Why is the transference theory of causation insufficient? The challenge of the Aharonov-Bohm effect

To be published in the Studies in History and Philosophy of Modern Physics

Vincent Ardourel   Alexandre Guay

October 4, 2017

Abstract

The transference theory reduces causation to the transmission (or regular manifestation) of physical conserved quantities, like energy or momenta. Although this theory aims at applying to all fields of physics, we claim that it fails to account for a quantum electrodynamic effect, viz. the Aharonov-Bohm effect. After having argued that the Aharonov-Bohm effect is a genuine counter-example for the transference theory, we offer a new physicalist approach of causation, ontic and modal, in which this effect is embedded.

1 Introduction

The transference theory reduces causation to the transmission (or regular manifestation) of physical conserved quantities, like energy or momenta (Dowe 2000, Kistler 2006). It is a major physicalist or ontic approach of causation, an approach that provides an account of causation as a physical process, based on our best scientific theories. As Dieks emphasizes, such an account maintains causation in "its rightful place as a category of physical ontology"(1986, p. 85). This theory of causation presumably applies to all fields of physics, like classical mechanics, relativistic physics as well as quantum physics. In particular, quantum electrodynamics is viewed as exemplifying the transference theory of causation: Interactions between charged particles and electromagnetic fields can be expressed via exchanges of physical quantities via photons. For instance, Salmon (1997), one of the first defenders of this account, argues for this claim:

> According to our best contemporary theory, quantum electrodynamics, the electromagnetic force is mediated by exchanges of photons. This means, in my terms, that whenever a photon is emitted or absorbed by a charged particle we have a causal interaction. Thus a charged particle undergoing acceleration in an electromagnetic field consists of a series of *causal processes* standing between frequent *causal interaction*. (Salmon 1997, p. 465. Our emphases)

In this paper, we nevertheless claim that the transference theory fails to account for a quantum electrodynamic effect, which is the Aharonov-Bohm (AB) effect. We have to carefully pay attention to this effect. First, it is not a singular effect but is, more generally, a particular case of class of different quantum phenomena (Berry 1984). More to the point, as we will argue for, the AB effect is a paradigmatic case of causal phenomena. This paper thus aims at showing that the AB effect is a counter-example for the transference theory in its current form but, also, at offering a new physicalist approach of causation in which this effect is embedded.

The AB effect is a quantum effect showing that an electronic interference pattern can be modified via an electromagnetic field that is completely shielded from the electrons themselves. This effect, predicted by Aharonov and Bohm (1959), has been well confirmed by experiments (Chambers 1960, Tonomura et al 1986), and has found various applications with materials (van Oudenaarden et al. 1998, Bachtold et al. 1999, Zaric et al. 2004). Such a phenomenon raises important questions for the foundations of physics. Healey (1997) has discussed in what sense this phenomenon exhibits *non-locality* and *non-separability*. On the other hand, Liu (1994, 1996) has argued for the reality of wave packets based on the AB effect. In this paper, we are interested in the AB effect with regard to the concept of causation and the transference theory. To our knowledge, there is a single discussion on the AB and its consequences on physical causation (Zangari 1992). However, it does not explicitly tackle the transference theory. Zangari rather defends that the notion of potential has to be added to account for physical causation. Although we agree with his approach, it is a worthy project to investigate in what sense the AB effect makes the transference theory controversial, and to show how it is possible to reconsider this theory in order to account for this phenomenon. Our project is not thus to reject merely the transference theory based on the analysis of the AB effect. Although we argue that the transference theory is not capable of accounting for a paradigmatic case of causation, we suggest how this theory could be revised for that purpose. We propose a view of physical causation based on the core notions of *propagation* and *interaction*.

The paper is structured as follows. First, we outline the main claims of the transference theory and sketch some possible issues (Section 2). We then turn to the AB effect by arguing that it cannot be a case of the transference theory (Section 3). Although the transference theory fails to account for this phenomenon, we then argue that the AB is nevertheless a paradigmatic case of causation (Section 4). Finally, we suggest how to reconsider the theory of transference, and offer a new physicalist approach of causation that includes the AB effect (Section 5).

2 The transference theory and possible counter-examples

Let us begin by introducing the transference theory and then its usual counter-examples as discussed in the literature.

2.1 The transference theory

According to the transference theory, causation reduces to the transmission (or regular manifestation) of a physical quantity from an event A to an event B. This theory comes from ideas of Aronson (1971a, 1971b) who suggests that "Prior to the time of the occurrence of B, the body that makes contact with the effect object possesses a quantity (e.g., velocity, momentum, kinetic energy, heat, etc.) which is *transferred* to the effect object (when contact is made) and manifested as B"(1971b, p. 422). Aronson makes clear that causation corresponds to the transference of a physical quantity. However, such a quantity is not clearly identified: it can be "heat" as well as "velocity", which are very different from a physical point of view since, for instance, heat can be dissipated and velocity is not a conserved quantity in elastic collisions. Fair (1979) offers a similar account of causation but focuses on "energy" and/or "momentum" as transferred quantities. The identification of these quantities comes from physicists' empirical investigations, as empirical facts.[1]

[1] It should be noted that the transference theory is just a new take on a old debate about modality, namely how to understand metaphysically the basic concepts of propagation and production (Schrenk forthcoming).

On the other hand, Salmon (1977, 1980, 1984) also provides a physicalist – even though quite different – theory of causation, the *theory of mark transmission*. In this approach, there are two distinct causal ingredients. On the one hand, there are causal processes that transmit marks, i.e., propagate some quantities, sometimes defined as processes that transmit energy (1984, p. 146). On the other hand, there are causal interactions, which correspond to the intersection of two causal processes. Following these different approaches, Dowe (1992a, 1992b, 2000) offers a unified theory, namely the *conserved quantity theory*, which is defined as follows:

> The conserved quantity theory can be expressed in just two propositions:
>
> CQ1. A *causal process* is a world line of an object that possesses a conserved quantity.
>
> CQ2. A *causal interaction* is an intersection of world lines that involves exchange of a conserved quantity. (2000, p. 90)

A *causal interaction*, like the collision between two billard balls, is thus defined via the exchange of conserved physical quantities, viz. energy and momentum. Similarly, Kistler (1998, 2006) argues for a transference theory based on the notion of transfer (or regular manifestation) of physical conserved quantities between distant events. His theory is defined by the statement (S) as follows:

> (S) Two events c and e are related as cause and effect if and only if there is at least one conserved quantity P, subject to a conservation law and exemplified in c and e, a determinate amount of which is transferred between c and e. (2006, p. 26)

By "transference", it is explicitly meant that "an amount A is said to be transferred between c and e, if and only if this very amount is present in both events" (Kistler 2006, p. 26). Despite several differences between Dowe's and Kistler's approaches, causation is defined in both cases as the *transfer* of a physical conserved quantity. Therefore, the decisive rebuttal for this theory would be to exhibit a paradigmatic case of causal relations between two events that would *not* involve any transference of conserved quantity.

2.2 Possible counter-examples

There have been many critical discussions against the transference theory of causation. Most of them have consisted in raising counter-examples in order to argue that the transference of conserved quantities is neither a necessary nor a sufficient condition for causation. For instance, counter-examples based on causation by *disconnection* (Schaffer 2000) have been raised. In those cases, although two events C and E are causally related – at least in a counterfactual sense – there is no exchange of a conserved quantity because of a lack of intrinsic connection between C and E. This could happen when (i) something prevents the exchange of a quantity – energy for instance – between two events C' and E, and (ii) C is an event that releases this prevention. As a result, C causes E although there is no transference of conserved quantities between C and C'. An paradigmatic example is a weight that accelerates because the stretched spring to which it is attached is unblocked. The usual reply for this kind of objection is to deny that those cases are genuinely causal. In particular, Aronson argues for the distinction between a *cause*, which involves transference of conserved quantity and an *occasion*, which is only "a condition that enables the cause to act"(1971a, p. 425). The release of a prevention is not a cause but what makes possible the cause to act.

Conversely, it has been argued that conserved quantities can be transmitted without characterizing a causal relation. This objection comes with *misconnection* (Dowe 2000, p. 147, Schaffer 2001). For instance, there is transmission of a certain quantity of billard chalk between a pool cue and a billard ball. However, this exchange of conserved quantity is not relevant for a causal relation between the motion of the cue and the motion of the ball. One can find different kinds of misconnections, viz. micro-connections and pseudo-connections. Against that kind of counter-example, it can be argued that the quantity of billard chalk involved is of such a negligible proportion to be regarded as causally relevant. It can also be argued that the line of billard chalk and the line of the billard balls moving on the table might be regarded as only coincident (Dowe 2000, chap. 7).

A very different kind of objection consists in paying attention to some physical phenomena for which the transference theory cannot be applied. These counter-objections are genuinely grounded in physics. We mention two important discussions that are relevant for our criticism.[2]

First, the case of quantum EPR experiments has been extensively discussed. In a nutshell, such experiments exhibit correlations between the measures of two entangled particules separated by a space-like interval, i.e., without any possibility of the transmission of a physical signal, and a fortiori,

[2]Reuger (1998) has also provided a criticism grounded in physics based on a case in general relativity.

any possibility of the transference of a physical quantity. The defenders of the transference theories try to overcome such issues in different ways. For instance, Dowe (2000) defends a *backwards* causality model, which relies on the hypothesis that "act of measurement on particle A brings about causal influence that propagates backwards-in-time to the source of the two particles"(2000, 185). In that case, the price to defend the transference theory is to endorse controversial claims about backwards-in-time processes. Nevertheless, although the EPR experiments are viewed as problematic with respect to the transference theory, the other option is merely to deny such phenomena as causal:

> The defender of the transference theory can argue that the correlation at a distance shown by the EPR phenomenon is not an instance of causation [...]
>
> The scientific debate on the explanation of this phenomenal is not yet over [...] So far, nothing prevents us from thinking that quantum mechanical correlations at a distance do not give rise to causal relations, but are a form of non-local and non-causal determination. Therefore, we can keep claiming that condition (S) is both sufficient and necessary. (Kistler 2006, p. 28-29)

In our opinion, although EPR experiments raise issues on causality, they are not a clearly decisive rebuttal for the transference theory. EPR experiments are based on quantum entanglement, which makes unclear whether there are two systems or rather only a single one. We will go back to EPR experiments in Section 4.

Second, in the context of classical mechanics, Lupher (2009) has pointed out that systems at equilibrium challenge the conserved quantity theory of causation (CQTC). He correctly emphasizes that this theory is, by design, too limited:

> The CQTC approach favors dynamical interactions as being causal interactions. This is evident from the definition of causal interactions involving the exchange of conserved quantities. But this rules out other types of interactions from counting as causal interactions, namely systems in equilibrium. (2009, p. 78)

For that purpose, he discusses the case of a ring with a mass at the center in gravitational equilibrium. There is no conserved quantities which are transferred between the two systems. Nevertheless, there is an obvious gravitational causal interaction between the two bodies. As Lupher stresses, "the fact that the net forces may sum to zero does not imply that no forces are present"(2009, p. 79).[3] We agree with this objection. It is a genuine counter-example for the transference theory of causation. Even if our paper mainly deals with quantum physics, our revised account of the transference theory of causation will allow us to represent systems at equilibrium as causal (see Section 5.2).

The previous objection is however restricted to static physical phenomena. In the next section, we show that a dynamical phenomenon also challenges the transference theory of causation, which is a quantum phenomenon different from EPR experiments.

3 The challenge of the Aharonov-Bohm effect

Let us now turn to the AB effect. We argue that this physical phenomenon, or more precisely the Aharonov-Bohm phase shift, might be a genuine counter-example to the transference theory of causation. For that purpose, we first introduce this phenomenon as it is presented in the usual textbooks and argue that it does not involve any transference of a conserved physical quantity.

3.1 The Aharonov-Bohm effect

The AB effect is a quantum interference phenomenon with electrons (or any charged particles) under the influence of an enclosed electromagnetic field. It is well illustrated by a two-slit experiment (see Figure 1).[4] Let us take such a two-slit experiment with a solenoid located between the two slits and the screen. The solenoid is perfectly isolated in a such way that there is a non-zero magnetic field $\mathbf{B}$ *only* into the solenoid and *no* magnetic field outside. Electrons are emitted from the source S and pass through two slits separated by a distance d. A pattern of bright and dark electronic fringes is observed on a screen at a distance l of the slits. This pattern comes from the interference between two possible paths for the electrons (path 1 and path 2) that meet at the point x on the screen.

Under these conditions, let us first assume that there is no solenoid – or the solenoid is unplugged. In that case, there is a certain pattern of fringes according to the standard quantum interference phenomenon (See Figure 1,

[3]Mattingly ironically stresses this point by reporting a quip of a colleague: "the assurance that no net forces are present will be cold comfort to a prisoner sentenced to be drawn and quartered – our victim is torn apart in any case"(2007, p. 904).

[4]This section is a summary of Healey's paper (1997, Section 2).

Original Pattern). Then, let us add a solenoid – or just let us plug in the solenoid. In that case, the pattern of fringes is modified. This is the AB effect that corresponds to a shift of the pattern:

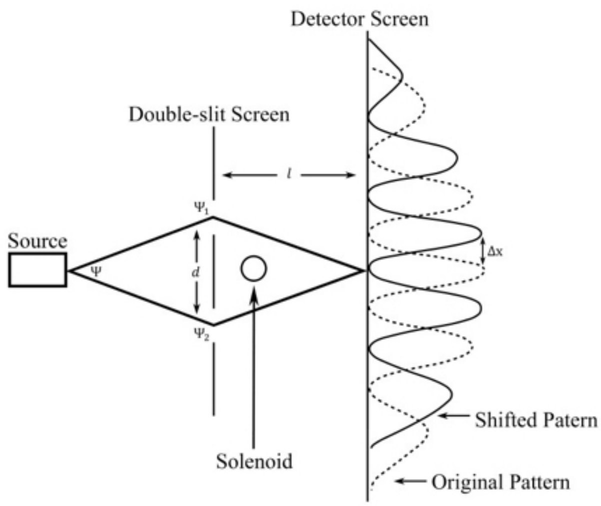

Figure 1: Aharonov-Bohm effect. Figure extracted from (Shech 2007).

$$\Delta x = \frac{l\lambda}{2\pi d} \frac{e}{\hbar} \Phi \tag{1}$$

where λ is the wave length of electrons, e their electric charge, $\hbar$ the reduced Planck constant, and Φ the magnetic flux produced by the solenoid (See Figure 1, Shifted Pattern).

The shift in the AB effect is explained by a change of the phase of electrons computed using the potential vector $\mathbf{A}$ that crosses the paths of electrons. When there is no solenoid, the usual difference of phase δ at the point x due to the two electronic paths 1 and 2 is :

$$\delta = \frac{2\pi x d}{l\lambda} \tag{2}$$

Interference fringes on the screen are bright when δ is a multiple of 2π. Then, because of the solenoid, there is a vector potential *outside* the solenoid. Apparently, this potential vector adds a phase change of the electrons wave function at the point $\mathbf{r}$ by an amount $-\frac{e}{\hbar}\mathbf{A}.d\mathbf{r}$. The resulting total phase change along a classical path is thus:

$$\delta = -\frac{e}{\hbar} \int \mathbf{A}.d\mathbf{r} \tag{3}$$

where the integral ranges over the path of electrons. Therefore, the difference of phase between the two paths from the source S to the screen is:

$$\Delta(\delta) = \delta_1 - \delta_2 = \left(-\frac{e}{\hbar}\int_1 \mathbf{A}.d\mathbf{r}\right) - \left(\frac{e}{\hbar}\int_2 \mathbf{A}.d\mathbf{r}\right) = -\frac{e}{\hbar}\oint \mathbf{A}.d\mathbf{r} \qquad (4)$$

where the last integral is taken on the closed curve corresponding to the sum of the paths 1 and 2. Then, applying the Stokes theorem to the definition of the magnetic field as $\mathbf{B} = \nabla \times \mathbf{A}$, one obtains:[5]

$$\Delta(\delta) = -\frac{e}{\hbar}\oint \mathbf{A}.d\mathbf{r} = -\frac{e}{\hbar}\iint \mathbf{B}.d\mathbf{S} = -\frac{e}{\hbar}\Phi \qquad (5)$$

Since this difference of phase does not depend on x, it leads to a global shift of the pattern by the amount:

$$\Delta x = \frac{l\lambda}{2\pi d}\Delta(\delta) = \frac{l\lambda}{2\pi d}\frac{e}{\hbar}\Phi \qquad (6)$$

3.2 A regular effect without transference of conserved quantities

Although the interference pattern is shifted by the application of an enclosed magnetic field, and thus a potential vector crossing the electrons' paths, we nevertheless claim that this phenomenon does not involve the transfer of any physical quantity.

First of all, let us focus on the usual physical conserved quantities discussed in the literature. None of them, viz. momentum, angular momentum, energy, and spin, are transferred (Olariu & Popescu 1985, p. 358; Zangari 1992, p. 271; Guay 2008a, p. 690). As Liu emphasizes:

> Does the [vector] potential have any effect on the expected values of a quantum object's observable properties? Does it have any effect on its average position, momentum, spin, and so on? Theoretically, the answer is again no. (Liu 1994, p. 994)

The theoretical reason put forth is that all the observables in quantum mechanics commute with the components $\mathbf{A}_{x,y,z}$ of the potential vector that appear in the additional phase change $\delta = -\frac{e}{\hbar}\mathbf{A}.d\mathbf{r}$. Therefore, the time

[5] We emphasize that this result comes from a first order approximation on the interaction between the electrons and the potential vector. For technical details, see (Guay 2004).

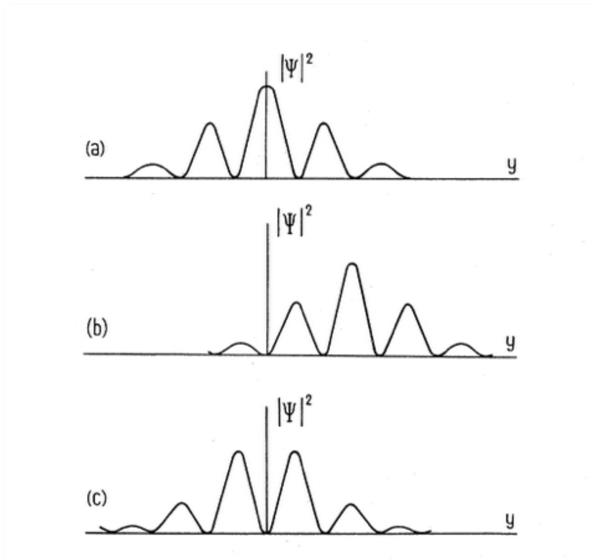

Figure 2: This figure is extracted from (Olariu and Popescu 1985, p. 350). It represents the electronic interference pattern in a two-slit experiment. In (a) there is no magnetic field. In (b), there is a magnetic field crossed by electrons. In (c) there is a magnetic field that is not crossed by electrons (A-B effect).

average value of any physical quantity associated to an observable is not modified by the potential vector (Liu 1994, p. 994).

Similarly, Olariu and Popescu (1985) claim that the usual conserved physical quantities are not modified by the potential vector within the AB effect:

> We show in this section that the changes in the average values of the aforementioned quantities [viz. energy and momenta] depend on the product of the fields strength times the probability density, whence we conclude that the average position, momentum, energy, and angular momentum of a charged particle are not affected by distributions of enclosed fluxes. (Olariu & Popescu 1985, p. 358)

Their proof is different and more detailed. They explicitly show that the equations describing the conservation of the kinetic momentum, the kinetic energy and the angular momentum only require the electromagnetic fields

E and **B**.[6] The potential vector **A** only appears in a such a way that it can be rewritten in terms of electromagnetic fields with respect to the usual equations between fields and potentials:

$$\mathbf{E} = \nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{B} = \nabla \times \mathbf{A} \tag{7}$$

with ϕ the electric potential. But, since there is no electromagnetic fields **E** and **B** crossing the paths of electrons, there is no modification of physical quantities of electrons, viz. energy, momenta, and spin.

In order to stress that any physical quantity is transmitted within the AB effect, let us contrast the AB effect with a two-slit experiment within a homogenous magnetic field instead of an enclosed field. Olariu and Popescu (1985, p. 350) show that the enveloppe of the interference pattern is modified with a homogenous electromagnetic field whereas it is not with an enclosed field. Both cases exhibit a shift of the interference pattern, corresponding to an additional phase change due to electromagnetic fields. But the enveloppe is modified only when electrons pass through a region with non-zero electromagnetic field (Figure 2). The average position of electrons is thus modified in the first case and not in the second. This means that no kinetic momentum or kinetic energy are exchanged in the second case.

None of the usual conserved physical quantities are thus modified within the AB effect. However, there is a quantity that is known to play a central role in predicting the AB effect (Wu and Yang 1975, Zangari 1992, Healey 1997, Belot 1998). It is the *holonomy*, also called the *Dirac phase factor* corresponding in the AB effect to:

$$S(C) = \exp\left(-ie/\hbar \oint_C \mathbf{A}(\mathbf{r}).d\mathbf{r}\right) \tag{8}$$

It corresponds to the factor of the wave function of the electron "interacting" with the potential vector along the curve C. According to Zangari, "[i]t therefore seems that the term which correctly characterizes the causal interaction in the Aharonov-Bohm effect is the phase factor" (1992, p. 269). Indeed, despite the question of which physical quantity is transmitted within

[6] For instance, "The conservation equation (1.89) [the equation of conservation at stake] shows that the average momentum of a charged particle can be modified only by a direct action of the field strengths. If the distribution of the field strengths is surrounded by finite barriers which render the probability of the presence of the particle in the region of the fields very small while keeping the wave function nonsingular, then by tracking the average of Eq. (1.89) we see that the total kinetic momentum is not changed by distributions of enclosed electromagnetic fields"(Olariu & Popescu 1985, p. 360).

the AB effect – if there is any – , there is a debate about which physical quantity is responsible for the AB effect. While the potential vector is a quite natural candidate, it cannot be viewed as a physical quantity since it is not gauge-invariant. In a nutshell, the electromagnetic potentials ϕ and $\mathbf{A}$ are not well-defined in the sense that they have gauge freedom. This means that there is a gauge transformation, i.e. the transformation that, for any smooth enough function f, corresponds to:

$$\mathbf{A} \longrightarrow \mathbf{A}' = \mathbf{A} + \nabla f \quad \text{and} \quad \phi \longrightarrow \phi' = \phi - \frac{\partial f}{\partial t} \tag{9}$$

which leaves unchanged the electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$. Such a transformation does not produce any physical observable differences. Instead, the holonomy $S(C)$ is gauge-invariant and therefore can be viewed as the physical quantity responsible for the AB effect.[7] Nevertheless, although the phase factor $S(C)$ is gauge-invariant, we argue below in Section 5 that it cannot be viewed as *transferred* or *transmitted*, undermining accordingly the transference theory of causation.

3.3 Is the Aharonov-Bohm phase shift a force-based effect?

Up until now, we have discussed the mainstream explanation of the effect, the one suggested by Aharonov and Bohm (1959): the phase shift of interference pattern is explained as occurring without any forces and transferred quantities. But it has to be noticed that this explanation is not the only one. There are discussions about the possibility of explaining the phase shift as a force-based effect. In this section, we make clear the consequences of this approach for the transference theory of causation.[8]

In several papers, Boyer (2000a, 2000b, 2006) proposes to explain the Aharonov-Bohm phase shift, i.e. the shift of the interference pattern, as an effect of electromagnetic forces. In a nutshell, there is an induced electromagnetic force between the passing charged particle and the solenoid. The exact forces are difficult to calculate, involving several particules and second order relativistic effects. Nevertheless, Boyer provides convincing arguments that could explain the Aharonov-Bohm phase shift as a force-based effect. More precisely, he shows that the Aharonov-Bohm phase shift might be recovered if we take into account these induced effects. Moreover, he stresses that his explanation is empirically distinguishable of the usual one. Unlike the standard predictions of the Aharonov-Bohm effect, induced

[7]We point out that this question is still controversial: see (Bunge 2015).

[8]We thank one of the anonymous reviewers for pointing us to this debate.

forces change the velocities of particules and their relative displacement. As a consequence, there might be experiments that could discriminate between the two explanations of the phase shift. Also, if Boyer's explanation were confirmed, the Aharonov-Bohm phase shift would not be viewed any longer as a counter-example of the transference theory of causation. However, as Boyer points out, there has been no such crucial experiments: "at present the only experimental evidence we have is the Aharonov-Bohm phase shift itself, and this phase shift is indistinguishable from those due to classical electromagnetic forces"(2000a, p. 898).

If Boyer is right, the Aharonov-Bohm phase shift implies the transference of linear momentum and kinetic energy between the solenoid and the passing particles. In consequence, it would not constitute a counter-example to the transference theory. But, in the meantime, since we are interested in "saving" the transference theory against possible objections, we focus on the most pessimistic option, the one for which there is no transmission of conserved quantities. This corresponds to the mainstream explanation of the Aharonov-Bohm phase shift.

4 The AB effect as a paradigmatic case of causality

Having discussed in what sense no transferred quantities might be involved in the AB phase shift, we now argue that the AB effect is still a paradigmatic case of a causal phenomenon.

First of all, we notice that the AB effect is generally described by philosophers as well as physicists with the vocabulary of causality. One precisely talks about the AB "effect" with regard to this quantum phenomenon unlike, for example, the EPR "experiment". It is indeed usual to describe the AB effect as the *effect* of an enclosed magnetic field on an interference pattern.[9] For instance, according to Healey, the AB effect involves a kind of causal relation, revealed by his use of the the word 'alter':

> As noted by Aharonov (1959), quantum mechanics predicts that the interference pattern produced by a beam of charged particles may be *altered* by the presence of a constant magnetic field, even though that field is confined to a region from which the particles are excluded. (Healey 1997, p. 19. Our emphasis.)

[9] We point out that, unlike this usual way of speaking, Bunge (2015, p. 132) claims that the AB effect is not really an *effect*.

Liu makes even clearer the idea that the AB effect lies on a causal relation:

> Roughly put, the AB effect unequivocally shows that there exists physical interaction between the electromagnetic potential and charged microscopic objects. [...] $\Delta S = (S_1 - S_2)$ [our $\Delta\delta = \delta_1 - \delta_2$, i.e., Eq. (5)] is the phase difference that *causes* the fringe shift. [...] Since the wave packets can never enter the solenoid nor can the **B**-field ever get out, if there is interaction between **B** and the electrons, it has to be action at a distance. Barring such action (a natural thing to do), the fringe shift must be *caused* by the interaction between **A** and the electrons. (Liu 1994, p. 989-990. Our emphases.)

And similarly, as we have seen, Zangari emphasizes that:

> It therefore seems that the term which correctly characterizes the *causal interaction* in the Aharonov-Bohm effect is the phase factor. (1992, p. 269)

However, the concept of causality on which the AB effect is based is still unclear. Healey (1997), Liu (1994), Zangari (1992) – among others – aim at describing in a physicalistic framework what happens in the AB effect. But, they do not tackle the very concept of *causality* to which their discussions refer. In Section 5, we shall offer such an account for causality in which the AB effect is embedded. But, before that, let us reinforce the idea that the AB effect should be viewed as causal.

In our opinion, the AB effect should be viewed as a causal phenomenon for mainly this reason: signaling or *information* can be sent via the AB effect. We stress this criterion since it allows us to contrast the AB effect with EPR experiments. Indeed, in the counterfactual approach to causation EPR experiments and the AB are not clearly distinguished. "Signaling" is not a novel criterion to argue for causal relations. Salmon already claimed that "it seems natural to refer to the genuine processes as "causal processes", for it is by virtue of the ability of such processes to transmit causal influences that they can transmit signals or information"(1997, p. 194). Even if we cannot provide here a definitive analysis of the AB effect from a signaling perspective, we make a first step towards this direction.

For that purpose, we suggest a very simple device for sending information with the AB effect. Let us take the two-slit experiments with, initially, no enclosed magnetic field in the solenoid – or no current in the solenoid. We have a bright fringe in the middle of the screen. Let us associate the state of the interference pattern to "0" (see Fig. 2a). Let us now apply a current in the solenoid in such a way that the interference pattern is shifted and a dark fringe is in the middle (Fig. 2b). Let us associate this state of the interference pattern to "1". With this code, it is thus possible to send a series of "0" and "1" from the solenoid to the screen.[10] In that sense, "signaling" is possible with the AB effect.

Further features might be required to argue that the AB effect allows signaling. If one adopts Salmon's view, one might require having a *continuous* process that transmit information. We thus should be able to ensure that all the information transmitted is in principle already available before the electrons arrive at the screen.[11] At this stage of our analysis, we do not see how to guarantee that such a criterion would be uncontroversially satisfied. Nevertheless, the AB experiment can be realized with a screen closer to the slits, with a distance $l' < l$ for instance (see Fig. 1). One would thus be able to send information with the same device at a different distance. How so? Because electrons have a continuous dynamics between the slits and the screen. There might be issues related to the status of quantum trajectories of electrons. But we stress that one deals here with a *dynamic* phenomenon, viz. the motion of electrons, which obeys to the Schrödinger equation. This is very different from the EPR experiments, which do not involve dynamics of electrons.

From this analysis, one might argue that the AB effect is actually an example of the transference theory since information is transmitted. But we stress that two notions are involved in the transference theory of causation: "transmission" and "conserved quantities". Although one can argue that information is transmitted, we emphasize that *information is not a physical quantity that satisfies laws of conservation.* As Kistler makes clear, "information is not transmitted in the physical sense required for the existence of causal relations because it does not obey to a conservation law" (2006, p. 233).[12] Similarly Lupher, who lists the 13 physical quantities associated with conservation laws in current scientific theory, does not mention information in the list (2009, p. 72). For that reason, one cannot view the AB effect as a genuine case of the transference theory.

Let us continue to contrast the AB effect with the EPR experiments.

[10] We could object that it is not possible to send two "0" since it is the same state. One can easily refine this protocol either by imposing a time sequence or by introducing a third state, for which the intensity of the fringe is between maximum and minimum, and take it as the neutral state from which one can either go to 0 or to 1.

[11] We thank an anonymous reviewer for pointing out this feature.

[12] See also Kistler (1998, p. 21).

Although one can doubt that the EPR experiments challenge the transference theory of causation, one should not view the AB effect as the same kind. First of all, it is unclear due to entanglement whether a single system or two systems occur in EPR experiments. Second, no information can be transmitted via these experiments, since no signal with finite speed can be transferred. Change in EPR experiment is based on *quantum measurement* – which is instantaneous – and not on a physical interaction described by a field, a tensor or another agent of interaction. As Zangari (1992) stresses this contrast:

> In the EPR case, strict locality is violated because the determination of an eigenvalue (i.e., the fixing of a kinematical quantity by measurement) associated with a particle at one point simultaneously determines the eigenvalue of another particle at a distant point. Since no time passes for a signal to propagate between the two particles, the correlation cannot be maintained by local degrees of freedom that satisfy strict locality. [...] By contrast, a non-local field-effect interpretation of the Aharonov-Bohm experiment has nothing to do with the measurement problem and would require the abandonment of locality, even while the system is "smoothly evolving" (i.e., described by Schrödinger's equation), rather than undergoing measurement. (Zangari 1992, p. 268)

The kind of non-locality involved with the AB effect is very different from the one in the EPR experiment. The interaction is described in a non-local way but nonetheless there is still a kind of locality since the electrons evolve according to Schrödinger's equation and do not require instantaneous change. Frisch (2005) also points out that the non-locality of the interaction within the AB effect is nevertheless compatible with a local propagation:

> Even though the phase factors are "spread out" through space, changes in their values do not propagate instantaneously. Since the vector potential propagates at a finite speed, the phase factor around paths far away from a disturbance in the field will change only after a finite time, when the disturbance has reached at least some point on the path. (Frisch 2005, p. 83)

Like the electromagnetic field, the potential vector satisfies a propagation equation with a finite speed. Even if it is not a physical quantity – because of its gauge non-invariance – a change of it that modifies $S(C)$ does not imply an instantaneous change of $S(C)$.

Before offering a new physicalist account for a causation in which the AB effect is embedded, let us go back to Salmon's original physicalist approach of causation. Salmon (1980) himself seems to leave open the possibility that a *causal interaction* might happen without transference of a physical quantity:

> When two causal processes intersect and suffer lasting modifications after the intersection, there is some correlation between the changes which occur in them. *In many cases – and perhaps all –* energy and/or momentum transfer occurs, and the correlations between the modifications are direct consequences of the respective conservation laws. (Salmon 1980, p. 60. Our emphasis.)

Salmon cautiously claims that, "in *many* cases", and not in "*all* cases", the correlation between the changes which occur when two causal processes intersect corresponds to the transfer of a conserved quantity. According to us, the AB effect is precisely such a phenomenon that cannot be denied as causal even if no conserved quantity is transmitted. Let us now make clear how such a phenomenon can be embedded in a physicalist approach of causation.

5 Towards a new account for causation

The transference theory of causation seems appropriate in classical contexts. In other words, in a classical setting, the metaphysical concepts of propagation and production seem, in most cases, well implemented by the notions of causal process and causal interaction. But the AB effect is a typically quantum effect with no classical equivalent. In the light of this observation, in order to incorporate all past successes, our goal, in this section, is to propose an extension, or more precisely a variation, of the theory of transference. This extension will be developed in two steps:

1. We replace the transference theory by a theory of causality based on the notion of interaction. This step will already expand the domain of physical processes that should be considered causal. However, it will be a conservative extension. All previous cases of causal relations will remain causal.

2. We expand the interactionist theory to include a modal aspect that will provide a sufficient base to model quantum causal processes.

In the last part of this section, we show how the AB effect is causal in this new causal framework.

5.1 Elements of alternative causal approaches

From an ontic or physicalist causal point of view, the AB effect has been explained in three ways:

1. The effect involves some kind of action at a distance from a local entity.
2. The effect is the result of the local action of a local entity, possibly a field.
3. The effect is the result of the local action of nonlocal entities or properties (Healey 2007, chapter 4).[13]

Let us briefly examine each of these possibilities. A natural way to implement the first option is to postulate that the charges confined in the solenoid act at a distance on electrons. Since the AB does not imply a transmission of energy, this causal relation does not violate the principle of local conservation of energy. Moreover, it has been shown in classical physics that an electromagnetic theory based on such an interaction can be formulated in such a way that it is empirically equivalent to the electrodynamics theory with a field (Wheeler and Feynman 1949).[14] At a first glance, it seems possible to extend this formulation to quantum physics. There is however a difficulty to this approach. In order to guarantee the local conservation of energy, we have to restrict the action at a distance to quantum phenomena like the AB effect. We will still have to add a field to the ontology to carry the energy during other kinds of interaction. To avoid this ontological inflation we could simply renounce to local conservation of energy, not a surprising position in the context of an action-at-a-distance theory.

A more problematic aspect of this conception is that even in the classical case this theory will be indeterministic in the Laplacian sense. Indeed, a complete description of the actual positions and momenta of all the charges is not sufficient to compute the future states. In order to save Laplacian determinism we have to add to the description of the state information

[13]It is important to mention that these are causal *interpretations* of the effect. The effect itself can be predicted using a diversity of models; using the gauge field or not, using the electromagnetic tensor or not, etc. See for example (DeWitt 1962). The goal of a causal interpretation is to identify the physical process these models try to represent. We thank an anonymous referee for this clarification.

[14]The main difference between these theories is that in the Wheeler and Feynman's approach there is no action of the charges upon themselves. Since this action is never verified without the use of an another charge, for all practical applications the theory with a field and the theory without a field are empirically equivalent.

about previous positions and momenta that could have a retarded impact.[15] This makes it very difficult to build a convincing causal implementation of the concepts of propagation and production that could be extended to the quantum realm.

The second option is a no go. As argued in (Healey 2007) and (Guay 2008b), we have reasons to believe that it is impossible, in the context of quantum physics, to define an ontologically plausible local bearer of electromagnetic interaction that would act locally. Either the bearer is not really gauge invariant (Leeds 1999) or it comes to choose a true gauge among all equivalent ones. This latter option is exemplified by the Liénard–Wiechert potential field $\mathbf{C}$. These potentials, such as $\mathbf{B} = \nabla \times \mathbf{C}$, "are completely definite and defined purely in terms of intrinsic, local properties" (Mattingly 2006, p. 246). This still raises the question of why one should choose this gauge among the other ones. In particular, $\mathbf{C}$ is still not empirically mesurable. Therefore, without strong philosophical reasons to choose it, this choice seems arbitrary. To model the AB effect as causal, some nonlocality or nonseparability seems unavoidable. Since these arguments are now well known, thanks to Healey, we will not discuss this point further.

The third option is more promising. For example, Healey (2007) argued that holonomies, understood as nonseparable processes, act locally on electrons. This interpretation can be extended to the classical realm. Holonomies propagate in conformity to the relativistic causality condition and allow local conservation of energy. Without denying the main insight of the holonomy interpretation, we have to assume that this proposition is not conservative. All previous cases of transference causation have to be reinterpreted as the effect of a nonseparable process. This bold move would be justified if this approach could increase substantially our comprehension of quantum causality. However this does not seem to be the case. From a causal point of view the local action of a non-local entity is not easily distinguishable from action at a distance. In the next subsection, we reinterpret the holonomy solution in a manner that will accommodate equally classical and quantum phenomena.

5.2 An interactionist account of causality in classical physics

In order to have a common language between classical and quantum physics, we will frame this discussion in the Lagrangian formalism. Let us begin by

[15] It is possible that the Wheeler-Feynman theory is deterministic in an ontological sense. Indeed we have reasons to believe that under realistic conditions, the solution for the evolution of charges is unique. For more details, see (Bauer, Deckert and Dürr 2013).

causation in classical cases. Once the variables describing states in the phase space are chosen, each system is characterised by Φ, the set of all possible histories. Φ includes the histories that satisfy and those that do not satisfy the dynamical equations (more on this below). On this set we can define a functional, the action $S : \Phi \to \mathbb{R}$, that will determine all the dynamical properties of the system. Indeed, between an initial state *in*, a certain state at a certain time, and a final state *out*, the physical history, the one that actually takes place (if it exists), is the one minimising S.[16] Let us assume that the dynamical equations are local in time, i.e. involving no time integrals and a finite number of time derivatives. In a relativistic theory, this implies that the equations are also local in space.[17] These two points put strong constraints on what the action S could be. We will assume the common answer: S is the integral of the Lagrange function or Lagrangian, L, plus boundary terms (DeWitt 2003, chapter 2).[18]

A few remarks before developing further the causation account. First, the Lagrangian formalism is part of a global approach of dynamics. Contrary to a local approach, where the starting points could be possible instantaneous states, in the Lagrangian formalism the starting point are possible histories of the system under study. Histories are extended in space-time.

Second, they are two kinds of possibilities involved in the Lagrangian formalism: kinematical and dynamical possibilities (Belot 2007). Kinematical possibilities are possible histories of the system. They consist in all the ways to map dependent variables describing the states of system based onto independent variables.[19] Together the kinematical possibilities form the set of possible histories Φ. The dynamical possibilities are a subset of the kinematical ones. They are the genuine physical histories. They are the possible histories that are compatible with dynamical laws. These histories could correspond to situations of the actual world.

It is tempting to associate each kinematical history to a possible world but this move should be resisted (Belot 2007). For example, a symmetry could associate a set of kinematical possibilities to the same possible world. It also possible that certain assumptions made (i.e. locality) are incompatible with other metaphysical claims about possible worlds. In consequence,

[16] It is a simplification, but it will do here.

[17] This assumption is not necessary but simplifies greatly the quantization process. For more details, see (Guay 2008b).

[18] For reason of simplicity, we will assume that L has an explicit dependence on a fixed space-time metric.

[19] In fact there are generally further constraints, for example differentlaibility, but this does not affect our argument here.

this direct association should be avoided.

As mentioned above, dynamical possibilities are histories that could happen according to the actual dynamical laws. They are the closest we have to actuality. They are a proper subset of kinematical possibilities. Therefore, there could exist kinematical possibilities that are incompatible with all dynamical laws. They would be not physically possible histories. Nevertheless, the set of kinematical possibilities should not be identified to the set of metaphysical possibilities. The former is apparently much more constraint than the later. The exact metaphysical interpretation of the set of kinematical possibilities will depend on the theory of natural laws sustained. For example, in the context of a governing law position, in which states can be defined independently of laws, a kinematical possibility that is not a dynamical possibility can be understood as an history that could have happened if the laws were different. However, in this context, a kinematical possibility could not refer to an eventual unlawful history, a history where no laws apply. We stop here and postpone a detailed discussion on modality in the Lagrangian formalism to a future paper.

Let us return to the basics of causation: propagation and production. In a typical Lagrangian we can find terms that describe free propagation and terms that represent interaction.[20] The first, the kinetic terms, describe the specific way energy is freely transmitted in this system from one location to another in spacetime. As Einstein argued (1970, p. 61), one of the main reasons to adopt an ontology in physics is in order to be able to formulate a reasonable statement of the principle of conservation of energy. In the context of electromagnetic systems discussed above, fields seem unavoidable. So what we mean by a causal process is the specific propagation process of entities involved in these kinetic terms, for exemple the free motion of charges or the free propagation of the electromagnetic field. Once we add the interaction terms, this picture becomes more complex since the entities identified in the free case can now interact, for example by exchanging energy. We have thus, using the interaction terms, a technique to represent the productive processes in this system. This presumes that the ontology identified in the free case survives more or less in the interactive case. This simplicity hypothesis is a good start but should not be assumed to always work, especially in the relativistic quantum case (Earman and Fraser 2006). The possibility of interaction can change the identity of the ontological entities identified in a non-interactive context. We have now a way to characterise each possible history, the propagation and the production story. If we want

[20]More on this below.

to know the causal story of the actual history of the system, we apply this analysis to the particular history that minimises the action.

In summary, our revised version of the transference theory of *classical* causation:

> CQ1'. A *causal process* is a continuous (or almost continuous) series of spacetime events governed by kinetic terms in L, conserving basic quantities, like energy.
>
> CQ2'. A *causal interaction* is a way the causal processes can interact with each other. This interaction is governed by the interaction terms of L.

We claim that the above described implementation is able to integrate all cases successfully analyzed by the transference theory of causation. Each case of causal process in the theory of transference, world line of a system that possesses a conserved quantity, can be described by a Lagrangian with only kinetic terms. The fact that this process in spacetime conserves energy or other quantities is the result of the particular structure of the dynamics of the system (the structure of the set of possible histories Φ and how is encoded the dynamics, the Lagrangian L). Each case of causal interaction in the usual theory, an intersection of world lines that involves exchange of a conserved quantity, is captured by specific interaction terms in L. Our Lagrangian approach manages to reproduce the transference results but also to extend them since we can capture the free propagation of entities that do not form clear world line, like fields. We can also model interaction that does not imply an asymmetrical exchange of a conserved quantity.

To illustrate this interactionist approach, let us apply it to a concrete example. Let us imagine N mass points interacting via the Newtonian gravitational force. In this case, the Lagrangian is

$$L = \sum_{i=1}^{N} \frac{m_i v_i^2}{2} - \sum_{1 \leqslant i < j \leqslant N} \frac{G m_i m_j}{\| \overrightarrow{r}_i - \overrightarrow{r}_j \|} \tag{10}$$

The first sum represents the kinetic terms. The second sum represents the gravitational interaction between material points.

According to CQ1', the kinetic terms determine what are the causal processes between interaction or more generally in the absence of interaction. If the dynamic was only governed by these terms the mass points would move inertially.[21] The interaction terms determine how the causal processes are

[21] For simplicity, we presume here that no collision would occur.

transformed. In this case all interactions are symmetric. In other words, L is not modified by indices permutation. It means that at a fundamental level there is no asymmetry in the causal interaction between material points. We can however in certain conditions obtain asymmetric causal claims. For example, if one of the points has a mass disproportionally bigger than all the others, we could justify the claim that this point is causally dominant. We can also explore how the dynamics would change if we modify the initial conditions or the values of the masses. This way we could even quantify roughly how this special material point is dominating the general behavior of the system.

In the context of Hamiltonian mechanics, Zangari (1992) has already made clear that asymmetric causal explanations –e.g., the fact that the change of atmospherical pressure casually explains the change of the value of the barometer, and not conversely– are compatible with symmetric interaction terms. He emphasized that the number of degrees of freedom of each system is decisive to obtain a causal dependance even if the interaction term is symmetric:

> In general, interaction terms determine the degree of "causal dependence" between interacting systems. In a loose sense, causal dependence can be quantified as a function of both the relative number of degrees of freedom from each system that contribute to the interaction, and the strength of the coupling between them. Roughly speaking, [...] only a very small number of the atmosphere [degrees of freedom] are involved in its interaction with the barometer. (Zangari 1992, p. 263)

We stress here that this interactionist account allows us to embed systems at equilibrium as displaying causal relations (See Section 2.2). In the case previously mentioned, of a ring and a mass point at its center, although the net force is zero, there is a gravitational potential and thus an interaction term in the Lagrangian, which reflects the causal relation between the two bodies.

Yet, given an arbitrary Lagrangian, can we always distinguish between kinetic and interaction terms? We do not have a general answer to this question. Nevertheless, we have a partial answer for a vast and useful class of systems: the natural ones. A Lagrangian system is called natural if its Lagrangian is equal to the difference between kinetic and potential energy. In these cases, the kinetic terms correspond to the quadratic form on each tangent space of the Riemannian manifold associated to the phase space

(Arnold 1989). Unless we have reason to believe the contrary, all other terms should be considered as interaction terms.[22]

Still, what about Lagrangians that are not natural, like the Einstein-Hilbert Lagrangian or the pure Chern-Simons Lagrangian? Our understanding is that in these cases causal interpretation is difficult independently of the conception of causality assumed. For example, even if the Einstein-Hilbert Lagrangian is apparently well behaved, the fact that the formulation of an associated Hamiltonian is far to be trivial (I am referring here to the famous problem of time in general relativity) is a sign that the meaning of propagation and production is not obvious. For the simple reason that the notion of change itself is not trivial. In the same vein, a Lagrangian that contains only Chern-Simons terms can represent certain aspects of a physical situation (i.e. the fractional quantum Hall effect). However, it is too poor to describe the dynamics of such quantum fluid if it is not very close to the ground state. If we enrich this Lagrangian to do so, we obtain a natural Lagrangian. We make the hypothesis that most of these problematic cases have similar roots. Either their dynamics is problematic and they can be considered not causal or they are approximations.

5.3 Feynman's functional integrals as a means for defining physical causation

How to move from this classical analysis to what it means to propagate and produce in quantum contexts? As already mentioned, one possibility is to follow Healey and keep the idea of a local action but allow nonlocal entities in order to represent the surprising aspects of quantum propagation and production. This option seems to us too little and at the same time too much. Too much in the sense that it does not preserve much of the classical version of causal process and interaction. Too little in the sense that it does not adequately account for the peculiarities of the quantum realm. It is still attached to a classical conception of what it is to *act*. This conception of classical action comes from classical physics, from the fact that world line of bodies or field propagations can be well defined. Quantum physics does not easily accommodate such an ontology.

To justify one or the other of the available quantum ontologies is beyond the objective of this paper. Rather we intend to develop a physicalist conception of causation that will apply generally. One trend we can follow from the classical to the quantum is to notice that this passage consists in

[22] In consequence, topological terms (terms that do not depend on the the spacetime metric) can have a causal effect, even if they do not involve energy transfer.

the replacement of the Hamilton functions and a phase space by transition amplitudes and an Hilbert space (Rovelli and Vidotto 2015). The Hamilton function $S(in, out)$ is defined as the value of the action on the physical history (if it exists) between *in* and *out* states, certain phase space states at certain specific times.[23] In the Feynman's functional formalism, a transition amplitude, between *in* and *out* states,[24] takes, in general, the form of in the DeWitt's condensed notation (Guay 2008b):

$$\langle out|in\rangle = \int e^{iS[\varphi]}\mu[\varphi][d\varphi], \quad [d\varphi] := \prod_i d\varphi^i, \tag{11}$$

where S is the action of the particular history of the entities represented by φ, $\mu[\varphi]$ plays the role of a volume density in the space of histories Φ and i is a general indice for all components of all entities (particles, fields...).[25] The transition amplitude is an integral that runs over possible histories. Each history contributes according to its action.

In classical physics, only histories that minimise the action are physically relevant. It is why we generally do not speak of probability in these cases. Between a state *in* and a state *out*, the physical history (if it exists) has a probability of 1. The other kinematical possibilities 0. In quantum physics, this is no longer the case. All possible histories are potentially conjointly responsible for the physical phenomenon. It is why this should be qualified as a physicalist modal conception of causation. To illustrate this proposition, we would like to discuss an interesting disanalogy. Let us imagine a machine M_1 that when receiving a 0 gives a 0 and another machine M_2 that gives a 1 in the same condition. We feed them both 0 and combine the result with an addition machine. Here we face an ordinary case of ontic causation, each machine consisting in a single causal process. These causal processes interact (in the addition machine) and a new process is generated. Let us now imagine that we have a *single* quantum machine. Each possible way the machine can process 0 is counted according to a certain protocol (addition). These ways to vote generate a result in the actual world. As in the first case, the dependance between the result and the machines is clear. However, the causal interpretation is different. In the classical case, the entire process of both machines is in the same causal regime. In the quantum case, only our world is quantum, the other possible histories are

[23]The Hamilton function is a complex mathematical object. See (Littlejohn 1992) for more details on its relation to the geometry of the phase space.

[24]In the context of a transition amplitude *in* and *out* do not represent classical states, but associated eigenstates of appropriate operators.

[25]For simplicity, we assume $c = \hbar = 1$.

classical. Nevertheless, they contribute causally to what happen in our world. While what happens in our world depends on what happens in other possible classical histories, the reverse is not true. We can imagine many cases, where there is no empirical difference between the result produced by a series of parallel classical causal processes and a modal sum of such processes. But this will not always be the case. Let us imagine that M_2 is not physically possible according to the dynamical laws. In this situation, the classical process and the quantum one will not provide the same output for the same input. The quantum process will still give 1 if feeding a 0 to M_2 and obtaining 1 is a possible process (a kinematical possibility). Apparently, the quantum machines give us an result classically impossible. In this context, it is not surprising that we would infer the existence of strange interaction. But we have to be careful. In this framework, the nonlocality is not necessarily the result of the action of non-local entities in our world. It could be the product of how the specific contributions of classical local causal processes are added. As an example, see the AB effect in the next subsection.

Let us summarize our proposition. In classical physics, we propose to identify causal process, propagation, by analysing kinetic terms of the Lagrangian. Causal interactions, which do not always reduce to the exchange of conserved quantity, are characterised by the interaction terms in the Lagrangian. The analysis can be applied to any element of the space of histories Φ. However the history that is actualised, knowing initial and final conditions, is, if it exists, the one that minimises the action, in other words the value of the Hamilton function $S(in, out)$.

In quantum physics, the Hamilton function $S(in, out)$ is replaced by the transition amplitude $\langle out|in\rangle$. The transition amplitudes are directly dependent on the action of each possible classical history between *in* and *out*. This dependence is formalized by the functional integral (Eq. 11). In our account of causation, a quantum causal process is the modal functional addition of classical causal process. In consequence, quantum causation is radically different from classical causation. However, there is a strong continuity between classical and quantum causation. For example, possible histories that are close to the actual classical one contribute constructively to the transition amplitude (MacKenzie 2000).

5.4 The causal interpretation of the Aharonov-Bohm effect

In the case of the AB effect, the transition amplitude takes the form of a Feynman propagator.[26] In particular, this propagator takes the form of an integral over all possible trajectories from *in* to *out*, $\langle out|in\rangle = \int D(\vec{q}(t))e^{iS[\vec{q}(t)]}$, where S is the classical action associated with path $\vec{q}$. With the free particle case as a reference, the action of the electromagnetic interaction is to multiply the contribution of each path $\vec{q}$ (of each possible history) by a non-integrable phase factor $U(out, in) = e^{-ie\int_q A_\mu dx^\mu}$ (called a Wilson line). If we compute the relative effect of the electromagnetic interaction on *two histories*, we obtain an holonomy (or a Dirac phase factor) already discussed above.

We could interpret these holonomies, as Healey does, as some kind of causal process in itself. More precisely as characterizing the local effect (the observed pattern of electrons' impact) of nonlocal electromagnetic properties. Of course this interpretation has important consequences on classical electromagnetic causality because the existence of these new properties has to be taken into account when we discuss the local effect of the classical local electromagnetic field. For us, the holonomies illustrate how the classical causal electromagnetic stories combine to generate the quantum causal process. In the end, both interpretations agree on the central role of holonomies. So what is at stake? To clarify this point we concentrate on the status of nonlocality. For Healey, the nonlocality of electromagnetism is an ontological fact concerning properties. Nonlocal electromagnetic properties actually exist. For us, it is quite possible that the apparent nonlocality in a quantum phenomenon is the modal result of the addition of local processes generated by local properties. This interpretation has no impact on classical causal analyses. No nonlocal causally powerful entity is necessary to explain the effect. We must only accept that distinct possible classical histories can be causally relevant at the same time.

The AB effect is a good example of this last point. The classical action of a charged particle on a path $\vec{q}$ is modified by turning on the current in the solenoid in the following way:

$$S[\vec{q}(t)] \rightarrow S[\vec{q}(t)] - e\int_q A_\mu dx^\mu \tag{12}$$

Classically, the new term has no impact. It adds the same constant to all trajectories that go in the same slit. The difference between the constant

[26] In this context, the Feynman propagator is the transition amplitude that a charged particle be present at a space-time point *out* if it was at *in*.

added to the top trajectories and the ones passing in the lower one (see Fig 1) is proportional to the electromagnetic flux in the solenoid. Computing the Euler-Lagrange equations shows that if a trajectory is not passing in a region where $F_{\mu\nu} \neq 0$, these constants make no dynamical difference. However, they should not be eliminated, since up to a gauge transformation, they describe the local properties of the electromagnetic flux. The quantum case is a different matter. Since all possible histories contribute to the phenomenon, the differences between these constants is causally significant. Thus the local properties of the enclosed electromagnetic field are contributing, not through an action at a distance, but in a modal way. This difference among classical histories is making the diffraction pattern shift. Nonlocality is the result of modality.

6 Conclusion

We claimed that the traditional transference theory of causation fails to account for the Aharonov-Bohm effect. None of the usual physical conserved quantities discussed in the literature, viz. momentum, angular momentum, energy, and spin, are transferred within this effect. Yet this effect is a paradigmatic case of causal phenomenon. In agreement with experimental results, it is possible to modify the interference pattern by changing continuously the magnetic field, allowing us in principle to send information from the solenoid to the screen.

We then argued for an extension of the traditional transference theory in order to embed the Aharonov-Bohm effect. The holonony or phase factor, which is the responsible for this quantum effect, is not strictly speaking a transferred quantity since there is no propagation of it. It is a non-local quantity. We thus offered a new version of traditional transference theory for which holonomies are the consequences of combined propagating processes. For that purpose, we interpreted the electrons' possible trajectories within Feynman's functional integrals formulation of quantum electrodynamics and deemed them to be propagating processes. In order to recover holonomies, all the possible paths have to be taken into account. This leads us to a modal extension of transference theory for which the Aharonov-Bohm effect is the result of the addition of possible causal processes.

Acknowledgements

References

[1] Aharonov, Y., & Bohm, D. 1959. *Physical Review*, 115 (3): 485-91.

[2] Arnold, Vladimir Igorevich. 1989. *Mathematical Methods of Classical Mechanics.* Translated by K. Vogtmann and A. Wienstein. Second Edition. New York: Springer.

[3] Aronson, J. L. 1971a. The Legacy of Hume's Analysis of Causation, *Studies in History and Philosophy of Science*, 2: 135–156.

[4] Aronson, J. L. 1971b. On the Grammar of 'Cause'. *Synthese* 22 (3): 414–430.

[5] Aronson, J. L. 1982. Untangling ontology from epistemology in causation. *Erkenntnis* 18 (3): 293-305.

[6] Bachtold, A. et al. (1999). Aharonov-Bohm oscillations in carbon nanotubes. *Nature* 397, 673–675.

[7] Bauer, G., D. -A. Deckert, and D. Drr. (2013). On the Existence of Dynamics in Wheeler–Feynman Electromagnetism. *Zeitschrift Fr Angewandte Mathematik Und Physik* 64 (4): 1087–1124.

[8] Belot, Gordon. 1998. Understanding Electromagnetism. *The British Journal for the Philosophy of Science* 49 (4): 531-55.

[9] Belot, Gordon. 2007. The Representation of Time and Change in Mechanics. In *Philosophy of Physics: Part A*, edited by Jeremy Butterfield and John Earman, 133?-227. Amsterdam: North-Holland.

[10] Berry, M.V. (1984). Quantal phase factors accompanying adiabatic changes, *Proceedings of the Royal Society of London*, Series A, 39 2: 45–57.

[11] Boyer, T. (2000a). Does the Aharonov–Bohm Effect Exist?. *Foundations of Physics*, 30(6), 893-.905.

[12] Boyer, T. (2000b). Classical Electromagnetism and the Aharonov–Bohm Phase Shift. *Foundations of Physics*, 30(6), 907-.932.

[13] Boyer, T. (2006). Darwin–Lagrangian analysis for the interaction of a point charge and a magnet: Considerations related to the controversy regarding the Aharonov–Bohm and Aharonov–Casher phase shifts. *Journal of Physics A*, 39, 3455–3477.

[14] Bunge, M.. 2015. Does the Aharonov-Bohm Effect Occur? *Foundations of Science*, 20 (2): 129-33. doi:10.1007/s10699-014-9354-2.

[15] Chambers, R. G. (1960). Shift of an electron interference pattern by enclosed magnetic flux. *Physical Review Letter* 5, 3–5.

[16] DeWitt, Bryce S. 1962. Quantum Theory without Electromagnetic Potentials. *Physical Review* 125 (6): 2189–91.

[17] DeWitt, Bryce. 2003. *The Global Approach to Quantum Field Theory.* Vol. 1. 2 vols. Oxford: Oxford University Press.

[18] Dowe P. (1992a), Process Causality and Asymmetry, *Erkenntnis* 37, pp. 179–196.

[19] Dowe P. (1992b), Wesley Salmon?s Process Theory of Causality and the Conserved Quantity Theory, *Philosophy of Science* 59, pp. 195–216.

[20] Dowe, P.. 2000. *Physical Causation.* Cambridge: Cambridge University Press.

[21] Earman, John, and Doreen Fraser. 2006. Haag?s Theorem and Its Implications for the Foundations of Quantum Field Theory. *Erkenntnis* 64 (3): 305-44. doi:10.1007/s10670-005-5814-y.

[22] Einstein, Albert. 1970. Autobiographical Notes. In *Albert Einstein: Philosopher-Scientist*, edited by Paul Arthur Schilpp. La Salle: Open Court.

[23] Fair, D.. 1979. Causation and the Flow of Energy. *Erkenntnis* 14 (3): 219-50.

[24] Felsager, B.. 1983. *Geometry, Particles, and Fields.* New York, NY: Springer-Verlag.

[25] Frisch, M. 2005. *Inconsistency, Asymmetry, and Non-Locality*, New York: Oxford University Press.

[26] Guay, Alexandre. 2004. *Symétrie: Réflexion Sur Les Formes Naturelles.* Montréal: University of Montréal.

[27] Guay, Alexandre. 2008a. Conceptual Foundations of Yang-Mills Theories. *Studies In History and Philosophy of Modern Physics*, 39: 687-93.

[28] Guay, Alexandre. 2008b. A partial elucidation of the gauge principle. *Studies In History and Philosophy of Modern Physics*, 39 (2): 346-63.

[29] Hausman, D. and J. Woodward (1999). Independence, Invariance and the Causal Markov Condition, *The British Journal for the Philosophy of Science*, 50, pp. 521-583.

[30] Healey, R. 1997. Nonlocality and the Aharonov-Bohm Effect, *Philosophy of Science*, 64 (1): 18–41.

[31] Healey, R. 2007. *Gauging What?s Real: The Conceptual Foundations of Contemporary Gauge Theories.* New York, NY: Oxford University Press.

[32] Kistler, M. 1998. Reducing causality to transmission. *Erkenntnis*, 48 (1): 125.

[33] Kistler, M. 2006. *Causation and Laws of Nature*. Oxon: Routledge.

[34] Kistler, M. 2013. The Interventionist Account of Causation and Non-Causal Association Laws. *Erkenntnis* 78 (S1): 65–84. doi:10.1007/s10670-013-9437-4.

[35] Leeds, Stephen. 1999. Gauges: Aharonov, Bohm, Yang, Healey. *Philosophy of Science* 66 (4): 606-27.

[36] Littlejohn, Robert G. 1992. The Van Vleck Formula, Maslov Theory, and Phase Space Geometry. *Journal of Statistical Physics* 68 (1-2): 7-50.

[37] Liu, C. 1994. The Aharonov-Bohm Effect and the Reality of Wave Packets, *The British Journal for the Philosophy of Science*, 45 (4): 977–1000.

[38] Lupher, T. 2009. A Physical Critique of Physical Causation. *Synthese*, 167 (1): 67-80. doi:10.1007/s11229-007-9289-z.

[39] MacKenzie, Richard. 2000. Path Integral Methods and Applications. `arXiv:quant-ph/0004090v1`.

[40] Mattingly, J. 2006. Which Gauge Matters? *Studies in History and Philosophy of Modern Physics* 37 (2): 243–62. doi:10.1016/j.shpsb.2005.08.001.

[41] Mattingly, J. 2007. Classical fields and quantum time-evolution in the Aharonov?Bohm effect *Studies in History and Philosophy of Modern Physics* 38: 888–905. doi:10.1016/j.shpsb.2005.08.001.

[42] Menzies, P., & Price, H.. 1993. Causation as a secondary quality. *British Journal for the Philosophy of Science*, 44 (2): 187-203.

[43] Olariu, S., & I. Iovitzu Popescu. 1985. The quantum effects of electromagnetic fluxes. *Reviews of Modern Physics*, 57 (2): 339-449.

[44] Rovelli, Carlo, and Francesca Vidotto. 2015. *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory.* Cambridge: Cambridge University Press.

[45] Rueger, A. (1998). Local Theories of Causation and the A Posteriori Identification of the Causal Relation, *Erkenntnis* 48: 25–38.

[46] Salmon, W. (1977). An "At-At" Theory of Causal Influence, *Philosophy of Science Association*, Vol. 44, No. 2 pp. 215–224.

[47] Salmon, W. (1980). Causality: Production and Propagation, *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1980, Volume Two: Symposia and Invited Papers (1980), pp. 49–69.

[48] Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

[49] Salmon, W. (1997). Causality and Explanation: A Reply to Two Critiques, *Philosophy of Science Association*, Vol. 64, No. 3 (Sep., 1997), pp. 461–477.

[50] Schaffer, J. (2000). Causation by Disconnection, *Philosophy of Science*, 67, pp. 285–300.

[51] Schaffer, J. (2001). Causes as Probability-Raisers of Processes, *Journal of Philosophy*, 98, pp. 75–92.

[52] Schrenk, Markus (forthcoming). *Metaphysics of Science: A Systematic and Historical Introduction.* Oxon: Routledge.

[53] Tonomura, A., Nobuyuki O., Tsuyoshi M., Takeshi K.i, Junji E., Shinichiro Y. and Hiroji Y. (1986). Evidence for Aharonov-Bohm Effect with Magnetic Field Completely Shielded from Electron Wave. *Physical Review Letters* 56 (8): 792-801.

[54] van Oudenaarden, A., Devoret, M., Nazarov, Y. & Mooij, J. (1998). Magneto-electric Aharonov-Bohm effect in metal rings. *Nature* 391, 768–770.

[55] von Wright, G. 1971. *Explanation and Understanding.* Ithaca, NY: Cornell University Press.

[56] WheelerFeynman1949 Wheeler, John Archibald, and Richard Phillips Feynman. 1949. Classical Electrodynamics in Terms of Direct Interparticle Action. *Reviews of Modern Physics* 21 (3): 425-33.

[57] Wu, Tai Tsun, & Chen Ning Yang. 1975. Concept of nonintegrable phase factors and global formulation of gauge fields. *Physical Review* D 12 (12): 3845-57.

[58] Zangari, Mark. 1992. Adding Potential to a Physical Theory of Causation. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1992, Volume One: Contributed Papers, 261-73.

[59] Zaric, S. et al. (2004). Optical signatures of the Aharonov-Bohm phase in single-walled carbon nanotubes. *Science* 304, 1129–1131.