Quantum theory as the most robust description of reproducible experiments: Application to a rigid linear rotator

H. De Raedt[a], M.I. Katsnelson[b], and K. Michielsen[c,d]

[a] Department of Applied Physics, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands
[b] Institute of Molecules and Materials, Radboud University of Nijmegen, NL-6525 ED Nijmegen, The Netherlands
[c] Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany
[d] RWTH Aachen University, D-52056 Aachen, Germany

ABSTRACT

Experiments in which there is uncertainty about each individual event are analyzed from the viewpoint of logical inference. A quantitative criterion of a robust description of the data generated by these experiments is presented. The requirement that the description is the most robust amounts to minimizing a functional of the plausibilities for observing the events. The analysis is, by construction, practical and minimal. By way of application to the case of a rigid linear rotator, it is shown how the Schrödinger equation emerges from minimization of this functional. From the viewpoint taken in this paper, quantum theory follows from inductive inference applied to a well-defined class of experiments, without any commitment to further notions of reality.

Keywords: logical inference, quantum theory, inductive logic, probability theory

1. INTRODUCTION

In a recent paper,[1] we have demonstrated that the mathematical framework of logical inference[2–6] applied to basic experiments such as the Stern-Gerlach and Einstein-Podolsky-Rosen-Bohm (EPRB) experiments yield descriptions that we know from quantum physics. The same paper also shows that the same approach can be used to derive the time-(in)dependent Schrödinger equation for a particle in a (gauge) potential. More specifically, it was shown that the basic equations of quantum theory directly follow from logical inference applied to experiments in which there is

(i) uncertainty about individual events,

(ii) the stringent condition that certain properties of the collection of events are robust, meaning that the observed distribution of frequencies of events are insensitive with respect to small changes in the conditions under which the experiments are carried out.

In the present paper, we give another illustration of the logical-inference approach by way of deriving the Schrödinger equation of a rigid linear rotator.

The basic idea behind our approach is largely captured by the following quotes of Niels Bohr:

1. 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.[7]

Further author information: (Send correspondence to H. De Raedt)
H. De Raedt : E-mail: h.a.de.raedt@rug.nl
M.I. Katsnelson: E-mail: M.Katsnelson@science.ru.nl
K. Michielsen: E-mail: k.michielsen@fz-juelich.de

The Nature of Light: What are Photons? V, edited by Chandrasekhar Roychoudhuri,
Al F. Kracklauer, Hans De Raedt, Proc. of SPIE Vol. 8832, 883212 · © 2013 SPIE
CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2026998

2. It is wrong to think that the task of physics is to find out how nature *is*. Physics concerns what we can *say* about nature.[8]

Similar views have been expressed by other fathers of quantum mechanics, e.g., Max Born and Wolfgang Pauli.[9] Roughly speaking, the statement is that "Quantum theory describes our *knowledge* of the atomic phenomena rather than the atomic phenomena themselves". Clearly, as it stands, this statement and the mentioned quotes of Niels Bohr are not directly useful to develop insight into the question why quantum theory has proven to be extraordinary powerful to describe a vast amount of very different experiments in (sub)-atomic, molecular and condensed matter physics, quantum optics and so on. Our work shows that the philosophical components contained in these quotes can be replaced by a mathematical framework from which the basic equations of quantum theory emerge. It is not concerned with the various interpretations[10–13] of quantum theory but, like Ref. 1, explores the possibility of exploiting logical inference to give a rational explanation for the success of quantum theory.

In line with Bohr's quote (2), let us take the point of view that the aim of physics is to provide a consistent description of relations between certain events that we perceive (usually with the help of some equipment) with our senses. Some of these relations express cause followed by an effect and others do not. If there are uncertainties about the individual events and the conditions under which the experiment is carried out, situations may arise in which it becomes difficult or even impossible to establish relations between individual events. In the case that the frequencies of these events are robust (see (ii)) it may still be possible to establish relations, not between the individual events, but between the frequency distributions of the observed events. It is precisely under these circumstances that the application of logical inference to (the abstraction of) the experiments yields the basic equations of quantum theory.[1] The theoretical description becomes unambiguous and independent of the individual subjective judgement and provides the rational for the extraordinary descriptive power of quantum theory: it is plausible reasoning, that is common sense, applied to reproducible experimental data.

The algebra of logical inference facilitates reasoning in a mathematically precise language which may be unambiguous and independent of the individual. It is the foundation for powerful tools such as the maximum entropy method and Bayesian analysis.[4,6] Although not formulated in the language of logical inference used in the present paper, Jaynes' papers on the relation between information and (quantum) statistical mechanics[14,15] are perhaps the first to "derive" theoretical descriptions using this general methodology of scientific reasoning. Elsewhere, we have shown that some of the basic equations of quantum theory also derive from the application of the algebra of logical inference.[1]

The rules of logical inference are not bound by "the laws of physics". Logical inference applies to situations where there may or may not be causal relations between the events.[4,6] The point of view taken (see also Bohr's quote (1)) in this paper is that the laws of physics should provide a consistent description of relations between certain events that we perceive by our senses and therefore they should conform to the rules of logical inference. Although extracting cause-and-effect relationships from empirical evidence by rational reasoning should follow the rules of logical inference, in general the latter cannot be used to establish cause-and-effect relationships.[6,16,17]

We emphasize that the logical-inference derivation of the quantum theoretical description does not, in any way, prohibit the construction of cause-and-effect mechanisms that, when analyzed in the same manner as in real experiments, create the *impression* that the system behaves as prescribed by quantum theory.[18–20] Work in this direction, for reviews see Ref. 21–24, has shown that it is indeed possible to build simulation models which reproduce, on an event-by-event basis, the results of interference and entanglement experiments with photons and neutrons, for instance. But as long as there is uncertainty about some aspects of the experiment under scrutiny (which, in real-life situations is always the case), the general rules by which we deduce whether a proposition about the experiment is true or false can no longer be used and inductive reasoning is unavoidable. However, although discovering the logical relationships between events gives us deeper insight into natural phenomena, history has shown over and over again that these discoveries are not essential for developing new technologies. In fact, it often is the development of new technologies that allow us to gain knowledge about processes that were previously "unaccessible" to us. In other words, the development of logical-inference (epistemological) and cause-and-effect (ontological) models cannot be disentangled but great care should be taken not to view epistemological models (which only describe logical relationships between events) as substitutes for the models of the actual cause-and-effect processes.

The paper is structured as follows. Section 2 contains a brief introduction to the algebra of logical inference,[2–6] a mathematical framework[2–6] which formalizes the patterns of plausible reasoning exposed by Pólya.[25] This mathematically precise formalism expresses what most people would consider to be rational reasoning. The key concept is the notion of the plausibility that a proposition is true given that another proposition is true. In Section 3, we demonstrate that the time-independent Schrödinger equation of a linear rigid rotator can be derived by logical inference from the assumption that the experiment yields robust data and that a weak form of the correspondence principle holds. A discussion of general aspects of our approach and conclusions are given in Section 4.

2. THE ALGEBRA OF LOGICAL INFERENCE

In this section, we briefly introduce the concepts that are necessary for the purpose of the present paper. For a detailed discussion of the foundations of plausible reasoning, its relation to Boolean logic and the derivation of the rules of logical inference, the reader is advised to consult the papers[2,5] and books[3,4,6] from which our summary has been extracted. In essence, we are only concerned about quantifying the truth of a proposition given the truth of another proposition. Then, it is possible to construct a mathematical framework, an extension of Boolean logic, that allows us to reason in a manner which is unambiguous and independent of the individual, in particular if there are elements of uncertainty in the description.[2–6]

We start by listing three so-called "desiderata" from which the algebra of logical inference can be derived.[3–6] The formulation which follows is taken from Ref. 5.

Desideratum 1. *Plausibilities are represented by real numbers.* The plausibility that a proposition A is true conditional on proposition B being true will be denoted by $P(A|B)$.

Desideratum 2. *Plausibilities must exhibit agreement with rationality.* As more and more evidence supporting the truth of a proposition becomes available, the plausibility should increase monotonically and continuously and the plausibility of the negation of the proposition should decrease monotonically and continuously. Moreover, in the limiting cases that the proposition A is known to be either true or false, the plausibility $P(A|B)$ should conform to the rules of deductive reasoning. In other words, plausibilities must be in qualitative agreement with the patterns of plausible reasoning uncovered by Pólya.[25]

Desideratum 3. *All rules relating plausibilities must be consistent.* Consistency of rational reasoning demands that if the rules of logical inference allow a plausibility to be obtained in more than one way, the result should not depend on the particular sequence of operations.

These three desiderata only describe the essential features of the plausibilities and definitely do not constitute a set of axioms which plausibilities have to satisfy. It is a most remarkable fact that these desiderata suffice to uniquely determine the set of rules by which plausibilities may be manipulated.[3–6]

Omitting the derivation, it follows that plausibilities may be chosen to take numerical values in the range $[0,1]$ and obey the rules[3–6]

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. $P(AB|Z)=P(A|BZ)P(B|Z)=P(B|AZ)P(A|Z)$ where the "product" BZ denotes the logical product (conjunction) of the propositions B and Z, that is the proposition BZ is true if both B and Z are true. This rule will be referred to as "product rule". It should be mentioned here that it is not allowed to define a plausibility for a proposition conditional on the conjunction of mutual exclusive propositions. Reasoning on the basis of two of more contradictory premises is out of the scope of the present paper.
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, that is the proposition $A+B$ is true if either A or B or both are true. These two rules show that Boolean algebra is contained in the algebra of plausibilities.

The algebra of logical inference, as defined by the rules (a–c), is the foundation for powerful tools such as the maximum entropy method and Bayesian analysis.[4,6] The rules (a–c) are unique.[4–6] 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 1–3.

The rules (1–3) are identical to the rules by which we manipulate probabilities.[6,26–28] However, the rules (1–3) were not postulated. They were derived from general considerations about rational reasoning and consistency only. Moreover, concepts such as sample spaces, probability measures etc., which are an essential part of the mathematical foundation of probability theory,[27,28] play no role in the derivation of rules (1–3). In fact, if Kolmogorov's axiomatic formulation of probability theory would have been in conflict with rules (1–3), we believe that this formulation would long have been disposed of because it would yield results which are in conflict with rational reasoning. Perhaps most important in the context of quantum theory is that in the logical inference approach uncertainty about an event does not imply that this event can be represented by a random variable as defined in probability theory.[28]

There is a significant conceptual difference between "mathematical" probabilities and plausibilities. Mathematical probabilities are elements of an axiomatic framework which complies with the algebra of logical inference. Plausibilities are elements of a language which also complies with the algebra of logical inference and serve to facilitate communication, in an unambiguous and consistent manner, about phenomena in which there is uncertainty.

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.[4] In general, $P(A|B)$ may express the degree of believe of an individual that proposition A is true, given that proposition B is true. However, in the present paper, we explicitly exclude applications of this kind because they do not comply with our main goal, namely to describe phenomena "in a manner independent of individual subjective judgment", see Bohr's quote (2).

To take away this subjective connotation of the word "plausibility", from now on we will simply call $P(A|B)$ the "inference-probability" or "i-prob" for short.

A comment on the notation used throughout this paper is in order. To simplify the presentation, we make no distinction between an event such as "detector D has fired" and the corresponding proposition "$D =$ detector D has fired". If we have two detectors, say D_x where $x = \pm 1$, we write $P(x|Z)$ to denote the i-prob of the proposition that detector D_x fires, given that proposition Z is true. Similarly, the i-prob of the proposition that two detectors D_x and D_y fire, given that proposition Z is true, is denoted by $P(x, y|Z)$. Obviously, this notation generalizes to more than two propositions.

2.1 Application to quantum phenomena

The theoretical description of "classical physics" applies to phenomena for which there is absolute certainty about the outcome of each individual experiment on each individual object.[29–31] In mapping the experimental data which are necessarily represented by a limited number of bits, that is by integers, onto the theoretical abstractions in terms of real numbers, it is assumed that the necessarily finite precision of the experiment can be increased without limit, at least in principle, and that there is a one-to-one mapping between the values of the variables in the theory and the values of the corresponding quantities measured in experiment.

In real experiments there is always uncertainty about some factors which may or may not influence the outcome of the measurements. In the realm of classical physics, standard techniques of statistical analysis are used to deal with this issue. It is postulated that these "imperfections" in the experimental data are not of fundamental importance but are technical in nature and can therefore be eliminated, at least in principle.[29–31]

Quantum theory is fundamentally different from classical theories in that there may be uncertainties about each individual event, uncertainties which cannot be eliminated, not even in principle.[29–31] Clearly, this is a statement about the theory, not about the observed phenomena themselves. The outcome of a real experiment, be it on "classical" or "quantum" objects, is always subject to uncertainties in the conditions under which the experiment is carried out. However, this issue is not of direct concern to us here because we only want to explore whether the quantum theoretical description, not the phenomena themselves, can be derived from logical inference applied to certain thought experiments. Such derivations seem to have a generic structure[1]

1. List the features of the experiment that are deemed to be relevant and to introduce the i-probs of the individual events.
2. Impose the condition that the experiment yields reproducible results, not on the level of individual events, but on the level of averages of many events. The result is a functional of the i-probs of the individual events.
3. Minimize the functional to obtain equations that determine the i-probs.

The i-probs obtained by solving these equations are identical to the corresponding probability obtained from the quantum theoretical description of the experiment.

3. SCHRÖDINGER EQUATION FOR A LINEAR ROTATOR

In Ref. 1 it was shown that by simply demanding that the recorded data sets of events yields robust, reproducible results for the i-probs, logical inference leads to expressions that are known from the quantum theoretical treatment of the Stern-Gerlach and Einstein-Podolsky-Rosen-Bohm experiments and also yields the Schrödinger equation for a particle moving in a time-dependent (gauge) field. In essence, these results derive from the following ideas:

(i) The i-probs for events to occur obey the rules of the algebra of logical inference.

(ii) The i-prob to observe an event depends explicitly on a variable condition.

(iii) Maximizing the robustness of the i-prob to observe the data with respect to small variations of the condition yields the functional dependence of the i-prob on this condition.

Here we demonstrate that a straightforward application of our approach yields the Schrödinger equation of a linear rotator. The key points are to formulate precisely what it means to perform a robust, reproducible experiment and to feed in knowledge about the Newtonian dynamics of the particle.

From the viewpoint of logical inference, there are two different ways to proceed. As we already showed that the time-(in)dependent Schrödinger equation for a particle moving in 3D space can be derived by logical inference, one way would be to generalize this derivation, which is almost trivial, to the many particle case and then use the resulting many-particle Schrödinger equation as the starting point for the treatment of the linear rotator. The second approach is to apply the logical-inference procedure directly to the case of the linear rotator, without making reference to the many-particle system. As this second approach provides an independent test of the logical-inference procedure itself, we will pursue this approach in what follows.

3.1 Experiment

We consider N repetitions of a thought experiment on a diatomic molecule which is assumed to be described by a linear rigid rotor model, consisting of two point masses separated by a fixed distance R. The kinematics of the classical linear rigid rotor is completely described in terms of the spherical polar coordinates θ (zenith) and φ (azimut), the distance R, and the masses m_1 and m_2. The expression of the classical kinetic energy reads

$$K=\frac{m_1+m_2}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)+\frac{I}{2}\left(\dot{\theta}^2+\dot{\varphi}^2\sin^2\theta\right), \tag{1}$$

where (x,y,z) are the coordinates of the center of mass in the fixed reference frame and θ and φ describe the rotation of the molecule in the frame of reference with its origin at the center of mass which does not rotate with respect to the fixed frame. We also introduce the symbol $I=m_1m_2R^2/(m_1+m_2)$ as a shorthand for the moment of inertia of the linear rotator. In the following, we focus on the rotational degrees of freedom, that is we consider experiments that are carried out in the moving frame of reference. Consequently, in the remainder of this section, we omit the kinetic energy of the center of mass.

We imagine that the rotating molecule is contained in a large sphere, uniformly covered by M detectors, the position of the centers of which are given by the spherical coordinates (u_j,v_j) for $j=1,\ldots,M$. The spatial extend of the detectors determines the resolution. In a later stage, we will take the continuum limit ($M\to\infty$) to solve the problem analytically. Next, we imagine that there is a source that sends N signals, each of which solicits a response of the molecule in the form of another signal which in turn causes one, and only one of the detectors to fire. The center of the detector which fires at the nth event is given by the spherical coordinates (u'_n,v'_n) for $1\le n\le N$. Note that up to this point, there is no relation between the unknown orientation (θ,φ) of the linear rotator and the position of the detectors which fired. In fact, as will become clear later, quantum theory does not provide any information about the unknown orientation (θ,φ).

The result of N repetitions of the experiment yields the data set

$$\Upsilon=\{u'_n,v'_n|n=1,\ldots,N\}, \tag{2}$$

or, denoting the total count of detector (u,v) by $0 \le k(u,v) \le N$, the experiment produces the data set of counts

$$\mathscr{D} \quad = \quad \Big\{k(u_j,v_j)\Big\}, \tag{3}$$

where, obviously $\sum_{j=1}^{M} k(u_j,v_j) = N$.

3.2 Inference-probability of the data produced by the experiment

Following the general procedure outlined above and explained in detail in Ref. 1, we introduce the i-prob $P(u,v|\theta,\varphi,Z)$ to describe the relation between the unknown orientation (θ,φ) of the linear rotator and the position (u,v) of the detector which fired. Except for the unknown orientation (θ,φ), all other experimental conditions are represented by Z and are assumed to be fixed and identical for all experiments. Unlike in the standard case of parameter estimation, in the case at hand both $P(u,v|\theta,\varphi,Z)$ and the "parameters" (θ,φ) are unknown. The key question is what the requirement of robustness tells us about the i-prob $P(u,v|\theta,\varphi,Z)$ as a function of (θ,φ).

First we list some plausible assumptions which, in essence, are the same as those listed in Ref. 1. We repeat them here with the minor modifications that are necessary to deal with rotational instead of translational motion.

1. It is assumed that each repetition of the experiment represents an identical event of which the outcome is logically independent of any other such event. By application of the product rule, the direct consequence of this assumption is that

$$P(\mathscr{D}|\theta,\varphi,N,Z) = N! \prod_{j=1}^{M} \frac{P(u_j,v_j|\theta,\varphi,Z)^{k(u_j,v_j)}}{k(u_j,v_j)!}. \tag{4}$$

   In probability theory, events with these properties are called Bernoulli trials, a concept that is central to many results in probability theory.[4,6,28]

2. We assume that space is isotropic, meaning that we do not expect the frequency distributions to change significantly if we rotate both the linear molecule and the detector array in the same manner. Denoting the unit vector in the (θ,φ) and (u_j,v_j) direction by $\boldsymbol{n}$ and $\boldsymbol{m}$, respectively, this assumed rotational invariance implies that the i-prob has the property

$$P(\mathscr{R}\boldsymbol{m}|\mathscr{R}\boldsymbol{n},Z) = P(\boldsymbol{m}|\boldsymbol{n},Z) = P(u_j,v_j|\theta,\varphi,Z), \tag{5}$$

   where $\mathscr{R}$ denotes an arbitrary rotation in three-dimensional space.

3.3 Condition for robustness

Although the data set Eq. (3) may change from run to run, we may expect that the frequencies with which the detectors fire exhibit some kind of robustness, smoothness with respect to small changes of the unknown values of (θ,φ). If this were not the case, these numbers would vary erratically with (θ,φ), most likely the results would be called "irreproducible", and the experimental data would be disposed of because repeating the run with slightly different values of (θ,φ) would often produce results that are very different from those of other runs.

Obviously, the important feature of robustness with respect to small variations of the conditions under which the experiment is carried out should be reflected in the expression for the i-prob to observe data sets which yield reproducible averages and correlations (with the usual statistical fluctuations). The next step therefore is to determine the expression for $P(u,v|\theta,\varphi,Z)$ which is most insensitive to small changes in (θ,φ).

It is expedient to formulate this question as an hypothesis test. Let H_0 and H_1 be the hypothesis that the data $\mathscr{D}$ is observed for (θ,φ) and $(\theta+\varepsilon_\theta,\varphi+\varepsilon_\varphi)$, respectively. The evidence Ev of hypothesis H_1, relative to hypothesis H_0, is defined by[4,6]

$$\mathrm{Ev} = \ln \frac{P(\mathscr{D}|\theta+\varepsilon_\theta,\varphi+\varepsilon_\varphi,N,Z)}{P(\mathscr{D}|\theta,\varphi,N,Z)}, \tag{6}$$

where the logarithm serves to facilitate the algebraic manipulations. If H_1 is more (less) plausible than H_0 then $\mathrm{Ev} > 0$ ($\mathrm{Ev} < 0$). In statistics, the r.h.s. of Eq. (6) is known as the log-likelihood function and used for parameter estimation. In contrast, in the present context, the function Eq. (6) is *not* used to estimate (θ,φ) but as will be shown below, is a vehicle to express the robustness with respect to the coordinates (θ,φ).

Making use of Eq. (4) and the assumption that the changes ε_θ and ε_φ of the conditions are small, we find

$$\begin{aligned}\mathrm{Ev} &= \sum_{j=1}^{M} k(u_j,v_j) \ln \frac{P(u_j,v_j|\theta+\varepsilon_\theta,\varphi+\varepsilon_\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \\ &= \sum_{j=1}^{M} k(u_j,v_j) \ln \left\{ 1+\varepsilon_\theta \frac{P_\theta(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{\varepsilon_\varphi}{\sin\theta}\frac{P_\varphi(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \right. \\ &\quad \left. +\varepsilon_\theta^2 \frac{P_{\theta\theta}(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + 2\frac{\varepsilon_\theta\varepsilon_\varphi}{\sin\theta}\frac{P_{\theta\varphi}(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{\varepsilon_\varphi^2}{\sin^2\theta}\frac{P_{\varphi\varphi}(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \dots \right\}, \end{aligned} \tag{7}$$

where the subscripts θ and φ denote differentiation with respect to θ and φ, respectively. Expanding the logarithm in Taylor series and retaining terms to second order only yields

$$\begin{aligned}\mathrm{Ev} &= \sum_{j=1}^{M} k(u_j,v_j) \left\{ \varepsilon_\theta \frac{P_\theta(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{\varepsilon_\varphi}{\sin\theta}\frac{P_\varphi(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \right\} \\ &\quad + \sum_{j=1}^{M} k(u_j,v_j) \left\{ \varepsilon_\theta^2 \frac{P_{\theta\theta}(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + 2\frac{\varepsilon_\theta\varepsilon_\varphi}{\sin\theta}\frac{P_{\theta\varphi}(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{\varepsilon_\varphi^2}{\sin^2\theta}\frac{P_{\varphi\varphi}(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \right\} \\ &\quad + \sum_{j=1}^{M} k(u_j,v_j) \left\{ \varepsilon_\theta \frac{P_\theta(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{\varepsilon_\varphi}{\sin\theta}\frac{P_\varphi(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \right\}^2 . \end{aligned} \tag{8}$$

According to our criterion of robustness, the evidence Eq. (8) should change as little as possible as θ and φ varies. From Eq. (8), it is clear this can be accomplished by minimizing, in absolute value, all the coefficients of the polynomial in ε_θ and ε_φ.

The first and second sum in Eq. (8) vanish identically if we choose $k(u_j,v_j)/N = P(u_j,v_j|\theta,\varphi,Z)$. Although this choice is motivated by the desire to eliminate contributions of order ε, it follows that our criterion of robustness suggests the intuitively obvious procedure to assign to $P(u_j,v_j|\theta,\varphi,Z)$ the value of the observed frequencies of occurrences $k(u_j,v_j)/N$.[4,6] The remaining contribution

$$\mathrm{Ev} = \sum_{j=1}^{M} \left\{ \varepsilon_\theta \frac{P_\theta(u_j,v_j|\theta,\varphi,Z)}{\sqrt{P(u_j,v_j|\theta,\varphi,Z)}} + \frac{\varepsilon_\varphi}{\sin\theta}\frac{P_\varphi(u_j,v_j|\theta,\varphi,Z)}{\sqrt{P(u_j,v_j|\theta,\varphi,Z)}} \right\}^2 . \tag{9}$$

vanishes identically if and only if $P_\theta(u_j,v_j|\theta,\varphi,Z)=0$ and $P_\varphi(u_j,v_j|\theta,\varphi,Z)=0$, implying that $P(u_j,v_j|\theta,\varphi,Z)$ can only describe experiments for which the data does not exhibit any dependence on θ or φ.

> Experiments which produce results that do not change with the conditions do not increase our knowledge about the relation between the conditions and the observed data. Therefore, in this paper, we explicitly exclude such experiments.

Thus, from now on, we explicitly exclude the class of experiments for which $P_\theta(u_j,v_j|\theta,\varphi,Z)=0$ and $P_\varphi(u_j,v_j|\theta,\varphi,Z)=0$.

The condition for robustness requires that we reduce the dependence of Ev on ε_θ and ε_φ as much as possible. This problem can be brought in a tractable form by invoking the Cauchy-Schwarz inequality, yielding

$$\mathrm{Ev} \le (\varepsilon_\theta^2+\varepsilon_\varphi^2) \sum_{j=1}^{M} \left\{ \frac{P_\theta^2(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{1}{\sin^2\theta}\frac{P_\varphi^2(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \right\} . \tag{10}$$

As ε_θ and ε_φ are arbitrary, it directly follows from Eq. (10) that minimizing the second factor in Eq. (10) is the best we can do to make sure that Ev is as small as possible. Therefore, we conclude that in order to impose the condition of robustness, we have to minimize the Fisher information

$$I_F = \sum_{j=1}^{M} \left\{ \frac{P_\theta^2(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} + \frac{1}{\sin^2\theta} \frac{P_\varphi^2(u_j,v_j|\theta,\varphi,Z)}{P(u_j,v_j|\theta,\varphi,Z)} \right\}, \tag{11}$$

subject to additional constraints that we impose (see below).

At this point, to simplify the mathematics, it is expedient to take the continuum limit, replacing sums by integrals over the spherical coordinates. As a result, the expression to be minimized reads

$$I_F = \int_0^{2\pi} dv \int_0^{\pi} \sin u\, du \left\{ \frac{P_\theta^2(u,v|\theta,\varphi,Z)}{P(u,v|\theta,\varphi,Z)} + \frac{1}{\sin^2\theta} \frac{P_\varphi^2(u,v|\theta,\varphi,Z)}{P(u,v|\theta,\varphi,Z)} \right\}. \tag{12}$$

The next step is to exploit rotational invariance, see Eq. (5), to replace the derivatives with respect to θ and φ by derivatives with respect to u and v. From Eq. (5), it follows that

$$P_\theta(u,v|\theta,\varphi,Z) = \frac{\partial P(u,v|\theta,\varphi,Z)}{\partial\theta} = -\frac{\partial P(u,v|\theta,\varphi,Z)}{\partial u} = -P_u(u,v|\theta,\varphi,Z), \tag{13}$$

and

$$\frac{1}{\sin\theta} P_\varphi(u,v|\theta,\varphi,Z) = \frac{1}{\sin\theta} \frac{\partial P(u,v|\theta,\varphi,Z)}{\partial\varphi} = -\frac{1}{\sin u} \frac{\partial P(u,v|\theta,\varphi,Z)}{\partial v} = -\frac{1}{\sin u} P_v(u,v|\theta,\varphi,Z), \tag{14}$$

and the expression to be minimized reads

$$I_F = \int_0^{2\pi} dv \int_0^{\pi} \sin u\, du \left\{ \frac{P_u^2(u,v|\theta,\varphi,Z)}{P(u,v|\theta,\varphi,Z)} + \frac{1}{\sin^2 u} \frac{P_v^2(u,v|\theta,\varphi,Z)}{P(u,v|\theta,\varphi,Z)} \right\}, \tag{15}$$

subject to the knowledge that the motion of the classical linear rotator is described by its classical kinetic energy Eq. (1), see below. Note that the "true" coordinates of the rotator, represented by (θ,φ) are unknown and remain unknown in what follows.

We include the knowledge that the motion of the classical linear rotator is described by its classical kinetic energy Eq. (1) by repeating the steps outlined in Ref. 1. First, we recall from classical Hamiltonian mechanics that the time-independent Hamilton-Jacobi equation for a linear rotator reads

$$\frac{1}{2I} \left[\left(\frac{\partial S(\theta,\varphi)}{\partial\theta} \right)^2 + \frac{1}{\sin^2\theta} \left(\frac{\partial S(\theta,\varphi)}{\partial\varphi} \right)^2 \right] = E, \tag{16}$$

where $S(\theta,\varphi)$ is the action and E denotes the rotational energy of the linear rotator. Next, we consider the case where there is uncertainty about (θ,φ) but not about the individual event, meaning that we assume (momentarily) that $P(u,v|\theta,\varphi,Z)$ is known and replace Eq. (16) by

$$\int_0^{2\pi} dv \int_0^{\pi} \sin u\, du \left[\left(\frac{\partial S(u,v)}{\partial u} \right)^2 + \frac{1}{\sin^2 u} \left(\frac{\partial S(u,v)}{\partial v} \right)^2 - 2IE \right] P(u,v|\theta,\varphi,Z) = 0, \tag{17}$$

In the limit that there is absolute certainty, that is if $P(u,v|\theta,\varphi,Z) \to \delta(u-\theta)\delta(v-\varphi)$, Eq. (17) reduces to Eq. (16). In a sense, Eq. (17) is reminiscent of the complementarity principle in quantum theory: it expresses the idea that if $P(u,v|\theta,\varphi,Z) \to \delta(u-\theta)\delta(v-\varphi)$, we want to recover classical Hamiltonian mechanics.

Finally, we consider the case that there is uncertainty in both the individual event and the conditions Then, the i-prob $P(u,v|\theta,\varphi,Z)$ is unknown but can be determined by requiring that the frequency distributions of the observed events are robust with respect to small changes in the conditions (the unknown coordinates θ and φ in the case at hand). Minimizing the Fisher information Eq. (15) with the constraint Eq. (17) amounts to minimizing the functional

$$F(\theta,\varphi)=\int_0^{2\pi}dv\int_0^{\pi}\sin u\,du\,\left\{\frac{1}{P(u,v|\theta,\varphi,Z)}\left(\frac{\partial P(u,v|\theta,\varphi,Z)}{\partial u}\right)^2+\frac{1}{\sin^2 u}\frac{1}{P(u,v|\theta,\varphi,Z)}\left(\frac{\partial P(u,v|\theta,\varphi,Z)}{\partial v}\right)^2\right.$$
$$\left.+\lambda\left[\left(\frac{\partial S(u,v)}{\partial u}\right)^2+\frac{1}{\sin^2 u}\left(\frac{\partial S(u,v)}{\partial v}\right)^2-2IE\right]P(u,v|\theta,\varphi,Z)\right\}, \tag{18}$$

where λ is a Lagrange parameter. The normalization of $P(u,v|\theta,\varphi,Z)$ can be taken care of by exploiting the invariance of the extrema of Eq. (18) with respect to the rescaling transformation $P(u,v|\theta,\varphi,Z)\to\alpha P(u,v|\theta,\varphi,Z)$. Note that in searching for the functions $P(u,v|\theta,\varphi,Z)$ and $S(u,v)$ which minimize Eq. (18), we explicitly exclude solutions for which $P_u(u,v|\theta,\varphi,Z)=0$ and $P_v(u,v|\theta,\varphi,Z)=0$, see above.

The extrema of F can be found by a standard variational calculation. The resulting differential equations are nonlinear in $P(u,v|\theta,\varphi,Z)$ and $S(u,v)$ and we do not know how to solve them directly. However, in analogy with Madelung's hydrodynamic-like formulation[32] or Bohm's interpretation[33] of quantum theory it follows that the extrema (and therefore also the minima) of Eq. (18) can be found by solving a time-independent Schrödinger equation, as we now show with a minimum of algebra.

We start from the functional

$$Q(\theta,\varphi)=\int_0^{2\pi}dv\int_0^{\pi}\sin u\,du\,\left(4\frac{\partial\psi^*(u,v|\theta,\varphi,Z)}{\partial u}\frac{\partial\psi(u,v|\theta,\varphi,Z)}{\partial u}+\frac{4}{\sin^2 u}\frac{\partial\psi^*(u,v|\theta,\varphi,Z)}{\partial v}\frac{\partial\psi(u,v|\theta,\varphi,Z)}{\partial v}\right.$$
$$\left.-2I\lambda E\,\psi^*(u,v|\theta,\varphi,Z)\psi(u,v|\theta,\varphi,Z)\right), \tag{19}$$

and substitute

$$\psi(u,v|\theta,\varphi,Z)=\sqrt{P(u,v|\theta,\varphi,Z)}e^{i\sqrt{\lambda}S(u,v)/2} \tag{20}$$

to find that $F(\theta,\varphi)=Q(\theta,\varphi)$. The extrema of $Q(\theta,\varphi)$ can be found by a standard variational calculation using the variation $\psi^*(u,v|\theta,\varphi,Z)\to\psi^*(u,v|\theta,\varphi,Z)+\delta\psi^*(u,v|\theta,\varphi,Z)$, and therefore it follows that the extrema of Eq. (19) are given by the solutions of the linear eigenvalue problem,

$$-\frac{1}{\sin u}\frac{\partial}{\partial u}\left(\sin u\frac{\partial\psi(u,v|\theta,\varphi,Z)}{\partial u}\right)-\frac{1}{\sin^2 u}\frac{\partial^2\psi(u,v|\theta,\varphi,Z)}{\partial v}=\frac{I\lambda}{2}E\;\psi(u,v|\theta,\varphi,Z). \tag{21}$$

If we set $\lambda=4/\hbar^2$, Eq. (21) is identical to the time-independent Schrödinger equation for the linear rotator model. Planck's constant $\hbar$ enters here because of dimensional reasons and it sets the energy scale of experiments. Of course, there is nothing in the logical inference approach that forces us to set $\lambda=4/\hbar^2$. From Eq. (21), it follows that λ only affects the energy scale, hence its value cannot be determined from the mathematical structure of the equations. The relation $\lambda=4/\hbar^2$ can only be established empirically by comparing the numerical values obtained by solving the Schrödinger equation with the corresponding experimental results.

The equivalence of Eq. (18) and Eq. (19) allows us to determine, from the solutions of the Schrödinger equation, the i-probs $\psi(u,v|\theta,\varphi,Z)$ which yield the most likely and most robust, reproducible data, collected in the experiment. Put differently, through the Schrödinger equation, quantum theory describes an experiment which yields data that is the most robust with respect to small variations of the external conditions, in the case at hand the unknown coordinates θ,φ of the linear rotator, under which the experiment is being performed.

In short, in the logical inference approach the Schrödinger equation is not postulated but it is derived by assuming that the theory describes reproducible experiments in the most robust possible way.

4. CONCLUSION

The derivations presented here and in our earlier work[1] demonstrate that logical inference, that is plausible reasoning, applied to experiments for which (i) there is uncertainty about each event, (ii) the conditions under which the experiment is carried out may be uncertain, and (iii) the frequencies with which events are observed are reproducible and robust against small changes in the conditions, yields two important, general results.

The first is the justification of the intuitive procedure to assign to the i-probs the frequencies for the events to occur. For *fixed* experimental conditions, the usual argument for adopting this assignment is that it maximizes the i-prob to observe these frequencies. On the other hand, it is quite natural to expect that under *variable* experimental conditions it is the most robust, reproducible experiment which produces the most likely frequencies of events. Obviously, the argument based on reproducibility under *variable* experimental conditions is more general as it contains the condition of *fixed* experimental conditions as a special case.

The second, and most important for the purpose of recovering the quantum theoretical description as an application of logical inference, are equations that determine the functional dependence of the i-probs on the condition that is considered to be variable. The key point in the derivation of the quantum theoretical description is to express precisely and unambiguously, using the mathematical framework of plausible reasoning,[2–6] the essential features of the experiment. No concepts of quantum theory enter in this derivation. Adding the requirement that the experimental results are insensitive to small changes of the conditions under which the experiment is carried out yields equations that are known from quantum theory.

The logical-inference methodology to derive the basic equations of quantum theory has some implications for interpretational aspects of quantum theory. First, although it supports Bohr's view expressed in quotes (1–2) reproduced in the introduction, it does not support the Copenhagen interpretation (in any form).[10] Indeed, the wave function Eq. (20) merely appears to be a purely mathematical vehicle to turn nonlinear differential equations into linear ones and it seems difficult to attribute more meaning to such a vehicle other than mathematical. Moreover, in the logical-inference description of EPRB- and Stern-Gerlach experiment the wave function does not appear at all.[1] On the other hand, there is no conflict with the statistical interpretation[34,35] if we ignore the conceptual difference between i-probs and "mathematical" probabilities.

Second, it follows that quantum theory is a "common sense" description of a vast class of experiments but it definitely does not describe what is happening to a particle, say. By construction, our approach only makes statements about logical relationships hence there can be no conflict with ontological descriptions of the same phenomena, as long as the latter conform to the rules of Boolean logic. This follows most clearly from our derivation of the Schrödinger equation, which shows that quantum theory does not provide *any* insight into the motion of a particle but instead describes all what can be *inferred* (within the framework of logical inference) from or, using Bohr's words, *said* about the observed data, in complete agreement with Bohr's view expressed in quotes reproduced in the introduction.

Acknowledgement

We would like to thank Koen De Raedt, Karl Hess, Fengping Jin, Andrei Khrennikov, Thomas Lippert, Seiji Miyashita, Theo Nieuwenhuizen, and Arkady Plotnitsky for many stimulating discussions. We are grateful to John Ralston for a critical reading of the manuscript and for suggestions that improve the presentation of our work.

REFERENCES

1. H. De Raedt, M. Katsnelson, and K. Michielsen, "Quantum theory as the most robust description of reproducible experiments," 2013. arXiv/quant-ph/1303.4574.
2. R. T. Cox, "Probability, Frequency and Reasonable Expectation," *Am. J. Phys.* **14**, pp. 1 – 13, 1946.
3. R. T. Cox, *The Algebra of Probable Inference*, Johns Hopkins University Press, Baltimore, 1961.
4. M. Tribus, *Rational Descriptions, Decisions and Designs*, Expira Press, Stockholm, 1999.
5. C. R. Smith and G. Erickson, "From Rationality and consistency to Bayesian probability," in *Maximum Entropy and Bayesian Methods*, J. Skilling, ed., pp. 29 – 44, Kluwer Academic Publishers, (Dordrecht), 1989.
6. E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, 2003.
7. N. Bohr, "XV. The Unity of Human Knowledge," in *Complementarity Beyond Physics (1928 –1962)*, D. Favrholdt, ed., *Niels Bohr Collected Works* **10**, pp. 155 – 160, Elsevier, Amsterdam, 1999.
8. A. Petersen, "The philosophy of Niels Bohr," *Bull. Atomic Scientists* **19**, pp. 8 – 14, 1963.
9. K. V. Laurikainen, *The Message of the Atoms. Essays on Wolfgang Pauli and the Unspeakable*, Springer, Berlin, 1997.
10. D. Home, *Conceptual Foundations of Quantum Physics*, Plenum Press, New York, 1997.

11. L. E. Ballentine, "Interpretations of Probability and Quantum Theory," in *Foundations of Probability and Physics*, A. Khrennikov, ed., pp. 71 – 84, World Scientific, (Singapore), 2001.
12. R. B. Griffiths, *Consistent Quantum Theory*, Cambridge University Press, Cambridge, 2002.
13. A. Y. Khrennikov, *Contextual Approach to Quantum Formalism*, Springer, Berlin, 2009.
14. E. T. Jaynes, "Information Theory and Statistical Mechanics," *Phys. Rev.* **106**, pp. 620 – 640, 1957.
15. E. T. Jaynes, "Information Theory and Statistical Mechanics. II," *Phys. Rev.* **108**, pp. 171 – 190, 1957.
16. J. Pearl, *Causality: models, reasoning, and inference*, Cambridge University Press, Cambridge, 2000.
17. A. Plotnitsky, ""Dark Materials to Create More Worlds": On Causality in Classical Physics, Quantum Physics, and Nanophysics," *J. Comput. Theor. Nanosci.* **8**, pp. 983 – 997, 2011.
18. G. 't Hooft, "Quantummechanical behaviour in a deterministic model," *Foundations of Physics Letters* **10**, pp. 105 – 111, 1997.
19. K. De Raedt, H. De Raedt, and K. Michielsen, "Deterministic event-based simulation of quantum interference," *Comp. Phys. Comm.* **171**, pp. 19 – 39, 2005.
20. G. 't Hooft, "The mathematical basis for deterministic quantum mechanics," in *Beyond the Quantum*, T. M. Nieuwenhuizen, B. Mehmani, V. Špička, M. J. Aghdami, and A. Y. Khrennikov, eds., pp. 3 – 19, World Scientific, Singapore, 2007.
21. K. Michielsen, F. Jin, and H. De Raedt, "Event-based Corpuscular Model for Quantum Optics Experiments," *J. Comp. Theor. Nanosci.* **8**, pp. 1052 – 1080, 2011.
22. H. De Raedt and K. Michielsen, "Event-by-event simulation of quantum phenomena," *Ann. Phys. (Berlin)* **524**, pp. 393 – 410, 2012.
23. H. De Raedt, F. Jin, and K. Michielsen, "Event-based simulation of neutron interferometry experiments," *Quantum Matter* **1**, pp. 1 – 21, 2012.
24. K. Michielsen and H. De Raedt, "Event-based simulation of quantum physics experiments," 2013. arXiv/quant-ph/13??.????
25. G. Pólya, *Mathematics and Plausible Reasoning*, Princeton University Press, Princeton, 1954.
26. J. M. Keynes, *A treatise on probability*, Macmillan, London, 1921.
27. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1, Wiley & Sons, New York, 1968.
28. G. R. Grimmet and D. R. Stirzaker, *Probability and Random Processes*, Clarendon Press, Oxford, 1995.
29. A. Plotnitsky, "The Art and Science of Experimentation in Quantum Physics," *AIP Conf. Proc.* **1232**, pp. 128 – 142, 2010.
30. A. Plotnitsky, ""Who Thinks Abstractly?:" Quantum Theory and the Architecture of Physical Concepts," *AIP Conf. Proc.* **1327**(1), pp. 207 – 220, 2011.
31. A. Plotnitsky, "On Foundational Thinking in Fundamental Physics, from Riemann to Einstein to Heisenberg," *AIP Conf. Proc.* **1424**(1), pp. 282 – 303, 2012.
32. E. Madelung, "Quantentheorie in hydrodynamischer Form," *Z. Phys.* **40**, pp. 322 – 326, 1927.
33. D. Bohm, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I," *Phys. Rev.* **85**, pp. 166 – 179, 1952.
34. L. E. Ballentine, "The Statistical Interpretation of Quantum Mechanics," *Rev. Mod. Phys.* **42**, pp. 358 – 381, 1970.
35. L. E. Ballentine, *Quantum Mechanics: A Modern Development*, World Scientific, Singapore, 2003.