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

BERNARD DIU

Statistical inference and distance between states in quantum mechanics

Annales de l'I. H. P., section A, tome 38, nº 2 (1983), p. 167-174
<http://www.numdam.org/item?id=AIHPA_1983__38_2_167_0>

Ann. Inst. Henri Poincaré,
Vol. XXXVIII, n° 2, 1983, p. 167-174.

Section A :
Physique théorique.

Statistical inference and distance between states in quantum mechanics

by

Bernard DIU
Laboratoire de Physique Théorique et Hautes Énergies (*), Paris (**)

RÉSUMÉ. — La notion de distance entre états quantiques, dont une proposition récente voulait faire la base d'un postulat général pour résoudre la question de l'inférence statistique relative en mécanique quantique, s'avère mal adaptée à ce but : dans une classe de situations simples (faisant intervenir des états « subjectifs » ou « objectifs », et une information « directe » ou « indirecte »), le nouveau postulat donne des réponses non-uniques ou incorrectes.

ABSTRACT. — The notion of a distance between quantum states, which was recently proposed as the basis for a general postulate to solve the relative statistical inference question in quantum mechanics, is shown to be inadequate for this purpose : in a class of simple situations (dealing with « subjective » or « objective » states, and with « direct » or « indirect » information), the new postulate leads either to non-unique answers, or to incorrect ones.

1. INTRODUCTION

A new point of view was recently put forward [*1*] concerning the question of relative statistical inference in quantum mechanics, i. e. the pro-

(*) Laboratoire associé au C. N. R. S., n° LA 280.
(**) *Postal address:* Université Paris VII, Tour 33, 1er étage, 2, place Jussieu, 75251 Paris Cedex 05, France.

blem of finding the density operator W_1 which describes the system under consideration after a measurement, knowing the result of the measurement and the density operator W_0 before it was performed. The solution proposed in ref. 1 is based on the use of a distance between quantum states. However, the corresponding prescription was subsequently shown [2] to yield ambiguous answers in some simple situations.

The purpose of the present note is to generalize the results of ref. 2 to other possible choices of the distance between quantum states and to other simple situations. The overall conclusion is that the notion of distance is inadequate for what concerns statistical inference: either it does not lead to a unique answer for the final density operator W_1, or it gives a wrong one.

After briefly recalling (section 2) the postulate proposed in ref. 1, we will apply it to simple problems in order to test it. Section 3 is concerned with cases of « direct information », described in § 3-*a*; it considers first (§ 3-*b*) two particular definitions for the distance, and then (§ 3-*c*) a general class of possible distances. An analogous problem of « indirect information » is analyzed in section 4. Finally, section 5 draws the general conclusion.

2. SUMMARY OF THE NEW PROPOSAL

The problem is as follows. The state of a particular quantum system is described by the density operator W_0. One then performs a measurement, which yields new information. What is the density operator W_1 incorporating this information together with that already contained in W_0?

The author of ref. 1 distinguishes, as is often done, between the « objective state » of a system, which is characterized by its preparation, and the « subjective state » or « state of knowledge », which represents the best estimation of an observer who does not possess complete information about the preparation. He also considers two different ways in which new information can be obtained through a measurement: either the latter is performed on the system itself (« direct information »), or it is performed on an ensemble of other identical systems, prepared in the same way as the one under study (« indirect information »). But the solution he proposes is fundamentally the same in all situations, which should certainly be considered a positive feature.

The starting point is to define a distance between states: to each pair of states, characterized by the density operators W and W′, one associates a number $d(W, W')$ possessing the classical properties of a distance:

$$d(W, W') = d(W', W) \geqslant 0 \ ; \qquad d(W, W') = 0 \Leftrightarrow W' = W . \tag{1}$$

(A particular definition is chosen in ref. 1, but the argument will be generalized to other distances). One then uses this distance as the basis for the solution of the statistical inference problem summarized above. Let S be the set of all density operators W compatible with the information brought by the measurement. The proposal is that the final density operator W_1 be the member of this set which lies closest to the initial density operator W_0, according to the distance previously defined:

$$d(W_1, W_0) = \inf_{W \in S} d(W, W_0). \tag{2}$$

It is then shown in ref. 1 that this new postulate, when applied to a pure initial state and to direct information, yields the usual « projection postulate » of quantum mechanics as a consequence.

3. THE CASE OF DIRECT INFORMATION

3-*a*. A simple problem.

Consider a quantum system, whose initial state is described by the density operator W_0. In the present section, we shall for definiteness deal only with objective states. A measurement is performed on the system, the result being characterized by a particular projector E onto a subspace of the Hilbert space of states [*3*]. The measurement is assumed not to have destroyed the system, and we look for the density operator W_1 describing its state after the measurement.

We now particularize to a very special and simple situation, in which W_0 and E are simultaneously diagonal. To be more concrete (without loosing the essential features of the problem), we will concentrate on a 3-dimensional state space and write simply

$$W_0 = \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{pmatrix}, \qquad E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3}$$

The density operators W compatible with the result of the measurement are then of the form

$$W = \left(\begin{array}{cc:c} \multicolumn{2}{c:}{\multirow{2}{*}{w}} & 0 \\ & & 0 \\ \hdashline 0 & 0 & 0 \end{array}\right). \tag{4}$$

Among these, it is clear (for instance on the basis of the classical « projection postulate ») that the final density operator W_1 is here

$$W_1 = \begin{pmatrix} \dfrac{p_1}{p_1+p_2} & 0 & 0 \\ 0 & \dfrac{p_2}{p_1+p_2} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5}$$

3-*b*. **Answer based on two particular distances.**

The distance advocated in ref. 1 is that originally introduced by Jauch, Misra and Gibson [*4*] (J. M. G.). It is shown in ref. 1 to have the simple expression

$$d_{\mathrm{JMG}}(W, W') = \frac{1}{2} \| W - W' \|_1, \tag{6}$$

where the norm $\| \cdot \|_1$ is given by the trace [*5*]:

$$\| A \|_1 = \mathrm{Tr} \, | A | = \sum_n | \lambda_n |, \tag{7}$$

λ_n being the eigenvalues of A. The simple problem of § 3-*a* was already analyzed in ref. 2 using this distance. To summarize, one finds

$$d_{\mathrm{JMG}}(W, W_0) = \frac{1}{2} [p_3 + | \lambda_+ | + | \lambda_- |], \tag{8}$$

λ_+ and λ_- being the eigenvalues of the 2×2 matrix $w - w_0$, where w is defined by (4) and w_0 by

$$w_0 = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix}. \tag{9}$$

These eigenvalues verify the relation

$$\lambda_+ + \lambda_- = p_3, \tag{10}$$

so that

$$d(W, W_0) \geqslant p_3, \tag{11}$$

the minimum being attained whenever λ_+ and λ_- are both positive.

The same result is found if one replaces the JMG distance by another one, first proposed by Gudder [*6*]:

$$d_G(W, W') = \inf \{ \alpha \in [0, 1]: (1-\alpha)W + \alpha X = (1-\alpha)W' + \alpha X'; \; X, X' \in \Sigma \}, \tag{12}$$

where Σ is the set of all possible density operators associated with the system being considered. One easily finds that $d_G(W, W_0)$ is such that

$$\frac{d_G(W, W_0)}{1 - d_G(W, W_0)} = \sup \{ | \lambda_+ |, | \lambda_- |, p_3 \}. \tag{13}$$

The minimum value of $d_G(W, W_0)$ then corresponds to that of the r. h. s. of (13), which is obvious:

$$\sup \{ | \lambda_+ |, | \lambda_- |, p_3 \} \geqslant p_3. \tag{14}$$

This minimum value is also reached here by all density operators W such that λ_+ and λ_- are both positive, due to formula (10).

Thus, if one applies the postulate of ref. 1 to the simple problem of § 3-*a*, with either the Jauch-Misra-Gibson distance or the Gudder distance, one obtains a whole family of possible final density operators. As discussed in ref. 2, this family contains in particular the following one-parameter subset:

$$W = \begin{pmatrix} q & 0 & 0 \\ 0 & 1 - q & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{15-a}$$

with

$$p_1 \leqslant q \leqslant 1 - p_2. \tag{15-b}$$

3-*c*. **Answer based on a more general class of distances.**

Suppose now that we try the same prescription with some other distance $d(W, W')$. Let us assume that, when applied to a pair of simultaneously diagonal density operators

$$W = \begin{pmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & q_3 \end{pmatrix}, \qquad W' = \begin{pmatrix} q'_1 & 0 & 0 \\ 0 & q'_2 & 0 \\ 0 & 0 & q'_3 \end{pmatrix}. \tag{16}$$

the definition chosen for $d(W, W')$ yields a continuous and differentiable function of the parameters q_i, q'_i. This function is necessarily of the form

$$d(q_i, q'_i) = f(| q_1 - q'_1 |, | q_2 - q'_2 |, | q_3 - q'_3 |), \tag{17}$$

where $f(x, y, z)$ is a positive differentiable function of its positive arguments x, y, z such that

$$f(0, 0, 0) = 0. \tag{18}$$

Since the ordering of the basis vectors has no intrinsic significance, f should be symmetric under interchange of its arguments: for instance,

$$f(y, x, z) \equiv f(x, y, z). \tag{19}$$

Let us for simplicity restrict our investigation to the one-parameter subset (15) (which the correct answer (5) belongs to). Due to the inequalities (15-b), we find

$$d(\mathrm{W}, \mathrm{W}_0) = f(q - p_1, 1 - q - p_2, p_3), \tag{20}$$

which we have to minimize with respect to q:

$$\frac{\partial f}{\partial x}(q - p_1, 1 - q - p_2, p_3) - \frac{\partial f}{\partial y}(q - p_1, 1 - q - p_2, p_3) = 0. \tag{21}$$

There are now two possible situations. In the first one, the function f is such that

$$\frac{\partial f}{\partial x}(x, y, z) \equiv \frac{\partial f}{\partial y}(x, y, z). \tag{22}$$

This is the case for the JMG and Gudder distances, because the corresponding functions f are linear, so that their first derivatives are constants. It is also the case when f is of the form

$$f(x, y, z) \equiv \varphi(x + y + z), \tag{23}$$

where $\varphi(t)$ is a function of a single variable t. In such situations, the conclusions are analogous to those of § 3-b: all density operators W belonging to the set (15) lie at the same distance from W_0, so that the new postulate cannot single out W_1 among them.

But there exists another class of possible f functions for which (21) is not an identity in q. This is for instance the case if

$$f(x, y, z) \equiv a(x^n + y^n + z^n), \tag{24}$$

with a a positive constant and $n > 1$. In such situations, the first derivatives of f satisfy the following identity, deduced from (19):

$$\frac{\partial f}{\partial y}(y, x, z) \equiv \frac{\partial f}{\partial x}(x, y, z). \tag{25}$$

Consequently, equation (21) writes

$$f'(q - p_1, 1 - q - p_2, p_3) = f'(1 - q - p_2, q - p_1, p_3), \tag{26}$$

f' being defined as

$$f'(x, y, z) \equiv \frac{\partial f}{\partial x}(x, y, z). \tag{27}$$

It is clear that equation (26) always admits the solution

$$q - p_1 = 1 - q - p_2, \tag{28}$$

i. e.

$$q = \frac{1}{2}(1 + p_1 - p_2). \tag{29}$$

This solution is in general different from the correct one (It coincides with it only if $p_1 = p_2$ or $p_3 = 0$). Whether or not equation (26) admits other solutions (It does not in the simplest cases, because the arguments are all positive), the prescription has to be rejected in the present case too.

4. THE CASE OF INDIRECT INFORMATION

The postulate proposed in ref. 1 leads to the same kind of difficulties also in the case of « subjective states » and « indirect information ».

Consider for instance a system of two spins 1/2. Suppose this system has undergone a definite preparation, about which we have nevertheless no information at all. If, despite this, we want to describe the system by a density operator, we have no other choice than

$$W_0 = \frac{1}{4}[|++\rangle\langle++| + |+-\rangle\langle+-| + |-+\rangle\langle-+| + |--\rangle\langle--|] \tag{30}$$

($|+-\rangle$, for instance, is the pure state in which the z-components of the two spins are $+\hbar/2$ and $-\hbar/2$ respectively). We know that the objective state of the system is probably different, but W_0 incorporates the information we have in the most unbiassed way.

In order to improve this « first guess », it is sometimes possible to perform measurements on a collection of other systems identical with the one we want to study, and prepared in the same way; the « indirect information » we eventually get from such measurements will then be used, together with the initial « subjective state » W_0, to deduce a « better » density operator W_1. Imagine for instance (This is surely an ideal situation) that a measurement of the z-component of the first spin on this collection of identical systems always yields $+\hbar/2$. Knowing this result and the fact that all four possible pure states are equally probable in W_0, we will write

$$W_1 = \frac{1}{2}[|++\rangle\langle++| + |+-\rangle\langle+-|]. \tag{31}$$

Now the postulate of ref. 1, when applied to this very simple situation, does not single out the result (31). Even if one considers only the diagonal density operators of the form

$$W = q|++\rangle\langle++| + (1-q)|+-\rangle\langle+-|, \tag{32}$$

those for which

$$\frac{1}{4} \leqslant q \leqslant \frac{3}{4} \tag{33}$$

all lie at the same minimal distance from W_0 (if one adopts either the JMG distance or the Gudder distance).

5. CONCLUSION

The postulate proposed in ref. 1, and based on the consideration of a distance between quantum states, was tested by applying it to some simple problems of relative statistical inference, the new information being provided either by a direct measurement, or by an indirect one. The results are in all cases unsatisfactory: depending on the definition chosen for the distance, one finds either a whole family of possible answers which the postulate does not distinguish from one another, or a unique but incorrect answer.

The notion of distance between states cannot then be accepted as the basis for the solution of the statistical inference problem in quantum mechanics.

REFERENCES

[1] N. HADJISAVVAS, *Ann. I. H. P.*, t. **A 35**, n° 4, 1982, p. 287, and Thèse, Université de Reims, octobre 1981.
[2] B. DIU, *Ann. I. H. P.*, t. **A 37**, n° 1, 1982, p. 59.
[3] J. VON NEUMANN, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955.
[4] J. M. JAUCH, B. MISRA and A. G. GIBSON, *Helv. Phys. Acta*, t. **41**, 1968, p. 513.
[5] R. SCHATTEN, *Theory of cross-norm spaces*, Princeton University Press, 1950.
[6] S. GUDDER, *Comm. Math. Phys.*, t. **29**, 1973, p. 249.

(Manuscrit reçu le 11 mars 1982)