qquad \varphi_s \in \mathcal{H}^{\text{ext}} \tag{9}$$

which are necessary for attaining the minimum of the expression (7) while (2) is a purification of (1).

Using the notations of chapter 8, in particular the decomposition (1) and its consequences, the following is obvious: For every hermitian operator B' of $\mathcal{B}(\mathcal{H}')$ one considers $\mathbf{1} \otimes B' = B$. Then with (9) the curve

$$s \mapsto U(s)\varphi_s \quad \text{with} \quad U(s) = \exp(isB) \tag{10}$$

will not give a value smaller than (9) if inserted into (7). The resulting inequality reads

$$0 \leq < B\varphi, B\varphi > + i[< \dot{\varphi}, B\varphi > - < \varphi, B\dot{\varphi} >] \tag{11}$$

If this ineqality is valid for all allowed B it implies

$$< \dot{\varphi}, B\varphi > = < \varphi, B\dot{\varphi} > \quad \text{for all} \quad B = I \otimes B' \tag{12}$$

Since the condition (12) is linear in B, and hence in B', and since the linear span of the hermitian operators contains all operators, (12) holds for all operators of the form $\mathbf{1} \otimes B' = B$. Remarkable enough every of the relations (12) is a transport condition like (4-7) which is only slightly masked by the arbitrariness of the norms. Namly if for a particular B one has $< \varphi, B\varphi > = \text{const.}$ then (12) sharpens to $< \dot{\varphi}, B\varphi > = 0$.

The necessary condition (12) does not reflect what is going on with the norm of the vectors. This is natural since it is the same with (2). If now (9) fulfills (12) and is normed, one can probe (7) with a new curve

$$s \mapsto \lambda_s \varphi_s, \qquad \lambda_s > 0 \tag{13}$$

In the expression (7) this will act as a substitution

$$\frac{< \dot{\varphi}, \dot{\varphi} >}{< \varphi, \varphi >} \quad \rightarrow \quad \frac{< \dot{\varphi}, \dot{\varphi} >}{< \varphi, \varphi >} + (\frac{\dot{\lambda}}{\lambda})^2 \tag{14}$$

The net result is: The necessary condition for (9) to minimize (7) such that (2) becomes a parallel lift of (1) is the following: The curve (9) fulfills (12) and the norm of its vectors is constant, i.e. is independent of the parameter.

But is this necessary condition also sufficient? To decide this question one needs more general deformations of a curve (9) which minimizes (7). One needs not only curves (10) of unitaries but smooth curves of arbitrary partial isometrics

$$U_s = \mathbf{1} \otimes U'_s, \quad \text{with} \quad U_s \in \mathcal{B}(\mathcal{H}') \tag{15}$$

They give rise to deformations

$$\varphi'_s = U_s \varphi_s \quad \text{with} \quad U_s^* \varphi'_s = \varphi_s \tag{16}$$

A certain role will be played by the curves of projection operators

$$Q_s = U_s^* U_s \quad \text{and} \quad Q'_s = U_s U_s^* \tag{17}$$

Its first role is in fixing U_s uniquely by one of the relations (16). Within all partial isometries fulfilling one (and therefore the other) relation (16) there is one for which the carriers Q_s are as small as possible. This choice will be made further on. Now

$$< \dot{\varphi}', \dot{\varphi}' > = < \dot{U}\varphi, \dot{U}\varphi > + < U\dot{\varphi}, U\dot{\varphi} > + < \dot{U}\varphi, U\dot{\varphi} > + < U\dot{\varphi}, \dot{U}\varphi > \tag{18}$$

The last two terms of (18) can be rewritten as following. By (12)

$$< U\dot{\varphi}, \dot{U}\varphi >=< \dot{\varphi}, U^*\dot{U}\varphi >=< \varphi, U^*\dot{U}\dot{\varphi} > \tag{19}$$

Hence

$$< \dot{U}\varphi, U\dot{\varphi} > + < U\dot{\varphi}, \dot{U}\varphi >=< \varphi, \dot{U}^*U\dot{\varphi} > + < \varphi, U^*\dot{U}\dot{\varphi} >=< \varphi, \dot{Q}\dot{\varphi} > \tag{20}$$

Inserting into (18) yields

$$< \dot{\varphi}', \dot{\varphi}' >=< \dot{U}\varphi, \dot{U}\varphi > + < U\dot{\varphi}, U\dot{\varphi} > + < \varphi, \dot{Q}\varphi > \tag{21}$$

Examination of the last two terms yields

$$< U\dot{\varphi}, U\dot{\varphi} > + < \varphi, \dot{Q}\varphi >=< \dot{\varphi}, Q\dot{\varphi} > + < \varphi, \dot{Q}\varphi >=< \dot{\varphi}, \dot{\varphi} > \tag{22}$$

where the last equality sign follows from differentiating $\varphi = Q\varphi$. Combining (21) and (22) provides

$$< \dot{\varphi}', \dot{\varphi}' >=< \dot{U}\varphi, \dot{U}\varphi > + < \varphi, \varphi > \tag{23}$$

Thus the condition (12) is not only necessary but also sufficient under the *additional assumption* that the projection operators (17) depend ***smoothly*** on its parameter s. In the case at hand, where the Hilbert spaces are all finite dimensional, this additional assumption is fulfilled iff (1) is smooth and of constant rank[11].

There is a further consequence if both, φ_s and φ'_s, are minimizing (7) because (23) then supplies $\dot{U}\varphi = 0$. The last is equivalent with $\dot{U}U^*\varphi'_s = 0$ as is shown by (16). Since the right support of U^* is choosen as small as possible it is the closed span of all $A\varphi'_s$ with $A \in \mathcal{B}(\mathcal{H})\otimes\mathbf{1}'$. As these operators A commute with $\dot{U}U^*$ one concludes $\dot{U}U^* = 0$. Multiplying from the left with U shows $\dot{U}Q = 0$. Now differentiating $UQ = U$ finally yields

$$\dot{U} = 0 \quad \text{or} \quad U_s = \text{const.} \tag{24}$$

This important results reads: If the curves of normed vectors $s \to \varphi_s$ and $s \to \varphi'_s$ both give parallel lifts of (1), i.e. if both curves minimize (7), then there is a s-independent partial isometry U of the form (15) such that

$$\varphi'_s = U\varphi_s \quad \text{with} \quad U^*\varphi'_s = \varphi_s \tag{25}$$

It follows that

$$< \varphi_0, A\varphi_1 >=< \varphi'_0, A\varphi'_1 > \quad \text{for all} \quad A \in \mathcal{B}(\mathcal{H}) \otimes \mathbf{1}' \tag{26}$$

In particular one can define correctly

$$\text{Berry}(s \mapsto \varrho_s) =< \varphi_0, \varphi_1 > \tag{27}$$

for smooth curves (1) of constant rank.

If (2) is a lift of (1) it is possible to switch to another representation. To every normed curve (9) there is according to (8-7) und (8-8) a curve

$$s \mapsto C_s, \quad 0 \leq s \leq 1 \tag{28}$$

of operators (maps) from $\mathcal{H}$ into $\mathcal{H}'$ satisfying

$$\varrho_s = C_sC_s^* \qquad \text{for all } s \tag{29}$$

It is now simply a matter of translation to rewrite some of the results above in terms of these operastors. (12) occurs to be equivalent to[23]

$$C_s^*\dot{C}_s - \dot{C}_s^*C_s = 0 \tag{30}$$

See also[28,29,30] for further discussions and calculations. As (29) implies norm one for the associated curve of vectors of $\mathcal{H}^{ext}$, (30) is another necessary and sufficient condition for paralleltiy if the supports of (1) change smoothly. If (30) is valid, i.e. for a parallel lift, (26) now reads

$$C_0\, C_1^* \qquad \text{depends only on } s \mapsto \varrho_s \tag{31}$$

and the value (27) is the trace of $C_0\, C_1^* \in \mathcal{B}(\mathcal{H})$.

Now I return to arbitrary lifts. Comparing (29) and (30) it is obvious that a *local gauge transformation*

$$C_s \mapsto C_s U_s, \qquad U_s \in \mathcal{B}(\mathcal{H}') \tag{32}$$

with unitary U_s will not change (29). (This is also true for the larger class of partial isometries if relevant support condition are fullfilled. But this will be ignored for simplicity.) Therefore one may ask wether the parallelity condition (30) can be understood as a parallel transport of a gauge theory living, so to say, somehow between $\mathcal{H}'$ and $\mathcal{H}$ but mostly on $\mathcal{H}'$. To make this vague idea handable one first converts (30) into an operator valued differential form

$$C_s^*\, \mathrm{d}C_s - \mathrm{d}C_s^*\, C_s \tag{33}$$

the integral curves of which are the parallel ones. This form, however, will not be a connection form with respect to a local gauge transformation (32).

Such a connection form, called $\mathbf{A}$ will be defined implicitely by[31]

$$C_s^*\, \mathrm{d}C_s - \mathrm{d}C_s^*\, C_s = C^*C\, \mathbf{A} + \mathbf{A}\, C^*C \tag{34}$$

$\mathbf{A}$ is a differential 1-form with values in $\mathcal{B}(\mathcal{H}')$. It depends by its very definition on operators mapping $\mathcal{H}'$ into $\mathcal{H}$ and their adjoints.

The definition is completed by

$$<\psi', \mathbf{A}\psi'> = 0 \quad \text{if} \quad C\psi' = 0 \quad \text{with} \quad \psi' \in \mathcal{H} \tag{35}$$

$\mathbf{A}$ is undefined (singular) at tangential elements $\dot{C}$ in the direction of which the rank of CC^* is changing. These directions must be excluded. With this restriction in mind, (34) and (35) determine $\mathbf{A}$ *uniquely*. This can be seen very easily by sandwiching (34) with a complete orthonormal frame of eigenvectors of CC^*.

The uniqueness yields

$$\mathbf{A}^* + \mathbf{A} = 0 \tag{35}$$

and, the main point of its introduction, $\mathbf{A}$ behaves with respect to *local* gauge transformations $C \to CU$ as

$$\mathbf{A} \mapsto U^*\mathbf{A}U + U^*\,\mathrm{d}U \tag{36}$$

which is proved by applying a local gauge to (34) and (35) and using the uniqueness property.

The connection form determines a *covariant differentiation* for (manifolds of) maps from $\mathcal{H}'$ into $\mathcal{H}$ and their adjoints. For instance one has

$$\mathrm{D}C := \mathrm{d}C - C\mathbf{A} \tag{37}$$

and

$$\mathrm{D}C^* := \mathrm{d}C^* + \mathbf{A}C^* \tag{38}$$

Of course (37) and (38) are adjoints one from another. The change in sign is due to (35).

It is of some use to introduce a further differential 1-form, $\mathbf{G}$ with values in $\mathcal{B}(\mathcal{H})$ by

$$\mathrm{D}\,C = \mathbf{G}\,C \tag{39}$$

By this $\mathbf{G}$ is defined for vectors of the form $C\psi'$ only, and the definition will be completed soon. Substituting

$$\mathrm{d}C = C\mathbf{A} + \mathbf{G}C \tag{40}$$

which comes from (37) and (40) into (34) results in

$$C^*(\mathbf{G} - \mathbf{G}^*)C = 0 \tag{41}$$

Thus $\mathbf{G}$ is is hermitian if resricted to the subspace of the vectors $C\psi'$.

The definition of $\mathbf{G}$ can now be completed without destroying (39) and (41) by

$$\mathbf{G} = \mathbf{G}^* \quad \text{and} \quad <\psi, \mathbf{G}\psi> = 0 \quad \text{if} \quad C^*\psi = 0 \tag{42}$$

($C^*\psi = 0$ is equivalent with the existence of a ψ' that produces $\psi = C\psi'$.)

$\mathbf{A}$ is obviously zero if restricted on a parallel curve. Because of (49) the parallel shifts are integral curves of

$$\mathrm{d}C - \mathbf{G}C \tag{43}$$

as well. It is therefore interesting to determine $\mathbf{G}$ alternatively.

Taking the total differential

$$\mathrm{d}(CC^*) = \mathrm{d}C\,C^* + C\,\mathrm{d}C^* \tag{44}$$

and replacing on the right hand side the total differential by the covariant ones, (37) and (38), one gets

$$\mathrm{d}(CC^*) = \mathrm{D}C\,C^* + C\,\mathrm{D}C^* \tag{45}$$

where the terms with $\mathbf{A}$ have cancelled. It remains to insert (39) and to replace CC^* by ϱ according to (8-8) or (30) to obtain

$$\mathrm{d}\varrho = \mathbf{G}\varrho + \varrho\mathbf{G} \tag{46}$$

REFERENCES

1. Fano, U., Rev. Mod. Phys. 29 (1957) 74
2. Wehrl, A., Rev. Mod. Phys. 50 (1978) 221
3. Bluhm, K.: Density Matrix Theory and Applications., Plenum Press, New York, 1981
4. Berry, M. V. Proc. Royal. Soc. Lond. A 392 (1984) 45
5. Simon, B., Phys. Rev. Lett. 51 (1983) 2167
6. Born, M., Z. Phys. 40 (1926) 167
7. Born, M. and Fock, V., Z. Phys. 51 (1928) 165
8. Born, M. and Oppenheimer, J., Ann. d. Phys. 84 (1927) 457
9. Aharonow, Y., Anandan, J., Phys. Rev. Lett. 58 (1987) 1593
10. Littlejohn, R.G., Phys. Rev. Lett. 61 (1988) 2159
11. Uhlmann, A., Annalen d. Phys. 46 (1989) 63
12. Berry, M., Quantum Adiabatic Holonomy. In: Anomalies, phases, defects (ed. M.Bregola, G.Marmo, G.Morandi), Bibliopolis, Naples, 1990
13. Geometric Phases in Physics. (ed. A.Shapere, F.Wilczek), World Scientific Publishing Co., Singapure, 1990
14. Akemann, C.A., Pedersen, G.K., Facial Structure in Operator Algebra Theory. Preprint, Kopenhavns, June 1990.
15. Uhlmann, A., Rep. Math. Phys. 7 (1975) 449
16. Uhlmann, A., Rep. Math. Phys. 9 (1976) 273
17. Raggio, G. A., Lett. Math. Phys. 6 (1982) 233
18. Araki, H. and Raggio, G. A., Lett. Math. Phys. 6 (1982) 237
19. Alberti, P. M., Lett. Math. Phys. 7 (1983) 25
20. Alberti, P. M. and Uhlmann, A., Lett. Math. Phys. 7 (1983) 107
21. Cantoni, V., Comm. Math. Phys. 44 (1975) 125
22. Uhlmann, A., Ann. d. Physik 42 (1985) 524
23. Uhlmann, A., Rep. Math. Phys. 24 (1986) 229

24. Uhlmann, A., Parallel Transport and Holonomy along Density Operators. In: "Differential Geometric Methods in Theoretical Physics" (ed. H.D.Doebner, J.D.Hennig), World Sci. Publ., Singapore 1987, p. 246 - 254
25. Bures, D. J. C., Trans. Amer. Math. Soc. 135 (1969) 199
26. Fock, V., Z. Phys. 49 (1928) 323
27. Wilczek, F. and Zee, A., Phys. Rev. Lett. 52 (1984) 2111
28. Dabrowski, L., Grosse, H., On Quantum Holonomy for Mixed States. Preprint, Wien, UWThPh-1988-36
29. Dabrowski, L., Jadcyk, A., Quantum Statistical Holonomy. Preprint, Trieste, 155/88/FM
30. Dabrowski, L. A Superposition Principle for Mixed States? Preprint, Trieste, 156/88/FM
31. Uhlmann, A., Gauge Field Governing Parallel Transport along Mixed States. Preprint, Wien, UWThPh-1990-25