

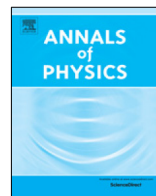


ELSEVIER

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: [www.elsevier.com/locate/aop](http://www.elsevier.com/locate/aop)



# Logical inference derivation of the quantum theoretical description of Stern–Gerlach and Einstein–Podolsky–Rosen–Bohm experiments

Hans De Raedt<sup>a</sup>, Mikhail I. Katsnelson<sup>b</sup>, Kristel Michielsen<sup>c,d,\*</sup>

<sup>a</sup> Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands

<sup>b</sup> Radboud University Nijmegen, Institute for Molecules and Materials, Heyendaalseweg 135, NL-6525 AJ Nijmegen, The Netherlands

<sup>c</sup> Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany

<sup>d</sup> RWTH Aachen University, D-52056 Aachen, Germany

## H I G H L I G H T S

- Logical inference connects data and their description in terms of concepts.
- It is used to derive the description of Stern–Gerlach experiments.
- Logical inference yields the description of Einstein–Podolsky–Rosen–Bohm experiments.
- These logical inference derivations do not rely on concepts of quantum theory.

## A R T I C L E I N F O

### Article history:

Received 26 April 2018

Accepted 9 July 2018

Available online 11 July 2018

### Keywords:

Logical inference

Quantum theory

Stern–Gerlach experiments

Einstein–Podolsky–Rosen–Bohm experiments

## A B S T R A C T

It is shown that the quantum-theoretical descriptions of idealized Stern–Gerlach and Einstein–Podolsky–Rosen–Bohm experiments result from a logical inference treatment of robust experiments, without relying on concepts of quantum theory. For concreteness, the derivation is given for the case that the Stern–Gerlach magnet deflects the beam in three different directions, which in quantum theory corresponds to a description in terms of spin-1 particles.

\* Corresponding author at: Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany.

E-mail address: [k.michielsen@fz-juelich.de](mailto:k.michielsen@fz-juelich.de) (K. Michielsen).

The logical inference description of the Einstein–Podolsky–Rosen–Bohm experiment is shown to contain the quantum theoretical description as a particular case.

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The algebra of logical inference (LI) provides a mathematical framework that facilitates rational reasoning when there is uncertainty [1–5]. A detailed discussion of the foundations of LI, its relation to Boolean logic and the derivation of its rules can be found in the papers [1,4] and books [2,3,5]. LI is the foundation for powerful tools such as the maximum entropy method and Bayesian analysis [3,5]. To the best of our knowledge, the first derivation of a theoretical description by this general methodology of scientific reasoning appears in Jaynes' papers on the relation between information and (quantum) statistical mechanics [6,7]. Some of the most basic equations of quantum theory, e.g. the Schrödinger, Pauli and Klein–Gordon equations, and the probability distributions of pairs of particles in the singlet or triplet state have been shown to emerge from the application of LI to (the abstraction of) robust experiments, without taking recourse to concepts of quantum theory [8–12]. The LI approach yields results that are unambiguous and independent of the individual subjective judgment and in addition, provides a rational explanation for the extraordinary descriptive power of quantum theory [9] and strong support for Bohr's statement [13] "The physical content of quantum mechanics is exhausted by its power to formulate statistical laws governing observations under conditions specified in plain language".

LI derivations of quantum theoretical descriptions are void of postulates regarding "wave functions", "observables", "quantization rules", "Born's rule", etc. 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 relations between those events that we, humans, discover. Whether or not these discrete events are "generated" according to certain (quantum) laws is irrelevant. The laws themselves appear as the result of (the best) inference (in a sense discussed below) based on data available in the form of discrete events. As our LI approach is built on the concept of an event (in the general sense of the word), it is not possible and it would even be a logical fallacy to "derive" the existence of a definite result of a measurement from the theory. In other words, in the LI approach there is no "measurement problem".

In the LI approach, the machinery of quantum theory appears as a result of transforming a non-linear, global optimization problem into a linear one which is (much) easier to handle. The wave function, spin, etc. are only mathematical concepts, vehicles that render a class of complicated nonlinear minimization problems into the minimization of quadratic forms. As products of our collective imagination, these concepts are extraordinarily useful but have no tangible existence, just like numbers themselves.

LI derivations of quantum theoretical descriptions also differ from the more traditional theoretical-physics approach in the sense that they follow directly from the description of the experimental scenario and do not rely on reduction from or approximation/simplification of a "parent" equation. For instance, our LI derivation of the Pauli-equation [10] for a neutral, magnetic particle proceeds directly from the measurement scenario to the equation whereas the conventional derivation of the Pauli equation from the Dirac equation for charged particles [14] does not even apply to electrically neutral particles.

A characteristic feature of quantum theory is the presence of Planck's constant  $\hbar$ . Clearly, Planck's constant cannot "emerge" from logical considerations only. In the LI derivation of the Schrödinger-, Pauli-, and Klein–Gordon equation, Planck's constant appears as a Lagrange parameter, a vehicle for solving the optimization problem [9–12]. In quantum theoretical language, this paper exclusively deals with spin operators. Although it is convention to represent these operators by dimensionless

matrices multiplied by  $\hbar$ , as long as there are no other degrees of freedom involved, there is no loss of generality by simply using the dimensionless matrices, which is the approach adopted in this paper.

In this paper, we present a LI treatment of idealized Stern–Gerlach (SG) and Einstein–Podolsky–Rosen–Bohm (EPRB) experiments with magnetic particles. We derive, not postulate, the quantum theoretical description of the SG experiment for magnetic particles directly from the LI description of the observed data and then reuse the LI description to derive the quantum theoretical description of EPRB experiments. It is important to mention here that we use the idealized SG and EPRB experiments as representative examples for experiments that, for each repetition, yield one definite outcome out of a finite number of different possibilities. No elements of e.g. motion enter the description, only the counting of events is involved, and therefore our treatment is void of physical units.

In short, a SG experiment involves a particle source, a magnet and a particle detector. The source emits neutral particles (atoms, neutrons, ...) which carry a magnetic moment. The particles are sent through a Stern–Gerlach magnet in which they experience an inhomogeneous magnetic field. As a result of the interaction with this field, the particles leave the Stern–Gerlach magnet in spatially well-separated directions, an experimental fact [15–19]. The directions of the deflections are assumed to depend on a unit vector  $\mathbf{a}$  which is along the main axis of inhomogeneity of the magnetic field. The observation of deflections in spatially well-separated directions is regarded as experimental evidence that the magnetic moment of the particles is quantized [15,20,21]. The quantized magnetization is called the spin of the particle and the number of different deflections is given by  $2S + 1$  where  $S$  is the magnitude of the spin. Following Feynman [21], we focus on the case that the particles are deflected in three distinct directions, that is on spin  $S = 1$  particles, and then show how the LI derivation generalizes, extending our earlier work on  $S = 1/2$  particles [9,12] to particles with larger spin. As an application, we use LI to derive the quantum theoretical description of the EPRB experiment with spin  $S = 1$ .

## 2. Double Stern–Gerlach experiment

As will become clear later on, to develop a consistent description, it is not sufficient to consider only the standard SG experiment consisting of a source, magnet and detector. It is necessary to consider a double SG experiment [20,21] such as the one sketched in Fig. 1. Following common practice in developing theoretical descriptions, we assume that the experiment is “perfect”, meaning that all particles leaving the source initially travel along the same direction (which we choose to be along the  $x$ -axis of the laboratory frame of reference), each particle is detected by one (and only one) of the nine detectors placed in the corresponding outgoing beams, that all SG magnets are identical and that their inhomogeneous magnetic fields do not change during the course of collecting data.

### 2.1. Data generated by the experiment

From Fig. 1, it is clear that for each particle passing through the double SG device, a click of one (and only one) of the nine detectors tells us the pair  $(k, l)$  which labels the path that the particle took. Thus, repeating the experiment with  $N$  particles (and assuming that no particles are lost in the process of leaving the source and being detected), yields the data set

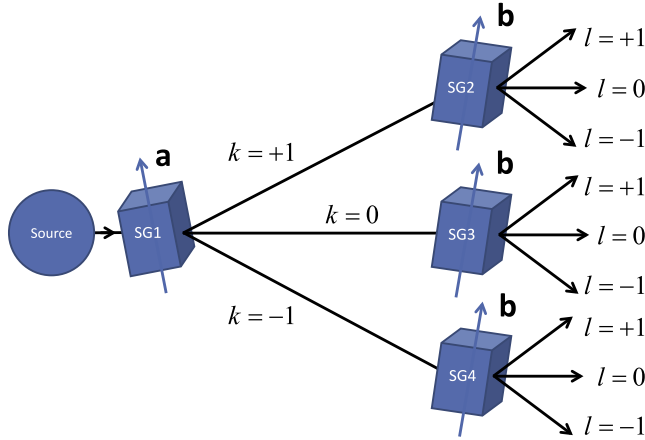
$$\mathcal{D} = \{(k_n, l_n) \mid k_n, l_n \in \mathcal{E}; n = 1, \dots, N\}, \quad (1)$$

where we introduced the set of elementary events  $\mathcal{E} = \{+1, 0, -1\}$ . From the data set Eq. (1) we compute the relative frequency of an event  $(k, l)$

$$f(k, l | \mathbf{a}, \mathbf{b}, N) = \frac{1}{N} \sum_{n=1}^N \delta_{k,k_n} \delta_{l,l_n}, \quad (2)$$

that a particle travels along the path labeled by  $(k, l)$ .

In general, knowledge of the relative frequencies  $f(k, l | \mathbf{a}, \mathbf{b}, N)$  does not suffice to fully characterize the data of the set  $\mathcal{D}$ . For instance, the relative frequencies cannot describe correlations between successive detection events, if any were present.



**Fig. 1.** (Color online) Diagram of the double Stern–Gerlach experiment with spin-1 particles. A source sends particles carrying a magnetic moment through a Stern–Gerlach magnet SG1 which maintains an inhomogeneous magnetic field characterized by a unit vector  $\mathbf{a}$ . Particles are deflected in three different directions labeled by  $k = +1, 0, -1$ . Each of the three beams of particles is sent through another SG magnet. The SG magnets SG2, SG3 and SG4 are assumed to be identical with their inhomogeneous magnetic fields characterized by a unit vector  $\mathbf{b}$ . The deflected beams emerging from these magnets are labeled by  $l = +1, 0, -1$ . Particle detectors (not shown) count the number of particles in each of the nine beams. The path of each particle through the double Stern–Gerlach device is uniquely determined by the labels  $(k, l)$ .

In this paper, we discarding all knowledge about the events that is not contained in the relative frequencies  $f(k, l|\mathbf{a}, \mathbf{b}, N)$ .

According to the text in this box, in this paper we take the viewpoint that all the knowledge that can be extracted from the data set  $\mathcal{D}$  is fully captured by the relative frequencies  $0 \leq f(k, l|\mathbf{a}, \mathbf{b}, N) \leq 1$  subject to the normalization constraint  $\sum_{k,l \in \mathcal{E}} f(k, l|\mathbf{a}, \mathbf{b}, N) = 1$ .

### 3. Logical inference approach

The key concept of the LI approach is the plausibility, denoted by  $P(A|B)$ , which in general, expresses the degree of believe of an individual that proposition  $A$  is true, given that proposition  $B$  is true [2,3,5,22]. 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 [3,5]. In this paper, the plausibility is regarded as the primary concept for the development of a theoretical description.

The algebra of LI can be derived from so-called “desiderata” which express, in words, what is generally regarded as rational reasoning [22]. It is a most remarkable fact that these desiderata suffice to uniquely determine the set of rules by which plausibilities may be manipulated [2–5]. It can be shown [2–5] that plausibilities may be chosen to take numerical values in the range  $[0, 1]$  and the values of these plausibilities are related by three rules, namely [2–5]

- (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”  $AB$  denotes the logical product (conjunction) of the propositions  $A$  and  $B$ .
- (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  $B$ .

The algebra of logical inference, as defined by the rules (1)–(3), contains Boolean algebra as a special case and provides the foundation for powerful tools such as the maximum entropy method and

Bayesian analysis [3,5]. The rules (1)–(3) are unique [3–5] in the sense that any other rule which applies to plausibilities represented by real numbers and is in conflict with rules (1)–(3) is also at odds with rational reasoning and consistency [3–5].

In spite of the apparent similarities of rules (1)–(3) with those of probability theory, the reader should not think of the plausibility as a frequency or probability in the traditional mathematical sense but merely as a numerical measure for proposition  $A$  to be true, given that proposition  $B$  is true [9]. Furthermore, plausibilities are concepts resulting from human reasoning about observed events and their relationships but are not the “cause” of these events. 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. In this paper, we only consider applications of LI which describe phenomena “in a manner independent of individual subjective judgment” [13].

For completeness, we mention here that it is not allowed to define a plausibility for a proposition conditional on the conjunction of mutual exclusive propositions: LI cannot be used to reason on the basis of two or more contradictory premises.

### 3.1. Application to the double SG experiment

The first step of the LI approach is to introduce a real number  $P(k, l|\mathbf{a}, \mathbf{b}, Z)$  which represents the plausibility that the detector labeled by the pair  $(k, l)$  fires, under fixed conditions specified by the unit vectors  $\mathbf{a}$  and  $\mathbf{b}$ , and proposition  $Z$ . In order to simplify the notation somewhat we make some abuse of notation by writing e.g.  $k, l$  as a shorthand for the proposition “the detector labeled by  $(k, l)$  fires”. The proposition  $Z$  represents the conjunction of all propositions about the experiment (e.g. temperature, humidity, ...) that are not deemed important for the description of the data.

The second step is to use the rules (1)–(3) to compute the plausibility  $P(\mathcal{D}|\mathbf{a}, \mathbf{b}, N, Z)$  that after collecting  $N$  detection events, we observe the data set  $\mathcal{D}$ . As stated above in the boxed text, in the description of the data, we ignore all correlations between detection events (if any). In terms of plausibilities, this means that the plausibility of an event  $(k, l)$  does not depend on an earlier or later event  $(k', l')$ . Invoking the product rule (2), the logical consequence of this independence is that we have

$$\begin{aligned}
 P(\mathcal{D}|\mathbf{a}, \mathbf{b}, N, Z) &= P(k_1, l_1|k_2, l_2, \dots, k_N, l_N, \mathbf{a}, \mathbf{b}, N, Z)P(k_2, l_2, \dots, k_N, l_N|\mathbf{a}, \mathbf{b}, N, Z) \\
 &= P(k_1, l_1|\mathbf{a}, \mathbf{b}, Z)P(k_2, l_2, \dots, k_N, l_N|\mathbf{a}, \mathbf{b}, N, Z) \\
 &= P(k_1, l_1|\mathbf{a}, \mathbf{b}, Z)P(k_2, l_2|k_3, l_3, \dots, k_N, l_N, \mathbf{a}, \mathbf{b}, N, Z) \\
 &\quad \times P(k_3, l_3, \dots, k_N, l_N|\mathbf{a}, \mathbf{b}, N, Z) \\
 &= P(k_1, l_1|\mathbf{a}, \mathbf{b}, Z)P(k_2, l_2|\mathbf{a}, \mathbf{b}, Z)P(k_3, l_3, \dots, k_N, l_N|\mathbf{a}, \mathbf{b}, N, Z) = \dots \\
 &= \prod_{n=1}^N P(k_n, l_n|\mathbf{a}, \mathbf{b}, Z), \tag{3}
 \end{aligned}$$

where  $P(k, l|\mathbf{a}, \mathbf{b}, Z)$  denotes the plausibility to observe a single event  $(k, l)$ . As the order in which pairs  $(k, l)$  appear is irrelevant (as a consequence of the statement in the boxed text), the data set  $\mathcal{D}$  can be, without loss of relevant information, be compressed to

$$\mathcal{N} = \{n_{+1,+1}, \dots, n_{-1,-1}|\mathbf{a}, \mathbf{b}, N, Z\}, \tag{4}$$

where we have made it explicit that  $\mathcal{N}$  depends on  $\mathbf{a}, \mathbf{b}, N$  and  $Z$ . It then follows from Eq. (3) that the plausibility  $P(\mathcal{N}|\mathbf{a}, \mathbf{b}, N, Z)$  to observe the data set  $\mathcal{N}$  is, by the usual counting argument, given by

$$P(\mathcal{N}|\mathbf{a}, \mathbf{b}, N, Z) = N! \prod_{k,l \in \mathcal{D}} \frac{1}{n_{k,l}!} (P(k, l|\mathbf{a}, \mathbf{b}, Z))^{n_{k,l}}. \tag{5}$$

which is the plausibility to observe the frequencies  $f(+1, +1|\mathbf{a}, \mathbf{b}, N), \dots, f(-1, -1|\mathbf{a}, \mathbf{b}, N)$ . From Eq. (5) it follows that the plausibility to observe the frequency  $f(k, l|\mathbf{a}, \mathbf{b}, N)$  is a function of  $P(k, l|\mathbf{a}, \mathbf{b}, Z)$  and  $N$  only.

The third step involves the formulation of assumptions about the symmetries of the problem and of other constraints we wish to impose. Thereby it is important to distinguish between the data and the description thereof. A symmetry expresses an exact relation between the **descriptions** of two or more different situations and not between the data sets themselves. Indeed, it is highly unlikely that data sets such as  $\mathcal{D}$ , collected in different experiments, exhibit an exact symmetry relation. In other words, we can impose symmetries on plausibilities (description) but not on frequencies (facts).

### 3.2. Rotational symmetries

In the case at hand, it seems reasonable to hypothesize that outcomes of the double SG experiment can be described by a model which depends on the relative orientation of the SG magnets, not on both  $\mathbf{a}$  and  $\mathbf{b}$  independently. This model is invariant under rotations of the frame of reference, implying that  $P(k, l|\mathbf{a}, \mathbf{b}, Z) = P(k, l|\mathbf{a} \cdot \mathbf{b}, Z) = P(k, l|\theta, Z)$  where  $\cos \theta = \mathbf{a} \cdot \mathbf{b}$ , from which it immediately follows that  $P(k, l|\theta, Z) = P(k, l|\theta + 2\pi, Z)$ . Furthermore, it seems reasonable to assume that changing the direction of SG1 from  $+\mathbf{a}$  to  $-\mathbf{a}$  has the same effect as changing  $+k$  to  $-k$  (see Fig. 1). Of course, the same argument holds for the SG magnets of the second stage. Therefore, assuming that this reflection symmetry holds implies the relations  $P(k, l|\mathbf{a} \cdot \mathbf{b}, Z) = P(-k, l|-\mathbf{a} \cdot \mathbf{b}, Z) = P(k, -l|-\mathbf{a} \cdot \mathbf{b}, Z) = P(-k, l|\theta + \pi, Z) = P(k, -l|\theta + \pi, Z)$ . To avoid misunderstandings, it may be important to mention here that our choice of labeling events by the numbers of the set  $\mathcal{E}$  and their relation to the direction of  $\mathbf{a}$  is very convenient but not essential. We could have equally well chosen to work with for instance the labels  $k \in \{r, g, b\}$  where  $r, g$ , and  $b$  denote the colors red, green and blue, respectively. Then, the assumption of reflection symmetry implies that changing  $+\mathbf{a}$  to  $-\mathbf{a}$  induces the permutation  $\{r, g, b\} \rightarrow \{b, g, r\}$  of the labels.

### 3.3. Ideal filtering device

The demonstration by SG that atom-size magnetic particles passing through an inhomogeneous magnetic field experience deflections in spatially well-separated directions was important for the development of quantum theory because it provided experimental evidence that not only the spectra of atoms but also the magnetic moment of the particles is quantized [15,20,21]. Since then, the SG experiment, or its conceptually equivalent experiment with single-photons passing through a birefringent crystal, is used in textbooks to motivate a few of the postulates of quantum theory [20,21,23].

The SG device is often taken as the prime example of a device that separates a beam of particles based on the direction of their magnetic moment. In its idealized form, it acts as a perfect filter.

At first sight, it sounds very reasonable to say that after a particle has left a SG magnet, its magnetic moment is aligned along the  $\mathbf{a}$  direction and is quantized with projection  $+1, 0$  or  $-1$  along this axis. Clearly, this quantization is in one-to-one correspondence with the observed deflection.

If we have no knowledge about the direction of magnetization of the particle as it leaves the source, the deflection by a single SG magnet with its magnetic field described by the vector  $\mathbf{a}$  is not, as a matter of principle, sufficient to make a meaningful statement about the direction of its magnetic moment.

A consistent assignment of a magnetic moment requires that if we send that particle through a second SG magnet with its magnetic field described by the same  $\mathbf{a}$ , the particle only emerges in the same beam, that is  $l$  must be equal to  $k$ .

If that is the case, the SG magnet can be viewed as a perfect filter. The direction of the magnetic moment defined by the direction  $\mathbf{a}$  of the first SG magnet is preserved if we send the particle through another SG magnet with its magnetic field along  $\mathbf{a}$ .

Establishing this filter property from LI principles obviously demands that we treat the double SG experiment, not the experiment with a single SG magnet. It is evident that the requirement that SG magnets act as ideal filters cannot be an intrinsic feature of the LI framework. The filtering property has to be imposed explicitly, just as we did with symmetry relations. The constraint on the plausibilities which enforces the filtering property is simple: we must have  $P(k, l|\mathbf{a}, \mathbf{a}, Z) = 0$  if  $k \neq l$ .

### 3.4. Relation to laboratory SG experiments

We are not aware of reports on double SG experiments which have been performed in the laboratory so we restrict the discussion that follows to SG experiments involving a single SG magnet [15–17]. The relevant data set for these experiments is

$$\mathcal{D}' = \{k_n \mid k_n \in \mathcal{E}; n = 1, \dots, N\}, \quad (6)$$

which clearly lacks information on how the particles form the different beams or how the particles are distributed over the detection area other than that they have been classified into three groups. Formulated differently, the data of the positions of where particles have been detected have been considerably compressed, i.e. reduced to a set of three-valued numbers.

A more detailed, quantum theoretical description starts from the Pauli equation for a neutral magnetic particle in an inhomogeneous magnetic field. The Hamiltonian reads

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \gamma\mathbf{B} \cdot \mathbf{S}, \quad (7)$$

where  $\hbar$ ,  $m$ ,  $\gamma$ , and  $\mathbf{S}$  are Planck's constant, the mass, gyromagnetic ratio and spin angular momentum of the particle, respectively.

Replacing  $\mathbf{B} \cdot \mathbf{S}$  by  $B^2(x, y)S^z$ , it follows immediately that the incoming beam splits according to the eigenvalues of the  $z$ -component of the spin operator [20,23,24]. In this case, the SG magnet performs an ideal, projective measurement, exactly as envisaged by von Neumann [25].

However, the Maxwell equation  $\nabla \cdot \mathbf{B} = 0$  does not admit inhomogeneous magnetic fields that only have one component. For instance, if the particles leave the source in the  $x$ -direction and we choose the magnetic field in this direction to be zero, the general form of the magnetic field is  $\mathbf{B} = (0, \partial f(\mathbf{r})/\partial z, -\partial f(\mathbf{r})/\partial y)$  where  $f(\mathbf{r})$  can be any nice function. Further restrictions on  $f(\mathbf{r})$  follow when we require that the electric field does not change with time, i.e.  $\nabla \times \mathbf{B} = 0$  (we assume that the electrical current is zero).

These constraints imposed by Maxwell's equations destroy the ideal filtering property of the SG magnet because the projections of the spin onto a particular direction do no longer commute with the Hamiltonian [26]. This means that in a two SG magnet setup with both magnetic fields along  $\mathbf{a}$ , a particle leaving SG1 in beam  $k$  has a non-zero probability to leave the second layer of SG magnets in a beam  $l \neq k$ . Numerical solutions of the Pauli equation with Hamiltonian equation (7) show that this is indeed what happens [26]. In conclusion, the LI derivation presented in this paper only concerns the idealized, projective-measurement version of the SG experiment.

### 3.5. Summary of constraints on the plausibilities

Symmetry considerations restrict the possible LI descriptions of the double SG experiment to plausibilities that satisfy the relations

$$\begin{aligned} P(k, l|\mathbf{a}, \mathbf{b}, Z) &= P(k, l|\mathbf{a} \cdot \mathbf{b}, Z) = P(k, l|\theta, Z) = P(k, l|\theta + 2\pi, Z) \\ &= P(-k, l|\theta + \pi, Z) = P(k, -l|\theta + \pi, Z), \end{aligned} \quad (8)$$

where  $\cos \theta = \mathbf{a} \cdot \mathbf{b}$ . Note that  $P(k, l|\theta, Z)$  can only depend on  $\theta$  through  $\cos \theta$ .

Imposing that the model of the SG magnet acts as a perfect filtering device requires that the plausibilities satisfy

$$P(k, l|\mathbf{a}, \mathbf{a}, Z) = 0 \quad \text{if } k \neq l. \quad (9)$$

### 3.6. Robust experiments

All quantum physics experiments have in common that it is impossible to predict the outcome of an individual event, in case a pair  $(k, l)$ . Therefore, in the case of the double SG experiment, the data set  $\mathcal{D}$  is expected to change significantly from run to run. However, if the events  $(k, l)$  occur with frequencies

that remain fairly constant (within the usual statistical fluctuations) upon repeating the experiment, it may be possible to construct a model that describes the dependence of the frequencies on the settings  $\mathbf{a}$  and  $\mathbf{b}$ . In previous work [8–12], it was shown that the most basic equations of quantum theory derive directly from the LI principles if we assume that the frequencies of occurrence are robust, i.e. smoothly change a little if the conditions vary a little. It seems to us that this is a general requirement for any successful scientific experiment. Indeed, if the frequencies of occurrence are not robust with respect to small changes of the conditions, they would vary erratically as conditions change and common practice is to discard experiments that produce such data.

In this paper, we derive the quantum theoretical description of the (double) SG experiment with spin-1 particles directly from LI principles and symmetry considerations, without taking recourse to the formalism of quantum theory. The basic assumptions of the LI derivation that follows can be summarized as follows:

- (i) There is uncertainty about each individual event  $(k, l)$  and knowledge about any such event does not change our knowledge about earlier or later events.
- (ii) Upon repetition of the experiment with  $N$  events, the frequencies  $f(k, l | \mathbf{a}, \mathbf{b}, N)$  for  $(k, l) = (+1, +1), \dots, (-1, -1)$  are, within the usual statistical fluctuations, reproducible.
- (iii) The frequencies  $f(k, l | \mathbf{a}, \mathbf{b}, N)$  for  $(k, l) = (+1, +1), \dots, (-1, -1)$  are **robust** with respect to small changes in  $\mathbf{a}$  or  $\mathbf{b}$ .

Obviously, assumptions (i) and (ii) are trivially satisfied when we adopt a description in terms of plausibilities. Expressing the notion of a robust experiment is the key element of the derivation.

Experiments which produce results that do not change with the conditions  $\mathbf{a}$  or  $\mathbf{b}$  are non-informative and therefore fairly pointless. Therefore, in the following, we explicitly exclude such experiments from further consideration. In terms of the plausibility, this means that we exclude the case in which  $\partial P(\mathcal{N} | \theta, N, Z) / \partial \theta = 0$ .

As explained above, if the outcome of the experiment can be described by the plausibility Eq. (5) and assuming that the experiment yields reproducible and **robust** results, small changes in  $\theta$  should not have a drastic effect on the description in terms of plausibilities. This means that it is reasonable to assume that  $P(\mathcal{N} | \theta, N, Z)$  changes smoothly with  $\theta$ .

Imagine that we carry out two experiments, one under condition  $\theta$  and another one under condition  $\theta + \varepsilon$  ( $\varepsilon$  small compared to  $\pi$ ) and assume that they produce the same data set  $\mathcal{N}$ . If the experiment is robust in the sense explained above, we may expect that the difference  $P(\mathcal{N} | \theta + \varepsilon, N, Z) - P(\mathcal{N} | \theta, N, Z)$  is small. Clearly, this difference is an obvious, very simple measure for the robustness.

In practice, it is often convenient to work with the logarithm of the plausibility because this turns the product rule (2) into a sum rule. This is allowed because we may replace all plausibilities by any continuous monotonic function of them without changing the content of the LI framework [5]. Therefore, we define

$$\mathcal{R} = \pm \ln \frac{P(\mathcal{N} | \theta + \varepsilon, N, Z)}{P(\mathcal{N} | \theta, N, Z)}, \tag{10}$$

to be the measure for the robustness. The  $\pm$  in Eq. (10) reflects the fact that instead of  $\mathcal{R} = \ln(P(\mathcal{N} | \theta + \varepsilon, N, Z) / P(\mathcal{N} | \theta, N, Z))$  we could have equally well used  $\mathcal{R} = \ln(P(\mathcal{N} | \theta, N, Z) / P(\mathcal{N} | \theta + \varepsilon, N, Z))$  as the measure of robustness. Of course, there may be other definitions for the measure of robustness. The rationale for adopting Eq. (10) is that it provides a path to derive, not postulate, the basic equations of quantum physics [9].

Our primary goal is to determine  $P(\mathcal{N} | \theta, N, Z)$  which minimizes the robustness measure  $|\mathcal{R}|$ . Thus,  $\mathcal{R}$  is to be regarded as a function of the set of plausibilities  $\{P(\mathcal{N} | \theta, N, Z) \mid 0 \leq \theta \leq \pi\}$ , i.e. as a functional of the plausibility. Therefore, the goal is to find the plausibility  $P(\mathcal{N} | \theta, N, Z)$ , not  $\theta$  or  $\varepsilon$ , which minimizes  $|\mathcal{R}|$ .

From Eq. (10) it follows immediately that if  $P(\mathcal{N} | \theta, N, Z)$  does not depend on  $\theta$  then  $\mathcal{R} = 0$  for all  $\theta$  which is certainly the most robust description one can ever have. However, as we explained above, a description of an experiment that does not show a  $\theta$ -dependence is of no interest to us. Therefore, it seems reasonable to relax the requirement  $\mathcal{R} = 0$  for all  $\theta$  and (small)  $\varepsilon$  a little and ask for the set



of plausibilities  $\{P(\mathcal{N}|\theta, N, Z) \mid 0 \leq \theta \leq \pi\}$  which do change with  $\theta$  and minimize the functional  $|\mathcal{R}| > 0$  for all  $\theta$ . In summary:

The LI description of the most robust (relevant) experiment is given by the set of plausibilities  $\{P(\mathcal{N}|\theta, N, Z) \mid 0 \leq \theta \leq \pi\}$  which minimizes  $|\mathcal{R}| > 0$  for all  $\theta$ .

Obviously, because of the qualifier “for all  $\theta$ ”, at the minimum,  $|\mathcal{R}|$  does not depend on  $\theta$ .

### 3.7. Robustness in terms of single-event plausibilities

Substituting Eq. (5) into Eq. (10) yields

$$\mathcal{R} = \sum_{k,l \in \mathcal{E}} n_{k,l} \ln \frac{P(k, l|\theta + \varepsilon, Z)}{P(k, l|\theta, Z)}, \tag{11}$$

where the advantage of using the logarithm of the plausibility becomes obvious as it turns all multiplications in Eq. (5) into additions. From Eq. (11) it follows immediately that our primary goal is to determine  $P(k, l|\theta, Z)$  by minimizing  $|\mathcal{R}| > 0$  for all  $\theta$ , where the measure of robustness  $\mathcal{R}$  is now given by Eq. (11).

Finding a solution of the global, robust optimization problem Eq. (11) looks like a formidable task. As explained earlier, robustness means that small variations of  $\mathbf{a}$  or  $\mathbf{b}$  result in small and smooth variations of  $P(k, l|\theta, Z)$ . Therefore, the optimization problem may be brought into a more tractable form by assuming that  $P(k, l|\theta, Z)$  varies smoothly with  $\theta$ , i.e. that  $P(k, l|\theta, Z)$  allows for a Taylor expansion as a function of  $\theta$ . A small change of  $\mathbf{a}$  or  $\mathbf{b}$  implies that  $\theta$  changes to  $\theta + \varepsilon$  where  $\varepsilon$  is small (compared to  $\pi$ ). Writing Eq. (11) as a Taylor series in  $\varepsilon$  we have

$$\mathcal{R} = \sum_{k,l \in \mathcal{E}} n_{k,l} \left\{ \varepsilon \frac{P'(k, l|\theta, Z)}{P(k, l|\theta, Z)} - \frac{\varepsilon^2}{2} \left( \frac{P'(k, l|\theta, Z)}{P(k, l|\theta, Z)} \right)^2 + \frac{\varepsilon^2}{2} \frac{P''(k, l|\theta, Z)}{P(k, l|\theta, Z)} \right\} + \mathcal{O}(\varepsilon^3), \tag{12}$$

and

$$\begin{aligned} |\mathcal{R}| \leq |\varepsilon| & \left| \sum_{k,l \in \mathcal{E}} n_{k,l} \frac{P'(k, l|\theta, Z)}{P(k, l|\theta, Z)} \right| + \frac{\varepsilon^2}{2} \sum_{k,l \in \mathcal{E}} n_{k,l} \left( \frac{P'(k, l|\theta, Z)}{P(k, l|\theta, Z)} \right)^2 \\ & + \frac{\varepsilon^2}{2} \left| \sum_{k,l \in \mathcal{E}} n_{k,l} \frac{P''(k, l|\theta, Z)}{P(k, l|\theta, Z)} \right| + |\mathcal{O}(\varepsilon^3)|, \end{aligned} \tag{13}$$

where a prime denotes the derivative with respect to  $\theta$ .

Discarding contributions of  $\mathcal{O}(\varepsilon^3)$  or higher, we minimize  $|\mathcal{R}|$  by making the three first sums in Eq. (13) as small as possible. If we set  $n_{k,l} = NP(k, l|\theta, Z)$ , then the first and third term vanish identically. Thus, searching for a solution of the global robust optimization problem leads us to make the intuitively reasonable assignment  $P(k, l|\theta, Z) \leftarrow n_{k,l}/N$ , i.e. to use the frequency  $n_{k,l}/N$  as the numerical value of the plausibility  $P(k, l|\theta, Z)$ . This is a first important consequence of minimizing  $|\mathcal{R}|$ : it is precisely this assignment which renders the LI description unambiguous and independent of the individual subjective judgment [9].

The problem has now reduced to that of minimizing the right-hand-side of

$$|\mathcal{R}| = \frac{\varepsilon^2 N}{2} \sum_{k,l \in \mathcal{E}} \frac{1}{P(k, l|\theta, Z)} (P'(k, l|\theta, Z))^2 > 0. \tag{14}$$

Clearly, the right-hand-side of Eq. (14), being non-negative, can only be zero if  $P'(k, l|\theta, Z) = 0$ , the case we have excluded because it can only describe non-informative experiments. Disregarding the irrelevant prefactor, the right-hand-side of Eq. (14) is the expression of the Fisher information [5,27–30] of the problem at hand.

Summarizing, to derive the quantum theoretical description of the double SG experiment with spin-1 particles directly from LI principles and symmetry considerations and without taking recourse to quantum theory, we search for the solutions of the following set of equations:

$$I_F = \sum_{k,l \in \mathcal{E}} \frac{1}{P(k, l|\theta, Z)} \left( \frac{\partial P(k, l|\theta, Z)}{\partial \theta} \right)^2 > 0 \quad , \quad \sum_{k,l \in \mathcal{E}} P(k, l|\theta, Z) = 1, \tag{15}$$

$$\frac{\partial P(k, l|\theta, Z)}{\partial \theta} \neq 0 \quad \text{for at least one pair } (k, l) \in \mathcal{E}, \tag{16}$$

$$P(k, l|\theta, Z) = P(k, l|\theta + 2\pi, Z) = P(-k, l|\theta + \pi, Z) = P(k, -l|\theta + \pi, Z). \tag{17}$$

$$P(k, l|\theta = 0, Z) = 0 \quad , \quad k \neq l. \tag{18}$$

The LI description of the double SG experiment is robust (in the sense explained above), informative and satisfies the symmetry considerations (see Section 3.1) if the plausibilities  $\{P(k, l|\theta, Z) \mid (k, l) \in \mathcal{E}\}$  are the solution of Eq. (15) subject to the constraints Eqs. (16)–(18). Note that we wrote  $I_F$  instead of  $I_F(\theta)$  to emphasize that at the minimum  $I_F$  (and therefore  $\mathcal{R}$  up to second order in  $\varepsilon$ ) does not depend on  $\theta$ .

### 3.8. Relation to other work

The idea that the Fisher information can be taken as the starting point for deriving the time-independent Schrödinger equation first appeared in a paper by Frieden [31]. Frieden’s work [30] has been an important source of inspiration for our work. However, conceptually, Frieden’s approach is very different from ours.

In Frieden’s approach, the Fisher information appears as a result of using concepts such as intrinsic fluctuations and “smart measurements” [30] in combination with estimation theory [27,30]. The conditions under which the experiments are performed (symbolized by  $\theta$  in the case at hand) are viewed as “parameter to be estimated”. This viewpoint takes the Fisher information as measure of the “resolving power” of the experiment, in our notation the answer to the question of how big the separation  $\varepsilon$  must be in order that the experiment can distinguish between  $\theta$  and  $\theta + \varepsilon$ .

This viewpoint has very little in common with the way in which actual quantum physics experiments are performed. In particular, we do not know of any SG experiment that aims at estimating the direction of the magnetic field by counting the deflected particles. On the contrary, the direction of the magnetic field is regarded as known (with limited precision of course). In the LI approach, the Fisher information appears quite naturally as a result of expressing the requirement that the experiment yields reproducible, robust results. The notion of robustness used in the present paper refers to the effect of small (systematic) changes of the condition ( $\theta$ ) on the state of knowledge encoded in the plausibilities.

There is an interesting conceptual link between the LI description of robust experiments and the theory of optimal experimental design [32]. The viewpoint taken by the latter is the following. Laboratory experiments cannot avoid errors and statistical methods may be very helpful for their design and analysis. The theory of optimal experimental design provides a framework to design experiments that are optimal with respect to some statistical criterion. It allows parameters to be estimated without bias and with minimum variance. Obviously, the optimality of a design depends on the statistical model and its assessment with respect to the statistical criterion chosen. There are several different optimality criteria, often formulated in terms of invariants of the information matrix [32], not to be confused with the Fisher information matrix, which in statistics parlance, is the covariance matrix of the derivative of the log-likelihood [27,29].

We have mentioned earlier that the LI derivation of quantum physics equations is not a kind of parameter-estimation problem but instead, there is some similarity with the optimal experimental design approach. Indeed, the LI description of a robust experiment defines an imaginary optimal

experimental design for which optimality is reached when the Fisher information (matrix) reaches its minimum. For instance, we know that the Schrödinger equation derives from the LI description of the appropriate robust experiment [9]. Turning this around we may state that the data collected by an experiment will comply with quantum theory in terms of the Schrödinger equation if the experiment satisfies the optimality criterion in terms of minimum Fisher information. If the design does not satisfy this criterion, the observations will be at odds with quantum theory.

### 3.9. Quadratic forms and Hilbert space

The plausibilities  $\{P(k, l|\theta, Z)|(k, l) \in \mathcal{E}\}$  are real-valued numbers in the range  $[0, 1]$ . Therefore, without loss of generality, we can write  $P(k, l|\theta, Z) = R^2(k, l|\theta, Z)$  where  $R(k, l|\theta, Z)$  is a real-valued number in the range  $[-1, +1]$ , and substitute this expression into Eq. (15) to obtain

$$I_F = 4 \sum_{k,l \in \mathcal{E}} \left( \frac{\partial R(k, l|\theta, Z)}{\partial \theta} \right)^2, \tag{19}$$

and

$$\sum_{k,l \in \mathcal{E}} R^2(k, l|\theta, Z) = 1. \tag{20}$$

Eq. (20) describes the surface of a hyper-sphere in a 9-dimensional space and the right-hand-side of Eq. (19) has the structure of an inner product of vectors in this space.

Alternatively, and motivated by the transformation which Madelung used to derive a hydrodynamic form of the Schrödinger equation [33] and which was also used in the derivation of the Schrödinger equation from LI principles [8–12], we introduce the complex-valued function

$$\Phi(k, l|\theta, Z) = \sqrt{P(k, l|\theta, Z)} e^{iS(k,l)} = R(k, l|\theta, Z) e^{iS(k,l)}, \tag{21}$$

where  $S(k, l)$  is a real-valued number. By straightforward algebra, it follows that

$$I_F = 4 \sum_{k,l \in \mathcal{E}} \left( \frac{\partial \Phi(k, l|\theta, Z)}{\partial \theta} \right)^* \left( \frac{\partial \Phi(k, l|\theta, Z)}{\partial \theta} \right) = 4 \left( \frac{\partial \Phi}{\partial \theta} \middle| \frac{\partial \Phi}{\partial \theta} \right), \tag{22}$$

where we have introduced a notation for the inner product

$$(\Phi | \Psi) \equiv \sum_{k,l \in \mathcal{E}} \Phi(k, l|\theta, Z)^* \Psi(k, l|\theta, Z), \tag{23}$$

of two vectors of nine elements to make it clear that the underlying mathematical structure of the problem is that of a linear vector space equipped with an inner product. As the inner product Eq. (23) is invariant under unitary transformations, performing a unitary transformation that does not depend on  $\theta$  leaves the Fisher information  $I_F$  unchanged. More specifically, if the set of functions  $\{\Phi(k, l|\theta, Z)|(k, l) \in \mathcal{E}\}$  is a solution of the problem and  $U((k, l), (k', l'))$  is a  $9 \times 9$  unitary matrix with entries that do not depend on  $\theta$ , then the set of functions

$$\tilde{\Phi}(k, l|\theta, U, Z) = \sum_{k', l' \in \mathcal{E}} U((k, l), (k', l')) \Phi(k', l'|\theta, Z), \tag{24}$$

is a solution too because unitary transformations do not change the inner product of two vectors, i.e.

$$4 \left( \frac{\partial \tilde{\Phi}}{\partial \theta} \middle| \frac{\partial \tilde{\Phi}}{\partial \theta} \right) = 4 \left( \frac{\partial \Phi}{\partial \theta} \middle| \frac{\partial \Phi}{\partial \theta} \right) = I_F, \tag{25}$$

Note that

$$|\tilde{\Phi}(k, l|\theta, U, Z)|^2 = \tilde{P}(k, l|\theta, U, Z) \neq P(k, l|\theta, U, Z) = |\Phi(k, l|\theta, U, Z)|^2, \tag{26}$$

implies that if we have a solution  $\{\Phi(k, l|\theta, Z)|(k, l) \in \mathcal{E}\}$  we can generate other, equivalent but different solutions by applying unitary transformations  $U((k, l), (k', l'))$ . From Eqs. (19) and (21)

and Eq. (26) it follows that it is sufficient to limit the search for solutions to real-valued functions  $R(k, l|\theta, U, Z)$  which satisfy

$$I_F = 4 \sum_{k,l \in \mathcal{E}} \left( \frac{\partial R(k, l|\theta, Z)}{\partial \theta} \right)^2, \quad \sum_{k,l \in \mathcal{E}} R^2(k, l|\theta, Z) = 1, \quad (27)$$

$$R(k, l|\theta, Z) \frac{\partial R(k, l|\theta, Z)}{\partial \theta} \neq 0 \quad \text{for at least one pair } (k, l) \in \mathcal{E}, \quad (28)$$

$$R(k, l|\theta, Z) = \pm R(k, l|\theta + 2\pi, Z) = \pm R(-k, l|\theta + \pi, Z) = \pm R(k, -l|\theta + \pi, Z). \quad (29)$$

$$R(k, l|\theta = 0, Z) = 0, \quad k \neq l. \quad (30)$$

where the  $\theta$ -dependence of  $R^2(k, l|\theta, Z)$  can only enter in terms of powers of  $\cos \theta$  and  $I_F$  must be independent of  $\theta$ .

### 3.10. Closed form solution

We do not know of a practical, deductive procedure to solve Eqs. (27)–(30) and obtain closed form expressions of the nine functions  $\{R(k, l|\theta, Z) | (k, l) \in \mathcal{E}\}$ . In [Appendix A](#), we reformulate the global optimization problem for a single SG magnet as a standard variational problem. The solutions of the latter all have an  $I_F$  which is independent of  $\theta$  but we also show that there exist solutions of the global optimization problem which cannot be found by solving the variational problem. In view of this, we take a pragmatic viewpoint and search for solutions of the global optimization problem in a non-deductive, constructive manner.

To this end, it is expedient to introduce a  $3 \times 3$  matrix  $\mathbf{R}(\theta)$  with elements  $R_{2-k,2-l}(\theta) = \sqrt{3}R(k, l|\theta, Z)$  (recall  $k, l = +1, 0, -1$ ) and write Eq. (27) as

$$I_F = \frac{4}{3} \text{Tr} \left( \frac{\partial \mathbf{R}(\theta)}{\partial \theta} \right)^2 \left( \frac{\partial \mathbf{R}^T(\theta)}{\partial \theta} \right)^2, \quad \frac{1}{3} \text{Tr} \mathbf{R}(\theta) \mathbf{R}^T(\theta) = 1, \quad (31)$$

where  $\text{Tr} \mathbf{X}$  denotes the trace of the  $3 \times 3$  matrix  $\mathbf{X}$ . If we set  $\mathbf{R}(\theta) = e^{\theta \mathbf{S}}$  where  $\mathbf{S} = -\mathbf{S}^T$  is a real, skew-symmetric matrix, then  $\mathbf{R}$  is an orthogonal matrix ( $\mathbf{R}^{-1}(\theta) = \mathbf{R}^T(\theta)$ ) and it follows that

$$I_F = \frac{4}{3} \text{Tr} \left( \frac{\partial \mathbf{R}(\theta)}{\partial \theta} \right)^2 = -\frac{4}{3} \text{Tr} e^{\theta \mathbf{S}} \mathbf{S} \mathbf{S} e^{-\theta \mathbf{S}} = \frac{4}{3} \text{Tr} \mathbf{S} \mathbf{S}^T, \quad (32)$$

$$\frac{1}{3} \text{Tr} \mathbf{R}(\theta) \mathbf{R}^T(\theta) = \frac{1}{3} \text{Tr} e^{\theta \mathbf{S}} e^{-\theta \mathbf{S}} = 1.$$

Clearly,  $I_F = (4/3) \text{Tr} \mathbf{S} \mathbf{S}^T$  does not depend on  $\theta$ . From the periodicity constraint  $R(k, l|\theta, Z) = \pm R(k, l|\theta + 2\pi, Z)$  it follows directly that  $e^{2\pi \mathbf{S}}$  must be equal to a diagonal matrix with elements  $\pm 1$ , implying that the eigenvalues of  $\mathbf{S}$  must be of the form  $i(n/2)$  where  $n$  can be any integer value. Therefore, in general,  $I_F = (n_1^2 + n_2^2 + n_3^2)/3$  where  $n_1, n_2$ , and  $n_3$  are integers, not all zero. Furthermore, as  $\text{Tr} \mathbf{S} = 0$  we have  $\det \mathbf{R}(\theta) = \det e^{\theta \mathbf{S}} = \exp(\theta \text{Tr} \mathbf{S}) = 1$ , meaning that the matrix  $\mathbf{R}(\theta)$  represents a rotation, not a reflection. In summary, every rotation matrix  $\mathbf{R}(\theta)$  is a solution of the global optimization problem Eq. (27) with  $I_F$  independent of  $\theta$  and each such matrix that satisfies the constraints Eqs. (28)–(30) is also a candidate for the robust description of the double SG experiment.

In general, the generators of rotations are the angular momentum operators  $J^x, J^y$ , and  $J^z$ , defined through the commutation relations  $[J^\alpha, J^\beta] = i \varepsilon_{\alpha\beta\gamma} J^\gamma$  ( $\alpha, \beta, \gamma = x, y, z$ , Einstein summation convention). The three operators  $J^x, J^y$ , and  $J^z$  are Hermitian and if the elements of two of them are

chosen to be real-valued the third one is purely imaginary. If we choose the elements of  $J^x$  and  $J^z$  to be real-valued, then  $ij^y$  is real-valued and skew-symmetric and we have  $\mathbf{S} = -iaj^y$  where  $a$  is a proportionality constant to be determined later. In the three-dimensional space spanned by the labels  $k$  or  $l$ , the angular momentum operator  $J^y$  has the matrix representation [24]

$$J^y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ +i & 0 & -i \\ 0 & +i & 0 \end{pmatrix}, \tag{33}$$

and satisfies  $(J^y)^3 = J^y$ . Using this latter property, it is easy to show that

$$\begin{aligned} e^{-ia\theta J^y} &= \mathbb{1} - ij^y \sin a\theta - (1 - \cos a\theta)(J^y)^2 \\ &= \frac{1}{2} \begin{pmatrix} 1 + \cos a\theta & -\sqrt{2} \sin a\theta & 1 - \cos a\theta \\ +\sqrt{2} \sin a\theta & 2 \cos a\theta & -\sqrt{2} \sin a\theta \\ 1 - \cos a\theta & +\sqrt{2} \sin a\theta & 1 + \cos a\theta \end{pmatrix}. \end{aligned} \tag{34}$$

From Eq. (34), it directly follows that  $a = 1, 3, 5, \dots$  because otherwise the matrix elements do not satisfy the symmetry relation  $R(k, l|\theta, Z) = \pm R(-k, l|\theta + \pi, Z) = \pm R(k, -l|\theta + \pi, Z)$ . Therefore, the eigenvalues of  $\mathbf{S} = -iaj^y$  are  $0, +ia$  and  $-ia$  and from Eq. (32) it follows that  $I_F = 8a^2/3$ . Clearly, we must choose  $a = 1$  to have the solution with the smallest, non-zero  $I_F$ . In summary, we have constructed a non-trivial solution of the global optimization problem Eqs. (27)–(30). The closed-form expression of this solution reads

$$\mathbf{R}(\theta) = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 + \cos \theta & -\sqrt{2} \sin \theta & 1 - \cos \theta \\ +\sqrt{2} \sin \theta & 2 \cos \theta & -\sqrt{2} \sin \theta \\ 1 - \cos \theta & +\sqrt{2} \sin \theta & 1 + \cos \theta \end{pmatrix}. \tag{35}$$

From Eq. (35), it follows that

$$\begin{aligned} P(k, l|\mathbf{a} \cdot \mathbf{b}, Z) &= P(k, l|\theta, Z) = R^2(k, l|\theta, Z) \\ &= \frac{1}{12} \begin{pmatrix} (1 + \mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & (1 - \mathbf{a} \cdot \mathbf{b})^2 \\ 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & 4(\mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) \\ (1 - \mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & (1 + \mathbf{a} \cdot \mathbf{b})^2 \end{pmatrix}_{2-k, 2-l} \\ &= \frac{1}{12} \begin{pmatrix} (1 + x)^2 & 2(1 - x^2) & (1 - x)^2 \\ 2(1 - x^2) & 4x^2 & 2(1 - x^2) \\ (1 - x)^2 & 2(1 - x^2) & (1 + x)^2 \end{pmatrix}_{2-k, 2-l}, \end{aligned} \tag{36}$$

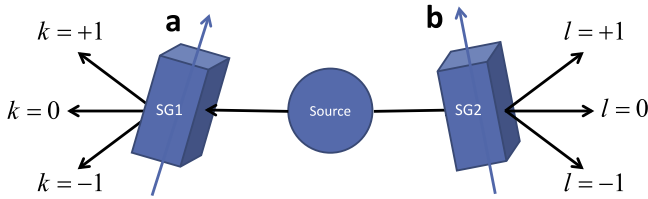
where  $x = \cos \theta = \mathbf{a} \cdot \mathbf{b}$  and the somewhat strange-looking indices of the matrix elements take care of the convention that these indices run from 1 to 3.

Substituting Eq. (35) into Eq. (27) yields  $I_F = 8/3$ , in agreement with the result obtained from Eq. (32). From Eq. (36) it follows that  $P(k, l|\mathbf{a}, \mathbf{a}, Z) = \delta_{k,l}/3$ , which is exactly the property that is needed for the SG magnet to function as an idealized filtering device, allowing us to assign a definite direction to the magnetization of the particle.

The expression Eq. (36) is in complete agreement with the result (see Appendix B, Eq. (B.8)) derived from the postulates of quantum theory. As we explained and also illustrated by a concrete example (see Appendix A), we do not have the tools to prove that the solution Eq. (36) yielding  $I_F = 8/3$  is also the solution of the global optimization problem with absolute minimum  $I_F$ , subject to the same symmetry relations. However, as we show in Appendix B, there does not exist a quantum theoretical description that satisfies the additional requirements and has an  $I_F$  that is smaller than  $8/3$ . Therefore, if there exists a solution of the global optimization problem, satisfying all additional requirements and having  $I_F < 8/3$ , this solution **cannot** be obtained from quantum theory. We conjecture that such a solution does not exist.

### 3.11. Generalization

For concreteness, this paper focuses on the case of three outcomes per SG magnet. In this particular case, the matrix Eq. (35) is the so-called Wigner- $d$  matrix for angular momentum  $J = 1$  [34],



**Fig. 2.** (Color online) Diagram of the EPRB thought experiment with spin-1 particles. A source emits a pair of particles. The particle moving to the left (right) passes through a Stern–Gerlach magnet SG1 (SG2) which maintains an inhomogeneous magnetic field characterized by a unit vector **a** (**b**). The particles moving left (right) are deflected in three different directions labeled by  $k = +1, 0, -1$  ( $l = +1, 0, -1$ ). Particle detectors (not shown) count the number of particles in each of the six beams.

suggesting that also for  $J \neq 0$ , the corresponding Wigner- $d$  matrices are the solutions of the global optimization problem Eqs. (27)–(30) when the set  $\mathcal{E}$  represents 2, 3, 4, ... outcomes. Indeed, in each step of the formulation of the problem and of the construction of the solution, with the exception of the step where we actually use the matrix representation of  $J^y$ , the number of outcomes, that is the number of elements in the set  $\mathcal{E}$ , is arbitrary. In other words, for any number of outcomes, the choice  $\mathbf{S} = -ij^y$ , the matrix elements of which are the Wigner- $d$  matrices, solves the global optimization problem Eqs. (27)–(30). If  $K$  denotes the number of elements in  $\mathcal{E}$ , it follows from Eq. (32) that for the solution  $\mathbf{S} = -ij^y$ , we have  $I_F = (K^2 - 1)/3$  for  $K = 2, 3, 4, \dots$ , respectively.

For instance, if the number of different outcomes is two instead of three, i.e.  $\mathcal{E} = \{+1, -1\}$ , we have

$$J^y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix}, \tag{37}$$

and  $(J^y)^2 = 1/4$  and the matrix  $\mathbf{S} = -iaJ^y$  reads

$$e^{-ia\theta J^y} = \mathbb{1} \cos\left(\frac{a\theta}{2}\right) - \frac{ij^y}{2} \sin\left(\frac{a\theta}{2}\right) = \begin{pmatrix} \cos(a\theta/2) & -\sin(a\theta/2) \\ \sin(a\theta/2) & \cos(a\theta/2) \end{pmatrix}. \tag{38}$$

As before, we use the symmetry relation  $R(k, l|\theta, Z) = \pm R(-k, l|\theta + \pi, Z) = \pm R(k, -l|\theta + \pi, Z)$  to determine  $a$  and find that  $a = 1$ . Therefore, for the case of a double SG experiment with two outcomes per SG magnet, the LI solution having  $I_F = 1$  reads

$$P(k, l|\theta, Z) = \frac{1}{2} \begin{pmatrix} \cos^2(\theta/2) & \sin^2(\theta/2) \\ \sin^2(\theta/2) & \cos^2(\theta/2) \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1+x & 1-x \\ 1-x & 1+x \end{pmatrix}_{i(k), i(l)}, \tag{39}$$

where  $x = \cos \theta = \mathbf{a} \cdot \mathbf{b}$ ,  $i(1) = 1$  and  $i(-1) = 2$ . Just as the LI solution for the case of 3 outcomes agrees with the  $S = 1$  expression postulated in quantum theory, Eq. (39) also agrees with the expression that is postulated in quantum theory for the  $S = 1/2$  case.

Finally, for illustrative purposes, we present the solutions for the case that each SG magnet splits the incoming beam into 4 beams. The expressions for the plausibilities read

$$P(k, l|\theta, Z) = \frac{1}{32} \begin{pmatrix} (1+x)^3 & 3(1-x)(1+x)^2 & 3(1+x)(1-x)^2 & (1-x)^3 \\ 3(1-x)(1+x)^2 & (1+x)(1-3x)^2 & (1-x)(1+3x)^2 & 3(1+x)(1-x)^2 \\ 3(1+x)(1-x)^2 & (1-x)(1+3x)^2 & (1+x)(1-3x)^2 & 3(1-x)(1+x)^2 \\ (1-x)^3 & 3(1+x)(1-x)^2 & 3(1-x)(1+x)^2 & (1+x)^3 \end{pmatrix}_{i(k), i(l)}, \tag{40}$$

where  $x = \cos \theta = \mathbf{a} \cdot \mathbf{b}$ ,  $i(2) = 1$ ,  $i(1) = 2$ ,  $i(-1) = 3$ , and  $i(-2) = 4$ .

#### 4. Einstein–Podolsky–Rosen–Bohm experiment

To head off possible misunderstandings, the material presented in this section does not contribute in any sense to the debate between Einstein and Bohr, about the foundations of quantum mechanics,

resulting in a Gedanken-experiment suggested by Einstein–Podolsky–Rosen [35] and later modified by Bohm [20].

The purpose of this section is to demonstrate that a straightforward application of the LI approach predicts exactly the same results as those obtained from the quantum theoretical treatment of the EPRB thought experiment, with the important ramification that the LI approach is void of concepts of quantum theory.

The layout of an EPRB thought experiment with spin-1 particles is shown in Fig. 2. For a fixed pair of settings  $(\mathbf{a}, \mathbf{b})$  of the SG magnets, the experiment produces a data set

$$\mathcal{D} = \{(k_n, l_n) \mid k_n, l_n \in \mathcal{E}; n = 1, \dots, N\}, \tag{41}$$

which has exactly the same structure as the data set obtained from a double SG experiment. Therefore, we only have to repeat the steps that led to the LI formulation of the double SG problem.

We begin by ignoring correlations between all pairs  $(k_n, l_n)$  and  $(k_{n'}, l_{n'})$  with  $n \neq n'$  and introduce the plausibility  $P(k, l | \mathbf{a}, \mathbf{b}, Z)$  to observe a pair  $(k, l)$  for the settings  $(\mathbf{a}, \mathbf{b})$  of the SG magnets. Then we assume that  $P(k, l | \mathbf{a}, \mathbf{b}, Z)$  does not depend on the orientation of the laboratory reference frame, implying that  $P(k, l | \mathbf{a}, \mathbf{b}, Z) = P(k, l | \mathbf{a} \cdot \mathbf{b}, Z) = P(k, l | \theta, Z)$ . Consistency of the description with the one of the double SG experiment further enforces the symmetry relations  $P(k, l | \theta + 2\pi) = P(k, l | \theta) = P(-k, l | \theta + \pi, Z) = P(k, -l | \theta + \pi, Z)$ . It should now be clear that, within the LI scheme, the mathematical formulation and solution of the global optimization problem for the case of the EPRB experiments is exactly the same as for the double SG experiment whereas their physical realization is very different.

We have already pointed out that in order to consistently assign a definite direction of the magnetization of a particle, it is necessary that a particle emerging from the SG magnet with direction  $\mathbf{a}$  in beam  $k$  and passing through a second SG magnet with direction  $\mathbf{b} = \mathbf{a}$ , emerges in beam  $l = k$  only. Only then it makes sense to say that the magnetization of the particle has a definite direction (determined by  $\mathbf{a}$ ). As explained in Section 3.10, from Eq. (36) it follows that  $P(k, l | \mathbf{a}, \mathbf{a}, Z) = \delta_{k,l}/3$ , as required for a consistent assignment of the direction of magnetization.

However, in the case of the EPRB experiment shown in Fig. 2, it is impossible to make such an assignment, as a matter of principle. For instance, if a particle leaves SG1 along beam  $k = +1$ , it is a logical fallacy to infer that the particle left the source with its magnetization along  $\mathbf{a}$ . Evidently, on the basis of the observation that a particle left SG1 along beam  $k = +1$  we cannot assign a definite direction to the magnetization of the second particle of the pair. In the absence of any other knowledge about the particles leaving the source, other than that they interact with a magnetic field, we can only speculate. Therefore, we have to leave open the possibility that the LI description of the EPRB experiment is in terms of solutions different from Eq. (36). One such other solution follows from Eq. (36) by making the substitution  $\mathbf{a} \cdot \mathbf{b} \rightarrow -\mathbf{a} \cdot \mathbf{b}$  (or  $\theta \rightarrow \theta + \pi$ ). It is easy to check that this change leaves the Fisher information unchanged, i.e.  $I_F = 8/3$ . In conclusion, we must consider the solutions  $P(k, l | \mathbf{a} \cdot \mathbf{b}, Z)$  and  $P(k, l | -\mathbf{a} \cdot \mathbf{b}, Z)$  as potential candidates for the LI description of the EPRB experiment.

In Appendix C, we give a rigorous proof that the LI solution  $P(k, l | \mathbf{a} \cdot \mathbf{b}, Z)$  cannot be obtained from quantum theory of a system of two  $S = 1/2$  or  $S = 1$  particles. This means that there exists a genuine probabilistic model, described by  $P(k, l | \mathbf{a} \cdot \mathbf{b}, Z)$ , which cannot be described within the framework of quantum theory. In Appendix C, we also show that the expressions for  $P(k, l | -\mathbf{a} \cdot \mathbf{b}, Z)$  are identical to those obtained from quantum theory for two spin-1 particles in the singlet state. In other words:

The LI framework includes quantum theory as a special case.

In the analysis of EPRB experiments, it is customary to compute the correlations between the detection event  $k$  and detection event  $l$ . From the two candidate solutions of the form Eq. (36) we find

$$\begin{aligned} \langle k \rangle &= \langle l \rangle = \langle k^2 l \rangle = \langle k l^2 \rangle = 0 \quad , \quad \langle k^2 \rangle = \langle l^2 \rangle = \frac{2}{3} \quad , \\ \langle k l \rangle &= \pm \frac{2}{3} \mathbf{a} \cdot \mathbf{b} \quad , \quad \langle k^2 l^2 \rangle = \frac{1}{3} (1 + (\mathbf{a} \cdot \mathbf{b})^2). \end{aligned} \tag{42}$$

According to quantum theory, see [Appendix C](#), we have  $\langle k \rangle = \langle \mathbf{a} \cdot \mathbf{S}_1 \rangle = 0$ ,  $\langle l \rangle = \langle \mathbf{b} \cdot \mathbf{S}_2 \rangle = 0$ ,  $\langle kl \rangle = \langle \mathbf{a} \cdot \mathbf{S}_1 \mathbf{b} \cdot \mathbf{S}_2 \rangle = -2 \mathbf{a} \cdot \mathbf{b}/3$ , etc. From Eq. (42) it follows that the correlation coefficient [28] is given by

$$\text{COR}(k, l) = \frac{\langle kl \rangle - \langle k \rangle \langle l \rangle}{\sqrt{\langle k^2 \rangle \langle l^2 \rangle}} = \pm \mathbf{a} \cdot \mathbf{b}, \quad (43)$$

independent of the value of the spin [36].

As explained above, in the case of EPRB experiments it is impossible, as a matter of principle, to assign definite directions of magnetization to the particles. However, the LI solution  $\text{COR}(k, l) = \pm \mathbf{a} \cdot \mathbf{b}$  implies that  $\text{COR}(k, l) = \pm 1$  if  $\mathbf{a} = \mathbf{b}$ , independent of  $\mathbf{a}$ . Therefore, we may say that each pair of particles leaves the source with their magnetic moments in an undefined but nevertheless perfectly correlated, parallel ( $\text{COR}(k, l) = 1$ ) or antiparallel ( $\text{COR}(k, l) = -1$ ) direction. On the basis of the data, this is all that can be inferred.

## 5. Conclusion

We have shown that the quantum-theoretical description of idealized Stern–Gerlach and Einstein–Podolsky–Rosen–Bohm experiments follow from a logical inference treatment, without having to resort to postulates or concepts of quantum theory. Following Feynman [21], a detailed derivation is given for the case that the Stern–Gerlach magnet deflects the beam in three different directions, i.e. to spin-1 particles. We also indicate how the derivation generalizes to arbitrary values of the spin.

Without going into detail, the main ingredients and results of our derivation may be summarized as follows:

1. It is assumed that each event is independent of a previous or later event.
2. It is assumed that the frequencies with which events occur are robust with respect to small changes of the conditions under which the data is being collected.
3. Knowledge about the physical aspects of the experiment enters the logical inference derivation through symmetry relations and other constraints.
4. The global minimum of the numerical measure for the robustness satisfying all constraints is found to coincide with the quantum theoretical description of the same experiment.
5. The logical inference description of the Einstein–Podolsky–Rosen–Bohm experiment is shown to contain the quantum theoretical description as a particular case.
6. The derivation directly proceeds from the data representing the events to the mathematical description in terms of the plausibilities to observe the data in robust experiments and does not suffer from the interpretational issues that bedevil quantum theory.

The power of the logical inference framework that we use stems from the fact that it is well-suited to bridge the gap between the experiences that we accumulate through our sensory system and the mental representations in terms of mathematical concepts that we construct and use to describe these experiences. In particular, it provides a solid mathematical basis to discuss Bohr’s philosophical view that “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 [13]”. The latter, we believe, is exactly what we have accomplished for the case of the Stern–Gerlach and Einstein–Podolsky–Rosen–Bohm experiments in this paper and for other basic quantum physics equations in our earlier work [8–12].

## Acknowledgments

The work of MIK is supported by the European Research Council (ERC) Advanced Grant No. 338957 FEMTO/NANO. We thank Dennis Willsch for helping us out with Mathematica<sup>®</sup>. We are grateful to Dennis Willsch and Madita Nocon for their critical reading of the manuscript.



**Appendix A. Variational solution of the LI problem of the SG experiment**

We focus on one particular beam in the double SG setup. According to the LI formulation of a robust experiment, we should search for the solutions  $\{R(l|\theta, Z)|l \in \mathcal{E}\}$  of

$$I_F = 4 \sum_{l \in \mathcal{E}} \left( \frac{\partial R(l|\theta, Z)}{\partial \theta} \right)^2, \tag{A.1}$$

with  $I_F$  independent of  $\theta$  and satisfying the normalization condition

$$\sum_{l \in \mathcal{E}} R^2(l|\theta, Z) = 1, \tag{A.2}$$

Let us search for a solution of the form

$$R(l|\theta, Z) = \begin{cases} \sin \alpha(\theta) \cos \beta(\theta) & , \quad l = +1 \\ \cos \alpha(\theta) & , \quad l = 0 \\ \sin \alpha(\theta) \sin \beta(\theta) & , \quad l = -1 \end{cases}, \tag{A.3}$$

which trivially satisfies  $\sum_{l \in \mathcal{E}} R^2(l|\theta, Z) = 1$  and upon substitution in Eq. (A.1) yields

$$I_F = 4 \left[ \left( \frac{d\alpha}{d\theta} \right)^2 + \sin^2 \alpha \left( \frac{d\beta}{d\theta} \right)^2 \right], \tag{A.4}$$

where we introduced the shorthands  $\alpha = \alpha(\theta)$  and  $\beta = \beta(\theta)$ . The problem is to determine  $\alpha$  and  $\beta$  such that  $I_F$  is a constant.

This particular problem can be solved by introducing

$$\widehat{I}_F = \frac{4}{2\pi} \int_0^{2\pi} \left[ \left( \frac{d\alpha}{d\theta} \right)^2 + \sin^2 \alpha \left( \frac{d\beta}{d\theta} \right)^2 \right] d\theta, \tag{A.5}$$

and employing standard variational calculus. By variation with respect to  $\alpha$  and  $\beta$ , we find the Euler-Lagrange equations

$$\frac{d^2 \alpha}{d\theta^2} - \sin \alpha \cos \alpha \left( \frac{d\beta}{d\theta} \right)^2 = 0 \quad , \quad \frac{d}{d\theta} \left( \sin^2 \alpha \frac{d\beta}{d\theta} \right) = 0, \tag{A.6}$$

from which it follows that

$$\frac{d^2 \alpha}{d\theta^2} - c^2 \frac{\cos \alpha}{\sin^3 \alpha} = 0 \quad , \quad \frac{d\beta}{d\theta} - \frac{c}{\sin^2 \alpha} = 0, \tag{A.7}$$

where  $c$  is a constant of integration. From Eq. (A.7) it follows that

$$\begin{aligned} \frac{d\alpha}{d\theta} \left( \frac{d^2 \alpha}{d\theta^2} - c^2 \frac{\cos \alpha}{\sin^3 \alpha} \right) &= \frac{1}{2} \frac{d}{d\theta} \left[ \left( \frac{d\alpha}{d\theta} \right)^2 + \frac{c^2}{\sin^2 \alpha} \right] \\ &= \frac{1}{2} \frac{d}{d\theta} \left[ \left( \frac{d\alpha}{d\theta} \right)^2 + \sin^2 \alpha \left( \frac{d\beta}{d\theta} \right)^2 \right], \end{aligned} \tag{A.8}$$

and therefore

$$E = \left( \frac{d\alpha}{d\theta} \right)^2 + \sin^2 \alpha \left( \frac{d\beta}{d\theta} \right)^2 = \frac{1}{4} I_F, \tag{A.9}$$

is independent of  $\theta$  whenever  $\alpha$  and  $\beta$  are solutions of Eq. (A.7). This is just a restatement of the well-known fact that the energy of a conservative classical mechanical system is a constant of motion. In the case at hand, the Lagrangian, the integrand in Eq. (A.5), consists of kinetic energy terms only. Hence the expression of the energy (Hamiltonian) and Lagrangian are the same.

Integrating the equation

$$\left(\frac{d\alpha}{d\theta}\right)^2 + \frac{c^2}{\sin^2\alpha} = E, \tag{A.10}$$

we find that

$$\cos\alpha = \sqrt{1 - \frac{c^2}{E}} \cos(\sqrt{E}\theta + \alpha_0), \tag{A.11}$$

where  $\alpha_0$  is a constant of integration. Similarly, integrating the equation

$$\frac{d\beta}{d\theta} = \frac{c}{\sin^2\alpha}, \tag{A.12}$$

we find that

$$\tan(\beta - \beta_0) = \frac{\sqrt{E}}{c} \tan(\sqrt{E}\theta + \alpha_0), \tag{A.13}$$

where  $\beta_0$  is a constant of integration.

It is instructive to consider a simple case. If  $c = 0$  then  $\beta = \beta_1$  does not depend on  $\theta$  and  $\cos\alpha = \cos(\sqrt{E}\theta + \alpha_0)$ . As  $P(l|\theta, Z)$  is a function of  $\cos^2\alpha$  (see Eq. (A.4) and must be equal to  $P(l|\theta + 2\pi, Z)$ , we must have  $E = n^2/4$  where  $n$  is an integer ( $n = 0$  is ruled out because of being non-informative). In other words, for  $c = 0$ , we found solutions of the global optimization problem with  $I_F = n^2$ , independent of  $\theta$ . However, the solution having  $I_F = 1$  reads

$$P(l|\theta, Z) = \begin{cases} (1 - \cos^2(\theta/2 + \alpha_0))\cos^2\beta_1 & , \quad l = +1 \\ \cos^2(\theta/2 + \alpha_0) & , \quad l = 0 \\ (1 - \cos^2(\theta/2 + \alpha_0))\sin^2\beta_1 & , \quad l = -1 \end{cases}, \tag{A.14}$$

and does not exhibit the reflection symmetry  $P(l|\theta, Z) = P(-l|\theta + \pi, Z)$ . Therefore it must be discarded. The solutions with  $I_F \geq 4$  satisfy all symmetry requirements.

Let us now try as a solution, say the first column

$$R(l|\theta, Z) = \begin{cases} (1 + \cos\theta)/2 & , \quad l = +1 \\ (1/\sqrt{2})\sin\theta & , \quad l = 0 \\ (1 - \cos\theta)/2 & , \quad l = -1 \end{cases}, \tag{A.15}$$

of the LI solution Eq. (35). It is easy to check that Eq. (A.15) is a valid LI solution of Eqs. (A.1) and (A.2) with  $I_F = 2$ . Having constructed a valid solution, we now ask ourselves how it relates to the solutions Eqs. (A.11) and (A.13) that were deduced by variational calculus. Equating Eqs. (A.3) and (A.15) we have

$$\cos\alpha = \frac{\sin\theta}{\sqrt{2}} = \sqrt{1 - \frac{c^2}{E}} \cos(\sqrt{E}\theta + \alpha_0) \Rightarrow E = 1, c = \frac{1}{\sqrt{2}}, \alpha_0 = -\frac{\pi}{2}, \tag{A.16}$$

which implies that this solution has  $I_F = 4E = 4$  and can therefore never match the solution Eq. (A.15) which has  $I_F = 2$ . This conclusion is confirmed by trying to match the solution for  $\beta$ . We have

$$\begin{aligned} \cos\beta &= \frac{1 + \cos\theta}{\sqrt{1 + \cos^2\theta}}, \quad \sin\beta = \frac{1 - \cos\theta}{\sqrt{1 + \cos^2\theta}}, \quad \tan\beta = \frac{1 - \cos\theta}{1 + \cos\theta}, \\ \tan(\beta - \beta_0) &= \frac{1}{\sqrt{2}} \cot\theta \Rightarrow \tan\beta_0 = \frac{\sqrt{2} \tan\beta - \cot\theta}{\sqrt{2} + \tan\beta \cot\theta} \\ \tan\beta_0 &= \frac{\sqrt{2} \sin\theta - \cos\theta - (\sqrt{2} \sin\theta + \cos\theta) \cos\theta}{\sqrt{2} \sin\theta + \cos\theta + (\sqrt{2} \sin\theta + \cos\theta) \cos\theta} \neq \text{constant}, \end{aligned} \tag{A.17}$$

from which it follows that the solution Eq. (A.15) of the global optimization problem is **not** a solution of the Euler–Lagrange equations (A.7).

Reformulating the global optimization problem as a standard variational problem, we can generate many **but apparently not all** solutions of the former. In other words, there are solutions to the problem “ $I_F$  independent of  $\theta$ ” which cannot be found by standard variational calculus.

**Appendix B. Quantum theory of the (double) SG experiment**

For the convenience of the reader, we briefly review the quantum theoretical description of the outcome of a (double) SG experiment with spin-1 particles. The description starts by introducing the  $S = 1$  matrices [21]

$$\begin{aligned}
 S^x &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad S^y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ +i & 0 & -i \\ 0 & +i & 0 \end{pmatrix}, \\
 S^z &= \begin{pmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},
 \end{aligned}
 \tag{B.1}$$

which represent the three components of the  $S = 1$  operator [21]. Furthermore we have

$$\begin{aligned}
 (S^x)^2 &= \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad (S^y)^2 = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \\
 (S^z)^2 &= \begin{pmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & +1 \end{pmatrix},
 \end{aligned}
 \tag{B.2}$$

and  $(S^x)^3 = S^x$ ,  $(S^y)^3 = S^y$ , and  $(S^z)^3 = S^z$ .

Next, we consider the matrices that describe only those particles which travel along a particular beam  $k$ . We first consider the case in which  $\mathbf{a} = \mathbf{e}_z$ . By construction we have

$$\begin{aligned}
 \mathbf{M}_k(\mathbf{e}_z) &= \mathbb{1} - (S^z)^2 + \frac{k}{2} S^z + \frac{k^2}{2} [3(S^z)^2 - 2\mathbb{1}] = \begin{pmatrix} \frac{k^2+k}{2} & 0 & 0 \\ 0 & 1-k^2 & 0 \\ 0 & 0 & \frac{k^2-k}{2} \end{pmatrix} \\
 &= \begin{cases} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & k = +1 \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & k = 0 \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & k = -1 \end{cases} .
 \end{aligned}
 \tag{B.3}$$

From Eq. (B.3) it is clear that a particle that is being deflected into the beam  $k$  can be associated with the projector  $\mathbf{M}_k(\mathbf{e}_z)$  for  $k \in \mathcal{E} = \{+1, 0, -1\}$  and that  $\mathbf{M}_k(\mathbf{e}_z)\mathbf{M}_l(\mathbf{e}_z) = \delta_{k,l}\mathbf{M}_k(\mathbf{e}_z)$ , i.e. the  $\mathbf{M}_k(\mathbf{e}_z)$ 's are three mutually orthogonal projectors.

Performing a rotation that changes  $\mathbf{e}_z$  into  $\mathbf{a}$  changes  $S^z$  into  $\mathbf{a} \cdot \mathbf{S}$  and therefore, the general expression of the spin projector along  $\mathbf{a}$  reads

$$\mathbf{M}_k(\mathbf{a}) = 1 - (\mathbf{a} \cdot \mathbf{S})^2 + \frac{k}{2} \mathbf{a} \cdot \mathbf{S} + \frac{k^2}{2} [3(\mathbf{a} \cdot \mathbf{S})^2 - 2\mathbb{1}],
 \tag{B.4}$$

which are three mutually orthogonal projectors as well. As Eq. (B.4) follows from Eq. (B.3) by unitary transformation,  $\mathbf{M}_k(\mathbf{a})$  is a Hermitian matrix and  $\text{Tr } \mathbf{M}_k(\mathbf{a}) = 1$ .

The particles that are sent to the first SG magnet yield a certain statistical distribution of the possible outcomes  $k \in \mathcal{E}$ . In general, this statistical distribution may depend on the direction  $\mathbf{a}$ . A complete description of the particles entering the SG magnet requires a specification of the statistical distribution for all possible  $\mathbf{a}$ . Quantum theory postulates such a specification by way of the density matrix  $\rho$ , a Hermitian matrix with  $\text{Tr } \rho = 1$  [24]. Then, still according to the postulates of quantum theory, the probability to observe a particle in the beam  $k$  is given by [24]

$$p(k|\mathbf{a}, \rho) = \text{Tr } \mathbf{M}_k(\mathbf{a})\rho\mathbf{M}_k(\mathbf{a}) = \text{Tr } \rho\mathbf{M}_k(\mathbf{a}). \tag{B.5}$$

Similarly, the probability to observe a particle in the beam  $(k, l)$  is given by [24]

$$p(k, l|\mathbf{a}, \mathbf{b}, \rho) = \text{Tr } \mathbf{M}_l(\mathbf{b})\mathbf{M}_k(\mathbf{a})\rho\mathbf{M}_k(\mathbf{a})\mathbf{M}_l(\mathbf{b}) = \text{Tr } \rho\mathbf{M}_k(\mathbf{a})\mathbf{M}_l(\mathbf{b})\mathbf{M}_k(\mathbf{a}). \tag{B.6}$$

As an illustrative example, we consider the choice  $\rho = \mathbb{1}/3$  and calculate the right-hand-sides of Eqs. (B.5) and (B.6). We have

$$p(k|\mathbf{a}, \rho = \mathbb{1}/3) = \frac{1}{3}, \quad \langle \mathbf{a} \cdot \mathbf{S} \rangle = \text{Tr } \rho \mathbf{a} \cdot \mathbf{S} = 0, \quad \langle (\mathbf{a} \cdot \mathbf{S})^2 \rangle = \text{Tr } \rho (\mathbf{a} \cdot \mathbf{S})^2 = 1, \tag{B.7}$$

all independent of the direction  $\mathbf{a}$ . Thus, we could say that the choice  $\rho = \mathbb{1}/3$  describes a situation in which an experiment with one SG magnet splits the incoming beam in three different beams which have the same intensity, independent of the direction  $\mathbf{a}$  of the SG magnet. From Eq. (B.6) we obtain

$$p(k, l|\mathbf{a}, \mathbf{b}, \rho = \mathbb{1}/3) = \frac{1}{12} \begin{pmatrix} (1 + \mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & (1 - \mathbf{a} \cdot \mathbf{b})^2 \\ 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & 4(\mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) \\ (1 - \mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & (1 + \mathbf{a} \cdot \mathbf{b})^2 \end{pmatrix}_{2-k, 2-l}. \tag{B.8}$$

Using the expression Eq. (B.8) to compute the Fisher information, we find that  $I_F = 8/3$ , in concert with the LI derivation.

Next, we prove that quantum theory of the double SG magnet cannot yield a value of  $I_F$  that is smaller than  $8/3$  if we require that  $p(k, l|\mathbf{a}, \mathbf{b})$  is a function of  $\mathbf{a} \cdot \mathbf{b}$  only. First, we note that by construction

$$\mathbf{M}_k(\mathbf{a}) = R(\mathbf{a})|k\rangle\langle k|R^\dagger(\mathbf{a}), \tag{B.9}$$

where  $|k\rangle$  denotes the eigenstate of the  $S^z$  matrix with eigenvalue  $k \in \mathcal{E}$  and  $R(\mathbf{a})$  is the  $3 \times 3$  matrix that corresponds to the rotation which changes the unit vector  $\mathbf{e}_z$  into the unit vector  $\mathbf{a}$ . From Eq. (B.6) it follows that

$$\begin{aligned} p(k, l|\mathbf{a}, \mathbf{b}, \rho) &= \text{Tr } R(\mathbf{b})|l\rangle\langle l|R^\dagger(\mathbf{b})R(\mathbf{a})|k\rangle\langle k|R^\dagger(\mathbf{a})\rho R(\mathbf{a})|k\rangle\langle k|R^\dagger(\mathbf{a})R(\mathbf{b})|l\rangle\langle l|R^\dagger(\mathbf{b}) \\ &= \langle l|R^\dagger(\mathbf{b})R(\mathbf{a})|k\rangle\langle k|R^\dagger(\mathbf{a})\rho R(\mathbf{a})|k\rangle\langle k|R^\dagger(\mathbf{a})R(\mathbf{b})|l\rangle \\ &= |\langle l|R^\dagger(\mathbf{b})R(\mathbf{a})|k\rangle|^2 \langle k|R^\dagger(\mathbf{a})\rho R(\mathbf{a})|k\rangle. \end{aligned} \tag{B.10}$$

If we take  $\rho = \mathbb{1}/3$ , the last factor in Eq. (B.10) is equal to  $1/3$  and from Eq. (B.8) it follows immediately that

$$|\langle l|R^\dagger(\mathbf{b})R(\mathbf{a})|k\rangle|^2 = 3p(k, l|\mathbf{a}, \mathbf{b}, \rho = \mathbb{1}/3), \tag{B.11}$$

is a function of  $\mathbf{a} \cdot \mathbf{b}$  only. In order that  $p(k, l|\mathbf{a}, \mathbf{b}, \rho)$  is a function of  $\mathbf{a} \cdot \mathbf{b}$  only for all  $\rho$ , we must have that the last factor in Eq. (B.10), i.e.  $\langle k|R^\dagger(\mathbf{a})\rho R(\mathbf{a})|k\rangle$ , is independent of  $\mathbf{a}$  for each of the three basis vectors  $|k\rangle$ . The only nonzero Hermitian matrix  $\rho$  with this property is  $\rho = c\mathbb{1}$  where  $c$  is a nonzero constant. This can be seen as follows. Consider an arbitrary Hermitian matrix  $A$  with elements  $A_{i,j}$  and apply a rotation that involves elements with row indices  $i_0 < i_1$  and column indices  $j_0 < j_1$ . We have

$$\begin{aligned} &\begin{pmatrix} \cos \varphi & e^{-i\gamma} \sin \varphi \\ -e^{+i\gamma} \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} A_{i_0 j_0} & A_{i_0 j_1} \\ A_{i_1 j_0}^* & A_{i_1 j_1} \end{pmatrix} \begin{pmatrix} \cos \varphi & -e^{-i\gamma} \sin \varphi \\ e^{+i\gamma} \sin \varphi & \cos \varphi \end{pmatrix} \\ &= \begin{pmatrix} A'_{i_0 j_0} & A'_{i_0 j_1} \\ A'_{i_1 j_0} & A'_{i_1 j_1} \end{pmatrix}, \end{aligned} \tag{B.12}$$

where  $A'_{i_0,j_0} = A_{i_0,j_0} \cos^2 \varphi + A_{i_1,j_1} \sin^2 \varphi + (A_{i_0,j_1} e^{+i\gamma} + A_{i_0,j_1}^* e^{-i\gamma}) \sin \varphi \cos \varphi$  and the expressions of the other elements are similar but, for the present purpose, not of interest. The requirement that the diagonal matrix elements are invariant under any rotation reads  $A'_{i_0,j_0} = A_{i_0,j_0}$  which can be rewritten as  $(A_{i_1,j_1} - A_{i_0,j_0}) \sin \varphi + [(A_{i_0,j_1} + A_{i_0,j_1}^*) \cos \gamma + i(A_{i_0,j_1} - A_{i_0,j_1}^*) \sin \gamma] \cos \varphi = 0$ . As the latter equation must hold for any choice of  $\varphi$  and  $\gamma$ , we must have  $A_{i_1,j_1} = A_{i_0,j_0}$  and  $A_{i_0,j_1} = 0$ . Repeating this procedure for all choices of  $(i_0, j_0)$  and  $(i_1, j_1)$  (with  $i_0 < i_1$  and  $j_0 < j_1$ ) completes the proof.

Finally, it is of interest to mention here that the double SG experiment provides a counter example to the folklore that in quantum theory, the eigenvalues of non-commuting observables cannot be measured simultaneously. In general,  $\mathbf{a} \cdot \mathbf{S}$  and  $\mathbf{b} \cdot \mathbf{S}$  do not commute yet in the double SG experiment, for each particle, we can read off the eigenvalue  $k$  of  $\mathbf{a} \cdot \mathbf{S}$  and the eigenvalue  $l$  of  $\mathbf{b} \cdot \mathbf{S}$  from the label  $(k, l)$  of the detector that fired.

### Appendix C. Quantum theory of the EPRB thought experiment with S=1 particles

According to quantum theory, the state of a system of two S=1 particles is represented by the density matrix  $\rho$ , a Hermitian  $9 \times 9$  matrix with trace equal to one. Unlike in the case of the double SG experiment, the requirement that the expectation values be rotational invariant, i.e. only depend on  $\mathbf{a} \cdot \mathbf{b}$ , does not imply that the density matrix is proportional to the unit matrix. As a concrete example, we consider the density matrix  $\rho_{12} = |\Psi\rangle\langle\Psi|$  where the pure state

$$|\Psi\rangle = \frac{1}{\sqrt{3}} (|-1, 1\rangle - |0, 0\rangle + |+1, -1\rangle), \tag{C.1}$$

is an eigenstate of the total spin  $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$  with eigenvalue zero and therefore rotational invariant.

Consistency demands that the expression of the projector

$$\mathbf{M}_k(\mathbf{a}, \mathbf{S}_i) = 1 - (\mathbf{a} \cdot \mathbf{S}_i)^2 + \frac{k}{2} \mathbf{a} \cdot \mathbf{S}_i + \frac{k^2}{2} [3(\mathbf{a} \cdot \mathbf{S}_i)^2 - 2 \mathbb{1}], \tag{C.2}$$

is the same as in the case of SG experiments. In Eq. (C.2), the spin  $\mathbf{S}_i$  appears as an argument because in an EPRB experiment, we have to distinguish between the spin operator that we assign to the left- and right-going particle, see Fig. 2.

The probability to observe a left-going particle in beam  $k$  (or  $l$ ) is given by

$$p(k|\mathbf{a}, \rho_{12} = |\Psi\rangle\langle\Psi|) = \text{Tr } \mathbf{M}_k(\mathbf{a}, \mathbf{S}_1) \rho_{12} \mathbf{M}_k(\mathbf{a}, \mathbf{S}_1) = \text{Tr } \rho \mathbf{M}_k(\mathbf{a}, \mathbf{S}_1) = \frac{1}{3}. \tag{C.3}$$

For the right-going particle, the same expression holds with  $k, \mathbf{a}$  and  $\mathbf{S}_1$  replaced by  $l, \mathbf{b}$  and  $\mathbf{S}_2$ , respectively. The joint probability to observe a particle in beam  $k$  and a particle in beam  $l$  is given by

$$\begin{aligned} p(k, l|\mathbf{a}, \mathbf{b}, \rho_{12} = |\Psi\rangle\langle\Psi|) &= \text{Tr } \mathbf{M}_l(\mathbf{b}, \mathbf{S}_2) \mathbf{M}_k(\mathbf{a}, \mathbf{S}_1) \rho_{12} \mathbf{M}_k(\mathbf{a}, \mathbf{S}_1) \mathbf{M}_l(\mathbf{b}, \mathbf{S}_2) \\ &= \text{Tr } \rho_{12} \mathbf{M}_k(\mathbf{a}, \mathbf{S}_1) \mathbf{M}_l(\mathbf{b}, \mathbf{S}_2) \\ &= \frac{1}{12} \begin{pmatrix} (1 - \mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & (1 + \mathbf{a} \cdot \mathbf{b})^2 \\ 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & 4(\mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) \\ (1 + \mathbf{a} \cdot \mathbf{b})^2 & 2(1 - (\mathbf{a} \cdot \mathbf{b})^2) & (1 - \mathbf{a} \cdot \mathbf{b})^2 \end{pmatrix}_{2-k, 2-l} \\ &= P(k, l - \mathbf{a} \cdot \mathbf{b}, Z), \end{aligned} \tag{C.4}$$

which is identical to Eq. (B.8) with  $\mathbf{a} \cdot \mathbf{b}$  replaced by  $-\mathbf{a} \cdot \mathbf{b}$ . With  $\rho_{12} = |\Psi\rangle\langle\Psi|$  and  $\Psi|$  given by Eq. (C.1), we have

$$\begin{aligned} \langle \mathbf{a} \cdot \mathbf{S}_1 \rangle &= \text{Tr } \rho_{12} \mathbf{a} \cdot \mathbf{S}_1 = 0, \quad \langle \mathbf{b} \cdot \mathbf{S}_2 \rangle = \text{Tr } \rho_{12} \mathbf{b} \cdot \mathbf{S}_2 = 0, \\ \langle \mathbf{a} \cdot \mathbf{S}_1 \mathbf{b} \cdot \mathbf{S}_2 \rangle &= \text{Tr } \rho \mathbf{a} \cdot \mathbf{S}_1 \mathbf{b} \cdot \mathbf{S}_2 = -\frac{2}{3} \mathbf{a} \cdot \mathbf{b}, \\ \langle (\mathbf{a} \cdot \mathbf{S}_1)^2 \mathbf{b} \cdot \mathbf{S}_2 \rangle &= \text{Tr } \rho_{12} (\mathbf{a} \cdot \mathbf{S}_1)^2 \mathbf{b} \cdot \mathbf{S}_2 = 0, \\ \langle \mathbf{a} \cdot \mathbf{S}_1 (\mathbf{b} \cdot \mathbf{S}_2)^2 \rangle &= \text{Tr } \rho_{12} \mathbf{a} \cdot \mathbf{S}_1 (\mathbf{b} \cdot \mathbf{S}_2)^2 = 0, \\ \langle (\mathbf{a} \cdot \mathbf{S}_1)^2 (\mathbf{b} \cdot \mathbf{S}_2)^2 \rangle &= \text{Tr } \rho_{12} (\mathbf{a} \cdot \mathbf{S}_1)^2 (\mathbf{b} \cdot \mathbf{S}_2)^2 = \frac{1}{3} (1 + (\mathbf{a} \cdot \mathbf{b})^2). \end{aligned} \tag{C.5}$$

The expressions Eq. (C.5) are identical to those obtained from Eq. (36) with  $\mathbf{a} \cdot \mathbf{b}$  replaced by  $-\mathbf{a} \cdot \mathbf{b}$ . Can quantum theory also yield the expressions that we obtained from Eq. (36) without this replacement? To study this question, we consider the inverse problem of determining the matrix  $\hat{\rho}$  such that all equations

$$\begin{aligned} \text{Tr } \hat{\rho} &= 1 \quad , \quad \text{Tr } \hat{\rho} \mathbf{a} \cdot \mathbf{S}_1 = 0 \quad , \quad \text{Tr } \hat{\rho} \mathbf{b} \cdot \mathbf{S}_2 = 0 \quad , \quad \text{Tr } \hat{\rho} \mathbf{a} \cdot \mathbf{S}_1 \mathbf{b} \cdot \mathbf{S}_2 = \frac{2}{3} \mathbf{Aa} \cdot \mathbf{b}, \\ \text{Tr } \hat{\rho} (\mathbf{a} \cdot \mathbf{S}_1)^2 \mathbf{b} \cdot \mathbf{S}_2 &= 0 \quad , \quad \text{Tr } \hat{\rho} \mathbf{a} \cdot \mathbf{S}_1 (\mathbf{b} \cdot \mathbf{S}_2)^2 = 0 \quad , \\ \text{Tr } \hat{\rho} (\mathbf{a} \cdot \mathbf{S}_1)^2 (\mathbf{b} \cdot \mathbf{S}_2)^2 &= \frac{1}{3} (1 + (\mathbf{Aa} \cdot \mathbf{b})^2), \end{aligned} \tag{C.6}$$

are satisfied for all  $\mathbf{a}$  and  $\mathbf{b}$ . The right-hand-sides in Eq. (C.6) have been obtained from the LI solution  $P(k, l | \mathbf{Aa} \cdot \mathbf{b}, Z)$  with  $A = +1$  or  $A = -1$ . With the help of Mathematica® we find

$$\hat{\rho} = \frac{1}{6} \begin{pmatrix} 1+A & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1+A & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-A & 0 & A-1 & 0 & 2 & 0 & 0 \\ 0 & 1+A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A-1 & 0 & 2 & 0 & A-1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1+A & 0 \\ 0 & 0 & 2 & 0 & A-1 & 0 & 1-A & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1+A & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1+A \end{pmatrix}, \tag{C.7}$$

which has eigenvalues  $(-1 - A, -1 - A, -1 - A, 4 - 2A, 1 + A, 1 + A, 1 + A, 1 + A)/6$ . For  $A = +1$ , three eigenvalues are negative implying that in this case,  $\hat{\rho}$  does not qualify as a density matrix in quantum theory. For  $A = -1$ , the eigenvalues are  $(0, 0, 0, 1, 0, 0, 0, 0)$  implying  $\hat{\rho}$  represents a pure state, i.e.  $\hat{\rho} = |\Psi\rangle\langle\Psi|$  with  $|\Psi\rangle$  given by Eq. (C.1). In summary: quantum theory cannot describe EPRB experiments for which the probability to observe a pair is given by  $P(k, l | \mathbf{a} \cdot \mathbf{b}, Z)$ . This is also true for the case of two instead of three outcomes per SG magnet (details of the proof are omitted). Whether it is possible to realize a laboratory experiment with  $S = 1/2$  particles yielding a correlation  $+\mathbf{a} \cdot \mathbf{b}$  instead of  $-\mathbf{a} \cdot \mathbf{b}$  is an open question. What is beyond doubt is that there exists a genuine probabilistic model, described by  $P(k, l | \mathbf{a} \cdot \mathbf{b}, Z)$  which cannot be described within the framework of quantum theory. In other words, the LI framework includes quantum theory, not vice versa.

**References**

- [1] R.T. Cox, Amer. J. Phys. 14 (1946) 1–13.
- [2] R.T. Cox, The Algebra of Probable Inference, Johns Hopkins University Press, Baltimore, 1961.
- [3] M. Tribus, Rational Descriptions, Decisions and Designs, Expira Press, Stockholm, 1999.
- [4] C.R. Smith, G. Erickson, in: J. Skilling (Ed.), Maximum Entropy and Bayesian Methods, Kluwer Academic Publishers, Dordrecht, 1989, pp. 29–44.
- [5] E.T. Jaynes, Probability Theory: The Logic of Science, Cambridge University Press, Cambridge, 2003.
- [6] E.T. Jaynes, Phys. Rev. 106 (1957) 620–640.
- [7] E.T. Jaynes, Phys. Rev. 108 (1957) 171–190.
- [8] H. De Raedt, M.I. Katsnelson, K. Michielsen, Proc. SPIE 8832 (2013) 883212–1–883212–11.
- [9] H. De Raedt, M.I. Katsnelson, K. Michielsen, Ann. Phys. 347 (2014) 45–73.
- [10] H. De Raedt, M.I. Katsnelson, H.C. Donker, K. Michielsen, Ann. Phys. 359 (2015) 166–186.
- [11] H. De Raedt, M.I. Katsnelson, H.C. Donker, K. Michielsen, Proc. SPIE 9570 (2015) 95700–1–95700–14.
- [12] H. De Raedt, M.I. Katsnelson, K. Michielsen, Phil. Trans. R. Soc. A 374 (2016) 20150233.
- [13] N. Bohr, in: David Favrholdt (Ed.), Complementarity beyond Physics (1928 – 1962), in: Niels Bohr Collected Works, vol. 10, Elsevier, Amsterdam, 1999, pp. 155–160.
- [14] V.B. Berestetskii, E.M. Lifshitz, L.P. Pitaevskii, Relativistic Quantum Theory, vol. 1, Pergamon, Oxford, 1971.
- [15] W. Gerlach, O. Stern, Z. Phys. 9 (1922) 349–352.
- [16] J.E. Sherwood, T.K. Stephenson, S. Bernstein, Phys. Rev. 96 (1954) 1546–1548.
- [17] B. Hamelin, N. Xiromeritis, P. Liaud, Nucl. Instrum. Methods 125 (1975) 79–84.
- [18] H. Batelaan, T.J. Gay, J.J. Schwendiman, Phys. Rev. Lett. 79 (1997) 4517–4521.
- [19] A. Diaz-Bachs, M.I. Katsnelson, A. Kirilyuk, New J. Phys. 20 (2018) 043042.
- [20] D. Bohm, Quantum Theory, Prentice-Hall, New York, 1951.
- [21] R.P. Feynman, R.B. Leighton, M. Sands, The Feynman Lectures on Physics, Vol. 3, Addison-Wesley, Reading MA, 1965.

- [22] G. Pólya, *Mathematics and Plausible Reasoning*, Princeton University Press, Princeton, 1954.
- [23] G. Baym, *Lectures on Quantum Mechanics*, W.A. Benjamin, Reading MA, 1974.
- [24] L.E. Ballentine, *Quantum Mechanics: A Modern Development*, World Scientific, Singapore, 2003.
- [25] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955.
- [26] G. Potel, F. Barranco, S. Cruz-Barrios, J. Gómez-Camacho, *Phys. Rev. A* 71 (2005) 052106.
- [27] H.L. Van Trees, *Detection, Estimation, and Modulation Theory (Part I)*, John Wiley, New York, 1968.
- [28] R. Barlow, *Statistics. a Guide to the Use of Statistical Methods in the Physical Sciences*, Wiley, Chichester, UK, 1989.
- [29] K.V. Mardia, J.T. Kent, J.M. Bibby, *Multivariate Analysis*, Academic Press, Amsterdam, 2003.
- [30] B.R. Frieden, *Science from Fisher Information: A Unification*, Cambridge University Press, Cambridge, 2004.
- [31] B.R. Frieden, *Amer. J. Phys.* 57 (1989) 1004–1008.
- [32] A. Atkinson, A. Donev, R. Tobias, *Optimum Experimental Designs, with SAS*, Oxford University Press, Oxford, UK, 2007.
- [33] E. Madelung, *Z. Phys.* 40 (1927) 322–326.
- [34] E.P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York, 1959.
- [35] A. Einstein, A. Podolsky, N. Rosen, *Phys. Rev.* 47 (1935) 777–780.
- [36] N.D. Mermin, *Phys. Rev. D* 22 (1986) 356–361.