Scientific Research

Explaining the Double-Slit Experiment

Moshe Wallace Callen, **Shaul Sorek**[1,2]

[1]*Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Be'er Sheva, Israel*
[2]*Department of Mechanical Engineering, Ben-Gurion University of the Negev, Be'er Sheva, Israel*
E-mail: *moshe.callen@gmail.com, sorek@bgu.ac.il*

Abstract

In response to Orion and Laitman's [1] explanation of the classic double-slit experiment of quantum mechanics, we propose an alternate explanation of that experiment by treating physical degrees of freedom as a conserved physical quantity, instead of referring to "vague terms" used in previous explanations, [1], that are not broadly applicable. Explanation in [1] refers to properties of groups of particles, even though the double-slit experiment's results should address only to a single particle. By using physical degrees of freedom and the application of Hamilton's principle, we obtain a single particle explanation of the double-slit experiment in terms of properties and via methods which apply equally in a quantum and a classical regime.

1. Introduction

The famous double-slit experiment involves either the observation or non-observation of an interference pattern in two physical situations which differ only by whether a certain measurement is not or is taken respectively [2,3]. An electron source projects electrons onto a screen with a double slit, through which an electron is then projected onto a second screen. When no measurement is taken to determine through which slit the electron passes, an interference pattern is observed on the second screen. When such a measurement is taken however, no interference pattern is observed. Various explanations have been proposed, but many of these explanations still regard this result as mysterious to varying degrees [4]. Perhaps the most recent such explanation is that proposed by Orion and Laitman [1] to which the current discussion serves as a specific response. The discussion in [1] proposes a *Kevutsa* or group interpretation of the double-slit experiment intended as an improvement on previous explanations which are claimed to have used "unclear terms".

Further, in [1] the interpretation depends on two proposed principles, an "equivalence of form" and "the particles connection to other particles, effectively functioning as a group." The two most noticeable problems with the proposed explanation provided in [1] are as follows:

1) A group-based and hence effectively multi-particle explanation is difficult to justify for a physical phe-nomenon which can be observed in a single particle situation as can the self-interference characteristic of the double slit experiment [2,3].

2) The proposed explanation simply replaces what are considered as ill defined terms with new terms namely a principle of form and the notion of *Kevutsa*.

For the sake of argument, one accepts the point of view that reference to a wave particle duality [2,3] is not especially useful in understanding the results of the double-slit experiment. After all, such a duality merely gives a name to the underlying observation that electrons and similar quantum particles have properties both of a wave and of a particle and that those properties will manifest themselves in different types of interactions; that is the definition of a wave particle duality. If then the wave particle duality is not useful in and of itself for understanding the results of the double-slit experiment, at least by assumption (again for the sake of argument), one needs a system by which to explain the results in terms which are well defined and without needing to worry about how an electron can be a wave, a particle or simultaneously both.

The physical degrees of freedom, namely the characteristic minimum number of variables needed to fully describe the physical situation, provide such an explanation. If one sends a single electron in a known physical state toward the screen with the two slits, by definition of a state that electron has no physical degrees of freedom before it reaches the slit-screen, In other words, the state

of the electron is a known quantity initially. Then it encounters the slit-screen. If no measurement is taken to determine which slit the electron went through, a degree of freedom is introduced at that point. If such a measurement is taken, no such degree of freedom is introduced, because the measurement removes the ambiguity of which slit the electron could have passed through. When the electron reaches the final screen (*i.e.*, the projection screen), then if a degree of freedom was introduced, some physical manifestation of that degree of freedom must occur (hence one sees an interference pattern), but if no such degree of freedom was introduced the electron remains in a known state and so no interference pattern is seen. Admittedly, the two physical events, namely the electron at the slit-screen and the electron at the projection screen, are not occurring at exactly the same moment in time, but a nearly trivial application of Hamilton's principle [5] surmounts that difficulty easily. Nothing anthropocentric is needed in this explanation, and all terms are clear. Moreover this same term, *i.e.*, physical degrees of freedom, is applicable to any arbitrary physical situation, not just the double-slit experiment and similar self-interference phenomena.

The resulting understanding of the double-slit experiment can then be summarized as follows. Quantum mechanics postulates that any particle such as an electron can also be described as a wave characterized by its de Broglie wavelength [1-3]. If so, then one should be able to create an interference pattern from an electron using a screen with a double-slit in it. That is in fact observed. The initially surprising aspect is the effect of a measurement taken at the slit-screen: namely, that such a measurement causes the interference pattern to disappear. That phenomenon can be understood by treating the number of physical degrees of freedom as a conserved quantity in the sense that the number of independent variables needed to describe any given physical situation is characteristic of that physical situation just like energy or momentum would be in collisions [5,6].

2. Re-Examination of the Double-Slit Experiment

The famous double-slit experiment [2,3] traditionally exemplifies the wave-particle duality of quantum mechanical systems. The double-slit experiment has been termed the fundamental quantum mechanical mystery [4,7]. The implications for the understanding of quantum mechanical systems makes it a topic of on-going research [8-11], albeit in varying forms, to this day, and by the same token it is used as a toy model for understanding more complex systems [12,13], similar to the manner in which one often uses the simple harmonic oscillator

[2,3]. In short, the double-slit experiment continues to both perplex and intrigue [14,15], encapsulating what is and is not understood about the quantum mechanical world.

The present discussion attempts to understand the usual double slit experiment in a novel manner. Namely, in any physical system, the number of degrees of freedom remains characteristic of the system [5]. For example, when describing a classical particle moving through space x, one is able to define a set of coordinate axes x' so that the linear momentum $p(x)$ becomes a vector $p_i'(x')$ parallel to one of the major axes x_i' thus effectively eliminating the need for the two other components of the position in space. Nevertheless, the particle retains three degrees of freedom so that the energy E and total linear momentum P become effectively parameters of the system rather than derivable quantities, in order that along with uniaxial position the system retains three degrees of freedom [5,6]. One has in this example transformed from generalized coordinates (x_1, x_2, x_3) to generalized coordinates (x', E, P). Conventional usage does not term the number of degrees of freedom a conserved quantity because it is not a directly measurable quantity itself, but in a sense the number of degrees of freedom is indeed a conserved quantity, albeit at a level of abstraction removed in that one does not have a degrees of freedom-meter with which to directly measure the total degrees of freedom. One may think of the number of degrees of freedom as a constraint upon the physical system, because this number is always constant. In the example already cited of the classical particle, the number of total degrees of freedom acts as a constraint in that it sets the number of input parameters from which other quantities may be derived. Similarly, in a quantum mechanical system, one may choose among alternate sets of what are termed "good" quantum numbers $\{n, l, m, s\}$, $\{n, j, j_z, s\}$, etc. (meaning quantum numbers which refer to simultaneously measurable quantities) [2,3], but the total number of quantum numbers needed to describe a system remains the same.

More formally, any local coordinate transformation of generalized coordinates q_h (defined as variables on which a specific Lagrangian depends) to another local mapping of generalized coordinates q_k' (defined as an alternate set of variables in terms of which that same Lagrangian can also be represented) takes the form [5,6].

$$q_k' = \sum_h \frac{\partial q_k'}{\partial q_h} q_h \qquad (1)$$

Yet at the same time one's choice of coordinates q_h or coordinates q_k' remains arbitrary, and so similarly one has

$$q_h = \sum_k \frac{\partial q_h}{\partial q_k'} q_k' \qquad (2)$$

Both (1) or (2) can be written as matrix equations, and so matrix $\partial q_k'/\partial q_h$ must be the inverse matrix of matrix $\partial q_h/\partial q_k'$. The existence of this pair of inverse matrices is only possible if the range of index h is identically the same as the range of index k. For unitary transformations, this common range becomes the trace $Tr(\partial q_k'/\partial q_h) = Tr(\partial q_h/\partial q_k')$ the value is terms the number of degrees of freedom. In the arbitrary case, the total number of degrees of freedom f tells one how many independent and simultaneously measurable variables q_h or q_k' are needed to describe the given physical situation. Algebraically, one then needs a similar number of equations to solve the system generally.

A classical example [5,6] is the orbital motion of a planet about the sun. To specify the location of the planet within the plane of its orbit at any time, one may use, respectively, a radial distance and an angle (r, β) or one may use Cartesian planar coordinates (x, y). One may also use energy and momentum (E, P), but one always will need two variables in this example.

Where the application becomes more interesting is in those cases where the Lagrangian describes both initial and final states. If one considers as an example a bucket of sand with a hole in the bottom, the physical situation takes the same number of variables to describe both just before the first sand begins to fall out of the bucket and throughout the process. In this example, using energy (E) as a single variable is simplest, but in principle one could for example also use the center of mass of the sand.

Yet, the total number of degrees of freedom has more profound implications than as a sort of bookkeeping device used to make sure one has the right number of equations. The notion that the total number of degrees of freedom remains constant and characteristic of any physical systems constrains that physical system to express all the degrees of freedom [5,6]. For example, a particle in its proper frame of reference defines the origin of the coordinate system and thus has a fixed position. Nevertheless, if in some other reference frame the motion of the particle is described by four independent equations, its motion must also be characterized by four equations in its proper frame of reference [5,6,16]. This is the aspect relevant to the double-slit experiment. Attempts to understand the connection between measurement and observation have spawned bizarre physical models such as the "many worlds" interpretation of quantum mechanics [4] and seemingly endless philosophical debate [7]. Yet, treatment of the total degrees of freedom as a conserved physical quantity, where meas-

urement is viewed as extraction of a particular degree of freedom, *i.e.*, removal of that degree of freedom from the system being considered, allows a physical interpretation of the results of the double-slit experiment which remains thoroughly physical and does not venture into speculative areas of discussion. This non-speculative interpretation of the double-slit experiment and of similar quantum mechanical phenomena which involve "self-interference" [2,3] is the viewpoint on the double-slit experiment expounded in this discussion.

The classic double-slit experiment consists of a source emitting single electrons (or photons), a screen with two narrow slits (to be termed the "slit-screen") and a screen onto which each electron is projected (to be termed the "projection-screen"). The three objects are placed so that each electron must pass through the slit-screen in order to be projected onto the projection-screen. Electrons may only pass through the two narrow slits in the slit-screen, but the width of the electrons (as defined by their de Broglie wavelength [1-3]) is still negligible compared to that of the slits, at least in the classic experiment [14]. The relative positions of the two screens determines the specific geometry of the interference pattern observed, as per the usual interference law [2]

$$d \sin \theta = m\lambda \qquad (3)$$

(in terms of separation d between the two slits, angle θ measured with respect to the central axis between the two screens, de Broglie wavelength λ and integer m which counts peaks of intensity from the center of the projection screen) but this level of detail can be ignored for the purposes of the present discussion. The key points of the present discussion are that when a single electron leaves the source in some known state, two possible results may be seen on the projection-screen, depending on whether or not a measurement is taken at the slit-screen (the nature of measurement as a physical interaction has been thoroughly discussed, [4,7], elsewhere in the literature). When no measurement is taken at the slit-screen, an interference pattern is observed on the projection-screen. Yet, when a measurement of which slit the electron passes through is taken at the slit-screen, no interference pattern is observed on the projection-screen. These relatively qualitative results are all that are needed to frame the present discussion; they imply an effectively static, *i.e.*, time-independent, one dimensional physical system. At any two positions x_1 and x_2 along the path of the electron, the number of degrees of freedom f must be the same so that degrees of freedom in the two positions x_1 and x_2 are identical as $f(x_1) = f(x_2)$. Here one is only demanding continuity.

Implicitly one has used Hamilton's principle [5] to construct an equation of generalized motion of a form

$$\frac{\partial L[f(x)]}{\partial f} - \frac{d}{dx}\frac{\partial L[f(x)]}{\partial (df/dx)} = 0 \qquad (4)$$

which uses a generalized Lagrangian L (linear in degrees of freedom f) and in which one in principle allows the number of degrees of freedom f to vary with position. Conservation of the number degrees of freedom f simplifies this equation of generalized motion to

$$\frac{\partial L[f(x)]}{\partial f} = 0 \qquad (5)$$

The slit-screen and projection-screen are not immediately adjacent, but no interaction takes place between them. So, one can equivalently simply use

$$df = 0 \qquad (6)$$

i.e., the number of degrees of freedom f for an isolated physical system remains invariant, regardless of the position or time at which that system is observed.

In terms of degrees of freedom as a conserved quantity, the results of the double-slit experiment mentioned above become understandable. The electron leaves the electron source in some known state so that one may define its initial number of degrees of freedom f_γ as zero $f_\gamma = 0$; its final number of degrees of freedom f_γ' will be treated as the variable to be determined. The slit-screen introduces in principle one degree of freedom into the system, *i.e.*, through which slit the electron passes, so that one may speak of the slit-screen as having one degree of freedom $f_{ss} = 1$ when no measurement is taken. The electron and slit-screen together constitute the physical system initially, and the electron and the projection-screen constitute the final physical system.

The projection-screen introduces no degrees of freedom since the relative geometry here does not matter, and so one defines the projection-screen's degrees of freedom as $f_{ps} = 1$. Conservation of the number of degrees of freedom at the position of the slit-screen then implies

$$\begin{aligned} f_\gamma + f_{ss} &= f_\gamma' + f_{ps} \\ 0 + 1 &= f_\gamma' + 0 \end{aligned} \qquad (7)$$

so that the electron has a final number of degrees of freedom $f_\gamma' = 1$. That degree of freedom must be expressed somehow. Energy-momentum conservation precludes variations in frequency ω or wavelength λ [3] because these are respectively proportional directly to energy and inversely to momentum by Planck's normalized constant $\hbar$,

$$E = \hbar\omega \qquad (8)$$

$$P = \hbar k \equiv \hbar\frac{2\pi}{\lambda} \qquad (9)$$

Only variation in intensity $I(x)$ or some equivalent variable is physically permitted because the state of the electron takes a general wave-form

$$\psi(x,t) = I(x)\exp\left[-i(\omega t - kx + \phi)\right], \qquad (10)$$

and the phase ϕ is purely arbitrary so that it can have no direct physical manifestation. An interference pattern, by definition, is a variation in intensity $I(x)$ of the electron on the projection screen. Similarly, when a measurement is taken at the slit-screen of through which slit the electron passes, the slit-screen does not introduce any actual degree of freedom. One may think of this alternately as taking a degree of freedom out of the system, analogous to the manner in which an unobserved neutrino carries energy out of the system, although total energy is conserved [17]. Here, conservation of the number of degrees of freedom at the position of the slit-screen leads to $f_\gamma + f_{ss} = f_\gamma' + f_{ps}$, namely,

$$0 + 0 = f_\gamma' + 0 \qquad (11)$$

so that the final number of degrees of freedom for the electron remains at zero, namely $f_\gamma' = 0$. The intensity cannot vary because this would require another degree of freedom which the physical system does not have. Thus, no interference pattern is observed. The double-slit experiment, the quintessentially quantum mechanical experiment, can be understood in terms of the number of degrees of freedom associated with the physical system when this is viewed as a conserved quantity.

Conceptually one can separate understanding of the double-slit experiment into two parts. First, this experiment tests whether or not a particle like an electron can be treated as a wave with a characteristic de Broglie wavelength [1-3]. This discussion takes for granted that the result of this part of the experiment is positive because the interference pattern is observed and as discussed elsewhere [2,3] that interference pattern is exactly what one should expect from a wave characterized by the electron's de Broglie wavelength. The second part of understanding the double-slit experiment is what this discussion addresses, namely why that interference pattern vanishes when a measurement is taken of which slit the electron passes through to get to the projection-screen. All that is needed to understand this phenomenon is the notion of physical degrees of freedom as characteristic of any physical system, a notion which applies both in quantum and in purely classical situations.

3. Conclusions

The proposed explanation [1] of the classical double-slit experiment [2,3], while perfectly valid in other ways,

     JMP

does not achieve the goal set forth. Namely, [1] simply replaces the mystery of a wave-particle duality with the mystery of why one needs to treat a single particle only as a member of a group of particles in order to understand a phenomenon that occurs when only the one particle is involved. Where other particles enter into the physical situation is not sufficiently explained. To address the latter point, we propose an explanation of the results of the double-slit experiment which requires no "vague terms" and does not invoke or replace a "quantum mystery" as an explanation [7]. Namely, the number of physical degrees of freedom $f \equiv f(x)$ at a location x is treated as a conserved quantity, i.e., as a quantity characteristic of the physical system at location x. By definition, so long as one has a closed system, that quantity will not change.

To summarize the experiment, an interference pattern occurs when a single electron is projected onto a screen, termed the projection-screen, if the electron is made to pass through a screen with a double-slit in it, termed the slit-screen, so long as no measurement is taken of which of the two slits in the slit-screen that electron passes through.

One has taken for granted that the electron can be treated as a wave characterized by its de Broglie wavelength [2,3]. The issue under discussion is simply why an interference pattern occurs if no measurement is taken at the slit-screen to determine which slit the electron passes through but no interference pattern is observed if such a measurement is taken.

The physical system first consists of the electron and the slit-screen and then consists of the electron and the projection-screen. These physical events are not immediately adjacent of course, but since no interaction occurs between them, a nearly trivial application of Hamilton's principle [5] relates the two events directly. One assumes the electron leaves its source in a known state because any extraneous degrees of freedom the electron might have when leaving the source have no bearing on this discussion. One then simply demands continuity. The physical degrees of freedom (DOF) can be summarized in the two cases then as follows, treating the final number f of DOF for the electron as a variable. One reads off on the two rows in **Table 1** so that $0+1 = f_{nm} + 0$ and $0+0 = f_m + 0$, for which f_{nm} and f_m denote, respectively, the final DOE of the electron if no measurement is imposed or when accounting for measurement.

The value of the variable f in the two cases is clear. The electron has a degree of freedom which must manifest itself at the projection-screen if no measurement was taken at the slit-screen, and otherwise it does not. The only degree of freedom potentially available to the elec

Table 1. Physical degrees of freedom (DOE): cases summary.

DOE	Electron Initially	Slit-screen	Electron Finally	Projection-screen
If no measurement	0	1	f_{nm}	0
If measurement	0	0	f_m	0

tron at the projection-screen is its wave intensity because its de Broglie wavelength fixes energy and equivalent variables while the projection-screen itself is at a fixed location. Phase of the wave is arbitrary and so cannot be physically manifested. This explanation actually addresses the notion that physical degrees of freedom are well understood and applied equally in arbitrary classical and quantum physical situations.

4. Acknowledgement

5. References

[1] Itzhak Orion and M. Laitman, "The Double-Slit Experiment and Wave-Particle Duality: Toward a Novel Quantum Interpretation," *Journal of Modern Physics*, 2010, Vol. 1, 90-92. doi:10.4236/jmp.2010.110013

[2] E. Merzbacher, "Quantum Mechanics," 3rd Edition, Wiley, New York, 1998.

[3] J. S. Townsend, "A Modern Approach to Quantum Mechanics," *University Science*, Sausalito, CA, 2000.

[4] R. B. Griffiths, "Consistent Quantum Mechanics," Cambridge University Press, New York, 2001.

[5] H. Goldstein *et al.*, "Classical Mechanics," 3rd Edition, Addison-Wesley, San Francisco, 2002.

[6] L. D. Landau and E. M. Lifshitz, "Mechanics," 3rd Edition, Butterworth-Heinemann, Boston, 2001.

[7] B. d'Espagnat, "Veiled Reality," 1st Edition, Addison-Wesley, New York, 1994.

[8] F. Lindner *et al.*, "Attosecond Double-Slit Experiment," *Physical Review Letters*, Vol. 95, No. 4, 2005.

[9] K. Wang and D. Cao, "Coincidence Subwavelength Interference by a Classical Thermal Light," *Physical Review A*, Vol. 70, 2004.

[10] S. P. Walborn *et al.*, "Double-slit Quantum Eraser," *Physical Review A*, Vol. 65, 2002. ArXiv: quantph/0106078 v. 1

[11] J. Xiong *et al.*, "Experimental Observation of Classical Sub-Wavelength Interference with Thermal-Like Light," 2004. ArXiv: quant-ph/0410020 v. 1

[12] G. Brida *et al.*, "A First Experimental Test of De Broglie-Bohm Theory Against Standard Quantum Mechanics,"

Journal of Physics B: *Atomic, Molecular and Optical Physics*, Vol. 35, 2002. ArXiv: quant-ph/0206196 v. 1, 2002)

[13] A. S. Sanz and F. Borondo, "A Quantum Trajectory Description of Decoherence," *European Physical Journal D*, Vol. 44, 2007. ArXiv: quantph/0310096 v. 4, 2004)

[14] M. Bozic *et al*., "An Asymmetric Double-Slit Interferometer for Small and Large Quantum Particles," 2003 ArXiv: quant-ph/0305189 v. 1

[15] A. G´ozdz and M. Pietrow, "Projection evolution in quantum mechanics," 2003. ArXiv: quantph/0303084 v. 1

[16] P. A. M. Dirac, "General Theory of Relativity," Princeton University Press, New Jersey, 1996.

[17] A. Pich, "The Standard Model of Electroweak Interactions," IFIC/05-13, FTUV/05-0201, 2005. ArXiv: hep-ph/0502010 v. 1