

2019

# Why Physical Understanding Should Precede the Mathematical Formalism—Conditional Quantum Probabilities as a Case-Study

Yakir Aharonov  
*Chapman University*, [aharonov@chapman.edu](mailto:aharonov@chapman.edu)

Eliahu Cohen  
*Bar Ilan University*

David H. Oaknin  
*Rafael Advanced Defense Systems Ltd.*

Follow this and additional works at: [https://digitalcommons.chapman.edu/scs\\_articles](https://digitalcommons.chapman.edu/scs_articles)

 Part of the [Other Physics Commons](#), and the [Quantum Physics Commons](#)

## Recommended Citation

Yakir Aharonov, Eliahu Cohen, and David H. Oaknin, "Why physical understanding should precede the mathematical formalism—Conditional quantum probabilities as a case-study," *Am. J. Phys.* **87**, 668–673 (2019). <https://doi.org/10.1119/1.5115980>

This Article is brought to you for free and open access by the Science and Technology Faculty Articles and Research at Chapman University Digital Commons. It has been accepted for inclusion in Mathematics, Physics, and Computer Science Faculty Articles and Research by an authorized administrator of Chapman University Digital Commons. For more information, please contact [laughtin@chapman.edu](mailto:laughtin@chapman.edu).

---

# Why Physical Understanding Should Precede the Mathematical Formalism—Conditional Quantum Probabilities as a Case-Study

## Comments

This article may be downloaded for personal use only. Any other use requires prior permission of the author and AIP Publishing. This article appeared in:

Yakir Aharonov, Eliahu Cohen, and David H. Oaknin, "Why physical understanding should precede the mathematical formalism—Conditional quantum probabilities as a case-study," *Am. J. Phys.* **87**, 668–673 (2019).

and may be found at DOI: [10.1119/1.5115980](https://doi.org/10.1119/1.5115980).

This article was originally published in *American Journal of Physics*, volume 87, in 2019.

## Copyright

American Association of Physics Teachers



# Why physical understanding should precede the mathematical formalism—Conditional quantum probabilities as a case-study

Yakir Aharonov<sup>a)</sup>

*Institute for Quantum Studies, Chapman University, Orange, California 92866; Schmid College of Science and Technology, Chapman University, Orange, California 92866; and School of Physics and Astronomy, Tel Aviv University, Tel Aviv 6997801, Israel*

Eliahu Cohen<sup>b)</sup>

*Faculty of Engineering and the Institute of Nanotechnology and Advanced Materials, Bar Ilan University, Ramat Gan 5290002, Israel*

David H. Oaknin<sup>c)</sup>

*Rafael Advanced Defense Systems Ltd., Haifa 61532, Israel*

(Received 23 February 2019; accepted 4 June 2019)

Conditional probabilities in quantum systems which have both initial and final boundary conditions are commonly evaluated using the Aharonov–Bergmann–Lebowitz rule. In this short note, we present a seemingly disturbing paradox that appears when applying the rule to systems with slightly broken degeneracies. In these cases, we encounter a singular limit—the probability “jumps” when going from perfect degeneracy to negligibly broken one. We trace the origin of the paradox and solve it from both traditional and modern perspectives in order to highlight the physics behind it: the necessity to take into account the finite resolution of the measuring device. As a practical example, we study the application of the rule to the Zeeman effect. The analysis presented here may stress the general need to first consider the governing physical principles before heading to the mathematical formalism, in particular, when exploring puzzling quantum phenomena. © 2019 American Association of Physics Teachers.

<https://doi.org/10.1119/1.5115980>

## I. INTRODUCTION

It is widely accepted that any physical theory must be formulated, once its fundamental principles have been understood and established, as a sound mathematical model in which the properties of all entities involved and their relationships are precisely described. Only then can the theory and its predictions be accurately confronted with the experimental evidence.

In some cases, the confrontation of the theoretical predictions with the data collected from experiments led to the conclusion that the tested theory is only an approximation valid within a limited range of applications of a more general theory. For example, Galileo’s transformation law defined with respect to different inertial frames is only an approximation of the more general Lorentz transformation law, classical electromagnetism is only an approximation of quantum electrodynamics, and Newton’s theory of gravitation is only an approximation of Einstein’s general relativity.

In other cases, the fundamental principles of the theory and its range of applicability are well established but the predictions obtained from them involve certain approximations whose range of validity may be limited. Hence, such predictions will be valid as long as the approximations involved are valid. In such cases, it is of utmost importance not to forget the physical understanding of the assumptions and approximations involved in the mathematical formalism that led to the predictions of interest. Otherwise, a misplaced use of the formalism may lead, as we shall show, to apparent paradoxes. In such cases, we believe that one should return to the fundamental physical principles lying at the heart of the problem and solve it carefully using an appropriate mathematical formalism that fits these principles. This perspective has been implicitly adopted in our previous works,<sup>1–5</sup> but here we wish to make it very explicit and further emphasize its importance when exploring quantum phenomena, especially when relying on a purely informational perspective.<sup>6–10</sup>

We shall base this perspective on a particular example concerned with conditional probabilities in quantum mechanics. This topic has previously been a source of puzzles and debates (see, for example, Refs. 11–18), but here we point out a new, at first perplexing paradox, which gives the impression of a singular limit.<sup>19</sup> We shall examine an apparent peculiarity related to the calculation of probabilities in slightly degenerate quantum systems which have both initial and final boundary conditions. These conditional probabilities are commonly evaluated using the Aharonov–Bergmann–Lebowitz (ABL) rule,<sup>20</sup> which will now be examined from both traditional and modern perspectives to highlight the physics behind it, especially when closely separated energy levels are involved.

To this aim we will resort to the modern notion of quantum measurements described by a set of positive-operator valued measure (POVM) elements, which generalizes the more traditional notion of projective measurements. Projective measurements are described by the set of projectors  $\{\hat{P}_k\}_{k \in K}$  onto the eigenspaces of each of the eigenvalues  $\{c_k\}_{k \in K}$  of the measured quantum observable  $\hat{C}$ , so that

$$\hat{C} = \sum_{k \in K} c_k \cdot \hat{P}_k, \quad (1)$$

while

$$\sum_{k \in K} \hat{P}_k = \mathbb{I}, \quad \hat{P}_k^\dagger = \hat{P}_k, \quad \hat{P}_k \cdot \hat{P}_{k'} = \delta_{k,k'} \hat{\mathbb{I}}. \quad (2)$$

Hence, the index  $k$  labels the possible different outcomes  $c_k$  of the projective measurement. The Born rule provides the probability for each of these outcomes to occur

$$p_k = \langle \Psi | \hat{P}_k | \Psi \rangle, \quad (3)$$

when the quantum system is prepared in a state  $|\Psi\rangle$ . It is then straightforward to test that

$$\sum_{k \in K} p_k = \left\langle \Psi \left| \sum_{k \in K} \hat{P}_k \right| \Psi \right\rangle = \langle \Psi | \hat{\mathbb{I}} | \Psi \rangle = 1. \quad (4)$$

Upon measurement, the quantum state of the system gets projected onto the corresponding eigenspace

$$|\Psi\rangle \rightarrow \frac{1}{\sqrt{p_k}} \hat{P}_k |\Psi\rangle. \quad (5)$$

A generalized measurement is described by a family  $\{\hat{\mathcal{E}}_r \equiv \hat{\mathcal{F}}_r^\dagger \hat{\mathcal{F}}_r\}_{r \in R}$  of positive-valued operators on the Hilbert space of the considered quantum system such that

$$\sum_{r \in R} \hat{\mathcal{E}}_r = \hat{\mathbb{I}}, \quad (6)$$

where  $R$  is the family of possible outcomes. Each one of them occurs with probability

$$p_r = \langle \Psi | \hat{\mathcal{E}}_r | \Psi \rangle = \langle \Psi | \hat{\mathcal{F}}_r^\dagger \hat{\mathcal{F}}_r | \Psi \rangle = |\hat{\mathcal{F}}_r |\Psi\rangle|^2 \geq 0, \quad (7)$$

when the quantum system is prepared in the state  $|\Psi\rangle$ , which generalizes the standard Born rule for projective measurements. As before, it is straightforward to test that

$$\sum_{r \in R} p_r = \sum_{r \in R} \langle \Psi | \hat{\mathcal{E}}_r | \Psi \rangle = \left\langle \Psi \left| \sum_{r \in R} \hat{\mathcal{E}}_r \right| \Psi \right\rangle = \langle \Psi | \hat{\mathbb{I}} | \Psi \rangle = 1. \quad (8)$$

Upon measurement, the quantum state of the system gets projected as

$$|\Psi\rangle \rightarrow \frac{1}{\sqrt{p_r}} \hat{\mathcal{F}}_r |\Psi\rangle, \quad (9)$$

which is not necessarily an eigenstate of the measured observable. The formalism of generalized measurements is needed, as we shall show below, when the projective measurement actually involves a system larger than the considered subsystem of interest. This larger system may include, for example, the measuring device, which is indeed the situation below.

## II. A CASE-STUDY CONCERNING CONDITIONAL QUANTUM PROBABILITIES

In the standard formalism of quantum mechanics, the most complete description of a closed system at a time  $t_i$  is given by a state vector  $|\Psi(t_i)\rangle$ , which is determined based on the outcomes of a complete set of compatible measurements performed on it until the given time. If the system is then left isolated, its state is determined at any later time  $t \geq t_i$  by the Schrödinger equation

$$|\Psi(t)\rangle = U(t, t_i) |\Psi(t_i)\rangle, \quad (10)$$

where  $U(t, t_i) = e^{-i\hat{H}(t-t_i)}$  and  $\hat{H}$  is the time-independent Hamiltonian of the system.

The description (10) holds over the time interval  $t_i \leq t < t_f$  until  $t_f$  at which a new set of compatible measurements is performed on the system, whose outcomes allow the state of the system to be redefined via  $|\Phi(t_f)\rangle$ . Similarly to the above description, we could say that during the time interval  $t_i < t \leq t_f$  the quantum system is described by the state vector given by

$$|\Phi(t)\rangle = U(t_f, t)^\dagger |\Phi(t_f)\rangle. \quad (11)$$

Given these initial and final conditions at times  $t_i$  and  $t_f$  we can ask about the probability to obtain a given outcome  $c_k$  in a projective measurement of a physical observable  $\hat{C}$  at any intermediate time  $t \in (t_i, t_f)$ , where  $c_k \in \{c_{k'}\}_{k' \in K}$  is one of the eigenvalues of the measured observable. This question was explored in Refs. 11 and 20, and it was found that

$$p(\hat{C} = c_k) = \frac{|\langle \Phi(t) | \hat{P}_k | \Psi(t) \rangle|^2}{\sum_{k' \in K} |\langle \Phi(t) | \hat{P}_{k'} | \Psi(t) \rangle|^2}, \quad (12)$$

where  $\hat{P}_k$  denotes the projector onto the eigenspace of the corresponding eigenvalue  $c_k$  for every  $k \in K$ . It can be readily verified that  $p(\hat{C} = c_k) \in [0, 1]$  and  $\sum_{k' \in K} p(\hat{C} = c_{k'}) = 1$ , hence the relation (12) can be thought of as a conditional probability, given the initial and final states of the system. The relation (12) is important in that it restores time-symmetry, even when collapse is involved in the description of quantum systems.

In this note we focus on the particular case in which the considered physical observable takes the form  $\hat{C} = \hat{C}_0 + \epsilon \hat{C}_1$ , where  $\epsilon \in \mathbb{R}$  is some small real number and the operators  $\hat{C}_0$  and  $\hat{C}_1$  do commute  $[\hat{C}_0, \hat{C}_1] = 0$ . In addition, we shall assume that  $\hat{C}_0$  has a degenerate spectrum, while the perturbation  $\hat{C}_1$  breaks some or all these degeneracies. It can be readily seen that the conditional probability as defined in Eq. (12) shows a discontinuity at  $\epsilon = 0$ .

As an example, we consider a three dimensional system and denote by  $\{|\alpha\rangle, |\beta\rangle, |\gamma\rangle\}$  an orthonormal basis. We define the linear operators  $\hat{C}_0$  and  $\hat{C}_1$  through the relationships

$$\hat{C}_0 |\alpha\rangle = -|\alpha\rangle, \quad \hat{C}_0 |\beta\rangle = +|\beta\rangle, \quad \hat{C}_0 |\gamma\rangle = +|\gamma\rangle, \quad (13)$$

and

$$\hat{C}_1 |\alpha\rangle = -|\alpha\rangle, \quad \hat{C}_1 |\beta\rangle = +|\beta\rangle, \quad \hat{C}_1 |\gamma\rangle = +2|\gamma\rangle. \quad (14)$$

Hence, according to Eq. (12) we have

$$p(\hat{C}(\epsilon = 0) = -1) = \frac{|\langle \Phi(t) | \alpha \rangle \cdot \langle \alpha | \Psi(t) \rangle|^2}{|\langle \Phi(t) | \alpha \rangle \cdot \langle \alpha | \Psi(t) \rangle|^2 + |\langle \Phi(t) | \beta \rangle \cdot \langle \beta | \Psi(t) \rangle + \langle \Phi(t) | \gamma \rangle \cdot \langle \gamma | \Psi(t) \rangle|^2},$$

$$p(\hat{C}(\epsilon \neq 0) = -1 - \epsilon) = \frac{|\langle \Phi(t) | \alpha \rangle \cdot \langle \alpha | \Psi(t) \rangle|^2}{|\langle \Phi(t) | \alpha \rangle \cdot \langle \alpha | \Psi(t) \rangle|^2 + |\langle \Phi(t) | \beta \rangle \cdot \langle \beta | \Psi(t) \rangle|^2 + |\langle \Phi(t) | \gamma \rangle \cdot \langle \gamma | \Psi(t) \rangle|^2}. \quad (15)$$

In the denominator of the expression for  $p(\hat{C}(\epsilon = 0) = -1)$  appears the term  $|\langle \Phi(t)|\beta\rangle\langle\beta|\Psi(t)\rangle + \langle \Phi(t)|\gamma\rangle\langle\gamma|\Psi(t)\rangle|^2$ , since the eigenvalue  $C = 1$  is degenerate when  $\epsilon = 0$ . On the other hand, when  $\epsilon \neq 0$  the degeneracy is broken into two different non-degenerate eigenvalues:  $C = 1 + \epsilon$  and  $C = 1 + 2\epsilon$  and, therefore, in the denominator of the expression for  $p(\hat{C}(\epsilon \neq 0) = -1 - \epsilon)$  appears the term  $|\langle \Phi(t)|\beta\rangle \cdot \langle\beta|\Psi(t)\rangle|^2 + |\langle \Phi(t)|\gamma\rangle \cdot \langle\gamma|\Psi(t)\rangle|^2$ , which is not necessarily equal to the former. Indeed, it is necessarily equal or larger, so that

$$p(\hat{C}(\epsilon \neq 0) = -1 - \epsilon) \leq p(\hat{C}(\epsilon = 0) = -1). \quad (16)$$

At first sight, the discontinuity at  $\epsilon = 0$  might be surprising and even disturbing, since it would imply that the hypothetical measurement at intermediate time would be able to detect, through a finite jump in the frequencies of its possible outcomes, any  $\epsilon \neq 0$  no matter how small it could be. For example, in the simple example that we have just discussed the measurement seems to be able to detect a difference between the two largest eigenvalues as small as, say,  $\epsilon = 10^{-100}$  or even smaller, which is of course physically unreasonable.

However, as implied above, one has to take into account the physical principles involved in order to understand this apparent conundrum. It turns out that the uncertainty of the measuring pointers has to be taken into account in order to understand the aforementioned discontinuity. Namely, the outcome of a measurement corresponds to a shift in the position of a pointer, whose wavefunction always has, in practice, a non-zero width. The ABL rule as stated in Eq. (12) is valid for projective measurements for which, by definition, the uncertainty of the measuring pointers is negligible compared to the gap between the different eigenvalues of the measured observable. The finite precision of the measuring pointer must obviously be taken into account when it is larger or comparable to the gap between different eigenvalues, since then they become practically indistinguishable and the system should be thought of as being effectively degenerate.

In order to take into account the uncertainty in the pointer's state, we can describe the measurement at time  $t$  by a set  $\{\hat{\mathcal{E}}_r \equiv \hat{\mathcal{F}}_r^\dagger \hat{\mathcal{F}}_r\}_{r \in R}$  of positive-operator valued measure (POVM) elements

$$\sum_{r \in R} \hat{\mathcal{E}}_r = \hat{\mathbb{I}}, \quad (17)$$

rather than as a projective measurement. The ABL rule can then be readily generalized as follows:

$$p(\hat{\mathcal{F}} = f_r) = \frac{|\langle \Phi(t)|\hat{\mathcal{F}}_r^\dagger \hat{\mathcal{F}}_r|\Psi(t)\rangle|^2}{\sum_{r' \in R} |\langle \Phi(t)|\hat{\mathcal{F}}_{r'}^\dagger \hat{\mathcal{F}}_{r'}|\Psi(t)\rangle|^2}. \quad (18)$$

Within this generalized framework, we can smoothly connect the case of a non-degenerate system in which the precision of the pointer is much larger than the gap between its energy eigenvalues to the case in which the width of the pointer is much smaller than the gap between eigenvalues.

For example, a generalized measurement at time  $t$  could be defined as follows:

$$\begin{aligned} \mathcal{F}_\alpha &= |\alpha\rangle\langle\alpha|, & \mathcal{F}_\beta &= \frac{1}{\sqrt{1+\zeta^2}}(|\beta\rangle\langle\beta| + \zeta|\gamma\rangle\langle\gamma|), \\ \mathcal{F}_\gamma &= \frac{1}{\sqrt{1+\zeta^2}}(\zeta|\beta\rangle\langle\beta| + |\gamma\rangle\langle\gamma|), \end{aligned} \quad (19)$$

where  $\zeta \equiv e^{-(\epsilon^2/4\Delta^2)}$  and  $\Delta$  is the resolution of the measurement, i.e., the width of the Gaussian pointer. It can be readily checked that

$$\begin{aligned} \mathcal{F}_\alpha^\dagger \mathcal{F}_\alpha &= |\alpha\rangle\langle\alpha|, \\ \mathcal{F}_\beta^\dagger \mathcal{F}_\beta &= \frac{1}{1+\zeta^2}(|\beta\rangle\langle\beta| + \zeta^2|\gamma\rangle\langle\gamma|), \\ \mathcal{F}_\gamma^\dagger \mathcal{F}_\gamma &= \frac{1}{1+\zeta^2}(\zeta^2|\beta\rangle\langle\beta| + |\gamma\rangle\langle\gamma|), \end{aligned}$$

since  $\langle\beta|\gamma\rangle = 0$ . Hence,

$$\mathcal{F}_\alpha^\dagger \mathcal{F}_\alpha + \mathcal{F}_\beta^\dagger \mathcal{F}_\beta + \mathcal{F}_\gamma^\dagger \mathcal{F}_\gamma = \hat{\mathbb{I}}, \quad (20)$$

as required.

Moreover, it is also straightforward to notice that in the limit of  $\epsilon \gg \Delta$ , we have  $\zeta \simeq 0$  and therefore,

$$\mathcal{F}_\alpha \simeq |\alpha\rangle\langle\alpha|, \quad \mathcal{F}_\beta \simeq |\beta\rangle\langle\beta|, \quad \mathcal{F}_\gamma \simeq |\gamma\rangle\langle\gamma|. \quad (21)$$

This limit describes a measurement whose resolution is able to distinguish between the eigenvalues  $C = 1 + \epsilon$  and  $C = 1 + 2\epsilon$  of the observable  $\hat{C}$  and, indeed, in this limit equation (18) reproduces the ABL rule for the non-degenerate case  $\epsilon \neq 0$ .

On the other hand, in the limit of  $\epsilon \ll \Delta$ , we have  $\zeta \simeq 1$  and, therefore,

$$\mathcal{F}_\alpha \simeq |\alpha\rangle\langle\alpha|, \quad \mathcal{F}_\beta \simeq \mathcal{F}_\gamma \simeq \frac{1}{\sqrt{2}}(|\beta\rangle\langle\beta| + |\gamma\rangle\langle\gamma|). \quad (22)$$

This limit describes a measurement whose resolution is not able to distinguish between the eigenvalues  $C = 1 + \epsilon$  and  $C = 1 + 2\epsilon$  associated to the eigenstates  $|\beta\rangle$  and  $|\gamma\rangle$ , respectively. In this limit, Eq. (18) reproduces the ABL rule for the degenerate case  $\epsilon = 0$ .

Equation (18), which trivially generalizes the original ABL formula for projective measurements (12), can be obtained as follows:

In the standard von-Neumann description of a projective measurement, the probed system and the measuring device (pointer) are prepared in a separable state and then interact for a short time interval, after which the state of the coupled system is described by the unitary transformation

$$\begin{aligned} e^{i\theta\hat{C} \otimes \hat{P}} &= e^{i\theta \sum_k c_k |k\rangle\langle k| \otimes \hat{P}} = \prod_k e^{i\theta c_k |k\rangle\langle k| \otimes \hat{P}} \\ &= \prod_k (\hat{\mathbb{I}} + |k\rangle\langle k| \otimes (e^{i\theta c_k \hat{P}} - \hat{\mathbb{I}})) \\ &= \hat{\mathbb{I}} + \sum_k |k\rangle\langle k| \otimes (e^{i\theta c_k \hat{P}} - \hat{\mathbb{I}}) \\ &= \sum_k |k\rangle\langle k| \otimes e^{i\theta c_k \hat{P}} = \sum_{k \in K} \hat{P}_k \otimes e^{i\theta c_k \hat{P}}, \end{aligned}$$

where, as before  $\hat{C}$  is the tested observable,  $|k\rangle$  and  $c_k$  are its eigenvectors and corresponding eigenvalues,  $\hat{P}$  is the

momentum operator of the pointer, and  $\theta$  describes the strength of the interaction. If before the interaction the pointer is prepared in an eigenstate  $|0\rangle$  of its position operator, the transformation  $e^{i\theta c_k \hat{P}}|0\rangle = |x_k\rangle$  describes a shift in the position of the pointer by an amount  $x_k = \theta c_k$ . Hence,

$$e^{i\theta \hat{C} \otimes \hat{P}}|0\rangle = \sum_{k \in K} \hat{P}_k \otimes e^{i\theta c_k \hat{P}}|0\rangle = \sum_{k \in K} \hat{P}_k \otimes |x_k\rangle. \quad (23)$$

Therefore, a shift  $x_k$  in the position of the pointer at the end of the measurement would imply that the state of the measured system has been projected by the projector  $\hat{P}_k$  onto the eigenspace of the corresponding eigenvalue  $c_k$ . This situation would correspond to the original ABL rule (12). Let us note, for the sake of clarity in the discussion that follows, that for any projector  $\hat{P}_k \hat{P}_k = \hat{P}_k^2 = \hat{P}_k$ .

Similarly, for any other orthonormal basis of states of the pointer  $\{|y_r\rangle\}_{r \in R}$ , such that its position eigenstates can be written as  $|x_k\rangle = \sum_{r \in R} \rho_{k,r} |y_r\rangle$  with  $\sum_{r \in R} |\rho_{k,r}|^2 = 1$ , we have

$$\begin{aligned} e^{i\theta \hat{C} \otimes \hat{P}}|0\rangle &= \sum_{k \in K} \hat{P}_k \otimes |x_k\rangle = \sum_{k \in K} \hat{P}_k \otimes \left( \sum_{r \in R} \rho_{k,r} |y_r\rangle \right) \\ &= \sum_{r \in R} \left( \sum_{k \in K} \rho_{k,r} \hat{P}_k \right) \otimes |y_r\rangle. \end{aligned} \quad (24)$$

In this case, an outcome  $y_r$  at the end of the measurement would imply that the state of the measured system has been transformed by the operator  $\mathcal{F}_r = \sum_{k \in K} \rho_{k,r} \hat{P}_k$ , which would lead to the generalized ABL rule (18). It is straightforward to test that

$$\begin{aligned} \sum_{r \in R} \mathcal{F}_r^\dagger \mathcal{F}_r &= \sum_{k \in K} \sum_{k' \in K} \sum_{r \in R} \rho_{k',r}^* \rho_{k,r} \hat{P}_{k'} \hat{P}_k \\ &= \sum_{k \in K} \sum_{r \in R} |\rho_{k,r}|^2 \hat{P}_k = \sum_{k \in K} \hat{P}_k = \hat{1}. \end{aligned} \quad (25)$$

### III. AN EXAMPLE: THE ZEEMAN EFFECT IN THE HYDROGEN ATOM

As a simple and deductive example, we consider the electronic degrees of freedom in a hydrogen atom. For the sake of simplicity, we ignore the electron's spin and neglect relativistic effects. In such approximation, the energy eigenstates of the electronic Hamiltonian  $\hat{C}$  (although the Hamiltonian is usually denoted by  $\hat{H}$  we keep using the above notation for the measured observable for pedagogic reasons) can be labelled by three quantum numbers as  $|n, l, m\rangle$ , where  $n = 1, 2, 3, \dots$  is the principal quantum number,  $l = 0, \dots, n-1$  labels the orbital angular momentum of the electron and  $m = -l, \dots, l$  is the projection of the orbital angular momentum of the electron along the  $Z$  axis.

In the absence of external magnetic fields, the corresponding eigenvalues of the electronic Hamiltonian  $\hat{C}_0$  are defined by

$$\hat{C}_0 |n, l, m\rangle = c_n |n, l, m\rangle, \quad (26)$$

with

$$c_n = -\frac{1}{n^2} \left( \frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m_e}{2\hbar^2} = -\frac{13.6 \text{ eV}}{n^2}, \quad (27)$$

where  $e$  is the elementary electric charge,  $m_e$  is the reduced mass of the electron,  $\epsilon_0$  is the permittivity of the vacuum, and  $\hbar$  is the reduced Planck constant. Thus, the eigenspace associated with the eigenvalue  $c_n$  has linear dimension  $\sum_{l=0}^{n-1} \sum_{m=-l}^l 1 = \sum_{l=0}^{n-1} (2l+1) = n^2$ , which is the degeneracy of this energy level.

Hence, according to the ABL rule, if the electron of the hydrogen atom is pre-selected in the state  $|\Psi(t)\rangle$  and post-selected in the state  $|\Phi(t)\rangle$ , the probability to obtain the outcome  $c_n$  in a hypothetical intermediate projective measurement at time  $t$  is given by

$$\begin{aligned} p(\hat{C}_0 = c_n) &= \frac{\left| \sum_{l=0}^{n-1} \sum_{m=-l}^l \langle \Phi(t) | n, l, m \rangle \langle n, l, m | \Psi(t) \rangle \right|^2}{\sum_{n'=1}^{\infty} \left| \sum_{l'=0}^{n'-1} \sum_{m'=-l'}^{l'} \langle \Phi(t) | n', l', m' \rangle \langle n', l', m' | \Psi(t) \rangle \right|^2}, \end{aligned} \quad (28)$$

which we can rewrite as

$$\begin{aligned} p(\hat{C}_0 = c_n) &= \frac{\left| \sum_{m=-(n-1)}^{n-1} \sum_{l=|m|}^{n-1} \langle \Phi(t) | n, l, m \rangle \langle n, l, m | \Psi(t) \rangle \right|^2}{\sum_{n'=1}^{\infty} \left| \sum_{m'=-n'+1}^{n'-1} \sum_{l'=|m'|}^{n'-1} \langle \Phi(t) | n', l', m' \rangle \langle n', l', m' | \Psi(t) \rangle \right|^2}. \end{aligned} \quad (29)$$

On the other hand, in the presence of a weak external magnetic field  $\epsilon B$  along the  $Z$  axis the electronic Hamiltonian gets the form  $\hat{C} = \hat{C}_0 - \epsilon(e/2m_e)B\hat{L}_z$ , which partially breaks the degeneracy of the electronic energy levels

$$\hat{C} |n, l, m\rangle = c_{n,m} |n, l, m\rangle = \left( c_n - \epsilon \frac{e}{2m_e} B m \right) |n, l, m\rangle. \quad (30)$$

Accordingly, the probability to obtain the outcome  $c_{n,m}$  in a hypothetical intermediate projective measurement at time  $t$  is given by

$$\begin{aligned} p(\hat{C} = c_{n,m}) &= \frac{\left| \sum_{l=|m|}^{n-1} \langle \Phi(t) | n, l, m \rangle \langle n, l, m | \Psi(t) \rangle \right|^2}{\sum_{n'=1}^{\infty} \sum_{m'=-n'+1}^{n'-1} \left| \sum_{l'=|m'|}^{n'-1} \langle \Phi(t) | n', l', m' \rangle \langle n', l', m' | \Psi(t) \rangle \right|^2}, \end{aligned} \quad (31)$$

which we can write as

$$p(\hat{C} = c_{n,m}) = \frac{|\langle \Phi(t) | \hat{P}_{n,m} | \Psi(t) \rangle|^2}{\sum_{n'=1}^{\infty} \sum_{m'=-n'+1}^{n'-1} |\langle \Phi(t) | \hat{P}_{n',m'} | \Psi(t) \rangle|^2}, \quad (32)$$

where

$$\hat{P}_{n,m} \equiv \sum_{l=|m|}^{n-1} |n, l, m\rangle \langle n, l, m| \quad (33)$$

is the projector onto the eigenspace associated with the energy eigenvalue  $c_{n,m}$ .

We can now define a generalized measurement  $\{\mathcal{F}_{n,m}\}_{n \in \mathbb{N}, |m| < n}$  as follows:

$$\mathcal{F}_{n,m} = \sum_{n',m'} \frac{\mathcal{Z}_{n',m',n,m}}{\mathcal{Z}_{n',m'}} \hat{P}_{n',m'}, \quad (34)$$

where

$$\sum_{n',m'} \equiv \sum_{n'=1}^{\infty} \sum_{m'=-n'+1}^{n'-1}, \quad (35)$$

$$\mathcal{Z}_{n',m',n,m} \equiv e^{-(c_{n',m'} - c_{n,m})^2 / 2\Delta^2}, \quad (36)$$

and

$$\mathcal{Z}_{n',m'} \equiv \sqrt{\sum_{n,m} (\mathcal{Z}_{n',m',n,m})^2}. \quad (37)$$

Hence, taking into account that

$$P_{n,m} \cdot P_{n',m'} = \delta_{n,n'} \delta_{m,m'} P_{n,m}, \quad (38)$$

we find that

$$\mathcal{F}_{n,m}^\dagger \mathcal{F}_{n,m} = \sum_{n',m'} \frac{(\mathcal{Z}_{n',m',n,m})^2}{(\mathcal{Z}_{n',m'})^2} \hat{P}_{n',m'}, \quad (39)$$

and

$$\sum_{n,m} \mathcal{F}_{n,m}^\dagger \mathcal{F}_{n,m} = \sum_{n',m'} \hat{P}_{n',m'} = \hat{\mathbb{I}}, \quad (40)$$

as required. Moreover, it can be readily seen from Eq. (34) that

$$\mathcal{F}_{n,m} \simeq \frac{1}{\sqrt{\mathcal{X}}} \left( \sum_{(n',m') : |c_{n',m'} - c_{n,m}| \leq \Delta} \hat{P}_{n',m'} \right), \quad (41)$$

where  $\mathcal{X} \sim \text{Card}\{(n', m') : |c_{n',m'} - c_{n,m}| \leq \Delta\}$  counts the number of energy levels whose distance to the level  $c_{n,m}$  is smaller than the resolution of the measurement  $\Delta$  and, therefore, cannot be actually distinguished from it, so that they should be taken as degenerate.

Only when  $\Delta \rightarrow 0^+$  is much smaller than the energy gap between the level  $c_{n,m}$  and all the others, this energy level can be completely resolved and we have

$$\mathcal{F}_{n,m} \simeq \hat{P}_{n,m}. \quad (42)$$

#### IV. DISCUSSION

Probabilities in quantum systems having both initial and final conditions are customarily calculated using the ABL formula,<sup>20</sup> which describes the probabilities to obtain a certain eigenvalue of the measured operator at any intermediate time.

When this rule is blindly applied to systems with negligibly broken degeneracies, it appears to lead to a paradox—no matter how small the difference between the almost degenerated eigenvalues is, it could be detected through a finite jump in the probabilities of each of the possible outcomes (with respect to the limiting case in which the degeneracy is exactly restored).

The paradox is solved upon completing the physical scenario and taking into account the finite resolution of the measuring device, either by applying the ABL rule to the measured system and the measuring device as a whole or by describing the intermediate measurement performed on the system as a set of positive-operator valued measure elements.

It can then be explicitly shown that eigenvalues of the measured observable whose difference is much smaller than the resolution of the intermediate measurement must be treated as effectively degenerate if we wish to use the ABL rule for the measured system alone. In contrast, eigenvalues whose difference is larger than the resolution of the measurement device must be treated as non-degenerate. As a practical example, we presented a detailed analysis of conditional probabilities in a hydrogen atom where the rotational symmetry is partially broken by a weak external magnetic field through the Zeeman effect.

This analysis may support a general viewpoint according to which when addressing a physical problem one should first understand the governing physical principles, as well as the involved approximations and assumptions, before heading to the mathematical formalism. Otherwise, a misguided use of the formalism could lead to apparent paradoxes. In such cases, one should return to the fundamental physical principles in order to trace the origin of the paradox and then solve it by employing a mathematical model which better accords with these principles. All this, of course, does not diminish the major importance of mathematical language in describing physical problems, but only calls for a careful use thereof which is aligned with the physics.

#### ACKNOWLEDGMENT

This work has been supported in part by the Israel Science Foundation Grant No. 1311/14.

<sup>a</sup>Electronic mail: yakir@post.tau.ac.il

<sup>b</sup>Electronic mail: eliahu.cohen@biu.ac.il

<sup>c</sup>Electronic mail: d1306av@gmail.com

<sup>1</sup>Yakir Aharonov and Daniel Rohrlich, *Quantum Paradoxes: Quantum Theory for the Perplexed* (Wiley-VCH, Weinheim, 2008).

<sup>2</sup>Yakir Aharonov, Alonso Botero, Sandu Popescu, Benni Reznik, and Jeff Tollaksen, "Revisiting Hardy's paradox: Counterfactual statements, real measurements, entanglement and weak values," *Phys. Lett. A* **301**(3–4), 130–138 (2002).

<sup>3</sup>Yakir Aharonov, Eliahu Cohen, Ariel Landau, and Avshalom C. Elitzur, "The case of the disappearing (and re-appearing) particle," *Sci. Rep.* **7**(531), 1–9 (2017).

<sup>4</sup>Yakir Aharonov, Eliahu Cohen, Avshalom C. Elitzur, and Lee Smolin, "Interaction-free effects between distant atoms," *Found. Phys.* **48**(1), 1–16 (2018).

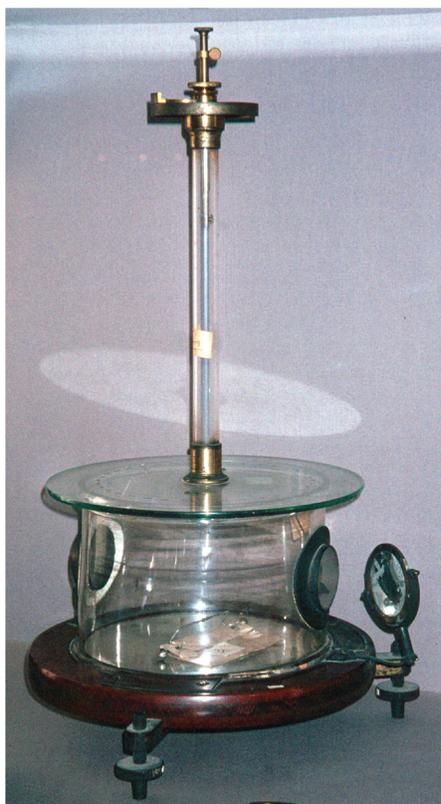
<sup>5</sup>Yakir Aharonov, Eliahu Cohen, Avishy Carmi, and Avshalom C. Elitzur, "Extraordinary interactions between light and matter determined by anomalous weak values," *Proc. R. Soc. A* **474**(2215), 20180030 (2018).

<sup>6</sup>Christopher A. Fuchs, "Quantum mechanics as quantum information (and only a little more)," e-print [arXiv:quant-ph/0205039](https://arxiv.org/abs/quant-ph/0205039).

<sup>7</sup>Carlton M. Caves, Christopher A. Fuchs, and Ruediger Schack, "Quantum probabilities as Bayesian probabilities," *Phys. Rev. A* **65**, 022305 (2002).

<sup>8</sup>Rob Clifton, Jeffrey Bub, and Hans Halvorson, "Characterizing quantum theory in terms of information-theoretic constraints," *Found. Phys.* **33**(11), 1561–1591 (2003).

- <sup>9</sup>Jeffrey Bub, “Quantum mechanics is about quantum information,” *Found. Phys.* **35**(4), 541–560 (2005).
- <sup>10</sup>Christopher A. Fuchs, N. David Mermin, and Ruediger Schack, “An introduction to QBism with an application to the locality of quantum mechanics,” *Am. J. Phys.* **82**(8), 749–754 (2014).
- <sup>11</sup>Yakir Aharonov and Lev Vaidman, “Complete description of a quantum system at a given time,” *J. Phys. A: Math. Gen.* **24**(10), 2315–2328 (1991).
- <sup>12</sup>Ruth E. Kastner, “The three-box ‘paradox’ and other reasons to reject the counterfactual usage of the ABL rule,” *Found. Phys.* **29**(6), 851–863 (1999).
- <sup>13</sup>Ulrich Mohrhoff, “What quantum mechanics is trying to tell us,” *Am. J. Phys.* **68**(8), 728–745 (2000).
- <sup>14</sup>Ruth E. Kastner, “Comment on ‘What quantum mechanics is trying to tell us,’ by Ulrich Mohrhoff, *Am. J. Phys.* 68(8), 728–745 (2000)],” *Am. J. Phys.* **69**(8), 860–863 (2001).
- <sup>15</sup>Ulrich Mohrhoff, “Objective probabilities, quantum counterfactuals, and the ABL rule—A response to R.E. Kastner,” *Am. J. Phys.* **69**(8), 864–873 (2001).
- <sup>16</sup>Matthew S. Leifer and Robert W. Spekkens, “Pre- and post-selection paradoxes and contextuality in quantum mechanics,” *Phys. Rev. Lett.* **95**, 200405 (2005).
- <sup>17</sup>Dmitri I. Sokolovski, Irene P. Gimenez, and R. Sala Mayato, “Path integrals, the ABL rule and the three-box paradox,” *Phys. Lett. A* **372**(44), 6578–6583 (2008).
- <sup>18</sup>Eliahu Cohen and Marcin Nowakowski, “Comment on measurements without probabilities in the final state proposal,” *Phys. Rev. D* **97**, 088501 (2018).
- <sup>19</sup>Michael V. Berry, “Singular limits,” *Phys. Today* **55**(5), 10–11 (2002).
- <sup>20</sup>Yakir Aharonov, Peter G. Bergmann, and Joel L. Lebowitz, “Time symmetry in the quantum process of measurement,” *Phys. Rev.* **134**, B1410–B1416 (1964).



### Coulomb Balance

A surprising number of these devices remain in museum and university collections. Since they are tricky to use, this is perhaps because they were rarely taken out for use! It was devised by Charles Coulomb (1736-1806) who made early measurements leading to what we now call Coulomb’s Law. One charge was placed on a small conducting sphere at one end of a T-bar hung at the end of the suspension, and the other one, of the same sign, was placed on a stationary sphere. The torque constant of the suspension was obtained from measurements of the period of the oscillating system. This example, by Elliott of London, is in the apparatus and ecclesiastical museum at St. Patrick’s College in Maynooth, Ireland. (Picture and text by Thomas B. Greenslade, Jr., Kenyon College)