

2010

Quantum Interference Experiments, Modular Variables and Weak Measurements

Jeff Tollaksen

Chapman University, tollakse@chapman.edu

Yakir Aharonov

Chapman University, aharonov@chapman.edu

Aharon Casher

Chapman University

Tirzah Kaufherr

Tel Aviv University

Shmuel Nussinov

Chapman University

Follow this and additional works at: https://digitalcommons.chapman.edu/scs_articles



Part of the [Atomic, Molecular and Optical Physics Commons](#), and the [Quantum Physics Commons](#)

Recommended Citation

Tollaksen, Jeff, Yakir Aharonov, Aharon Casher, Tirzah Kaufherr, and Shmuel Nussinov. "Quantum interference experiments, modular variables and weak measurements." *New Journal of Physics* 12.1 (2010): 013023. doi: 10.1088/1367-2630/12/1/013023

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.

Quantum Interference Experiments, Modular Variables and Weak Measurements

Comments

This article was originally published in *New Journal of Physics*, volume 12, issue 1, in 2011. DOI: [10.1088/1367-2630/12/1/013023](https://doi.org/10.1088/1367-2630/12/1/013023)

Creative Commons License



This work is licensed under a [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/).

Copyright

IOP Publishing Ltd.

Quantum interference experiments, modular variables and weak measurements

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2010 New J. Phys. 12 013023

(<http://iopscience.iop.org/1367-2630/12/1/013023>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 206.211.139.126

This content was downloaded on 10/03/2015 at 16:20

Please note that [terms and conditions apply](#).

Quantum interference experiments, modular variables and weak measurements

Jeff Tollaksen¹, Yakir Aharonov^{1,2}, Aharon Casher^{1,2},
Tirzah Kaufherr² and Shmuel Nussinov^{1,2}

¹ Department of Physics, Computational Science, and Engineering,
Schmid College of Science, Chapman University, 1 University Dr,
Orange, CA 92866, USA

² School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
E-mail: Tollaksen@Chapman.edu

New Journal of Physics **12** (2010) 013023 (29pp)

Received 25 September 2008

Published 28 January 2010

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/12/1/013023

Abstract. We address the problem of interference using the Heisenberg picture and highlight some new aspects through the use of pre-selection, post-selection, weak measurements and modular variables. We present a physical explanation for the different behaviors of a single particle when the distant slit is open or closed; instead of having a quantum wave that passes through all slits, we have a localized particle with non-local interactions with the other slit(s). We introduce a Gedanken experiment to measure this non-local exchange. While the Heisenberg and Schrödinger pictures are equivalent formulations of quantum mechanics, nevertheless, the results discussed here support a new approach to quantum mechanics which has led to new insights, new intuitions, new experiments and even the possibility of new devices that were missing from the old perspective.

Contents

1. Introduction	2
2. Brief review of loss of interference from the Schrödinger perspective	6
2.1. Analyzing changes in probability distribution using method 1: moments of the conserved quantity	7
2.2. Analyzing changes in probability distributions using method 2: Fourier transform of the conserved quantity	9
3. Interference phenomena from the Heisenberg perspective: modular variables	10
3.1. Modular variables are the observables that are sensitive to the relative phase	11
3.2. For the interference phenomena, modular variables satisfy non-local equations of motion	11
3.3. Non-local exchange of modular variables in the double-slit set-up	12
3.4. Why does the interference pattern disappear when the particle is localized?	13
4. The Gedanken experiment to measure non-local equations of motion	14
4.1. Information gain without disturbance: safety in numbers	15
4.2. Pre-selection, post-selection and weak measurements	17
4.3. Applying pre- and post-selection and weak measurements to the interference phenomena	18
5. Discussion	20
6. Conclusion	23
Acknowledgments	23
Appendix A. Proof that the moments do not depend on the relative phase	23
Appendix B. Non-locality in configuration space	24
Appendix C. The non-locality of modular variables is generic to all interference phenomena	24
Appendix D. Conservation law for modular variables	27
References	28

1. Introduction

The two-slit experiment is the quintessential example of the dual character of quantum mechanics. The initial incoming particle seems to behave as a wave when falling on the (left and right) slits, but when recorded on the screen its wavefunction ‘collapses’ into that of a localized particle. By repeating the experiment for an ensemble of many particles, the interference pattern manifests through the density of hits along the screen (aligned with, say, the x -direction): $dn(x)/dx \sim |\psi_L(x) + e^{i\alpha} \psi_R(x)|^2$ with $\psi_L(x)$ coming from the left slit, $\psi_R(x)$ from the right (located a distance D away), and α the relative phase between the left and right parts of the wavefunction (see figure 1).

There are two ways of thinking about such phenomena.

The first accepts the Schrödinger description as given, with wavepackets evolving in time. Indeed, the Schrödinger description has been extremely useful, having served, for example, as the starting point for the Feynman path integral. The apparent analogy between Schrödinger wave interference and classical wave interference (arising from the use of identical calculations) presents a conceptually simple interpretation of quantum phenomena in terms of our classical

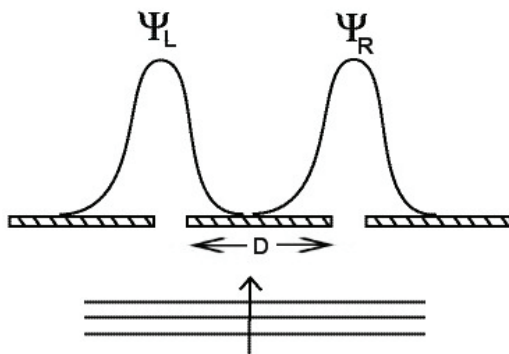


Figure 1. Wavefunction for a single particle in double-slit set-up when we do not know which slit the particle has passed through.

picture. One is often advised to apply this consistent formalism for statistical predictions (providing, in this case, probability distributions for the positions of many particles) *without* asking questions about its interpretation. In fact, the belief that the Schrödinger picture is the *only* way by which the interference and relative phase can be inferred played a central role in the development of the probability amplitude interpretation in the quantum formalism.

The second way of thinking maintains that this is not the end of the matter, and advocates further inquiry. For example, Feynman [3] stated that such phenomena are at the ‘heart of quantum mechanics.’ Which, in reality, ‘contains the *only* mystery.’ Such proponents often seek to obtain as close a correspondence as possible between theory and measurement. As a consequence, they try to weed out ‘classical’ notions when they have been mis-applied to the quantum realm. For example, classical waves involve many degrees of freedom (e.g. field phenomena such as sound and electromagnetic waves) and their phase *can* of course be measured by local experiments. But the meaning of a quantum phase is very different. Multiplying the wavefunction $\psi_L(x) + e^{i\alpha}\psi_R(x)$ by an overall phase ϕ does not change the relative phase α and thus does not yield a different state. Furthermore, it seems that the relative phase α cannot be measured directly on a single particle since it cannot be represented by a Hermitian operator. That is, $\psi_L(x) + e^{i\alpha}\psi_R(x)$ and $\psi_L(x) + e^{i\beta}\psi_R(x)$ are not generally orthogonal and thus cannot be eigenstates belonging to different eigenvalues of a Hermitian operator. In further contrast to the classical phase, a change in the relative quantum phase—say from $\psi_L(x) + \psi_R(x)$ to $\psi_L(x) - \psi_R(x)$ —would not result in a measurable change in any local properties. The change only shows up in certain non-local properties, or much later when the two separate components $\psi_L(x)$ and $\psi_R(x)$ eventually overlap and interfere. It seems that the relative phase cannot be thought of simply as the difference between a local phase at $\psi_L(x)$ and another local phase at $\psi_R(x)$.

Another aspect of this second way of thinking is the realization that the Schrödinger wave has a measurable meaning only for an ensemble of particles, not generally for a single particle. This therefore leaves important questions unanswered concerning the *physics* of interference from the perspective of a single particle: if physics obeys local dynamics, then how does the localized particle passing through the right slit sense whether or not the distant left slit is open (closed), causing it to scatter (or not scatter) into a region of destructive interference? Interference experiments have been performed with electron/photon beams whose intensity is so small that only one electron/photon traverses the interference apparatus at a

time³. The interference pattern with light and dark bands is nevertheless built up successively, mark by mark, with each individual ‘particle-like’ electron/photon [24]. One is then confronted with the fact that a single degree of freedom created the interference pattern. This mystery led Feynman to declare: ‘Nobody knows how it can be like that’ [3].

We follow the second way of thinking and offer a fresh approach to this time-honored problem [8]–[10], [18] and [27]. To motivate the first step, involving a fundamental shift in the types of observables utilized, we make several observations.

- Firstly, most discussions of this problem are based on measurements that disturb the interfering particle. This is one of the main reasons that quantum interference is generally considered to be intimately associated with the problems that stem from the statistical character of the quantal description.
- Secondly, the observables studied to date have been simple functions of position and momentum. These observables, however, are *not* sensitive to the relative phase between different ‘lumps’ of the wavefunction (centered around each slit). Nevertheless, the subsequent interference pattern of course is entirely determined by the relative phase between these ‘lumps’, suggesting that simple moments of position and momentum are not the most appropriate dynamical variables to describe quantum interference phenomena.
- Thirdly, operators that *are* sensitive to the relative phase are exponentials of the position and momentum.

We address the first observation with non-disturbing measurements. To date, several non-disturbing measurements, such as weak measurements and protective measurements, have stimulated lively debates and have proven useful for separating various aspects of quantum theory from the probabilistic aspects [26]. The underlying framework for the approach to interference presented in this paper is based on another kind of non-disturbing measurement, the ‘set of deterministic operators’ or ‘deterministic experiments’ [8, 18, 27]. This set involves measurement of only those variables for which the state of the system under investigation is an eigenstate. This set answers the question ‘What is the set of Hermitian operators \hat{A}_ψ for which ψ is an eigenstate?’ for any state ψ , i.e. $\hat{A}_\psi = \{\hat{A}_i \text{ so that } \hat{A}_i|\psi(t)\rangle = a_i|\psi(t)\rangle, a_i \in \mathfrak{R}\}$. This question is dual to the more familiar question ‘What are the eigenstates of a given operator?’ Measurement of these operators \hat{A}_ψ does not collapse the wavefunction, since the wavefunction is initially an eigenstate of the operator being measured. Elaboration of this framework is left to the existing and forthcoming literature [8]–[10], [27]. The essential point needed for this article is the relevance of deterministic experiments for a *single* particle since they can be performed without causing a disturbance.

We address the second and third observations by performing yet another kind of non-disturbing measurement, namely weak measurements, on the observables that *are* sensitive to the relative phase. These observables that are sensitive to the relative phase are functions of modular variables. For the case of interference in space, as considered here, the relevant modular variable is modular momentum, not ordinary momentum. These observables are also members of the ‘deterministic set of operators’ and are relevant for an individual particle. We then see that in the context of interference phenomena, the Heisenberg equations of motion for these modular

³ ‘The most beautiful experiment in physics, according to a poll of Physics World readers, is the interference of single electrons in a Young’s double slit.’ Second place went to Galileo’s experiment with falling bodies. Third place went to Millikan’s oil experiment. Fourth went to Newton’s prism, etc [36].

variables are *non-local*. The non-locality of these observables is quite intuitive; the operators sensitive to the relative phase simply translate the different ‘lumps’ of the wavefunction. The appropriate translation may cause one lump to overlap with another lump or to overlap simply with the region where the distant slit is either open (or closed). This provides a physical explanation for the different behaviors of a *single* particle when the distant slit is open or closed. It therefore provides the under-pinnings for a new ontology based on localized particles with non-local interactions, rather than a less physical Schrödinger ‘wave of probability’ traveling throughout all of space.

This kind of non-locality, which is revealed in the equations of motion, is *dynamical* non-locality, to distinguish it from kinematic non-locality [17, 23]⁴ implicit in quantum correlations. These two kinds of non-locality are fundamentally different. Kinematic non-locality arises from the structure of Hilbert space and does not create any change in probability distributions, causes and effects cannot be distinguished, and therefore ‘action-at-a-distance’ cannot manifest. Kinematic non-locality has been extremely useful, having catalyzed, e.g., much of the progress in quantum information science⁵. On the other hand, dynamical non-locality arises from the structure of the equations of motion and *does* create explicit changes in probability, though in a ‘causality-preserving’ manner. This approach was first introduced by Aharonov, Pendelton and Petersen (APP) [8] in order to explain the non-locality of topological phenomena such as the Aharonov–Bohm (AB) effect [1, 25]⁶. The AB effect conclusively proved that a magnetic (or electric) field inside a confined region can have a measurable impact on a charged particle that never traveled inside the region. In order to represent the closest correspondence between measurement and theory, APP introduced non-local interactions between the particle and field. This was in contrast to the prevailing approach of reifying local interactions with (unphysical) non-gauge invariant quantities outside the confined region, such as the vector (and/or scalar) potential.

Both dynamic and kinematic non-locality are generic and can be found in almost every type of quantum phenomenon [4]. Prior to APP, dynamical non-locality was avoided due to the possibility that it could violate causality. However, in a beautiful theorem, APP proved that the dynamical non-locality they introduced could never violate causality. They considered the general set of conditions necessary to see the non-local exchange of modular variables, for example when the left slit is either monitored or closed and the particle is localized around the right slit. APP proved that these are precisely the same conditions that make the non-local exchange completely uncertain and therefore ‘unobservable’.

Although it was an attractive quality of quantum mechanics that it allowed ‘action-at-a-distance’ to ‘peacefully coexist’ with causality, this theorem nevertheless proved to be somewhat anti-climatic; if we cannot actually observe the non-local exchange of modular variables, then have we not violated the dictum of maintaining the closest correspondence between measurement and theory by claiming the existence of a new kind of non-local—yet unobservable—effect?

⁴ Shimony also states ‘the first confirmation of entanglement ... antedated Bell’s work, since Bohm and Aharonov [17]’, see Shimony [23].

⁵ Los Alamos Quantum Computing Road map [13]: ‘At least two important precursors to this [quantum computing] paradigm shift had critical influence’, citing non-locality, e.g. the Einstein–Podolsky–Rosen–Bohm (EPRB) [23] and Aharonov–Bohm (AB) effects and developments in quantum information theory.

⁶ Yakir Aharonov described this general approach to interference to Heisenberg personally. Professor Heisenberg had never thought before about interference phenomenon in the Heisenberg picture and was extremely pleased by it.

One of the principal new results presented in this paper is to show, for the first time, that these non-local interactions *can* be observed⁷. This has to be done in a causality-preserving manner. Therefore, in order to measure this non-locality, we must utilize various tools such as pre-selection, post-selection and weak measurements. Although some of the components utilized in the present analysis were published long ago, they are not generally known and are therefore briefly reviewed.

With this development, we have thereby underscored a fundamental difference between classical mechanics and quantum mechanics that is easily missed from the perspective of the Schrödinger picture; the equations of motion for observables relevant to quantum mechanical interference phenomena can be non-local in a peculiar way that preserves causality. These novel results motivate a new approach to quantum mechanics, starting from the Heisenberg picture and involving the set of deterministic operators. While the new framework and its associated language is, in principle, equivalent to the Schrödinger formulation, it has led to new insights, new intuitions, new experiments and even the possibility of new devices that were missing from the old perspective. These types of development are indicative of a successful re-formulation.

Although further elaboration of this new approach is left to a future article [18], we briefly mention one important conceptual shift: when quantum mechanics is compared to classical mechanics, often the uncertainty or indeterminism of quantum mechanics is emphasized and the profound, fundamental differences in the dynamics are ignored. This is perhaps a result of the similarity between the classical dynamical description (Poisson bracket) and the quantum dynamical description (commutator) for simple functions of momentum or position. Furthermore, uncertainty is viewed in a kind of ‘negative’ light; as a result of the uncertainty in quantum mechanics, we have lost the ability that we had in classical mechanics to predict the future. Not only is nature ‘capricious’, but it seems that we do not even gain anything from the uncertainty.

The new approach allows us to change this perspective by deriving uncertainty from principles that we argue are more fundamental, namely from non-locality and causality. This changes the meaning of uncertainty from one with a negative connotation to one with a positive connotation. Something similar happened with special relativity when the axioms of relativity were discovered. This inspired a modification of the old language, e.g. that light has the same velocity in all reference frames is certainly highly unusual, but everything works in a self-consistent way due to the axiomatic framework, and because of this, special relativity is rather easy to understand. Similarly, we are convinced that the new approach arising from this paper will lead to a deeper understanding of the nature of quantum mechanics.

2. Brief review of loss of interference from the Schrödinger perspective

We begin by reviewing past attempts to analyze the disappearance of interference in situations where it is possible to detect which slit the particle passes through. The original debate was famously conducted by Einstein and Bohr. Einstein attempted to challenge the consistency of quantum mechanics by arguing that a Which Way Measurement (WWM) could be performed without destroying the interference pattern by measuring the transverse recoil (i.e. the transverse momentum kick) of the double-slit screen after the particle passed through. Bohr maintained that the consistency of quantum mechanics depended on the destruction of the interference pattern

⁷ These results were first presented in [37].

when WWM information is obtained. He showed that the measurement-induced uncertainty created in the transverse position of the screen by an accurate measurement of the transverse momentum was sufficient to destroy the interference pattern.

This reasoning leads to a paradox, the resolution of which sets the stage for some of the basic elements in our approach. It has been argued (borrowing from the discussion of the ‘Heisenberg microscope’) that if the particles were ‘observed’ at the right slit, then the photon involved in this observation should have a wavelength $\lambda \leq D/2$ and a corresponding momentum uncertainty $\Delta p > 2\hbar/D$. This momentum uncertainty is imparted to the particle making its wave number $k = p/\hbar$ uncertain, thereby destroying the interference pattern.

This argument is incorrect. To see this, assume that a sensitive detector, placed at the left slit, failed to detect any particle. We then know that all particles passed through the right slit. The interference pattern will then be completely destroyed despite the fact that there was no interaction with the detector [5, 8]. One might suppose that since the action of opening/closing the left slit never caused an interaction with the particle at the right slit, nothing associated with the particle should change. But it was first pointed out by APP [8] that in this scenario when a WWM is performed without actually interacting with the interfering particle, the probability distribution of the momenta *does* change, although none of the moments of the momenta change.

To best resolve this paradox, we need to take a step back. We note that the effect of a generic interaction or collision between any two quantum systems can be characterized by a change in the probability distribution of the momentum, i.e. going from an initial probability distribution, $\rho_i(p)$, to a final distribution, $\rho_f(p)$. We can analyze this change in two ways⁸.

1. Look at moments such as $\langle p^n \rangle = \int \rho(p) p^n dp$ and calculate $\delta \langle p^n \rangle = \langle p^n \rangle_f - \langle p^n \rangle_i$, and thus ask how the interaction affected these averages. This is the usual approach.
2. Alternatively, we may look at the Fourier transform of the probability distribution $\int \rho(p) e^{(i/\hbar)pD} dp$. (We will later see that these functions, $\langle e^{(i/\hbar)pD} \rangle$, are precisely the observables that are sensitive to the relative phase.) To analyze the effect of the interaction, we calculate $\langle e^{(i/\hbar)pD} \rangle_f - \langle e^{(i/\hbar)pD} \rangle_i$ and ask how the interaction affected these averages.

In principle, one can discuss the effect of interactions using (1) or (2), since knowing (2) for all D is equivalent to knowing (1) for all n .

2.1. Analyzing changes in probability distribution using method 1: moments of the conserved quantity

Scully *et al* [5] and Storey *et al* [6] further debated the issues introduced by APP, resulting in many hundreds of cited papers.

Scully *et al* were dissatisfied with Bohr’s original response to Einstein. They suggested that a microscopic pointer (i.e. a micro-maser) could be used in such a way that the interference in a WWM is destroyed without imparting any momentum to the particle (just as we alluded to earlier in the discussion of the case in which a sensitive detector failed to find the particle at the left slit).

However, Storey *et al* countered this, stating that the momentum distribution does change when WWMs are made. They noted that having a plane wave with initial $\Delta x = \infty$ and $\Delta p = 0$ impinge on the two slits projects the initial plane wave on to ‘lumps’, which therefore have a significant Δp .

⁸ We consider momentum here, but our comments apply to any conserved quantity.

The principal components of both camps' arguments were previously put forward in APP, i.e. there is both a change in probability and no change in the moments. But can we actually observe the change in the probability of the momentum when the left slit is open or closed? To determine whether the momentum is disturbed by the WWM, the momentum of the particle must be known before the WWM and after. However, if an ideal measurement is made of the momentum before the WWM, then we have effectively measured the interference, rendering useless the subsequent WWM.

The techniques of weak measurement have proven very useful in scenarios like this requiring manifestation of two opposing situations, i.e. to have a 'have-your-cake-and-eat-it' solution. Weak measurements have had a direct impact on the central 'mystery' alluded to by Feynman concerning indeterminism, namely the fact that the past does not completely determine the future. This mystery was accentuated by an assumed 'time-asymmetry' within quantum mechanics, namely the assumption that measurements only have consequences *after* they are performed, i.e. toward the future. Nevertheless, a positive spin was placed on quantum mechanics' non-trivial relationship between initial and final conditions by Aharonov, Bergmann and Lebowitz (ABL) [2] who showed that the new information obtained from future measurements was also relevant for the *past* of quantum systems and not just the future. This inspired ABL to reformulate quantum mechanics in terms of *pre- and post-selected ensembles*. The traditional paradigm for ensembles is to simply prepare systems in a particular state and thereafter subject them to a variety of experiments. These are 'pre-selected-only-ensembles'. For *pre-and-post-selected*-ensembles, we add one more step, a subsequent measurement or post-selection. By collecting only a subset of the outcomes for this later measurement, we see that the 'pre-selected-only-ensemble' can be divided into sub-ensembles according to the results of this subsequent 'post-selection-measurement'. Because pre- and post-selected ensembles are the most refined quantum ensemble, they are of fundamental importance and have revealed novel aspects of quantum mechanics that were missed before, particularly the weak value that has been confirmed in numerous weak measurement experiments. Weak values have led to quantitative progress in many questions in the foundations of physics [15], including interference [10], field theory, tunneling, quantum information such as the quantum random walk, foundational questions, the discovery of new aspects of mathematics, such as Super-Fourier or super-oscillations, and so on. It has also led to generalizations of quantum mechanics that were missed before.

While it is standard lore that the wave and particle nature cannot manifest at the same time, weak measurements on pre- and post-selected ensembles *can* provide information about both the (pre-selected) interference pattern and the (post-selected) direction of motion for each particle. This aspect of weak measurements formed the basis for the first application of weak measurements to study the change in momentum for WWM within the double-slit set-up as presented by Wiseman [7]. This was followed by an experiment (Mir *et al* [19]). Besides clarifying the different definitions and different measurements (etc) used by both sides of the debate, Wiseman [7] and Mir *et al* [19] show that the momentum transfer can be observed for the spatial wavefunction used in the two slits (as opposed to momentum eigenstates) by using weak measurements.

They implemented the weak measurement with position shifts and polarization rotations in a large optical interferometer. Plotting the conditional probability to obtain a particular momentum (given the appropriate post-selection) and integrating over all possible post-selections, they were able to verify both the Scully and Storey viewpoints. With respect to

Scully [5], they show that none of the moments of the momentum change. With respect to Storey [6], they show that the momentum does extend beyond a certain width.

However, there are inherent limitations to any approach based on analyzing changes in the probability for momenta through changes in the moments. For example, while momentum is of course conserved, there is no definite connection between the probability of an individual momentum before and after an exchange between the interfering particle and the slit. Furthermore, the analysis in terms of moments does not offer any intuition as to *how* or *why* the probability of momentum changes.

2.2. Analyzing changes in probability distributions using method 2: Fourier transform of the conserved quantity

When compared to the first (traditional) approach based on the moments, the second approach focusing on the Fourier transform of the probability distribution has many advantages, both mathematical and physical. In this section, we briefly review some of the mathematical advantages, leaving most of the physical advantages to the rest of the article.

The first ‘moments’ approach to interference derived from intuitions developed with wavefunctions consisting of just one ‘lump’. In these cases, the averages of x (or of p) evolve according to *local* classical equations of motion. Also the uncertainties $(\Delta x)^2 \equiv (\bar{\hat{x}})^2 - \bar{\hat{x}}^2$ and $(\Delta p)^2 \equiv (\bar{\hat{p}})^2 - \bar{\hat{p}}^2$, describing the spread in these variables, have properties similar to those of the spread of variables in a classical situation with unsharply defined initial conditions and which evolve according to diffusion-like rules.

This drastically changes when we have two or more separate ‘lumps’ of the wavefunction. Indeed, the wavefunction, after passing through the symmetric two slits, consists of a superposition of two identical, but physically disjointed ‘lumps’, ψ_L and ψ_R (see figure 1):

$$|\Psi_\alpha\rangle = \frac{1}{\sqrt{2}}\{|\psi_L\rangle + e^{i\alpha}|\psi_R\rangle\}. \quad (1)$$

Collapsing it to just $\psi_R(x) \equiv \langle x|\psi_R\rangle$ does *not* change either Δp or the expectation values of any finite-order polynomial in p , as none of these local operators has a non-vanishing matrix element between the disjointed ‘lumps’ of the wavefunction. In other words, measuring which slit the particle passes through does not have to increase the uncertainty in momentum. Later in this article we will review another uncertainty relationship that is more relevant for this issue.

So far we have focused on the disappearance of interference upon WWM. But the other fundamental mystery highlighted by Feynman remains; namely, how does a particle localized at the right slit ‘know’ whether the left slit is open or closed? The first approach based on moments tells us nothing about this mystery. The decisive importance of the second ‘Fourier transform’ approach for this mystery is best illustrated through a basic theorem that characterizes all interference phenomenon: all *moments* of both position and momentum are *independent* of the relative phase parameter α (until the wavepackets overlap).

Theorem 1. *Let $\Psi_\alpha = \psi_L(x, t) + e^{i\alpha}\psi_R(x, t)$ such that there is no overlap of $\psi_L(x, 0)$ and $\psi_R(x, 0)$. If n and m are integers, then for all values of t , and choices of α, β :*

$$\int [\Psi_\alpha^*(x, t)\Psi_\alpha(x, t) - \Psi_\beta^*(x, t)\Psi_\beta(x, t)]x^m p^n dx = 0. \quad (2)$$

For the particular double-slit wavefunction, it is easy to see that if there is no overlap between ψ_L and ψ_R then nothing of the form $\int_{-\infty}^{\infty} \Psi^* x^m p^n \Psi dx$ will depend on α for any value of m and n . Furthermore, expanding $\int \{\psi_L + e^{-i\alpha} \psi_R\}^* x^m p^n \{\psi_L + e^{i\alpha} \psi_R\} dx$, we see that only the cross terms, i.e. $\langle \psi_L | x^m p^n | e^{i\alpha} \psi_R \rangle$, have the possibility of depending on α ; but operators of the form $x^m p^n$ cannot change the fact that ψ_R and ψ_L do not overlap. When integrated, these terms vanish and are therefore insensitive to the relative phase.

This suggests that these dynamical variables (e.g. $\langle x \rangle$, $\langle p \rangle$, Δx , Δp) are not the most appropriate to describe quantum interference phenomena. What observables, then, are sensitive to this interference information, which appears to be stored in a subtle fashion? To fully capture the physics of these scenarios with wavefunctions composed of multiple lumps, non-polynomial and *non-local* operators, connecting the disjoint parts, are required. For many, equidistant slits, these are the discrete translation by $\pm D$, namely $\exp\{\pm(i/\hbar)\hat{p}D\}$, effecting $\exp\{-(i/\hbar)\hat{p}D\}\psi_R(x) \rightarrow \psi_R(x - D)$, which overlaps with $\psi_L(x)$. The expectation value of the translation operator $\exp\{(i/\hbar)\hat{p}D\}$ does depend on α : $\langle \Psi_\alpha | \exp\{i\hat{p}D/\hbar\} | \Psi_\alpha \rangle = e^{-i\alpha}/2$.

This provides the basis for a mechanism to explain *how* the particle at the right ‘knows’ what is happening at the left slit. As we will see, the second ‘Fourier transform’ approach even provides us with the parameters relevant for this question (namely the distance between the slits), while the first ‘moments’ approach remains silent.

Before proceeding, in the next section, to the *physics* of interference for single particles, we briefly mention two additional mathematical advantages concerning the second ‘Fourier transform’ approach.

Firstly, all the moments $\langle p^n \rangle$ are averages of unbounded quantities, while $\langle \exp\{(i/\hbar)\hat{p}D\} \rangle$ are averages of bounded quantities. There are problems with unbounded quantities (as pointed out by Mir *et al*). Infinitesimal changes in $\rho(p)$ can cause very large changes in the moments $\langle p^n \rangle$. To see this, consider a negligible change, $\delta\rho(p)$, in $\rho(p)$. By ‘negligible’, we mean there is only a small change in the probability distribution. If we calculate $\delta\langle p^n \rangle = \int \delta\rho(p) p^n dp$, we can get a finite change if $\delta\rho(p)$ differs from zero at a sufficiently large p . In the limit, we could in fact consider $p \rightarrow \infty$ and $\delta\rho(p) \rightarrow 0$, in such a fashion that Δp^n is finite. Then clearly $\delta\langle p^{n+1} \rangle$ diverges as do all higher moments. The second ‘Fourier transform’ approach never has these kinds of problems and is always finite.

The other significant ‘mathematical’ difference concerns the utility of conservation laws. As mentioned in section 2.1, while conservation of momenta is certainly maintained for the averages of moments, there is no definite connection between an individual momentum before and after an exchange in this general kind of set-up. As we shall see below, the second ‘Fourier transform’ approach uncovers an exchange of a new conserved quantity. The conservation law for these quantities can be expressed in a ‘product-form’ rather than a sum (as occurs for ordinary momentum). This product-form conservation law is more relevant for many situations such as a change in relative phase.

3. Interference phenomena from the Heisenberg perspective: modular variables

As we have argued previously, the basic gauge symmetry would be violated if any quantum experiment could measure the local phase in $|\Psi_\alpha\rangle$ and therefore there is no *locally accessible phase* information in $|\Psi_\alpha\rangle$. The relative phase is a truly *non-local* feature of quantum mechanics. This point is often missed when the Schrödinger picture is taught and classical intuitions are

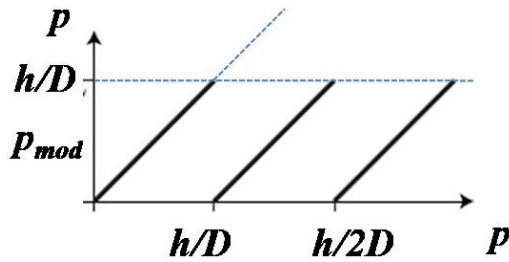


Figure 2. Stacking an integer number (N_p) of h/D on top of the modular portion of p (p_{mod}).

applied to interference. For this and other reasons, we maintain that the non-local aspect of interference is clearer in the Heisenberg picture.

3.1. Modular variables are the observables that are sensitive to the relative phase

In section 2.2, we pointed to the significance of the Heisenberg translation operator, $\exp\{\pm(i/\hbar)\hat{p}D\}$, effecting $\exp\{-(i/\hbar)\hat{p}D\}\psi_R(x) \rightarrow \psi_R(x - D)$ overlapping with $\psi_L(x)$. Therefore, the expectation value of the translation operator $\exp\{(i/\hbar)\hat{p}D\}$ does depend on α : $\langle \Psi_\alpha | \exp\{i\hat{p}D/h\} | \Psi_\alpha \rangle = e^{-i\alpha}/2$.

But exactly *what* information about α does $\langle \exp\{\pm(i/\hbar)\hat{p}D\} \rangle$ reveal? It is easy to see that if we replace p with $p - (nh/D)$ (n is the largest integer such that $n(h/D) < p$ (i.e. satisfying $0 \leq \hat{p} - n(h/D) \leq h/D$), then $e^{(i/\hbar)\hat{p}D}$ changes by $e^{(iD/\hbar)(nh/D)} = e^{in2\pi} = 1$, i.e. nothing changes. This means that $e^{(i/\hbar)\hat{p}D}$ gives us information about the remainder after this integer number of h/D is subtracted from p . This is otherwise known as the modular momentum $p_{\text{mod}} \equiv \hat{p}$ modulo h/D (see figure 2) defined by \hat{p} modulo $h/D \equiv \hat{p} - n(h/D)$.

It is clear that $p \text{ mod } h/D$ has the topology of a circle, like any periodic function. Every point on the circle is another possible value for p_{mod} . We deal with modular quantities every time we look at a wristwatch that displays the time modulo 12.

We can get back to ordinary momentum through the relation

$$p = N_p \frac{h}{D} + p_{\text{mod}}. \quad (3)$$

We can see this (figure 2) if we stack an integer number (N_p) of h/D on top of the modular portion of p (p_{mod} is the lower portion of figure 2). Note that the eigenstates of the translation operator $\exp\{(i/\hbar)\hat{p}D\}$ are also eigenstates of the modular momentum p_{mod} .

3.2. For the interference phenomena, modular variables satisfy non-local equations of motion

The key to our explanation of interference from the single particle perspective is the non-local equations of motion satisfied by these modular variables. Thus, using $H = (p^2/2m) + V(x)$ and $e^{(i/\hbar)\hat{p}D}V(x)e^{-(i/\hbar)\hat{p}D} = V(x + D)$, we find *non-local* [8, 10] Heisenberg equations of motion for modular variables:

$$\frac{d}{dt}e^{(i/\hbar)\hat{p}D} = \frac{i}{\hbar}[H, e^{(i/\hbar)\hat{p}D}] = \frac{i}{\hbar}[V(x) - V(x + D)]e^{(i/\hbar)\hat{p}D}, \quad (4)$$

with $e^{(i/\hbar)\hat{p}D}$ changing even when $\partial V/\partial x = 0$.

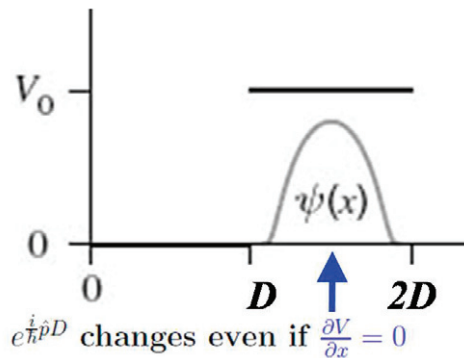


Figure 3. A potential with two values and a wavepacket with support only in the interval $D < x < DL$.

This, essentially quantum, phenomenon has no classical counterpart. The classical equations of motion for any function $f(p)$ derive from the Poisson bracket

$$\frac{df(p)}{dt} = \{f(p), H\}_{\text{PB}} = -\frac{\partial f}{\partial p} \frac{\partial H}{\partial x} + \underbrace{\frac{\partial f}{\partial x} \frac{\partial H}{\partial p}}_{=0} = 0, \quad (5)$$

i.e. $f(p)$ changes only if $\partial V/\partial x \neq 0$ at the particle's location.

Unlike the Poisson bracket in classical mechanics, quantum mechanics has non-trivial and unique solutions to the commutator, $[f(p), g(x)] = 0$ if $f(p) = f(p + p_o)$, $g(x) = g(x + x_o)$ and $x_o p_o = h$. This leads us to a new structure within the quantum mechanics of periodic functions which leads naturally to the concept of modular variables. In our particular case, $x_o = D$, $p_o = \frac{h}{D}$ and $f(p) = e^{i/hbar p D}$, so $g(x)$ only depends on the function $g(x)$ modulus D and $f(p)$ only depends on the function $f(p)$ modulus $\frac{h}{D}$. The non-local equations of motion that $e^{i/hbar p D}$ satisfies show how the potential at the left slit *does* affect the evolution of the modular variable even when we consider a particle located at the right slit (and vice versa, see figure 3). Modular variables obey non-local equations of motion independent of the specific state of the Schrödinger wavefunction, whether it is localized around one slit or in a superposition. Nevertheless, the modular momentum may change (non-locally) even if the wavefunction experiences no force. We can therefore see that the non-local effect of the open or closed slit is to produce a shift in the modular momentum of the particle while leaving the expectation values of moments of its momentum unaltered.

3.3. Non-local exchange of modular variables in the double-slit set-up

For the special double-slit case, a set of spin-like observables can be identified as members of the set of deterministic operators. For simplicity (without affecting the generality of our arguments), we can express the relevant modular variable as the parity (exchange) operation \hat{P} (effecting $\hat{P}|\psi_L\rangle = |\psi_R\rangle$ and $\hat{P}|\psi_R\rangle = |\psi_L\rangle$). It is sensitive to the relative phase α between the disjoint lumps of equation (1) [8, 10, 27]:

$$\begin{aligned} \langle \Psi_\alpha | \hat{P} | \Psi_\alpha \rangle &= \frac{1}{2} \{ \langle \psi_L | + e^{-i\alpha} \langle \psi_R | \} \hat{P} \{ | \psi_L \rangle + e^{i\alpha} | \psi_R \rangle \} \\ &= \frac{1}{2} \{ e^{i\alpha} + e^{-i\alpha} \} = \langle \cos \alpha \rangle. \end{aligned} \quad (6)$$

To simplify further, we will focus on the ± 1 eigenstates of \hat{P} : $\psi_L(x) + \psi_R(x)$ and $\psi_L(x) - \psi_R(x)$. A measurement of which slit the particle goes through (i.e. a WWM) will change the value of $\langle \hat{P} \rangle$. For example, if the initial state is $|\psi_L\rangle + |\psi_R\rangle$, then $\langle \cos \alpha \rangle = 1$, i.e. $\langle \hat{P} \rangle = 1$. If we collapse the state to $|\psi_R\rangle$, then $\langle \cos \alpha \rangle = 0$ and $\langle \hat{P} \rangle = 0$.

We can also see from equation (4) that if the left slit is open, then $V(x) - V(x + D) = 0$, and therefore p_{mod} is conserved. However, if the left slit is closed, then $V(x) - V(x + D) \neq 0$ and p_{mod} is not conserved.

3.4. Why does the interference pattern disappear when the particle is localized?

When we obtain WWM information, we collapse the superposition from $|\Psi_\alpha\rangle$ to $|\psi_L\rangle$ or $|\psi_R\rangle$ (in the Schrödinger picture). In the Heisenberg picture, however, we cannot describe the collapse of a superposition. The wavefunction is still of course relevant as a boundary condition, but it does not evolve in time. Only the operators evolve in time according to the Heisenberg equation of motion:

$$\frac{dA_H}{dt} = \frac{i}{\hbar} [H, A_H] + U^{-1}(t) \frac{\partial A_S}{\partial t} U(t).$$

But which operators become uncertain when WWM information is obtained?

Suppose again that the particle travels through the right slit and that we choose either to open or close the left slit. This action causes a non-local exchange of modular momentum between the potential at the left slit and the particle going through the right slit. Is this observable?

Until the time of writing of this paper, it was believed that this could not be observed. The reason is that modular momentum (unlike ordinary momentum) becomes, upon detecting (or failing to detect) the particle at a particular slit, *maximally* uncertain. In other words, the effect of introducing a potential at a distance D from the particle (i.e. of opening a slit) is equivalent to a rotation in the space of the modular variable—let us call it θ —that is exchanged non-locally. Suppose the amount of non-local exchange is given by $\delta\theta$ (i.e. $\theta \rightarrow \theta + \delta\theta$). Now ‘maximal uncertainty’ means that the probability to find a given value of θ is independent of θ , i.e. $P(\theta) = \text{constant} = 1/2\pi$. Under these circumstances, the shift in θ to $\theta + \delta\theta$ will introduce no observable effect, since the probability to measure a given value of θ , say θ_1 , will be the same before and after the shift, $P(\theta_1) = P(\theta_1 + \delta\theta_1)$. We shall call a variable that satisfies this condition a ‘completely uncertain variable’. Using this, APP proved a stronger *qualitative* uncertainty principle for the modular momentum, instead of the usual quantitative statement of the uncertainty principle (e.g. $\Delta p D \geq \hbar$); if the non-local exchange of any modular variable θ came close to violating causality, then the probability distribution for all averages of that modular variable flattens out, i.e. every value for θ became equally probable and change in θ becomes unmeasurable.

Theorem 2. *Qualitative uncertainty principle for modular variables: if $\langle e^{in\theta} \rangle = 0$ for any integer $n \neq 0$ and if θ is a periodic function with period τ , then θ is completely uncertain if θ is uniformly distributed on the unit circle.*

Proof. We expand the probability density $\text{Prob}(\theta)$ to a Fourier series $\text{Prob}(\theta) = \sum_{n=-\infty}^{+\infty} a_n e^{in\theta}$ (integer n is a requirement for the function to be periodic in θ), where $a_n = \int \text{Prob}(\theta) e^{in\theta} d\theta = \langle e^{in\theta} \rangle$ (since the average of any function is given by the integral of the function with the probability). We see that $\text{Prob}(\theta) = \text{const}$ if and only if $a_n = 0$ for all $n \neq 0$, and therefore $\langle e^{in\theta} \rangle = 0$ for $n \neq 0$. \square

Consider how this works in the double-slit set-up. Let us start with a particular $|\Psi_\alpha\rangle$, namely the symmetric ($\alpha = 0$, $|\psi_L\rangle + |\psi_R\rangle$) or anti-symmetric ($\alpha = \pi$, $|\psi_L\rangle - |\psi_R\rangle$) states. The parity \hat{P} then has sharp eigenvalues ± 1 . However, \hat{P} becomes maximally uncertain when the state is localized at one slit: by definition, $\bar{P}^2 = 1$; however, $\langle \psi_L | \hat{P} | \psi_L \rangle = \langle \psi_L | \psi_R \rangle = 0$ so $\bar{P} = 0$, and therefore $\Delta \hat{P} \equiv \sqrt{(\bar{P})^2 - \bar{P}^2} = 1$, i.e. we have maximal uncertainty when the particle is localized at one slit. Stated differently, when the particle is at the right (or left) slit, its wave function is a superposition with *equal* weights of the two parity eigenstates $|\psi_L\rangle \pm |\psi_R\rangle$ with ± 1 eigenvalues, which by definition is the state of maximal variance of the operator involved.⁹

The vanishing of the expectation value of the modular momentum variable (as per theorem 2) is the manifestation in our present picture of the loss of information on α and of the interference pattern, once we localize the particle at the left or right slit.¹⁰

This brings us to what we believe to be a more physical answer (from the perspective of an individual particle) for the disappearance of interference: the momentum exchange with the left slit and the resulting momentum uncertainty (destroying the interference pattern when the left slit is closed) are not that of ordinary momentum since, as we noted, Δp does not change. Rather, the closing of the left slit and localization of the particle at the right slit involve a non-local exchange of *modular* momentum. This phenomenon can also be demonstrated for any refinement of the double slit. For example, any measurement at the left slit introduces an uncertain potential there. As a result of the non-local equations of motion, this introduces complete uncertainty in the modular variable. Thus, detecting which slit a particle passes through destroys all information about the modular momentum.

It therefore appears that no observable effect of one slit acting on the particle traveling through the other slit can be obtained via the non-local equations of motion of the modular variable and, therefore, this non-locality ‘peacefully co-exists’ with causality. Have we not violated the dictum of maintaining the closest correspondence between measurement and theory by claiming the existence of a new kind of non-local—yet unobservable—effect?

The key novel observation we next make is that this non-locality *does have* an observable meaning in the context of weak measurements on pre- and post-selected ensembles [12, 14].

4. The Gedanken experiment to measure non-local equations of motion

What are the general issues involved in any attempt to measure this kind of non-locality?

- *First.* If we start with the state $\psi_R + \psi_L$, i.e. a wavepacket around each slit, then the modular momentum is known but we cannot argue that the particle goes through one slit and is affected non-locally by the other slit. Therefore, we need to start with a state that is localized around one slit.
- *Second.* But under these circumstances when the particle is localized around one slit, the modular variable is completely uncertain and therefore unobservable. How can we get around this fact in order to observe this non-locality?
- *Third.* If we are able to get around this fact, then how is causality not violated?

⁹ In passing, we note that this is readily extended from the $Z(2)$ case of just two slits to the $Z(N)$ case of N equidistant, equal slits with periodic boundary conditions (see appendix B).

¹⁰ Although much of the discussion in this article focuses on the simplest interference example with two slits, our approach becomes clearer when it is applied to an infinite number of slits (see appendix C).

As we mentioned previously, weak measurements allow us to ‘have our cake and eat it’ to a certain extent. To address the first issue, we use pre- and post-selection to arrange for a localized particle property (pre-selection). To address the second issue, we later post-select a definite state of modular momentum. (We are interested in particular post-selections, rather than averages over all pre- (and/or) post-selections as done in Mir *et al* [19].) We may perform a weak measurement in order to see the weak value of the modular momentum. This weak measurement has a negligible probability to kick a particle centered around the right slit to the left slit, so we still satisfy the first criterion. Finally, because we must rely on a post-selection and because of the nature of the weak measurement, it is impossible to violate causality with this method.

We proceed now to address each of these issues.

4.1. Information gain without disturbance: safety in numbers

Traditionally, it was believed that if a measurement interaction is weakened so that there is no disturbance on the system, then no information will be obtained. However, it has been shown that information can be obtained even though not a single particle (in an ensemble) was disturbed. To make this article as self-contained as possible, we briefly recall some simple components of this argument [12, 26]. Let us start by considering a general theorem for any vector (state) in Hilbert space.

Theorem 3. $\hat{A}|\psi\rangle = \langle\hat{A}\rangle|\psi\rangle + \Delta A|\psi_{\perp}\rangle$, where $\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle$, $|\psi\rangle$ is any vector in Hilbert space, $\Delta A^2 = \langle\psi|(\hat{A} - \langle\hat{A}\rangle)^2|\psi\rangle$, and $|\psi_{\perp}\rangle$ is a vector (state) in the perpendicular Hilbert space so that $\langle\psi|\psi_{\perp}\rangle = 0$.

Proof. Left multiplication by $\langle\psi|$ yields the first term; evaluating $|(A - \langle A\rangle)|\psi\rangle|^2 = \Delta A^2$ yields the second. \square

So far, this is a completely general geometric property. To actually make a measurement of an observable \hat{A} , we switch on an interaction [16] $H_{\text{int}} = \lambda g(t)\hat{q}\hat{A}$ with a normalized time profile $\int g(t)dt = 1$. The pointer, namely the momentum \hat{p}_q conjugate to \hat{q} , shifts by $\lambda\langle\hat{A}\rangle$.

Now, the average of any operator $\langle\hat{A}\rangle \equiv \langle\Psi|\hat{A}|\Psi\rangle$ that appears in theorem 3 can be measured in three distinct ways [33, 34]:

1. *Statistical method with disturbance.* The traditional approach is to perform ideal measurements of \hat{A} on each particle, obtaining a variety of different eigenvalues, and then manually calculate the usual statistical average to obtain $\langle\hat{A}\rangle$.
2. *Statistical method without disturbance.* The interaction $H_{\text{int}} = -\lambda(t)\hat{q}\hat{A}$ is weakened by minimizing $\lambda\Delta q$. For simplicity, we consider $\lambda \ll 1$ (assuming without lack of generality that the state of the measuring device is a Gaussian with spreads $\Delta p_q = \Delta q = 1$). We may then set $e^{-i\lambda\hat{q}\hat{A}} \approx 1 - i\lambda\hat{q}\hat{A}$ and use theorem 3 to show that the system state is

$$\begin{aligned} e^{-i\lambda\hat{q}\hat{A}}|\Psi_{\text{in}}\rangle &= (1 - i\lambda\hat{q}\hat{A})|\Psi_{\text{in}}\rangle \\ &= (1 - i\lambda\hat{q}\langle\hat{A}\rangle)|\Psi_{\text{in}}\rangle - i\lambda\hat{q}\Delta\hat{A}|\Psi_{\text{in}\perp}\rangle. \end{aligned} \quad (7)$$

Using the norm of this state $\|(1 - i\lambda\hat{q}\hat{A})|\Psi_{\text{in}}\rangle\|^2 = 1 + \lambda^2\hat{q}^2\langle\hat{A}^2\rangle$, the probability to leave $|\Psi_{\text{in}}\rangle$ unchanged after the measurement is

$$\frac{1 + \lambda^2\hat{q}^2\langle\hat{A}\rangle^2}{1 + \lambda^2\hat{q}^2\langle\hat{A}^2\rangle} \longrightarrow 1 \quad (\lambda \rightarrow 0), \quad (8)$$

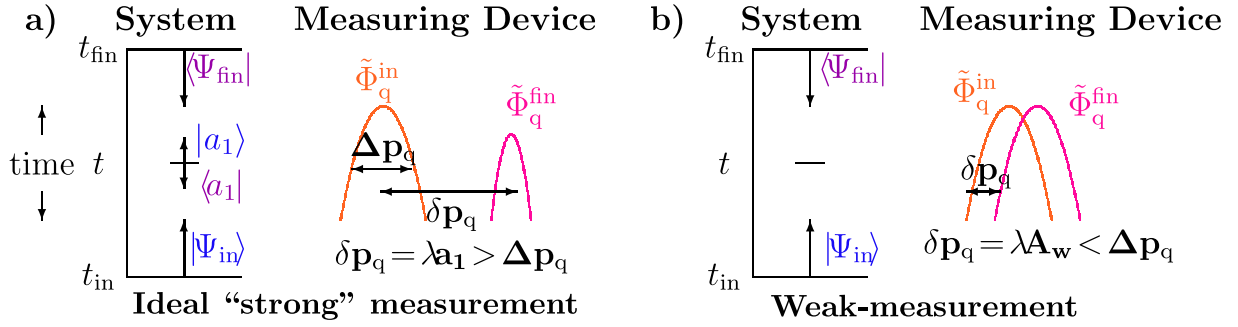


Figure 4. a) With an ideal or ‘strong’ measurement at t (characterized e.g. by $\delta P_{\text{md}} = \lambda a_1 \gg \Delta P_{\text{md}}$), ABL then gives the probability to obtain a collapse on to eigenstate a_1 by propagating $\langle \Psi_{\text{fin}} |$ backwards in time from t_{fin} to t and $| \Psi_{\text{in}} \rangle$ forwards in time from t_{in} to t ; in addition, the collapse caused by ideal measurement at t creates a new boundary condition $| a_1 \rangle \langle a_1 |$ at time $t \in [t_{\text{in}}, t_{\text{fin}}]$. b) If a weak-measurement is performed at t (characterized e.g. by $\delta P_{\text{md}} = \lambda A_w \ll \Delta P_{\text{md}}$), then the outcome of the weak-measurement, the weak-value, can be calculated by propagating the state $\langle \Psi_{\text{fin}} |$ backwards in time from t_{fin} to t and the state $| \Psi_{\text{in}} \rangle$ forwards in time from t_{in} to t ; the weak-measurement does not cause a collapse and thus no new boundary condition is created at time t .

while the probability to disturb the state (i.e. to obtain $| \Psi_{\text{in}\perp} \rangle$) is

$$\frac{\lambda^2 \hat{q}^2 \Delta \hat{A}^2}{1 + \lambda^2 \hat{q}^2 \langle \hat{A}^2 \rangle} \rightarrow 0 \quad (\lambda \rightarrow 0). \quad (9)$$

The final state of the measuring device is now a superposition of many substantially overlapping Gaussians with probability distribution given by

$$\text{Pr}(p_q) = \sum_i |\langle a_i | \Psi_{\text{in}} \rangle|^2 \exp \left\{ -\frac{(p_q - \lambda a_i)^2}{2 \Delta p_q^2} \right\}.$$

This sum is a Gaussian mixture, so it can be approximated by a single Gaussian

$$\tilde{\Phi}_{\text{md}}^{\text{fin}}(p_q) \approx \langle p_q | e^{-i \lambda \hat{q} \langle \hat{A} \rangle} | \Phi_{\text{md}}^{\text{in}} \rangle \approx \exp \left\{ -\frac{(p_q - \lambda \langle \hat{A} \rangle)^2}{\Delta p_q^2} \right\}$$

centered on $\lambda \langle \hat{A} \rangle$.

It follows from equation (9) that the probability for a collapse decreases as $O(\lambda^2)$, but the measuring device’s shift grows linearly $O(\lambda)$ with $\delta p_q = \lambda a_i$ [34]. That is, for a sufficiently weak interaction (e.g. $\lambda \ll 1$), the probability for a collapse can be made arbitrarily small, while the measurement still yields information. However, the measurement is less precise than the first case because the shift in the measuring device is much smaller than its uncertainty $\delta p_q \ll \Delta p_q$ (see figure 4).

3. *Non-statistical method without disturbance.* This is the case where $\langle \Psi | \hat{A} | \Psi \rangle$ is the ‘eigenvalue’ of a single ‘collective operator’, $\hat{A}^{(N)} \equiv \frac{1}{N} \sum_{i=1}^N \hat{A}_i$ (with \hat{A}_i being the same operator \hat{A} acting on the i th particle). Using this, we are able to obtain information

about $\langle \Psi | \hat{A} | \Psi \rangle$ without causing disturbance (or a collapse) and without using a statistical approach because any product state $|\Psi^{(N)}\rangle$ becomes an eigenstate of the operator $\hat{A}^{(N)}$. To see this, we apply theorem 3 to the N particle product state $|\Psi^{(N)}\rangle = |\psi\rangle_1 |\psi\rangle_2 \cdots |\psi\rangle_N$ with all particles in the same state $|\psi\rangle$. We see that

$$\hat{A}^{(N)} |\Psi^{(N)}\rangle = \frac{1}{N} \left[N \langle \hat{A} \rangle |\Psi^{(N)}\rangle + \Delta A \sum_i |\Psi_{\perp}^{(N)}(i)\rangle \right], \quad (10)$$

where $\langle \hat{A} \rangle$ is the average for any one particle and the N states $|\Psi_{\perp}^{(N)}(i)\rangle = |\psi\rangle_1 |\psi\rangle_2 \cdots |\psi_{\perp}\rangle_i \cdots |\psi\rangle_N$ are mutually orthogonal. With a normalized state, $|\Psi_{\perp}^{(N)}\rangle = \sum_i \frac{1}{\sqrt{N}} |\Psi_{\perp}^{(N)}(i)\rangle$, the last term of equation (10) is $\frac{\Delta A}{\sqrt{N}} |\Psi_{\perp}^{(N)}\rangle$ and $|\frac{\Delta A}{\sqrt{N}} |\Psi_{\perp}^{(N)}\rangle|^2 \propto \frac{1}{N}$. The probability that measuring \hat{A}_i/N changes the state of the i th system is proportional to $1/N^2$ and therefore the probability that it changes the state of any system is proportional to $1/N$. Thus, as $N \rightarrow \infty$, $|\Psi^{(N)}\rangle$ becomes an eigenstate of $\hat{A}^{(N)}$ with value $\langle \hat{A} \rangle$ and not even a single particle has been disturbed (as $\hat{N} \rightarrow \infty$) [11].

To perform an actual measurement in this case, we fix Δp_q (the width of the initial pointer momentum distribution) to be 1. We can then take $\lambda \gg 1$, in order to distinguish the shift, $\lambda \langle \hat{A} \rangle$ from the width. In addition, fixing $\lambda \ll \sqrt{N}$ along with $|\hat{A}_i| < 1$ ensures that the measurement does not shift any particle into an orthogonal state. The coupling to any individual member of the ensemble is reduced by $1/N$. When N is very large, the coupling to individual systems is very weak, and in the limit $N \rightarrow \infty$, the coupling approaches zero. Although the probability that a measurement will disturb any member of the ensemble approaches zero as $1/N$, information about the average is obtained.

4.2. Pre-selection, post-selection and weak measurements

By adding a post-selection to these ordinary—yet weakened—von Neumann measurements, the measuring device will register a weak value [12],

$$\hat{A}_w = \frac{\langle \Psi_{\text{fin}} | \hat{A} | \Psi_{\text{in}} \rangle}{\langle \Psi_{\text{fin}} | \Psi_{\text{in}} \rangle}, \quad (11)$$

with $|\Psi_{\text{in}}\rangle$ and $|\Psi_{\text{fin}}\rangle$ being the initial and final (post-selected) states. The weak value, A_w , is an unusual quantity and is not, in general, an eigenvalue of \hat{A} . We have used such limited disturbance measurements to explore many paradoxes (see, e.g. [15, 27, 35]). A number of experiments have been performed to test the predictions made by weak measurements and results have proven to be in very good agreement with theoretical predictions [28]–[32].

Equation (11) can also be motivated by inserting a complete set of states $\{|\Psi_{\text{fin}}\rangle_j\}$ into $\langle \hat{A} \rangle$

$$\langle \hat{A} \rangle = \langle \Psi_{\text{in}} | \hat{A} | \Psi_{\text{in}} \rangle = \sum_j |\langle \Psi_{\text{fin}} | \Psi_{\text{in}} \rangle_j|^2 \underbrace{\frac{\langle \Psi_{\text{fin}} | \hat{A} | \Psi_{\text{in}} \rangle_j}{\langle \Psi_{\text{fin}} | \Psi_{\text{in}} \rangle_j}}_{A_w^j \equiv \text{weak value}}, \quad (12)$$

with $|\Psi_{\text{fin}}\rangle_j$ being the states corresponding to the outcome of a final ideal measurement on the system (i.e. the post-selection). The average $\langle \hat{A} \rangle$ over all post-selections j is thus constructed out of pre- and post-selected sub-ensembles in which the weak value (A_w^j) is multiplied by a probability to obtain the particular post-selection $|\Psi_{\text{fin}}\rangle_j$.

To see more precisely how the weak value arises naturally from this weakened measurement interaction with post-selection, we consider here the final state of the measuring device in the third ‘non-statistical’ method:

$$\begin{aligned} |\Phi_{\text{fin}}^{\text{MD}}\rangle &= \prod_{j=1}^N \langle \Psi_{\text{fin}} |_j \exp\left(\frac{-i\lambda}{N} \hat{q} \sum_{k=1}^N \hat{A}_k\right) \prod_{i=1}^N |\Psi_{\text{in}}\rangle_i |\Phi_{\text{in}}^{\text{MD}}\rangle \\ &= \prod_{j=1}^N \langle \Psi_{\text{fin}} |_j \exp\left(\frac{-i\lambda}{N} \hat{q} \hat{A}_j\right) |\Psi_{\text{in}}\rangle_j |\Phi_{\text{in}}^{\text{MD}}\rangle. \end{aligned} \quad (13)$$

Since the particles do not interact with each other, we calculate one term and take the result to the N th power. (In the following, we substitute the parity operator, \hat{P} , for \hat{A} .) Using $\hat{P}^2 = 1$, equation (13) becomes

$$\begin{aligned} |\Phi_{\text{fin}}^{\text{MD}}\rangle &= \left\{ \langle \Psi_{\text{fin}} | \left(\cos \frac{\lambda \hat{q}}{N} - i \hat{P} \sin \frac{\lambda \hat{q}}{N} \right) | \Psi_{\text{in}} \rangle \right\}^N |\Phi_{\text{in}}^{\text{MD}}\rangle \\ &= (\langle \Psi_{\text{fin}} | \Psi_{\text{in}} \rangle)^N \left\{ \cos \frac{\lambda \hat{q}}{N} - i P_w \sin \frac{\lambda \hat{q}}{N} \right\}^N |\Phi_{\text{in}}^{\text{MD}}\rangle, \end{aligned} \quad (14)$$

$$\approx \left\{ 1 - i P_w \frac{\lambda \hat{q}}{N} + \dots \right\}^N |\Phi_{\text{in}}^{\text{MD}}\rangle \approx \exp(-i \lambda \hat{q} P_w) |\Phi_{\text{in}}^{\text{MD}}\rangle. \quad (15)$$

The first bracket of equation (14) can be neglected since it does not depend on \hat{q} and thus can only affect the normalization. Equation (15) represents a shift in the pointer by the weak value, P_w , i.e. $\Phi_{\text{in}}^{\text{MD}}(p_q) \rightarrow \Phi_{\text{fin}}^{\text{MD}}(p_q - \lambda P_w)$.

4.3. Applying pre- and post-selection and weak measurements to the interference phenomena

How can we use these tools to perform measurements of dynamical non-locality? Briefly, we start with particles sent through the right slit. Before they encounter the double slit, we perform a weak measurement of the modular momentum (which, again, is sensitive to the relative phase). We then choose whether to open the left slit or to close it. After the particles pass the double-slit set-up, we perform an ideal measurement of the modular momentum and post-select only those particles in a particular eigenstate of this modular momentum. When we analyze the earlier weak measurement (assuming the post-selection is satisfied), we see two dramatically different results, one result if the left slit is closed and a very different result if the left slit is open.

In this section, we elaborate on the details by using the third ‘non-statistical’ method and will later discuss the use of the second ‘statistical’ method. Consider the following sequences.

- We send toward the slits, N consecutive particles, each in the same state centered around the right slit, Ψ_R , i.e. we pre-select $|\Psi_{\text{in}}\rangle = |\psi_R\rangle$ rather than $|\Psi_{\alpha}\rangle$.
- After the pre-selection, but prior to encountering the slits, we measure weakly the average modular variable, $\hat{P}^{(N)}$ (the parity) yielding an outcome of $\cos \alpha$. In order to perform this measurement, we utilize (following von Neumann) the interaction Hamiltonian $H_{\text{int}} = \frac{1}{N} \sum_{i=1}^N \lambda g(t) \hat{q} \hat{P}_i$, thereby generating the evolution $\prod_{i=1}^N \exp\{-i \lambda \hat{q} \hat{P}_i\}$, which simply sums the displacements of the ‘pointer’ due to the interactions with each of the N particles, namely a shift proportional to $\hat{P}^{(N)}$.

- (c) Finally, we post-select an eigenstate of the same modular variable observable previously measured weakly. In particular, we post-select the symmetric state: $|\Psi_{\text{fin}}\rangle = \frac{1}{\sqrt{2}}\{|\psi_L\rangle + |\psi_R\rangle\}$, which we note is an eigenstate of the corresponding parity operator \hat{P} with eigenvalue +1.

What, then, are the results of the above weak measurement, (i) when both slits are open and (ii) when the left slit is closed?

Case 1. With the left slit open, parity is conserved since in this symmetric slit arrangement, the Hamiltonian commutes with the parity operator. (Furthermore, as noted in section 3.2 by equation (4), $(d/dt)e^{(i/\hbar)\hat{p}D} = 0$ when the left slit is open because $V(x) - V(x + D) = 0$.) This can also be seen by evolving the post-selected state backward in time. This yields $\hat{P}_i = +1$ for each of the N particles (both before and after the double slit) and the measuring device then registers the weak value $\mathbf{P}_w^{(N)} = 1$. More specifically, the wavefunction of the measuring device evolves as

$$\begin{aligned}\Phi_{\text{fin}}^{\text{MD}}(p_q) &\approx \left\{ \{ \langle \psi_L | + \langle \psi_R | \} \exp \left\{ -i \frac{\lambda}{N} \hat{q} \mathbf{P}_w \right\} | \psi_R \rangle \right\}^N \Phi_{\text{in}}^{\text{MD}}(p_q) \\ &= \left[\exp \left\{ -i \frac{\lambda}{N} \hat{q} \right\} \right]^N \Phi_{\text{in}}^{\text{MD}}(p_q) = \Phi_{\text{in}}^{\text{MD}}(p_q - \lambda).\end{aligned}\quad (16)$$

Case 2. With the left slit closed, the results of the weak measurement described above are *drastically* changed. Parity is now maximally violated and there is no connection between the +1 parity of the post-selected state and the results of the weak parity measurements performed prior to entering the slits. (Again, as noted by equation (4) in section 3.2, $(d/dt)e^{(i/\hbar)\hat{p}D} \neq 0$ when the left slit is closed, because $V(x) - V(x + D) \neq 0$.)

We next show that with this second ‘slit-closed’ case, the weak value of the parity is centered around 0. Only $|\psi_R\rangle$ can now propagate through the system of slits (any component of $|\psi_L\rangle$ generated by the weak measurement is always reflected by the closed slit¹¹). The pointer shift is given by equation (14)

$$\Phi_{\text{fin}}^{\text{MD}}(p_q) = \left\{ \langle \psi_R | \cos \frac{\lambda \hat{q}}{N} - i \hat{P} \sin \frac{\lambda \hat{q}}{N} | \psi_R \rangle \right\}^N \Phi_{\text{in}}^{\text{MD}}(p_q).\quad (17)$$

Using $\langle \psi_R | \hat{P} | \psi_R \rangle = \langle \psi_R | \psi_L \rangle = 0$, only the cosine part remains. The pointer state $\Phi_{\text{in}}^{\text{MD}}(p_q)$ then shifts by

$$\Phi_{\text{fin}}^{\text{MD}}(p_q) = \left\{ \frac{1}{2} \right\}^N \left\{ e^{(i\lambda/N)\hat{q}} + e^{-(i\lambda/N)\hat{q}} \right\}^N \Phi_{\text{in}}^{\text{MD}}(p_q),\quad (18)$$

which upon binomial expansion becomes

$$\Phi_{\text{fin}}^{\text{MD}}(p_q) = \left\{ \frac{1}{2} \right\}^N \sum_k \binom{N}{k} \Phi_{\text{in}}^{\text{MD}} \left(p_q + \frac{\lambda(2k - N)}{N} \right).\quad (19)$$

Since the binomial coefficients $\binom{N}{k}$ peak around $k = N/2$, the effect of the shifts vanishes in the $N \rightarrow \infty$ limit, and $\langle \mathbf{P}^{(N)} \rangle_w = 0$ as claimed.

All incoming particles are initially in the state $|\psi_R\rangle$, so we would not expect that closing the left slit should have any effect on the result of any weak measurement (and in particular weak

¹¹ Unitarity requires a reflected state $|\psi_3\rangle$. Unlike $|\psi_L\rangle$ and $|\psi_R\rangle$, moving in the +y-direction, $|\psi_3\rangle$ moves in the -y-direction. Since $\langle \psi_3 | \psi_L \rangle = 0$, omitting it does not affect the argument.

measurements performed prior to the opening or closing of the slits). Nevertheless, the *earlier* weak measurement of the parity yields $\langle \hat{P}^{(N)} \rangle_w = 0$ (case 2) if the left slit is (*later*) closed and $\langle \hat{P}^{(N)} \rangle_w = 1$ (case 1) if the left slit is (*later*) opened.

5. Discussion

How do we understand these two results? In principle, a weak measurement with finite N shifts particles from the right slit to the left slit so that the evolving wavepacket has $|\psi_L\rangle$ components and therefore may ‘sense’ whether the left slit is open or closed. However, all modular operators and parity in particular have norms ≤ 1 . The exponents in the von Neumann interaction Hamiltonian are thus bound by $\frac{\lambda}{N} \hat{P}^{(N)} \hat{q} < \frac{\lambda}{N} \hat{q}$ and, hence, it suffices to expand the binomial to the order of a few λ . This implies then that the weak measurement can shift at most a few λ of the N particles from the right slit to the left slit. But how can the $N - \lambda$ particles that were not shifted and did not go through the left slit still be influenced non-locally so that we will have the dramatic (and large) change from case 1 (*each* particle shifts the measuring device by $\langle \hat{P} \rangle_w = 1$) to case 2 (each particle shifts the measuring device by $\langle \hat{P} \rangle_w = 0$)?

We do not see any reasonable way to use local interactions at the left-hand slit to account for the different subsequent behaviors of the particles going through the right slit.

We can, however, make sense of the results by considering the non-local behavior of modular variables. In particular, the results are calculated using the non-local exchange of modular momentum. The ‘open-slit’ case is calculated by using *conservation* of modular momentum, which follows from this non-local exchange. The use of this conservation principle is one of the crucial features that distinguishes our procedure from observations that are done with ordinary momentum.

These issues are not just ‘academic’, as this article sets the stage for a forthcoming paper [20] describing an actual quantum optics experiment to measure the non-local exchange. For illustrative purposes, we mention an experimentally simpler example using the second method of section 4.1, i.e. the statistical weak measurement; consider two consecutive Mach–Zehnder interferometers (see figure 5). The first Mach–Zehnder prepares the pre-selection; by adjusting the arm lengths, it is possible to arrange that the photon emerges at R_4 , which corresponds to the pre-selection $|\psi_R\rangle$ (localized at the right slit).¹² In addition, the weak measurement of parity is performed within the first Mach–Zehnder by measuring small transverse shifts in the position of the photon produced by inserting thin glass plates [21] on to both the R_2 and L_2 arm. The regime of weak measurement is obtained by adjusting the tilt of the plates so that the transverse spatial shift is small compared to the uncertainty in the transverse position of the photon. The second Mach–Zehnder is the analogue of the double-slit; e.g. blocking the L_4 path corresponds to closing the left-slit of the double-slit set-up.

Note that when the particle passed through the two-slits (after BS_2), parity was a non-local operator. Later in time (after BS_3), the previously non-local parity was converted into a local quantity. Because of this, we are able to perform an ideal measurement—i.e. a post-selection—of the parity.

¹² We label a left-pointing arm as $|L_n\rangle$, where a subscript 1 refers to the photon before entering BS_1 ; a 2 refers to the photon after entering BS_1 and before M_1 (etc); similarly $|R_n\rangle$ is used for a right-pointing arm. When put into the MZI in the right arm (without any weak measurement), the photon will exclusively exit at R_4 . Specifically, $|R_1\rangle \xrightarrow{BS_1} \frac{1}{\sqrt{2}}\{i|L_2\rangle + |R_2\rangle\} \xrightarrow{M_1/M_2} \frac{1}{\sqrt{2}}\{i|L_3\rangle - |R_3\rangle\} \xrightarrow{BS_2} \frac{1}{2}\{i|L_4\rangle - |R_4\rangle\} - \frac{1}{2}\{|R_4\rangle + i|L_4\rangle\} = -|R_4\rangle$.

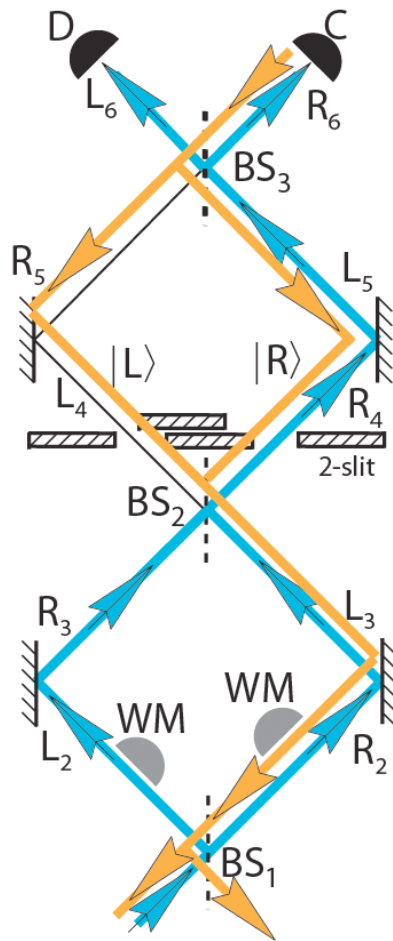


Figure 5. The Mach-Zehnder analogue of the two-slit set-up. Also depicted is the preselected wavefunction (evolving forward in time in blue) and the post-selected wavefunction (evolving backward in time in orange).

If we post-select the +1 parity (after the photon passes BS_3) and if L_4 is open, then the earlier weak measurement of parity will register +1 (meaning that the weak value of the number of particles within arm R_2 is $(N_{R_2})_w = 1$ and within arm L_2 is $(N_{L_2})_w = 0$). However, if we now close the left-hand slit, i.e. semi-arm L_4 , then the earlier weak measurement will register $(N_{R_2})_w = (N_{L_2})_w = \frac{1}{2}$ even though no particles took the path where the slit was closed!¹³

In this ‘statistical’ weak measurement method, although each particle shifts its associated measuring device by the calculated weak value, for any individual measuring device the shift is less than the uncertainty. To register the shift, a statistical calculation is necessary.

¹³ Again, the relation with modular variables is somewhat clearer if we consider many slits since modular momentum is exactly conserved and we can directly speak about modular momentum rather than parity in the two-slit case. Suppose that we send particles that are localized around a single slit toward the infinite slits and we then perform a weak measurement of $\sum \exp\{(i/\hbar)\hat{p}D\}$. We then consider two situations: either all the other slits are open or all the other slits are closed. In those cases where the post-selection yields an eigenstate of modular position, and if all the slits are open, then we see that modular momentum is conserved and the earlier weak measurement will equal the post-selected ideal measurement. If the other slits are closed, then modular momentum is not conserved, and the non-local equations of motion will reflect this.

The third ‘non-statistical’ method discussed in section 4.1 *sharply* measures the ‘collective operator’ $\hat{A}^{(N)} \equiv \frac{1}{N} \sum_{i=1}^N \hat{A}_i$, with each \hat{A}_i operating on a state $|\psi\rangle_i$ in a product state $|\Psi^{(N)}\rangle = |\psi\rangle_1 |\psi\rangle_2 \dots |\psi\rangle_N$ of N non-interacting particles. The measurement corresponds to summing the weak values for the ensemble of N particles, each of which goes through the same procedure in encountering the two-slits. We also conceive of situations where the collective operator $\hat{A}^{(N)}$ is directly measured by having the pointer of the measuring apparatus integrate the small shifts due to all repeated interactions. A case in point corresponds to photons exerting pressure on a mirror.

This third ‘non-statistical’ method is extremely important for many reasons. For example:

- The physics is clearer in this case; although rare¹⁴, the measuring device registers a large shift in a single trial ($\sim N$, the calculated weak value), which is much larger than the uncertainty. The number of particles that get shifted to the left slit due to the weak measurement is λ —independent of N . Furthermore, we can make λ as small as we like. We only require $\lambda > \frac{1}{\sqrt{N}}$ in order to obtain a measurable effect. The effect is proportional to N , i.e. all the particles know whether the left slit is open or closed, contrary to what would be required in an attempt to argue that the effect is due to a local interaction at the left slit.
- Another reason (to be elaborated in a future article) is that the collective observable leads to a broad new class of measurable Hermitian observables involving highly non-local properties (in both space and time). For example, instead of looking only at the special case of parity, we may go back to the spin-like observables with $\hat{\sigma}^z$ telling us about which slit the particle is centered around, etc. The non-local relative phase can then be measured with

$$\alpha \sim \frac{\frac{1}{N} \sum_{i=1}^N \hat{\sigma}_i^x}{\sqrt{1 - (\frac{1}{N} \sum_{i=1}^N \hat{\sigma}_i^x)^2}}. \quad (20)$$

We believe that this can only be measured using the technique of weak measurement involving the collective observable.

One final point about the second method concerning analysis of changes in probability distributions discussed in section 2.2 (i.e. using the Fourier transform of the conserved quantity). It has many advantages when compared to the traditional analysis with respect to moments discussed in section 2.1. For example, the Fourier transform method provides us with the parameters relevant to the physical problem (e.g. the distance D between the slits), whereas the first ‘moments’ approach remains silent. (We note that, in effect, looking at the modular variables is asking how the Fourier transform of the momentum distribution changes.) In addition, there are different conservation laws involved with the second method, which are more relevant and useful. One of the basic notions used in the analysis of conserved quantities in any interaction is that as the probability of one conserved quantity changes ($\text{prob}(A)$), the probability of another should also change ($\text{prob}(B)$), such that the probability of the sum ($\text{prob}(A + B)$) does not change. As we pointed out, there are situations where the probability of one variable does not change ($\text{prob}(B)$), while the probability of the other does change ($\text{prob}(A)$). This, as per theorem 2, can only happen if that variable (e.g. B) is completely uncertain. That is, this can only happen if the Fourier transform of $\text{prob}(B)$ below some value remains unaffected while the Fourier transform of $\text{prob}(B)$ above some value is affected. This means that there is a whole range of modular variables that are being exchanged non-locally and a large number of conservation laws that can be utilized.

¹⁴ In [33], a new weak measurement procedure was introduced for finite samples which yields accurate weak values that are outside the range of eigenvalues and which do not require an exponentially rare ensemble.

6. Conclusion

The purpose of this self-contained article is to highlight some new aspects of interference using pre- and post-selection and weak measurements. We have emphasized the importance of analyzing quantum mechanical interference through the use of states which are more ‘non-classical,’ namely with states comprised of multiple lumps. We argued that the most appropriate picture for these situations is given by modular variables. Exciting new features of quantum mechanics were obtained using this approach. For example, applying the Heisenberg picture to multiple-lump states lead us to a physical explanation for the different behaviors of a *single* particle when the distant slit is open or closed; this helps us to shift from a picture involving a quantum wave that passes through all slits, to one with more ‘localized’ particles which interact *non-locally* with the other slit(s). This contrasts with classical mechanics where the equations of motion are local. Furthermore, we showed that although particles localized around the right slit can exchange modular momentum (non-locally) with the ‘barrier’ at the left slit, the uncertainty in quantum mechanics appears to protect causality. In other words, under exactly the conditions where the non-local exchange of modular variables *could potentially* violate causality, the variable exchanged non-locally becomes completely uncertain and therefore unobservable, thereby protecting causality. Nevertheless, we showed that the non-local equations can have a measureable effect without violating causality. We showed this with the following calculation. If the left slit is later closed, then the probability that the earlier weak measurement shifts any particles from the right slit to the left slit is $O(\frac{1}{N})$. Therefore, in the limit of large N , the weak measurement does not shift even a single particle from the right slit to the left slit. This can be confirmed by placing a photographic plate at the left slit. On the other hand, if the left slit is later opened (i.e. after the weak measurement), then we calculate that a small number of particles (independent of N), are shifted from the right slit to the left slit. However, all N particles contribute to the dramatically different weak measurement results. This is explained by the non-local equations of motion satisfied by the modular variables¹⁵. Finally, we note that while the Heisenberg and the Schrödinger pictures are equivalent formulations of quantum mechanics, nevertheless, the results discussed here support a new approach which has led to new insights, new intuitions, new experiments, and even the possibility of new devices that were missing from the old perspective. These types of development are indicative of a successful re-formulation.

Acknowledgments

We thank J Gray, A D Parks, S Spence and J Troupe.

Appendix A. Proof that the moments do not depend on the relative phase

A basic theorem that characterizes all interference phenomena is that all moments of both position and momentum are independent of the relative phase parameter α . It is easy to see this for the particular double-slit wavefunction equation (1); assuming that there is no overlap of $\psi_L(x, 0)$ and $\psi_R(x, 0)$ and that n is an integer, then for all values of t and choices of α, β :

$$\int [\Psi_\alpha^*(x, t)\Psi_\alpha(x, t) - \Psi_\beta^*(x, t)\Psi_\beta(x, t)]x^n dx = 0. \quad (\text{A.1})$$

¹⁵ These observables, $p \bmod \frac{\hbar}{D}$, do not have a classical limit: e.g. when D is kept fixed and $\hbar \rightarrow 0$, then these operators oscillate infinitely fast. Therefore, although these operators continue to obey non-local equations of motion in the classical limit, the non-locality loses its observable meaning in the classical limit.

We see that $\Psi_\alpha^*(x, 0)\Psi_\alpha(x, 0)$ is independent of α , and hence $\Psi_\alpha^*(x, 0)\Psi_\alpha(x, 0) - \Psi_\beta^*(x, 0)\Psi_\beta(x, 0) = 0$. Therefore, at $t = 0$, $\langle x^n \rangle$ is independent of α , as is $\langle p^n \rangle$. The latter follows from the non-overlapping nature of $\psi_L(x, 0)$ and $\psi_R(x, 0)$. It is also easy to show that $\langle x^n p^m \rangle$ at $t = 0$ is independent of α by using the Heisenberg representation $\langle x^n(t) \rangle = \int \Psi_\alpha^*(x, 0)x^n(t)\Psi_\alpha(x, 0)dx$; noting that $x(t) = x(0) + p(0)(t/m)$ and $p(t) = p(0)$ in this representation, we must have $\langle x^n(t) \rangle = \int \Psi_\alpha^*(x, 0)[x(0) + p(0)(t/m)]^n(t)\Psi_\alpha(x, 0)dx$. This is clearly independent of α , since term by term it is independent of α . Equation (A.1) then follows, and holds for p^n , as long as we retain the proper $\Psi_\alpha^*p^n\Psi_\alpha$ order.

Appendix B. Non-locality in configuration space

The ‘configuration’ space states are those of a particle at N discrete locations: $|\Psi(1)\rangle$ at slit 1, $|\Psi(2)\rangle$ at slit two, etc, and the N discrete modular momentum eigenstates are the appropriate linear combinations:

$$\begin{aligned} |\chi(1)\rangle &= \frac{1}{\sqrt{N}}[|\Psi(1)\rangle + |\Psi(2)\rangle + \dots + |\Psi(N)\rangle], \\ |\chi(2)\rangle &= \frac{1}{\sqrt{N}}[|\Psi(1)\rangle + b|\Psi(2)\rangle + \dots + b^{N-1}|\Psi(N)\rangle], \\ |\chi(3)\rangle &= \frac{1}{\sqrt{N}}[|\Psi(1)\rangle + b^2|\Psi(2)\rangle + \dots + b^{2(N-1)}|\Psi(N)\rangle], \\ &\vdots \\ |\chi(n+1)\rangle &= \frac{1}{\sqrt{N}}[|\Psi(1)\rangle + \dots + b^{n(N-1)}|\Psi(N)\rangle], \\ &\vdots \end{aligned} \tag{B.1}$$

where $b = \exp(-i2\pi/N)$, and each of the $|\chi(k)\rangle$ is an eigenstate of the cyclic shift operator $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots \rightarrow N \rightarrow 1$, namely the relevant modular operator with eigenvalues b^{k-1} . The inverse of the above relates each of the configuration eigenstates to an *equal* weight combination of the $|\chi(k)\rangle$ states, which again is a state with maximal angular momentum uncertainty. (This in turn is an of the Dirac $\delta(x)$ function being an equal weight superposition of all regular continuous momenta p , for the discrete Kronecker $\delta(j)$ in the present case.)

Appendix C. The non-locality of modular variables is generic to all interference phenomena

Although much of the discussion in this article focuses on the simplest interference example with two slits, our approach becomes clearer when it is applied to an infinite number of slits. For example, the two-slit set-up involves slightly more complicated functions of the modular momentum while an infinite number of slits is characterized very simply in terms of modular momentum.

Furthermore, in the infinite slit case, both the non-local equation of motion and the conservation laws for modular momentum are exactly satisfied. To (briefly) see how this

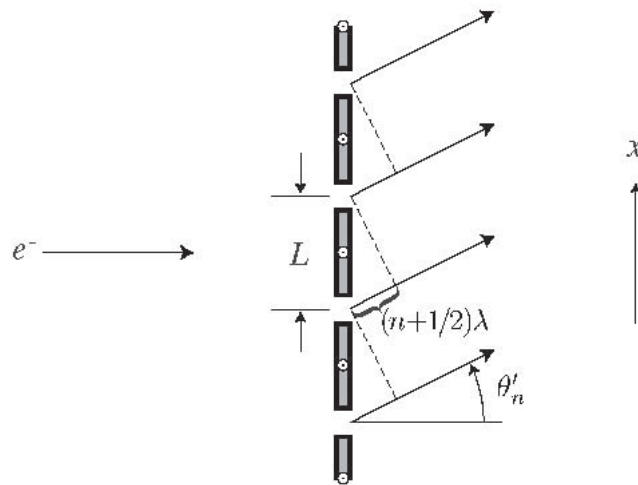


Figure C.1. A grating and a lattice of solenoids with electrons diffracting through them. If the length of the path of the electron emerging from different slits differs by $(n + \frac{1}{2})\lambda$, then the electrons will interfere constructively.

works, consider a system of infinite slits [10] that is freely moving in the x -direction (see figure C.1). Suppose electrons are sent towards the slits in a momentum eigenstate $p_y \equiv p_o$ and $p_x = 0$. After passing through the slits, one can prove that the transverse momentum of the electrons is $p_x = n \frac{h}{D}$, with n an integer. Therefore, the particle and slit-grating can only exchange transverse momentum in integer multiples of $\frac{h}{D}$. With a Hamiltonian $H = \frac{p^2}{2m} + V(x)$ (where $V(x + D) = V(x)$), one can also prove that $[H, e^{\frac{i}{\hbar} \hat{p} D}] = 0$, i.e. modular momentum is conserved.

As we mentioned earlier, we can observe the change in modular momentum by opening or closing other slits, or in close analogy, by performing some operation such as applying an uncertain potential. It is easier to see the non-locality of modular variables if we perform the later, as in an Aharonov-Bohm set-up. Suppose that we place solenoids with magnetic flux $\Phi = \frac{1}{2} \Phi_o$ inside the slit-gratings (see figure C.1) so that there is no contact between the electrons and fields generated by the magnetic flux inside the solenoids. Furthermore, suppose that we connect all the solenoids together so that they could move independent of the slit-gratings. One can prove that the condition for constructive interference is satisfied if the transverse momentum exchange is

$$\frac{h}{D} \left\{ n + \frac{\Phi}{\Phi_o} \right\} = \frac{h}{D} \left\{ n + \frac{1}{2} \right\}. \quad (\text{C.1})$$

We know that without the solenoids, only an integer multiple of $\frac{h}{D}$ can be exchanged between the electrons and the slit-grating. But equation (C.1) states that a non-integer value is exchanged. Therefore, using the conservation laws, we conclude that $\frac{h}{2D}$ must be exchanged non-locally between the magnetic flux shielded inside the solenoids and the electrons.

This non-locality seems to be paradoxical: classically there is no local interaction between the fields of the magnetic flux and the electrons. Therefore, solenoid-system should not accumulate momentum in such a way as to violate the correspondence principle. With a single electron there is no problem because the amount exchanged, $\frac{h}{2D}$, is less than the uncertainty.

To see this, consider that by definition, Δ (position of the solenoids) $\leq D$ and therefore Δ (momentum of solenoids) $\geq \frac{\hbar}{D}$. If we send a single electron, then the $\frac{\hbar}{2D}$ exchanged non-locally with the solenoids is less than the uncertainty $\frac{\hbar}{D}$ and is therefore unobservable. However, one might wonder what happens if we send many electrons. Each successive electron exchanges $\pm \frac{\hbar}{2D}$ non-locally and since there is no memory, one would expect that the total amount of momentum exchanged between the electrons and the solenoids should grow as $\sim \sqrt{N} \frac{\hbar}{2D}$ as occurs in a random walk. The paradox is resolved by the fact that it is modular momentum, not ordinary momentum, that is exchanged non-locally. Ordinary momentum accumulates like a random walk on a line. But modular momentum accumulates like a random walk on a *circle* and this of course can never grow more than the circumference of the circle. The solenoids do not change the average of the momentum or the average of the square of the momentum (etc). Rather, the solenoids change the *modular momentum* of the electrons.

Earlier, we asked ‘which operators become uncertain when WWM information is obtained?’ When the electron passes through the slit-gratings then we do not know through which slit it passed. We only know its position modulus the distance D between the slits. We can define modular position as $x_{\text{mod}} \equiv x \text{ modulus } D$. Ordinary position then can be expressed as: $x = N_x D + x_{\text{mod}}$. The grating is providing us with information about both $x \text{ mod } D$ and also $p \text{ mod } \frac{\hbar}{D}$. These states, comprised of multiple lumps, are perhaps some of the most non-classical states. They are best described by the non-locality of the modular variables. This contrasts strongly with the more ‘classical’ states described by single lumps. To better understand the relationship between these 2 extremes, we derive the following relations:

$$[e^{\frac{i}{\hbar} p_{\text{mod}} D}, N_x] = e^{\frac{i}{\hbar} p_{\text{mod}} D} \quad (\text{C.2})$$

and

$$[e^{i2\pi x_{\text{mod}}/D}, N_p] = e^{i2\pi x_{\text{mod}}/D}. \quad (\text{C.3})$$

These two commutation relations are similar to the angular momentum–angle relation $[L_z, e^{i\phi}] = e^{i\phi}$: if the angular momentum is known precisely then the angle is completely uncertain. This leads us to think of uncertainty in phase space in a new way with p_{mod} conjugate to N_x and x_{mod} conjugate to N_p :

$$\Delta N_x \Delta p_{\text{mod}} \geq \frac{\hbar}{D}, \quad (\text{C.4})$$

which means, e.g. that if the system is localized in space, i.e. N_x is known, then the modular momentum is completely uncertain. Another uncertainty relation that can be derived is:

$$\Delta N_p \Delta x_{\text{mod}} \geq D, \quad (\text{C.5})$$

i.e. if the momentum is well known, then the modular position is uncertain.

The standard Heisenberg uncertainty principle does not allow us to localize a particle in phase space to anything less than an area of \hbar . The uncertainty principle for modular variables allows us to precisely locate a spot within a cell of area \hbar , i.e. x_{mod} and p_{mod} are certain, but it gives no information about which cell it is in, i.e. N_x and N_p are completely uncertain. This suggests that we can have precise but partial information about the momentum, namely the modular momentum, and simultaneously precise but partial information about the location, namely the modular position (see figure C.2(a)). We can also have intermediate situations, i.e. less precise information of the exact point within a phase cell but more information of which cell we’re in (see figure C.2(b)).

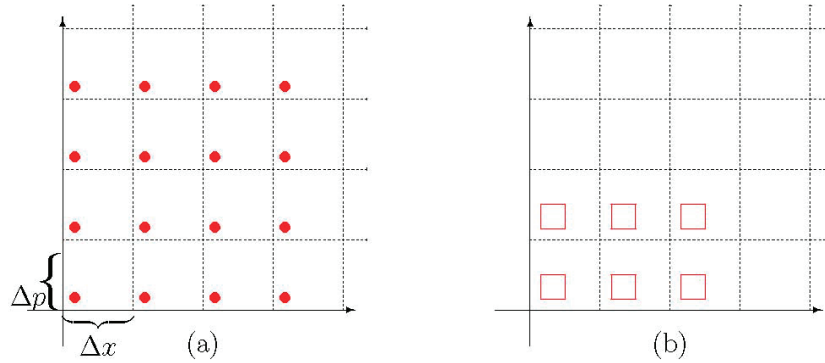


Figure C.2. (a) Phase space of particles with definite modular position, $x \bmod D = D/4$ and definite modular momentum $p \bmod h/D = \frac{1}{2}h/D$ and cell size $\Delta x \Delta p = \hbar$; (b) intermediate situation: more knowledge of which cell we are in means less knowledge of modular variable.

With single wavepackets, N_p and N_x are well known and therefore x_{mod} and p_{mod} are almost completely uncertain. However, as these uncertainty relations indicate, when we have wavefunctions with more than one ‘lump’, then x_{mod} and p_{mod} are more relevant: the electron passes the grating, so we have precise information about x_{mod} (in the transverse direction); x_{mod} and p_{mod} commute, so we have precise information about p_{mod} . But x_{mod} and N_p do not commute (see equations (C.4) and (C.5)) so we have no information about N_p . The interaction of the electron with the grating conserves p_{mod} so these facts determine the interference pattern completely: p_{mod} fixes the position of the fringes relative to the grating; N_p is completely uncertain and therefore the fringes are equally dense. Now consider the effect of a lattice of solenoids. The solenoids affect the modular momentum in the same way as the potential $V(x)$ of figure 3. The non-local interaction of the electrons with the solenoids changes p_{mod} of the diffracting electrons; hence the diffraction pattern shifts.

Appendix D. Conservation law for modular variables

Modular variables have different kinds of conservation laws that are enforced by the non-local equations of motion, and this will prove to be crucially important for this article. For the two-slit set-up, conservation of modular momentum is particularly easy. If we start with $|\psi_L\rangle + |\psi_R\rangle$, then we will end up with $|\psi_L\rangle + |\psi_R\rangle$. If we start with $|\psi_L\rangle - |\psi_R\rangle$, then we will end up with $|\psi_L\rangle - |\psi_R\rangle$. More generally, the modular momentum analogy to conservation of ordinary momentum (e.g. $P_1^{\text{in}} + P_2^{\text{in}} = P_1^{\text{fin}} + P_2^{\text{fin}}$) can be derived as follows. Using $\pi_1(P_1) = \cos(2\pi P_1/P_0)$ and $\pi_2(P_2) = \cos(2\pi P_2/P_0)$ (another expression for p_{mod}), we see that

$$\cos[2\pi\{P_1 + P_2\}/P_0] = \cos[2\pi\{P'_1 + P'_2\}/P_0]; \quad (\text{D.1})$$

in other words

$$\pi_1\pi_2 - \sqrt{1 - \pi_1^2}\sqrt{1 - \pi_2^2} = \pi'_1\pi'_2 - \sqrt{1 - (\pi'_1)^2}\sqrt{1 - (\pi'_2)^2}, \quad (\text{D.2})$$

which gives

$$\begin{aligned} & (\pi'_1)^2 + (\pi'_2)^2 - 2\cos[2\pi\{P_1 + P_2\}/P_0]\pi'_1\pi'_2 \\ & = 1 - \cos[2\pi\{P_1 + P_2\}/P_0]^2 \end{aligned} \quad (\text{D.3})$$

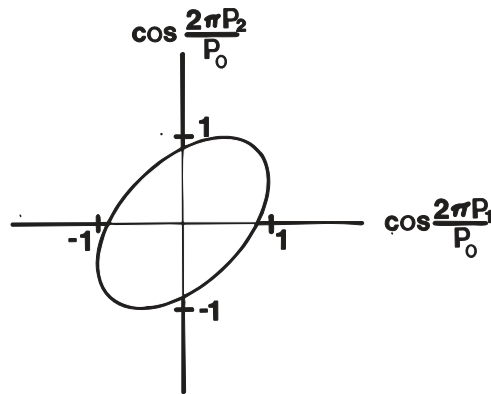


Figure D.1. Conservation law for modular momentum.

Thus, instead of a line $P_1 + P_2 = \text{constant}$, the conservation law for modular variables is an ellipse (see figure D.1).

If we know the initial values of the modular momentum of the two interacting systems, then we may represent their initial state by a point on the conserved ellipse of figure D.1. As the interaction between the two systems proceeds, the point representing the system will move along the ellipse and eventually come back to its original position. We see then how the periodicity of the non-local phenomena is reflected in the conservation laws for the relevant modular variables. We also note that in the classical limit, $p_o \rightarrow 0$, so that $p \bmod p_o$ changes so rapidly as a function of p as to become entirely unobservable.

References

- [1] Aharonov Y and Bohm D 1959 Significance of electromagnetic potentials in the quantum theory *Phys. Rev.* **115** 485
- [2] Aharonov Y, Bergmann P G and Lebowitz J L 1964 *Phys. Rev. B* **134** 1410
Reprinted in *Quantum Theory and Measurement* 1983 ed J A Wheeler and W H Zurek (Princeton, NJ: Princeton University Press) pp 680–6
Aharonov Y, Popescu S, Tollaksen J and Vaidman L 2007 arXiv:0712.0320
- [3] Feynman R, Leighton R and Sands M 1965 *The Feynman Lectures on Physics* vol III (Reading, MA: Addison-Wesley) pp 1–9
- [4] Popescu S and Rohrlich D 1992 Generic quantum nonlocality *Phys. Lett. A* **166** 293
- [5] Scully M O, Englert B G and Walther H 1991 *Nature* **351** 111
- [6] Storey E P, Tan S M, Collett M J and Walls D F 1994 *Nature* **367** 626
- [7] Wiseman H 2003 *Phys. Lett. A* **311** 285
- [8] Aharonov Y, Pendelton H and Petersen A 1969 *Int. J. Theor. Phys.* **2** 213
Aharonov Y, Pendelton H and Petersen A 1970 *Int. J. Theor. Phys.* **3** 443
- [9] Tollaksen J 2007 *J. Phys.: Conf. Ser.* **70** 012014
Aharonov Y and Kaufherr T 2004 *Phys. Rev. Lett.* **92** 070404
- [10] Aharonov R and Rohrlich 2005 *Quantum Paradoxes* (New York: Wiley-VCH)
- [11] Note that this measurement must last a period of time longer than D/c in order not to violate causality; see also Tollaksen J 2007 *Proc. SPIE* **6573** 65730
- [12] Aharonov Y, Albert D and Vaidman L 1988 *Phys. Rev. Lett.* **60** 1351
Aharonov Y and Vaidman L 1990 *Phys. Rev. A* **41** 11

- Aharonov Y and Vaidman L 1991 *J. Phys. A: Math. Gen.* **24** 2315
- Aharonov Y and Botero A 2005 *Phys. Rev. A* **72** 052111
- Aharonov Y, Casher A, Albert D and Vaidman L 1987 *Phys. Lett. A* **124** 199
- [13] Doolen G and Whaley B 2002 *A Quantum Information Science and Technology Road Map* Section 6.8, December 1, Version 1.0
- [14] Reznik B and Aharonov Y 1995 *Phys. Rev. A* **52** 2538
- [15] Aharonov Y, Massar S, Popescu S, Tollaksen J and Vaidman L 1996 *Phys. Rev. Lett.* **77** 983
- Aharonov Y, Botero A, Popescu S, Reznik B and Tollaksen J 2002 *Phys. Lett. A* **301** 130–8
- Tollaksen J 2007 *J. Phys. A: Math. Theor.* **40** 9033–66
- Aharonov Y, Botero A and Scully M 2001 *Z. Naturf. A* **56** 5–15
- Nussinov S N and Tollaksen J 2008 *Phys. Rev. D* **78** 036007
- [16] von Neumann J 1983 *Mathematical Foundations of Quantum Theory* (Princeton, NJ: Princeton University Press)
- [17] Bohm D and Aharonov Y 1957 *Phys. Rev.* **108** 1070–6
- [18] Tollaksen J and Aharonov Y 2010 *J. Phys.: Conf. Ser.* in press
- [19] Mir R, Lundeen J S, Mitchell M W, Steinberg A M, Garretson J L and Wiseman H 2007 *New J. Phys.* **9** 287
- Hosten O and Kwiat P 2008 *Science* **319** 787
- Solli D R *et al* 2004 *Phys. Rev. Lett.* **92** 043601
- [20] Spence S E and Parks A D 2010 Experimental Evidence for a Dynamical Non-locality Induced Effect in Quantum Interference using Weak Values *in preparation*
- [21] Resch K J, Lundeen J S and Steinberg A M 2004 *Phys. Lett. A* **324** 125–31
- [22] Furry W H and Ramsey N F 1960 *Phys. Rev.* **118** 623
- [23] Shimony A 1984 *Proc. Int. Symp. on Foundations of Quantum Mechanics in Light of New Technology 1983 (Tokyo)* ed S Kamefuchi, H Ezawa, Y Murayama, M Namiki, S Nomura, Y Ohnuki and T Yajima (Tokyo: Physics Society of Japan) pp 225–30, cf p 227
- Shimony <http://plato.stanford.edu/entries/bell-theorem/>
- [24] Feynman R P and Hibbs A R 1965 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill) pp 2–9
- Tonomura A, Endo J, Matsuda T, Kawasaki T and Ezawa H 1989 *Am. J. Phys.* **57** 117
- [25] Batelaan N and Tonomura A 2009 *Phys. Today* **62** 38
- [26] Aharonov Y and Tollaksen J 2009 New insights on time-symmetry in quantum mechanics *VISIONS OF DISCOVERY: New Light on Physics, Cosmology And Consciousness* ed R Y Chiao, M L Cohen, A J Leggett, W D Phillips and C L Harper Jr (Cambridge: Cambridge University Press)
- [27] Tollaksen J 2001 Quantum reality and nonlocal aspects of time *PhD Thesis* Boston University (thesis advisor: Professor Yakir Aharonov)
- [28] Parks A D, Cullin D W and Stoudt D C 1979 *Proc. R. Soc. A* **454** 2997–3008
- [29] Pryde G J, O’Brien J L, White A G, Ralph T C and Wiseman H M 2005 *Phys. Rev. Lett.* **94** 220405
- [30] Ritchie N W M, Story J G and Hulet R G 1991 *Phys. Rev. Lett.* **66** 1107
- [31] Ahnert S E and Payne M C 2004 *Phys. Rev. A* **70** 042102
- [32] Wiseman H M 2002 *Phys. Rev. A* **65** 032111
- [33] Tollaksen J 2007 Robust weak measurements on finite samples *J. Phys.: Conf. Ser.* **70** 012015
- Brandt H, Kim Y S and Man’ko M A (ed) 2007 arXiv:quant-ph/0703038
- [34] Tollaksen J 2007 Non-statistical weak measurements *Quantum Information and Computation V (Proc. SPIE vol 6573)* ed E Donkor, A Pirich and H Brandt (Bellingham, WA: SPIE) CID 6573–33
- [35] Brout R, Massar S, Parentani R, Popescu S and Spindel Ph 1995 *Phys. Rev. D* **52** 1119
- Nussinov S and Tollaksen J 2008 *Phys. Rev. D* **78** 036007
- [36] Robert P Crease *Physics World* Sept 2002, p. 19
- [37] Tollaksen J 2007 Novel Relationships between Superoscillations, Weak Values, and Modular Variables *J. Phys. Conf. Series* **012016**