Annales de la Fondation Louis de Broglie, Volume 18, n° 1, 1993

Quantum mechanical correlations between subsystems as an aspect of tensor algebra

S. Bergia

Dip. di Fisica, Univ. di Bologna and I.N.F.N., Sez. di Bologna*
Via Irnerio 46, 40126, Bologna, Italy

ABSTRACT. I give a new discussion of a 'diagonalization theorem' on composite systems, described quantum-mechanically, established by J. von Neumann in connection with his analysis of the measuring process. My approach stresses the algebraic aspects of the proof, which are made evident by considering the state "vector" of the composite system as a $\binom{0}{2}$ tensor, whose nature, similar to that of a manifold's fundamental tensor, a non-degenerate, not necessarily symmetric $\binom{0}{2}$ tensor, is pointed out. Since it is independent of specific assumptions, as emphasized in my presentation, the theorem sheds light on the universal character of the Einstein-Podolsky-Rosen correlations, analysed in recent papers by L. Accardi and H. Primas. Such correlations are in fact akin to those established by a manifold's fundamental tensor between vectors and co-vectors. The history of the diagonalization theorem is briefly outlined, and its connection with generalizations of Bell's theorem is pointed out.

RESUME. Je donne une nouvelle discussion d'un "théorème de diagonalisation", valable pour les systèmes composés en mécanique quantique, qui fut établi par J. von Neumann dans le cadre son analyse du processus de mesure. La discussion souligne les aspects algébriques de la démonstration qui sont mis en évidence si l'on considère le vecteur d'état du système composé comme un tenseur $\binom{0}{2}$, dont on souligne l'affinité avec le tenseur fondamental d'une variété, c'est-à-dire un tenseur non dégénéré, non nécessairement symétrique. Puisque le théorème, comme il résulte de ma démonstration, ne dépend d'aucune hypothèse spécifique, il éclaircit le caractère universel des corrélations d'Einstein-Podolsky-Rosen, analysé dans des articles récents de L. Accardi et H. Primas. Ces corrélations sont

* partially supported with M.P.I. 60% funds

en effet proches de celles établies entre vecteurs et covecteurs par le tenseur fondamental. L'histoire du théorème de diagonalisation est esquissée en soulignant sa relation avec des généralisations du théorème de Bell.

1. Introduction.

Correlations between subsystems of a physical system described in terms of quantum mechanics are best analysed in terms of a general theorem established by von Neumann in his treatise [1].

The theorem can be formulated as follows [2]: let H_1 and H_2 be two separable Hilbert spaces and let $\mathbf{\Phi}$ be a unit vector in $H_1 \otimes H_2$. Then there exist two orthomormal bases $\{\psi_i\}, \{\eta_i\}, i = 1, 2, \ldots$, respectively of H_1 and H_2 , and positive numbers w_i such that:

$$\mathbf{\Phi} = \sum_i \sqrt{w_i}\psi_i \otimes \eta_i \tag{1.1}$$

Since, given two bases, $\{\phi_m\}$ and $\{\xi_n\}, m, n = 1, 2, \ldots$, respectively in H_1 and H_2 , an expansion of the form

$$\mathbf{\Phi} = \sum_{m,n} c_{mn}\phi_m \otimes \xi_n \tag{1.2}$$

is obviously valid, the emphasis is on the one-to-one correlation that expansion (1.1) establishes between the elements of the two bases.

A comparison between expansions (1.1) and (1.2) suggests that the essence of the theorem amounts to nothing more than the diagonalization of the matrix C of elements c_{mn}.

In the general case, however, C is an "infinite dimensional matrix", and the algebraic aspects of the proof are likely to be overshadowed (von Neumann original demonstration is no exception, as can be checked by general inspection). The analytical aspects, related to the polar decomposition of Hilbert-Schmidt operators [3], are of course important. Nevertheless, as I will argue, it is the peculiarities of the algebraic aspects what makes the theorem an indispensable tool in discussions about the foundations of quantum mechanics. Since the main purpose of this paper, whose scope is therefore mainly pedagogical[1], is to illustrate these peculiarities, I shall avoid the complications arising as a consequence of the (possible) infinite dimensionality of H_1 and H_2. Actually H_1 and H_2 will be here assumed to be of the same finite dimensionality n, so that they are isomorphic to C^n.

Since the presentation will be rather lengthy due to my wish to be exhaustive, it is perhaps worthwhile to outline the peculiarities of the diagonalization procedure in advance. They are best seen if the tensorial character of the entities involved is clearly indicated from the start. To be definite, and for future reference, let us assume that, in Dirac's terminology, ϕ_m, ξ_n and $\mathbf{\Phi}$ are bra's; then $\mathbf{\Phi}$ is in fact a $\binom{0}{2}$ tensor, and not a $\binom{1}{1}$ tensor corresponding to a linear operator of C^n. The specific tensorial nature of $\mathbf{\Phi}$ makes it reminiscent of a metric. A metric is, however, a non-degenerate[2] symmetric $\binom{0}{2}$ tensor. Here, however, we must give up symmetry from the start, since we are interested in the case in which elements of of the two sub-spaces are physically distinguishable. A non-degenerate, not necessarily symmetric, $\binom{0}{2}$ tensor on a space-time manifold is generally called the manifold's fundamental tensor. Non-degeneracy is a more basic property than symmetry: it is in fact a necessary and sufficient condition for the fundamental tensor to establish a one-to-one correspondence between the manifold's vectors and covectors (one-forms). $\mathbf{\Phi}$ is generally neither symmetric nor non-degenerate. The lack of non-degeneracy forbids in general the establishment of a one-to-one correspondence between kets of the first (second) subsystems and bras of the second (first) one: in other words, the mappings described above have no inverse. The first clue of the demonstration is to deal first with the particular case of a non-degenerate $\mathbf{\Phi}$, i. e. of a "non-degenerate" (i.e. regular) matrix C; complications arising in the general case are then easily disposed of.

The case of a fundamental tensor g *stricto sensu*, in which no complication due to the complex nature of the state space arises, is recalled in section 2, to stress similarities and point out a basic difference. Such a tensor is certainly diagonalizable by a transformation of the class

$$g \longrightarrow g' = O^T g O, \tag{1.3}$$

[1] Its first motivation was in fact the wish to make more explicit and formally more transparent the discussion presented in the 1985 paper by Accardi quoted above.

[2] The term is customary in differential geometry and General Relativity; see footnote 3 below for details.

with O an orthogonal matrix[3], only if it is symmetric. It may thus appear that in general "the fundamental tensor in the state space" will not be diagonalizable, even if non-degeneracy is assumed. It is worth stressing, as the essential peculiarity of the procedure, that this *is not* the kind of diagonalization which is at stake here. In other words, what is sought for is not a basis transformation, such as would be described by Eq. (1.3), but rather a "segment" transformation, involving only one of the basic sets. It is easily shown that this problem has one and only one solution. The essential reason is that the non-symmetric fundamental tensor γ determines a one-to-one relationship between bases in the sub-spaces 1 and 2: the natural choice in 2, say, which can always be made in a unique way, is that of a basis $\{\bar{e}_j^2\}$, which, together with $\{\bar{e}_i^1\}$, diagonalizes γ. The discussion extends straightforwardly to the case of the "fundamental tensor" $\mathbf{\Phi}$ (section 3).

It is somewhat suggestive to read the conclusions of this discussion as follows: the typical non-local quantum mechanical one-to-one correlations holding for entangled[4] states are due to the overall state "vector" acting as a fundamental tensor (a "metric") in the product Hilbert space independently of the causal structure of spacetime.

Section 4 contains a short discussion of the physical meaning of operators appearing in the demonstration; in section 5, the complications arising in the general case, when the non-degeneracy condition is released, are briefly discussed.

As already mentioned, the purpose of this paper is mainly pedagogical. I do not pretend that something really new is implied here, with a possible exception as regards the use of a tensorial language. Actually, it is perhaps of some interest to highlight some episodes in the history and chronicle of the diagonalization theorem. This is briefly done in section 6. Finally, in section 7, the connections with generalizations of Bell's theorem are pointed out.

---

[3] The transformation corresponds to the similarity transformation which applies to $\binom{1}{1}$ tensors. Note that, since a fundamental tensor g is a mapping from $R^n \otimes R^n$ into R, not from R^n into itself, the notion of an eigenvector of g is strictly meaningless, and so is, properly speaking, that of an eigenvalue. One may however agree to call the eigenvalues of g the elements of the matrix in the diagonal form (if it is diagonalizable at all). The matrix is non-degenerate (see footnote 2) if it has no vanishing "eigenvalues".

[4] The term seems to have suddenly become popular: it is, by the way, a fair translation of the German "verschränkt" which was first used by Schrödinger referring to such states.

Notations are, in general, as in Schutz [4]: in particular, vectors are denoted with a bar, co-vectors with a tilde; tensors other than $\binom{0}{1}$ and $\binom{1}{0}$ are in bold-face; elements of a basis (co-basis) are denoted with a subscript (superscript) index.

2. The fundamental tensor analogy.

A fundamental tensor is a $\binom{0}{2}$ non-degenerate tensor, i.e. a bilinear mapping

$$\gamma : \mathbf{R}^n \otimes \mathbf{R}^n \longrightarrow \mathbf{R} \tag{2.1,a}$$

by

$$\bar{U} \in \mathbf{R}^n, \bar{V} \in \mathbf{R}^n \longmapsto \gamma(\bar{U}, \bar{V}), \tag{2.1,b}$$

such that:

$$\gamma(\bar{U}, \bar{V}) = 0 \quad \forall \bar{V} \in \mathbf{R}^n \quad \Rightarrow \bar{U} = 0. \tag{2.2}$$

We shall be interested in the case in which elements of the first and second sub-spaces, hereafter to be denoted as $\mathbf{R}_1^n$ and $\mathbf{R}_2^n$, in Eq.(2.1,a) are physically distinguishable. Then it does not make sense to ask of γ to be symmetric, as it would be asked of a metric tensor.

Bases of vectors $\{\bar{e}_i^1\}, \{\bar{a}_j^2\}$, and dual bases of co-vectors $\{\tilde{\omega}_1^i\}, \{\tilde{\sigma}_2^j\}$, $i, j = 1, \ldots, n$, are introduced in $\mathbf{R}_1^n$ and $\mathbf{R}_2^n$ such that

$$\gamma = \gamma_{ij} \tilde{\omega}_1^i \otimes \tilde{\sigma}_2^j \tag{2.3}$$

with

$$\gamma_{ij} = \gamma(\bar{e}_i^1, \bar{a}_j^2), \tag{2.4}$$

$$< \tilde{\omega}_1^i, \bar{e}_k^1 >= \tilde{\omega}_1^i(\bar{e}_k^1) = \delta_k^i \tag{2.5,a}$$

$$< \tilde{\sigma}_2^j, \bar{a}_l^2 >= \tilde{\sigma}_2^j(\bar{a}_l^2) = \delta_l^j, \tag{2.5,b}$$

where the brackets indicate saturation and the letters denoting bases of the same tensorial character in the two spaces are kept distinguished to stress the independence of the basis choice in $\mathbf{R}_1^n$ and $\mathbf{R}_2^n$. The requirement of non-degeneracy, Eq.(2.2), is equivalent to the condition $\mid \gamma \mid = \det\ (\gamma_{ij}) \neq 0$.

With the tensor γ, other mappings may be associated in a natural way. If it is stipulated that one of the basis forms in Eq.(2.3), say the first one, does not operate, then

$$\gamma_{ij} \tilde{\omega}_1^i \otimes \tilde{\sigma}_2^j(\quad) \tag{2.6}$$

(where the bracket indicates that $\tilde{\sigma}_2^j$ is going to take on its argument) is a mapping

$$\mathbf{R}_2^n \longrightarrow \mathbf{R}_1^{n*} \tag{2.7}$$

($\mathbf{R}_1^{n*}$ is the dual space of $\mathbf{R}_1^n$) by:

$$\bar{V} \in \mathbf{R}_2^n \longmapsto \gamma_{ij}\tilde{\omega}_1^i \tilde{\sigma}_2^j(\bar{V}) = \gamma_{ij} V^j \tilde{\omega}_1^i \tag{2.8}$$

The result is a covector belonging to $\mathbf{R}_1^{n*}$,which we may denote as

$$\tilde{V} = V_i \tilde{\omega}_1^i, \tag{2.9}$$

with

$$V_i = \gamma_{ij} V^j. \tag{2.10}$$

The antilinear operator (2.6) is not an operator *of* $\mathbf{R}_2^n$ (it maps a vector of $\mathbf{R}_2^n$ into a covector of $\mathbf{R}_1^{n*}$). An alternative notation for it is $\gamma(-,\quad)$.

Similarly, one can introduce

$$\gamma_{ij}\tilde{\omega}_1^i(\quad) \otimes \tilde{\sigma}_2^j, \tag{2.11}$$

or, in alternative notation $\gamma(\quad,-)$, which is a mapping

$$\mathbf{R}_1^n \longrightarrow \mathbf{R}_2^{n*} \tag{2.12}$$

by:

$$\bar{U} \in \mathbf{R}_1^n \longmapsto \gamma_{ij}\tilde{\omega}_1^i(\bar{U})\tilde{\sigma}_2^j = \gamma_{ij} U^i \tilde{\sigma}_2^j. \tag{2.13}$$

The result is a covector belonging to $\mathbf{R}_2^{n*}$, which we may denote as

$$\tilde{U} = U_j \tilde{\sigma}_2^j, \tag{2.14}$$

with

$$U_j = \gamma_{ij} U^i. \tag{2.15}$$

Side by side with γ, we introduce a $\binom{2}{0}$ (also non-degenerate) tensor, i.e. a bilinear mapping

$$\gamma^* : \mathbf{R}_1^{n*} \otimes \mathbf{R}_2^{n*} \longrightarrow \mathbf{R} \tag{2.16,a}$$

by:

$$\tilde{U} \in \mathbf{R}_1^{n*}, \tilde{V} \in \mathbf{R}_2^{n*} \longmapsto \gamma^*(\tilde{U}, \tilde{V}) \tag{2.16,b}$$

γ^* may be analysed according to

$$\gamma^* = \gamma^{ij} \bar{e}_i^1 \otimes \bar{a}_j^2, \tag{2.17}$$

with:

$$\gamma^{ij} = \gamma^*(\tilde{\omega}_1^i, \tilde{\sigma}_2^j). \tag{2.18}$$

With the tensor γ^* other mappings may be associated in a natural way. If it is stipulated that one of the basis vectors in Eq. (2.17), say the first one, does not operate, then:

$$\gamma^{ij} \bar{e}_i^1 \otimes \bar{a}_j^2(\quad) \tag{2.19}$$

is a mapping

$$\mathbf{R}_2^{n*} \longrightarrow \mathbf{R}_1^n \tag{2.20,a}$$

by:

$$\tilde{U} \in \mathbf{R}_2^{n*} \longmapsto \gamma^{ij} \bar{e}_i^1 \bar{a}_j^2(\tilde{U}) = \gamma^{ij} U_j \bar{e}_i^1. \tag{2.20,b}$$

The result is a vector belonging to $\mathbf{R}_1^n$, which we may denote as

$$\bar{U} = U^i \bar{e}_i^1, \tag{2.21}$$

with

$$U^i = \gamma^{ij} U_j. \tag{2.22}$$

The antilinear operator (2.19) is not an operator *of* $\mathbf{R}_2^{n*}$ (it maps a covector of $\mathbf{R}_2^{n*}$ into a vector of $\mathbf{R}_1^n$). An alternative notation for it is $\gamma^*(-, \quad)$.

Similarly, one can introduce

$$\gamma^{ij} \bar{e}_i^1(\quad) \otimes \bar{a}_j^2. \tag{2.23}$$

or, in alternative notation, $\gamma^*(\quad, -)$, which is a mapping

$$\mathbf{R}_1^{n*} \longrightarrow \mathbf{R}_2^n \tag{2.24,a}$$

by:

$$\tilde{V} \in \mathbf{R}_1^{n*} \longmapsto \gamma^{ij} \bar{e}_i^1(\tilde{V}) \bar{a}_j^2 = \gamma^{ij} V_i \bar{a}_j^2. \tag{2.24,b}$$

The result is a vector belonging to $\mathbf{R}_2^n$, which we may denote as

$$\bar{V} = V^j \bar{a}_j^2, \tag{2.25}$$

with

$$V^j = \gamma^{ij} V_i \tag{2.26}$$

The diagonalization of the matrix γ may be looked at from different points of view. Let us preliminarily make the obvious remark that, since γ is a mapping from $\mathbf{R}^n \otimes \mathbf{R}^n$ into $\mathbf{R}$ (Eq.2.1,a), and not from $\mathbf{R}^n$ into itself, the notion of eigenvector of γ is meaningless, and so is, properly speaking, that of eigenvalue. We may however agree to call eigenvalues of γ the elements of the matrix in the diagonal form (if it is diagonalizable at all).

With γ in the form of Eqs.(2.3) and (2.4), the matrix γ can be diagonalized if a new basis $\{\tilde{\omega}_2^{j'}\}$ can be found, for instance in subspace $\mathbf{R}_2^{n*}$,

$$\tilde{\sigma}_2^j = \Lambda^j_{j'} \tilde{\omega}_2^{j'} \tag{2.27}$$

(with $\Lambda jj'$ a regular matrix), such that

$$\gamma = \gamma_{ij'} \tilde{\omega}_1^i \otimes \tilde{\omega}_2^{j'} \tag{2.28}$$

with

$$\gamma_{ij'} = \gamma_{ij} \Lambda^j_{j'} = p_{j'} \delta_{ij'}, \tag{2.29}$$

that is:

$$\gamma = p_i \tilde{\omega}_1^i \otimes \tilde{\omega}_2^i \tag{2.30}$$

Note that we cannot claim at this stage that the p_i are real numbers, since γ is not assumed to be symmetric.

For the same reason, in general γ will not be diagonalizable by the trasformation that for a $\binom{0}{2}$ tensor corresponds to a similarity transformation, that is

$$\gamma \longrightarrow \gamma' = O^T \gamma O, \tag{2.31}$$

with O an orthogonal matrix. But this is not what is at stake here, since the procedure outlined does not involve an overall basis transformation, but only a "segment" transformation in one of the subspaces. To show that this problem is always solvable, which is the essential point of our discussion, we observe that $\gamma_{ij'}$, Eq.(2.29), is given by

$$\gamma_{ij'} = \gamma(\bar{e}_i^1, \bar{e}_{j'}^2), \tag{2.32}$$

where $\{\bar{e}^1_i\}$ and $\{\bar{e}^2_{j'}\}$ are the dual bases of $\{\tilde{\omega}^i_1\}$ and $\{\tilde{\omega}^{j'}_2\}$.This may be read as the scalar product, $\bar{e}^1_i \cdot \bar{e}^2_{j'}$, of the vectors $\bar{e}^1_i, \bar{e}^2_{j'}$, generated by the fundamental tensor γ. Eq.(2.29), written as

$$\bar{e}^1_i \cdot \bar{e}^2_{j'} = p_{j'}\delta_{ij'}, \tag{2.33}$$

expresses the weighted orthogonality[5] of the two sets $\{\bar{e}^1_i\}$, $\{\bar{e}^2_{j'}\}$. Given the set $\{\bar{e}^1_i\}$, the problem of finding a set $\{\bar{e}^2_{j'}\}$, such that Eqs.(2.33) hold, has one and only one solution.

Things may be looked at form a slightly different point of view. The scalar product in Eq.(2.33) can be described, in a non-explicitly metric language, as a two-step process as follows.As the first step, the $\mathbf{R}^n_1 \longrightarrow \mathbf{R}^{n*}_2$ mapping of Eq.(2.11) connects with a basis $\{\bar{e}^1_k\}$ a co-basis in $\mathbf{R}^{n*}_2$ according to:

$$\gamma_{ij}\tilde{\omega}^i_1(\bar{e}^1_k)\tilde{\sigma}^j_2 = \gamma_{ij}\delta^i_k\tilde{\sigma}^j_2. \tag{2.34}$$

The right-hand side of Eq.(2.34) identifies indeed a set of co-vectors labeled by k. Let us set

$$\tilde{\omega}^k_2 = \frac{1}{p_k}\gamma_{ij}\delta^i_k\tilde{\sigma}^j_2 \tag{2.35}$$

(no summation over k in the right-hand side! $p_k \neq 0$ as a consequence of the assumption of non-degeneracy). We are thus mapping a basis in $\mathbf{R}^n_1$ into a co-basis in $\mathbf{R}^{n*}_2$ according to

$$\mathbf{f} \equiv \gamma_{ij}\tilde{\omega}^i_1(\quad) \otimes \tilde{\sigma}^j_2 : \mathbf{R}^n_1 \longrightarrow \mathbf{R}^{n*}_2, \tag{2.36,a}$$

by:

$$\bar{e}^1_k \longmapsto \tilde{\omega}^k_2 = \frac{1}{p_k}\gamma_{ij}\delta^i_k\tilde{\sigma}^j_2 \equiv \frac{1}{p_k}\mathbf{f}\bar{e}^1_k, \quad k = 1, \ldots, n. \tag{2.36,b}$$

---

[5] see Courant and Hilbert [5], p. 404; Morse and Feshbach [6], p. 885; in the terminology of these authors, the normalization factor of Eq. (2.37) determines the “biorthogonality relations” expressed by Eq. (2.38). The notion of “biorthogonal sets” apparently first arose in connection with Sturm-Liouville problems for non-Hermitian operators L, in which case the eigenvectors of L and $L^\dagger$ do not coincide, but satisfy biorthogonality relations; it will prove appropriate to our case as soon as the choice $\gamma^{lj} \equiv \gamma_{lj}$ will be made (see below)

As the second step, the saturations

$$< \tilde{\omega}^k, \bar{e}_{j'} > \equiv \tilde{\omega}_2^k(\bar{e}_{j'}^2) = \frac{1}{p_k}\gamma(\bar{e}_k^1, \bar{e}_{j'}^2) \tag{2.37}$$

are performed.

With $\{\tilde{\omega}_2^k\}$ the dual basis of $\{\bar{e}_{j'}^2\}$, it is:

$$< \tilde{\omega}_2^k, \bar{e}_{j'}^2 > = \delta_{j'}^k \tag{2.38}$$

Eqs.(2.37) and (2.38) agree with Eqs.(2.32) and (2.33).

The above discussion may be summarized in the following terms. A fundamental tensor γ determines a natural relationship between bases in the sub-spaces $\mathbf{R}_1^n$ and $\mathbf{R}_2^n$. According to line 1, the natural choice in $\mathbf{R}_2^n$, say, is that of a basis $\{\bar{e}_{j'}^2\}$, which, together with $\{\bar{e}_i^1\}$, diagonalizes γ, much in the same way as a proper metric chooses an orthonormal basis. According to line 2, the natural choice in $\mathbf{R}_2^n$ is that of a basis $\{\bar{e}_{j'}^2\}$, dual to a co-basis $\{\tilde{\omega}_2^k\}$, that γ connects with $\{\bar{e}_k^1\}$; if $\{\bar{e}_{j'}^2\}$ is such as to diagonalize γ, the normalization of $\{\tilde{\omega}_2^k\}$ is fixed (by Eq.(2.36,b)).

As to now, we have given no indication as to the evaluation of the eigenvalues p_i. We notice however the following. The composition of the mappings

$$\mathbf{f} \equiv \gamma_{ij}\tilde{\omega}_1^i(\quad) \otimes \tilde{\sigma}_2^j : \mathbf{R}_1^n \longrightarrow \mathbf{R}_2^{n*}$$

and

$$\mathbf{f}^* \equiv \gamma^{ij}\bar{e}_i^1 \otimes \bar{a}_j^2(\quad) : \mathbf{R}_2^{n*} \longrightarrow \mathbf{R}_1^n \tag{2.39}$$

is $\binom{1}{1}$ tensor, i.e. a mapping of $\mathbf{R}_1^n$ into itself

$$\mathbf{f}^* \circ \mathbf{f} : \mathbf{R}_1^n \longrightarrow \mathbf{R}_1^n \tag{2.40}$$

by:

$$\bar{e}_k^1 \longmapsto \gamma^{ij}\bar{e}_i^1\bar{a}_j^2(\tilde{\sigma}_2^l)\gamma_{rl}\delta_k^r = \gamma^{ij}\gamma_{rl}\delta_j^l\delta_k^r\bar{e}_i^1 = \gamma^{ij}\gamma_{kj}\bar{e}_i^1. \tag{2.41}$$

That is, $\mathbf{f}^* \circ \mathbf{f}$ connects with each vector of the basis $\{\bar{e}_k^1\}$ in subspace $\mathbf{R}_1^n$, the vector $\gamma^{ij}\gamma_{kj}\bar{e}_i^1$, and we may write:

$$\mathbf{f}^* \circ \mathbf{f}\bar{e}_k^1 = \gamma^{ij}\gamma_{kj}\bar{e}_i^1. \tag{2.42}$$

It is thus achieved in one step what was obtained in two steps in the procedure outlined above. According to Eq.(2.42), $\mathbf{f}^* \circ \mathbf{f}$ has matrix elements:

$$\begin{aligned}(\mathbf{f}^* \circ \mathbf{f})lk &=< \tilde{\omega}_1^l, \mathbf{f}^* \circ \mathbf{f}\bar{e}_k^1 >= \gamma^{ij}\gamma_{kj} < \tilde{\omega}_1^l, \bar{e}_i^1 > \\ &= \gamma^{ij}\gamma_{kj}\delta_i^l = \gamma^{lj}\gamma_{kj}\end{aligned} \tag{2.43}$$

If we take, in particular, the matrix elements of γ^* as

$$\gamma^{lj} = \gamma_{lj}, \tag{2.44}$$

then, as a matrix:

$$\mathbf{f}^* \circ \mathbf{f} = \gamma \cdot \tilde{\gamma}; \tag{2.45}$$

Since $\gamma \cdot \tilde{\gamma}$ is symmetric positive semi-definite, it is diagonalizable with non negative eigenvalues.

$\mathbf{R}_1^n$ can be made a metric (Euclidean) space via a scalar product $(\quad,\quad)$; in the basis which diagonalizes $\mathbf{f}^* \circ \mathbf{f}$, it is:

$$(\bar{e}_i^1, \mathbf{f}^* \circ \mathbf{f}\bar{e}_j^1) = w_j\delta_{ij}, \quad w_j \geq 0. \tag{2.46}$$

Alternatively, one may write:

$$< \tilde{\omega}_1^i, \mathbf{f}^* \circ \mathbf{f}\bar{e}_j^1 >= w_j\delta_j^i. \tag{2.47}$$

For the well-defined square root H of the operator $\mathbf{f}^* \circ \mathbf{f}$ holds, in the same basis,

$$H_\beta^\alpha e_k^\beta = \sqrt{w_k} e_k^\alpha, \tag{2.48}$$

where greek indices label the vector components and we have suppressed the superscript for simplicity.

A completeness relation for the basis $\{\bar{e}_{k'}^2\}$ in subspace $\mathbf{R}_2^n$ may be formally written as:

$$\sum_k \bar{e}_{k'}^2 >< \tilde{\omega}_2^{k'} = I. \tag{2.49}$$

To use it in connection with Eq. (2.47), one should see $e_{k'}^2 >$ $(< \omega_2^{k'})$ as acted upon by $\mathbf{f}(\mathbf{f}^*)$ from the left (right). One should also let $\mathbf{f}$ and $\mathbf{f}^*$ act on two arguments, rather than on a single one as in Eqs. (2.36,a and 2.39).

Then, from Eq. (2.36,b), and using dual bases, one gets:

$$< \mathbf{f}, \bar{e}^1_j \bar{e}^2_{k'} >= p_j < \tilde{\omega}^j_2, \bar{e}^2_{k'} >= p_j \delta^j_{k'}. \tag{2.50}$$

Since, with $\mathbf{f}^*$ (Eq. (2.41)) written as

$$\mathbf{f}^* = \gamma^{lj} \bar{e}^1_l (\quad) \otimes \bar{a}^2_j, \tag{2.51}$$

one has the mapping

$$\tilde{\omega}^i_1 \mapsto \bar{e}^2_i = \frac{1}{p_i} \gamma^{lj} \delta^i_l \bar{a}^2_j \equiv \frac{1}{p_i} \tilde{\omega}^i_1 \mathbf{f}^*, \tag{2.52}$$

it is also:

$$< \tilde{\omega}^{k'}_2 \tilde{\omega}^i_1, \mathbf{f}^* >= p_i < \omega^{k'}_2, e^1_i >= p_i \delta^{k'}_i. \tag{2.53}$$

Inserting Eq. (2.49) into Eq. (2.47), and using Eqs. (2.50), (2.53), one obtains:

$$< \tilde{\omega}^{k'}_2 \tilde{\omega}^i_1, \mathbf{f}^* >< \mathbf{f}, \bar{e}^1_j \bar{e}^2_{k'} >= p_i p_j \delta^j_{k'} \delta^{k'}_i = (p_j)^2 \delta^i_j = w_j \delta^i_j. \tag{2.54}$$

The "eigenvalues" of the fundamental tensor γ, Eq. (2.30), are therefore the square roots of the positive eigenvalues of $\mathbf{f}^* \circ \mathbf{f}$, the eigenvalues of the square root operator H. Note that the assumption of non-degeneracy of the fundamental tensor (Eq. (2,36,b)) is equivalent to strict positiveness of $\mathbf{f}^* \circ \mathbf{f}$. This completes the essential of the discussion. For completeness and for further reference, we also consider the mapping

$$\mathbf{f} \circ \mathbf{f}^* : \mathbf{R^{n*}_2} \longrightarrow \mathbf{R^{n*}_1} \tag{2.55}$$

by

$$\tilde{\sigma}^k_2 \longmapsto \gamma_{ij} \tilde{\omega}^i_1 (\bar{e}^1_l) \gamma^{lr} \tilde{\sigma}^j_2 \bar{a}^2_r (\tilde{\sigma}^k_2) = \gamma_{ij} \gamma^{lr} \delta^i_l \delta^k_r \tilde{\sigma}^j_2 = \gamma_{ij} \gamma^{ik} \tilde{\sigma}^j_2. \tag{2.56}$$

That is, $\mathbf{f} \circ \mathbf{f}^*$ connects with the basis $\{\tilde{\sigma}^j_2\}$ the basis $\gamma_{ij} \gamma^{ik} \tilde{\sigma}^j_2$ in subspace $\mathbf{R^{n*}_2}$, and we may write:

$$\mathbf{f} \circ \mathbf{f}^* \tilde{\sigma}^k_2 = \gamma_{ij} \gamma^{ik} \tilde{\sigma}^j_2. \tag{2.57}$$

According to Eq.(2.57),$\mathbf{f} \circ \mathbf{f}^*$ has matrix elements:

$$(\mathbf{f} \circ \mathbf{f}^*) lk =< \bar{a}^2_l, \mathbf{f} \circ \mathbf{f}^* \tilde{\sigma}^k_2 >= \gamma_{ij} \gamma^{ik} < \bar{a}^2_l, \tilde{\sigma}^k_2 >= \gamma_{ij} \gamma^{ik} \delta^k_l = (\tilde{\gamma} \cdot \gamma^*)^k_l \tag{2.58}$$

With the choice of Eq.(2.44), as a matrix:

$$\mathbf{f} \circ \mathbf{f}^* = \tilde{\gamma} \cdot \gamma. \tag{2.59}$$

Of course, $\mathbf{f} \circ \mathbf{f}^*$ is also a symmetric positive semi-definite operator, coinciding, as a matrix, with $\mathbf{f}^* \circ \mathbf{f}$. Hence, it has the same eigenvalues w_i.

3. The case of the theorem.

The state vector of a quantum system made up of two subsystems 1 and 2, with state subspaces $H_1 \equiv \mathbf{C_1^n}, H_2 \equiv \mathbf{C_2^n}$ is represented by a bra $\mathbf{\Phi}$.

$\mathbf{\Phi}$ can be considered as a $\binom{0}{2}$ tensor, i.e. a bilinear mapping

$$\mathbf{\Phi_2^0} : \mathbf{C_1^n} \otimes \mathbf{C_2^n} \longrightarrow \mathbf{C}, \tag{3.1,a}$$

by

$$|\psi_1 > \in \mathbf{C_1^n}, |\psi_2 > \in \mathbf{C_2^n} \longmapsto \Phi_2^0(|\psi_1 >, |\psi_2 >). \tag{3.1,b}$$

If bases of vectors $\{|e_i^1 >\}$, $\{|a_j^2 >\}$, and co-vectors, $\{< \omega_i^1|\}$, $\{< \sigma_j^2|\}$, are introduced in $\mathbf{C_1^n}$ and $\mathbf{C_2^n}$, one may conventionally write

$$\mathbf{\Phi_2^0} \equiv \Phi_{ij} < \omega_1^i| \otimes < \sigma_2^j| \tag{3.2}$$

with

$$\Phi_{ij} = \mathbf{\Phi_2^0}(|e_i^1 >, |e_j^2 >), \tag{3.3}$$

$$< \omega_1^i|e_k^1 >= \delta^i{}_k \tag{3.4}$$

$$< \sigma_2^j|a_l^2 >= \delta^j{}_l. \tag{3.5}$$

The scalar products in Dirac's notation in Eqs.(3.4) and (3.5) may be looked at as saturations.

With the tensor $\mathbf{\Phi}$, other mappings may be associated in a natural way. If it is stipulated that one of the basic bras in Eq. (3.2), say the first one, does not operate, then

$$\Phi_{ij} < \omega_i^1| \otimes < \sigma_2^j \tag{3.6}$$

(where the omission of the bar indicates that σ_2^j is going to take on its argument) is a mapping

$$\mathbf{C_2^n} \longrightarrow \mathbf{C_1^{n*}} \tag{3.7}$$

($\mathbf{C_2^{n*}}$ is the dual space of $\mathbf{C_2^n}$) by:

$$|\psi>\in \mathbf{C_2^n} \longmapsto \Phi_{ij} < \omega_1^i| < \sigma_2^j|\psi>= \Phi_{ij}\psi^j < \omega_1^i| \tag{3.8}$$

The result is a co-vector (bra) belonging to $\mathbf{C_1^{n*}}$, which we may denote as

$$<\psi| = \psi_i < \omega_1^i| \tag{3.9}$$

with:

$$\psi_i = \Phi_{ij}\psi^j. \tag{3.10}$$

The antilinear operator of Eq. (3.6) is not an operator *of* $\mathbf{C_2^n}$ (it maps a vector of $\mathbf{C_2^n}$ into a co-vector of $\mathbf{C_1^{n*}}$).

Similarly, one can introduce:

$$\Phi_{ij} < \omega_1^i \otimes < \sigma_2^j|, \tag{3.11}$$

which is a mapping

$$\mathbf{C_1^n} \longrightarrow \mathbf{C_2^{n*}} \tag{3.12}$$

by:

$$|\psi>\in \mathbf{C_1^n} \longmapsto \Phi_{ij} < \omega_1^i|\phi>< \sigma_2^j| = \Phi_{ij}\phi^i < \sigma_2^j|. \tag{3.13}$$

The result is a co-vector (bra), belonging to $\mathbf{C_2^{n*}}$, which we may denote as

$$<\phi| = \phi_j < \sigma_2^j|, \tag{3.14}$$

with:

$$\phi_j = \Phi_{ij}\phi^i. \tag{3.15}$$

On the other hand, if the state vector of the composite system is represented by a ket Φ, this can be considered as a $\binom{2}{0}$ tensor, i.e. a bilinear mapping

$$\mathbf{\Phi_0^2} : \mathbf{C_1^{n*}} \otimes \mathbf{C_2^{n*}} \longrightarrow \mathbf{C} \tag{3.16,a}$$

by

$$<\psi_1| \in \mathbf{C_1^{n*}}, <\psi_2| \in \mathbf{C_2^{n*}} \longmapsto \mathbf{\Phi_2^0}(<\psi_1|, <\psi_2|). \tag{3.16,b}$$

$\mathbf{\Phi_0^2}$ may be analysed according to

$$\mathbf{\Phi_0^2} = \Phi^{ij}|e_i^1>\otimes|a_j^2>, \tag{3.17}$$

with:

$$\Phi^{ij} = \mathbf{\Phi_0^2}(< \omega_1^i|, < \sigma_2^j|). \tag{3.18}$$

With the tensor $\mathbf{\Phi_0^2}$, other mappings may be associated in a natural way. If it is stipulated that one of the basis vectors in Eq. (3.17), say the first one, does not operate, then:

$$\Phi^{ij}|e_i^1 > \otimes a_j^2 > \tag{3.19}$$

is a mapping

$$\mathbf{C_2^{n*}} \longmapsto \mathbf{C_1^n} \tag{3.20,a}$$

by:

$$< \psi| \in \mathbf{C_2^{n*}} \longmapsto \Phi^{ij}|e_i^1 >< \psi|a_j^2 >= \Phi^{ij}\psi_j|e_i^1 > \tag{3.20,b}$$

The result is a vector (ket) belonging to $\mathbf{C_1^n}$, which we may denote as

$$|\psi >= \psi^i|e_i^1 >, \tag{3.21}$$

with:

$$\psi^i = \Phi^{ij}\psi_j. \tag{3.22}$$

The antilinear operator of Eq. (3.19) is not an operator *of* $\mathbf{C_2^{n*}}$ (it maps a co-vector of $\mathbf{C_2^{n*}}$ into a vector of $\mathbf{C_1^n}$).

Similarly, one can introduce

$$\Phi^{ij}e_i^1 > \otimes|a_j^2 >, \tag{3.23}$$

which is a mapping

$$\mathbf{C_1^{n*}} \longrightarrow \mathbf{C_2^n} \tag{3.24,a}$$

by:

$$< \phi| \in \mathbf{C_1^{n*}} \longmapsto \Phi^{ij} < \phi|e_i^1 > |a_j^2 >= \Phi^{ij}\phi_i|a_j^2 > . \tag{3.24,b}$$

The result is a vector (ket) belonging to $\mathbf{C_2^n}$, which we may denote as:

$$|\phi >= \phi^j|a_j^2 >, \tag{3.25}$$

with:

$$\phi^j = \Phi^{ij}\phi_i. \tag{3.26}$$

Let us now see how the discussion of section 2 may be adapted to the case of the diagonalization of the matrix Φ of elements Φ_{ij}. Considerations similar to those made in section 2 apply to the notion of eigenvalue of the matrix here discussed.

In this case, with $\mathbf{\Phi_2^0}$ in the form of Eqs. (3.2), (3.3), the matrix Φ can be diagonalized if a new basis, $\{< \omega_2^{j'}|\}$, can be found, for instance in subspace $\mathbf{C_2^{n*}}$,

$$< \sigma_2^j| = T^j{}_{j'} < \omega_2^{j'}|, \tag{3.27}$$

such that,

$$\mathbf{\Phi_2^0} = \Phi_{ij'} < \omega_1^i| \otimes < \omega_2^{j'}| \tag{3.28}$$

with

$$\Phi_{ij'} = p_{j'}\delta_{ij'} = \Phi ij T^j{}_{j'}, \tag{3.29}$$

that is:

$$\Phi_2^0 = p_i < \omega_1^i| < \omega_2^i|. \tag{3.30}$$

Again, we cannot claim at this stage that the p_j are real numbers. Nor can we claim that they are all different from 0. Throughout this section, it will be assumed that it is indeed so. This is the equivalent of the assumption of non-degeneracy of the fundamental tensor of section 2.

Recalling that the symmetry of $\mathbf{\Phi_2^0}$ is not requested, in general Φ will not be diagonalizable by the transformation for a $\binom{0}{2}$ tensor which corresponds to a similarity transformation , that is:

$$\Phi \longrightarrow \Phi' = U^\dagger \Phi U, \tag{3.31}$$

with U a unitary matrix. Again, this is not, however, what is involved here, since the procedure does not imply an overall basis transformation, but only a “segment” transformation in one of the subspaces. Since $\Phi_{ij'}$, Eq.(3.29), is given by

$$\Phi_{ij'} = \mathbf{\Phi_2^0}(|e_i^1>, |e_{j'}{}^2>), \tag{3.32}$$

where $\{|e_i^1>\}, \{|e_{j'}^2>\}$ are the dual bases of $\{<\omega_i^1|\}, \{<\omega_2^{j'}|\}$, Φ connects, very much as the fundamental tensor of section 2, with the vectors $|e_i^1>, |e_{j'}{}^2>$, a number, which may be taken as their scalar product. Eq.(3.29) may then be read as expressing the weighted orthogonality of the two sets $\{|e^1_i>\}, \{|e^2_{j'}>\}$. Again, given the set $\{|e^1_i>\}$, the problem of finding a set $\{|e^2_{j'}>\}$, such that Eq.(3.29) hold, has one and only one solution.

Much like in the fundamental tensor analogy, things may be looked at from a slightly different point of view. The scalar product can be described, in a non-explicitly metric language, as a two-step process as follows. As the first step, the $\mathbf{C^n_1} \longrightarrow \mathbf{C^{n*}_2}$ mapping of Eq.(3.12) connects with a basis $\{|e^1_k>\}$ a co-basis in $\mathbf{C^{n*}_2}$ according to:

$$|e^1_k> \longmapsto \Phi_{ij} < \omega^i_1|e^1_k>< \sigma^j_2| = \Phi_{ij}\delta^i{}_k < \sigma^j_2|. \tag{3.33}$$

The right-hand side of Eq.(3.33) identifies indeed a set of co-vectors labeled by k. Recalling that we have assumed $p_k \neq 0, \forall k$, let us set:

$$< \omega^k_2| = \frac{1}{p_k}\Phi_{ij}\delta^i{}_k < \sigma^j_2| \tag{3.34}$$

(no summation over k on the right-hand side!). We are thus mapping a basis in $\mathbf{C^n_1}$ into a co-basis in $\mathbf{C^{n*}_2}$ according to

$$\mathbf{F}^\dagger \equiv \Phi_{ij} < \omega^i_1 \otimes < \sigma^j_2| : \mathbf{C^n_1} \longrightarrow \mathbf{C^{n*}_2}, \tag{3.35,a}$$

by:

$$|e^1_k> \longmapsto < \omega^k_2| = \frac{1}{p_k}\Phi_{ij}\delta^i{}_k < \sigma^j_2| = \frac{1}{p_k}\mathbf{F}^\dagger|e^1_k> . \tag{3.35,b}$$

As the second step, the saturations:

$$< \omega^k_2|e^2_{j'}> = \frac{1}{p_k}\mathbf{\Phi^0_2}(|e^1_k>, |e^2_{j'}>) \tag{3.36}$$

are carried through. With $\{< \omega^k_2|\}$ the "dual basis" of $\{|e^2_{j'}>\}$, it is:

$$< \omega^k_2|e^2_{j'}> = \delta^k{}_{j'}. \tag{3.37}$$

The discussion following Eq. (1.88) can be immediately adapted to these results .

The above discussion may be summarized in the following terms: the state vector $\mathbf{\Phi}$ of a composite system determines a natural relationship between bases in the state spaces of the subsystems $\mathbf{C^n_1}$ and $\mathbf{C^n_2}$. According to line 1, the natural choice in, say, $\mathbf{C^n_2}$ is that of a basis

$\{|e_{j'}{}^2>\}$, which, together with $\{|e^1_i>\}$, diagonalizes Φ, much in the same way as a proper metric chooses an orthonormal basis. According to line 2, the natural choice in $\mathbf{C^n_2}$ is that of a basis $\{|e^2_{j'}>\}$,dual to a co-basis $\{<\omega^k_2|\}$ that $\mathbf{\Phi^0_2}$ connects with $\{e^2_{k'}\}$; if $\{e^2_{j'}\}$ is such as to diagonalize Φ, the normalization of $<\omega^k_2|\}$ is fixed (by Eq.3.35,b).

As regards the determination of the eigenvalues p_i, we note the following. The composition of the mappings:

$$\mathbf{F}^\dagger = \Phi_{ij} < \omega^i_1 \otimes < \sigma^j_2| : \mathbf{C^n_1} \longrightarrow \mathbf{C^{n*}_2}$$

and

$$\mathbf{F} = \Phi^{ij}|e^1_i > \otimes a^2_j : \mathbf{C^{n*}_2} \longrightarrow \mathbf{C^n_1} \tag{3.38}$$

is a $\binom{1}{1}$ tensor, i.e. a mapping of $\mathbf{C^n_1}$ into itself

$$\mathbf{F} \circ \mathbf{F}^\dagger : \mathbf{C^n_1} \longrightarrow \mathbf{C^n_1} \tag{3.39}$$

by

$$\begin{aligned} |e^1_k > \longmapsto & \Phi^{ij}|e^1_i >< \sigma^l_2|a^2_j \Phi_{rl}\delta^r{}_k \\ &= \Phi^{ij}\Phi_{rl}\delta^l{}_j\delta^r{}_k|e^1_i > \\ &= \Phi^{ij}\Phi_{kj}|e^1_i > . \end{aligned} \tag{3.40}$$

That is, $\mathbf{F} \circ \mathbf{F}^\dagger$ connects with each vector of the basis $\{|e^1_k>\}$ in subspace $\mathbf{C^n_1}$, the vector $\Phi^{ij}\Phi_{kj}|e^1_i>$, and we may write:

$$\mathbf{F} \circ \mathbf{F}^\dagger|e^1_k >= \Phi^{ij}\Phi_{kj}|e^1_i > . \tag{3.41}$$

According to Eq.(3.41), $\mathbf{F} \circ \mathbf{F}^\dagger$ has matrix elements:

$$\begin{aligned} (\mathbf{F} \circ \mathbf{F}^\dagger)^l{}_k &=< \omega^l_1|\mathbf{F} \circ \mathbf{F}^\dagger)|e^1_k >= \Phi^{ij}\Phi_{kj} < \omega^l_1|e^1_i > \\ &= \delta^l{}_i = \Phi^{lj}\Phi_{kj} \end{aligned} . \tag{3.42}$$

If we take, as we must in quantum mechanics:

$$\Phi_{kj} = \bar{\Phi}^{kj} \tag{3.43}$$

then, as a matrix:

$$\mathbf{F} \circ \mathbf{F}^\dagger = \Phi \cdot \bar{\tilde{\Phi}} = \Phi \cdot \Phi^\dagger. \tag{3.44}$$

An operator of this form is in general self-adjoint positive semi-definite, and therefore diagonalizable with non-negative eigenvalues.

In the basis which diagonalizes $\mathbf{F} \circ \mathbf{F}^\dagger$, it is:

$$< \omega_1^i|\mathbf{F} \circ \mathbf{F}^\dagger|e_j^1 >= w_j {\delta^i}_j. \tag{3.45}$$

For the well-defined square root H of the operator $\mathbf{F} \circ \mathbf{F}^\dagger$, in the same basis, one has the eigenvalue equation:

$$H^\alpha{}_\beta e^\beta{}_k = \sqrt{w_k} e^\alpha{}_k \tag{3.46}$$

(notations in Eq.(2.48)).

A completeness relation for the basis $\{|e_{k'}^2 >\}$ in subspace $\mathbf{C_2^n}$ reads:

$$\sum_{k'} e_{k'}^2 >< \omega_2^{k'} = I. \tag{3.47}$$

Its use in connection with Eq.(3.45), along lines similar to those following Eq.(2.47), leads to the conclusion that $\mathbf{F}$ and $\mathbf{F}^\dagger$ are "diagonal in mixed (1,2) bases", with real eigenvalues $p_j = \sqrt{w_j}$, the eigenvalues of the square-root operator H.

Note, in particular, that, with $\mathbf{F}$ written as:

$$\mathbf{F} = \Phi^{lj}|e_l^1(\quad) > |a_j^2 >= e_l^1 > \otimes|a_j^2 >, \tag{3.48}$$

one has the mapping:

$$< \omega_1^i| \longmapsto |e_i^2 >= \frac{1}{p_i}\Phi^{lj}{\delta_l}^i|a_j^2 >\equiv \frac{1}{p_i} < \omega_1^i|\mathbf{F}. \tag{3.49}$$

Again, for completeness and for further reference, we also consider the mapping

$$\mathbf{F}^\dagger \circ \mathbf{F} : \mathbf{C_2^{n*}} \longrightarrow \mathbf{C_1^{n*}} \tag{3.50,a}$$

by:

$$\begin{aligned} < \sigma_k^2| \longmapsto & \bar{\Phi}_{ij} < \omega_1^i|e_l^1 > \Phi^{lr} < \sigma_2^j| < \sigma_2^k|a_r^2 > \\ & = \bar{\Phi}_{ij}\Phi^{lr}{\delta^i}_l{\delta^k}_r < \sigma_2^j| = \bar{\Phi}_{ij}\Phi^{ik} < \sigma_2^j| \end{aligned} . \tag{3.50,b}$$

That is, $\mathbf{F} \circ \mathbf{F}^\dagger$ connects with the basis $\{< \sigma_k^2|\}$ the basis $\bar{\Phi}_{ij}\Phi^{ik}\{< \sigma_2^j|\}$ in subspace $\mathbf{C_2^{n*}}$, and we may write:

$$< \sigma_2^k|\mathbf{F}^\dagger \circ \mathbf{F} =< \sigma_2^j|\bar{\Phi}_{ij}\Phi^{ik}. \tag{3.51}$$

According to Eq. (3.51), $\mathbf{F}^\dagger \circ \mathbf{F}$ has matrix elements:

$$(\mathbf{F}^\dagger \circ \mathbf{F})_l^{\ k} =< \sigma_2^k | \mathbf{F}^\dagger \circ \mathbf{F} | a_l^2 >= \bar{\Phi}_{ij} \Phi^{ik} \delta^j{}_l = (\Phi^\dagger \cdot \Phi)_l^{\ k}. \tag{3.52}$$

Therefore, as a matrix:

$$\mathbf{F}^\dagger \circ \mathbf{F} = \Phi^\dagger \cdot \Phi. \tag{3.53}$$

Of course, $\mathbf{F}^\dagger \circ \mathbf{F}$ is also a Hermitian positive semi-definite operator, coincident, as a matrix, with $\Phi \circ \Phi^\dagger$. It has therefore the same eigenvalues w_i.

4. Physical interpretation of the operators $\mathbf{F} \circ \mathbf{F}^\dagger$ and $\mathbf{F}^\dagger \circ \mathbf{F}$.

The composition $\mathbf{\Phi} \circ \mathbf{\Phi}^\dagger$ determines a mapping from $\mathbf{C_1^n} \otimes \mathbf{C_1^n}$ into $\mathbf{C_1^{n*}} \otimes \mathbf{C_2^{n*}}$, by:

$$\begin{aligned}
&|\psi_1 > \in \mathbf{C_1^n}, |\psi_2 > \in \mathbf{C_2^n} \\
&|\psi_1 >, |\psi_2 > \mapsto \Phi^{ij} |e_l^1 > \otimes |a_j^2 > \bar{\Phi}_{kl} (< \omega_1^k \otimes < \sigma_2^l)(|\psi_1 >, |\psi_2 >) \\
&\qquad = \Phi^{ij} \bar{\Phi}_{kl} |e_i^1 > \otimes |a_j^2 >< \omega_k^1 | \psi_1 >< \sigma_l^2 | \psi_2 > .
\end{aligned} \tag{4.1}$$

$\Phi \circ \Phi^\dagger$ is by definition the density operator of the composite system:

$$\rho |\Phi >< \Phi| \equiv \Phi \circ \Phi^\dagger$$

On the other hand, the density operators for subsystems 1 and 2 are obtained according to:

$$\begin{aligned}
\rho_1 &= \sum_m < \sigma_2^m | \rho | a_m^2 >= \sum_m \Phi^{ij} \bar{\Phi}_{kl} \delta_j{}^m \delta^l{}_m |e_i^1 >< \omega_k^1 | \\
&= \Phi^{lj} \bar{\Phi}_k j |e_l^1 >< \omega_1^k | = (\mathbf{F} \circ \mathbf{F}^\dagger)^l{}_j |e_l^1 >< \omega_1^k |
\end{aligned} \tag{4.2}$$

$$\begin{aligned}
\rho_2 &= \sum_n < \omega_1^n | \rho | e_n^1 >= \sum_n \Phi^{ij} \bar{\Phi}_{kl} \delta_i{}^n \delta_n{}^k |a_j^2 >< \sigma_l^2 | \\
&= \Phi^{ik} \bar{\Phi}_{il} |a_k^2 >< \sigma_2^l | = (\mathbf{F}^\dagger \circ \mathbf{F})_l{}^k |a_k^2 >< \sigma_2^l |
\end{aligned} \tag{4.3}$$

(see Eqs. 3.42 and 3.52). We see then that $\mathbf{F} \circ \mathbf{F}^\dagger$ and $\mathbf{F}^\dagger \circ \mathbf{F}$ represent respectively the density operators for the subsystems 1 and 2. We can therefore conclude that the solution of the diagonalization problem for Φ stems from the solution of the eigenvalue problems for ρ_1 and ρ_2.

5. Complications arising from degeneracy and infinite dimensionality.

In section 3 it was explicitly assumed the strict positiveness of the eigenvalues of $\mathbf{F} \circ \mathbf{F}^\dagger$ and $\mathbf{F}^\dagger \circ \mathbf{F}$, i.e. of ρ_1 and ρ_2. These operators are however, in general, only positive semi-definite. Therefore, in the general case, the procedure outlined in section 3 must be modified.

The essential point is that the co-vectors $< \omega_2^i|$ of Eq.(3.35.b)

$$< \omega_2^i| = \frac{1}{p_i}\mathbf{F}^\dagger |e_i^1 >, \tag{5.1}$$

and the vectors $|e_2^i >$ of Eq. (3.49)

$$|e_i^1 >= \frac{1}{p_i} < \sigma_2^i|\mathbf{F}, \tag{5.2}$$

can only be defined for $p_i \neq 0$. Starting from a complete set in subspace 1 (2), we cannot define in this way a corresponding complete set in subspace 2 (1).

However, the sets may be completed by adding the eigenvectors of ρ_1 and ρ_2 corresponding to vanishing eigenvalues.

The result for the diagonalization theorem is not altered, since no further term appears in Eq. (1).

The complications arising from the infinite dimensionality of the (separable) Hilbert spaces H_1 and H_2 are of analytic character, and concern only the validity of the spectral decompositions:

$$\rho_1 = \sum_{k=1}^{\infty} w'_k P_{\psi_{k'}} \tag{5.3}$$

$$\rho_2 = \sum_{k=1}^{\infty} w''_k P_{\psi_{k''}} \tag{5.4}$$

where $P_{\psi_{k'}}$ and $P_{\psi_{k''}}$ are projection operators onto the eigenvectors of ρ_1 and ρ_2 corresponding to the eigenvalues w'_k, w''_k. However, no problem arises, in particular in connection with the proof that $w'_k = w''_k = w_k$ and with the strict positiveness of ρ_1 and ρ_2.

6. History.

An expansion such as that expressed by Eq. (1.1) may be called a Schmidt decomposition of the state, after the mathematician E. Schmidt [7], who discussed a related problem arising in the theory of integral equations with asymmetric kernels, when quantum mechanics had not yet been heard of. Schmidt's paper is referred to in the treatise of Courant and Hilbert[6]. That such a decomposition is possible in the widest generality for any quantum mechanical state was first shown by von Neumann [1] in his treatise (Schmidt's paper is quoted by von Neumann), as initially recalled. The theorem provided actually the basis of von Neumann's theory of measurement.

The original demonstration by von Neumann contains undoubtedly the essential of what has been here developed in some detail. In particular, operators such as $\mathbf{F}$, $\mathbf{F}^\dagger$ and their compositions are introduced and used much in the same way as here, with only minor differences. However, as already mentioned, the basically algebraic character of the proof remained hidden, partly due to the necessary attention paid to analytic aspects, partly because of the fact that no mention was made of dual spaces and of the tensorial character of the operators introduced.

It is interesting to recall that Schrödinger, as far as one can judge independently of von Neumann (Schrödinger referred to Schmidt through Courant-Hilbert's treatise, and not to von Neumann), had achieved, as early as in 1935 [8] the conclusion that an expansion of the type of Eq. (1.1) should hold in general for a composite quantum system, by arguing in a way which is very close to the one followed here. Schrödinger had in mind the Einstein, Podolsky, Rosen (EPR) paper [9] appeared earlier that year, and his argument intended to make more manifest what was implied by it. Actually, the expansion:

$$\Psi(x,y) = \sum_n c_n g_n(x) f_n(y) \tag{6.1}$$

was his starting point. The meaning he attached to it is the following: $f_n(y)$ is the set of orthonormal eigenfunctions of some set of commuting observables in subsystem 2; Eq.(6.1) is nothing but an expansion of the overall state $\Psi(x,y)$ over the functions $f_n(y)$, whose coefficients $c_n g_n(x)$

---

[6] [5]. p. 134.

depend on x only. The coefficients c_n are introduced to allow for the normalization of the g_k:

$$\int \bar{g}_k(x) g_k(x) dx = 1 \tag{6.2}$$

$|c_k|^2$ then represents the probability that a measurement on 2 leaves it in the state $f_k(y)$. The essential point, from our point of view, comes with the equations

$$c_k g_k(x) = \int \bar{f}_k(y) \Psi(x, y) dy, \tag{6.3}$$

which, together with Eq.(6.2), determine both the coefficients c_k and the functions g_k, except for an inessential phase factor in g_k and for the indetermination of some of the g_k, if the integral on the right hand side vanishes identically for some value of k. This presentation clearly shows that the overall state vector $\Psi(x, y)$ connects naturally a wave function $g_k(x)$ to any wave function $f_k(y)$ of subsystem 2. Schrödinger went on observing that there is no reason why the g_k should be mutually orthogonal. The question is, how the functions f_k must be chosen in order that this happens. Schrödinger found the answer that the inverses of the $|c_k|^2$ and the functions $f_k(y)$ must respectively be the eigenvalues and the eigenfunctions of an homogeneous integral operator, whose kernel is what we have called the density matrix for subsystem 2. This shows that, in some way, our approach is a paraphrase os Schrödinger's.

In his paper, as we said, Schrödinger had in mind the EPR paper, and he stated explicitly that his aim was to show that, even limiting oneself to measurements on one subsystem, the function representative of the other is by no means independent of the particular choice of the (completely arbitrary) observations adopted for the purpose. Schrödinger referred to this peculiarity of the quantum mechanical description of subsystems - seen as physical systems having interacted in the past - as *the* characteristic feature of quantum mechanics[7]. This point has recently repeatedly stressed by Primas [12], [13], who has also insisted on the *universal character* of the EPR correlations. We hope that this note

[7] The must undoubtedly be in the literature many references to, and elaborations upon, these early studies: I have only been made aware of work by Jauch [10] and Kochen [11].

will contribute to clarify, by further developing Accardi's [2] analysis, that this universality stems from the universality of von Neumann's expansion of Eq.(1), i.e. is of algebraic origin, and is ultimately rooted in the fact that the state vector of a composite tensor acts as a fundamental tensor in the tensor product of the subspaces.

The basic diagonalization theorem holds for the most general quantum state of a composite system. The Bohm system [14] is on the other hand a very particular case: the singlet state is in fact a state of vanishing total angular momentum, a quantity which, as a consequence of the invariance under rotation, is assumed to be conserved. Immediate generalizations are still rather specific: the most natural extension appears indeed to be a generic two-particle state of vanishing angular momentum. The way to further extensions is indicated by the observation that one is here dealing with a stationary state, a simultaneous eigenstate of the Hamiltonian, of the total angular momentum and of one of its components (not necessarily corresponding to a vanishing eigenvalue of the latter: a one-to-one correlation obtains also in the case of a non-vanishing component of the overall angular momentum). For such a state, the standard Schmidt decomposition is the Clebsch-Gordan expansion. More generally, one may wish to consider generic stationary states, simultaneous eigenstates of the Hamiltonian and of the set formed by the compatible generators and the Casimir operators of the widest symmetry group of the system (which might be a direct product). The Schmidt expansion will coincide in this case with the corresponding generalized Clebsch-Gordan expansion. In all these cases the quantum one-to-one correlations are a direct consequence of it.

The question arises as to the relation between the Schmidt decomposition and a Clebsch-Gordan expansion. The former is evidently more general, since it does not depend on the assumption that the overall state is a stationary state, nor on the validity of a conservation law. What the theorem shows is that a Schmidt decomposition is nonetheless always possible. Since in the case of a Clebsch-Gordan expansion one-to-one correlations appear to be a consequence of a conservation law, one wonders in the first place as to what is the mechanism which in the general case plays its role. The answer goes as follows: one may associate infinitely many observables $A = \sum_m a_m|\psi_m >< \psi_m|$ and $B = \sum_n b_n|\eta_n >< \eta_n|$, defined by their spectra $\{a_m\}$, $\{b_n\}$, with the orthonormal bases diagonalizing the matrix of the overall state vector. They can always be choosen such that $a_\nu + b_\nu = K$, with K a constant, for each ν. The overall state "vector" is then an eigenstate of the observable $M = A \otimes I + I \otimes B$ corresponding to the eigenvalue K. The observable M is then "conserved" in the state, and for each K a one-to-one correspondence is established between the eigenvalues of the subsystem observables A and B.

On the other hand, in the case the system physical nature is specified, and an explicit conservation law is active, one would like to check on an explicit example the way the general Schmidt decomposition reduces to the corresponding Clebsch-Gordan expansion. This aspect was analysed in Bergia *et al.* [15].

7. Connections with extensions of Bell's theorem.

One aspect of Bell's original philosophy was the intent, specified in the very Introduction of his 1964 paper [16], to provide a set-up in which the EPR argument, in Bohm's version [14], could be experimentally checked. Now, what the EPR experiment points out, is first of all, the peculiarity of the one-to-one quantum correlations between subsystems. Thus an extension of Bell's original framework to any entangled quantum state of a two-body system[8], no matter what the dimensionality of the subsystems state space is [18], implies examining likewise the way in which EPR correlations extend to general two-particle systems. The diagonalization theorem establishes first of all the pattern for this extension. This was clear enough to Schrödinger, who was already perfectly aware that "entanglement" is one of the essential, if not *the* essential feature of quantum mechanics, and that the Schmidt decomposition is the instrument which makes its consequences transparent, and permits the generalization of the EPR situation to *any* quantum mechanical system [8], but has been often overlooked in the recent literature.

The diagonalization theorem proves its usefulness in another respect. It is now agreed that Bell's theorem extends to the general situations just mentioned, that is to say that quantum mechanics predicts results that are inconsistent with local realism for any entangled two-particle state [18, 19, 20, 21, 22]. While everybody agrees that violations of Bell's inequality depend, in the general case, on the choice of the observables and not on the state, provided it is entangled, little emphasis

[8] Further independent extensions refer to decays producing more than two particles [17].

has been given to the role the diagonalization theorem has in achieving proofs of Bell's theorem, though such proofs do make use of it [20, 21].

In conclusion, I express the hope that the stress I have put on the algebraic aspects of the diagonalization theorem may shed light on the essential reasons of its relevance in these fields.

Acknowledgments.

I am indebted to F.Cannata for informing me about references to "biorthogonality" in the books by Courant and Hilbert and Morse and Feshbach and about Schrödinger's paper. I also wish to thank my student C.Piccinni for suggestions about notations. This paper was partly written during a stay at CERN. I wish to thank A.Vitale and G.Turchetti, of the Sezione di Bologna of INFN for giving me this opportunity, and G.Giacomelli for the hospitality offered by his group during my stay at CERN.

References

[1] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, (Springer, Berlin, 1932); English translation: *Mathematical Foundations of Quantum Mechanics*, (Princeton University Press, Princeton, 1955), p. 422 ff.

[2] L. Accardi, *On the universality of the Einstein, Podolsky, Rosen phenomenon*, Princeton (1985, unpublished); see also: L. Accardi, *Foundations of Quantum Mechanics. A Quantum Probabilistic Approach*, in: *The Nature of Quantum Paradoxes*, G. Tarozzi, A. van der Merwe eds., (Kluwer Academic Publishers, Dordrecht, Boston, London, 1988).

[3] M. Reed, B. Simon, *Methods of Modern Mathematical Physics. 1. Functional Analysis*, (Academic Press, New York and London, 1972), Th. VI.10, p. 197.

[4] B. Schutz, *Geometrical methods in mathematical physics*, (Cambridge University Press, Cambridge, 1980).

[5] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, (Interscience Publishers, New York, London, Sidney, 1966).

[6] P. H. Morse, H. Feshbach, *Methods of Theoretical Physics*, (McGraw-Hill, New York, Toronto, London, 1953).

[7] E. Schmidt, Math. Ann. **63**, 433-476 (1907).

[8] E. Schrödinger, Proc. Cambridge Philosophical Society **31**, 555 (1935).

[9] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777 (1935).

[10] J. M. Jauch, *Foundations of Quantum Mechanics* (1968).

[11] S. Kochen, *New interpretation of quantum mechanics*, in: *Symposium on the Foundations of Modern Physics*, P. Lahti, P. Mittelstaedt, eds. (World Scientific, Singapore, 1985).

[12] H. Primas, *Chemistry, Quantum Mechanics and Reductionism*, (Springer-Verlag, Berlin, Heidelberg, New York, 1983).

[13] H. Primas, *Kann Chemie auf Physik reduziert werden?*, manuscript for *Chemie in unserer Zeit* (1983).

[14] D. Bohm, *Quantum Theory*, (Prentice Hall, New York, 1951), pp. 614-622.

[15] S. Bergia, F. Cannata, S. Ruffo, M. Savoia, *Group theoretical interpretation of von Neumann's theorem on composite systems*, Am. J. Phys. **47**, 548 (1979).

[16] J. Bell, *On the Einstein-Podolsky-Rosen paradox*, Physics **1**, 195 (1964).

[17] D. M. Greenberger, M. H. Horne, A. Zeilinger, in: Bell's Theorem, Quantum Theory, and Conceptions of the Universe, M. Kafatos, ed., (Kluwer Academic Publishers, Dordrecht, Boston, London, 1989), p.69; D. M. Greenberger, M. H. Horne, A. Shimony, A. Zeilinger, *Bell's theorem without inequalities*, Am. J. Phys. **58**, 1131 (1990); N. D. Mermin, *What's wrong with these elements of reality?*, Phisics Today, June, 9 (1990); *Quantum mysteries revisited*, Am. J. Phys. **58**, 731 (1990); *Extreme quantum entanglement in a superposition of macroscopically distinct states*, Phys. Rev. Lett. **65**, 1838 (1990); *Simple unified form for the major no-hidden variable theorems*, Phys. Rev. Lett. **65**, 3373 (1990); A. Peres, *Incompatible results of quantum measurements*, Phys. Lett. **A151**, 107 (1990).

[18] A. Baracca, S. Bergia, R. Livi, M. Restignoli, *Reinterpretation and extension of Bell's inequality for multivalued observables*, Int. J. Theor. Phys. **15**, 473 (1976); A. Baracca, S. Bergia, F. Cannata, S. Ruffo, M. Savoia, *Is a Bell-type inequality for nondichotomic observables a good test of quantum mechanics?*, Int. J. Theor. Phys. **16**, 491 (1977); S. Bergia and F. Cannata, *Higher-order tensors and tests of quantum mechanics*, Found. Phys. **12**, 723 (1982).

[19] V. Capasso, D. Fortunato, and F. Selleri, *Sensitive observables of quantum mechanics*, Int. J. Theor. Phys. **7**, 319 (1973); D. Fortunato and F. Selleri, *Sensitive observables on infinite dimensional Hilbert spaces*, Int. J. Theor. Phys. **15**, 333 (1976).

[20] S. Popescu and D. Rohrlich, *Generic quantum non-locality*, NBI-HE-91-35/TAUP-1870-91 (1991).

[21] N. Gisin, *Bell's inequality holds for all non-product states*, Phys. Lett. **A154**, 201 (1991).

[22] S. Bergia and F. Cannata, *Two-particle quantum mechanical correlations and local-realistic theories for non-binary observables*, in: *Bell's Theorem and the Foundations of Modern Physics*, A. van der Merwe and F. Selleri, eds. (World Scientific, Singapore), in press.

(Manuscrit reçu le 8 avril 1992)