Underlining some limitations of the statistical formalism in quantum mechanics

F Fratini\textsuperscript{1} and A G Hayrapetyan\textsuperscript{2}

\textsuperscript{1}Ruprecht-Karls-Universität, 69120 Heidelberg, Germany
\textsuperscript{2}Max-Planck-Institut für Kernphysik, Postfach 103980, 69029 Heidelberg, Germany

E-mail: fratini@physi.uni-heidelberg.de

Received 20 July 2011
Accepted for publication 27 July 2011
Published 31 August 2011
Online at stacks.iop.org/PhysScr/84/035008

Abstract
We show that two chosen ensembles of spin states, which are differently prepared but are described by the same density matrix in quantum mechanics, do not fully share the same measurable characteristics. One characteristic in which they differ is shown to be the variance of the spin along a given direction. We conclude that the statistical description of an ensemble of states as given by its density matrix, although sufficient in many cases, should be considered incomplete, as it does not fully describe the measurable characteristics of the ensemble. A discussion \textit{a posteriori} on the problem is provided.

PACS numbers: 03.65.−w, 03.65.Ca, 05.30.−d, 05.30.Ch

1. Introduction

Density matrices are naturally used in quantum mechanics for describing physical systems whose quantum states are not equal and to which a single wavefunction cannot be therefore assigned. These systems are said to be in a mixed state and are described, within the density matrix formalism, by means of an incoherent superposition of pure states. By definition, incoherent superposition means that the probability of getting a certain experimental result, as a consequence of a measurement on the ensemble, is found by taking the average of the results that are obtained for the pure states the ensemble is made of, where the weights are the population fractions (or numbers) of the pure states. On the basis of this principle, the density matrix formalism was introduced into quantum mechanics in the late 1920s by von Neumann, Dirac and Landau, who developed and applied it to quantum information theory, statistical thermodynamics and wave mechanics [1–3].

Density matrices have been extensively applied for studying numerous physics fields. Some examples include quantum phase transitions [4], many-particle systems [5], entanglement measures [6, 7] and scattering processes [8]. Thorough reviews of the density matrix formalism and its applications can be found in [9–12].
2. Ensemble definitions and observables to investigate

The ensembles we consider are made up of non-interacting spin-1/2 particles. To some extent, spin-1/2 particles are the most representative particles in quantum mechanics, since any measurement of their spin along whichever axis can result in only two possible outcomes: +\( h/2 \) or \(-h/2\).

After having defined any system of coordinate axis, we consider two ensembles of particles defined as follows:

- ensemble \( A \): \( N \) particles whose spin states are defined along the \( \hat{x} \) axis, \( N/2 \) of which are eigenstates of the spin operator \( \hat{S}_z \) with eigenvalue \(+h/2\), while the remaining \( N/2 \) are eigenstates of the same operator with eigenvalue \(-h/2\);

- ensemble \( B \): \( N \) particles whose spin states are defined along the \( \hat{z} \) axis, \( N/2 \) of which are eigenstates of the spin operator \( \hat{S}_x \) with eigenvalue \(+h/2\), while the remaining \( N/2 \) are eigenstates of the same operator with eigenvalue \(-h/2\).

Both \( A \) and \( B \) are totally unpolarized ensembles. In the language of statistical quantum mechanics, which will be encountered in section 4, they are also said to be in a maximally mixed state.

We shall focus on deriving, for both ensembles, the expectation value and the variance of the spin along the \( \hat{x} \) direction. Both of them are measurable quantities.

We recall that the expectation value of some discrete variable \( y \) distributed according to a certain probability function \( f(y) \) is defined as [17]

\[
E(y) = \sum_i f(y_i) y_i, \tag{1}
\]

where \( i \) labels the values that the variable \( y \) is allowed to assume.

The variance of \( y \), which is a measure of how widely \( y \) is spread about its mean value, is obtained as [17]

\[
\text{Var}(y) = E((y - E(y))^2) = E(y^2) - (E(y))^2. \tag{2}
\]

Normally, the distribution function \( f(y) \) is experimentally measured by performing many measurements of the observable \( y \). However, in many cases \( f(y) \) can be theoretically derived, so that predictions for the expectation value and the variance can be provided.

If we have some operator \( \hat{O} \) and its discrete spectrum of eigenstates \( |\alpha_i\rangle \) which satisfy the equation

\[
\hat{O} |\alpha_i\rangle = \alpha_i |\alpha_i\rangle, \tag{3}
\]

in quantum mechanics it is postulated that the expectation value of the variable \( o \) over a certain state \( |\beta\rangle \) can be theoretically obtained by calculating [18, 19]

\[
E(o) = \langle \beta | \hat{O} | \beta \rangle = \sum_i (|\alpha_i|\beta\rangle)^2 \alpha_i, \tag{4}
\]

where the identity

\[
\sum_i |\alpha_i\rangle \langle \alpha_i| = \hat{I} \tag{5}
\]

and the normalization equation

\[
\langle \alpha_i | \alpha_j \rangle = \delta_{i,j} \tag{6}
\]

have been used in the last step of equation (4). By comparing equation (4) with equation (1), we notice that the probability function is evidently given by

\[
f(o_i) = |\langle \alpha_i|\beta\rangle|^2. \tag{7}
\]

In the next section, we shall show that although the expectation values of the spin along the \( \hat{x} \) direction coincide for ensembles \( A \) and \( B \), the variances do not.

3. Expectation value and variance for the ensembles

In this section, we derive the expectation value and the variance of the spin along the \( \hat{x} \) direction for the ensembles of states \( A \) and \( B \) defined in section 2. We will work separately on the two ensembles, without turning to the quantum statistical description. For this purpose, we can pragmatically think of using, for example, a Stern and Gerlach (SG) apparatus with the magnetic field along the \( \hat{x} \) direction [20]. The SG apparatus would measure the spin of each single particle of the ensemble, so that the total spin of the ensemble would then be obtained as the sum of all the single-particle spin measurements.

Since this paper is theoretically oriented, we will consider the SG apparatus as ideal, without any experimental limitation. We will furthermore denote by \( |S_x, \pm 1\rangle \) and \( |S_x, \pm 1\rangle \) the eigenstates of the operators \( \hat{S}_x \) and \( \hat{S}_z \), respectively. The corresponding eigenvalues are \( \pm h/2 \) in both cases.

3.1. Ensemble \( A \)

Because of the characteristics chosen for ensemble \( A \), the SG apparatus would exactly separate particle flux into two equal parts, or, what is the same, it will measure \( N/2 \) particles having spin along the \( +\hat{x} \) direction and \( N/2 \) particles having spin along the \(-\hat{x} \) direction. The probability of registering the outcome \( \pm h/2 \), when the spin along \( \hat{x} \) is measured on state \( |S_x, \pm 1\rangle \), is in fact 1.

Consequently, the total spin of ensemble \( A \), as measured along the \( \hat{x} \) direction, will always be 0 in any measurement that is carried out on the ensemble, independent of the number of particles contained in \( A \). Since, as already recalled in section 2, the variance is a measure of how widely the single


measurements are spread about the mean value, the variance for the spin measurement on ensemble \( A \) also vanishes.

We can summarize what was previously stated by writing

\[
E(S_x)_A = 0 \\
\text{Var}(S_x)_A = 0 \implies S^A_x = 0 \pm 0, \tag{8}
\]

where subscript \( A \) denotes the ensemble on which the expectation value and the variance are evaluated. \( S^A_x \) represents our theoretical prediction for the measurement of the spin along the \( \hat{x} \) direction on ensemble \( A \).

### 3.2. Ensemble \( B \)

Differently from ensemble \( A \), each particle of ensemble \( B \), independent of one the other, has one-half probability of being registered with spin along the \( +\hat{x} \) direction and one-half probability of being registered with spin along the \( -\hat{x} \) direction. This obvious statement can be easily proved by writing the spin states that are defined along the \( \hat{x} \) direction as [19]

\[
|S_x, \pm 1\rangle = \frac{1}{\sqrt{2}} (|S_x, +1\rangle \pm |S_x, -1\rangle), \tag{9}
\]

and by noticing that the probability of registering the outcome \( \pm h/2 \), in a measurement of the spin along the \( \hat{x} \) axis of the state in equation (9), can be obtained by using equation (7):

\[
|\langle S_x, \pm 1|S_x, \pm 1\rangle|^2 = \frac{1}{2}. \tag{10}
\]

In light of the above, the measurement outcomes from the SG apparatus must follow a binomial distribution. This means that \( N/2 \) particles are expected to be measured with spin along the \( +\hat{x} \) direction and \( N/2 \) particles with spin along the \( -\hat{x} \) direction. Being a binomial distribution, the variance of these expected particle numbers can be easily calculated as \( N/4 \) [17]. If we adopt the standard deviation (\( \sigma = \sqrt{\text{Var}} \) as indeterminacy of the measurement, we may write that the number of particles measured with spin equal to \( +h/2 \) (as well as the number of particles measured with spin equal to \( -h/2 \)) must

\[
\frac{N}{2} \pm \frac{\sqrt{N}}{2}. \tag{11}
\]

Consequently, the total spin of ensemble \( B \) along the \( \hat{x} \) direction, as measured by the SG apparatus, must be

\[
\hat{S}_x^B = \frac{\hbar}{2} \left( 0 \pm \frac{\sqrt{N}}{2} \right) = 0 \pm \frac{\hbar\sqrt{N}}{2}. \tag{12}
\]

The factor ‘2’ that multiplies the standard deviation in equation (12) arises from the fact that ‘not detecting a particle with spin \( +h/2 \)’ results in ‘detecting that particle with spin \( -h/2 \)’. Summarizing what was obtained for ensemble \( B \),

\[
E(S_x)_B = 0 \\
\text{Var}(S_x)_B = \frac{\hbar^2N}{4} \implies S^B_x = 0 \pm \frac{\hbar\sqrt{N}}{2}, \tag{13}
\]

where \( S^B_x \) represents our theoretical prediction for measurement of the spin along the \( \hat{x} \) direction on ensemble \( B \).

We can certainly conclude that the two ensembles \( A \) and \( B \) are not equal, as they show a measurable difference, which is the variance of the spin along the \( \hat{x} \) direction. The latter can be naturally measured in experiments by extracting the width of the measured spin distribution.

Since only a few solid principles of quantum mechanics and of statistics have been used up to now, the above predictions are considered reliable: any measurement is reasonably expected to agree with them.

### 4. Quantum statistical description

The overall state of an admixture of particles with different states is formally described in statistical quantum mechanics by means of the so-called ‘density operator’. The normalized density operator for an ensemble of states reads [11, 12, 19]

\[
\hat{\rho} = \sum_i W_i |a_i\rangle \langle a_i|, \tag{14}
\]

where \( W_i \) is the statistical weight of the state \( |a_i\rangle \) or, more practically, the fraction of particles of the ensemble that share the same pure state \( |a_i\rangle \). If \( W_i = \delta_{ij} \), then the ensemble is said to be in the pure state \( |a_j\rangle \). In any other case, the ensemble is said to be in a mixed state. If \( W_i = W_0 \), where \( W_0 \) is constant for all \( i \), the ensemble is said to be in a maximally mixed state.

The density matrix that describes the ensemble, in a given representation \( |b_i\rangle \), is obtained as

\[
\rho_{ij} = \langle b_j|\hat{\rho}|b_i\rangle. \tag{15}
\]

Since it does not make any substantial difference between describing an ensemble of states by means of its density matrix or by means of its density operator, we choose to deal, in the following, with density operators.

The density operator for ensemble \( A \) can be explicitly written by applying the definition given in section 2 to equation (14):

\[
\hat{P}_A = 0.5 |S_x, +1\rangle \langle S_x, +1| + 0.5 |S_x, -1\rangle \langle S_x, -1|. \tag{16}
\]

Similarly, the density operator for ensemble \( B \) reads

\[
\hat{P}_B = 0.5 |S_x, +1\rangle \langle S_x, +1| + 0.5 |S_x, -1\rangle \langle S_x, -1| \\
= 0.5 |S_x, +1\rangle \langle S_x, +1| + 0.5 |S_x, -1\rangle \langle S_x, -1|. \tag{17}
\]

where, in the last step, we made use of equation (9), which links spin states defined along the \( \hat{x} \) axis with spin states defined along the \( \hat{z} \) axis [19].

By comparing equation (16) with equation (17), we notice that the density operators related to ensembles \( A \) and \( B \) are analytically the same. Since the density operator is supposed to contain the whole information of the ensemble [11, 12], we should conclude that \( A \) and \( B \) show exactly the same polarization features, i.e. they are identical from the point of view of any polarization measurement. However, such a conclusion leads to a contradiction. As mentioned in section 3, ensembles \( A \) and \( B \) show at least one measurable difference, which is the variance of the spin along the \( \hat{x} \) direction. It can be easily understood that the reason for this difference lies in the different physical characteristics that
have been chosen \textit{a priori} to characterize the two ensembles, i.e. on information concerning preparation of the ensembles.

To complete the analysis, we derive, similarly to equations (8) and (13), the expectation value and the variance that the quantum statistical formalism provides for the ensembles. For this purpose, we need to first recall that, in such a formalism, the expectation value of some variable \( o \) is given by [11, 12, 19]

\[
E(o) = \text{Tr}[\hat{P} \hat{O}],
\]

(18)

where \( \hat{O} \) is the quantum mechanical operator related to \( o \) as in equation (3). \( \text{Tr} \) denotes the trace over any set of quantum states and \( \hat{P} \) is the density operator that describes the ensemble on which the variable \( o \) is measured.

By combining equations (2) and (16)–(18), we obtain

\[
\begin{align*}
E(S_x)_A &= E(S_x)_B = 0 \\
\text{Var}(S_x)_A &= \text{Var}(S_x)_B = \frac{\hbar^2}{4} \\
\Rightarrow S^4_x &= S^6_x = 0 \pm \frac{\hbar}{2}.
\end{align*}
\]

These results are in stark contrast with equations (8) and (13).

We now note that the definition of the normalized density operator in equation (14) does not account for the number of particles that the ensemble contains. By virtue of this, any measurable quantity that depends on such a number, like the variance of the spin measurement over the ensemble along whichever axis, cannot be adequately described by using the normalized density operator. Consequently, we are led to think that the use of the \textit{not normalized} density operator for the two ensembles, which reads [10]

\[
\hat{P}_A = \hat{P}_B = \frac{N}{2}(|S_z, +1\rangle\langle S_z, +1| + |S_z, -1\rangle\langle S_z, -1|),
\]

(20)

might benefit the calculation. Indeed, by using equations (18) and (20), the results we obtain in this case are

\[
\begin{align*}
E(S_x)_A &= E(S_x)_B = 0 \\
\text{Var}(S_x)_A &= \text{Var}(S_x)_B = \frac{\hbar^2}{4} \\
\Rightarrow S^4_x &= S^6_x = 0 \pm \frac{\hbar\sqrt{N}}{2},
\end{align*}
\]

which are correct for ensemble \( B \), as they match equation (13), but still not correct for ensemble \( A \), as they do not match equation (8).

Since the density operators for ensembles \( A \) and \( B \) are the same while the correct results for the variance of the spin along the \( \hat{x} \) direction are not, it is evidently not possible to recover the right description for both ensembles within the quantum statistical formalism.

5. Discussion

To analyze the discrepancy that has been found in the previous sections, we first start by recalling how the density operator is normally introduced into quantum mechanics.

The leading idea in the formulation of the density operator is the following. Given an ensemble \( \chi \) made up of \( N \) sub-ensembles, where each one of the latter is composed of particles that share the same quantum state \( |\beta_i\rangle \) (\( i = 1, \ldots, N \)), the expectation value of a certain variable \( o \) over the ensemble must reasonably be a statistical average of the expectation values of the same variable over the sub-ensembles [10]. The statistical weight of each addend will naturally be the particle fraction (for intensive variables) or the particle number (for extensive variables) of the sub-ensemble. For any observable \( o \), this supposition implies

\[
E(o)_\chi = \sum_{i=1}^{N} W_i \langle \beta_i | \hat{O} | \beta_i \rangle
\]

(22)

where \( \hat{O} \) is related to \( o \) as in equation (3), \( W_i \) is the fraction or the number of particles of the \( i \)th sub-ensemble and \( |\beta_i\rangle \) are states that form a complete basis in the quantum space where \( |\beta_i\rangle \) are defined. The relation \( \sum_i |a_j\rangle \langle a_j| = \hat{1} \) has been used in the last step.

Now, from equation (22), if \( \hat{O} \) does not depend on the sub-ensembles’ states \( |\beta_i\rangle \), we clearly see that what depends on the characteristics of the ensemble is enclosed in parentheses. In this case, such a quantity can be assigned to be representative for the ensemble, similarly to the ket vector for a particle state, and is therefore assigned the name ‘density operator’. As a consequence of this assignment, equation (22) becomes nothing but the explicit form of equation (18).

In order to apply equation (22) to the variance and, consequently, to obtain the variance over an ensemble of states as an expectation value, we should first find an operator that correctly represents the variance in quantum mechanics. For this purpose, we notice that the correct value for the variance of the spin along the \( \hat{x} \) direction, over an ensemble of \( N \) particles whose states are \( |\beta\rangle \), may be obtained as

\[
\text{Var}(S_x)_\beta = N \langle |\beta\rangle \hat{O}_\xi |\beta\rangle = N \langle \beta | \hat{O}_\xi |\beta\rangle.
\]

(23)

where the operator \( \hat{O}_\xi \) reads

\[
\hat{O}_\xi = (\hat{S}_x - E(S_x)_\xi)^2 = \frac{\hbar^2}{4} + (E(S_x)_\xi)^2 - 2\hat{S}_x E(S_x)_\xi.
\]

(24)

The operator \( \hat{S}_z \), as is evident from equation (23), is defined such that, by acting on any state \( |\beta\rangle \), it becomes the index which identifies the state. The evaluation of the expectation value contained in the definition of \( \hat{O}_\xi \) is then performed over that state.

We could therefore attempt to associate, in quantum mechanics, the operator \( \hat{O}_\xi \) to the variance of the spin along the \( \hat{x} \) direction. However, the definition of the operator \( \hat{O}_\xi \) as given in equation (24), leads to contradictions. For instance, it follows from equation (24) that the eigenstates of the operator
would be only the two states $|S_z, \pm 1\rangle$, for which the corresponding eigenvalues would be vanishing. Because of this, the operator $\hat{O}_z$ would then be equivalent to the null operator in quantum mechanics, although the expectation value on $\hat{O}_z$ would not be vanishing over some states, such as $|S_z, \pm 1\rangle$ or $|S_z, \pm 1\rangle$. This is certainly nonsense!

The problem may be identified in the definition of the operator $\hat{O}_z$. The definition of the operator $\hat{O}_z$, as given in equation (24), is certainly atypical as it does not allow the action of the operator $\hat{O}_z$ on a given ket state $|\beta\rangle$ to be disentangled from the statistical evaluation of the variable $S_z$ over an ensemble of states $|\beta\rangle$.

In short, we can say that it does not seem possible to assign a well-defined quantum operator to the variance, and consequently equation (22) together with the definition of the density operator cannot be used to predict the variance for an ensemble of states. This is the reason for the wrong prediction provided by the statistical formalism for the variance of the spin along the $\hat{x}$ direction in section 4.

The above discussion strengthens the fact that the description of ensembles of states as given by their density operators should be considered incomplete, as it does not allow the description of some statistical measurable quantities of the ensembles, like variance.

To conclude, we briefly clarify some issues related to unpolarized ensembles. Since the two ensembles considered in this paper are both unpolarized and have been shown to behave differently in experiments, our reasoning raises the question of which ensemble to theoretically consider when unpolarized particles or systems are the object of experiments. However, as explained above, the expectation value of physical observables, for which a quantum operator can be safely assigned, does not change when the unpolarized ensemble is changed, but rather it is equally well described by any unpolarized ensemble. Since such quantities represent what is normally aimed to be measured in experiments, the problem of choosing the right unpolarized ensemble of states is without foundation. In contrast, when a full description of the ensemble is required, including observables such as the variance, then the choice of the unpolarized ensemble to theoretically consider should be constrained by information concerning preparation of the ensemble. When the experimental preparation of the ensemble is not under control, the unpolarized ensemble to theoretically consider should be averaged over the possible representations, or, what is effectively the same, the phase that determines the used representation should be randomly defined, as pointed out by Tolman in the late 1930s [21].

In forthcoming papers, we will further investigate how the theoretical and experimental results depend on preparation of the ensemble.

6. Summary

In summary, we have shown that ensembles of states that share the same density operator in quantum mechanics, i.e. the same density matrix, can behave differently in experiments. In order to prove this statement, we started out by defining two ensembles of spin states and by showing, without using any quantum statistical mean, that these two ensembles are characterized by a measurable difference, which is the variance of the spin along a given direction. Then, we analyzed the two ensembles within the statistical formalism of quantum mechanics. Since, in this formalism, the two ensembles turn out to be identically described, the variance of the spin is found to be equal for both. The contradiction is then solved by showing that the statistical formalism of quantum mechanics cannot be applied at all to evaluate the variance of the spin of ensembles. We therefore concluded that the description of ensembles given by their density operators (or by their density matrices) should be considered incomplete, as it cannot be applied to predict some measurable statistical quantities and as it lacks measurable information given by the preparation of the ensembles.

References

[1] von Neumann J 1927 Nachr. Akad. Wiss. Göttingen Math.-physik. Kl. I 245
[2] Dirac P A M 1929 Proc. Camb. Phil. Soc. 25 62
[3] Landau L D 1927 Z. Phys. 45 430
[4] Wu L A, Sarandy M S and Lidar D A 2004 Phys. Rev. Lett. 93 250404
[5] Löwdin P O 1955 Phys. Rev. 97 1474
[6] Wootters W K 1998 Phys. Rev. Lett. 80 2245
[7] Fratini F, Tichy M C, Jahnzetz T, Buchleitner A, Fritzche S and Surzhykov A 2011 Phys. Rev. A 83 032506
[8] Dombey O 1969 Rev. Mod. Phys. 41 236
[9] Fano U 1957 Rev. Mod. Phys. 29 74
[10] McWeeny R 1960 Rev. Mod. Phys. 32 335
[11] Blum K 1996 Density Matrix Theory and Applications (New York: Plenum)
[12] Balashov V V, Grum-Grzhimailo A N and Kabachnik N M 2000 Polarization and Correlation Phenomena in Atomic Collisions (New York: Kluwer)
[13]Bloch F 1932 Z. Phys. 74 295
[14] Avakyan R M, Hayrapetyan A G, Khachatryan B V and Petrosyan R G 2007 Phys. Lett. A 372 77
[15] Glauber R J 1963 Phys. Rev. Lett. 10 84
[16] Glauber R J 1963 Phys. Rev. 131 2766
[17] Cowan G 1998 Statistical Data Analysis (Oxford: Clarendon)
[18] Dirac P A M 1982 The Principles of Quantum Mechanics (Oxford: University Press)
[19] Sakurai J J 1994 Modern Quantum Mechanics (Reading, MA: Addison-Wesley)
[20] Gerlach W and Stern O 1922 Z. Phys. 9 353
[21] Tolman R C 1938 The Principles of Statistical Mechanics (Oxford: Clarendon) p 344