

Review



Cite this article: De Raedt H, Katsnelson MI, Michielsen K. 2016 Quantum theory as plausible reasoning applied to data obtained by robust experiments. *Phil. Trans. R. Soc. A* **374**: 20150233.
<http://dx.doi.org/10.1098/rsta.2015.0233>

Accepted: 4 November 2015

One contribution of 14 to a theme issue
'Quantum foundations: information approach'.

Subject Areas:

quantum physics

Keywords:

logical inference, quantum theory,
inductive logic, probability theory

Author for correspondence:

H. De Raedt
e-mail: h.a.de.raedt@rug.nl

Quantum theory as plausible reasoning applied to data obtained by robust experiments

H. De Raedt¹, M. I. Katsnelson² and K. Michielsen^{3,4}

¹Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

²Institute of Molecules and Materials, Radboud University, 6525 ED Nijmegen, The Netherlands

³Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany

⁴Department of Physics, RWTH Aachen University, 52056 Aachen, Germany

 HDR, 0000-0001-3461-4015

We review recent work that employs the framework of logical inference to establish a bridge between data gathered through experiments and their objective description in terms of human-made concepts. It is shown that logical inference applied to experiments for which the observed events are independent and for which the frequency distribution of these events is robust with respect to small changes of the conditions under which the experiments are carried out yields, without introducing any concept of quantum theory, the quantum theoretical description in terms of the Schrödinger or the Pauli equation, the Stern–Gerlach or Einstein–Podolsky–Rosen–Bohm experiments. The extraordinary descriptive power of quantum theory then follows from the fact that it is plausible reasoning, that is common sense, applied to reproducible and robust experimental data.

1. Introduction

Quantum theory is unsurpassed as a description of the data produced by many very different experiments in (sub)-atomic, molecular and condensed matter physics, quantum optics, etc. A large body of work focuses on different interpretations of the quantum formalism [1–8] and its derivations from different sets of axioms [9–20] but offers no explanation of the success of quantum theory that goes beyond 'that is because of the way it is'.

In this paper, we review recent work that de-mystifies the extraordinary power of quantum theory [21–23] by formalizing the human thought processes by which relationships between data gathered through experiments and objective descriptions in terms of human-made concepts may be established. A basic premise of this approach is that scientific theories are the result of cognitive processing of the discrete events which are registered by our sensory system, of the logical and/or cause-and-effect relationships between those events, and the use of metaphors to make abstractions and construct concepts. This form of cognitive processing may be expressed in terms of the algebra of logical inference (LI), a mathematical framework that facilitates rational reasoning when there is uncertainty [24–28]. Statistical mechanics can be given an information-theoretical justification by viewing the former as a problem of LI, thereby establishing a relationship between the information-theoretic entropy [29] and the thermodynamic entropy [30,31]. The resulting maximum-entropy principle [26,28,30,31] has recently been generalized to a principle of entropic dynamics, a framework in which dynamical laws can be derived [32,33].

The LI approach, being extended logic, is not bound by ‘laws of physics’ and does not require assumptions such as ‘the observed events are signatures of an underlying objective reality—which may or may not be mathematical in nature’, ‘all things physical are information-theoretic in origin’, ‘the universe is participatory’, etc. It yields results that are unambiguous and independent of the individual subjective judgment, providing a rational explanation for the extraordinary descriptive power of quantum theory, and it also provides strong support for Bohr’s statement [34] that ‘The physical content of quantum mechanics is exhausted by its power to formulate statistical laws governing observations under conditions specified in plain language’.

LI applies to situations where there may or may not be causal relationships between the events [26,28]. Extracting cause-and-effect relationships from empirical evidence is a highly non-trivial problem. In general, LI does not establish cause-and-effect relationships [28,35], although rational reasoning about these relations should comply with the rules of LI. Furthermore, a derivation of a quantum theoretical description from LI principles does not prohibit the construction of cause-and-effect mechanisms that create the *impression* that these mechanisms produce data that can be described by quantum theory [36,37]. In fact, there is a substantial body of work demonstrating that it is indeed possible to construct simulation models which reproduce, on an event-by-event basis, the results of interference/entanglement/uncertainty experiments with single photons/neutrons [38–42]. This demonstration does not imply the reality of hidden variables or something like that.

The LI approach which we review here leads to the view that quantum theory is a phenomenological theory which can be derived from a set of simple general principles, not axioms, in a way that is independent of any (strictly speaking, unknown) ‘more microscopic’ level of description. Therefore, its power does not depend on whether there exists an underlying classical world with some hidden variables, or not. In this sense, there is a clear parallel with Einstein’s view on thermodynamics. Einstein did not regard thermodynamics as a constructive theory, an attempt to build a picture of complex phenomena out of some relatively simple propositions, but rather as a theory of principles based on empirically observed properties of phenomena, independent of a particular underlying model [43].

The paper is structured as follows. Section 2 briefly recapitulates the basic elements of LI and reviews applications to the Stern–Gerlach (SG) and Einstein–Podolsky–Rosen–Bohm (EPRB) experiment and a particle in a potential. It shows how the LI approach directly leads to the probabilities for observing the events without invoking any concept of quantum theory. In §3, we discuss two methods for transforming the solutions obtained through LI into the equations that we know from quantum theory. A summary and discussion of more general aspects of the work presented in this paper are given in §4.

2. Logical inference

The key concept of the LI approach is the plausibility, denoted by $P(A|B)$, which, in general, expresses the degree of belief of an individual that proposition A is true, given that proposition

B is true [25,26,28,44]. The plausibility $P(A|B)$ is an intermediate mental construct that serves to carry out inductive logic, that is rational reasoning, in a mathematically well-defined manner [26,28].

The algebra of LI can be derived from three so-called ‘desiderata’, namely (i) *plausibilities are represented by real numbers*, (ii) *plausibilities must exhibit agreement with rationality* and (iii) *all rules relating plausibilities must be consistent* [25–28]. These three desiderata only describe the essential features of the plausibilities and are not a set of axioms that plausibilities have to satisfy. It is a most remarkable fact that these three desiderata suffice to uniquely determine the set of rules by which plausibilities may be manipulated [25–28]. It can be shown [25–28] that plausibilities may be chosen to take numerical values in the range $[0, 1]$ and these values are related by three rules, namely: (1) $P(A|Z) + P(\bar{A}|Z) = 1$, where \bar{A} denotes the negation of proposition A and Z is a proposition assumed to be true; (2) the ‘product rule’ $P(AB|Z) = P(A|BZ)P(B|Z) = P(B|AZ)P(A|Z)$, where the ‘product’ BZ denotes the logical product (conjunction) of the propositions B and Z ; and (3) $P(A\bar{A}|Z) = 0$ and $P(A + \bar{A}|Z) = 1$, where the ‘sum’ $A + B$ denotes the logical sum (inclusive disjunction) of the propositions A and B [25–28]. The algebra of LI, as defined by the rules (1)–(3), contains Boolean algebra as a special case and is the foundation for powerful tools such as the maximum entropy method and Bayesian analysis [26,28]. The rules (1)–(3) are unique [26–28]: any other rule which applies to plausibilities represented by real numbers and is in conflict with rules (1)–(3) will be at odds with rational reasoning and consistency, as embodied by the desiderata (i)–(iii). It should be mentioned here that it is not allowed to define a plausibility for a proposition conditional on the conjunction of mutually exclusive propositions: reasoning on the basis of two or more contradictory premises is out of the scope of LI.

The applications of LI which we review in the present paper describe phenomena in a manner which is independent of individual subjective judgment. Therefore, to differentiate between the ‘objective’ and ‘subjective’ mode of application of LI, we will refer to the plausibility as ‘inference-probability’ or ‘i-prob’ for short. A more extensive discussion and arguments for distinguishing between plausibility, inference probability and Kolmogorov probability can be found in [22]. For the purpose of the present review, it is sufficient if the reader does not think of the i-prob as a frequency or probability in the traditional mathematical sense but merely as a numerical measure for the fact that proposition A is true, given that proposition B is true.

In real experiments, there is always uncertainty about some factors which may or may not influence the outcome of the measurements: it is presumptuous to assume that we know ‘everything’ about these factors. In particular, if experiment shows that (a) there is uncertainty about each individual event and (b) the conditions under which the experiment is carried out are also subject to uncertainties, then the data collected in such an experiment cannot be described by the traditional theories of classical physics, the reason being that the theoretical description of ‘classical physics’ assumes that there is absolute certainty about the outcome of each individual experiment on each individual object. By contrast, the LI approach is well suited to deal with uncertainties but, as will be explicitly shown later, to render the resulting description free of individual subjective judgement, it is necessary to assume that (c) the frequencies with which events are observed are reproducible and robust (to be discussed later) against small changes in the conditions. Furthermore, the LI approach only yields a quantum theoretical description if in addition we assume that (d) individual events are independent, meaning that knowing any (necessarily finite) set of events does not help to increase the certainty by which we can predict another (past or future) event that does not belong to the set.

If the experimental data comply with requirements (a)–(d), application of LI rather straightforwardly yields basic equations of quantum theory. The LI derivation of these equations has a generic structure. The first step is to list the features of the experiment that are deemed to be relevant and to introduce the i-probs of the individual events. The second step is to impose the condition that the experiment yields robust, reproducible results, not on the level of individual events, but on the level of the frequencies of observing many events and, depending on the problem, to impose other constraints about, for example, the fact that the particle moves, etc. As an example of such a constraint, we will use a natural requirement that the equations of

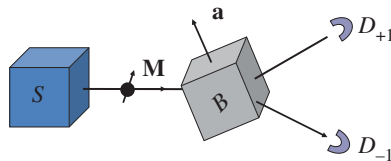


Figure 1. Diagram of the SG experiment. The source S , activated at times labelled by $i = 1, 2, \dots, N$, sends a particle carrying a magnetic moment \mathbf{M} to the magnet B with its (inhomogeneous) magnetic field oriented along the direction \mathbf{a} . Depending on the relative directions of \mathbf{a} and \mathbf{M} , the particle is detected with 100% certainty by either D_{+1} or D_{-1} . (Online version in colour.)

classical mechanics should be as accurate as possible for the *average* results of many quantum experiments. The result of the second step is a functional of the i-prob. The third step is to solve the robust optimization problem defined in terms of this functional and, optionally, to transform the formal solution in terms of i-probs into linear equations which we recognize as basic equations of quantum theory.

(a) Application: Stern–Gerlach experiment

The application of LI to event-based phenomena follows a particular pattern which is best illustrated by considering the simplest case, namely the SG experiment of which the schematic is shown in figure 1. In the SG experiment, there are two different outcomes which we label by the variable x taking the values $x = +1$ or $x = -1$. In an SG experiment, the source is activated at discrete times labelled by $i = 1, \dots, N$, resulting in time series of detection events $x_i = \pm 1$. The first step in the LI treatment is to assign to an individual event $x = \pm 1$, an i-prob $P(x | \mathbf{a}, \mathbf{M}, Z)$ to observe that event. Here, \mathbf{a} and \mathbf{M} are shorthands for the proposition that (within a small range) the directions of the magnet and of the magnetic moment of the particle are indeed \mathbf{a} and \mathbf{M} , respectively, and that the proposition Z , representing all the other conditions under which the experiment is performed, is true. It is assumed that the conditions represented by Z are fixed and identical for all experiments.

Assuming that the observed counts do not depend on the orientation of the chosen reference frame, $P(x | \mathbf{a}, \mathbf{M}, Z)$ can only depend on $\mathbf{a} \cdot \mathbf{M}$ (by construction $|\mathbf{a}| = 1$ and $|\mathbf{M}| = 1$). Hence, we must have $P(x | \mathbf{a}, \mathbf{M}, Z) = P(x | \mathbf{a} \cdot \mathbf{M}, Z) = P(x | \theta, Z)$, where $\cos \theta = \mathbf{a} \cdot \mathbf{M}$. This assumption is necessary to consider \mathbf{M} as the direction of the magnetic moment of the particle, whereas only \mathbf{a} is known from experiment. Such symmetry requirements are very important for our construction, as they establish relationships between what is measured (position of detector) and what is supposed to be measured (characteristics of a particle). It is expedient to write $P(x | \theta, Z)$ as

$$P(x | \theta, Z) = \frac{1 + xE(\theta)}{2}, \quad E(\theta) = \sum_{x=\pm 1} xP(x | \theta, Z). \quad (2.1)$$

According to assumption (d), there is no relationship between the actual values of x_n and $x_{n'}$ if $n \neq n'$. With this assumption, repeated application of the product rule yields

$$P(x_1, \dots, x_N | \theta, Z) = \prod_{i=1}^N P(x_i | \theta, Z). \quad (2.2)$$

Repeating the experiment N times yields n_x events of the type $\{x\}$ ($n_{+1} + n_{-1} = N$) and the i-prob to observe the compound event $\{n_{+1}, n_{-1}\}$ is given by [22]

$$P(n_{+1}, n_{-1} | \theta, N, Z) = N! \prod_{x=\pm 1} \frac{P(x | \theta, Z)^{n_x}}{n_x!}. \quad (2.3)$$

Although the individual events may be expected to change from run to run, for sufficiently large N the numbers $\{n_{+1}, n_{-1}\}$ should exhibit some kind of robustness with respect to small changes of θ . Otherwise, the $\{n_{+1}, n_{-1}\}$ would vary erratically with θ and these ‘irreproducible’ experiments would be discarded. Obviously, the expected robustness with respect to small variations should be reflected in the expression of the i-prob to observe the data (within the usual statistical fluctuations).

If the outcome of the experiment is indeed described by the i-prob equation (2.3) and the experiment is supposed to yield reproducible, robust results, small changes of θ should not have a drastic effect on the outcome. Let H_0 and H_1 be the hypotheses that the data $\{n_{+1}, n_{-1}\}$ are observed if the angle between the unit vector \mathbf{a} is θ and $\theta + \varepsilon$, respectively. The evidence Ev of hypothesis H_1 , relative to hypothesis H_0 , is defined by [26,28]

$$Ev = \ln \frac{P(n_{+1}, n_{-1} | \theta + \varepsilon, N, Z)}{P(n_{+1}, n_{-1} | \theta, N, Z)}, \quad (2.4)$$

where the logarithm serves to facilitate the algebraic manipulations. If H_1 is more (less) plausible than H_0 then $Ev > 0$ ($Ev < 0$). The absolute value of the evidence, $|Ev|$ is a measure for the robustness of the description (the sign of Ev is arbitrary, hence irrelevant): the smaller $|Ev|$ the more robust the experiment is for small changes of θ .

The problem of determining the most robust description of the experimental data may now be formulated as follows: search for the i-probs $P(n_{+1}, n_{-1} | \theta, N, Z)$ which minimize $|Ev|$ for all possible ε (ε small) and for all possible θ . The clauses ‘for all possible ε and θ ’ render the minimization problem an instance of a robust optimization problem. The robust optimization problem has a trivial solution, namely $P(n_{+1}, n_{-1} | \theta, N, Z) = P(n_{+1}, n_{-1} | N, Z)$, which can only describe experiments for which $\{n_{+1}, n_{-1}\}$ show no dependence on θ . Experiments which produce results that do not depend on the conditions seem fairly pointless and therefore we explicitly exclude i-probs that are constant with respect to changes of the conditions. It is not difficult to show [22] that our concept of a robust experiment implies that the i-probs which describe such an experiment can be found by minimizing $|Ev|$, subject to the constraints that (C1) ε is small but arbitrary, (C2) not all i-probs are independent of θ and (C3) $|Ev|$ is independent of θ [21–23]. Omitting terms of $\mathcal{O}(\varepsilon^3)$, minimizing $|Ev|$ while taking into account the constraints (C2) and (C3) amounts to finding the i-probs $P(x | \theta, Z)$ which minimize [22],

$$I_F = \sum_{x=\pm 1} \frac{1}{P(x | \theta, Z)} \left(\frac{\partial P(x | \theta, Z)}{\partial \theta} \right)^2, \quad (2.5)$$

subject to the constraint that $\partial P(x | \theta, Z) / \partial \theta \neq 0$.¹ The r.h.s. of equation (2.5) is the Fisher information for the problem at hand and, because of the constraint (C3), should not depend on θ .

Using equation (2.1), we can rewrite equation (2.5) as $I_F = (\partial E(\theta) / \partial \theta)^2 / (1 - E^2(\theta))$, yielding $E(\theta) = \cos(\theta \sqrt{I_F} + \phi)$, where ϕ is an integration constant. As $E(\theta)$ is a periodic function of θ we must have $\sqrt{I_F} = K$, where K is an integer and hence $E(\theta) = \cos(K\theta + \phi)$. The solution $K = I_F = 0$ is excluded from further consideration because it describes an experiment in which the frequency distribution of the observed data does not depend on θ (see constraint (C2)). Therefore, the physically relevant, non-trivial solution with minimum Fisher information corresponds to $K = 1$. Furthermore, as $E(\theta)$ is a function of $\mathbf{a} \cdot \mathbf{M} = \cos \theta$ only, we must have $\phi = 0, \pi$. Therefore, for the SG experiment, the solution of the robust optimization problem reads

$$P(x | \mathbf{a} \cdot \mathbf{M}, Z) = P(x | \theta, Z) = \frac{1 \pm x \mathbf{a} \cdot \mathbf{M}}{2}. \quad (2.6)$$

The \pm sign in equation (2.6) reflects the fact that the mapping between $x = \pm 1$ and the two different directions is only determined up to a sign.

¹In the course of deriving equation (2.5), our criterion of robustness enforces the intuitively obvious assignment $P(x | \theta, Z) = n_x / N$, establishing the relationship between the epistemological concept (i-prob) and the physically measurable quantity (frequency of outcomes). It is at this point that the possibility to view the i-prob as a ‘subjective’ assignment is eliminated [22].

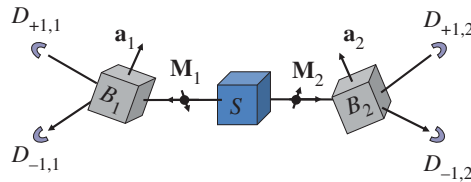


Figure 2. Diagram of the EPRB thought experiment. The source S , activated at times labelled by $i = 1, 2, \dots, N$, sends a particle with magnetic moment, represented by the unit vector \mathbf{M}_1 , to the SG magnet B_1 and another particle with magnetic moment, represented by the unit vector \mathbf{M}_2 , to the SG magnet B_2 . The orientations of the magnets, represented by unit vectors \mathbf{a}_1 and \mathbf{a}_2 , affect the deflection of the particles by the magnets. Each particle going to the left (right) is detected with 100% certainty by either $D_{+1,1}$ or $D_{-1,1}$ ($D_{+1,2}$ or $D_{-1,2}$). (Online version in colour.)

Comparing equation (2.6) with the quantum theoretical expression (which is exactly the same) demonstrates that Born's rule, one of the postulates of quantum theory, appears as a consequence of LI applied to robust experiments [22,23]. In the LI approach, equation (2.6) is not postulated but follows from the assumption that the (thought) experiment that is being performed yields the most reproducible results, revealing the conditions for an experiment to produce data which are described by quantum theory.

(b) Application: Einstein–Podolsky–Rosen–Bohm experiment

The LI treatment of the EPRB experiment is, in essence, the same as that of the SG experiment. Therefore, we only discuss the main assumptions and present the results. Technical details can be found elsewhere [22].

Referring to the schematic shown in figure 2, the i-prob to observe a pair $\{x, y\}$ is denoted by $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z)$, where Z represents all the conditions under which the experiment is performed, with the exception of the directions \mathbf{a}_1 and \mathbf{a}_2 of the SG magnets B_1 and B_2 , respectively. It is important to note that $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z)$ does not depend on \mathbf{M}_1 and \mathbf{M}_2 . In concert with the general assumption (d), it is assumed that there is no relationship between the actual values of the pairs $\{x_i, y_i\}$ and $\{x_{i'}, y_{i'}\}$ if $i \neq i'$, meaning that each repetition of the experiment represents an identical event of which the outcome is logically independent of any other such event. Invoking the product rule, the logical consequence of this assumption is that $P(x_1, y_1, \dots, x_N, y_N | \mathbf{a}_1, \mathbf{a}_2, Z) = \prod_{i=1}^N P(x_i, y_i | \mathbf{a}_1, \mathbf{a}_2, Z)$ meaning that the i-prob $P(x_1, y_1, \dots, x_N, y_N | \mathbf{a}_1, \mathbf{a}_2, Z)$ to observe the compound event $\{\{x_1, y_1\}, \dots, \{x_N, y_N\}\}$ is completely determined by the i-prob $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z)$ to observe the pair $\{x, y\}$.

We also assume that the i-prob $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z)$ to observe a pair $\{x, y\}$ does not change if we apply the same rotation to both magnets B_1 and B_2 . Expressing this invariance with respect to rotations of the coordinate system (Euclidean space and Cartesian coordinates are used throughout this paper) in terms of i-probs yields $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z) = P(x, y | \mathcal{R}\mathbf{a}_1, \mathcal{R}\mathbf{a}_2, Z)$, where \mathcal{R} denotes an arbitrary rotation in three-dimensional space which is applied to both magnets B_1 and B_2 , implying that $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z)$ is a function of the inner product $\mathbf{a}_1 \cdot \mathbf{a}_2$ only. Therefore, we must have $P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z) = P(x, y | \mathbf{a}_1 \cdot \mathbf{a}_2, Z) = P(x, y | \theta, Z)$, where $\theta = \arccos(\mathbf{a}_1 \cdot \mathbf{a}_2)$ denotes the angle between the unit vectors \mathbf{a}_1 and \mathbf{a}_2 . Note that for any integer value of K , $\theta + 2\pi K$ represents the same physical arrangement of the magnets M_1 and M_2 . From the algebra of LI, it follows that the i-prob to observe x , irrespective of the observed value of y , is given by $P(x | \mathbf{a}_1, \mathbf{a}_2, Z) = \sum_{y=\pm 1} P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z)$. In the context of the EPRB experiment, it is assumed that observing $x = +1$ is as likely as observing $x = -1$, independent of the observed value of y . This implies that we must have $P(x = +1 | \mathbf{a}_1, \mathbf{a}_2, Z) = P(x = -1 | \mathbf{a}_1, \mathbf{a}_2, Z)$ which, in view of the fact that $P(x = +1 | \mathbf{a}_1, \mathbf{a}_2, Z) + P(x = -1 | \mathbf{a}_1, \mathbf{a}_2, Z) = 1$ implies that $P(x = +1 | \mathbf{a}_1, \mathbf{a}_2, Z) = P(x = -1 | \mathbf{a}_1, \mathbf{a}_2, Z) = \frac{1}{2}$. Applying the same reasoning to the assumption that, independent of the observed values of x , observing $y = +1$ is as likely as observing $y = -1$ yields $P(y | \mathbf{a}_1, \mathbf{a}_2, Z) = P(x = +1, y | \mathbf{a}_1, \mathbf{a}_2, Z) + P(x = -1, y | \mathbf{a}_1, \mathbf{a}_2, Z) = \frac{1}{2}$. Note that we did not assign any prior i-prob

nor did we make any reference to concepts such as the singlet state. Although the symmetry properties which have been assumed are reminiscent of those of the singlet state, this is deceptive: without altering the assumptions, the LI approach also yields the correlations for the triplet states [22].

Adopting the same reasoning as in §2b, it follows directly from assumptions (a)–(d) that the i -prob to observe a pair $\{x, y\}$ takes the form [22] $P(x, y | \theta, Z) = [1 + xyE_{12}(\theta)]/4$, where $E_{12}(\theta) = E_{12}(\mathbf{a}_1, \mathbf{a}_2, Z)$ is a periodic function of θ . Minimization of the corresponding expression of $|E_V|$ while taking into account the constraints (C2) and (C3) (see §2a) is tantamount to finding the i -probs $P(x, y | \theta, Z)$ which minimize [22]

$$I_F = \sum_{x, y = \pm 1} \frac{1}{P(x, y | \theta, Z)} \left(\frac{\partial P(x, y | \theta, Z)}{\partial \theta} \right)^2 = \frac{1}{1 - E_{12}^2(\theta)} \left(\frac{\partial E_{12}(\theta)}{\partial \theta} \right)^2, \quad (2.7)$$

subject to the constraint that $\partial P(x, y | \theta, Z) / \partial \theta \neq 0$ for some pairs (x, y) . Equation (2.7) is readily integrated to yield $E_{12}(\theta) = \cos(\theta\sqrt{I_F} + \phi)$, where ϕ is an integration constant. As $E_{12}(\theta)$ is a periodic function of θ , we must have $\sqrt{I_F} = K$, where K is an integer and hence $E_{12}(\theta) = \cos(K\theta + \phi)$. Because of constraint (C2), we exclude the case $K = I_F = 0$ from further consideration. Hence the physically relevant, non-trivial solution corresponds to $K = 1$. Furthermore, as $E_{12}(\theta)$ is a function of $\mathbf{a}_1 \cdot \mathbf{a}_2 = \cos \theta$ only, we must have $\phi = 0, \pi$, reflecting an ambiguity in the definition of the direction of B_1 relative to the direction of B_2 . Choosing the solution with $\phi = \pi$, we find

$$P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z) = \frac{1 - xy\mathbf{a}_1 \cdot \mathbf{a}_2}{4}, \quad \langle xy \rangle = \sum_{x, y = \pm 1} xy P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z) = -\mathbf{a}_1 \cdot \mathbf{a}_2, \quad (2.8)$$

$\langle x \rangle = \sum_{x, y = \pm 1} x P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z) = 0$ and $\langle y \rangle = \sum_{x, y = \pm 1} y P(x, y | \mathbf{a}_1, \mathbf{a}_2, Z) = 0$, all in complete agreement with the quantum theoretical description of two $S = \frac{1}{2}$ particles in the singlet state [45,46]. As the LI treatment of a robust EPRB experiment directly yields the probabilistic description that we know from quantum theory without invoking the notions of the latter, it follows that the concept of quantum entanglement cannot be essential for describing the data produced by EPRB experiments.

It may be of interest to mention here that, in spite of the widely spread claims that real EPRB experiments have proved quantum theory correct, none of the three different experiments for which data have been made available [47–49] survives the confrontation with the five-standard-deviation-criterion hypothesis test that the data comply with the quantum theoretical description given by equation (2.8) [50,51]. It seems that, for the time being, only computer experiments are able to generate data that are not in conflict with the quantum theoretical description of the EPRB thought experiment [39,50].

(c) Application: particle in a potential

For simplicity of notation and presentation, in the present review, we only discuss the problem of inferring the plausibility that the particle is at a certain position X on a line and produces a click on a detector at position x , also on a line. The fully fledged three-dimensional derivation including electromagnetic potentials and/or spin can be found elsewhere [22,23].

The measurement scenario is as follows. We imagine N repetitions ($n = 1, \dots, N$) of an experiment performed on a particle moving on a line of linear extent $[-L, L]$. Nothing is known about the direction of motion of the particle. In each such experiment, a source emits a signal at discrete times labelled by the integer $\tau = 1, \dots, M$. It is assumed that for each repetition, the particle is at the unknown position $-L \leq X_\tau \leq L$. The signal solicits a response of the particle, generating a click of the detector at discrete position $j_{n,\tau}$ with $-K \leq j_{n,\tau} \leq K$. The detectors $-K, \dots, K$ have spatial extent $\Delta = L/K$ and are placed next to each other, completely covering the line segment $[-L, L]$. It is assumed that for each signal emitted by the source, one and only one of the $2K + 1$ detectors fires.

The result of N repetitions of the experiment yields the dataset

$$\Upsilon = \{j_{n,\tau} \mid -K \leq j_{n,\tau} \leq K; n = 1, \dots, N; \tau = 1, \dots, M\}, \quad (2.9)$$

or, denoting the total count of clicks of detector j at time τ by $0 \leq k_{j,\tau} \leq N$, we have

$$\mathcal{D} = \left\{ k_{j,\tau} \mid N = \sum_{j=-K}^K k_{j,\tau}; \tau = 1, \dots, M \right\}. \quad (2.10)$$

Following the general procedure, the next step is to introduce the i-prob $P(j \mid X_\tau, \tau, Z)$, expressing the relation between a particle at unknown location X_τ at discrete time τ and the click of the detector at position j . The conditions represented by Z are fixed and identical for all experiments. Note that we will not try to estimate the unknown position X but rather determine the i-prob $P(j \mid X_\tau, \tau, Z)$ which yields the most robust set of data, robust with respect to small changes of X_τ for all τ . According to basic assumption (d), there is no relation between the actual values of $j_{n,\tau}$ and $j_{n',\tau'}$ if $n \neq n'$ or $\tau \neq \tau'$. Hence, for fixed positions X_τ , the i-prob to observe all the data \mathcal{D} is given by

$$P(\mathcal{D} \mid X_1, \dots, X_N, N, Z) = N! \prod_{\tau=1}^M \prod_{j=-K}^K \frac{P(j \mid X_\tau, \tau, Z)^{k_{j,\tau}}}{k_{j,\tau}!}. \quad (2.11)$$

It is now straightforward to repeat the steps that led to equation (2.4) to find that the measure of a robust experiment is given by

$$\text{Ev} = -\frac{N\varepsilon^2}{2} \sum_{j=-K}^K \sum_{\tau=1}^M \frac{1}{P(j \mid X_\tau, \tau, Z)} \left(\frac{\partial P(j \mid X_\tau, \tau, Z)}{\partial X_\tau} \right)^2 + \mathcal{O}(\varepsilon^3), \quad (2.12)$$

and that the most non-trivial robust experiment is described by the i-prob $P(j \mid X_\tau, \tau, Z)$ which minimizes the Fisher information

$$I_F = \sum_{j=-K}^K \sum_{\tau=1}^M \frac{1}{P(j \mid X_\tau, \tau, Z)} \left(\frac{\partial P(j \mid X_\tau, \tau, Z)}{\partial X_\tau} \right)^2, \quad (2.13)$$

subject to the constraint that not all $\partial P(j \mid X_\tau, \tau, Z) / \partial X_\tau$ are zero and the additional constraints to be discussed below.

As the Schrödinger equation is formulated in continuum space, it is necessary to replace equation (2.13) by its continuum limit,

$$\begin{aligned} I_F &= \int dx \int dt \frac{1}{P(x \mid X(t), t, Z)} \left(\frac{\partial P(x \mid X(t), t, Z)}{\partial X(t)} \right)^2 \\ &= \int dx \int dt \frac{1}{P(x \mid X(t), t, Z)} \left(\frac{\partial P(x \mid X(t), t, Z)}{\partial x} \right)^2, \end{aligned} \quad (2.14)$$

where we assumed that it does not matter where in space we perform the experiment (homogeneity of space), implying that $P(x \mid X(\tau), \tau, Z) = P(x + \zeta \mid X(\tau) + \zeta, \tau, Z)$, where ζ is an arbitrary real number. As before, it is a symmetry requirement which allows us to regard the unknown quantity X as the 'coordinate of a particle' based on measurement of the coordinate of the detector that clicks. Technically speaking, after passing to the continuum limit, $P(x \mid X(t), t, Z)$ denotes the probability density, not the probability itself, but as there can be no confusion about which case, discrete or continuum, we are considering, we use the same symbol for the probability density and the probability.

In general, if there is no uncertainty about individual events, we expect the description to agree with classical theoretical mechanics. We use this 'correspondence principle' to incorporate classical theoretical mechanics into the LI approach [22,23]. In the absence of uncertainty and in line with the basic ideas of classical mechanics, the observed detector clicks form smooth trajectories. One such trajectory can always be represented by $dx(t)/dt = U(x(t), t)$ but the function

$U(.,.)$ may not be ‘universal’ in the sense that it may change from experiment to experiment, e.g. with the initial conditions. However, if $U(.,.)$ is universal and sufficiently ‘nice’ (we ignore technical details related to differentiability, etc.) we may write [52]

$$\frac{dx(t)}{dt} = U(x(t), t) = \frac{\partial S(x, t)}{\partial x}, \quad (2.15)$$

where $S(x, t) \equiv S(x(t), t)$ and it follows that

$$\frac{d^2x(t)}{dt^2} = \frac{\partial}{\partial x} \left[\frac{\partial S(x, t)}{\partial t} + \frac{1}{2} \left(\frac{\partial S(x, t)}{\partial x} \right)^2 \right] \equiv -\frac{\partial V(x, t)}{\partial x}, \quad (2.16)$$

showing that if there exists a universal function $U(x(t), t)$ which describes the data according to equation (2.15), then there exists a potential $V(x, t)$ such that the Hamilton–Jacobi equation,

$$\frac{\partial S(x, t)}{\partial t} + \frac{1}{2} \left(\frac{\partial S(x, t)}{\partial x} \right)^2 + V(x, t) = 0, \quad (2.17)$$

holds [52]. Thus, the assumption that in the absence of uncertainty, all the possible trajectories $x(t)$ can be described by one function $U(.,.)$ quite straightforwardly yields equation (2.17), that is one of the formulations of classical theoretical mechanics.

In the presence of uncertainty about individual events, we can combine the notion of a robust experiment and the desire to recover equations of classical mechanics as a limiting case by searching for the robust (i.e. for all $X(t)$) minima of the functional [22,23]

$$F = \int dx \int dt \left\{ \frac{1}{P(x | X(t), t, Z)} \left(\frac{\partial P(x | X(t), t, Z)}{\partial x} \right)^2 + 2m\lambda \left[\frac{\partial S(x, t)}{\partial t} + \frac{1}{2m} \left(\frac{\partial S(x, t)}{\partial x} \right)^2 + V(x, t) \right] P(x | X(t), t, Z) \right\}, \quad (2.18)$$

where λ is a parameter having dimension $\text{s}^2/\text{kg m}^4$ and, for convenience of comparing with quantum theory, we introduced the mass m of the particle by substituting $S(x, t) \rightarrow S(x, t)/m$ and $V(x, t) \rightarrow V(x, t)/m$.

Standard variational calculus yields the extrema of equation (2.18) in terms of two coupled nonlinear first-order differential equations of the functions $P(x | X(t), t, Z)$ and $S(x, t)$ which are identical to the (one-dimensional version of the) equations that appear in Madelung’s hydrodynamical form [53] or Bohm’s interpretation [1] of quantum theory. However, equation (2.18) was not derived from quantum theory but was obtained through logical inference from data produced by robust experiments and a correspondence principle, without invoking concepts of quantum theory. Therefore, in principle we do not need the latter to describe these experiments but we can use the equivalence of equation (2.18) and the mathematical framework of quantum theory to great advantage for turning the nonlinear equations into linear ones which can be solved by the powerful machinery of linear algebra. Technical details of the derivation of the functionals analogous to F for the multidimensional Schrödinger equation and Pauli equation for a particle with spin can be found in the original papers [22] and [23], respectively.

3. Connecting with quantum theory

The LI approach yields descriptions of robust experiments in terms of i-probs. In this section, we discuss two different methods of transforming these i-probs into the wavefunction formalism of quantum theory.

The first method is based on the general observation that in scientific reasoning it is good practice to reduce the complexity of the description of the whole by separating the description of data into several parts. We consider different ways of organizing the observed data and scrutinize the conditions under which a description of the various parts of the experiment can be separated (as much as possible). Then we show, in the case of the SG and EPRB experiments,

how the wavefunction description naturally emerges as a result of this separation procedure. It automatically follows that the wavefunction (or density matrix) description is less general than the one in terms of conditional probabilities in the sense that the former can only describe situations in which the separation procedure can actually be carried out.

The second method employs a polar representation of the i-prob to bring the nonlinear robust optimization problem into a linear form and is most useful for cases that involve dynamics, yielding Schrödinger-like equations.

(a) Separation procedure [54]

Consider again the SG experiment (figure 1) yielding the dataset

$$\mathcal{D} = \{x_1, \dots, x_N \mid x_i = \pm 1, i = 1, \dots, N\}, \quad (3.1)$$

where N is the total number of recorded events. Suppose that the analysis of correlations among the observed x_i indicates that the x_i are independent events, in line with assumption (d) (see §2). Then the counts $N(\pm 1 \mid \mathbf{a}, \mathbf{M}, Z)$ of outcomes with $x = \pm 1$ ($N = N(+1 \mid \mathbf{a}, \mathbf{M}, Z) + N(-1 \mid \mathbf{a}, \mathbf{M}, Z)$) give a complete characterization of the data. In essence, all datasets having the same average

$$\langle x \rangle = \frac{1}{N} \sum_{x=\pm 1} xN(x \mid \mathbf{a}, \mathbf{M}, Z) \equiv \sum_{x=\pm 1} xf(x \mid \mathbf{a}, \mathbf{M}, Z) \quad (3.2)$$

are equivalent. Assuming (as in §2a) that the observed counts do not depend on the orientation of the chosen reference frame, $f(x \mid \mathbf{a}, \mathbf{M}, Z)$ can only depend on $\mathbf{a} \cdot \mathbf{M}$ (by construction $|\mathbf{a}| = 1$ and $|\mathbf{M}| = 1$). Hence, we must have $f(x \mid \mathbf{a}, \mathbf{M}, Z) = f(x \mid \mathbf{a} \cdot \mathbf{M}, Z)$.

Equation (3.2) is a holistic description of the data in terms of $\mathbf{a} \cdot \mathbf{M}$ and it is by no means obvious how to construct, if possible at all, a description in terms of a part that refers to the object (represented by \mathbf{M}) and another part that refers to the magnet (represented by \mathbf{a}). To explore the possibilities of separating in parts, it is expedient to consider alternative ways of writing equation (3.2). Let us first organize the data and frequencies in vectors $\mathbf{x} = (+1, -1)^T$ and $\mathbf{f} = (f(+1 \mid \mathbf{a}, \mathbf{M}, Z), f(-1 \mid \mathbf{a}, \mathbf{M}, Z))^T$, respectively. Then, we trivially have

$$\langle x \rangle = \mathbf{x}^T \cdot \mathbf{f} = \text{Tr } \mathbf{x}^T \mathbf{f} = \text{Tr } \mathbf{f} \mathbf{x}^T, \quad (3.3)$$

where $\mathbf{f} \mathbf{x}^T$ is a 2×2 matrix and $\text{Tr } A$ denotes the trace of the matrix A . Now note that *any* rewriting of \mathbf{x} and \mathbf{f} in terms of vectors, matrices, \dots , $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{f}}$ such that $\text{Tr } \tilde{\mathbf{f}} = 1$ and $\text{Tr } \tilde{\mathbf{f}} \tilde{\mathbf{x}} = \langle x \rangle$ does not change $\langle x \rangle$; that is, it yields the same complete description of the data. Therefore, with this in mind, we consider the rearrangement of the data into 2×2 (diagonal, hermitian) matrices X and F with elements $X(x, x') = x\delta_{x,x'}$ and $F(x, x') = f(x \mid \mathbf{a}, \mathbf{M}, Z)\delta_{x,x'}$, respectively, and rewrite equation (3.3) as

$$\langle x \rangle = \text{Tr } FX = \text{Tr } \hat{\rho} \hat{X}, \quad (3.4)$$

where $\hat{\rho}$ and \hat{X} can be any pair of 2×2 matrices that satisfies equation (3.4). Clearly, a formal rewriting of equation (3.3) such as equation (3.4) cannot, by itself, bring anything new but representation (3.4) offers the flexibility that allows us to perform the separation by using some elementary linear algebra, as we now show.

We know from linear algebra that any hermitian 2×2 matrix can be written as a linear combination of four hermitian 2×2 matrices. Without loss of generality, we may choose the Pauli-spin matrices $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ and the unit matrix $\mathbb{1}$ as the orthonormal basis set for the vector space of 2×2 matrices with an inner product defined by $(A, B) = \text{Tr } A^\dagger B$. Without loss of generality, we may write

$$\hat{\rho} = \frac{\mathbb{1} + \boldsymbol{\rho} \cdot \boldsymbol{\sigma}}{2} \quad \text{and} \quad \hat{X} = u_0 \mathbb{1} + \mathbf{u} \cdot \boldsymbol{\sigma}, \quad (3.5)$$

where $\boldsymbol{\rho} = (\rho_x, \rho_y, \rho_z)$, u_0 and $\mathbf{u} = (u_x, u_y, u_z)$ are all real-valued. It is now straightforward to show [23] that the desired separation can be realized by requiring that $u_0 = u_0(\mathbf{a}, Z)$, $u_x = u_x(\mathbf{a}, Z)$, $u_y = u_y(\mathbf{a}, Z)$, $u_z = u_z(\mathbf{a}, Z)$, $\rho_x = \rho_x(\mathbf{M}, Z)$, $\rho_y = \rho_y(\mathbf{M}, Z)$ and $\rho_z = \rho_z(\mathbf{M}, Z)$ (recall that Z is considered to represent all fixed conditions which are important to the actual experiment but are

not of immediate interest). Assuming that the observed counts do not depend on the orientation of the reference frame (see earlier), $\langle x \rangle$ is a function of $\mathbf{a} \cdot \mathbf{M}$ only. This requirement enforces $\boldsymbol{\rho} = \mathbf{M}$ and $\mathbf{u} = \mathbf{a}$. Hence, we have $\langle x \rangle = u_0 + \mathbf{M} \cdot \mathbf{a}$ and as $|\langle x \rangle| \leq 1$ it follows that $-1 \leq u_0 + \mathbf{M} \cdot \mathbf{a} \leq 1$; for $\mathbf{a} = \mathbf{M}$ and $\mathbf{a} = -\mathbf{M}$ we have $u_0 \leq 0$ and $0 \geq u_0$, respectively, hence $u_0 = 0$. Note that we could equally well have made the choice $\boldsymbol{\rho} = \mathbf{a}$ and $\mathbf{u} = \mathbf{M}$ instead of $\boldsymbol{\rho} = \mathbf{M}$ and $\mathbf{u} = \mathbf{a}$. However, the former choice leads to inconsistencies for instance when we consider an experiment in which we place several SG magnets in succession or consider the EPRB experiment.

Thus, we have shown that only the desire to represent the data equation (3.1) such that the description of the whole experiment separates into a description of the ‘source’ (\mathbf{M}) and a description of the ‘measurement device’ (\mathbf{a}) together with some elementary linear algebra leads to the unique description of the SG experiment in terms of 2×2 matrices,

$$\hat{\rho} = \frac{(\mathbb{1} + \mathbf{M} \cdot \boldsymbol{\sigma})}{2} \quad \text{and} \quad \hat{X} = \mathbf{a} \cdot \boldsymbol{\sigma}, \quad (3.6)$$

conditional on the assumptions that the individual outcomes of an SG experiment are independent and that the frequency distribution of these outcomes does not depend on the orientation of the reference frame. From equation (3.6), it follows immediately that $\hat{\rho}^2 = \hat{\rho}$, that is $\hat{\rho}$ is a projection. This implies that we can write [46]

$$\hat{\rho} = |\Psi\rangle\langle\Psi|, \quad |\Psi\rangle = a_{\uparrow}|\uparrow\rangle + a_{\downarrow}|\downarrow\rangle, \quad (3.7)$$

where the vector $|\Psi\rangle$ is expressed in the basis of the eigenstates ($|\uparrow\rangle, |\downarrow\rangle$) of the σ^z matrix.

Summarizing: changing the representation of the data in combination with the desire to separate as much as possible the description of the source and measurement devices automatically enforces the Hilbert space structure that is a characteristic signature of quantum theory [23]. No postulates of quantum theory are required to derive (or postulate) equations (3.6) or (3.7). Furthermore, it is straightforward to extend the description to include mixed states [23].

There is nothing that forbids an experiment to yield for instance $f(x|\mathbf{a} \cdot \langle \mathbf{M} \rangle, Z) = (1 + x(\mathbf{a} \cdot \mathbf{M})^k)/2$ with $k = 2$ (we certainly can generate such data using a digital computer, a metaphor of a physical device on which we carry out experiments). However, the data produced by such an experiment cannot be represented by equation (3.6). In other words, the class of conceivable SG experiments is significantly larger than the class of experiments that allows for the separation: the class of realizable SG experiments is (much) larger than the class of SG experiments describable by quantum theory.

The fact that the separation procedure leads, in such a simple manner, to the quantum theoretical description (3.6) of the SG experiment provokes the question: ‘What is so special about the case in which the separation procedure can be carried out?’ The answer is given in §2. Using equation (3.6), according to the postulate of quantum theory [46], the probability to observe an event x is given by $P(x) = \text{Tr} \hat{\rho}(\mathbb{1} + x\mathbf{a} \cdot \boldsymbol{\sigma})/2 = (1 + x\mathbf{a} \cdot \mathbf{M})/2$, which is exactly the same expression as the one obtained by LI treatment of a robust SG experiment. In other words, if the SG experiment is robust, it may be equally well described by quantum theory.

The application of the separation procedure to the EPRB experiment is an almost trivial extension of the application to the SG experiment. We start by writing the observations $xy = (+1, -1, +1, -1)$ and frequencies $(f(+1, +1|\theta, Z), f(-1, +1|\theta, Z), f(+1, -1|\theta, Z), f(-1, -1|\theta, Z))$ as 4×4 matrices X, Y and F with elements $X([x, y], [x', y']) = x\delta_{x,x'}\delta_{y,y'}$, $Y([x, y], [x', y']) = y\delta_{x,x'}\delta_{y,y'}$ and $F([x, y], [x', y']) = f(xy|\mathbf{a}, \mathbf{M}, Z)\delta_{x,x'}\delta_{y,y'}$, respectively. Here we use the notation $[x, y] = (1-x)/2 + (1-y)$ to indicate that the pairs (x, y) and (x', y') specify the row, respectively, the column index (running from 0 to 3) of the matrices X and F . We search for 4×4 matrices $\hat{\rho}, \hat{X}$ and \hat{Y} which satisfy

$$\text{Tr} \hat{\rho} = 1, \quad \text{Tr} \hat{\rho} \hat{X} = \langle x \rangle, \quad \text{Tr} \hat{\rho} \hat{Y} = \langle y \rangle \quad \text{and} \quad \text{Tr} \hat{\rho} \hat{X} \hat{Y} = \langle xy \rangle, \quad (3.8)$$

and allow for the desired separation. Using the direct product of the Pauli-spin matrices $\sigma_j = (\sigma_j^x, \sigma_j^y, \sigma_j^z)$ for $j = 1, 2$ and the unit matrix $\mathbb{1}$ as the orthonormal basis set for the vector space of 4×4 matrices, we may write (without loss of generality) $\hat{\rho} = \rho_0 \mathbb{1} + \rho_1 \cdot \sigma_1 \otimes \mathbb{1}_2 + \mathbb{1}_1 \otimes \rho_2 \cdot \sigma_2 +$

$\sigma_1 \cdot \rho_{12} \cdot \sigma_2$, where the number ρ_0 , the vectors ρ_j and the matrix ρ_{12} are all real-valued. As each of the two sides of the EPRB experiment contains an SG magnet, consistency with the separated description of the SG experiment demands that we choose $\hat{X} = \mathbf{a}_1 \cdot \sigma_1 \otimes \mathbb{1}_2$ and $Y = \mathbb{1}_1 \otimes \mathbf{a}_2 \cdot \sigma_2$. We find the explicit expression of $\hat{\rho}$ by requiring that equation (3.8) holds. Focusing on the case of the EPRB experiment for which $\langle x \rangle = \langle y \rangle = 0$ and $\langle xy \rangle = -\mathbf{a}_1 \cdot \mathbf{a}_2$, it follows that $\rho_0 = \frac{1}{4}$, $\rho_1 = \rho_2 = 0$ and that $\hat{\rho}$ takes the form [23]

$$\hat{\rho} = \frac{\mathbb{1} - \sigma_1 \cdot \sigma_2}{4}. \quad (3.9)$$

It is not difficult to verify that $\hat{\rho}^2 = \hat{\rho}$, hence equation (3.9) is the density matrix of a pure quantum state [46]. Computing the matrix elements of $\hat{\rho}$ in the spin-up, spin-down basis of both spins, we find

$$\hat{\rho} = \left(\frac{|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle}{\sqrt{2}} \right) \left(\frac{\langle\uparrow\downarrow| - \langle\downarrow\uparrow|}{\sqrt{2}} \right) \quad (3.10)$$

and

$$\langle xy \rangle = \text{Tr} \hat{\rho} \hat{X} \hat{Y} = \text{Tr} \hat{\rho} \mathbf{a}_1 \cdot \sigma_1 \mathbf{a}_2 \cdot \sigma_2 = -\mathbf{a}_1 \cdot \mathbf{a}_2, \quad (3.11)$$

which we recognize as the quantum theoretical description of two spin- $\frac{1}{2}$ objects in the singlet state. Therefore, we have shown that rewriting the data gathered in an ideal EPRB thought experiment in a manner that allows for the envisaged separation naturally leads, without invoking postulates of quantum theory and/or probability theory, to the quantum theoretical description of two $S = \frac{1}{2}$ spins in the singlet state.

As in the case of the ideal SG experiment, the representation in parts puts a severe restriction on the kind of data that we can describe, again provoking the question: ‘What is so special about the case in which the separation procedure can be carried out?’ The answer is the same as in the case of the SG experiment: it is precisely for the special case of the robust EPRB experiment.

(b) Equivalence with quadratic forms

In the case of SG or EPRB experiment, the LI approach yields equations for the i-probs which are easy to solve directly. As the i-probs describe the data produced by robust experiments, the connection to the quantum formalism is mainly of pedagogical interest. However, if the equations for the i-probs are nonlinear, as in the case of a particle in a potential discussed in §2c, and not easy to solve, it is expedient to search for alternative equations that are much easier to solve. Fortunately, in the case at hand, we can make good use of the large body of work that explores mathematically equivalent forms of quantum theory.

Consider the quadratic functional

$$Q = \int dx \int dt \left[2im\sqrt{\lambda} \left(\psi \frac{\partial \psi^*}{\partial t} - \psi^* \frac{\partial \psi}{\partial t} \right) + 4 \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x} + 2m\lambda V(x, t) \psi^* \psi \right], \quad (3.12)$$

with the shorthand notation $\psi \equiv \psi(x | X(t), t, Z)$. Substitute $\psi = \sqrt{P(x | X(t), t, Z)} e^{iS(x, t)\sqrt{\lambda}/2}$ to obtain equation (2.18), demonstrating that equation (2.18) and equation (3.12) are equivalent (the ambiguity in the phase of ψ can be shown to be irrelevant [22]). On the other hand, the extrema of equation (3.12) are given by the solution of the linear partial differential equation

$$\frac{2i}{\sqrt{\lambda}} \frac{\partial \psi}{\partial t} = - \frac{2}{m\lambda} \frac{\partial^2 \psi}{\partial x^2} + V(x, t) \psi, \quad (3.13)$$

which turns into the time-dependent Schrödinger equation if we set $\lambda = 4/\hbar^2$.

From our derivation of equation (3.13) from LI principles, it is clear that (i) the actual value of λ can only be determined by comparing the outcome of calculations based on equation (2.18) or equation (3.13) with experimental data and that (ii) the wavefunction $\psi(x | X(t), t, Z)$ is just a mathematical concept, a vehicle to solve a class of complicated nonlinear minimization problems through the minimization of quadratic forms. As a product of human imagination, this concept is

an extraordinarily useful tool that serves no purpose other than transforming nonlinear equations into linear ones.

4. Conclusion

Using the simplest, non-trivial examples, it was shown how the application of LI to experiments for which the observed events are independent and for which the frequency distribution of these events is robust with respect to small changes of the conditions under which the experiments are carried out yields, without introducing any concept of quantum theory, some of the most basic equations of quantum theory. More extensive discussions of applications to the time-(in)dependent phenomena with or without spin can be found elsewhere [21,22,41]. Work to include relativistic effects is in progress [55].

The key point of an LI application to quantum physics experiments is to express precisely and unambiguously, using the mathematical framework of plausible reasoning [24–28], the conditions of a robust experiment; see §2. This translates into a global optimization problem for the i-prob, the solution of which may be very simple as in the case of the SG and EPRB experiment or may yield a fairly complicated nonlinear set of equations. The mathematical machinery of quantum theory appears as a result of transforming a set of nonlinear equations into a set of linear ones or emerges from the desire to separate the description into various parts.

It will not have escaped the reader that the LI approach reviewed in the present paper is void of postulates regarding ‘wavefunctions’, ‘observables’, ‘quantization rules’, ‘quantum measurements’ [56], ‘Born’s rule’, etc., nor that there are ‘interpretational’ issues. This is a direct consequence of the basic premise of the LI approach, namely that current scientific knowledge derives, through cognitive processes in the human brain, from the discrete events which are observed in laboratory experiments and from relationships between those events that we, humans, discover. These discrete events are not ‘generated’ according to certain quantum laws: instead these laws appear as the result of (the best) LI from the data.

This viewpoint seems completely in line with Bohr’s view [34]: ‘Physics is to be regarded not so much as the study of something a priori given, but rather as the development of methods of ordering and surveying human experience. In this respect, our task must be to account for such experience in a manner independent of individual subjective judgment and therefore objective in the sense that it can be unambiguously communicated in ordinary human language.’ This, in our opinion, is exactly what the LI approach allows us to do. The extraordinary descriptive power of quantum theory then follows from the fact that it is plausible reasoning, that is common sense, applied to robust experiments.

From our LI derivations of some of the most basic equations of quantum theory, it follows that the latter describes only robust experiments. This is best illustrated by comparing the (high) accuracy by which quantum theory predicts, say, the ratios of the wavelengths of the Balmer absorption/emission lines of hydrogen [57] with the comparably low accuracy of say EPRB experiments that purport to provide evidence for the singlet state of two spin- $\frac{1}{2}$ particles [39,50]. In the former case, the high accuracy originates from doing a massive amount of experiments on a very large collection of identical atoms and, as in any statistical experiment, what we observe most of the time is the most robust response. Thus, the solution of the LI problem (e.g. the Schrödinger equation in the case at hand) is the one that is ‘observed’ most frequently. By contrast, in experiments that provide data on an event-by-event basis, the statistical samples are much smaller and the external conditions may vary significantly from one experiment to the next. In other words, these experiments are not as robust as the spectroscopic experiments. From the LI viewpoint, it is therefore natural that these experiments produce data that show (much) larger deviations from the quantum-theoretical prediction than spectroscopic data. By the same argument, the LI approach offers a rational explanation for the observation that it seems to take considerable effort to engineer nanoscale devices that operate in a regime such that the experimental data comply with quantum theory.

Authors' contributions. All authors contributed to the material presented in and writing of this manuscript.

Funding. M.I.K. acknowledges financial support by the European Research Council, project 338957 FEMTO/NANO.

Competing interests. The authors have no competing interests.

References

- Bohm D. 1952 A suggested interpretation of the quantum theory in terms of 'hidden' variables. I. *Phys. Rev.* **85**, 166–179. (doi:10.1103/PhysRev.85.166)
- de la Peña L, Cetto AM. 1996 *The quantum dice: an introduction to stochastic electrodynamics*. Dordrecht, The Netherlands: Kluwer.
- Ballentine LE. 1970 The statistical interpretation of quantum mechanics. *Rev. Mod. Phys.* **42**, 358–381. (doi:10.1103/RevModPhys.42.358)
- Home D. 1997 *Conceptual foundations of quantum physics*. New York, NY: Plenum Press.
- Ballentine LE. 2001 Interpretations of probability and quantum theory. In *Foundations of probability and physics* (ed. A Khrennikov), pp. 71–84. Singapore: World Scientific.
- Caves CM, Fuchs CA, Schack R. 2007 Subjective probability and quantum certainty. *Stud. Hist. Philos. Sci. B* **38**, 255–274. (doi:10.1016/j.shpsb.2006.10.007)
- Bub J. 2007 Quantum probabilities as degrees of belief. *Stud. Hist. Philos. Sci. B* **38**, 232–254. (doi:10.1016/j.shpsb.2006.09.002)
- Fuchs CA, Schack R. 2013 Quantum-Bayesian coherence. *Rev. Mod. Phys.* **85**, 1693–1715. (doi:10.1103/RevModPhys.85.1693)
- Khrennikov AY. 2005 On the representation of contextual probabilistic dynamics in the complex Hilbert space: linear and nonlinear evolutions, Schrödinger dynamics. *Nuovo Cimento* **120**, 353–366. (doi:10.1393/ncb/i2004-10142-6)
- Hardy L. 2007 Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure. *J. Phys. A: Math. Theor.* **40**, 3081–3099. (doi:10.1088/1751-8113/40/12/S12)
- Khrennikov AY. 2009 *Contextual approach to quantum formalism*. Berlin, Germany: Springer.
- Khrennikov A. 2010 *Ubiquitous quantum structure: from psychology to finance*. Berlin, Germany: Springer.
- Chiribella G, D'Ariano GM, Perinotti P. 2010 Probabilistic theories with purification. *Phys. Rev. A* **81**, 062348. (doi:10.1103/PhysRevA.81.062348)
- Chiribella G, D'Ariano GM, Perinotti P. 2011 Informational derivation of quantum theory. *Phys. Rev. A* **84**, 012311. (doi:10.1103/PhysRevA.84.012311)
- Brukner Č. 2011 Questioning the rules of the game. *Physics* **4**, 55. (doi:10.1103/Physics.4.55)
- Masanes L, Müller MP. 2011 A derivation of quantum theory from physical requirements. *New J. Phys.* **13**, 063001. (doi:10.1088/1367-2630/13/6/063001)
- Oreshkov O, Costa F, Brukner Č. 2012 Quantum correlations with no causal order. *Nat. Commun.* **3**, 1–8. (doi:10.1038/ncomms2076)
- Klein U. 2010 The statistical origins of quantum mechanics. *Phys. Res. Int.* **2010**, 808424. (doi:10.1155/2010/808424)
- Kapustin A. 2013 Is quantum mechanics exact? *J. Math. Phys.* **54**, 062017. (doi:10.1063/1.4811217)
- de la Peña L, Cetto AM, Valdes-Hernandes A. 2015 *The emerging quantum: the physics behind quantum mechanics*. Cham, Switzerland: Springer International.
- De Raedt H, Katsnelson MI, Michielsen K. 2013 Quantum theory as the most robust description of reproducible experiments: application to a rigid linear rotator. *Proc. SPIE* **8832**, 883212–1–11. (doi:10.1117/12.2026998)
- De Raedt H, Katsnelson MI, Michielsen K. 2014 Quantum theory as the most robust description of reproducible experiments. *Ann. Phys.* **347**, 45–73. (doi:10.1016/j.aop.2014.04.021)
- De Raedt H, Katsnelson MI, Donker HC, Michielsen K. 2015 Quantum theory as a description of robust experiments: derivation of the Pauli equation. *Ann. Phys.* **359**, 166–186. (doi:10.1016/j.aop.2015.04.017)
- Cox RT. 1946 Probability, frequency and reasonable expectation. *Am. J. Phys.* **14**, 1–13. (doi:10.1119/1.1990764)

25. Cox RT. 1961 *The algebra of probable inference*. Baltimore, MD: Johns Hopkins University Press.
26. Tribus M. 1999 *Rational descriptions, decisions and designs*. Stockholm, Sweden: Expira Press.
27. Smith CR, Erickson G. 1989 From rationality and consistency to Bayesian probability. In *Maximum entropy and Bayesian methods* (ed. J Skilling), pp. 29–44. Dordrecht, The Netherlands: Kluwer Academic.
28. Jaynes ET. 2003 *Probability theory: the logic of science*. Cambridge, UK: Cambridge University Press.
29. Shannon CE, Weaver W. 1949 *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
30. Jaynes ET. 1957 Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–640. (doi:10.1103/PhysRev.106.620)
31. Jaynes ET. 1957 Information theory and statistical mechanics. II. *Phys. Rev.* **108**, 171–190. (doi:10.1103/PhysRev.108.171)
32. Caticha A. 2014 Entropic dynamics: an inference approach to quantum theory, time and measurement. *J. Phys.: Conf. Series* **504**, 012009. (doi:10.1088/1742-6596/504/1/012009)
33. Caticha A. 2015 Entropic dynamics. *Entropy* **17**, 6110–6128. (doi:10.3390/e17096110)
34. Bohr N. 1961 [1999] XV. The unity of human knowledge. In *Niels Bohr collected works*, vol. 10, *Complementarity beyond physics (1928–1962)* (ed. D Favrholt), pp. 155–160. Amsterdam, The Netherlands: Elsevier [reprint 1999]. (doi:10.1016/S1876-0503(08)70208-1)
35. Pearl J. 2000 *Causality: models, reasoning, and inference*. Cambridge, UK: Cambridge University Press.
36. 't Hooft G. 2007 The mathematical basis for deterministic quantum mechanics. In *Beyond the quantum* (eds TM Nieuwenhuizen, B Mehmami, V Špička, MJ Aghdami, AY Khrennikov), pp. 3–19. Singapore: World Scientific.
37. De Raedt K, De Raedt H, Michielsen K. 2005 Deterministic event-based simulation of quantum interference. *Comput. Phys. Commun.* **171**, 19–39. (doi:10.1016/j.cpc.2005.04.012)
38. Michielsen K, Jin F, De Raedt H. 2011 Event-based corpuscular model for quantum optics experiments. *J. Comput. Theor. Nanosci.* **8**, 1052–1080. (doi:10.1166/jctn.2011.1783)
39. De Raedt H, Michielsen K. 2012 Event-by-event simulation of quantum phenomena. *Ann. Phys. (Berlin)* **524**, 393–410. (doi:10.1002/andp.201100299)
40. De Raedt H, Jin F, Michielsen K. 2012 Event-based simulation of neutron interferometry experiments. *Quantum Matter* **1**, 1–21. (doi:10.1166/qm.2012.1001)
41. De Raedt H, Michielsen K. 2014 Discrete-event simulation of uncertainty in single-neutron experiments. *Front. Phys.* **2**, 14.1–14.12. (doi:10.3389/fphy.2014.00014)
42. Michielsen K, De Raedt H. 2014 Event-based simulation of quantum physics experiments. *Int. J. Mod. Phys. C* **25**, 01430003. (doi:10.1142/S0129183114300036)
43. Klein MJ. 1967 Thermodynamics in Einstein's thought. *Science* **157**, 509–516. (doi:10.1126/science.157.3788.509)
44. Pólya G 1954 *Mathematics and plausible reasoning*. Princeton, NJ: Princeton University Press.
45. Bohm D. 1951 *Quantum theory*. New York, NY: Prentice-Hall.
46. Ballentine LE. 2003 *Quantum mechanics: a modern development*. Singapore: World Scientific.
47. Weihs G. 2000 Ein Experiment zum Test der Bellschen Ungleichung unter Einsteinscher Lokalität. PhD thesis, University of Vienna. See <http://www.uibk.ac.at/exphys/phonik/people/gwdiss.pdf>.
48. Shih Y. 2011 *An introduction to quantum optics: photon and biphoton physics*. Boca Raton, FL: CRC Press.
49. Vistnes A, Adenier G. 2012 There may be more to entangled photon experiments than we have appreciated so far. *AIP Conf. Proc.* **1508**, 326–333. (doi:10.1063/1.4773143)
50. De Raedt H, Jin F, Michielsen K. 2013 Data analysis of Einstein–Podolsky–Rosen–Bohm laboratory experiments. *Proc. SPIE* **8832**, 88321N. (doi:10.1117/12.2021860)
51. Hess K. 2015 *Einstein was right!* Singapore: Pan Stanford.
52. Ralston JP. 2013 Emergent mechanics, quantum and un-quantum. *Proc. SPIE* **8832**, 88320W. (doi:10.1117/12.2025000)
53. Madelung E. 1927 Quantentheorie in hydrodynamischer Form. *Z. Phys.* **40**, 322–326. (doi:10.1007/BF01400372)

54. De Raedt H, Katsnelson MI, Donker HC, Michielsen K. 2015 Quantum theory as a description of robust experiments: application to Stern–Gerlach and Einstein–Podolsky–Rosen–Bohm experiments. *Proc. SPIE* **9570**, 957002. (doi:10.1117/12.2185704)
55. Donker H, Katsnelson M, De Raedt H, Michielsen K. In preparation. Logical inference approach to relativistic quantum mechanics: derivation of the Klein–Gordon equation.
56. Allahverdyan AE, Balian R, Nieuwenhuizen TM. 2013 Understanding quantum measurement from the solution of dynamical models. *Phys. Rep.* **525**, 1–166. (doi:10.1016/j.physrep.2012.11.001)
57. Mohr PJ, Taylor BN, Newell DB. 2012 CODATA recommended values of the fundamental physical constants: 2010*. *Rev. Mod. Phys.* **84**, 1527–1605. (doi:10.1103/RevModPhys.84.1527)