

Interpretations of quantum mechanics, and interpretations of violation of Bell's inequality

Willem M. de Muynck
Theoretical Physics,
Eindhoven University of Technology,
Eindhoven, the Netherlands

January 14, 2001

Abstract

The discussion of the foundations of quantum mechanics is complicated by the fact that a number of different issues are closely entangled. Three of these issues are i) the interpretation of probability, ii) the choice between realist and empiricist interpretations of the mathematical formalism of quantum mechanics, iii) the distinction between measurement and preparation. It will be demonstrated that an interpretation of violation of Bell's inequality by quantum mechanics as evidence of non-locality of the quantum world is a consequence of a particular choice between these alternatives. Also a distinction must be drawn between two forms of realism, viz. a) realist interpretations of quantum mechanics, b) the possibility of hidden-variables (sub-quantum) theories.

1 Realist and empiricist interpretations of quantum mechanics

In *realist* interpretations of the mathematical formalism of quantum mechanics state vector and observable are thought to refer to the microscopic object in the usual way presented in most textbooks. Although, of course, preparing and measuring instruments are often present, these are not taken into account in the mathematical description (unless, as in the theory of measurement, the subject is the interaction between object and measuring instrument). In figure 1 this is symbolized by dotted lines, indicating that these objects are not explicitly represented in the mathematical formalism.

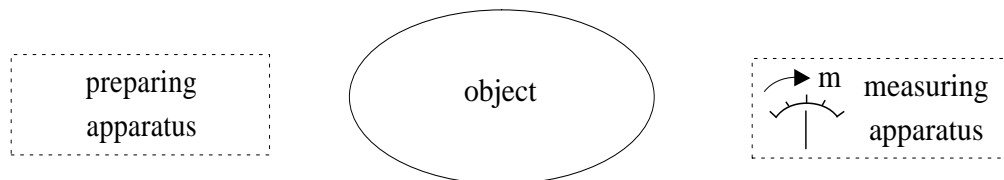


Figure 1: Realist interpretation of quantum mechanics.

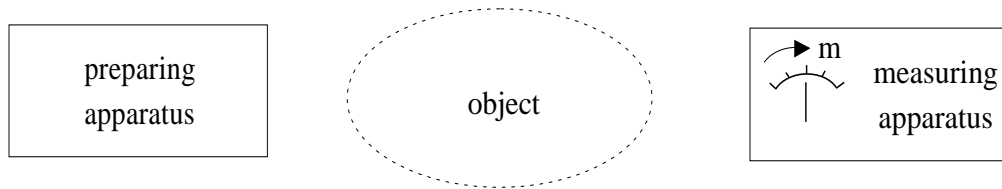


Figure 2: Empiricist interpretation of quantum mechanics.

In an *empiricist* interpretation quantum mechanics is thought to describe relations between input and output of a measurement process. A state vector is just a label of a preparation procedure; an observable is a label of a measuring instrument. In an empiricist interpretation (cf. figure 2) quantum mechanics is not thought to describe the microscopic object. This, of course, does not imply that this object would not exist; it only means that it is not described by quantum mechanics. Explanation of relations between input and output of a measurement process should be provided by another theory, e.g. a hidden-variables (sub-quantum) theory. This is analogous to the way the theory of rigid bodies describes the empirical behavior of a billiard ball, or to the description by thermodynamics of the thermodynamic properties of a volume of gas, explanations being relegated to theories describing the microscopic (atomic) properties of the systems.

Although a term like ‘observable’ (rather than ‘physical quantity’) is evidence of the empiricist origin of quantum mechanics (compare Heisenberg [1]), there has always existed a strong tendency toward a realist interpretation in which observables are considered as properties of the microscopic object, more or less analogous to classical ones. Likewise, many physicists use to think about electrons as wave packets flying around in space, without bothering too much about the “Unanschaulichkeit” that for Schrödinger [2] was such a problematic feature of quantum theory. Without entering into a detailed discussion of the relative merits of either of these interpretations (e.g. de Muynck [3]) it is noted here that an empiricist interpretation is in agreement with the operational way theory and experiment are compared in the laboratory. Moreover, it is free of the paradoxes, which have their origin in a realist interpretation. As will be seen in the next section, the difference between realist and empiricist interpretations is highly relevant when dealing with the EPR problem.

2 EPR experiments and Bell experiments

In figure 3 the experiment is depicted, proposed by Einstein, Podolsky and Rosen [4] to study (in)completeness of quantum mechanics. A pair of particles (1 and 2) is prepared in an entangled state and allowed to separate. A measurement is performed on particle 1. It is essential to the EPR reasoning that particle 2 does *not* interact with any measuring instrument, thus allowing to consider so-called ‘elements of physical reality’ of this particle, that can be considered as *objective* properties, being attributable to

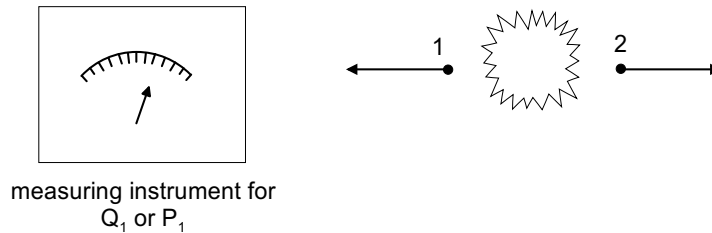


Figure 3: EPR experiment.

particle 2 independently of what happens to particle 1. By EPR this arrangement was presented as a way to perform a measurement on particle 2 *without in any way disturbing this particle*.

The EPR experiment should be compared to correlation measurements of the type performed by Aspect et al. [5, 6] to test Bell's inequality (cf. figure 4). In these latter experiments also particle 2 is interacting with a measuring instrument. In the literature these experiments are often referred to as EPR experiments, too, thus neglecting the fundamental difference between the two measurement arrangements of figures 3 and 4. This negligence has been responsible for quite a bit of confusion, and should preferably be avoided by referring to the latter experiments as Bell experiments rather than EPR ones. In EPR experiments particle 2 is not subject to a *measurement*, but to a (conditional) *preparation* (conditional on the measurement result obtained for particle 1). This is especially clear in an empiricist interpretation, because here measurement results cannot exist unless a measuring instrument is present, the pointer positions of which corresponding to the measurement results.

Unfortunately, the EPR experiment of figure 3 was presented by EPR as a *measurement* performed on particle 2, and accepted by Bohr as such. That this could happen is a consequence of the fact that both Einstein and Bohr entertained a *realist* interpretation of quantum mechanical observables (note that they differed with respect to the interpretation of the state vector), the only difference being that Einstein's realist interpretation was an *objectivistic* one (in which observables are considered as properties of the object, possessed independently of any measurement: the EPR 'elements of physical reality'), whereas Bohr's was a *contextualistic* realism (in which observables are only well-defined within the context of the measurement). Note that in Bell experiments the EPR reasoning would break down because, due to the interaction of particle 2 with its measuring instrument, there cannot exist 'elements of physical reality'.

Much confusion could have been avoided if Bohr had maintained his *interactional*

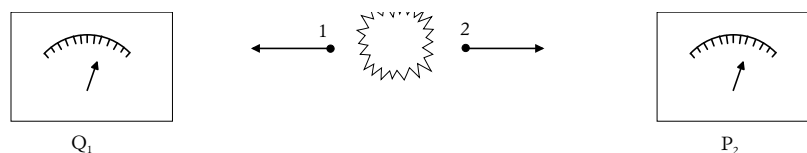


Figure 4: Bell experiment.

view of measurement. However, by accepting the EPR experiment as a measurement of particle 2 he had to weaken his interpretation to a *relational* one (e.g. [7, 8]), allowing the observable of particle 2 to be (co-)determined by the measurement context for particle 1. This introduced for the first time non-locality in the interpretation of quantum mechanics. But this could easily have been avoided if Bohr had required that for a measurement of particle 2 a measuring instrument should be actually interacting with this very particle, with the result that an observable of particle n ($n = 1, 2$) can be (co-)determined *in a local way* by the measurement context of that particle only. This, incidentally, would have completely made obsolete the EPR ‘elements of physical reality’, and would have been quite a bit less confusing than the answer Bohr [9] actually gave (to the effect that the definition of the EPR ‘element of physical reality’ would be ambiguous because of the fact that it did not take into account the measurement arrangement for the other particle), thus promoting the non-locality idea.

Summarizing, the idea of EPR non-locality is a consequence of i) a neglect of the difference between EPR and Bell experiments (equating ‘elements of physical reality’ to measurement results), ii) a realist interpretation of quantum mechanics (considering measurement results as properties of the microscopic object, i.c. particle 2). In an *empiricist* interpretation there is no reason to assume any non-locality.

It is often asserted that non-locality is proven by the Aspect experiments, because these are violating Bell’s inequality. The reason for such an assertion is that it is thought that non-locality is a necessary condition for a derivation of Bell’s inequality. However, as will be demonstrated in the following, this cannot be correct since this inequality can be derived from quite different assumptions. Also, experiments like the Aspect ones, -although violating Bell’s inequality,- do not exhibit any trace of non-locality, because their measurement results are completely consistent with the postulate of local commutativity, implying that relative frequencies of measurement results are independent of which measurements are performed in causally disconnected regions. Admittedly, this does not logically exclude a certain non-locality at the individual level, being unobservable at the statistical level of quantum mechanical probability distributions. However, from a physical point of view a peaceful coexistence between locality at the (physically relevant) statistical level and non-locality at the individual level is extremely implausible. Unobservability of the latter would require a kind of conspiracy not unlike the one making unobservable 19th century world aether. For this reason the ‘non-locality’ explanation of the experimental violation of Bell’s inequality does not seem to be very plausible, and does it seem wise to look for alternative explanations.

Since non-locality is never the *only* assumption in deriving Bell’s inequality, such alternative explanations do exist. Thus, Einstein’s assumption of the existence of ‘elements of physical reality’ is such an additional assumption. More generally, in Bell’s derivation [10] the existence of hidden-variables is one. Is it still possible to derive Bell’s inequality if these assumptions are abolished? Moreover, even assuming the possibility of hidden-variables theories, are there in Bell’s derivation no hidden assumptions, additional to the locality assumption.

Bell’s inequality refers to a set of four quantum mechanical observables, A_1, B_1, A_2 and B_2 , observables with different/identical indices being compatible/incompatible. In

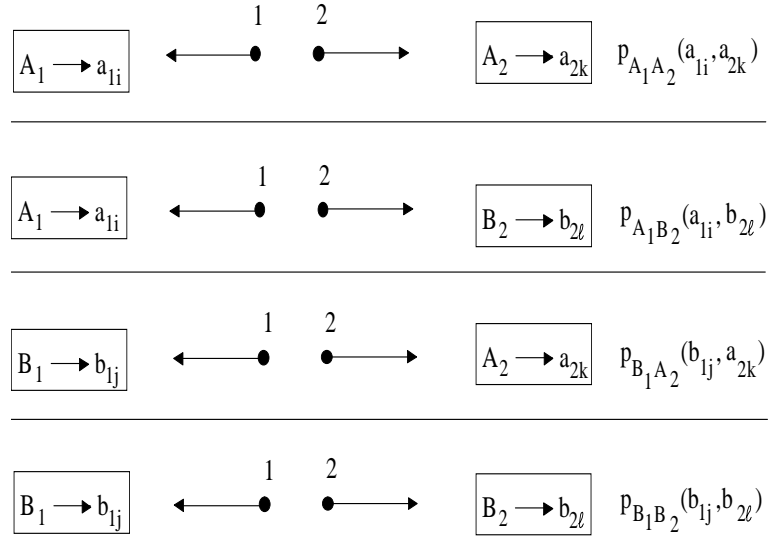


Figure 5: The four Bell experiments conventionally involved in the Bell inequality.

the Aspect experiments measurements of the four possible compatible pairs are performed (cf. figure 5; in these experiments A_n and B_n refer to polarization observables of photon n , $n = 1, 2$, respectively). Bell's inequality can typically be derived for the stochastic quantities of a classical Kolmogorovian probability theory. Hence, violation of Bell's inequality is an indication that observables A_1, B_1, A_2 and B_2 are not stochastic quantities in the sense of Kolmogorov's probability theory. In particular, there cannot exist a quadrivariate joint probability distribution of these four observables. Such a non-existence is a consequence of the *incompatibility* of certain of the observables. Since incompatibility is a *local* affair, this is another reason to doubt the 'non-locality' explanation of the violation of Bell's inequality.

In the following derivations of Bell's inequality will be scrutinized to see whether the non-locality assumption is as crucial as was assumed by Bell. In doing so it is necessary to distinguish derivations in quantum mechanics from derivations in hidden-variables theories.

3 Bell's inequality in quantum mechanics

For dichotomic observables, having values ± 1 , Bell's inequality is given according to

$$|\langle A_1 A_2 \rangle - \langle A_1 B_2 \rangle| - \langle B_1 B_2 \rangle - \langle B_1 A_2 \rangle \leq 2. \quad (1)$$

A more general inequality, being valid for arbitrary values of the observables, is the BCHS inequality

$$-1 \leq p(b_1, a_2) + p(b_1, b_2) + p(a_1, b_2) - p(a_1, a_2) - p(b_1) - p(b_2) \leq 0 \quad (2)$$

from which (1) can be derived for the dichotomic case. Because of its independence of the values of the observables inequality (2) is by far preferable to inequality (1). Bell's inequality may be violated if some of the observables are *incompatible*: $[A_1, B_1]_- \neq O$, $[A_2, B_2]_- \neq O$.

I shall now discuss two derivations of Bell's inequality, which can be formulated within the quantum mechanical formalism, and which do not rely on the existence of hidden variables. The first one is relying on a 'possessed values' principle, stating that

$$\left\{ \begin{array}{l} \text{'possessed} \\ \text{values'} \\ \text{principle} \end{array} \right. = \left\{ \begin{array}{l} \text{values of quantum mechanical observables} \\ \text{may be attributed to the object as} \\ \text{objective properties, possessed by the object} \\ \text{independent of observation} \end{array} \right.$$

The 'possessed values' principle can be seen as an expression of the objectivistic-realist interpretation of the quantum mechanical formalism preferred by Einstein (compare the EPR 'elements of physical reality'). The important point is that by this principle well-defined values are simultaneously attributed to *incompatible* observables. If $a_i^{(n)}$, $b_j^{(n)} = \pm 1$, are the values of A_i and B_j for the n^{th} of a sequence of N particle pairs, then we have

$$-2 \leq a_1^{(n)} a_2^{(n)} - a_1^{(n)} b_2^{(n)} - b_1^{(n)} b_2^{(n)} - b_1^{(n)} a_2^{(n)} \leq 2,$$

from which it directly follows that the quantities

$$\langle A_1 A_2 \rangle = \frac{1}{N} \sum_{n=1}^N a_1^{(n)} a_2^{(n)}, \text{ etc.}$$

must satisfy Bell's inequality (1) (a similar derivation has first been given by Stapp [11], although starting from quite a different interpretation). The essential point in the derivation is the assumption of the existence of a quadruple of values (a_1, b_1, a_2, b_2) for each of the particle pairs.

From the experimental violation of Bell's inequality it follows that an objectivistic-realist interpretation of the quantum mechanical formalism, encompassing the 'possessed values' principle, is impossible. Violation of Bell's inequality entails failure of the 'possessed values' principle (no quadruples available). In view of the important role measurement is playing in the interpretation of quantum mechanics this is hardly surprising. As is well-known, due to the incompatibility of some of the observables the existence of a quadruple of values can only be attained on the basis of a doubtful counterfactual reasoning. If a realist interpretation is feasible at all, it seems to have to be a contextualistic one, in which the values of observables are (co-)determined by the measurement arrangement. In the case of Bell experiments non-locality does not seem to be involved.

As a second possibility to derive Bell's inequality within quantum mechanics we should consider derivations of the BCHS inequality (2) from the existence of a quadri-variate probability distribution $p(a_1, b_1, a_2, b_2)$ by Fine [12] and Rastall [13] (also de Muynck [14]). Hence, from violation of Bell's inequality the non-existence of a quadri-variate joint probability distribution follows. In view of the fact that incompatible observables are involved, this, once again, is hardly surprising.

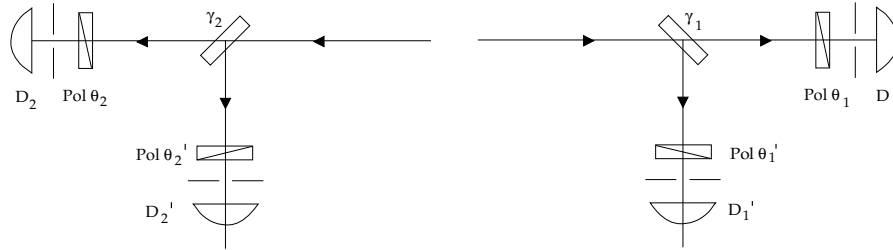


Figure 6: Generalized Aspect experiment.

A priori there are two possible reasons for the non-existence of the quadrivariate joint probability distribution $p(a_1, b_1, a_2, b_2)$. First, the limit $\lim_{N \rightarrow \infty} N(a_1, b_1, a_2, b_2)/N$ of the relative frequencies of quadruples of measurement results might not exist. Since, however, Bell's inequality already follows from the existence of relative frequency $N(a_1, b_1, a_2, b_2)/N$ with *finite* N , and the limit $N \rightarrow \infty$ is never involved in any experimental implementation, this answer does not seem to be sufficient. Therefore the reason for the non-existence of the quadrivariate joint probability distribution $p(a_1, b_1, a_2, b_2)$ can only be the non-existence of relative frequencies $N(a_1, b_1, a_2, b_2)/N$. This seems to reduce the present case to the previous one: Bell's inequality can be violated because quadruples $(A_1 = a_1, B_1 = b_1, A_2 = a_2, B_2 = b_2)$ do not exist.

Could non-locality explain the non-existence of quadruples $(A_1 = a_1, B_1 = b_1, A_2 = a_2, B_2 = b_2)$. Indeed, it could. If the value of A_1 , say, is (co-)determined by the measurement arrangement of particle 2, then non-locality could entail

$$a_1(A_2) \neq a_1(B_2), \quad (3)$$

thus preventing the existence of one single value of observable A_1 for the two Aspect experiments involving this observable. This, precisely, is the 'non-locality' explanation referred to above. This explanation is close to Bohr's 'ambiguity' answer to EPR referred to in section 2, stating that the definition of an 'element of physical reality' of observable A_1 must depend on the measurement context of particle 2.

As will be demonstrated next, there is a more plausible *local* explanation, however, based on the inequality

$$a_1(A_1) \neq a_1(B_1), \quad (4)$$

expressing that the value of A_1 , say, will depend on whether either A_1 or B_1 is measured. Inequality (4) could be seen as an implementation of Heisenberg's disturbance theory of measurement, to the effect that observables, incompatible with the actually measured one, are disturbed by the measurement. That such an effect is really occurring in the Aspect experiments, can be seen from the generalized Aspect experiment depicted in figure 6. This experiment should be compared with the Aspect switching experiment [6], in which the switches have been replaced by two semi-transparent mirrors (transmissivities γ_1 and γ_2 , respectively). The four Aspect experiments, depicted in figure 5, are special cases of the generalized one, having $\gamma_n = 0$ or 1 , $n = 1, 2$.

Restricting for a moment to one side of the interferometer, it is possible to calculate the joint detection probabilities of the two detectors:

$$(p_{\gamma_1}(a_{1i}, b_{1j})) = \begin{pmatrix} 0 & \gamma_1 \langle E^{(1)}_+ \rangle \\ (1 - \gamma_1) \langle F^{(1)}_+ \rangle & 1 - \gamma_1 \langle E^{(1)}_+ \rangle - (1 - \gamma_1) \langle F^{(1)}_+ \rangle \end{pmatrix}, \quad (5)$$

in which $\{E^{(1)}_+, E^{(1)}_-\}$ and $\{F^{(1)}_+, F^{(1)}_-\}$ are the spectral representations of the two polarization observables (A_1 and B_1) in directions θ_1 and θ'_1 , respectively. The values $a_{1i} = +/-, b_{1j} = +/-$ correspond to yes/no registration of a photon by the detector. $p_{\gamma_1}(+, +) = 0$ means that, like in the switching experiment, only one of the detectors can register photon 1. There, however, is a fundamental difference with the switching experiment, because in this latter experiment the photon wave packet is sent either toward one detector or the other, whereas in the present one it is split so as to interact coherently with both detectors. This makes it possible to interpret the right hand part of the generalized experiment of figure 6 as a joint non-ideal measurement of the incompatible polarization observables in directions θ_1 and θ'_1 (e.g. [15]), the joint probability distribution of the observables being given by (5).

It is not possible to extensively discuss here the relevance of experiments of the generalized type for understanding Heisenberg's disturbance theory of measurement, and its relation to the Heisenberg uncertainty relations (see e.g. [16]). The important point is that such experiments do not fit into the standard (Dirac-von Neumann) formalism in which a probability is an expectation value of a projection operator. Indeed, from (5) it follows that $p_{\gamma_1}(a_{1i}, b_{1j}) = Tr \rho R^{(1)}_{ij}$ is yielding operators $R^{(1)}_{ij}$ according to

$$(R^{(1)}_{ij}) = \begin{pmatrix} O & \gamma_1 E^{(1)}_+ \\ (1 - \gamma_1) F^{(1)}_+ & \gamma_1 E^{(1)}_- + (1 - \gamma_1) F^{(1)}_- \end{pmatrix}. \quad (6)$$

The set of operators $\{R^{(1)}_{ij}\}$ constitutes a so-called positive operator-valued measure (POVM). Only generalized measurements corresponding to POVMs are able to describe joint non-ideal measurements of incompatible observables. By calculating the marginals of probability distribution $p_{\gamma_1}(a_{1i}, b_{1j})$ it is possible to see that for each value of γ_1 information is obtained on both polarization observables, be it that information on polarization in direction θ_1 gets more non-ideal as γ_1 decreases, while information on polarization in direction θ'_1 is getting more ideal. This is in perfect agreement with the idea of mutual disturbance in a joint measurement of incompatible observables. The explanation of the non-existence of a single measurement result for observable A_1 , say, as implied by inequality (4), is corroborated by this analysis.

The analysis can easily be extended to the joint detection probabilities of the whole experiment of figure 6. The joint detection probability distribution of all four detectors is determined by a quadrivariate POVM

$$p_{\gamma_1 \gamma_2}(a_{1i}, b_{1j}, a_{2k}, b_{2\ell}) = Tr \rho R_{ijkl}, \quad (7)$$

that can be expressed in terms of the POVMs of the left and right interferometer arms according to

$$R_{ijkl} = R^{(1)}_{ij} R^{(2)}_{kl}. \quad (8)$$

It is important to note that the existence of the quadrivariate joint probability distribution (7), and the consequent satisfaction of Bell's inequality, is a consequence of the existence of quadruples of measurement results, available because it is possible to determine for each individual particle pair what is the result of each of the four detectors. Although, because of (8), also locality is assumed, this does not play an essential role. Under the condition that a quadruple of measurement results exists for each individual photon pair Bell's inequality would be satisfied also if, due to non-local interaction, R_{ijkl} were not a product of operators of the two arms of the interferometer. The reason why the standard Aspect experiments of figure 5 do not satisfy Bell's inequality is the non-existence of a quadrivariate joint probability distribution yielding the bivariate probabilities as marginals. Such a non-existence is strongly suggested by Heisenberg's idea of mutual disturbance in a joint measurement of incompatible observables. This is corroborated by the easily verifiable fact that the quadrivariate joint probability distributions of the standard Aspect experiments, obtained from (7) and (8) by taking γ_n to be either 1 or 0, are all distinct. Moreover, in general the quadrivariate joint probability distribution (7) for one standard Aspect experiment does not yield the bivariate ones of the other experiments as marginals. Although it is not strictly excluded that a quadrivariate joint probability distribution might exist having the bivariate probabilities of the standard Aspect experiments as marginals (hence, different from the ones referred to above), does the mathematical formalism of quantum mechanics not give any reason to surmise its existence. As far as quantum mechanics is concerned, the standard Aspect experiments need not satisfy Bell's inequality.

4 Bell's inequality in stochastic and deterministic hidden-variables theories

In stochastic hidden-variables theories the quantum mechanical probabilities are usually given as

$$p(a_1) = \int_{\Lambda} d\lambda \rho(\lambda) p(a_1|\lambda), \quad (9)$$

in which Λ is the space of hidden variable λ (to be compared with classical phase space), and $p(a_1|\lambda)$ is the conditional probability of measurement result $A = a_1$ if the value of the hidden variable was λ , and $\rho(\lambda)$ the probability of λ . It should be noticed that expression (9) fits perfectly into an empiricist interpretation of the quantum mechanical formalism, in which measurement result a_1 is referring to a pointer position of a measuring instrument, the object being described by the hidden variable. Since $p(a_1|\lambda)$ may depend on the specific way the measurement is carried out, the stochastic hidden-variables model corresponds to a contextualistic interpretation of quantum mechanical observables. Deterministic hidden-variables theories are just special cases in which $p(a_1|\lambda)$ is either 1 or 0. In the deterministic case it is possible to associate in a unique way (although possibly dependent on the measurement procedure) the value a_1 to the phase space point λ the object is prepared in. A disadvantage of a deterministic theory is that the physical interaction of object and measuring instrument is left out of consideration, thus suggesting measurement result a_1 to be a (possibly contextually

determined) property of the object. In order to have maximal generality it is preferable to deal with the stochastic case.

For Bell experiments we have

$$p(a_1, a_2) = \int_{\Lambda} d\lambda \rho(\lambda) p(a_1, a_2 | \lambda), \quad (10)$$

a condition of conditional statistical independence,

$$p(a_1, a_2 | \lambda) = p(a_1 | \lambda) p(a_2 | \lambda), \quad (11)$$

expressing that the measurement procedures of A_1 and A_2 do not influence each other (so-called locality condition).

As is well-known the locality condition was thought by Bell to be the crucial condition allowing a derivation of his inequality. This does not seem to be correct, however. As a matter of fact, Bell's inequality can be derived if a quadrivariate joint probability distribution exists [12, 13]. In a stochastic hidden-variables theory such a distribution could be represented by

$$p(a_1, b_1, a_2, b_2) = \int_{\Lambda} d\lambda \rho(\lambda) p(a_1, b_1, a_2, b_2 | \lambda), \quad (12)$$

without any necessity that the conditional probability be factorizable in order that Bell's inequality be satisfied (although for the generalized experiment discussed in section 3 it would be reasonable to require that $p(a_1, b_1, a_2, b_2 | \lambda) = p(a_1, b_1 | \lambda) p(a_2, b_2 | \lambda)$). Analogous to the quantum mechanical case, it is sufficient that for each individual preparation (here parameterized by λ) a quadruple of measurement results exists. If Heisenberg measurement disturbance is a physically realistic effect in the experiments at issue, it should be described by the hidden-variables theory as well. Therefore the explanation of the non-existence of such quadruples is the same as in quantum mechanics.

However, with respect to the possibility of deriving Bell's inequality there is an important difference between quantum mechanics and the stochastic hidden-variables theories of the kind discussed here. Whereas quantum mechanics does not yield any indication as regards the existence of a quadrivariate joint probability distribution returning the bivariate probabilities of the Aspect experiments as marginals, local stochastic hidden-variables theory does. Indeed, using the single-observable conditional probabilities assumed to exist in the local theory (compare (11)), it is possible to construct a quadrivariate joint probability distribution according to

$$p(a_1, a_2, b_1, b_2) = \int_{\Lambda} d\lambda \rho(\lambda) p(a_1 | \lambda) p(a_2 | \lambda) p(b_1 | \lambda) p(b_2 | \lambda), \quad (13)$$

satisfying all requirements. It should be noted that (13) does not describe the results of any joint measurement of the four observables that are involved. Quadruples (a_1, a_2, b_1, b_2) are obtained here by combining measurement results found in different experiments, assuming the same value of λ in all experiments. For this reason the

physical meaning of this probability distribution is not clear. However, this does not seem to be important. The existence of (13) as a purely mathematical constraint is sufficient to warrant that any stochastic hidden-variables theory in which (10) and (11) are satisfied, must require that the standard Aspect experiments obey Bell's inequality. Admittedly, there is a possibility that (13) might not be a valid mathematical entity because it is based on multiplication of the probability distributions $p(a|\lambda)$, which might be distributions in the sense of Schwartz' distribution theory. However, the remark made with respect to the existence of probability distributions as infinite- N limits of relative frequencies is valid also here: the reasoning does not depend on this limit, but is equally applicable to relative frequencies in finite sequences.

The question is whether this reasoning is sufficient to conclude that no *local* hidden-variables theory can reproduce quantum mechanics. Such a conclusion would only be justified if locality would be the *only* assumption in deriving Bell's inequality. If there would be any additional assumption in this derivation, then violation of Bell's inequality could possibly be blamed on the invalidity of this additional assumption rather than on locality. Evidently, one such additional assumption is the existence of hidden variables. A belief in the completeness of the quantum mechanical formalism would, indeed, be a sufficient reason to reject this assumption, thus increasing pressure on the locality assumption. Since, however, an empiricist interpretation is hardly reconcilable with such a completeness belief, we have to take hidden-variables theories seriously, and look for the possibility of additional assumptions within such theories.

In expression (9) one such assumption is evident, viz. the existence of the conditional probability $p(a_1|\lambda)$. The assumption of the applicability of this quantity in a quantum mechanical measurement is far less innocuous than appears at first sight. If quantum mechanical measurements really can be modeled by equality (9), this implies that a quantum mechanical measurement result is determined, either in a stochastic or in a deterministic sense, by an *instantaneous* value λ of the hidden variable, prepared independently of the measurement to be performed later. It is questionable whether this is a realistic assumption, in particular, if hidden variables would have the character of rapidly fluctuating stochastic variables. As a matter of fact, every individual quantum mechanical measurement takes a certain time, and it will in general be virtually impossible to determine the precise instant to be taken as the initial time of the measurement, as well as the precise value of the stochastic variable at that moment. Hence, hidden-variables theories of the kind considered here may be too specific.

Because of the assumption of a non-contextual preparation of the hidden variable, such theories were called *quasi-objectivistic* stochastic hidden-variables theories in [17] (dependence of the conditional probabilities $p(a_1|\lambda)$ on the measurement procedure preventing complete objectivity of the theory). In the past attention has mainly been restricted to quasi-objectivistic hidden-variables theories. It is questionable, however, whether the assumption of quasi-objectivity is a possible one for hidden-variables theories purporting to reproduce quantum mechanical measurement results. The existence of the quadrivariate probability distribution (13) is only excluding quasi-objectivistic local hidden-variables theories (either stochastic or deterministic) from the possibility of reproducing quantum mechanics. As will be seen in the next section, it is far more reasonable to blame quasi-objectivity than locality for this, thus leaving the possibility

of local hidden-variables theories that are not quasi-objectivistic.

5 Analogy between thermodynamics and quantum mechanics

The essential feature of expression (9) is the possibility to attribute, either in a stochastic or in a deterministic way, measurement result a_1 to an *instantaneous* value of hidden variable λ . The question is whether this is a reasonable assumption within the domain of quantum mechanical measurement. Are the conditional probabilities $p(a_1|\lambda)$ experimentally relevant within this domain? In order to give a tentative answer to this question, we shall exploit the analogy between thermodynamics and quantum mechanics, considered already a long time ago by many authors (e.g. [18, 19, 20, 21, 22]).

$$\begin{array}{ccc}
 \text{Quantum mechanics} & \rightarrow & \text{Hidden variables theory} \\
 (A_1, A_2, B_1, B_2) & & \lambda \\
 \updownarrow & & \updownarrow \\
 \text{Thermodynamics} & \rightarrow & \text{Classical statistical mechanics} \\
 (p, T, S) & & \{q_i, p_i\}
 \end{array}$$

In this analogy thermodynamics and quantum mechanics are considered as phenomenological theories, to be reduced to more fundamental “microscopic” theories. The reduction of thermodynamics to classical statistical mechanics is thought to be analogous to a possible reduction of quantum mechanics to stochastic hidden-variables theory. Due to certain restrictions imposed on preparations and measurements within the domains of the phenomenological theories their domains of application are thought to be contained in, but smaller than the domains of the “microscopic” theories.

In order to assess the nature and the importance of such restrictions let us first look at thermodynamics. As is well known (e.g. [23]) thermodynamics is valid only under a condition of molecular chaos, assuring the existence of local equilibrium¹ necessary for the ergodic hypothesis to be satisfied. Thermodynamics only describes measurements of quantities (like pressure, temperature, and entropy) being defined for such equilibrium states. From an operational point of view this implies that measurements within the domain of thermodynamics do not yield information on the object system, valid for one particular instant of time, but it is *time-averaged* information, time averaging being replaced, under the ergodic hypothesis, by ensemble averaging. In the Gibbs theory this ensemble is represented by the canonical density function $Z^{-1}e^{-H(\{q_n, p_n\})/kT}$ on phase space. This state is called a *macrostate*, to be distinguished from the *microstate* $\{q_n, p_n\}$, representing the point in phase space the classical object is in at a certain instant of time.

The restricted validity of thermodynamics is manifest in a two-fold way: i) through the restriction of all possible density functions $\rho(\{q_n, p_n\}) \geq 0$, $\int \rho(\{q_n, p_n\}) \Pi_n dq_n dp_n = 1$ on phase space to the canonical ones; ii) through the restriction of thermodynamical

¹In “equilibrium thermodynamics” the equilibrium is assumed to be even global.

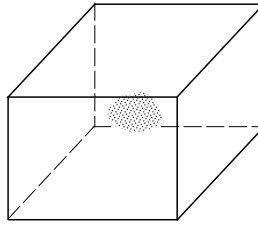


Figure 7: Example of a non-thermodynamic microstate.

quantities (observables) to functionals on the space of thermodynamic states. Physically this can be interpreted as a restriction of the domain of application of thermodynamics to those measurement procedures probing only properties of the macrostates. This implies that such measurements only yield information that is averaged over times exceeding the relaxation time needed to reach a state of (local) equilibrium. Thus, it is important to note that thermodynamic quantities are quite different from the physical quantities of classical statistical mechanics, the latter ones being represented by functions of the microstate $\{q_n, p_n\}$ and, hence, referring to a particular instant of time². Only if it is possible to perform measurements faster than the relaxation time would it be necessary to consider such non-thermodynamic quantities. Such measurements, then, are outside the domain of application of thermodynamics. Thus, if we have a cubic container containing a volume of gas in a microstate concentrated initially at its center (cf. figure 7), and if we could measure at a single instant of time either the total kinetic energy or the force exerted on the boundary of the container, then these results would not be equal to thermodynamic temperature and pressure³, respectively, because this microstate is not an equilibrium state. Only after the gas has reached equilibrium within the volume defined by the container (equilibrium) thermodynamics becomes applicable.

Within the domain of application of thermodynamics the microstate of the system may change appreciably without the macrostate being affected. Indeed, a macrostate is equivalent to an (ergodic) trajectory $\{q_n(t), p_n(t)\}_{ergodic}$. We might exploit as follows the difference between micro- and macrostates for characterizing objectivity of a physical theory. Whereas the microstate is thought to yield an objective description of the (microscopic) object, the macrostate just describes certain phenomena to be attributed to the object system only while being observed under conditions valid within the domain of application of the theory. In this sense classical mechanics is an objective theory, all quantities being instantaneous properties of the microstate. Thermodynamic quantities, only being attributable to the macrostate (i.e., to an ergodic trajectory), can not be seen, however, as properties belonging to the object at a certain instant of time. Of course, we might attribute the thermodynamic quantity to the event in space-time represented by the trajectory, but it should be realized that this

²Note that a “definition” of an instantaneous temperature by means of the equality $3/2nkT = \sum_i \mathbf{p}_i^2/2m_i$ does not make sense, as can easily be seen by applying this “definition” to an ideal gas in a container freely falling in a gravitational field.

³Thermodynamic pressure is defined for the canonical ensemble by $p = kT\partial/\partial V \log Z$.



Figure 8: Incompatible thermodynamic arrangements.

event is not determined solely by the preparation of the microstate, but is determined as well by the macroscopic arrangement serving to define the macrostate.

In order to illustrate this, consider two identical cubic containers differing only in their orientations (cf. figure 8). The same microstate as depicted in figure 7 may be prepared, in principle, in the two containers. Because of the different orientations, however, the macrostates, evolving from this microstate during the time the gas is reaching equilibrium with the container, are different (for different orientations of the container we have $H_1 \neq H_2$, and, hence, $e^{-H_1/kT}/Z_1 \neq e^{-H_2/kT}/Z_2$, since $H = T + V$, and $V_1 \neq V_2$ because potential energy is infinite outside a container). This implies that thermodynamic macrostates may be different even though starting from the same microstate. Macrostates in thermodynamics have a *contextual* meaning. It is important to note that, since the container is part of the preparing apparatus, this contextuality is connected here to *preparation* rather than to measurement. Consequently, whereas classical quantities $f(\{q_n, p_n\})$ can be interpreted as objective properties, thermodynamic quantities are non-objective, the non-objectivity being of a contextual nature.

Let us now suppose that quantum mechanics is related to hidden-variables theory analogous to the way thermodynamics is related to classical mechanics, the analogy maybe being even closer for non-equilibrium thermodynamics (only *local* equilibrium being assumed) than for the thermodynamics of global equilibrium processes. Support for this idea was found in [17], where it was demonstrated that in the Husimi representation of quantum mechanics by means of non-negative probability distribution functions on phase space an analogous restriction to a “canonical” set of distributions obtains as in thermodynamics. In particular, it was demonstrated that the dispersion-free states $\rho(q, p) = \delta(q - q_0)\delta(p - p_0)$ are not “canonical” in this sense. This implies that within the domain of quantum mechanics it does not make sense to consider the preparation of the object in a “microstate” with a well-defined value of the hidden variables (q, p) .

In the analogy quantum mechanical observables like A_1, A_2, B_1, B_2 should be compared to thermodynamic quantities like pressure, temperature, and entropy. The central issue in the analogy is the fact that thermodynamic quantities like pressure and temperature *cannot* be conditioned on the instantaneous phase space variable $\{q_n, p_n\}$ (microstate). Expressions like $p(\{q_n, p_n\})$ and $T(\{q_n, p_n\})$ are meaningless within thermodynamics. Thermodynamic quantities are conditioned on macrostates, corresponding to ergodic paths in phase space. Analogously, a quantum mechanical observable

might not correspond to an instantaneous property of the object, but might have to be associated with an (ergodic) path in hidden-variables space Λ (macrostate) rather than with an instantaneous value λ (microstate). This allows to extend the analogy between thermodynamics and quantum mechanics in the following way.

	<u>Thermodynamics</u>	<u>Quantum mechanics</u>
<u>Observables:</u>	(p, T, S)	(A_1, A_2, B_1, B_2)
<u>States (macrostates):</u>	$e^{-H/kT}/Z$ (canonical density function)	ρ (density operator)
<u>Reduction to:</u>	\downarrow Classical mechanics	\downarrow Hidden variables theory
<u>Microstates:</u>	$\{q_i, p_i\}$	λ
<u>Macrostates (ergodicity):</u>	$e^{-H/kT}/Z \sim \overline{\{q_i(t), p_i(t)\}}^t$	$\rho \sim \overline{\lambda(t)}^t$

On the basis of the analogy between thermodynamics and quantum mechanics it is possible to state the following conjectures:

- Quantum mechanical measurements (analogous to thermodynamic measurements) do not probe microstates but macrostates.
- Quantum mechanical quantities (analogous to thermodynamic quantities) should be conditioned on macrostates.

A hidden-variables macrostate will be symbolically indicated by $\bar{\lambda}^t$. For quantum mechanical measurements the conditional probabilities $p(a_1|\lambda)$ of (9) should then be replaced by $p(a_1|\bar{\lambda}^t)$. Concomitantly, quantum mechanical probabilities should be represented in the hidden-variables theory by a functional integral,

$$p(a_1) = \int d\bar{\lambda}^t \rho(\bar{\lambda}^t) p(a_1|\bar{\lambda}^t), \quad (14)$$

in which the integration is over all possible macrostates consistent with the preparation procedure.

By itself conditioning of quantum mechanical observables on macrostates rather than microstates is not sufficient to prevent derivation of Bell's inequality. As a matter of fact, on the basis of expression (14) a quadrivariate joint probability distribution can be defined, analogous to (13), according to

$$p(a_1, a_2, b_1, b_2) = \int d\bar{\lambda}^t \rho(\bar{\lambda}^t) p(a_1|\bar{\lambda}^t) p(a_2|\bar{\lambda}^t) p(b_1|\bar{\lambda}^t) p(b_2|\bar{\lambda}^t), \quad (15)$$

from which Bell's inequality can be derived just as well. There is, however, one important aspect that up till now has not sufficiently been taken into account, viz. contextuality. In the construction of (15) it is assumed that the macrostate $\bar{\lambda}^t$ is applicable

in each of the measurement arrangements of observables $A_1, A_2, B_1,$ and B_2 . Because of the incompatibility of some of these observables this is an implausible assumption. On the basis of the thermodynamic analogy it is to be expected that macrostates $\bar{\lambda}^t$ will depend on the measurement context of a specific observable. Since $[A_1, B_1]_- \neq O$, we will have

$$\bar{\lambda}^{tA_1} \neq \bar{\lambda}^{tB_1}, \quad (16)$$

and analogously for A_2 and B_2 . Then, for the Bell experiments measuring the pairs (A_1, A_2) and (A_1, B_2) , respectively, we have

$$p(a_1, a_2) = \int d\bar{\lambda}^{tA_1A_2} \rho(\bar{\lambda}^{tA_1A_2}) p(a_1|\bar{\lambda}^{tA_1A_2}) p(a_2|\bar{\lambda}^{tA_1A_2}), \quad (17)$$

$$p(a_1, b_2) = \int d\bar{\lambda}^{tA_1B_2} \rho(\bar{\lambda}^{tA_1B_2}) p(a_1|\bar{\lambda}^{tA_1B_2}) p(b_2|\bar{\lambda}^{tA_1B_2}). \quad (18)$$

Now, the contextuality expressed by inequality (16) prevents the construction of a quadrivariate joint probability distribution analogous to (15). Hence, like in the quantum mechanical approach, also in the local non-objectivistic hidden-variables theory a derivation of Bell's inequality is prevented due to the local contextuality involved in the interaction of the particle and the measuring instrument it is directly interacting with.

6 Conclusions

Our conclusion is that if quantum mechanical measurements do probe macrostates $\bar{\lambda}^{tA}$ rather than microstates λ , then Bell's inequality cannot be derived *for quantum mechanical measurements*. Both in quantum mechanics and in hidden-variables theories is Bell's inequality a consequence of the assumption that the theory is yielding an objective description of reality in the sense that the preparation of the microscopic object, *as far as relevant to the realization of the measurement result*, can be thought to be independent of the measurement arrangement. The important point to be noticed is that, although in Bell experiments the preparation of the particle pair at the source (i.e. the microstate) can be considered to be independent of the measurement procedures to be carried out later (and, hence, one and the same microstate can be assumed in different Bell experiments), the measurement result is only determined by the macrostate, which is co-determined by the interaction with the measuring instruments. It really seems that the Copenhagen maxim of the impossibility of attributing quantum mechanical measurement results to the object as objective properties, possessed independently of the measurement, should be taken very seriously, and implemented also in hidden-variables theories purporting to reproduce the quantum mechanical results. The *quantum mechanical dice* is only cast *after* the object has been interacting with the measuring instrument, even though its result can be deterministically determined by the (sub-quantum mechanical) microstate.

The thermodynamic analogy suggests which experiments could be done in order to transcend the boundaries of the domain of application of quantum mechanics. If

it would be possible to perform experiments that probe the microstate λ rather than the macrostate $\bar{\lambda}^{tA}$, then we are in the domain of (quasi-)objectivistic hidden-variables theories. Because of (13) it, then, is to be expected that Bell's inequality should be satisfied for such experiments. In such experiments preparation and measurement must be completed well within the relaxation time of the microstates. Such times have been estimated by Bohm [24] “for the sake of illustration” as the time light needs to cover a distance of the order of the size of an atom (10^{-18} s, say). If this is correct, then all present-day experimentation is well within the range of quantum mechanics, thus explaining the seemingly universal applicability of this latter theory. By hindsight, this would explain why Aspect's switching experiment [6] is corroborating quantum mechanics: the applied switching frequency (50 MHz), although sufficient to warrant locality, has been far too low to beat the local relaxation processes in each of the measuring instruments separately.

It has often been felt that the most surprising feature of Bell experiments is the possibility (in certain states) of a strict correlation between the measurement results of the two measured observables, without being able to attribute this to a previous preparation of the object (no ‘elements of physical reality’). For many physicists the existence of such strict correlations has been reason enough to doubt Bohr's Copenhagen solution to renounce causal explanation of measurement results, and to replace ‘determinism’ by ‘complementarity’. It seems that the urge for causal reasoning has been so strong that even within the Copenhagen interpretation a certain causality has been accepted, even a *non*-local one, in an EPR experiment (cf. figure 3) determining a measurement result for particle 2 by the measurement of particle 1. This, however, should rather be seen as an internal inconsistency of this interpretation, caused by a tendency to make the Copenhagen interpretation as realist as possible. In a consistent application of the Copenhagen interpretation to Bell experiments such experiments could be interpreted as measurements of bivariate correlation observables. The certainty of obtaining a certain (bivariate) eigenvalue of such an observable would not be more surprising than the certainty of obtaining a certain eigenvalue of a univariate one if the state vector equals the corresponding eigenvector.

It is important to note that this latter interpretation of Bell experiments takes seriously the Copenhagen idea that quantum mechanics need not *explain* the specific measurement result found in an individual measurement. Indeed, in order to compare theory and experiment it would be sufficient that quantum mechanics just *describe* the relative frequencies found in such measurements. In this view quantum mechanics is just a phenomenological theory, in an analogous way describing (not explaining) observations as does thermodynamics in its own domain of application. Explanations should be provided by “more fundamental” theories, describing the mechanisms behind the observable phenomena. Hence, the Copenhagen ‘completeness’ thesis should be rejected (although this need not imply a return to determinism).

This approach has important consequences. One consequence is that the non-existence, within quantum mechanics, of ‘elements of physical reality’ does not imply that ‘elements of physical reality’ do not exist at all. They could be elements of the “more fundamental” theories. In section 5 it was discussed how an analogy between quantum mechanics and thermodynamics could be exploited to spell this out. ‘El-

ements of physical reality' could correspond to hidden-variables microstates λ . The determinism necessary to explain the strict correlations, referred to above, would be explained if, within a given measurement context, a microstate would define a unique macrostate $\bar{\lambda}^{tA}$. This demonstrates how it could be possible that quantum mechanical measurement results cannot be attributed to the object as properties possessed prior to measurement, and there, yet, is sufficient determinism to yield a local explanation of strict correlations of quantum mechanical measurement results in certain Bell experiments.

Another important aspect of a dissociation of phenomenological and fundamental aspects of measurement is the possibility of an *empiricist* interpretation of quantum mechanics. As demonstrated by the generalized Aspect experiment discussed in section 3, an empiricist approach needs a generalization of the mathematical formalism of quantum mechanics, in which an observable is represented by a POVM rather than by a projection-valued measure corresponding to a self-adjoint operator of the standard formalism. Such a generalization has been very important in assessing the meaning of Bell's inequality. In the major part of the literature of the past this subject has been dealt with on the basis of the (restricted) standard formalism. However, some conclusions drawn from the restricted formalism are not cogent when viewed in the generalized one (for instance, because von Neumann's projection postulate is not applicable in general). For this reason we must be very careful when accepting conclusions drawn from the standard formalism. This, in particular, holds true for the issue of non-locality.

References

- [1] W. Heisenberg, *Zeitschr. f. Phys.* **33**, 879 (1925).
- [2] E. Schrödinger, *Naturwissenschaften* **23**, 807, 823, 844 (1935) (English translation in *Quantum Theory and Measurement*, eds. J.A. Wheeler and W.H. Zurek, Princeton Univ. Press, 1983, p. 152).
- [3] W.M. de Muynck, *Synthese* **102**, 293 (1995).
- [4] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
- [5] A. Aspect, P. Grangier, and G. Roger, *Phys. Rev. Lett* **47**, 460 (1981).
- [6] A. Aspect, J. Dalibard, and G. Roger, *Phys. Rev. Lett.* **49**, 1804 (1982).
- [7] K.R. Popper, *Quantum theory and the schism in physics*, Rowman and Littlefield, Totowa, 1982.
- [8] M. Jammer, *The philosophy of quantum mechanics*, Wiley, New York, 1974.
- [9] N. Bohr, *Phys. Rev.* **48**, 696 (1935).
- [10] J.S. Bell, *Physics* **1**, 195 (1964).
- [11] H.P. Stapp, *Phys. Rev. D* **3**, 1303 (1971); *Il Nuovo Cim.* **29B**, 270 (1975).

- [12] A. Fine, *Journ. Math. Phys.* **23**, 1306 (1982); *Phys. Rev. Lett.* **48**, 291 (1982).
- [13] P. Rastall, *Found. of Phys.* **13**, 555 (1983).
- [14] W.M. de Muynck, *Phys. Lett. A* **114**, 65 (1986).
- [15] W.M. de Muynck, W. De Baere, and H. Martens, *Found. of Phys.* **24**, 1589 (1994).
- [16] W.M. de Muynck, *Found. of Phys.* **30**, 205 (2000).
- [17] W.M. de Muynck and J.T. van Stekelenborg, *Ann. der Phys., 7. Folge*, **45**, 222 (1988).
- [18] L. de Broglie, *La thermodynamique de la particule isolée*, Gauthier–Villars, 1964; L. de Broglie, *Diverses question de mécanique et de thermodynamique classiques et relativistes*, Springer-Verlag, 1995.
- [19] D. Bohm, *Phys. Rev.* **89**, 458 (1953).
- [20] D. Bohm and J.-P. Vigier, *Phys. Rev.* **96**, 208 (1954).
- [21] E. Nelson, *Dynamical theories of Brownian motion*, Princeton University Press, 1967.
- [22] E. Nelson, *Quantum fluctuations*, Princeton University Press, 1985.
- [23] H.B. Hollinger and M.J.Zenzen, *The Nature of Irreversibility* (Dordrecht: D. Reidel Publishing Company, 1985), sect. 4.4.
- [24] D. Bohm, *Phys. Rev.* **85**, 166, 180 (1952).