ANNALES DE L'I. H. P., SECTION A

N. HADJISAVVAS

Distance between states and statistical inference in quantum theory

Annales de l'I. H. P., section A, tome 35, nº 4 (1981), p. 287-309
<http://www.numdam.org/item?id=AIHPA_1981__35_4_287_0>

© Gauthier-Villars, 1981, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section A » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Ann. Inst. Henri Poincaré,
Vol. XXXV, n° 4, 1981, p. 287-309.

Section A :
Physique théorique.

Distance between states and statistical inference in quantum theory

by

N. HADJISAVVAS
Laboratoire de Mécanique Quantique, BP 347, 51062 Reims, France

---

ABSTRACT. — We define the distance between two quantal states as a global measure of the difference of their respective statistical predictions. The expression of this distance in the usual Hilbert space formalism is found. This concept is then used to study the projection postulate and the statistical inference problem in Quantum Theory.

RÉSUMÉ. — Nous définissons la distance entre deux états quantiques comme une mesure globale de la différence de leurs prédictions statistiques respectives. L'expression de cette distance dans le formalisme Hilbertien habituel est déterminé. Par la suite, ce concept est utilisé dans l'étude du postulat de projection et du problème de l'inférence statistique en Mécanique Quantique.

I. INTRODUCTION

Quantum Mechanics, as it stands today, is a physical theory describing the microsystems exclusively from the observational point of view. It would be thus natural for it to be connected with other theories treating the observation on a general level and in purely abstract terms, as Estimation Theory and Information theory. However, if numerous works on this subject appeared lately [*1*] [*2*] [*3*], the elaboration of this connection is far from been achieved. In particular, only few concrete results have been found [*4*] [*5*] [*6*]; other results have no clear physical interpretation, even if they are sometimes very suggestive [*7*].

© Gauthier-Villars

Much of this effort is principally concentrated on the problem of specifying (or estimating) the state of a quantal system by using the available information about it. This is a crusial question, for Quantum Mechanics gives precise statistical predictions on the results of future measurements when the state is known, but on the other hand it says little about the determination of that state. This is especially true for the quantum statistical mechanics, where in general the available information is not sufficient to determine the state.

The study of this question was originated by Jaynes [*8*]. He proposed the so-called « maximum entropy principle » as a way to specify the state which best represents some statistical data. Later the study of a similar subject was undertaken by the present author [*9*], and Gudder, Marchand and Wyss [*10*], i. e. that of finding a state using the available information *and* an initial state.

The scope of our paper is to advance the study towards this later direction. We divide the work in two parts: in the first, we define a concept of *distance* between states, we discuss its physical interpretation, and derive its mathematical properties. In the second part, we use this concept for the study of the problem of the change of the state provoqued by a new information on the system. We consider two cases, depending on the nature of this information. As a result, we derive the so-called « projection postulate » from another, physically much more appealing postulate.

II. THE DISTANCE BETWEEN TWO STATES OF A SYSTEM

1. The definition of the distance.

Let us first specify some constraints to which an appropriate distance should obey. From a mathematical point of view, the distance $d(p_1, p_2)$ of two states p_1 and p_2 should satisfy the following conditions:

a) $$d(p_1, p_2) \geqslant 0\,, \qquad d(p_1, p_2) = d(p_2, p_1)$$

b) $$d(p_1, p_2) = 0 \Leftrightarrow p_1 = p_2$$

It would be also desirable for the distance to satisfy the triangle inequality:

c) $$d(p_1, p_3) \leqslant d(p_1, p_2) + d(p_2, p_3)$$

i. e. to be a metric. We note however that there exist some « distances », as the information discrimination of Kullback [*11*], which are not metrics [*12*].

Several different distances satisfying conditions *a*), *b*) and *c*) have been defined up to now [*13-17*]. Among these, there are two which shall retain our attention because they are the only ones defined on the basis of physical considerations, namely: the distance $d_G(p_1, p_2)$ of Gudder [*13*] and the distance $d(p_1, p_2)$ defined by Jauch, Misra, Gibson [*18*] [*19*], and independently, by Kronfli [*20*] and Hadjisavvas [*9*].

Gudder's distance is introduced as follows. Let Σ be the set of all « statistical » (or « density ») operators (i. e. positive operators of unit trace) in the Hilbert space corresponding to a physical system. As well known, the statistical operators represent the states of the system. Σ is known to be convex: If $0<\alpha<1$ and $W_1, W_2 \in \Sigma$ then the « mixture » $\alpha W_1 + (1-\alpha)W_2$ also belongs to Σ. Now Gudder qualified two states p_1, p_2 represented by the statistical operators W_1, W_2 as « close » if there exists a mixture containing mostly W_1 equal to a mixture containing mostly W_2: that is, if there exist $W_3, W_4 \in \Sigma$ and a small λ such that

$$(1 - \lambda)W_1 + \lambda W_3 = (1 - \lambda)W_2 + \lambda W_4 .$$

Afterwards, based on this conception of closeness, he defined the distance between the two states p_1, p_2 by

$$d_G(p_1, p_2) = \inf\{ \lambda \in [0, 1] : (1-\lambda)W_1 + W_3 = (1-\lambda)W_2 + \lambda W_4 ;\ W_3, W_4 \in \Sigma \} \quad (1)$$

The property of the quantal states to form mixtures corresponds to a certain physical fact. But it seems highly improbable *a priori* that a distance which takes into account only *this* secondary property while neglecting all the other aspects of the concept of state, can properly express the difference between two states. That is why we shall rather be concerned with another distance put forward by Jauch, Misra and Gibson. The definition of this second distance makes use of the von Neumann-Jauch [*21*] [*22*] formalism of « quantum logic » which we briefly recall:

To any physical system corresponds a set L of « propositions ». If the system does not possess any superselection rules, one admits that this set L is isomorphic to the set of all projectors in a complex separable Hilbert space. In this formalism, the states are probability measures on L and are represented, via the Gleason's theorem [*23*], by statistical operators. The relation connecting the physical concepts of proposition and state with their mathematical representatives is the following: if to the proposition « a » and to the state p there correspond, respectively, the projection operator E and the statistical operator W, then the probability to find, by making a suitable measurement, that the proposition « a » is true, is given by the formula:

$$p(a) = \mathrm{Tr}\,(WE) \quad (2)$$

We now have:

DEFINITION 1. — The distance between any two states p_1, p_2 is given by the expression:

$$d(p_1, p_2) = \sup_{a \in L} | p_1(a) - p_2(a) |$$

The justification of this definition can be based on the following argument which, if not unshakable, is much stronger than the usual arguments (when they exist) in favour of other definitions. As we already pointed out in the introduction, Quantum Theory describes the microsystems exclusively through their behaviour upon measurement. In particular, a quantal state does not only yield statistical predictions, but it can also be completely defined by these predictions (this is one of the aspects of Gleason's theorem). Accordingly, the distance between two states p_1, p_2 can be measured by the difference between the statistical predictions yielded by p_1, p_2. Now for a given proposition $a \in \mathrm{L}$, the difference of the statistical predictions of p_1, p_2 about a is $| p_1(a) - p_2(a) |$. Accordingly, an appropriate distance should be based on the study of the whole set of values of the function

$$f(a) = | p_1(a) - p_2(a) | , \qquad a \in \mathrm{L}$$

We shall prove in the following subsection that the set of values of this function is an interval of the form $[0, d]$. Now if d is small, then the difference of predictions $| p_1(a) - p_2(a) |$ is small for any a and thus the two states p_1, p_2 should be considered as close to each other. If on the contrary d is big, then $| p_1(a) - p_2(a) |$ is close to d for an infinite number of propositions a, so that the distance between p_1, p_2 should be big. That is why d seems to be an appropriate measure of this distance, and this explains the choise of the definition 1.

Jauch, Misra and Gibson put forward this distance for the sole purpose of defining a topology in the set of states, in order to take correctly limits in a rigorous study of the scattering process. They determined the topology introduced by $d(p_1, p_2)$, but they did not give the expression of the distance in the Hilbert space formalism. Our scope here is to use the distance in the statistical inference problem in Quantum Theory. We need for this the expression of $d(p_1, p_2)$ for any pair of states, pure or mixtures. This expression, together with some other properties, is given in the following subsection.

2. Explicit form of $d(p_1, p_2)$ in the Hilbert space formalism.

We start by briefly recalling some definitions and results of trace theory [*24*]. An operator A in a complex separable Hilbert space H is said to belong to the trace class iff the series $\sum_{n=1}^{\infty} (f_n, | \mathrm{A} | f_n)$ converges for at least one basis $\{ f_n \}_{n \in \mathrm{N}}$. In that case, the expression $\mathrm{Tr}\, \mathrm{A} = \sum_{n=1}^{\infty} (f_n, \mathrm{A} f_n)$ is independent of the particular basis $\{ f_n \}_{n \in \mathrm{N}}$. The trace class is denoted by $L_1(H)$ and is known to be a Banach space with respect to the norm $\| A \|_1 = \mathrm{Tr} \, | A |$. The statistical operators are by definition, the positive elements with unit norm in $L_1(H)$.

If a self-adjoint operator belongs to $L_1(H)$, then its spectrum is discrete: $\{ \lambda_n \}_{n \in N}$ and one has: $\mathrm{Tr}\, A = \sum_n \lambda_n$ and $\| A \|_1 = \sum_n | \lambda_n |$. Furthermore, the usual bound norm of operators is given in this case by $\| A \| = \sup_{n \in N} | \lambda_n |$.

Some conventions on notations: L(H) is the Banach space of all continuous operators on H. For any subset, $\Gamma \subset H$, we denote by $[\Gamma]$ the algebraic subspace generated by Γ-that is, the set of all finite linear combinations of elements of Γ – and by $\overline{\Gamma}$ the topological closure of Γ. For any vector g, we denote by $P_{[g]}$ the projector on the one dimensional subspace containing g. For any operator A, R(A) will be his range. If A is self adjoint, then A^+ and A^- are his positive and negative parts.

We now give the explicit form of $d(p_1, p_2)$ in the most general case:

THEOREM 1. — Let p_1, p_2 be any two states, represented by the statistical operators W_1, W_2 respectively. Then

$$d(p_1, p_2) = \frac{\| W_1 - W_2 \|_1}{2}$$

Proof. — Since $W_1, W_2 \in L_1(H)$ one also has $W_1 - W_2 \in L_1(H)$. Let $W_1 - W_2 = \sum_n \lambda_n P_{[f_n]}$ be the spectral decomposition of $W_1 - W_2$, where λ_n are its eigenvalues and $\{ f_n \}_{n \in N}$ a complete orthonormal set of eigenvectors. Then, by the definition of $d(p_1, p_2)$ one has:

$$\begin{aligned} d(p_1, p_2) &= \sup \{ | \mathrm{Tr}\,(W_1 E) - \mathrm{Tr}\,(W_2 E) | : E \text{ any projector} \} \\ &= \sup_E | \mathrm{Tr}\,((W_1 - W_2)E) | = \sup_E \left| \sum_n \lambda_n (f_n, E f_n) \right| \end{aligned} \tag{3}$$

On the other hand,

$$\forall n, \quad \forall E : 0 \leqslant (f_n, E f_n) \leqslant 1 \;\Rightarrow\; \forall \lambda_n > 0 : 0 \leqslant \lambda_n (f_n, E f_n) \leqslant \lambda_n$$

and

$$\forall \lambda_n < 0 : \lambda_n \leqslant \lambda_n (f_n, E f_n) \leqslant 0$$

Thus:

$$\sum_{\lambda_n < 0} \lambda_n \leqslant \sum_n \lambda_n (f_n, E f_n) \leqslant \sum_{\lambda_n > 0} \lambda_n \tag{4}$$

Since Tr W_1 = Tr W_2 = 1, we have $\sum_n \lambda_n = \text{Tr}(W_1 - W_2) = 0$ so that

$$\sum_{\lambda_n<0} \lambda_n = -\sum_{\lambda_n>0} \lambda_n .$$

Substitution in (4) gives

$$\forall E : \left| \sum_n \lambda_n(f_n, Ef_n) \right| \leqslant \sum_{\lambda_n>0} \lambda_n \Rightarrow \sup_E \left| \sum_n \lambda_n(f_n, Ef_n) \right| \leqslant \sum_{\lambda_n>0} \lambda_n \quad (5)$$

Set $E_0 = \sum_{\lambda_n>0} P_{[f_n]}$. Then

$$\left| \sum_n \lambda_n(f_n, E_0 f_n) \right| = \sum_{\lambda_n>0} \lambda_n \quad (6)$$

Using (3), (5), (6) we obtain:

$$d(p_1, p_2) = \sum_{\lambda_n>0} \lambda_n \quad (7)$$

But $||W_1 - W_2||_1 = \sum_n |\lambda_n| = 2\sum_{\lambda_n>0} \lambda_n$. Substitution in (7) then gives the desired result. ■

We note a remarquable fact: the distance which seemed to us as the most « natural » among all distances defined up to now is given, up to a constant, by the distance $|| W_1 - W_2 ||_1$ introduced by the natural norm of the mathematical space $L_1(H)$ associated to states. However, in spite of the fact that the expression $|| W_1 - W_2 ||_1/2$ is general and elegant, it is hard to calculate in practice. Fortunately, when at least one of the states is pure, the expression of the distance becomes simpler as shown by the following result:

PROPOSITION 1. — Let p_1 be a pure state represented by a normalized vector f, and p_2 any state represented by a statistical operator W. Then:

a) The space $\overline{R(P_{[f]} - W)^+}$ is one-dimensional.

b) $d(p_1, p_2) = || P_{[f]} - W ||$.

Proof. — Let F be the projector on $\overline{R(P_{[f]} - W)^+}$. Then

$$P_{[f]} - W = (P_{[f]} - W)^+ - (P_{[f]} - W)^- \Rightarrow F(P_{[f]} - W)F = F(P_{[f]} - W)^+F$$
$$\Rightarrow FP_{[f]}F = (P_{[f]} - W)^+ + FWF \quad (8)$$

Now $FP_{[f]}F = | Ff \rangle \langle Ff |$ has one-dimensional range, and the same must be true for the right member of (8). Now by corollary 1, ref. [25] one has

$$\overline{R((P_{[f]} - W)^+ + FWF)} = \overline{R((P_{[f]} - W)^+)} + \overline{R(FWF)}$$

and thus $(P_{[f]} - W)^+$ has also a 1-dimensional range. This proves *a*). Now *b*) is a obvious consequence of *a*) and of relation (7). ■

When the states p_1, p_2 are both pure, their mutual distance acquires a very simple form:

COROLLARY 1. — If p_1, p_2 are two pure states represented by the vectors f and g, then

$$d(p_1, p_2) = \sqrt{1 - |(f, g)|^2}$$

Proof. — We just have to calculate the eigenvalues of $P_{[f]} - P_{[g]}$. This has been made, for exemple, in ref. [18] and gives the result:

$$\pm \sqrt{1 - |(f, g)|^2}$$

which, combined with relation (7), proves the corollary. ■

We finally prove the proposition concerning the values of $|p_1(a) - p_2(a)|$ mentionned in the preceding subsection:

PROPOSITION 2. — For any pair of states p_1, p_2 the range of the function f defined on the set of all propositions L by

$$\forall a \in L : f(a) = |p_1(a) - p_2(a)|$$

is the entire interval $\left[0, \frac{\|W_1 - W_2\|_1}{2}\right]$.

Proof. — We already know, by virtue of theorem 1, that

$$\max_{a \in L} f(a) = d(p_1, p_2) = \frac{\|W_1 - W_2\|_1}{2}$$

Now let $\{f_i\}_{i \in I}$ be a basis of $\overline{R(W_1 - W_2)^+}$ and $\{g_j\}_{j \in J}$ a basis of $\overline{R(W_1 - W_2)^-}$. We denote by R_0 the one between these two subspaces that has the smallest dimension, say $R_0 = \overline{R(W_1 - W_2)^+}$. Then we can suppose that $I \subset J$. The idea of the proof is to « turn » continuously R_0 from its initial position until it falls on $\overline{R(W_1 - W_2)^-}$.

For any $i \in I$ and $t \in [0, 1]$ define:

$$f_i(t) = \sqrt{1 - t^2} f_i + t g_i, \qquad R(t) = \overline{[\{f_i(t)\}_{i \in I}]}, \qquad E(t) = \sum_{I} P_{[f_i(t)]}$$

Obviously

$$(f_i(t), f_j(t')) = \delta_{ij}(\sqrt{(1 - t^2)(1 - t'^2)} + tt') \tag{9}$$

and also

$$\| E(t)-E(t') \|^2 = \sup_{\|e\|=1} \| (E(t)-E(t'))e \|^2 = \sup_{\|e\|=1} \sum_I \| (P_{[f_i(t)]} - P_{[f_i(t')]})e \|^2$$
$$= \sup_{\|e\|=1} \sum_{i\in I} \| (P_{[f_i(t)]} - P_{[f_i(t')]}).(P_{[f_i]} + P_{[g_i]})e \|^2$$
$$\leqslant \sup_{\|e\|=1} \sum_{i\in I} \| P_{[f_i(t)]} - P_{[f_i(t')]} \|^2 . \| (P_{[f_i]} + P_{[g_i]})e \|^2 \quad (10)$$

Now corollary 1 and proposition 1 imply the relation

$$\| P_{[f_i(t)]} - P_{[f_i(t')]} \|^2 = 1 - | (f_i(t), f_i(t') |^2$$

Inserting this in relation (10) and using (9), we get $\forall t, t' \in [0, 1]$:

$$\| E(t) - E(t') \|^2 \leqslant 1 - (\sqrt{(1-t^2)(1-t'^2)} + tt')^2$$

which shows that the function $t \to E(t)$ is continuous. Hence the function $t \to h(t) := \text{Tr}\,((W_1 - W_2)E(t))$ is also continuous and, consequently, its range is an interval. For $t = 0$, E(0) is the projector on $R_0 = \overline{R(W_1 - W_2)^+}$, thus

$$h(0) = \frac{\| W_1 - W_2 \|_1}{2}.$$

For $t = 1$, E(1) is the projector on $R_1 \subset \overline{R(W_1 - W_2)^-}$, thus $h(1) < 0$. If a_t is the proposition corresponding to E(t) then $f(a_t) = | h(t) |$ so that f takes on all the values from 0 to $\frac{\| W_1 - W_2 \|_1}{2}$ as was to be proved. ■

With proposition 2 we concluded our study of the distance $d(p_1, p_2)$ and turn now to applications.

III. STATISTICAL INFERENCE IN QUANTUM THEORY

The purpose of this section is to give an application of the defined distance to the problem of statistical inference in Quantum Theory. Before doing so, we shall try to elucidate further what we mean by the denomination « state of the system ». This will help us to justify in the following subsections our choices of definitions and methods. However, we think that these choices and methods may be justified and useful even for an understanding of the concept of state different from ours, or in different contexts. Their use is not necessarily bound to our personal conceptions on the interpretation of Quantum Theory.

1. The « subjective » and the « objective » state.

There are many different concepts which have been called by the commun name « state of the system » by quantum theorists. These concepts differ by the organisation level to which they apply-level of an individual system or of a statistical ensemble of systems— but also by the different conceptions that they express. Throughout this work we shall follow the conventional Quantum Theory.

The states are usually seen as a result of a certain *preparation* of the system. Let us reproduce J. M. Jauch's definition [*21*] « A state is the result of a series of physical manipulations which constitute the preparation of the state. Two states are identical if the relevant conditions in the preparation are identical ». And Jauch explains furthermore that the relevance or not of a condition is not *a priori* known, but is a question of physics.

It is clear that the concept of state so defined has an objective character: a preparation defines a state independently of what the observer thinks or knows about it. In the framework of Quantum Theory, this state is represented, as well known, by a statistical operator. Accordingly, the fundamental problem which a physicist has to solve if he wants to predict the future behaviour of systems produced by a given preparation, is to determine the corresponding statistical operator. In order to do this, he can make use of all the information available about the preparation. This information is usually of two kinds:

a) It may consist of statistical results of measurements made on a large set of systems produced by that preparation. In that case, using induction, he assumes that these results will also be observed in future measurements, i. e. characterize the state produced by the preparation and not the particular ensemble of systems which was used for the measurement. Such information will be called *indirect*.

b) It may concern the preparation itself (assembling of the instruments, boundary conditions, etc.), i. e. have a *direct* character.

The ensemble of all available information, direct or indirect, will be denoted by J.

In practice, the available information on the preparation is rarely sufficient to determine uniquely the corresponding state. If, for exemple, it is indirect and consists in the knowledge of mean values, dispersions, etc. of some observables, it is well known that there exist in general infinitely many states in which these observables have the given mean values and dispersions. We shall call these states « states compatible with J » and note the set of corresponding statistical operators by S_J.

The available information J on a preparation will be called « complete » if it determines the state uniquely, i. e. if the set S_J contains only one element. When J is incomplete, as it usually happens, the state corresponding to the preparation is unknown and one can be only sure of its belonging to S_J. In this case, it is in general not possible to give *objective* predictions concerning the results of measures on systems so prepared, i. e. specify the relative frequencies which will be actually realized. However, in practice one always tries to make use of the available information, even if it is incomplete, in order to give predictions who correspond to a « reasonable degree of belief » [*8*] [*26*] even if it is not certain that they will be verified by the experience. In other words, one tries to find between all elements of S_J, the « most probable » or the one that constitutes the « best estimation » with respect to the available information. This research is in the classical probability theory, the subject of the problem of statistical inference. But one cannot simply transpose to quantum mechanics the methods used in probability theory, for two distinct reasons. First, the quantum-theoretical algorithms are often different from those of the theory of probability. Second, quantum mechanics is a physical theory, so that our available information may have a non statistical character (direct information about the aparata used in the preparation, etc.).

At this point we think that it would be useful to enlarge our language. Given a preparation, we shall say that our information J about this preparation defines a certain « *subjective state* » of the physicist or a state of knowledge. By opposition to the subjective state, we shall call from now on the state of the system defined by the preparation itself « objective state ». We saw that each objective state has a mathematical representative (i. e. a statistical operator) with the help of which all objective statistical predictions concerning the system may be obtained. Now we *postulate* that the subjective state also has a mathematical representative which can be used to obtain statistical predictions, but in this case these predictions will be subjective and they will simply reflect our knowledge about the system in the best possible way. Thus this representative will be the statistical operator described earlier as the « most probable » between the elements of S_J, or the « best estimation » of the objective state.

We summarize the content of this subjection in the following mnemonic definitions:

DEFINITION 2 (Direct and indirect information). — An information concerning the preparation of a system is called indirect if it consists of statistical results of measurements performed on a large number of systems produced by the preparation. It is called direct if it concerns the preparation itself. The total amount of the available information will be denoted by J.

DEFINITION 3 (Objective and subjective state). — We shall say that a preparation defines the objective state of the system, and that an amount of information J about the given preparation defines the « subjective state » or « state of knowledge » of the observer. Both of these states are represented by statistical operators.

DEFINITION 4 (Complete information). — An amount of information J concerning a given preparation is called complete if it determines uniquely the objective state. In this case the subjective and the objective state are represented by the same statistical operator.

In this new language, the statistical inference problem is translated as follows: for a given incomplete amount of information J, find the statistical operator that represents our state of knowledge.

After these remarks on the concept of state, we turn now to the application of the distance to the statistical inference problem. From now on states and statistical operators will be denoted by the same symbol. Since we shall primarily be concerned with subjective states, we shall sometimes drop the adjective « subjective ».

2. Modification of the subjective state due to additional information.

Jaynes was one of the first to study a particular case of the statistical inference problem in Quantum theory, namely the case of information of purely statistical character (i. e. indirect). He defined a new concept, the *entropy* of a quantal state, and proposed the following solution to the problem: for a given amount of information J, let S be the set of all statistical operators compatible with J in the sense defined previously. Then the subjective state is represented by the element of S which has the greatest entropy. This solution is nowadays considered as correct. In fact, as we have shown elsewhere [*27*], this « maximum entropy principle » proposed by Jaynes is implied by the principle of Laplace, at least in the classical case. Yet it has been understood that it is applicable exclusively on the case of indirect information (i. e. mean values of some observables are given). Its application to other cases produces paradoxes [*28*].

We shall study a different case of the statistical inference problem which can be stated as follows; suppose that our state of knowledge on a physical system is W_0. At a given time we acquire new data which change our state of knowledge. The problem is to determine the final state W_1 in terms of the state W_0 and of the new data. In what follows, we shall use the Heisenberg picture, so that there will be no dynamical evolution of the state in time. States can change only when our information on the system changes.

We shall consider two different cases, according to the nature of the new information which changes the state, direct or indirect (Cf. def. 3).

2.1. The case of indirect information.

2.1.1. *Physical investigation.*

When the new information is indirect, we can state our problem as follows. For a given preparation, suppose that our information about it permitted to choose in any way (which we do not specify) a statistical operator W_0 representing this particular state of knowledge. In order to test whether or not W_0 represents also the objective state which this preparation defines, we can mesure some observables on a great number of systems produced by the preparation. This provides some statistical data which may be identical or not to those predicted by W_0. If they are identical, then we have an indication that W_0 is close to the statistical operator representing the objective state, or even that it might be equal to it. Therefore, we maintain W_0 as a representative of the subjective state. If the statistical data differ from those predicted by W_0, then we are obliged to choose another W_1 as representing our state of knowledge, since this later is changed. The problem of the statistical inference is here to find this W_1 on the basis of W_0 and of the new statistical data.

The solution we propose to this problem is based on the following qualitative argument. Let S be the set of all density operators who give predictions in accordance with the new statistical data. Obviously W_1 must be one of them. But on the other hand W_1 has to take into account, as far as possible, the knowledge contained in W_0. This condition can be fulfilled if W_1, while belonging to S, is chosen so as to give the closest possible statistical predictions to those given by W_0, and this for all imaginable measurements. By definition 1 of the distance between states, this is equivalent to saying that $d(W_0, W_1)$ should be as small as possible. In other words, W_1 should minimize the distance $d(W_0, W)$ taken between W_0 and an element $W \in S$.

It may be that this argument, because of its qualitative character, is not decisively convincing. However, it is quite in the spirit of the arguments used in classical statistical inference. We thus raise its conclusion to the status of a postulate.

POSTULATE 1. — *Let* W_0, W_1 *represent the subjective states before and after the acquiring of an information consisting of some statistical data. Let furthermore* S *be the set of all density operators giving predictions in accordance with these statistical data. Then* W_1 *is an element of* S *which satisfies the relation:*

$$d(W_0, W_1) = \inf_{W \in S} d(W_0, W) \tag{11}$$

This is the translation of the physical problem into mathematical language. Indeed now we have to specify S and find between its elements the one which satisfies relation (11).

The first part of the problem, i. e. the determination of S in the case of an indirect information, is simple. Indeed, quantitative statistical data are generally of three kinds:

a) Mean values, dispersions, etc. of some observables are given.

b) The probabilities of the answer « yes » are determined for some propositions.

c) One may determine completely the probability density of the values of some observables.

Now all these results can be expressed as mean values of some self-adjoint operators. In case *a*) this is evident, since dispersions are expressed by the mean values of squares of operators. For case *b*), it is sufficient to remember that propositions are also represented by operators (projectors). As for the case *c*), let us note that if $C = \int \lambda dE_\lambda$ is the spectral decomposition of the operator C, then the probability distribution of the values of C in the state W is given by the formula: $f(\lambda) = \mathrm{Tr}\,(E_\lambda W)$. But Tr (AW) for any self adjoint operator A is, by definition, the mean value of A in the state W. Thus the giving of the probability distribution of C is equivalent to the giving of all the mean values of the projectors E_λ.

We conclude that, indeed, any indirect information can be expressed by the giving of the mean values k_i, $i \in I$ of some observables C_i. Accordingly, the set S of all states compatible with an indirect information has the form:

$$S = \{ W \in \Sigma : \mathrm{Tr}\,(WC_i) = k_i \}, \qquad k_i \in \mathbb{C} \tag{12}$$

where Σ is the set of all density operators and C_i are observables defined for the studied system.

The first part of the mathematical problem raised by the postulate i. e. the determination of the set S for the case of indirect information—is thus solved. The second part—the finding of the operator W_1 in S that satisfies rel (11)—is much more difficult. In fact, we have only found some partial results. In order to facilitate the mathematical treatment, we shall suppose from now on that C_i are continuous operators. If this is the case, then the structure of S is given by the following proposition.

PROPOSITION 3. — The set S defined by (12) is convex and closed (in the $\| \cdot \|_1$ − topology).

Proof. — *a*) For any W′, W″ ∈ S and any $a, b \in \mathbb{R}^+$ such that $a + b = 1$, one has:

$$\mathrm{Tr}\,((aW' + bW'')C_i) = a\,\mathrm{Tr}\,(W'C_i) + b\,\mathrm{Tr}\,(W'C_i) = k_i\,, \qquad \forall i \in I$$

Consequently, $aW' + bW'' \in S$, so that S is convex.

b) The functions

$$f_i : L_1(H) \ni W \rightarrow \mathrm{Tr}\,(WC_i) \in \mathbb{C}$$

being continuous [*29*] it follows that the sets $f^{-1}(k_i)$ are closed. But $S = \bigcap_I f_i^{-1}(k_i)$ so that S is closed. ■

The result of the proposition allows us to apply the theory of best approximation of an element of a Banach space by elements of a closed, convex subset [*30*] [*31*]. In our problem, the Banach space will be $L_1(H)$, and we shall have to find, according to our postulate, the best approximation of W_0 by an element W_1 of S, with respect to the norm $\|\cdot\|_1$ (cf. Theorem 1). The theory of best approximation for convex sets is still far from being sufficiently developed; yet it yields general criteria characterizing W_1, which are much simpler that condition (11). These criteria are given in the mathematical section 2.1.2 by the following theorems: for a general S (cf. relation (12)) by the theorems 3 and 4. For the special case of an S defined by the mean value of one observable C, cf. Corollary 2. Finally, these criteria enable us, in some special cases, to completely determine W_1 (cf. Theorem 6).

2.1.2. *Approximations in the space* $L_1(H)$.

Throughout this subsection, S will be a closed convex subset of Σ and W_0 a density operator not belonging to S. Our scope is to give criteria characterizing the « element of best approximation », i. e. the unknown statistical operator W_1 which satisfies rel. (11).

We shall make use of the basic result of the approximation theory of convex sets in normed spaces, due to I. Singer [*31*], which runs as follows:

THEOREM 2 (I. Singer). — Let E be a normed space, G a convex subset of E, $x \in E \setminus \bar{G}$ and $g_0 \in G$. Then $\|x - g_0\| = \inf_{g \in G} \|x - g\|$ if and only if there exists an $f \in E^*$ (the topological dual of E) such that

$$\|f\| = 1, \quad \operatorname{Re}(f(x-g_0)) = \|x-g_0\|, \quad \operatorname{Re}(f(g_0-g)) \geqslant 0, \quad \forall g \in G.$$

In the case we study, Singer's theorem implies the following result:

THEOREM 3. — If $W_1 \in S$, then a necessary and sufficient condition in order to have

$$\|W_0 - W_1\|_1 = \min_{W \in S} \|W_0 - W\|_1 \tag{13}$$

is the following. There exists a hermitian operator A such that

$$\|A\| = 1, \qquad \operatorname{Tr}(A(W_0 - W_1)) = \|W_0 - W_1\|_1 \tag{14}$$

and

$$\operatorname{Tr}(A(W_1 - W)) \geqslant 0, \qquad \forall W \in S \tag{15}$$

Proof. — We apply Singer's theorem to the case $E = L_1(H)$, $G = \bar{G} = S$, $x = W_0, g_0 = W_1$. It is well known [*29*] that the dual of $L_1(H)$ is isomorphic to L(H). The duality is established via the continuous bilinear form

$$L(H) \times L_1(H) \in (B, V) \rightarrow \operatorname{Tr}(BV) \in \mathbb{C}$$

where

$$|\operatorname{Tr}(BV)| \leqslant \|B\| . \|V\|_1$$

Thus, if relation (13) is true, then an element $B \in L(H)$ exists such that:

$$\| B \| = 1, \qquad \mathrm{Re}\,[\mathrm{Tr}\,(B(W_0 - W_1))] = \| W_0 - W_1 \|_1$$

and $\forall W \in S$: $\mathrm{Re}\,[\mathrm{Tr}\,(B(W_1 - W))] \geqslant 0$. Setting $A = \dfrac{B + B^*}{2}$ we find:

$$\mathrm{Tr}\,(A(W_0 - W_1)) = \mathrm{Re}\,[\mathrm{Tr}\,(B(W_0 - W_1))] = \| W_0 - W_1 \|_1$$
$$\mathrm{Tr}\,(A(W_1 - W)) = \mathrm{Re}\,[\mathrm{Tr}\,(B(W_1 - W))] \geqslant 0$$
$$\| A \| \leqslant \frac{\| B \| + \| B^* \|}{2} = 1$$

And also

$$\| W_0 - W_1 \|_1 = \mathrm{Tr}\,(A(W_0 - W_1)) \leqslant \| A \| . \| W_0 - W_1 \|_1 \Rightarrow \| A \| \geqslant 1$$

Thus $\| A \| = 1$. ■

The problem of finding the element of best approximation is thus transformed by Theorem 3 to a pair of simpler problems. We should first find the general solution of the pair of equations (14) and then examine if there exists one between these solutions that satisfies the condition (15). We establish now a theorem which solves the first of these two problems.

THEOREM 4. — The general solution of the system of equations (24) with respect to the hermitian operator A is the following:

$$A = F - F' + B \tag{16}$$

where F, F′ are, respectively, the projectors on

$$\overline{R(W_0 - W_1)^+} \qquad \text{and} \qquad \overline{R(W_0 - W_1)^-}$$

and B is a hermitian operator the range of which is orthogonal to $R(W_0 - W_1)$ and such that $\| B \| \leqslant 1$.

Proof. — We first show that if B is as described above, then A defined by (16) satisfies the relations (14).

If $x \in R(F)$ then $Bx = 0$ and thus $(x, Ax) = (x, Fx) = \| x \|^2$ which shows that $\| A \| \geqslant 1$. On the other hand, for any $x \in H$ write

$$x = x_1 + x_2 + x_3 + x_4$$

where $x_1 \in R(F)$, $x_2 \in R(F')$, $x_3 \in \overline{R(B)}$ and $x_4 \perp R(F) \oplus R(F') \oplus R(B)$. Then

$$| (x, Ax) | = | (x_1, Fx_1) - (x_2, F'x_2) + (x_3, Bx_3) |$$
$$\leqslant \| F \| . \| x_1 \|^2 + \| F' \| . \| x_2 \|^2 + \| B \| . \| x_3 \|^2 \leqslant \| x \|^2,$$

thus proving that $\| A \| = 1$. Besides, if we set for simplicity $\tilde{W} = W_0 - W_1$ then

$$\mathrm{Tr}\,(A\tilde{W}) = \mathrm{Tr}\,((F - F')\tilde{W}).$$

Now obviously $\tilde{W} = (F - F')|\tilde{W}|$ is the polar decomposition of $\tilde{W}$ so that

$$\mathrm{Tr}\,((F - F')\tilde{W}) = \mathrm{Tr}\,|\tilde{W}| = \|\tilde{W}\|_1$$

which together with the former relation, shows that conditions (14) are satisfied.

Let us now prove that any solution of (14) has the form (16). If

$$\tilde{W} = \sum_n p_n P_{[\psi_n]}$$

is the spectral decomposition of $\tilde{W}$, then (14) implies

$$\sum_n p_n(\psi_n, A\psi_n) = \mathrm{Tr}\,(\tilde{W}A) = \|\tilde{W}\|_1 = \sum_n |p_n| \tag{17}$$

Since $|(\psi_n, A\psi_n)| \leqslant \|A\| \leqslant 1$, we deduce that

$$p_n > 0 \Rightarrow (\psi_n, A\psi_n) = 1\,, \qquad p_n < 0 \Rightarrow (\psi_n, A\psi_n) = -1 \tag{18}$$

Now if $A = \int_{-1}^{1} \lambda dE_\lambda$ is the spectral decomposition of A one has:

$$p_n > 0 \Rightarrow 1 = (\psi_n, A\psi_n) = \int_{-1}^{1} \lambda d(\psi_n, E_\lambda\psi_n) \leqslant \int_{-1}^{1} 1d(\psi_n, E_\lambda\psi_n) = 1$$
$$\Rightarrow (\psi_n, E_\lambda\psi_n) = 0\,, \qquad \forall\lambda < 1 \Rightarrow A\psi_n = \psi_n$$

and analogously,

$$p_n < 0 \Rightarrow A\psi_n = -\psi_n$$

Therefore, if we set by definition $B = A - (F - F')$, we have $B\psi_n = 0$ whenever $p_n \neq 0$. This implies that R(B) is orthogonal to $\overline{R(F-F')} = \overline{R(\tilde{W})}$. Since $A = B + (F - F')$ and $\|A\| = 1$, this orthogonality implies in its turn that $\|B\| \leqslant 1$ as was to be demonstrated. ■

The following proposition is useful for proving the unicity of the element of best approximation for the cases of interest.

PROPOSITION 4. — If the inequality (13) characterizing the elements of best approximation has two solutions with respect to W_1, say W_1, W_1', then any A obeying conditions (14) and (15) obeys also the same conditions with W_1' in the place of W_1.

Proof. — Suppose that W_1, W_1' are solutions of (13) and that A obeys conditions (14), (15). Then one has

$$\|W_0 - W_1\|_1 = \|W_0 - W_1'\|_1$$

and, by virtue of (15),

$$\mathrm{Tr}\,(AW_1) \geqslant \mathrm{Tr}\,(AW_1')$$

Replacing these expressions in (14) we find

$$\mathrm{Tr}\,(\mathrm{A}(\mathrm{W}_0 - \mathrm{W}_1')) \geqslant \mathrm{Tr}\,(\mathrm{A}(\mathrm{W}_0 - \mathrm{W}_1)) = \| \mathrm{W}_0 - \mathrm{W}_1 \|_1 = \| \mathrm{W}_0 - \mathrm{W}_1' \|_1$$

which easily implies

$$\mathrm{Tr}\,(\mathrm{A}(\mathrm{W}_0 - \mathrm{W}_1')) = \| \mathrm{W}_0 - \mathrm{W}_1' \|_1 , \qquad \mathrm{Tr}\,(\mathrm{AW}_1') = \mathrm{Tr}\,(\mathrm{AW}_1) . \quad \blacksquare$$

Let us sum up the results obtained up to now.

Theorem 3 shows that W_1 is an element of best approximation if and only if there exists a common solution A of the equations (14) which also satisfies relation (15). The theorem 4 gives the general solution of (14). Remains now the second part of the problem, that is of finding for which W_1 there exists among the solutions of (14) one which satisfies condition (15). This came out to be considerably more difficult. We have only established some partial results, replacing condition (15) by simpler ones, as we pass from the general to more special cases. These new conditions will be sufficiently simple to enable us, for example, to prove the projection postulate.

We considered up to now a quite general set of states S, which was only bound to be closed and convex. From now on we shall make a particularization, by defining S as the set of all states for which the mean value of an observable C is a number k. We shall make use of the following theorem.

THEOREM 5. — Let $S = \{ W \in \Sigma : \mathrm{Tr}\,(WC) = k \}$, where C is a given hermitian (continuous) operator. Then any element of S is a mixture of pure states belonging to S.

In other words, the theorem states that if $\mathrm{Tr}\, WC = k$, then W can be written as a mixture $\sum_n \lambda_n P_{[\psi_n]}$ such that $\forall \psi_n : (\psi_n, C\psi_n) = k$. The proof of this statement is easy for finite-dimensional spaces. For the more general case of a separable Hilbert space considered here, we gave a proof in reference [*25*].

Theorem 5 has an immediate corollary which provides simpler alternatives to condition (15).

COROLLARY 2. — Let $S = \{ W \in \Sigma : \mathrm{Tr}\,(WC) = k \}$ as above and $W_1 \in S$. Then condition (15) is equivalent to each of the following two conditions:

a) $\mathrm{Tr}\,(AW_1) \geqslant (f, Af), \quad \forall P_{[f]} \in S$.

b) W_1 can be written as a mixture of pure states $P_{[f]} \in S$ which maximize the expression (f, Af) for $P_{[f]} \in S$.

Proof. — (15) $\Rightarrow$ *a*): Obvious.

a) $\Rightarrow$ (15): Any $W \in S$ can be written by virtue of Theorem 5, in the form: $W = \sum_j \alpha_j P_{[g_j]}$ where $P_{[g_j]} \in S$. Condition *a*) now gives

$$\operatorname{Tr}(AW_1) \geqslant (g_j, Ag_j) \qquad \forall j \;\Rightarrow\; \operatorname{Tr}(AW_1) \geqslant \operatorname{Tr}(AW)$$

i. e. condition (15).

a) $\Rightarrow$ *b*): Using again Theorem 5, we can write: $W_1 = \sum_i \lambda_i P_{[f_i]}$ with $P_{[f_i]} \in S$. Applying condition *a*) we get

$$\sum_i \lambda_i (f_i, Af_i) \geqslant (f, Af), \qquad \forall P_{[f]} \in S$$

which easily implies $(f_i, Af_i) \geqslant (f, Af)$, $\forall i$, $\forall P_{[f]} \in S$.

b) $\Rightarrow$ *a*): Obvious. ■

Combining now Theorems 3 and 4 with the preceding corollary, we obtain two criteria characterizing the element of best approximation. The first is condition *a*) above, which should be fulfilled by the element of best approximation W_1 and the operator A, defined by relation (16). It is considerably simpler than condition (15), since it involves comparaison of W_1 with *pure* states only. The second, condition *b*), does not involve any comparaison at all; it expresses a relation between A, which depends on W_1, and W_1 itself. We shall see in the next subsection that, in simple cases, finding a W_1 which satisfies *b*) is a kind of eigenvalue problem which can be easily solved.

2.2. A case of direct information. The projection postulate.

We now come to the second of the two cases considered in the beginning of the Section III-2, in which the information that changes our state of knowledge is direct (cf. Definition 2) i. e. it concerns the preparation itself. In fact, we shall consider only a very special subcase, since our goal is to study the « projection postulate » in quantum mechanics. This subcase can be stated as follows: suppose that the subjective state of a quantal system is W_0. If we carry out on this system a measurement of the first kind, which means, in von Neumann's terminology, that the system is not destroyed by the measurement and that an eventual second measurement, immediately subsequent to the first would have the same result, our problem is to find the subjective state W_1 after the measurement.

We shall now bring forward some explanatory remarks.

a) The initial subjective state W_0 is entirely arbitrary. In particular, it may correspond to a « complete » amount of information (cf. Definition 4) so that it can be mathematically identified to the initial objective state of the system. Here, we stay in the general case where the subjective state may differ from the objective state.

b) The two subjective states W_0 and W_1 correspond to different preparations, and thus to different objective states, in opposition to the case studied in subsec III-2.1. Indeed, the measurement *is part of the preparation* corresponding to W_1.

The problem has a trivial solution only when the measurement is complete, that is when it is a simultaneous measurement of a complete commuting set of observables. The state W_1 is then represented by the unique common eigenvector of these observables which corresponds to the result found. Unfortunately, this is usually not the case. All measurements of observables with continuous spectrum—in particular the position, momentum, the energy of a free particle, etc.—are necessarily incomplete. This is also the case for the observables with degenerate spectrum, when only one such observable is measured. There exist then infinitely many states for which the result observed is true, and the Quantum Mechanical formalism generally applied [*22*] cannot determine W_1 uniquely.

The choice of the state after the measurement W_1 depends, of course, on the particular experimental arrangement which has been used. However we shall make the hypothesis that we ignore completely the experimental arrangement, and that our information consists only of the nature of the observable measured and of the observed result. This hypothesis is not too restrictive nor as abstract as it may sound. Indeed, some information about the experimental arrangement can be included in the result of the measurement. For instance, the position of the apparatus in space implies a localization of the particle, and this can be considered as a part of the result. On the other hand, the Quantum Mechanical formalism never makes a reference to the experimental arrangement when stating its fundamental principles.

Before we propose a solution to that problem, we shall expose briefly some opinions expressed by other authors. We shall see that each one is implicitely or explicitely attached to à different concept of state.

According to the opinion of von Neumann [*22*], Prygovecki [*32*] and others, the problem has no solution: one cannot determine the state after an incomplete measurement. The authors are right, if by the name « state » one understands « objective state » (Definition 3).

Contrarily to von Neumann, some other authors [*33*] [*36*] supply the quantum mechanical formalism with an additional postulate which permits to determine the state W_1 after the measurement uniquely, if the initial state was pure; it is the so-called projection postulate. According to this, if the initial state was represented by the vector *f*, then the final state will be represented by the projection of *f* onto the subspace formed by all the state vectors for which the result of the measurement would be certain i. e., in the vocabulary of subsec. III-1, by all the state vectors compatible with the result of the measurement. As for the domain of applicability of this postulate, some authors apply it without distinction to all cases. Others, like A. Messiah and C. Piron affirm that the measurement should be « ideal » (cf. [*35*], p. 167 and [*36*], p. 68 for the definition of « ideal » measurement given by these two authors). However, Messiah and Piron (and especialy Piron who does it explicitly) use the name « state » to designate only *pure* states, and in addition these later are supposed to have an objective character and describe individual particles (cf. [*36*], p. 19).

In the present context, where « state » means subjective state, the problem is more difficult. Indeed, by definition, to *any* amount of information corresponds a unique subjective state. Accordingly, we cannot affirm, as von Neumann did, that the problem has no solution. On the other hand, we should not make any additional hypothesis concerning the nature, ideal or not, of the measurement.

Our amount of information is precisely given: we know the quantity measured, the result of the measurement, the fact that the measurement was of the first kind, and that the subjective state was W_0 before the measurement takes place. The statistical operator W_1 corresponding to the subjective state after the measurement should represent *this* amount of information.

The solution we shall propose and the arguments on which it is based will be analogous to those in the preceding subsection 2.1. We first translate the problem into a mathematical language: we carry out a measurement of the first kind on a physical system. If the initial (i. e. before the measurement) subjective state was W_0, we seek to determine the final state W_1. Now the result of the measurement of a quantity K can be described by the fact that a certain proposition « *a* » namely « the value of K belongs to a Borel set B »—was found true. Since the measurement is of the first kind, we know that a second measurement would give the same result. Accordingly, the probability of the same proposition « *a* » in the final state W_1 is equal to 1. By the von Neumann isomorphism, proposition « *a* » corresponds to a projector E, and the relation (2) implies that

$$\mathrm{Tr}\,(W_1 E) = 1 \tag{19}$$

We found this condition by taking into account the nature of the observable K and the result found (used to determine the proposition « *a* » and the projector E) as well as the fact that the measurement was of the first kind. On the other hand, we know that the subjective state before the measurement is W_0. Consequently, W_1 has to take into account, as far as possible, the information contained in W_0. By an argument identical to that used in the preceding subsection, we conclude that this will be realized if W_1, while satisfying condition (19) is the closest possible to W_0 in the sense of the distance *d*. We thus propose the following postulate:

POSTULATE 2. — *Let a measurement of the first kind be carried out on a* *physical system. Suppose that the initial subjective state was* W_0 *and that the result of the measurement is represented by a projector* E. *Then the final subjective state* W_1 *will be the closest to* W_0 *between all states satisfying condition* (19). *In order worlds,*

$$\mathrm{Tr}\,(W_1E) = 1 \quad \text{and} \quad d(W_0, W_1) = \inf\{\, d(W_0, W) : \mathrm{Tr}\,(WE) = 1 \,\} \quad (20)$$

If we accept this postulate on the basis of the physical arguments formulated above, then the projection postulate becomes a theorem. Indeed, one can prove the following.

THEOREM 6. — Given a pure state represented by the vector f, and a projector E such that $Ef \neq 0$, then the only state satisfying conditions (20) is also pure, and is represented by the normalized projection $\dfrac{Ef}{\|Ef\|}$.

Proof. — The theorem is an immediate application of the results of the preceding subsection. To see this, set $W_0 = P_{[f]}$, $W_1 = P_{[g]}$ where $g = \dfrac{Ef}{\|Ef\|}$, $B = -(E - P_{[g]})$, $A = F - F' + B$ where F and F′ are the projectors on the one dimensional subspaces corresponding to the eigenvalues $\pm\sqrt{1 - |(f, g)|^2}$ of $P_{[f]} - P_{[g]}$ (cf. Corollary 1). By Theorem 4, W_1 satisfies rel. (14). On the other hand, for any $h \in R(E)$ we have:

$$(h, (F - F')h) = (h, E(F - F')Eh) = \left(h, \frac{E(P_{[f]} - P_{[g]})Ef}{\sqrt{1 - |(f, g)|^2}}\right)$$
$$= -\sqrt{1 - |(f, g)|^2}(h, P_{[g]}, h)$$

and thus

$$(h, Ah) = -1 + (1 - \sqrt{1 - |(f, g)|^2})(h, P_{[g]}h)$$

Consequently, (h, Ah) takes its maximal value on R(E) for $h = g$. By Corollary 2, W_1 satisfies (13) and by Theorem 1 it satisfies rel. (20). Finally $W_1 = P_{[g]}$ is the *unique* solution of (20) because of Proposition 4 and the fact that (h, Ah) is maximized on R(E) only for $h = g$. ∎

This theorem is the exact mathematical translation of the « projection postulate ». Indeed, Theorem 6 and Postulate 2 imply the following corollary.

COROLLARY 3 (projection postulate). — Let a measurement of the first kind be performed on a physical system. Suppose the initial subjective state was a pure state represented by the vector f. Then the final subjective state is also pure and is represented by the projection of f onto the subspace formed by all the state vectors for which the result of the measurement would be certain.

Proof. — Let the result of the measurement be represented by a projection operator E according to our previous discussion. By the von Neumann isomorphism, E is exactly the projector on the subspace formed by all the state vectors for which the result of the measurement would be certain. On the other hand, $\| Ef \|^2 = \mathrm{Tr}\,(WP_f)$ was the *a priori* probability of the observed result in the initial state f. Therefore, $Ef \neq 0$. Now by Postulate 2, W_1 satisfies (20), and by Theorem 6, $W_1 = P_{[g]}$ where $g = \dfrac{Ef}{\| Ef \|}$. ■

IV. PERSPECTIVES

We have shown how the concept of distance can be used to solve a particular case of the statistical inference problem. However, if the method is clearly indicated by a postulate, the concrete results are not general enough. Two lines of research are thus open. The first one is to try to find the final state W_1 for non-pure initial states and for more general informational situations; the second, to compare physically the results with these given by other distances (in any case, it is sure that, for instance, the Bures distance [*14*] gives results different from ours).

There is another important question left. As we have shown elsewhere [*9*] [*37*], when the amount of information is too poor, the subjective state cannot be represented by statistical operators. In many cases, they are correctly represented by positive continuous operators of infinite trace, via a generalization of Gleason's theorem. If would be thus interesting to incorporate in some way these elements as possible initial states W_0 in our formalism.

ACKNOWLEDGMENT

This work has much profited from thorough discussions with Professor M. Mugur-Schächter.

REFERENCES

[*1*] A. Wehrl, *Rev. Mod. Phys.*, t. **50**, 1978, p. 221-260.
[*2*] M. B. Ruskai, *Ann. Inst. H. Poincaré*, t. **19**, 1973, p. 357.
[*3*] J. Cyranski, *Found. Phys.*, t. **8**, 1978, p. 805.
[*4*] W. ochs, *Rep. Math. Phys.*, t. **8**, 1975, p. 109.
[*5*] A. Uhlmann, *Rep. Math. Phys.*, t. **1**, 1970, p. 147.
[*6*] R. Ingarden, *Rep. Math. Phys.*, t. **10**, 1976, p. 43.
[*7*] I. Bialynicki-Birula and J. Mycielski, *Comm. Math. Phys.*, t. **44**, 1975, p. 129.
[*8*] E. T. Jaynes, *Phys. Rev.*, t. **106**, 1957, p. 620; t. **108**, 1957, p. 171.
[*9*] N. Hadjisavvas, *Ann. Fond. Louis de Broglie*, t. **3**, 1978, p. 155.
[*10*] S. Gudder, J. P. Marchand and W. Wyss, *J. Math. Phys.*, t. **20**, 1979, p. 1963.
[*11*] S. Kullback, *Information theory and statistics*. Wiley, New York, 1959.
[*12*] I. Csiszar, *Stud. Scient. Math. Hung.*, t. **2**, 1967, p. 299.
[*13*] S. Gudder, *Comm. Math. Phys.*, t. **29**, 1973, p. 249.

[14] S. GUDDER, J. P. MARCHAND and W. WYSS, *J. Math. Phys.*, t. **20**, 1979, p. 1963.
[15] V. CANTONI, *Comm. Math. Phys.*, t. **44**, 1975, p. 125.
[16] S. GUDDER, *Comm. Math. Phys.*, t. **63**, 1978, p. 265.
[17] H. HADJISAVVAS, Commun. *Math. Phys.*, 1981 *(to be published)*.
[18] J. M. JAUCH, B. MISRA and A. G. GIBSON, *Helv. Phys. Acta*, t. **41**, 1968, p. 513.
[19] B. MISRA, in *Physical reality and Mathematical description. C. Enz and J. Mehra eds, D. Reidel publishing company, Dortrecht, Holland, 1974.*
[20] N. S. KRONFLI, *Int. J. Theor. Phys.*, t. **3**, 1970, p. 191.
[21] J. M. JAUCH, *Foundations of Quantum Mechanics*. Addisson-Wesley, Reading, Mass, 1968).
[22] J. VON NEUMANN, *Mathematical Foundations of Quantum Mechanics*, Princeton Univers. press, Princeton, 1955.
[23] A. M. GLEASON, *J. of Rat. mech. and Analysis*, t. **6**, 1957, p. 885.
[24] R. SCHATTEN, *Theory of cross-norm spaces*, Princeton, 1950.
[25] N. HADJISAVVAS, *Lett. math. Phys.*, t. **5**, 1981, p. 327.
[26] B. D'ESPAGNAT, *Conceptual Foundations of Quantum Mechanics*. W. A. Benjamin Inc., 1971.
[27] N. HADJISAVVAS, *J. Stat. Phys.*, t. **26**, 1981, p. 693.
[28] K. FRIEDMANN and A. SHIMONY, *J. Stat. Phys.*, t. **3**, 1971, p. 301.
[29] J. DIXMIER, *Les algèbres d'opérateurs dans l'espace Hilbertien, Gauthier-Villars*, Paris, 1969.
[30] I. SINGER, *Best approximation in normed linear spaces by elements of linear subspaces*. Springer Verlag, 1970.
[31] I. SINGER, *The theory of best approximation and functional analysis*, S. I. A. M., Philadelphia, 1974.
[32] E. PRYGOVECKI, *Quantum Mechanics in Hilbert space* (Academic Press, 1971).
[33] K. BUGAJSKA and S. BUGASKI, *Bull. Acad. Pol. Sci.*, t. **21**, 1976, p. 873.
[34] E. G. BELTRAMENTI and G. CASSINELLI, *Riv., Nuov. Cim.*, t. **6**, 1976, p. 321.
[35] A. MESSIAH, *Mécanique Quantique*, Dunod, Paris, 1965.
[36] C. PIRON, *Foundations of Quantum Physics*. W. A. Benjamin, Reading, Mass., 1976.
[37] N. HADJISAVVAS and M. MUGUR-SCHÄCHTER, *Lett. Nuovo Cim.*, t. **23**, 1978, p. 439.

(Manuscrit reçu le 23 mars 1981)