New Perspectives on the Aharonov-Bohm Effect

Mostafa E. El Demery (memrae2@cam.ac.uk)
St Edmund's College

Supervised by: Dr. Berry Groisman
DAMTP, Cambridge

03 May 2013

Submitted in partial fulfillment of a Master of Advanced Studies (MASt) at Cambridge

Abstract

The Aharonov-Bohm effect is a quantum mechanical effect, that is, has no classical counterpart. The effect was predicted in 1959 in a seminal paper of Y. Aharonov and D. Bohm [AB59] in which they demonstrated that a beam of electrons is affected by the existence of the electric/magnetic field even though electrons travel through field-free regions. Aharonov and Bohm carried out two hypothetical experiments to support their claim that potentials are more fundamental than fields and they are responsible of the effect. Since then, the debate has arisen around whether potentials are mathematical tools or fundamental entities in physics. Different arguments have been set to explain the results predicted by Aharonov and Bohm and experimentally confirmed. Amongst these arguments, the first argument adopted by Aharonov and Bohm was that potentials are physically significant. Many have claimed that fields do have non-local features, i.e. action at a distance. Others have claimed that topological effects may interpret the effect in which potentials are modeled as connections in higher-dimensional fiber bundle geometries. The most recent argument has been proposed by Vaidman [Vai12] who claimed that the the composite system is represented by one state, an entangled state, and due to the electromagnetic interactions part of this state is changed, hence, the total state. In the present essay, I will discuss the latter argument as well as reviewing some other arguments.

Contents

1. Introduction

Since it was first discovered in 1959, the Aharonov-Bohm (AB) effect [AB59] has raised a lot of discussions concerning our understanding of the foundations of the quantum theory. Much of these discussions has been dedicated to the role and significance of electromagnetic potentials in the quantum theory, in other words, whether electromagnetic potentials do exhibit measurable effects or are just mathematical auxiliaries. In addition, the notion of non-locality, in other words action at a distance, in the quantum theory has been revisited after it was first proposed by Einstein, Podolsky and Rosen [EPR35] in 1935. The AB effect has led to a strong debate and divided the community between supporters and opponents of the idea that potentials are the fundamental entities in physics rather than fields. Different interpretations have been given in order to fully explain the striking results predicted by Aharonov and Bohm [AB59]. After nearly half a century new perspectives [Pop10, Vai12] are still represented in order that full understanding may be founded. Though its predictions are critically important as well as seemingly controversial, the effect itself is surprisingly simple[1].

a- A beam of electrons passes through a Young's double slit apparatus and produces interference patterns.

b- A magnet, or any other source of magnetic flux, is placed behind the slitted barrier in such a way that it does not intercept the paths of the electrons. The magnetic flux, however, will change the phases of the electrons' wave functions causing the interference patterns to be disturbed.

c- The magnetic flux is blocked by surrounding the magnet by an extremely effective shield in order to leave the interference patterns undisturbed. Aharonov and Bohm predicted, however, that the interference patterns will be distorted because of the existence of the magnetic flux even though the electrons pass through field-free regions.

Such predictions for their controversy have raised two possible standpoints. Firstly, electromagnetic potentials do exhibit measurable effects and are no longer mathematical tools which means that they may be more fundamental than electromagnetic fields. Secondly, electromagnetic fields might have non-local action, i.e. action at a distance which was less favored by physicists before Bell's theory [Bel64] in 1964. In fact, great efforts, both theoretical and experimental, have been made to critique and/or confirm the existence of the AB effect. The first experimental work to report the detection of the AB effect was carried out by Chambers [Cha60]. Motivated by Chambers' work and in order to answer the arguments raised against the conclusion of their first paper, Aharonov and Bohm [AB61] extended their treatment to include the sources of the potentials quantum mechanically asserting for their previous results [AB59]. Many authors questioned the accuracy of the experiments carried out claiming that the field lines were not completely shielded. However, from a purely theoretical point of view, the debate around the existence of the AB effect waged out and many authors claimed that Aharonov and Bohm are in error [BL78, Hen81]. Such view points in addition to others were reviewed by Olariu and Popescu [OP85]. After the first precise experiment carried out by Tonomura and his team at Hitachi using electron holography [TMS+82] followed by his work [TOM+86] using superconducting shields, almost everyone agrees that the AB effect is confirmed and that it is a genuine feature of the standard quantum mechanics [PT89].

---

[1]In the present context we will give a simple description of the AB effect, however the thought experiments carried out by Aharonov and Bohm [AB59] will be described in detail in the next chapter.

From a practical point of view, the AB effect proved to be very important for many applications to electron phase microscopy [cAT06]. Now the debate around the existence of the AB effect is settled down, however, a debate around its interpretation is continuously active. In the present essay, I will try to shed some light on one of these interpretations, the Vaidman perspective [Vai12]. The aim of this essay is to discuss and review the AB effect from a conceptual point of view, therefore we will focus on the philosophical sides rather than the mathematical derivations.

The present essay is organized as follows. The next chapter will give a brief description of the AB effect. Both the magnetic AB and electric AB effects will be presented in the light of the thought experiments proposed by Aharonov and Bohm [AB59]. The phenomenological features of the AB effect will be briefed. Moreover, different objections will be highlighted and the position of Aharonov and Bohm will be summarized over a definite period of time. In the third chapter, the Vaidman argument [Vai12] will be discussed and set to critique. Conclusions will be drawn at the final chapter.

2. The Aharonov-Bohm Effect

The AB effect is a quantum-mechanical phenomenon, i.e. does not have a classical counterpart, which claims that charged particle are affected, in the form of a phase shift, by the existence of electromagnetic fields even though the charged particles travel through field-free regions.

2.1 Magnetic AB Effect

In their original paper published in 1959, Aharonov and Bohm [AB59] carried out thought experiments in order to show that electrons are affected by electromagnetic fields even if they travel through field-free regions. In this section I will present the magnetic AB effect. Consider the following interference setup, shown in Fig. 2.1. A beam of electrons enters from the left and is split coherently into two beams in a two-arm interferometer. In principle, any change in the relative phase between the beams in the two arms can be observed as a shift in the interference pattern when the two beams are gathered at the right. As a source of magnetic flux, an electric current flows in a cylindrical solenoid with radius R, center at the origin and axis along the z direction. A magnetic field $\mathbf{B}$ is created and confined within the solenoid, while the vector potential $\mathbf{A}$ will be non-zero everywhere outside the solenoid. The total magnetic flux through every circuit containing the origin is

$$\Phi = \int \mathbf{B} \cdot d\mathbf{s} = \int \mathbf{A} \cdot d\mathbf{x}. \tag{2.1.1}$$

The Hamiltonian H of such setup is given by

$$H = \frac{1}{2m} \left[-i\hbar \nabla + \left(\frac{e}{c} \right) \mathbf{A} \right]^2. \tag{2.1.2}$$
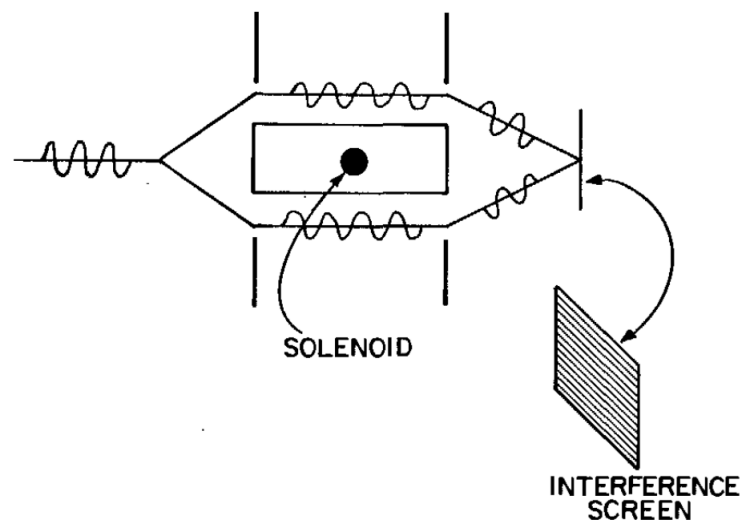

Figure 2.1: Magnetic AB effect. The axis of the solenoid is perpendicular to the page. Taken from [PT89].

3

The wave function associated with this Hamiltonian, if we are in singly connected regions, would be $\psi = \psi_0 exp[-iS/\hbar]$ where $S = (e/c) \int \mathbf{A} \cdot d\mathbf{x}$. Since we are in multiply connected regions, ψ is no longer single-valued which can not represent a solution for the Schrödinger equation. However, such solution can be used for each beam of electrons, upper and lower, which stays in a simply connected region and the total wave function is $\psi = \psi_1 + \psi_2$. Therefore, we can write down

$$\psi_1 = \psi_1^0 e^{-iS_1/\hbar}, \psi_2 = \psi_2^0 e^{-iS_2/\hbar}, \tag{2.1.3}$$

where S_1 and S_2 are equal to $(e/c) \int \mathbf{A} \cdot d\mathbf{x}$ along the paths of the first and second beams, respectively. The interference between the two beams depends upon the phase difference between them, therefore the AB phase is expressed as

$$\phi_{AB} = \frac{1}{\hbar}(S_1 - S_2) = \frac{e}{\hbar c} \int \mathbf{A} \cdot d\mathbf{x} = \frac{e}{\hbar c}\Phi. \tag{2.1.4}$$

2.2   Electric AB Effect

Aharonov and Bohm hypothetically proposed the following setup as shown in Fig. 2.2. A beam of electrons is split into two coherent beams each of which passes through ideal conducting tubes which shield the electrons from the electric field. Using electrical shutters, the beams are chopped into wave packets that are long relative to the wavelength λ and short relative to the length of the tubes. Voltages $V_{e1}(t)$ and $V_{e2}(t)$ are set on the two tubes during a limited time interval when the wave packets are inside the tubes.
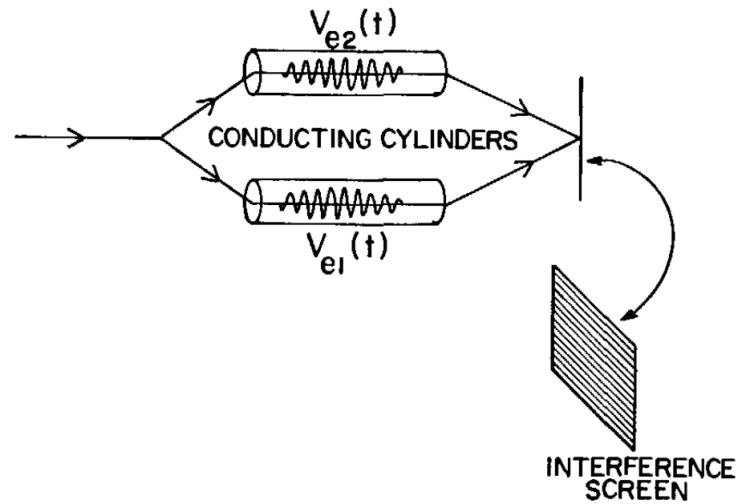

Figure 2.2: Electric AB effect. $V_{e1} = V_{e2} = 0$ except when the electrons' wave packet is shielded from the electric field. Taken from [PT89].

The total Hamiltonian of the present system is

$$H = H_0 - eV(x, t), \tag{2.2.1}$$

where $H_0 = -(\hbar^2/m)\nabla^2$ is the free Hamiltonian and the solution for the wave function attached to this Hamiltonian is

$$\psi = \psi_0 e^{-iS/\hbar} \tag{2.2.2}$$

where ψ_0 is the wave packet in the absence of the external potential $V_e(x,t)$ and $S = -e \int V_e(x,t)dt$. The total wave function is the sum of the solutions for each beam, such that

$$\psi_1 = \psi_1^0 e^{-iS_1/\hbar}, \psi_2 = \psi_2^0 e^{-iS_2/\hbar},$$
$$\psi = \psi_1 + \psi_2. \tag{2.2.3}$$

When the two packets interferat some time t after the potential has returned to zero everywhere, thier relative phase is shifted by the AB phase

$$\phi_{AB} = \frac{1}{\hbar}(S_1 - S_2) = \frac{e}{\hbar c} \oint V dt, \tag{2.2.4}$$

where $V = |V_{e1} - V_{e2}|$ is the potential difference between the two cylinders.

2.3   Phenomenology of the AB Effect

The AB effect described above has been extensively studied in order to a plausible explanation for it. Different theories were proposed in order to explain the disturbance of the interference pattern predicted by Aharonov and Bohm [AB59]. For every point of view, a tedious mathematical work has been done in order to account for this prediction. Since this essay is conceptually oriented, I shall concentrate on the physics behind every point of view and the reader is urged to review the mathematical work where referred. Before doing so, I need to summarize the phenomenological characteristics of the AB effect. Much of the work done in the present section and the comming sections is covered in Chapter 3 in [Ken93].

2.3.1 Phase shift of the interference Fringes. The main result of the AB effect is the shift in the interference fringes due to the existence of the fields. The AB phase in the magnetic and electric versions of the AB effect is given by Eqs. (2.1.4) and (2.2.4) respectively and measure the total phase difference accumulated by the two beams as they pass by the enclosure.

2.3.2 Stability of the Interference Envelope. Boyer [Boy73b, Boy73a] has shown that the envelope of the interference pattern is stable, and that the interference fringes, i.e, the minima and maxima of the pattern, shift without disturbing the envelope as a whole. This is important since the averages of all dynamic variables for the system depend only upon the envelope.

2.3.3 Fringe Shifts Independent of the Effectiveness of Shielding. The AB effect has been observed in a wide variety of the experiments [OP85]. If the fringe shifts were due to flux leakage, the size of the shifts should diminish with rising effectiveness of the shields which has not been observed.

2.3.4 Shifts Persist with Near-Ideal Shields. In his experiment, Tonomura [TOM⁺86] used several layers of shielding to prevent overlap between the enclosed flux and the passing electrons. In addition

to copper shielding, superconducting shields were used to surround a flux-bearing torus. Since a field line cannot cross a superconductor, Tonomura concluded that the ineliminable but minute field leakage could not produce the observed fringe shifts.

2.4   Timeline of the AB Effect 1959 - 1962

(i) *Aharonov and Bohm, 1959*

In [AB59], Aharonov and Bohm tried to demonstrate the existence of the effect described earlier. Their calculations served their central claim that the potentials are responsible for the effect, as the title indicates: "Significance of Electromagnetic Potentials in the Quantum Theory." Since the discussion below will be concerned to trace the modifications in Aharonov and Bohm's Central argument, a detailed statement of the argument is needed here.

(a) Existence of AB Effect. "An electron (for example)" can be influenced by the potentials even if all the field regions are excluded from it. . . "

(b) Local Field Theories. "According to current relativistic notions, all fields must interact only locally. And since the electrons cannot reach the regions where the fields are, we cannot interpret such effect as due to the fields themselves." The concept of local interactions is not here precisely defined.

(c) Ineliminability of the Potentials. The paper begins with a very important statement: "It is true that in order to obtain a canonical formalism, the potentials are needed. . . In quantum mechanics, however, the potentials cannot be eliminated from the basic equations." In classical electrodynamics, it is shown that the mathematical trick of introducing a generalized potential (whose interpretation was unclarified) would preserve the form of Lagrange's equations for electromagnetism only if the Lorentz force were re-written in terms of the potentials. Aharonov and Bohm's claim that the potentials are necessary rests upon this procedure for constructing the Hamiltonian.

(d) Potentials are Fundamental. Since the fields are disqualified by the locality requirement and the potentials appear indispensable, the AB effect suggests quite a strong conclusion:

> "It would seem natural at this point to propose that, in quantum mechanics, the fundamental physical entities are the potentials, while the fields are derived from them by differentiations."

In classical mechanics, the potentials had been (after Hertz and Heaviside "purely mathematical auxiliaries." In a complete reversal, Aharonov and Bohm suggest the fields are now mere manifestations of the potentials.

(e) Gauge Dependence Objection. Aharonov and Bohm make two responses to the objection that the potentials, as partly arbitrary gauge-dependent quantities, cannot represent physical magnitudes. The first is merely to quote again the need for local interactions. Only the potentials permit local descriptions, and therefore must be accepted despite their arbitrariness. More helpfully, they suggest secondly that different gauges may have undiscovered observable consequences. In fact, the tolerance of gauge dependence is the cost of the locality requirement.

(f)Summery. The 1959 paper which announced the AB effect initiated a tradition of ascribing a new physical significance to the electromagnetic potentials. At this stage, their arguments rest on only two important ideas: locality and the ineliminability of the potentials.

(ii) *Aharonov and Bohm's Attack on Forces, 1961*

In order to answer the arguments raised against the conclusion of their first paper, Aharonov and Bohm [AB61] extended their treatment to include the sources of the potentials quantum mechanically asserting for their previous results [AB59]. The central claim is again the new significance of the potentials necessitated by locality:

> "… we see that the fields in the excluded region cannot be regarded as interacting directly with the electron. Indeed, this interaction goes by the intermediary of the potentials; and it is only when this is taken into account that the essential feature of the locality of interaction of electromagnetic fields with the electron can be brought properly into the theory." [AB, 1961; 1522 - 23]

Aharonov and Bohm begin by noting that the Bohr-Sommerfeld rule for eigenvalues

$$\oint \mathbf{p} \cdot d\mathbf{q} = n\hbar, \tag{2.4.1}$$

clearly cannot be the result of electromagnetic forces. In fact, for non-zero vector potential $\mathbf{A}$ the canonical momentum in Eq. (2.4.1) can be rewritten to give

$$\oint [m\dot{\mathbf{x}} + \frac{e}{c}\mathbf{A}] \cdot d\mathbf{x}, \tag{2.4.2}$$

which shows that the extra ingredient determining the behavior of the electron is just the potentials themselves.

By slightly shifting from the above-mentioned line of argument, Aharonov and Bohm further noted that the calculation of the average quantum force with Eherenfest's theorem

$$\frac{m d\mathbf{r}}{dt} = \mathbf{f} = \int \psi^* \left[-\nabla V + e\nabla\phi - \left(\frac{e}{c}\right)(\mathbf{r} \times \mathbf{B}) \right] \psi d\mathbf{r}, \tag{2.4.3}$$

requires knowledge of the psi-function. This, however, requires in turn construction of the Hamiltonian and the attendant introduction of the potentials. Thus even calculations of the quantum equivalent of the Lorentz force demands use of the potentials, and forces therefore seem less fundamental than in classical mechanics. Therefore, Aharonov and Bohm concluded:

> "Thus, in quantum mechanics, force is an extremely abstract concept, having at best a highly indirect significance, which is of only secondary importance… It is clear that … the electromagnetic potentials (and not the fields) are what play a fundamental role in the expression of the laws of physics."

(iii) *Challenging the Ineliminability of the Potentials*

In [AB59], constructing the Euler-Lagrange equations led Aharonov and Bohm to claim that the potentials could not be eliminated from the quantum theory. In 1962 no less than four papers appeared, all denouncing this claim: DeWitt [DeW62], Belinfante [Bel62], Noerdlinger [Noe62], and Mandlestan

[Man62]. DeWitt and Mandlestan independently suggested that the minimal coupling rule through which the potentials enter the Schrdinger equation can be rewritten

$$\frac{\partial}{\partial x^\mu} \rightarrow -ie \int_{-\infty}^{0} F_{\nu\sigma}(z) \frac{\partial z^\nu}{\partial \zeta} \frac{\partial z^\sigma}{\partial x^\mu} d\zeta \tag{2.4.4}$$

so that a path integral indexed by ζ replaces the potential. Mandlestan used this idea to produce a gauge-invariant formulation of QED.

Noerdlinger [Noe62] suggested the possibility of eliminating the potentials by the straightforward substitution

$$A_\mu = \int \left[\frac{j_\mu}{r}\right] d^3x \tag{2.4.5}$$

where the brackets indicate retardation. Noerdlinger suggested that the Hamiltonian

$$H = \frac{1}{2m} \frac{e}{4\pi} \int \left[\frac{\nabla \cdot \mathbf{E}}{r}\right] d^3x + \left[P - \frac{e}{4\pi c} \int \left[\frac{\nabla \times \mathbf{B}}{r}\right] d^3x\right]^2 \tag{2.4.6}$$

introduces a "non locality" into the theory. It means that one physical quantity, A_μ, is a function of physical quantities outside the neighborhood of x. Thus he can conclude, "the non-locality in the theory only relates present events to those in the backward light come, i.e., no violation of causality is implied." Thus Noerdlinger's suggestion leaves open the question of whether a description of the physical situation in a local neighborhood can account for all local physics.

To summarize these important exchanges of 1962:

(a) Potentials were proved eliminable.

(b) A "physical significance" to the potentials became the locality requirement.

(c) The choice between accepting non-locality or elevating the potentials came into sharp focus.

(d) As locality is distinguished from causality, the physical basis for a locality requirement seems less secure and optional to future developments.

(iv) *The Principle of Localizability , 1963*

After Belinfante [Bel62], Aharonov and Bohm [AB63] restated their position toward locality:

> The main point that we wished to stress in our earlier paper was that when QED is expressed in terms of the potentials, one obtains a complete set of local commuting [operators] and a set of purely local equations of motion. It was certainly not our intention, of course, to insist that a correct theory must be local in this sense. Indeed, we explicitly pointed out that physics is currently in a state of flux, such that there are even good reasons to suppose that this requirement of locality must somehow eventually either be changed or given up altogether.

The significance of the potentials is now reduced to the fact that they appear in such local forms of QED. Aharonov and Bohm shift their interpretation of "locality" from being a "requirement determining the

from" of acceptable theories to a mathematical characteristic of "current QED theories." They named it the new *principle of localizability*.

Aharonov and Bohm defended this principle on grounds which go beyond the common intuitions at its basis:

> In this way one obtains a clear and natural mathematical expression for the common "intuitive" notion that the world can be analysed into "local" entities (i.e., the fields and charges) interacting by a sort of "contact" (which notion is , in fact, implicit in modern field theories).

(v) *Summary*

The evolution of Aharonov and Bohm's support for the potentials theory of the AB effect may be drawn as:

(a)1959: The AB effect shows that the potentials are the fundamental entity in electromagnetism . They cannot be eliminated from the formalism and reflect the locality of physical interactions.

(b)1962: The potentials can be eliminated , but may be involved in future theories , whether local or non-local.

(c)1963: Potentials exhibit a mathematical feature of QED (which may be non local): the theory is capable of transformation back to a local form.

The nature of "locality" for Aharonov and Bohm underwent a parallel evolution:

(a) 1959: All physical interactions must be local.

(b)1962: Locality is a "requirement" determining the form of physical theories, which may have to be abandoned.

(c)1963: The principle of localizability corresponds to a mathematical characteristic of non-local theories which guarantees compatibility with local commutation relations.

To sum up, Aharonov and Bohm effectively abandoned the potentials theory and accept non-local theories which satisfy an abstract constraint. The principle of localizability is a puzzling feature of the quantum theory which gathers non-local interactions to local causality and Lorentz invariance.

2.5   Non-local Effects

2.5.1 Greenberger's Argument for Non-Locality. Greenberger [Gre81] claimed that non locality would be the reason behind the AB effect. He proposed a thought experiment in which long solenoid in the idealized AB experiment has two holes through which a pencil beam can be shot directly at the center as well as to left and right. Three distinct interference experiments can be performed: (i) beams at left and center, (ii) beams at right and center, and (iii) beams at left and right. Greenberger found that the right-hand beam leads the left-hand beam since the former leads the central beam which in turn leads the latter. Thus, he concludes that the relative phase difference must be due to some non-local effect since neither beam traverses a field.

2.5.2 Aharonov's Analysis of Non-Locality. In this section, we discuss Aharonov's definition of non-local properties and his particular arguments for the non-locality of the AB effect.

In 1984, Aharonov [Aha84] adopted the non-local interpretation of the AB effect and denied his opinions about the potentials theory. He states in the introduction:

> "In recent years it has become clear that non-local phenomena and topological effects play an important role in many areas of physics . . . [the AB effect] is common to all gauge theories, and provides a particularly clear example of non- local phenomena."

Aharonov defines "non-local" properties for an extended system, a system " which occupies two separate regions of space." The critical part of the definition is

> "The essential difference between local and non-local properties of the system is that in the former case all possible information can be obtained by independent measurements made in the two regions, while in the latter case this is not true."

Aharonov offers a collection of arguments for the non-locality of the AB effect in particular, most of which are negative. He criticizes various local theories of the AB effect, namely, (i) the dependence of the vector potential on the gauge and therefore not being measurable, (ii) leakage cannot produce the phase shifts because calculations for the ideal case with zero leakage nonetheless predict the shifts, (iii) the psi-function must be not multi-valued, and (iv) the apparently local, non linear equations of the hydrodynamics formulation conceal non-local effects.

The final argument for non-locality considers the derivations of the AB effect from scattering theory [OP85]. If the AB experiment is regarded as a scattering one, a beam of particles hits the solenoid with some specific radius. The particles can either pass safely to the other side, deflect away off the solenoid and continue on or reflect back in the opposite direction. The prediction of the AB effect can be recovered by application of standard scattering methods. Surprisingly, Aharonov noted that the predicted phase shift hold even when the radius of the solenoid approaches zero in contrast to what usually happens in a typical scattering experiment where we find that the effect of scattering is gone when the dimensions of the target are reduced due to lack of collisions. However, in the case of AB effect regarded as as a scattering problem, the calculations predict an effect even when collisions have a near-zero probability. This feature led Aharonov to conclude that the AB is not caused by collision, but rather is due to some non-local action. Although formally flawless, this argument is a negative one against explanations of the AB effect by direct collision with the solenoid for it does not give more evidence against the potentials theory.

To sum up, Aharonov uses a method of draining for non-locality; he strikes theories of locality one after another leaving only non-locality views.

3. Vaidman Argument

3.1 Introduction

It is widely agreed that the AB effect exists after its experimental confirmation by Tonomura and coworkers [TMS$^+$82, TOM$^+$86]. However, the interpretation of the role of the potentials remains obscure for the reason whether they have some physical meaning or being just mathematical tools. The potentials themselves are not gauge-invariant, which means that they are mathematical notions, however their line integrals are gauge-invariant which are used in the calculations through the Schrödinger formalism of the standard quantum mechanics. Recently, Lev Vaidman has proposed that the treatment of the AB effect may be given without the usage of the potentials at all implying that electrodynamics before the AB effect still holds [Vai12]. The argument of Vaidman can be briefed as follows. The composite system of both the electron and the sources of the fields are represented by one state, an entangled state represented by the total wave function

$$\frac{1}{\sqrt{2}} \left(|L\rangle_e \, |\Psi_L\rangle_S + |R\rangle_e \, |\Psi_R\rangle_S \right), \tag{3.1.1}$$

where $|L\rangle_e$ $(|\Psi_L\rangle_S)$ and $|R\rangle_e$ $(|\Psi_R\rangle_S)$ are the left and right states of the electron (source) respectively. Due to the electromagnetic interaction of the electron with the sources, the states of the sources are identical except for the AB phase. Therefore, the total wave function becomes

$$\frac{1}{\sqrt{2}} |\Psi\rangle_S \left(|L\rangle_e + e^{i\phi_{AB}} |R\rangle_e \right), \tag{3.1.2}$$

and the AB phase of the leads to interference patterns. In section 3.2, the thought experiment proposed by Vaidman to present the magnetic AB effect will be briefed while in section 3.3 the electric case will be discussed.

3.2 The magnetic AB effect

Regardless to its feasibility in laboratory, Vaidman [Vai12] proposed the following setup to discuss the magnetic AB effect. Consider a solenoid that consists of two cylinders of radius r, mass M, large length L, and charges Q and Q homogenously spread on their surfaces. The cylinders rotate in opposite directions with surface velocity v. As shown in Fig. 3.1, the electron encircles the solenoid with velocity u in a superposition of being in the left and in the right sides of the circular trajectory of radius R.

The flux in the solenoid due to the two cylinders is

$$\Phi = 2\pi r^2 \frac{4\pi}{c} \frac{Qv}{2\pi rL},$$
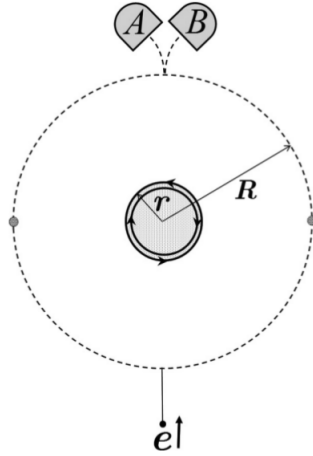$$= \frac{4\pi Qvr}{cL}. \tag{3.2.1}$$

Figure 3.1: The magnetic AB effect. Taken from [Vai12]

Thus, the AB phase, i.e., the change in the relative phase between the left and the right wave packets due to the electromagnetic interaction, is

$$\begin{aligned}
\phi_{AB} &= \frac{e\Phi}{c\hbar}, \\
&= \frac{4\pi eQvr}{c^2L\hbar}.
\end{aligned} \tag{3.2.2}$$

Eq. (3.2.2) gives the AB phase according to using electromagnetic potentials through the calculations. We shall now calculate the AB phase, ϕ_{AB}, using Vaidman's calculation based on direct action of the electromagnetic field and compare it with its definition from Aharonov and Bohm [AB59], Eq. (3.2.2). For the sake of simplicity we assume that $r \ll R \ll L$. The strategy will be as follows. When the electron moves toward the solenoid in a straight path, before following a circular path, the total magnetic flux through any cross section of the solenoid is zero. When the electron follows a circular path, the magnetic flux through a cross section of the solenoid at a distance z from the perpendicular drawn from the electron is

$$\Phi(z) = \frac{\pi r^2 euR}{c(R^2 + z^2)^{3/2}}. \tag{3.2.3}$$

Due its motion, the electron changes the magnetic flux and induces an electromotive force on the charged cylinders which changes their angular velocity v by δv. The change in the velocity δv is [Vai12]

$$\delta v = \frac{1}{M} \int_{-L/2}^{L/2} \frac{\pi r^2 euR}{c^2(R^2 + z^2)^{3/2}} \frac{1}{2\pi r} \frac{Q}{L} dz \simeq \frac{uQer}{c^2MRL}. \tag{3.2.4}$$

Therefore, the shift of the wave packet of a cylinder during the motion of the electron is

$$\delta x = \delta v \frac{\pi R}{u} = \frac{\pi Qer}{c^2ML}. \tag{3.2.5}$$

Therefore, the AB phase, ϕ_{AB}, is given as

$$\phi_{AB} = 4\frac{2\pi\delta x}{\lambda}, \tag{3.2.6}$$

where $\lambda = h/Mv$ is the de Broglie wavelength of each cylinder. The 4 factor arises from the fact that both cylinders are shifted and that they are shifted in the two branches. Substituting from Eq. (3.2.5) into Eq. (3.2.6), we get the expression for the AB phase as follows

$$\phi_{AB} = \frac{4\pi eQvr}{c^2 L\hbar}, \tag{3.2.7}$$

which is in agreement with Eq. (3.2.2). If the uncertainty in the velocity of the cylinders is smaller than δv, then, during the electron circular motion, the electron and the cylinders become entangled. But when the electron leaves the circular trajectory, it exerts an opposite impulse on the cylinders and this entanglement disappears.

3.3   The electric AB effect

In order to represent the electric AB effect, Vaidman proposed a one-dimensional interferometer as shown in Fig. 3.2. The interferometer consists of two mirrors on the same axis and a 50 % barrier at the middle between them. The interferometer is tuned in such a way that the electron always reaches mirror B. The electron wave packet starts moving toward the barrier from the right which transmits equal-weight wave packets to mirror A and B. Near Mirror A, the electron experiences a potential U, as shown in Fig. 3.3, which is a function of the distance between the electron and the mirror. The behavior of this function is described as follows. It goes to infinity at the surface of the mirror, at a certain distance from the mirror d it drops smoothly to a constant value V, $x \in (0, d)$, dieing out to zero for distances greater than d. The dimensions of the interferometer are much larger than d and we state that the electron is near the mirror when $x \in (0, d)$. The reason behind using such a potential is that the electron spends more time τ near mirror A in order to electromagnetically interact with the sources which are represented by identical particles of mass M and charge Q. The sources are placed symmetrically on the perpendicular axis of mirror A and are moving toward it by equal initial velocity v and have their own mirrors at a distance r from mirror A that make them spend more time T near mirror A. It is worth noting that $T < \tau$ such that the sources are near their mirrors at the same time the electron's wave packets are near mirror A. It must be noted that the electron does not experience any electric field anywhere at any time, the interference pattern, however, differs from that for any neutral particle. Therefore, this difference in the interference patter is due to the electromagnetic interaction. For this setup, the Aharonov-Bohm phase is given as

$$\phi_{AB} = \frac{-2eQT}{r\hbar}. \tag{3.3.1}$$

Now we follow Vaidman's argument to calculate the AB phase, Eq. (3.3.1). The composite system as a whole, electron and sources, is represented by a superposition state of the wave packets of the electron and the sources as well. Since the electron does not experience any electric field, its wave packets are not shifted. On the other hand, the wave packets of the sources are shifted due to their electromagnetic interaction with the electron and we shall now calculate this shift which is responsible for the AB phase.
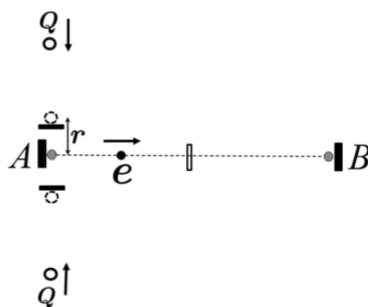
Figure 3.2: The electric AB effect. Taken from [Vai12]

The shift of the wave packets of the sources is generated during time T when they are near their mirror and is given by [Vai12]

$$\delta x = \frac{-eQT}{Mvr}. \tag{3.3.2}$$

It is worth while noting that Vaidman assumed that the shift δx should be *much* smaller than the position uncertainty of the charges in order that the interference in the AB setup be observed. The reason behind this is that, when the position uncertainty is smaller, the entanglement between the electron and the charges remains leading to decoherence washing out the AB effect. Since both charges Q are shifted in the same way, we get a factor of 2 in the relation between the AB phase and this shift such that $\phi_{AB} = 2 \times (\delta x/\lambda) \times 2\pi$, where $\lambda = h/Mv$ is the de Broglie wavelength. Therefore, the AB phase is expressed as

$$\phi_{AB} = \frac{-2eQT}{r\hbar}, \tag{3.3.3}$$
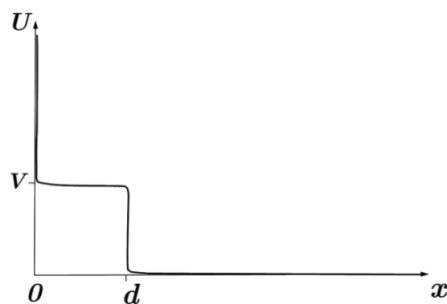
which is in agreement with Eq. (3.3.1).

Figure 3.3: The potential of the mirror forces. The potential energy of the particle as a function of its distance from the mirror. The particle with an energy slightly higher than V spends long time near the mirror. Taken from [Vai12]

4. Conclusions

In the previous chapters, I have reviewed the AB effect as well as discussed some of the interpretations of it. The direct and straightforward conclusion of the AB effect is that potentials are the fundamental physical entities rather than the fields which are merely given by differentiating the potentials. It turns out, however, that this is strikingly controversial and not widely acceptable. Although being experimentally verified and being useful in a very wide range of applications, the AB effect is still one of the puzzles in the quantum theory. Finding a plausible explanation of the AB effect, may help us to unravel some other mysteries of the quantum theory which certainly will lead to enlighten our way to better understand the quantum physics. In the subsequent sections, I will summarize some of the different arguments discussed in the previous chapters.

4.1   Arguments in Favor of Potentials

Positive argument in favor of the potentials can be listed as follows:

(i) *Locality.* Intuitions about locality,even if they were implicit, have been the most important reason for supporting the potentials view.

(ii) *Simplicity.* There is also another view which finds potentials to be the "natural" language of modern physics. We can see that relativistic electrodynamics can be expressed by the 4-potentials as well as both relativistic quantum theory, QED, and gauge theory (Yang-Mills fields)which take on this 4-potential as fundamental. This way we can see that potentials appear notably in both textbook presentations of most contemporary theory and in the daily work of physicists. Despite its validity, this kind of arguments is difficult to measure.

In favor of this point of view, Feynman in his textbook *Lectures on Physics* [FLS64] says in a section entitled "$\mathbf{B}$ versus $\mathbf{A}$"

> Is the vector potential merely a device which is useful in making calculations ... or is the vector potential a real field? ... What we mean here by a "real" field is this; a real field is a mathematical function we use for avoiding the idea of action at a distance... A "real" field is then a set of numbers we specify in such a way that what happens *at a point* depends only on the numbers *at that point*. We do not need to know any more about what's going on at other places... for a long time it was believed that [the potential] was not a real field. It turns out, however, that there are phenomena involving quantum mechanics which show that the potential is in fact a "real" field in the sense we have defined it. [Feynman 1964; 15-7, 8]

In addition, Feynman also repeats the assertion of Aharonov and Bohm that the potentials cannot be eliminated from the formalism of the quantum theory and replaced by the fields:

> The fact that the vector potential appears in ... quantum mechanics was obvious from the day it was written. That it cannot be replaced by the magnetic field in any easy way was observed by one man after another who tried to do so... It seems strange in retrospect that no one thought of discussing this experiment until 1959, when Bohm and Aharonov first suggested it and made the whole question crystal clear. [Feynman 1964; 15-12]

On the other hand, many arguments, however, claim that potentials cannot be regarded more than mathematical tools because of many reasons which are represented in the next section.

4.2   Arguments Against the Potential

(i)*Potentials are Gauge Dependent, Locally Arbitrary, and Unobservable.* The fact that any arbitrary gauge function can be added to the value of the potential counts against the potentials theory. Because we cannot "fix the gauge", we can never measure the value of a potential experimentally. This means that at the expense of locality, the potentials theory introduces the concept of the unobservable.

(ii) *Gauge invariant functions of the potentials are not local.* In the regions where no fields exist, the gauge invariant quantities constructed from the vector potentials are loop integrals. These integrals depend in an explicit way on the potential at many points implying that locality is not strictly conserved in the potentials.

(iii) *Potentials can be excluded.* In the quantum theory of electromagnetism, referring to potentials can be omitted from physical descriptions. For example, the AB phase shift can be written as

$$\phi_{AB} = \frac{e}{c\hbar} \oint \mathbf{A} \cdot d\mathbf{x} = \frac{e}{c\hbar} \int_{\Gamma} \mathbf{B} \cdot d\hat{\mathbf{s}}. \tag{4.2.1}$$

This way the phase shift can be expresses as a function of the field strength only in the second integral rather than depending on the potential in the first integral. The fact that potentials can be eliminated implies that the potentials are merely mathematical devices not necessary to physical descriptions. It is obvious that we cannot accept the idea of eliminating potentials completely, as in this case we will end up accepting non-local interactions.

4.3   Vaidman's Argument

The assumptions of Lev Vaidman discussed in the previous chapter is merely one of the arguments against the potentials theory summarized in the previous section. I, however, distinguished between them and that of Vaidman because the latter may represent a new class of interpretations that focuses on the composite system as a whole. Vaidman proposed two hypothetical experiments in order to show that we can produce the AB phase shift without using either the potentials or the fields. He assumes that the composite system, the electron and the sources, is represented by one state, and due to the electromagnetic interactions part of this state is changed (the sources' wave packets), hence, the total state. Such argument may be valid, however, it needs to be examined both theoretically and experimentally. In my opinion, I believe that Vaidman's treatment of producing the AB phase, though is considerably simple, needs to be applied to the Aharonov and Bohm's interference experiments which are experimentally verified. It may represent a formidable task to apply Vaidman's treatment, as far as I think, it is, however, necessary in order to prove for its generalization.

Acknowledgement

References

[AB59]   Y. Aharonov and D. Bohm, *Significance of electromagnetic potentials in the quantum theory*, Phys. Rev. **115** (1959), 485.

[AB61]   _____ , *Further considerations on electromagnetic potentials in the quantum theory*, Phys. Rev. **123** (1961), 1511.

[AB63]   _____ , *Further discussion of the role of electromagnetic potentials in the quantum theory*, Phys. Rev. **130** (1963), 1625.

[Aha84]   Y. Aharonov, *Aharonov-Bohm effect and non local phenomena*, Proc. Int. Symp. Foundations of Quantum Mechanics and their Technical Implications, Tokyo (1984), 10.

[Bel62]   F. J. Belinfante, *Consequences of the postulate of a complete commuting set of observables in quantum electrodynamics*, Phys. Rev. **128** (1962), 2832.

[Bel64]   J. Bell, *On the einsteinpodolskyrosen paradox. physics i (1964) 195*, Phys. (1964), 195.

[BL78]   P. Bocchieri and A. Loinger, *Nonexistence of the Aharonov-Bohm effect*, Il Nuovo Cimento **47** (1978), 475.

[Boy73a]   T. H. Boyer, *Classical electromagnetic deflections and lag effects associated with quantum interference pattern shifts: Considerations related to the Aharonov-Bohm effect*, Phys. Rev. D **8** (1973), 1679.

[Boy73b]   _____ , *Classical electromagnetic interaction of a charged particle with a constant-current solenoid*, Phys. Rev. D **8** (1973), 1667.

[cAT06]   cf. A. Tonomura, *The Aharonov-Bohm effect and its applications to electron phase microscopy*, Proceedings of the Japan Academy, Series B **82** (2006), 45 and references therin.

[Cha60]   R. G. Chambers, *Shift of an electron interference pattern by enclosed magnetic flux*, Phys. Rev. Lett. **5** (1960), 3.

[DeW62]   B. S. DeWitt, *Quantum theory without electromagnetic potentials*, Phys. Rev. **125** (1962), 2189.

[EPR35]   A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47** (1935), 777.

[FLS64]   R. P. Feynman, R. B. Leighton, and M. Sands, *Lectures on physics*, Addison-Wesley, 1964.

[Gre81]   D. M. Greenberger, *Reality and significance of the aharonov-bohm effect*, Phys. Rev. D **23** (1981), 1460.

[Hen81]   W. Henneberger, *AharonovBohm scattering and the velocity operator*, J. Math. Phys. **22** (1981), 116.

[Ken93]   J. B. Kennedy, *The Aharonov-Bohm effect and the non-locality debate*, Ph.D. thesis, Stanford University, 1993.

[Man62]   S. Mandelstam, *Quantum electrodynamics without potentials*, Annals of Physics **19** (1962), 1.

[Noe62]   P. D. Noerdlinger, *Elimination of the electromagnetic potentials*, Il Nuovo Cimento **23** (1962), 158.

[OP85]   S. Olariu and I. Popescu, *The quantum effects of electromagnetic fluxes*, Rev. Mod. Phys. **57** (1985), 339.

[Pop10]   S. Popescu, *Dynamical quantum non-locality*, Nature Phys. **6** (2010), 151.

[PT89]   M. Peshkin and A. Tonomura, *The Aharonov-Bohm effect*, Lecture notes in physics, Springer-Verlag, 1989.

[TMS+82] A. Tonomura, T. Matsuda, R. Suzuki, A. Fukuhara, N. Osakabe, H. Umezaki, J. Endo, K. Shinagawa, Y. Sugita, and H. Fujiwara, *Observation of Aharonov-Bohm effect by electron holography*, Phys. Rev. Lett. **48** (1982), 1443.

[TOM+86] A. Tonomura, N. Osakabe, T. Matsuda, T. Kawasaki, and J. Endo, *Evidence for Aharonov-Bohm effect with magnetic field completely shielded from electron wave*, Phys. Rev. Lett. **56** (1986), 792.

[Vai12]   L. Vaidman, *Role of potentials in the Aharonov-Bohm effect*, Phys. Rev. A **86** (2012), 040101.