The Deterministic set of Operators, Quantum Interference Phenomena, and Quantum Reality

Jeff Tollaksen and Yakir Aharonov

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

E-mail: `tollaksen@chapman.edu`

Abstract.
   We develop an approach to quantum interference phenomenon which makes the characteristics of the problem more conspicuous and brings out more physical concepts. We do this by discussing the idea of deterministic quantum experiments using the Heisenberg picture. The variables that are relevant for describing interference are modular variables. These variables obey non-local equations of motion and can be used to follow the time evolution of a system without disturbing it. 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). While the Heisenberg picture and the Schrodinger 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 missed from the old perspective.

Written to celebrate the 75th birthday of E.C.G. Sudarshan: deep and subtle

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): $dn(x)/dx \sim |\psi_{\rm L}(x) + e^{i\alpha}\psi_{\rm R}(x)|^2$ with $\psi_{\rm L}(x)$ coming from the left slit, $\psi_{\rm R}(x)$ from the right (located a distance D away), and α the relative phase between the left and right parts of the wavefunction.

There are two ways to think about such phenomenon:

   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 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 it's 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 story and advocates further inquiry. For example, Feynman [32] stated that such phenomena "..have in it the heart of quantum mechanics. In reality, it 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 phenomenon 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 $\psi_\mathrm{L}(x) + e^{i\alpha}\psi_\mathrm{R}(x)$ 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, $\psi_\mathrm{L}(x) + e^{i\alpha}\psi_\mathrm{R}(x)$ and $\psi_\mathrm{L}(x) + e^{i\beta}\psi_\mathrm{R}(x)$ 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 $\psi_\mathrm{L}(x) + \psi_\mathrm{R}(x)$ to $\psi_\mathrm{L}(x) - \psi_\mathrm{R}(x)$ - would not result in a measureable change in any local properties. The change only shows up in certain non-local properties or much later when the two separate components $\psi_\mathrm{L}(x)$ and $\psi_\mathrm{R}(x)$ eventually overlap and interfere. It seems that the relative phase cannot be thought of simply as the difference between a local phase at $\psi_\mathrm{L}(x)$ and another local phase at $\psi_\mathrm{R}(x)$.

Another aspect of this second way of thinking is the realization that the Schrödinger wave only has a measureable meaning for an ensemble of particles, not for a single particle (generally speaking). 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 sufficiently small such that only one electron/photon traverse the interference apparatus at a time. The interference pattern with light and dark bands is nevertheless built up successively, mark by mark, with each individual "particle-like" electron/photon, [37]. 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." [32]

We follow the second way of thinking and offer a fresh approach to this time honored problem [13] [12] [26] [29]. To motivate the first step, involving a fundamental shift in the types of observables utilized, we make several observations:

First, most discussions of this problem are based on measurements which 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.

Second, 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.

Third, 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 in separating various aspects of quantum theory from the probabilistic aspects [14]. The underlying framework for the approach to interference presented in this paper is based on another kind of non-disturbing measurement on the "set of deterministic operators" or "deterministic experiments" [12] [29]. This set $(\hat{A}_\psi)$ involves measurement of only those variables for which the state of the system under investigation is an eigenstate. This question is dual to the more familiar question "what are the eigenstates of a given operator?" Measurement of these operators $\hat{A}_\psi$ does not collapse the wavefunction, since the wavefunction is initially an eigenstate of the operator being measured. The essential point emphasized here is the relevance of deterministic experiments for a *single* particle since they can be performed without causing a disturbance. As an example, consider a spin-1/2 particle placed in a magnetic field $\vec{B}$ pointing in the x-direction and with $\sigma_z = +1$ initially. If at some time t later we measure σ_z again, the result is not predictable; we get fluctuations and the experiment is not a deterministic one. However, if instead we rotate our apparatus and measure the spin in the yz-plane in a direction $\alpha = \omega t$, where ω is the Larmor frequency, then the result is predictable and we have a deterministic experiment. As a second example, consider a free particle of mass m which at $t = 0$ is in the state $\psi = C \exp\{-\frac{x^2}{\Delta x^2}\}$. This state is an eigenstate of the Hermitian operator $\hat{A}(0) = \hat{x}^2 + \hat{p}^2$. At time t we can measure the operator $\hat{A}(t) = \left\{\hat{x} - \frac{\hat{p}}{m}t\right\}^2 + \hat{p}^2$ with a predictable outcome, and thus we have again a deterministic experiment.

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 senstive to the relative phase are functions of **modular variables**. For the special 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 phenomenon, the Heisenberg equations of motion for these modular variables are *non-local*. The nonlocality 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 behavior 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 an unphysical 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 [36] [16] 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 science. 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) [12] in order to explain the nonlocality of topological phenomena such as the Aharonov-Bohm (AB) effect [4] [38]. The AB effect conclusively proved that a magnetic (or electric) field inside a confined region can have a measureable impact on a charged particle which never traveled inside the region. In order to represent the closest correspondence between measurement and theory, APP introduced nonlocal 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. [33] Prior to APP, dynamical nonlocality was avoided due to the possibility that it could violate causality. However, in a beautiful theorem, APP proved that the dynamical nonlocality 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 which make the non-local exchange completely uncertain and therefore "un-observable."

While it was beautiful that quantum mechanics 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 nonlocal exchange of modular variable, 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 nonlocal - yet un-observable - effect?

It was recently pointed out [28], that these non-local interactions ***can*** be observed. To measure this nonlocality (when it occurs) in a causality-preserving manner, various tools must be utilized such as pre-selection, post-selection, and weak measurements.

This development thereby underscores 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 phenomenon 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 associated language are, 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 missed from the old perspective. These types of developments are signatures of a successful re-formulation.

We briefly mention one important conceptual shift evident in this new approach: 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 is 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 mechanicics, 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. The Deterministic Set of Operators (DSOs)

The approach to the issue of the quantum reality for a single particle is based on the notion of "deterministic" experiments. We call an experiment a deterministic experiment when we measure only variables for which the state of the system under investigation is an eigenstate. In other words, for any state ψ it is possible to ask the following question: "What is the set of Hermitian operators $\hat{A}_\psi$ for which ψ is an eigenstate?" That is, which satisfy:

$$\hat{A}_\psi = \{\hat{A}_i \; such \; that \; \hat{A}_i|\psi(t)\rangle = a_i|\psi(t)\rangle \,, a_i \in \Re\} \tag{2.1}$$

where $\Re$ is the set of real numbers. This question is dual to the more familiar question which asks "What are the eigenstates of a given operator?" A measurement of such operators $\hat{A}_\psi$ would not lead to collapse of the wave function, since the wave function is initially an eigenstate of the operator being measured. A succession of such experiments can be performed in time and still, no statistics need be associated with the experimental results. The results are completely predictable, i.e. the experiments are deterministic. In deterministic experiments, it is possible to give a detailed physical description of the evolution of the system.

The set $\hat{A}_\psi$ is closed in the sense that if $\hat{A}_i, \hat{A}_j \in \hat{A}_\psi$ then:

(i) $[\hat{A}_i, \hat{A}_j] = \hat{A}_k \in \hat{A}_\psi$ since $\hat{A}_k\psi = 0$. $\hat{A}_k$ is an operator proportional to a projector perpendicular to ψ.

(ii) $\hat{A}_i + \hat{A}_j$ and $\hat{A}_i\hat{A}_j$ are in $\hat{A}_\psi$

The set $\hat{A}_\psi$ is also unique in the sense that it defines a single state $|\psi(t)\rangle$ which satifies the set membership criteria of eq. (2.1)[1]. This suggests that we can equivalently specify either the state $|\psi(t)\rangle$ or the set $A_\psi(t)$ of all time dependent eigenoperators of $|\psi(t)\rangle$. It follows that, since we know H and $A_\psi(0)$, that we know $A_\psi(t)$. The logic is straightforward: knowing the set $A_\psi(0)$ gives us uniquely $|\psi(0)\rangle$. $|\psi(0)\rangle$ together with H gives us $|\psi(t)\rangle$, which has a unique set $A_\psi(t)$ of eigenoperators. The proof of the uniqueness is trivial in a finite (and denumerably infinite) dimensional Hilbert space.

If we have a Hilbert space of dimension n then we can choose a basis such that a state vector $|\psi\rangle$ can be represented by an $n \times 1$ column vector (a unitary transformation can always bring us to this basis) given by 2.2.a and orthogonal vectors of the form given by 2.2.b:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \; a) \qquad\qquad \begin{pmatrix} 0 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \; b) \tag{2.2}$$

The Hermitian operators that operate on this space, represented (in that same basis) by an $n \times n$ matrix are in turn an n^2 dimension space. Operators of the form eq. 2.3 can be the DSO (where $a_{ij} = a_{ij}^*$) with dimensionality $(n-1)^2 + 1$.

The extension of these results to denumerably infinite spaces is straightforward: any matrix can be written as a sum over a complete set of vectors $|\Psi_n\rangle\langle\Psi_m|$ i.e. $\sum a_{nm}|\Psi_n\rangle\langle\Psi_m|$. An infinite matrix representing DSO is then given by:

$$a_{11}|\Psi_1\rangle\langle\Psi_1| + \sum_{n,m\neq 1} a_{nm}|\Psi_n\rangle\langle\Psi_m| \tag{2.4}$$

---

[1] This is so because any other statevector will be disturbed by at least one operator in $\hat{A}_\psi$. In addition, any other operator not in $\hat{A}_\psi$ will disturb $|\psi(t)\rangle$ (see §2.1)

$$\begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & a_{32} & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix} \tag{2.3}$$

dimension 1

$dim(n-1) \ form \ \sum_{n,m \neq 1} a_{nm}|\Psi_n\rangle\langle\Psi_m|$

Figure 1. Deterministic subset of operators with dimension $(n-1)^2 + 1^2$

There is a single unique eigenvector associated with this that is given by eq. (2.2.a) with either finite n or infinite $n \to \infty$. This is the only vector that is not disturbed by operators of the form $\sum_{n,m \neq 1} a_{nm}|\Psi_n\rangle\langle\Psi_m|$ in eq. (2.4). If the only operators being used or measured are those for which eq. (2.1) is satisfied, then the wavefunction evolves in a deterministic way, i.e. if we define ψ at one time, then we know ψ uniquely at some later time. Hence, if we define DSO at one time, it must evolve in a completely deterministic fashion to the DSO at a later time.

2.1. The Completely Uncertain Operators

In addition to the set $\hat{A}_\psi$ given by 2.1, there is also a set B_ψ of operators whose results are completely uncertain. If we define $\hat{\Pi}_\psi$ such that $\hat{\Pi}_\psi|\psi\rangle = |\psi\rangle$ where $\hat{\Pi}_\psi^2 = \hat{\Pi}_\psi$ and $\hat{\Pi}_\psi|\psi'\rangle = 0$ if and only if $\langle\psi'|\psi\rangle = 0$, then this operator is the projector on $|\psi\rangle$, $\hat{\Pi}_\psi = |\psi\rangle\langle\psi|$. We can also decompose any operator $\hat{C}(x,p)$ as follows:

$$\hat{C}(x,p) \equiv \hat{A}_\psi^{(C)} + \hat{B}_\psi^{(C)} \tag{2.5}$$

where $\hat{A}_\psi^{(C)} = \hat{\Pi}_\psi\hat{C}\hat{\Pi}_\psi + (1 - \hat{\Pi}_\psi)\hat{C}(1 - \hat{\Pi}_\psi)$ and $\hat{B}_\psi^{(C)} = (1 - \hat{\Pi}_\psi)\hat{C}\hat{\Pi}_\psi + \hat{\Pi}_\psi\hat{C}(1 - \hat{\Pi}_\psi)$. In the above basis, the completely uncertain operators for 2.2.a have the matrix form:

$$\underbrace{\begin{pmatrix} 0 & a_{12} & a_{13} & \ldots & a_{2n} \\ a_{21} & 0 & 0 & \ldots & 0 \\ a_{31} & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ a_{n1} & 0 & 0 & \ldots & 0 \end{pmatrix}}_{completely \ uncertain \ operators} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix} \tag{2.6}$$

Note that the uncertain operators have non-zero entries a_{12}, a_{13}, $...a_{2n}$ and a_{21}, $a_{31}...a_{n1}$ which take the state given by 2.2.a to orthogonal states 2.2.b.

2.2. Geometric interpretation

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 [21] that information can be obtained even though not a single particle (in an ensemble) was disturbed. To begin to introduce this point, let us consider a general theorem for any vector (state) in Hilbert space:

Theorem I: $\hat{A}|\psi\rangle = \langle\hat{A}\rangle|\psi\rangle + \Delta A|\psi_\perp\rangle$ 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 such that $\langle\psi|\psi_\perp\rangle = 0$ (see fig. 2.a). Proof: left multiplication by $\langle\psi|$ yields the first term; evaluating $|(A - \langle A\rangle)|\psi\rangle|^2 = \Delta A^2$ yields the second.

Using this theorem, it is easy to see that in the basis of eq. 2.2, the diagonal elements of 2.3 correspond to the averages of the observables, e.g. $a_{11} = \langle \hat{A} \rangle$ for state 2.2.a, which again, can be measured without uncertainty. We can also see how the relation $\hat{A}|\Psi\rangle = \langle \hat{A} \rangle |\Psi\rangle + \Delta A |\Psi_{\perp}\rangle$ (Theorem I) works in the DSO picture. As a result of the Hermiticity and the closed algebra, we can think of these operators as vectors in a real vector space. We can also consider a product, $\hat{A}_1 \odot \hat{A}_2 = trace(\hat{A}_1 \cdot \hat{A}_2)$ which defines the length of the vectors. The normalization of 2.6 then is $N \equiv |a_{21}|^2 + |a_{31}|^2 + \cdots + |a_{n1}|^2$. Using $a_{ij} = a_{ij}^*$, we see that this is the same as $\sqrt{N}$, which is the uncertainty in the operator, ΔA. We can therefore construct the geometric interpretation for any operator 2.5 illustrated in figure 2.b: the length of the projection onto the DSO is given by the average of the operator, and the length of the projection onto the completely uncertain operators is given by the uncertainty in the operator. Subtracting the dimensionality of the whole set of Hermitian operators, n^2, from the dimensionality of the DSO, $(n-1)^2 + 1$, gives us $O(n)$ the dimensionality of the completely uncertain operators. Therefore, as $n \to \infty$, most of the parameters belong to the DSO.

Figure 2. a) geometry of $\hat{A}|\Psi\rangle = \langle \hat{A} \rangle |\Psi\rangle + \Delta A |\Psi_{\perp}\rangle$, b) geometry of $\hat{C}(x,p) \equiv \hat{A}_{\psi}^{(C)} + \hat{B}_{\psi}^{(C)}$, i.e. any operator can be decomposed into it's projection (proportional to $\langle \hat{C} \rangle$) onto the DSO $\hat{A}_{\psi}$ and its projection (proportional to ΔC) onto the completely uncertain set B_{ψ}

3. Interference phenomena from the traditional perspective

To show the utility of DSOs in interference situations, we review past attempts to analyze the disappearance of interference whenever it is possible to detect through which slit the particle passes. 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 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 which helps to motivate our approach. It has been argued (borrowing from the discussion of the "Heisenberg microscope") that if the particle 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! [12] [34] One might suppose that since the action of opening/closing the left slit never caused an interaction with the particle at the right slit, then nothing associated

with the particle should change. But, it was first pointed out by APP [12] that in this scenario when a WWM is performed without actually interacting with the interfering particle, then 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 ways[2]:

(i) 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.

(ii) Or, we may look at the fourier transform of the probability distribution $\int \rho(p) e^{\frac{i}{\hbar} pD} dp$. (We will later see that these functions, $\langle e^{\frac{i}{\hbar} pD} \rangle$, are precisely the observables that are sensitive to the relative phase.) To analyze the effect of the interaction, we calculate $\langle e^{\frac{i}{\hbar} pD} \rangle_f - \langle e^{\frac{i}{\hbar} pD} \rangle_i$ 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.

The first approach has been used universally to analyze interference phenomena [34]. 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.

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 much 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 $(\Delta x)^2 \equiv \overline{(\hat{x})^2} - \overline{\hat{x}}^2$ and $(\Delta p)^2 \equiv \overline{(\hat{p})^2} - \overline{\hat{p}}^2$, 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 disjoint "lumps," ψ_{L} and ψ_{R}:

$$|\Psi_\alpha\rangle = \frac{1}{\sqrt{2}} \{ |\psi_{\mathrm{L}}\rangle + e^{i\alpha} |\psi_{\mathrm{R}}\rangle \} \tag{3.7}$$

Collapsing it to just $\psi_{\mathrm{R}}(x) \equiv \langle x | \psi_{\mathrm{R}} \rangle$ does *not* change Δp nor the expectation values of any finite order polynomial in p, as none of these local operators have a non-vanishing matrix element between the disjoint "lumps" of the wavefunction. In fact, the superposition of any countable number of such wavepackets (so that Δx becomes arbitrarily large) does not require the uncertainty in momentum, Δp, to increase. To see this, consider a simple example: let $\psi_o(x)$ be a wave packet centered at $x = 0$ that is exactly equal to zero for $x \in [-\frac{L_o}{2}, \frac{L_o}{2}]$. For this wave function we have $\Delta p \approx \frac{\hbar}{\Delta x}$. We now define $\psi_n(x) = \psi_o(x + nL)$ where $n = 1, 2, ...$ and $L > L_o$.

---

[2] We consider momentum here, but our comments apply to any conserved quantity.

We now show that for the class of states $\Psi(x) = \sum_n a_n \psi_n(x)$ with $\sum_n |a_n|^2 = 1$, the uncertainty in momentum is independent of a_n. We have $\Delta p = (\langle p^2 \rangle - \langle p \rangle^2)^{1/2}$. For $\langle p \rangle$ we have:

$$\langle p \rangle = \int_{-\infty}^{+\infty} dx \, \Psi^*(x)(-i\hbar)\frac{\partial}{\partial x}\Psi(x) = \int_{-\infty}^{+\infty} dx \sum_{n,m} a_n^* a_m \Psi_n^*(x)(-i\hbar)\frac{\partial}{\partial x}\psi_m(x) \qquad (3.8)$$

But, by definition ψ_m and ψ_n do not overlap, so that ψ_m and its derivatives are different from zero only where $\psi_n = 0$.[3] Thus

$$\langle p \rangle = \sum_{n,m} a_n^* a_m \frac{\hbar}{i} \int_{-\infty}^{+\infty} dx \, \psi_n^*(x)\frac{\partial}{\partial x}\psi_n(x) \qquad (3.9)$$

Further, since $\psi_n(x) = \psi_o(x + nL)$, we have

$$\int_{-\infty}^{+\infty} dx \, \psi_n^*(x)\frac{\partial}{\partial x}\psi_n(x) = \int_{-\infty}^{+\infty} dx \, \psi_o^*(x+nL)\frac{\partial}{\partial x}\psi_o(x+nL) = \int_{-\infty}^{+\infty} dx \, \psi_o^*(x)\frac{\partial}{\partial x}\psi_o(x) \quad (3.10)$$

Therefore,

$$\langle p \rangle = \sum_n |a_n|^2 \frac{\hbar}{i} \int_{-\infty}^{+\infty} dx \, \psi_o^*(x)\frac{\partial}{\partial x}\psi_o(x) = \frac{\hbar}{i} \int_{-\infty}^{+\infty} dx \, \psi_o^*(x)\frac{\partial}{\partial x}\psi_o(x) \qquad (3.11)$$

I.e. $\langle p \rangle$ is independent of the choice of a_n. In a similar way, one can show that

$$\langle p^2 \rangle = \sum_n |a_n|^2 \frac{\hbar}{i} \int_{-\infty}^{+\infty} dx \, \psi_o^*(x)\left\{\frac{\partial}{\partial x}\right\}^2 \psi_o(x) = \frac{\hbar}{i} \int_{-\infty}^{+\infty} dx \, \psi_o^*(x)\left\{\frac{\partial}{\partial x}\right\}^2 \psi_o(x) \qquad (3.12)$$

I.e. $\langle p^2 \rangle$ is independent of a_n and thus $(\Delta p)^2 = (\langle p^2 \rangle - \langle p \rangle^2)$ is independent of the choice of a_n.

From this surprising result we learn, for example, that if in a typical quantum interference experiment we want to measure through which slit the particle passes, then we do not have to increase the uncertainty in momentum. In fact, what our intended experiment does is to change the situation from one in which $a_1 = a_2 = \frac{1}{\sqrt{2}}$ to one in which, e.g., $a_1 = 1$, $a_2 = 0$. In both situations Δp is the same. Therefore, it is misleading to say that in measuring which slit the particle passes through we destroy the interference pattern by increasing the uncertainty in momentum.

Up until now 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 tell us nothing about this mystery. The decisive importance of the second "fourier transform" approach for this mystery is best illustrated through a basic theorem which characterizes all interference phenomenon: all **moments** of both position and momentum are *independent* of the relative phase parameter α (until the wavepackets overlap):

Theorem II: 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:

$$\int [\Psi_\alpha^*(x,t)\Psi_\alpha(x,t) - \Psi_\beta^*(x,t)\Psi_\beta(x,t)]x^m p^n dx = 0 \qquad (3.13)$$

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.

---

[3] $\psi^*(x)\frac{\partial}{\partial x}\psi(x+L) = 0$ for all x if $\psi(x)$ vanishes outside the interval, then so does its derivatives.

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.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 $\langle x^n(t) \rangle = \int \Psi_\alpha^*(x,0) x^n(t) \Psi_\alpha(x,0) dx$ and noting that $x(t) = x(0) + p(0)\frac{t}{m}$ and $p(t) = p(0)$ in this representation, we must have $\langle x^n(t) \rangle = \int \Psi_\alpha^*(x,0)[x(0) + p(0)\frac{t}{m}]^n(t) \Psi_\alpha(x,0) dx$. This is independent of α, since term by term it is independent of α. Eq. (3.13) then follows, and holds for p^n, as long as we retain the proper $\Psi_\alpha^* p^n \Psi_\alpha$ order.

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, equi-distant slits, these are the discrete translation by $\pm D$, namely $\exp\{\pm \frac{i}{\hbar}\hat{p}D\}$, effecting $\exp\{-\frac{i}{\hbar}\hat{p}D\}\psi_R(x) \rightarrow \psi_R(x - D)$ which overlaps with $\psi_L(x)$. The expectation value of the translation operator $\exp\{\frac{i}{\hbar}\hat{p}D\}$ ***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 three additional mathematical advantages concerning the second "fourier transform" approach.

First, all the moments $\langle p^n \rangle$ are averages of unbounded quantities, while $\langle \exp\{\frac{i}{\hbar}\hat{p}D\} \rangle$ 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 could 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 second significant "mathematical" difference concerns the utility of conservation laws. 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 setup. 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.

Finally, the Fourier transform method provides us with the parameters relevant to physical problem (e.g. the distance D between the slits), while 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 ($prob(A)$), then the probability of another should also change ($prob(B)$), such that the probability of the sum ($prob(A + B)$) does not change. As we pointed out, there are situations where the probability of one variable does not change ($prob(B)$), while

the probability of the other does change ($prob(A)$). This, per Theorem II, can only happen if that variable (e.g. B) is completely uncertain. That is, this can only happen if the fourier transform of $prob(B)$ below some value remains unaffected while the fourier transform of $prob(B)$ above some value is affected. This means that there is a whole range of modular variables that are being exchanged nonlocally and a large number of conservation laws which can be utilized.

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

4.1. **Modular variables are the observables that are sensitive to the relative phase**

We previously pointed to the significance of the Heisenberg translation operator, $\exp\{\pm\frac{i}{\hbar}\hat{p}D\}$, effecting $\exp\{-\frac{i}{\hbar}\hat{p}D\}\psi_\mathrm{R}(x) \to \psi_\mathrm{R}(x - D)$ overlapping with $\psi_\mathrm{L}(x)$. Therefore, the expectation value of the translation operator $\exp\{\frac{i}{\hbar}\hat{p}D\}$ ***does*** depend on α: $\langle\Psi_\alpha \mid \exp\{i\hat{p}D/h\} \mid \Psi_\alpha\rangle = e^{-i\alpha}/2$. More generally:

$$\langle e^{i\frac{pD}{\hbar}}\rangle = \int_{-\infty}^{+\infty} dx \sum_m a_m^* \psi_m^* e^{i\frac{pD}{\hbar}} \sum_n a_n^* \psi_n^* = \int_{-\infty}^{+\infty} dx \sum_m a_m^* \psi_m^* \sum_n a_n^* \psi_{n+1}^* = \sum_n a_{n+1}^* a_n \quad (4.14)$$

Similarly for $\langle e^{-i\frac{pD}{\hbar}}\rangle = \sum_n a_{n-1}^* a_n$ and thus:

$$\langle\cos\frac{pD}{\hbar}\rangle = \frac{1}{2}\sum_n \{a_{n+1}^* a_n + a_{n-1}^* a_n\} = \frac{1}{2}\sum_n \{a_{n+1}^* + a_{n-1}^*\} a_n \qquad (4.15)$$

which obviously depends on the choice of a_n. In particular, if all the a_n are equal, then $\langle\cos\frac{pD}{\hbar}\rangle = 1$, while if only one a_n is different from zero, then $\langle\cos\frac{pD}{\hbar}\rangle = 0$. Thus, a measurement to find through which slit the particle passes in an interferometer will change the value of $\langle\cos\frac{pD}{\hbar}\rangle$.

But, exactly *what* information about α does $\langle\exp\{\pm\frac{i}{\hbar}\hat{p}D\}\rangle$ reveal? It is easy to see that if we replace p with $p - \frac{nh}{D}$, then $e^{\frac{i}{\hbar}\hat{p}D}$ changes by $e^{\frac{iD}{\hbar}\frac{nh}{D}} = e^{in2\pi} = 1$, i.e. nothing changes. Furthermore, suppose n is the largest integer such that $n\frac{h}{D} < p$ (i.e. satisfying $0 \le \hat{p} - n\frac{h}{D} \le \frac{h}{D}$). This means that $e^{\frac{i}{\hbar}\hat{p}D}$ gives us information about the remainder after this integer number of $\frac{h}{D}$ is subtracted from p. This is otherwise known of as the modular momentum $p_{\mathrm{mod}} \equiv \hat{p}$ modulo $\frac{h}{D}$ (see fig. 3) defined by: $\hat{p}$ modulo $\frac{h}{D} \equiv \hat{p} - n\frac{h}{D}$.

It is clear that $p \bmod \frac{h}{D}$ has the topology of a circle, as would 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 which displays the time modulo 12.

We can get back to ordinary momentum through the relation: $p = N_p\frac{h}{D} + p_{\mathrm{mod}}$. We can see this (fig 3) if we stack an integer number (N_p) of $\frac{h}{D}$ on top of the modular portion of p (p_{mod} is the lower portion of fig 3). Note that the eigenstates of the translation operator $\exp\{\frac{i}{\hbar}\hat{p}D\}$ are also eigenstates of the modular momentum p_{mod}.

4.2. **Modular variables satisfy non-local equations of motion**

The key to our explanation of interference from the single particle perspective are the non-local equations of motion satisfied by these modular variables. Thus, using $H = \frac{p^2}{2m} + V(x)$ and
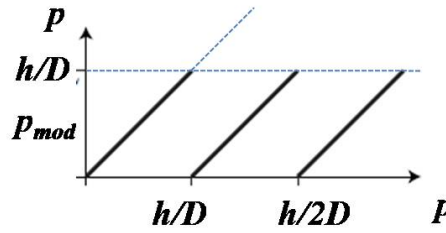
Figure 3. Stacking an integer number (N_p) of $\frac{h}{D}$ on top of the modular portion of p (p_{mod}.

$e^{\frac{i}{\hbar}\hat{p}D}V(x)e^{-\frac{i}{\hbar}\hat{p}D} = V(x+D)$, we find **non-local** [13, 12] Heisenberg equations of motion for modular variables:

$$\frac{d}{dt}e^{\frac{i}{\hbar}\hat{p}D} = \frac{i}{\hbar}[H, e^{\frac{i}{\hbar}\hat{p}D}] = \frac{i}{\hbar}[V(x) - V(x+D)]e^{\frac{i}{\hbar}\hat{p}D} \tag{4.16}$$

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

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

$$\frac{df(p)}{dt} = \{f(p), H\}_{PB} = -\frac{\partial f}{\partial p}\frac{\partial H}{\partial x} \quad + \underbrace{\frac{\partial f}{\partial x}\frac{\partial H}{\partial p}}_{=0} = 0 \tag{4.17}$$

i.e. $f(p)$ changes only if $\frac{\partial V}{\partial x} \neq 0$ 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 $x_o p_o = h$. These are easy to miss in the Schrödinger picture.

However, in the Heisenberg picture, the non-local equations of motion that the modular variables satisfy 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). 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.

Let's consider a particular example with $f(p) \equiv \sin\frac{p}{p_o}$ (setting $\frac{\hbar}{D} = p_o$). Classically we have:

$$\frac{d}{dt}\sin\frac{p}{p_o} = \frac{dp}{dt}\frac{1}{p_o}\cos\frac{p}{p_o} = -\frac{\partial V}{\partial x}\frac{1}{p_o}\cos\frac{p}{p_o} \tag{4.18}$$

If the force vanishes ($\frac{\partial V}{\partial x} = 0$) at the position of the particle, then $\sin\frac{p}{p_o}$ will be a constant of the motion for every value of p_o. The quantum equations of motion for $\sin\frac{p}{p_o}$ is (using $\sin\frac{p}{p_o} = \frac{1}{2i}\left\{e^{i\frac{p}{p_o}} - e^{-i\frac{p}{p_o}}\right\}$):

$$\frac{d}{dt}e^{i\frac{p}{p_o}} = \frac{i}{\hbar}[H, e^{i\frac{p}{p_o}}] = \frac{i}{\hbar}\left\{V(x)e^{i\frac{p}{p_o}} - e^{-i\frac{p}{p_o}}V(x)\right\} = \frac{i}{\hbar}\left\{V(x) - V(x+D)\right\}e^{i\frac{p}{p_o}} \tag{4.19}$$

Similarly $\frac{d}{dt}e^{-i\frac{p}{p_o}} = \frac{i}{\hbar}\left\{V(x) - V(x-D)\right\}e^{-i\frac{p}{p_o}}$ and therefore:

$$\frac{d}{dt}\sin\frac{p}{p_o} = \frac{1}{2i}\left\{\frac{d}{dt}e^{i\frac{p}{p_o}} - \frac{d}{dt}e^{-i\frac{p}{p_o}}\right\} = \frac{1}{2i}\frac{i}{\hbar}\left\{V(x)e^{i\frac{p}{p_o}} - V(x+D)e^{i\frac{p}{p_o}} - V(x)e^{-i\frac{p}{p_o}} + V(x-D)e^{-i\frac{p}{p_o}}\right\}$$

$$= \frac{1}{2\hbar} \left\{ V(x{-}D){-}V(x{+}D) \right\} \left\{ \frac{e^{i\frac{p}{p_o}}+e^{-i\frac{p}{p_o}}}{2} \right\} + i \left\{ 2V(x){-}V(x{-}D){-}V(x{+}D) \right\} \left\{ \frac{e^{i\frac{p}{p_o}}-e^{-i\frac{p}{p_o}}}{2} \right\}$$

$$= \frac{1}{2\hbar} \left\{ V(x{-}D){-}V(x{+}D) \right\} \cos\frac{p}{p_o} + i \left\{ 2V(x){-}V(x{-}D){-}V(x{+}D) \right\} \sin\frac{p}{p_o} \tag{4.20}$$

We see again that for variables like $\sin\frac{p}{p_o}$ there is a remarkable difference between the classical and the quantum equations of motion [4]. While the quantum equations of motion for the average of the momentum itself is identical to the classical equations of motion for the momentum, there is no such correspondence in the case of $\sin\frac{p}{p_o}$ [5] because quantum interference involves a non-local exchange of $\sin\frac{p}{p_o}$ between the particle and the screen with the slits.

4.3. **Non-local exchange of modular variables in the double-slit setup**
For the special case of two-slits, a set of spin-like observables can be identified as members of the set of deterministic operators. To obtain these operators, we take $\psi_{\rm L}$ and $\psi_{\rm R}$ as base states and consider the Hilbert space of all vectors:

$$\Psi(x) = a_1\psi_{\rm L}(x) + a_2\psi_{\rm R}(x) \tag{4.21}$$

with $|a_1|^2 + |a_2|^2 = 1$. Our purpose is to find the class of operators $\hat{A}_{a_1 a_2}(x,p)$ which satisfy the condition:

$$\hat{A}_{a_1 a_2}(x,p) \left\{ a_1\psi_{\rm L}(x) + a_2\psi_{\rm R}(x) \right\} = a \left\{ a_1\psi_{\rm L}(x) + a_2\psi_{\rm R}(x) \right\} \tag{4.22}$$

This will be the operators pertaining to the deterministic interference experiments. Since our Hilbert space is isomorphic to that of a spin-1/2 particle we expect that the $\hat{A}_\psi \equiv \hat{A}_{a_1 a_2}(x,p)$ operators will have the same properties as the Pauli spin operators and therefore satisfy a spin algebra $[\tilde{\sigma}_i, \tilde{\sigma}_j]_- = i\epsilon_{ijk}\tilde{\sigma}_k$ and $[\tilde{\sigma}_i, \tilde{\sigma}_j]_+ = \delta_{ij}$. One of the three operators should give us information about which slit the particle went through:

$$\tilde{\sigma}_3 = \frac{\sin(\pi x/D)}{|\sin(\pi x/D)|} \tag{4.23}$$

E.g., if $\tilde{\sigma}_3 = 1$ is obtained, then this means the particle went through the right slit (i.e. $a_1 = 0$ and $a_2 = 1$) while $\tilde{\sigma}_3 = -1$ refers to the particle going through the left slit because $\sin(\pi x/D)$ in eq. (4.23) is positive for $\psi_{\rm R}$ and negative for $\psi_{\rm L}$. To derive the other two spin-like operators $\tilde{\sigma}_1$

---

[4] Consider also a potential $V(x)$, such that $V(x+D) = V(x)+V_o$. Suppose we have a single particle characterized by a wavefunction that is in a superposition of being isolated in region $(0,D)$ and $(2D,3D)$ (etc). Initially we take $\exp\frac{i}{\hbar}\{pD\} = 1$ and thus, initially, $p(t)\,mod\,\frac{h}{D} = 0$. Subsequently, $\frac{d}{dt}e^{\frac{i\hat{p}D}{\hbar}} = V_o e^{\frac{i\hat{p}D}{\hbar}}$ leading to $e^{\frac{i\hat{p}D}{\hbar}}(t) = \exp\frac{i}{\hbar}\{pD - V_o t\}$ and thus $p(t)\,mod\,\frac{h}{D} = V_o t\frac{h}{D}$. Therefore, irrespective of the system state and even though the potential vanishes and no forces are acting in the region $(0,D)$, the modular momentum [12, 26, 13, 29] is changing at a constant rate based on the particles' nonlocal interaction with the potential in other distant regions. We showed that these functions will depend on exponentials of the form $e^{\frac{i\hat{p}D}{\hbar}}$ and therefore obey non-local equations of motion.

[5] There was an attempt to give a counter-example to the non-local phenomenon (e.g. the Aharonov-Bohm effect) described by modular variables by considering a formulation of QM in terms of the density $\rho(x,p) = \psi^*(x,p)\psi(x,p)$ and the current $\vec{j} = \psi^*\nabla\psi - \psi\nabla\psi^*$. The equations of motion for ρ and $\vec{j}$ are non-linear and only involve the fields (rather than potentials and are thus gauge invariant) and are local. Consider again the superposed states with relative phase α between the two wavepackets. Because $\vec{j} = \rho\nabla\alpha$, the relative phase can be found by integrating $\int \frac{\vec{j}}{\rho}$ over all space. However, for wavefunctions that go to zero in between the wavepackets, this approach does not work. Nevertheless, this approach does reveal an interesting connection between non-locality and chaos. For example, there is an instability characterized by infinitesimal changes in ρ and/or $\vec{j}$ in the region between the wavepackets which will result in a large change later when the wavepackets overlap.

and $\tilde{\sigma}_2$, we will need an operator $\tilde{\sigma}_+$ such that $\tilde{\sigma}_+|\psi_{\mathrm{L}}\rangle = |\psi_{\mathrm{R}}\rangle$ and $\tilde{\sigma}_+|\psi_{\mathrm{R}}\rangle = 0$, e.g. $\tilde{\sigma}_+ = \tilde{\sigma}_3 e^{-i\frac{pD}{\hbar}}$ and $\tilde{\sigma}_- = e^{i\frac{pD}{\hbar}}\tilde{\sigma}_3$. From this we construct $\tilde{\sigma}_1 = \tilde{\sigma}_+ + \tilde{\sigma}_-$ and $\tilde{\sigma}_2 = i(\tilde{\sigma}_+ - \tilde{\sigma}_-)$ and therefore:

$$\tilde{\sigma}_1 = \cos\frac{pD}{\hbar} - i\sin\frac{pD}{\hbar}\frac{\sin(\pi x/D)}{|\sin(\pi x/D)|} \tag{4.24}$$

$$\tilde{\sigma}_2 = \sin\frac{pD}{\hbar} + i\cos\frac{pD}{\hbar}\frac{\sin(\pi x/D)}{|\sin(\pi x/D)|} \tag{4.25}$$

Eq. 4.25 can be understood as translating the particle to the left when eq. 4.23 is positive (i.e. when the particle is located at the right slit then the operator is just $e^{\frac{ipD}{\hbar}}$). If $\tilde{\sigma}_1$ or $\tilde{\sigma}_2$ are measured, then information about the relative phase (i.e. the parameter α between the 2 wave-packets) is obtained, e.g. if we measure $\tilde{\sigma}_1$ and obtain $+1$, the state is $\frac{1}{\sqrt{2}}\{\psi_{\mathrm{L}} + \psi_{\mathrm{R}}\}$, i.e., $a_1 = 1 = a_2$; the result $\tilde{\sigma}_1 = -1$ yields $a_1 = 1$ and $a_2 = -1$. If we have the state: $\Psi = \frac{1}{\sqrt{2}}\left\{e^{i\frac{\alpha}{2}}\psi_{\mathrm{L}} + e^{-i\frac{\alpha}{2}}\psi_{\mathrm{R}}\right\}$ then the corresponding operator is $\tilde{\sigma}_1\cos\alpha + \tilde{\sigma}_2\sin\alpha = 1$. For any choice of a_1 and a_2 we can make a corresponding rotation in our state space, in full analogy to the ordinary spin-1/2 case.

For simplicity (without affecting the generality of our arguments), we can also express the relevant modular variable as the parity (exchange) operation $\hat{\mathsf{P}}$ (effecting $\hat{\mathsf{P}}|\psi_{\mathrm{L}}\rangle = |\psi_{\mathrm{R}}\rangle$ and $\hat{\mathsf{P}}|\psi_{\mathrm{R}}\rangle = |\psi_{\mathrm{L}}\rangle$). It **is** sensitive to the relative phase α between the disjoint lumps of eq. 3.7 [12] [13] [29]:

$$\langle\Psi_\alpha|\hat{\mathsf{P}}|\Psi_\alpha\rangle = \frac{1}{2}\left\{\langle\psi_{\mathrm{L}}| + e^{-i\alpha}\langle\psi_{\mathrm{R}}|\right\}\hat{\mathsf{P}}\left\{|\psi_{\mathrm{L}}\rangle + e^{i\alpha}|\psi_{\mathrm{R}}\rangle\right\}$$

$$= \frac{1}{2}\left\{e^{i\alpha} + e^{-i\alpha}\right\} = \langle\cos\alpha\rangle \tag{4.26}$$

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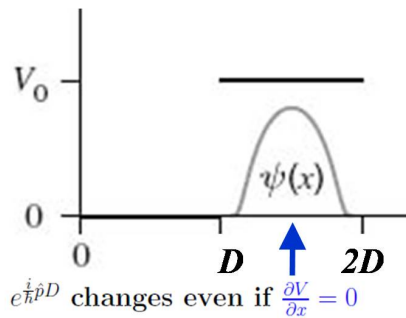

Figure 4. A potential with 2 values and a wave-packet with support only in a limited interval.

We can also see from eq. 5.27 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.

4.4. 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 $|\psi_{\mathrm{L}}\rangle$ or $|\psi_{\mathrm{R}}\rangle$ (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 the particle travels through the right slit and 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?

Up until this paper, it was believed that this could not be observabed. 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's call it θ - that is exchanged nonlocally. Suppose the amount of nonlocal exchange is given by $\delta\theta$ (i.e. $\theta \to \theta + \delta\theta$). Now "maximal uncertainty" means that the probability to find a given value of θ is independent of θ, i.e. $P(\theta) = constant = \frac{1}{2\pi}$. 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 nonlocal 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 un-measureable:

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

Consider how this works in the double slit setup. 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, $\overline{\hat{P}^2} = 1$, however, $\langle \psi_L|\hat{P}|\psi_L\rangle = \langle \psi_L|\psi_R\rangle = 0$ so $\overline{\hat{P}} = 0$, and therefore $\Delta\hat{P} \equiv \sqrt{\overline{(\hat{P})^2} - \overline{\hat{P}}^2} = 1$, 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.

The vanishing of the expectation value of the modular momentum variable 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. [6]

---

[6] Although much of the discussion in this article focuses on the simplest interference example with 2-slits, our approach becomes clearer when it is applied to an infinite number of slits with (interfering) particles that are initially in an eigenstate of momentum. In this case, we can directly speak about modular momentum (instead of slightly more complicated functions for the double-slit setup). Also, both the non-local equation of motion for modular momentum is exact as is the conservation of modular momentum.

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 resulting momentum uncertainty (destroying the interference pattern when the left slit is closed) is 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 involves 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 nonlocal 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 nonlocal - yet un-observable - effect?

5. Gedanken-experiment to measure non-local equations of motion

The key novel observation we next make is that this non-locality *does have* an observable meaning. We discuss two different methods.

5.1. Method 1: von Neuman measurement after the non-local interaction

It is thus clear that if we want to design a deterministic interference experiment we must find a method to measure $\tilde{\sigma}$-type operators (which are diagonal in the given two-slit set-up). According to the von Neumann method [39] this can be done provided a potential proportional to $\tilde{\sigma}$ can be switched on. From the point of view of non-relativistic quantum theory we may in principle assume an arbitrary potential V which is a function of both $\hat{x}$ and $\hat{p}$, and thus a potential proportional to, for example, $\tilde{\sigma}_1$. However, in reality we are restricted to potentials which are local, i.e. functions of positions alone. The existence of potentials which are arbitrary functions of momentum would lead to non-local motion of a particle. To see this, consider a situation where the Hamiltonian contains a term proportional to $\sin \frac{pD}{\hbar}$. We then have

$$i\hbar \frac{\partial \Psi(x,t)}{\partial t} = \sin \frac{pD}{\hbar} \Psi(x,t) = \frac{1}{2i} \left\{ e^{i\frac{pD}{\hbar}} - e^{-i\frac{pD}{\hbar}} \right\} \Psi(x,t) = \frac{1}{2i} \left\{ \Psi(x+D) - \Psi(x-D) \right\} \quad (5.27)$$

By way of example, suppose that at $t = 0$ the wave function Ψ is different from zero only in a region around $x = 0$ which is small compared to D. Then immediately after application of eq. 5.27, Ψ will begin to grow around the positions $x = +D$ and $x = -D$. This violates the idea of locality or continuity of motion. Therefore, if we want to admit only deterministic experiments which do not violate locality we must find *local* potentials to measure the $\tilde{\sigma}$-type variables.

We do not have to look very far for such potentials: when the particle in the two-slit experiment has passed the slits, it moves again as a free particle. Thus, its position at time t after the passage is related to the position at the time of passage by

$$x(t) = x(0) + \frac{p}{m}t = x(0) + \frac{p(0)}{m}t \quad (5.28)$$

If a measurement of $\sin[x(t)]$ is equivalent to a measurement of $\sin[x(0) + \frac{p(0)}{m}t]$. But

$$\sin[x(0) + \frac{p(0)}{m}t] = \frac{1}{2i} \left\{ \exp\left\{ i[x(0) + \frac{p(0)}{m}t] \right\} - \exp\left\{ -i[x(0) + \frac{p(0)}{m}t] \right\} \right\} \quad (5.29)$$

and using the identity $\exp(\hat{u} + \hat{v}) = \exp(\hat{u}) \exp(\hat{v}) \exp(-\frac{1}{2}[\hat{u}, \hat{v}])$ (which holds for any pair of operators, u and v, that commute with $[u, v]$), it is easy to see that

$$\sin[x(0) + \frac{p(0)}{m}t] = \exp\left\{\frac{it\hbar}{2m}\right\}\left\{\sin[x(0)]\cos[\frac{p(0)}{m}t] + \cos[\frac{p(0)}{m}t]\sin[x(0)]\right\} \tag{5.30}$$

Thus, measurement of $\sin[x(t)]$ is related to measurement of sine and cosine of $x(0)$ and $p(0)$. This suggests that we can reach our goal by introducing a periodic potential in the x-direction. To exhibit the essential aspects of the procedure, consider a simple one-dimensional example of two disjoint wavepackets approaching each other, with a velocity $2\vec{v}$. We let $\Psi_2(x) = \Psi_1(x - 2D)$, and at $t = 0$ let the two wavepackets be

$$\Psi_1(x, 0) = N \int_{-\infty}^{\infty} e^{-\frac{(k-k_o)^2}{\Delta k^2}} e^{ik(x+D)} dk \quad and \quad \Psi_2(x, 0) = N \int_{-\infty}^{\infty} e^{-\frac{(k-k_o)^2}{\Delta k^2}} e^{ik(x-D)} dk \tag{5.31}$$

Then at time t we have

$$\Psi_1(x, t) = N \int_{-\infty}^{\infty} e^{-\frac{(k-k_o)^2}{\Delta k^2}} e^{ik(x+D)} e^{-i\frac{k^2 t}{2m}} dk \quad and \quad \Psi_2(x, t) = N \int_{-\infty}^{\infty} e^{-\frac{(k-k_o)^2}{\Delta k^2}} e^{ik(x-D)} e^{-i\frac{k^2 t}{2m}} dk \tag{5.32}$$

Again, these wavefunctions define an entire class of wavefunctions given by: $\Psi_\alpha = \Psi_1 + e^{i\alpha}\Psi_2$ While different α's yield different wavefunctions, the wavefunction $e^{-i\alpha}\Psi_\alpha$ is equivalent to Ψ_α (since absolute phase cannot be measured). This property, combined with the assumption that Ψ_1 and Ψ_2 do not overlap, means that no local experiment can measure α: i.e., if an apparatus only interacts with Ψ_2, it cannot determine α, unless it is later brought into contact with an apparatus which has interacted with Ψ_1. We could also wait a time $t > \frac{D}{c}$ for the packets to overlap without requiring 2 measuring devices (where $2D$ is the separation between the packets), thereby producing interference and the possibility of determining α. Let us further assume that we can neglect the spread of the wavepackets during the time it takes for the packets to become completely overlapping. Then, putting $k - k_o = \bar{k}$ and $t = T = \frac{Dm}{k_o}$, we have:

$$\Psi_1(x, t) = N \int_{-\infty}^{\infty} \exp\left\{-\frac{\bar{k}^2}{2(\Delta k)^2} - i\frac{\bar{k}^2 T}{2m} + ik_o \frac{D}{2} + i\bar{k}x\right\} d\bar{k} e^{-ik_o x} = F(x)e^{ik_o x} \tag{5.33}$$

Similarly for Ψ_2. We consider now the two states

$$X_1 = \frac{1}{\sqrt{2}}\left\{\Psi_1(x, t) + \Psi_2(x, t)\right\} \quad and \quad X_2 = \frac{1}{\sqrt{2}}\left\{\Psi_1(x, t) - \Psi_2(x, t)\right\} \tag{5.34}$$

We have

$$X_1 = \frac{1}{\sqrt{2}}\left\{F(x)e^{ik_o x} + F(x)e^{-ik_o x}\right\} = \frac{2}{\sqrt{2}}F(x)\cos k_o x \tag{5.35}$$

Similarly, $X_2 = \frac{2i}{\sqrt{2}}F(x)\sin k_o x$. We want to find a method to distinguish between X_1 and X_2 that is analogous to the Stern-Gerlach experiment on a spin-1/2 particle, i.e. that makes X_1 move "up" and X_2 move "down".

The method that achieves this is the following: at the time when the complete overlap occurs we adiabatically switch on a weak potential of the form: $V(x) = A\cos k_o x$. This makes the state X_1 go into a new state X_1' which is an eigenstate of the Hamiltonian $H = \frac{p^2}{2m} + V(x)$. Similarly $X_2 \to X_2'$ which is a different eigenstate of H. The periodicity of the potential gives rise to the well known energy bands (see fig. 5) and X_1' and X_2' correspond to the points 1 and 2 in the (k, ω) diagram. We now add a constant force in the x-direction. This will shift the states X_1' and X_2' from the points 1 and 2 to the points 1 and 2. Since the group velocity $\frac{dw}{dk}$ is negative at 1 and positive at 2, the two states will be pushed in opposite directions in space. Thus, we have a situation analogous to that of the Stern-Gerlach experiment in which the spin-up receives an "upward" kick and the spin-down receives a "downward" kick.
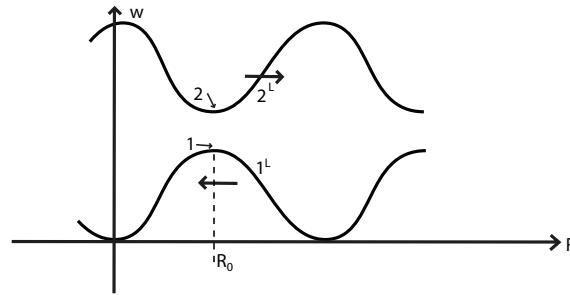
Figure 5.

5.2. Method 2: weak measurement at the time of the nonlocal interaction

What are the general issues involved in the second method to measure dynamical nonlocality?

First: if we start with the state $\psi_{\mathrm{R}} + \psi_{\mathrm{L}}$, 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 nonlocally by the other slit. Therefore we need to start with a state which 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 un-observable. How can we get around this fact in order to observe this nonlocality?

Third: if we are able to get around this fact, then how is causality not violated?

Weak measurements allows 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 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 criteria. 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.

5.2.1. Information gain without disturbance: safety in numbers

So far, Theorem I describes a completely general geometric property. How is it related to experiments? To actually make a measurement of an observable $\hat{A}$, we switch on an interaction [39] $H_{\mathrm{int}} = \lambda g(t)\hat{q}\hat{A}$ 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 $\langle\hat{A}\rangle \equiv \langle\Psi|\hat{A}|\Psi\rangle$ which appears in Theorem I, can be measured in three distinct ways [24, 21]:

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 $H_{\mathrm{int}} = -\lambda(t)\hat{q}\hat{A}$ 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 I to show that the system state is:

$$
\begin{aligned}
e^{-i\lambda\hat{q}\hat{A}}|\Psi_{\mathrm{in}}\rangle &= (1-i\lambda\hat{q}\hat{A})|\Psi_{\mathrm{in}}\rangle \\
&= (1-i\lambda\hat{q}\langle\hat{A}\rangle)|\Psi_{\mathrm{in}}\rangle - i\lambda\hat{q}\Delta\hat{A}|\Psi_{\mathrm{in}\perp}\rangle
\end{aligned}
\tag{5.36}
$$

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

$$\frac{1 + \lambda^2\hat{q}^2\langle\hat{A}\rangle^2}{1 + \lambda^2\hat{q}^2\langle\hat{A}^2\rangle} \longrightarrow 1 \quad (\lambda \to 0) \tag{5.37}$$

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

$$\frac{\lambda^2\hat{q}^2\Delta\hat{A}^2}{1 + \lambda^2\hat{q}^2\langle\hat{A}^2\rangle} \longrightarrow 0 \quad (\lambda \to 0) \tag{5.38}$$

The final state of the measuring device is now a superposition of many substantially overlapping Gaussians with probability distribution given by $Pr(p_q) = \sum_i |\langle a_i|\Psi_{\text{in}}\rangle|^2 \exp\left\{-\frac{(p_q - \lambda a_i)^2}{2\Delta p_q^2}\right\}$. 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 eq. 5.38 that the probability for a collapse decreases as $O(\lambda^2)$, but the measuring device's shift grows linearly $O(\lambda)$, so $\delta p_q = \lambda a_i$ [21]. 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 becomes less precise because the shift in the measuring device is much smaller than its uncertainty $\delta p_q \ll \Delta p_q$ (see fig 6).

3.   Non-statistical method without disturbance   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 $\hat{A}_i$ 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 I to the N particle product state $|\Psi^{(N)}\rangle = |\psi\rangle_1|\psi\rangle_2....|\psi\rangle_N$ with all particles in the same state $|\psi\rangle$. We see that:

$$\hat{A}^{(N)}|\Psi^{(N)}\rangle = \frac{1}{N}\left[N\langle\hat{A}\rangle|\Psi^{(N)}\rangle + \Delta A\sum_i |\Psi_\perp^{(N)}(i)\rangle\right] \tag{5.39}$$

where $\langle\hat{A}\rangle$ is the average for any one particle and the N states $|\Psi_\perp^{(N)}(i)\rangle = |\psi\rangle_1|\psi\rangle_2...|\psi_\perp\rangle_i...|\psi\rangle_N$ 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 eq. (5.39) 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 $N \to \infty$, $|\Psi^{(N)}\rangle$ becomes an eigenstate of $\hat{A}^{(N)}$ with value $\langle\hat{A}\rangle$ and not even a single particle has been disturbed (as $\hat{N} \to \infty$).

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 $\lambda \gg 1$, allowing us to distinguish the result by having the shift, $\lambda\langle\hat{A}\rangle$, exceed the width $\Delta p_q = 1$ of the distribution of the pointer. 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 $\frac{1}{N}$. When N is very large, the coupling to individual systems is very weak, and in the limit $N \to \infty$, the coupling approaches zero. Although the probability that a measurement will disturb any member of the ensemble approaches zero as $\frac{1}{N}$, nevertheless, information about the average is obtained.

5.2.2. **Pre-selection, post-selection and weak measurements**   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. towards
the future.   Nevertheless, a positive spin was placed on quantum mechanic's non-trivial
relationship between initial and final conditions by Aharonov, Bergmann and Lebowitz (ABL) [5]
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 re-formulate
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." Post-selection has also been a very useful technique (see fig. 6.a). For
example, 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 about the (post-selected) direction of motion
for each particle.

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 which has been confirmed in numerous weak measurement
experiments. Weak values have led to quantitative progress on many questions in the foundations
of physics [10] [11] including interference, etc. [13] [28] field theory, in tunneling, in quantum
information such as the quantum random walk, in foundational questions, in the discovery of
new aspects of mathematics, such as Super-Fourier or super-oscillations[7].

What is the weak value? By adding a post-selection to the ordinary -yet weakened- von
Neumann measurements mentioned in §5.2.1, the measuring device will register a weak value [7]:

$$\hat{A}_w = \frac{\langle \Psi_{\text{fin}} | \hat{A} | \Psi_{\text{in}} \rangle}{\langle \Psi_{\text{fin}} | \Psi_{\text{in}} \rangle} \tag{5.40}$$

with $|\Psi_{\text{in}}\rangle$ and $|\Psi_{\text{fin}}\rangle$ 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. [10, 29]). 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.

Eq. 5.40 can also be motivated by inserting a complete set of states $\{|\Psi_{\text{fin}}\rangle_j\}$ into $\langle \hat{A}\rangle$

$$\langle\hat{A}\rangle=\langle\Psi_{\text{in}}|\hat{A}|\Psi_{\text{in}}\rangle=\sum_j |\langle\Psi_{\text{fin}}|_j \Psi_{\text{in}}\rangle|^2 \underbrace{\frac{\langle\Psi_{\text{fin}}|_j \hat{A} | \Psi_{\text{in}}\rangle}{\langle\Psi_{\text{fin}}|_j \Psi_{\text{in}}\rangle}}_{A_{\text{w}}^j \equiv weak\ value} \tag{5.41}$$

with $|\Psi_{\text{fin}}\rangle_j$ 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

---

[7] Super-fourier phenomenon have also been useful in analyzing tachyons [3]. Sudarshan is a member of the lineage
of Fourier: Marshak→ Bethe → Sommerfeld → Lindemann → Klein → Lipschitz → Dirichlet → (Fourier and
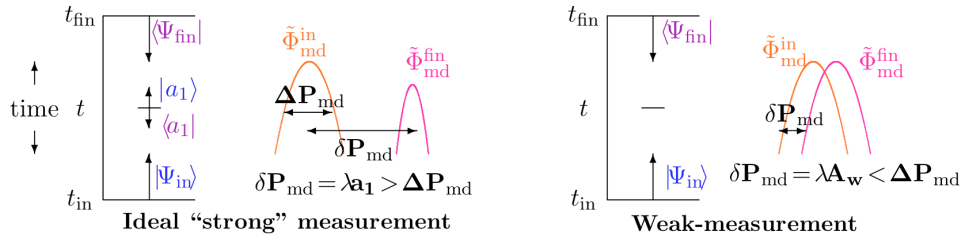Poisson)→ Lagrange → Euler → Bernoulli →Leibniz

Figure 6. a) with an ideal or "strong" measurement at t (characterized e.g. by $\delta P_{md} = \lambda a_1 \gg \Delta P_{md}$), then ABL gives the probability to obtain a collapse onto eigenstate a_1 by propagating $\langle\Psi_{fin}|$ backwards in time from t_{fin} to t and $|\Psi_{in}\rangle$ 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 $t \in [t_{in}, t_{fin}]$; b) if a weak-measurement is performed at t (characterized e.g. by $\delta P_{md} = \lambda A_w \ll \Delta P_{md}$), then the outcome of the weak-measurement, the weak-value, can be calculated by propagating the state $\langle\Psi_{fin}|$ backwards in time from t_{fin} to t and the state $|\Psi_{in}\rangle$ 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.

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 $|\Psi_{fin}\rangle_j$.

To see more precisely how the weak value arises naturally from this weakened measurement with post-selection, we consider the final state of the measuring device after the above described procedure described in the third "non-statistical" method:

$$
\begin{aligned}
|\Phi_{fin}^{MD}\rangle &= \prod_{j=1}^{N} \langle\Psi_{fin}|_j \exp(\frac{-i\lambda}{N}\hat{q}\sum_{k=1}^{N}\hat{A}_k)\prod_{i=1}^{N}|\Psi_{in}\rangle_i|\Phi_{in}^{MD}\rangle \\
&= \prod_{j=1}^{N}\langle\Psi_{fin}|_j \exp(\frac{-i\lambda}{N}\hat{q}\hat{A}_j)|\Psi_{in}\rangle_j|\Phi_{in}^{MD}\rangle
\end{aligned}
\tag{5.42}
$$

Since the particles do not interact with each other, we calculate one term and take the result to the Nth power. (In the following, we substitute the parity operator, $\hat{P}$, for $\hat{A}$.) Using $\hat{P}^2 = 1$, eq. 5.42 becomes:

$$
\begin{aligned}
|\Phi_{fin}^{MD}\rangle &= \left\{\langle\Psi_{fin}|(\cos\frac{\lambda\hat{q}}{N} - i\hat{P}\sin\frac{\lambda\hat{q}}{N})|\Psi_{in}\rangle\right\}^N|\Phi_{in}^{MD}\rangle \\
&= (\langle\Psi_{fin}|\Psi_{in}\rangle)^N\left\{\cos\frac{\lambda\hat{q}}{N} - iP_w\sin\frac{\lambda\hat{q}}{N}\right\}^N|\Phi_{in}^{MD}\rangle
\end{aligned}
\tag{5.43}
$$

$$
\approx \left\{1 - iP_w\frac{\lambda\hat{q}}{N} + \cdots\right\}^N|\Phi_{in}^{MD}\rangle \approx \exp(-i\lambda\hat{q}P_w)|\Phi_{in}^{MD}\rangle
\tag{5.44}
$$

The first bracket of eq. 5.43 can be neglected since it does not depend on $\hat{q}$ and thus can only affect the normalization. Eq. 5.44 represents a shift in the pointer by the weak value, P_w, i.e. $\Phi_{in}^{MD}(p_q) \rightarrow \Phi_{fin}^{MD}(p_q - \lambda P_w)$.

5.2.3. **Applying pre- and post-selection and weak measurements to interference phenomenon** In [28], these tools were used to perform measurements of dynamical nonlocality. To briefly summarize our procedure, 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 setup, 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 opened. The slit is open or closed only after the weak measurement has been completed and the results recorded.

All incoming particles are initially in the state $|\psi_{\mathrm{R}}\rangle$, 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 measurements performed prior to the opening or closing of the slits). Nevertheless, [28] predicted that the *earlier* weak measurement of the parity yields $\langle\hat{\mathsf{P}}\rangle_w = 0$ if the left slit is (*later*) closed and $\langle\hat{\mathsf{P}}\rangle_w = 1$ if the left slit is (*later*) opened. The quantum optics experiment suggested in [28] to measure the non-local exchange has been performed successfully [35].

How do we understand these two results? In principle, a weak measurement with a finite number of particles (N) shifts particles from the right slit to left slit so that the evolving wave-packet has $|\psi_{\mathrm{L}}\rangle$ components and therefore may "sense" whether the left slit is open or closed. However, in [28] it was shown that the weak measurement **can shift at most a few of the N particles from the right slit to the left slit.** But, how can the $\sim N$ particles which 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 the left slit open case (in which *each* particle shifts the measuring device by $\langle\hat{\mathsf{P}}\rangle_w = 1$) to the slit-closed case (each particle shifts the measuring device by $\langle\hat{\mathsf{P}}\rangle_w = 0$)?

We do not see any reasonable way to use local interactions at the left-hand slit to account for the different subsequent behavior 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 nonlocal exchange. The use of this conservation principle is one of the crucial features that distinguishes this procedure from observations that are done with ordinary momentum.

6. Discussion

One of the main points of this article is to emphasize the utility of non-disturbing measurements in probing questions concerning quantum reality. In other publications [14], we have emphasized other classes of non-disturbing measurements, such as the weak measurement [24, 25, 21, 27]. Another purpose of this article is to highlight some new aspects of interference. Our use of the Deterministic Set of Operators (DSOs) in the Heisenberg picture lead us to 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). 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 be just right to protect causality.

While DSOs discussed here is a new conceptual point-of-view that has predicted novel, verified effects which **seem** counter-intuitive from the old perspective, the DSO framework is in fact a *re-formulation* of quantum mechanics, just as the Heisenberg and Schrödinger pictures are equivalent formulations of quantum mechanics. Therefore, experiments cannot prove the DSO framework over standard quantum mechanics (or vice-versa). The motivation to pursue such re-formulations, then, depends on their ability to generate new insights, new intuitions, new experiments, etc. In fact, the DSO framework fulfills several criterion which any re-formulation of quantum mechanics should satisfy in order to be useful and interesting:

- DSO is consistent with all the predictions made by standard quantum mechanics,

- DSO has revealed new features and effects of quantum mechanics that were missed before

(for example, DSOs suggests that a fundamental difference between classical mechanics and quantum mechanics is that the equations of motion for observables relevant to quantum mechanical interference phenomena can be dynamically **non-local**),

- DSO has stimulated discoveries in other fields,
- DSO suggested generalizations of quantum mechanics not easily articulated in the old language

6.1. Quantum Reality

We would like to conclude with a brief discussion concerning the relevance of these considerations to questions of the nature of quantum reality.

6.1.1. The Zeno effect of Sudarshan-Misra and its generalization by Aharonov-Vardi

The Zeno effect by Sudarshan and Misra [2] considers systems which naturally evolve away from a particular state. The fact that a rapid (dense) sequence of measurements prevents it from doing so has thus been nicknamed "watched pot doesn't boil." The generalization of the Zeno effect by Aharonov and Vardi (AV) [19] is related to DSOs. For example, the first example involving a spin-1/2 particle placed in a magnetic field $\vec{B}$ pointing in the x-direction and with $\sigma_z = +1$ initially. If at some time t later we rotate our apparatus and measure the spin in the yz-plane in a direction $\alpha = \omega t$, where ω is the Larmor frequency, then the result is predictable and we have a deterministic experiment. AV generated a time evolution determined by the *observer's* choice of a sequence of measurements, rather than the natural evolution of the system. We suggest that AV be called "watched pot with no fire can boil." By way of example, we consider again the spin-1/2 particle rotating clockwise with $\hat{\sigma}_x(\Delta t) = \hat{\sigma}_x(t=0)\cos\omega\Delta t + \hat{\sigma}_y(t=0)\sin\omega\Delta t$. However, we can force the particle to rotate in the *opposite* direction with certainty if measurements in a counter-clockwise rotating measuring device are made rapidly enough. [8]

What can Zeno and AV tell us about quantum reality? Bohr claimed that we cannot ignore the disturbance on the system (since $\hbar$ cannot be reduced to 0) and therefore we cannot assign a reality to a quantum system between classical measurements because there will always be counterfactual paradoxes similar to the Hardy paradox [10]. John Wheeler elegantly describes "observer-participants" by describing the game of 20 questions in its surprise version: [20] "While I was out of the room, they explain, they had agreed not to agree on a word. There was no word in the room when I entered." He concludes, "First, the word already existed in the room – we thought – independent of any question that we might or might not ask. But it didn't. Likewise the electron has a position and a momentum inside the atom – physics once thought – independent of any act of observation. But it doesn't." To what extent is AV, i.e. generation by an observer of an arbitrary time evolution, a manifestion of this idea? One could argue that the outcome of AV cannot be considered as a pre-existing property of the system, rather it was dictated by the choice of measurement of the observer.

6.1.2. The relevance of non-disturbing measurements to the nature of quantum reality

Nevertheless, the basic observer-participancy mentioned above has fundamental underpinnings in quanum mechanics. For example, "contextuality" means that whenever we speak

---

[8] The probability to make the particles rotate in the opposite direction is $\{1 - \frac{1}{N^2}\}^N \to e^{-\frac{1}{N}} \to 1$ as $N \to \infty$ where $\Delta t = \frac{T}{N}$. It can also be used to give the Feynman Path picture a physical meaning [19]. If AV is performed on any of the Feynman paths, then if the verification measurements are dense enough, then the Feynman path can be verified with probability one. If a measuring device is in a superposition such that one part of the superposition verifies one Feynman path and another superposition of the measuring device verifies a different Feynman path, then the relative phase between the 2 paths is the classical action utilized in the Feynman formalism. Therefore, AV can be used to give an operational meaning to Feynman's claim that each path has equal probability.

about an "element of reality" we must also include the conditions of observation. There cannot be any contextuality in a single DSO since (by eq. (2.1)) all the functions of any given member of DSO is also a member of the DSO. However, as shown in [30, 31], if we consider the DSO for the pre-selected state and another DSO for the post-selected state (one for the pre-selected vector and one for the post-selected vector), then by combining these DSOs, we can create a novel kind of disturbance which reveals contextuality. By analyzing the concept of contextuality (Bell-Kochen-Specker) in terms of pre-and-post-selection and DSOs, it is possible to assign definite values to observables in a new way [30, 31]. When measurements are performed which do not disturb the pre- and post-selection (i.e. weak measurements), then novel *experimental* aspects of contextuality can be demonstrated including a proof that every pre- and post-selection-paradox with definite predictions implies contextuality. In [30, 31], it is also shown how certain results of these measurements (eccentric weak values with e.g. negative values outside the spectrum), cannot be explained by a "classical-like" hidden variable theory. Additional connections between DSOs and pre- and post-selected systems are explored in [26]

We emphasize that the DSOs discussed here involve those variables which are relevant for a *single* particle without causing a disturbance and can therefore, in principle, be used to probe quantum reality without the usual reliance on ensembles. The third "non-statistical" method (using the collective observable discussed in §5.2.1) also leads to a broad new class of measureable Hermitian observables involving highly nonlocal properties (in both space and time). For example, consider again the spin-like observables discussed in §4.3 with $\hat{\sigma}^z$ telling us about which slit the particle is centered around, etc. The nonlocal relative phase can then be measured with:

$$\alpha \sim \frac{\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}_i^x}{\sqrt{1 - (\frac{1}{N}\sum_{i=1}^{N}\hat{\sigma}_i^x)^2}} \tag{6.45}$$

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

In the discussions above we referred to quantum states as being simply the mathematical tools for describing the system. However, states also have an ontological dimension. This is a highly debated issue, which even for the simple case of a standard state $|\Psi\rangle$ is very controversial. Does the state have a "reality" of its own, or is it just a mathematical tool for making predictions? Does the state actually collapse, or is the collapse simply our updating the mathematical description following the acquisition of new data (the results of new measurement). Is the state a physical entity (such as in Bohm's pilot wave model)? Discussing the ontological status of the second post-selected vector is bound to be even more controversial. It was not our intention to dwell too much on this issue here other than to emphasize that when applied to the Schroedinger picture, weak measurements dispense of the statistical aspect of the wavefunction (though still requiring an ensemble [21]).

Nevertheless, we conclude with a comment on a significant interpretative issue concerning weak values which was originally pointed out by Sudarshan [1]. Namely, there are two ways to view the "weak value" of an observeable in which the measuring device is shifted by strange values which can be any complex number or can be arbitrarily large: 1) the weak value is a result of a complex interference effect in the measuring device, 2) the measuring device is registering the weak value of the **system** as characterized by the 2-vectors. While this first interpretation of weak values is technically available, in a subsequent publication, we will show that it is often unnatural when compared to the second interpretation which regards the weak value as a property of the system. For example, when the system is post-selected, the measuring device measures a single value the weak value that depends only on the measured system and not on measuring device. With post-selection, there is no superposition of measuring device positions, i.e. there is no need to monitor the quantum phases in the measuring device. In fact the second is the

only interpretation if the measuring device is classical.

References

[1] Duck I M, Stevenson P M, and Sudarshan E C G 1989 Phys. Rev.**D 40**, 2112
[2] Sudarshan E C G, Misra B 1977 J. Math Phys. 18(4) 756-763
[3] Sudarshan E C G, Bilaniuk O M P, and Deshpande V K 1962 Amer. J. Phys. 30, 718-723; Aharonov Y, Reznik, B, Stern B 1998 A. Phys. Rev. Lett. 81 2190-2193
[4] Aharonov Y, Bohm D 1959, Phys. Rev. **115**, 485
[5] Aharonov Y, Bergmann P G, and Lebowitz J L 1964 Phys. Rev. **134**, B1410, reprinted in *Quantum Theory and Measurement*, eds. Wheeler J A and Zurek W H 1983 (Princeton University Press) pp. 680-686
[6] Aharonov Y and Albert D Z 1980 Phys. Rev. **D12** 3316, ibid. 1981 **D24** 359
[7] Aharonov Y, Albert D Z, and Vaidman L 1988 Phys. Rev. Lett. **60**, 1351; Aharonov Y, Vaidman L 1990, Phys. Rev. A, **41**, 11; Aharonov Y, Vaidman L 1991, J. Phys. A **24**, 2315; Aharonov Y, Botero A 2005, Phys. Rev. A 72, 052111; Aharonov Y, Casher A, Albert D, Vaidman L 1987, Phys. Lett. A124, 199
[8] Aharonov Y, Vaidman L 1990 Phys. Rev. A, **41**, 11
[9] Aharonov Y and Vaidman L 1991 J. Phys. A **24**, 2315
[10] Aharonov Y, Massar S, Popescu S, Tollaksen J, and Vaidman L 1996 Phys. Rev. Lett., **77**, p. 983; Tollaksen, J 2007, Jrnl of Phys A: **40** 9033-9066; Aharonov Y, Botero A, Scully M 2001, Zeit Natur Sec A, **56**, Issue: 1-2, pg 5-15; Nussinov SN, Tollaksen J 2008, Phys. Rev. D, **78** 036007
[11] Aharonov Y, Botero A, Popescu S, Reznik B, Tollaksen J 2002, "Revisiting Hardy's paradox: counterfactual statements, real measurements, entanglement and weak values," Phys. Lett. A 301 (3-4): 130-138; This novel effect has been verified in several independent experiments such as Yokota, et al 2009, *New Journal of Physics*, vol 11, 033011 and Lundeen JS, Steinberg AM 2009, *Phys. Rev. Lett.*, 102:020404. This work has also been featured in popular science journals including the Asian edition of Scientific American (August, 2009), New Scientist "They said it couldnt be done but now we can see inside the quantum world," etc.
[12] 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.
[13] Aharonov Y and Rohrlich D 2005 *Quantum paradoxes:* Weinheim Cambridge: Wiley-VCH
[14] Aharonov Y, Tollaksen J 2009, "New insights on Time-Symmetry in Quantum Mechanics," in 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 (in press).
[15] Aharonov Y, in *Proc. Int. Symp. Foundations of QM*, Tokyo, pp. 10-19.
[16] Bohm D, Aharonov Y 1957, Phys. Rev. **108** 1070-1076
[17] Aharonov Y, Vaidman L 1990 Phys. Rev. A, **41**, 11
[18] Aharonov Y and Vaidman L 1991 J. Phys. A **24**, 2315
[19] Aharonov Y, Vardi Y 1980 Phys. Rev. D 21 (8): 2235-2240
[20] Wheeler J 1986 *How Come the Quantum?,"* New Techniques and Ideas in Quantum Measurement Theory, ed. Daniel Greenberger, *Ann. NYAcad. Sci.***755**, 304
[21] Tollaksen, J 2007 *Non-statistical weak measurements, Quantum Information and Computation V*, edited by Eric Donkor, Andrew Pirich, Howard Brandt, Proceedings of SPIE Vol. 6573 (SPIE, Bellingham, WA), CID 6573-33
[22] Tollaksen J 2007 *Novel Relationships between Superoscillations and Modular Variables,* forthcoming in J. Phys. Conf. Series vol. 70, Editors: Brandt H, Kim Y S, and Man'ko M A
[23] Tollaksen J 2007 *Robust Weak Measurements on Finite Samples,* forthcoming in J. Phys. Conf. Series vol. 70,, Editors: Brandt H, Kim Y S, and Man'ko M A, quant-ph/0703038
[24] Tollaksen J 2007 "Robust Weak Measurements on Finite Samples," J. Phys. Conf. Series, vol. 70, 012015, Editors: Brandt H, Kim YS, and Man'ko MM, quant-ph/0703038. (featured in Hosten O, Kwiat P 2008, Science, "Observation of the Spin Hall Effect of Light via Weak Measurements," 319, 787 which was featured in Resch KJ 2008, Science, "Amplifying a Tiny Optical Effect," 319, 733, both of which were featured in Physics Today , April 2008, Day C, "Light exhibits a spin Hall effect," Volume 61, Issue 4, pp. 8-96, see also Popescu S 2009, "Weak measurements just got stronger," Physics 2, 32
[25] Tollaksen J, Ghoshal D 2007 *Weak Measurements, Weak Values and Entanglement, Quantum Information and Computation V*, Ed by Donkor E, Pirich A, Brandt H, Proc of SPIE Vol. 6573 (SPIE, Bellingham, WA), CID 6573-36
[26] Tollaksen J 2007, "Novel Relationships between Superoscillations and Modular Variables," Jrnl Phys: Conf, 70, Inst of Phys, 012016, Editors: H. Brandt, Y. S. Kim, and M. A. Man'ko; Aharonov T, Kaufherr T 2004, Phys Rev Lett 92, 7, 070404
[27] Tollaksen J 2007 *Quantum properties that are extended in time, Quantum Information and Computation V*, Ed by Donkor E, Pirich A, Brandt H, Proc of SPIE Vol. 6573 (SPIE, Bellingham, WA), CID 6573-35

[28] Tollaksen J, Aharonov Y, Casher A, Kaufherr T, Nussinov N 2004, "Quantum interference experiments, modular variables and weak measurements," arXiv:0910.4227.

[29] Tollaksen J 2001 PhD thesis, Boston University; thesis advisor: Yakir Aharonov.

[30] Tollaksen J 2007 *Probing Contextuality with pre-and-post-selection,* forthcoming in J. Phys. Conf. Series vol. 70,, Editors: Brandt H, Kim Y S, and Man'ko M A

[31] Tollaksen J 2007 *Pre-and-post-selection, weak-values, and contextuality,* to appear in Journal of Physics A: Mathematical and General; quant-ph/0602226

[32] Feynman R , Leighton R, Sands M 1965, The Feynman Lectures on Physics, Vol. III, AddisonWesley, p. 19.

[33] Popescu S, Rohrlich D 1992, "Generic quantum nonlocality," Phys. Lett. A 166, 293

[34] Scully MO, Englert BG, and Walther H 1991, Nature 351, 111; Storey EP, Tan SM, Collett MJ, and Walls DF 1994, Nature 367, 626; Wiseman H 2003, *Physics Letters A*, 311, 285; Mir R, Lundeen JS, Mitchell MW, Steinberg AM, Garretson, Wiseman H 2007, New Journal of Phys, **9**, 287

[35] Gray, Parks, Spence, Tollaksen, Troupe, forthcoming.

[36] Shimony A 1983, in *Foundations of Quantum Mechanics in Light of New Technology*, Proc of the Intnl Symp, Tokyo, eds. Kamefuchi S, Ezawa H, Murayama Y, Namiki M, Nomura S, Ohnuki Y, Yajima T 1984, (Tokyo: Phys Soc of Japan), pp. 22530; cf. p. 227; Shimony also states: "the first confirmation of entanglement ... antedated Bell's work, since Bohm and Aharonov [16]" see Shimony A "Bells Theorem," *Stanford Encyclopedia of Philosophy*.

[37] Feynman RP, Hibbs AR 1965, Quantum Mechanics and Path Integrals (New York: McGraw Hill) , pp. 29; Tonomura A, Endo J, Matsuda T, Kawasaki T, Ezawa H 1989, Am. J. Phys. 57 117.

[38] Batelaan N, Tonomura A 2009, "Testing Aharonov-Bohm effects," *Physics Today*.

[39] von Neumann J 1983 *Mathematical Foundations of Quantum Theory*, Princeton Univ Press, New Jersey