A DYNAMICAL FORMULATION OF THE AHARONOV–BOHM EFFECT

Mark D. Semon

Department of Physics, Bates College, Lewiston, ME 04240, and Department of Physics, Amherst College, Amherst, MA 01002

John R. Taylor

Department of Physics, University of Colorado, Boulder, CO 80309

In 1969 Aharonov, Pendleton and Petersen introduced a new dynamical variable, the modular momentum, and discussed how several nonlocal interactions could be characterized as resulting from its exchange. In this paper we review and extend their approach by applying it to the Aharonov–Bohm effect with an electromagnetic scalar potential. First we review the nonclassical nature of the effect, showing that the interaction involved does not change the linear momentum of any particle creating the interference pattern. We then use the modular momentum to describe the interaction in the Heisenberg picture. Studying the equation of motion for the modular momentum, we prove the assertion of Aharonov et al. that modular momentum provides a dynamical description of the Aharonov–Bohm effect.

I. INTRODUCTION

In 1959 Aharonov and Bohm[1] proposed two experiments that called into question our understanding of what constitutes an interaction. In both cases an interference pattern was predicted to shift even though the particles creating it were never acted upon by any force. Since then equivalent effects have been predicted to occur in gauge fields,[2] Yang–Mills fields,[2,3] gravitational fields,[2,4] and inertial fields.[5-7] Nonetheless, our understanding of *how* these effects occur has remained incomplete. Explanations have been offered in terms of local effects of potentials,[8] nonlocal effects of field strengths,[9-11] and field–field interactions.[12,13] The effects even have been attributed to the time-dependent processes involved in setting up the experimental conditions.[14,15] Although each of these explanations illuminates certain aspects of the effect, a complete picture has yet to emerge from any one of them.

In 1969, Aharonov, Pendleton and Petersen[16] introduced a new dynamical variable, the modular variable, and discussed how several nonlocal interactions could be characterized as resulting from its exchange. Their work indicated that modular variables could provide a new formalism for describing the Aharonov–Bohm effect, and more generally, could give a new approach to nonlocality. Because their paper introduced the variables, it developed them in the context of several examples and only sketched proofs of the main results. After studying their work we felt that modular variables were interesting enough to review and expand upon in a simpler context. What we present here is a discussion of our approach and a summary of what we have proved so far.

II. REVIEW OF THE AB EFFECT WITH AN ELECTROMAGNETIC SCALAR POTENTIAL

In order to better understand modular variables, we decided to study them in the context of the simplest Aharonov–Bohm effect, the one involving the electromagnetic scalar potential. Following Aharonov *et al.*[16] we refer to this as the potential effect. In this section we give a brief review of the effect and its nonclassical nature.[16]

Consider a charged particle inside a conducting cylinder. The Schrödinger equation in this case is

$$H_0\psi_0 = i\,\partial\psi_0/\partial t \tag{2.1}$$

where H_0 is the free particle Hamiltonian, and we take $\hbar = 1$. If we begin putting charge on the outside of the cylinder then a time-dependent potential is created inside, but no **E** or **B** fields. In this case the Schrödinger equation becomes

$$H\psi = i\,\partial\psi/\partial t \tag{2.2}$$

with

$$H = H_0 + q\phi(t) \tag{2.3}$$

and ϕ the time-dependent potential, which is independent of position inside the cylinder. It is easily shown that solutions to this equation have the form

$$\psi = \psi_0 \exp(-i\alpha) \tag{2.4}$$

where ψ_0 is any solution to Eq. (2.1), and α is a phase given by

$$\alpha = q \int^{t} \phi(t')\,dt'. \tag{2.5}$$

Since α is independent of position, the state of any particle whose wavefunction lies entirely within the cylinder is changed only by an overall phase factor. Thus, the presence of $\phi(t)$ inside the conducting cylinder has no observable consequences, which is what we would expect classically, since the region is force-free.

In 1959, however, Aharonov and Bohm showed that, contrary to our classical expectations, there are situations in which the presence of potentials in force-free regions can result in observable effects. They proposed placing a conducting cylinder behind each slit in a two-slit interference experiment. As the particle passes through the slits its state is given by

$$\psi_0 = \psi_1 + \psi_2 \tag{2.6}$$

where ψ_1 is the probability amplitude for going through slit one, and ψ_2 the probability amplitude for going through slit two. If no potentials are applied during the particle's passage through the cylinders, the interference pattern on the screen is described by

$$\psi_0^*\psi_0 = |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1||\psi_2|\cos\delta \tag{2.7}$$

where δ is the relative phase between ψ_1 and ψ_2. However, if a potential is applied to the second cylinder during the particle's time within the cylinders, and removed before the particle exits, a new situation results. When the particle hits the screen its wavefunction can be written as

$$\psi_\alpha = \psi_1 + \exp(-i\alpha)\psi_2 \tag{2.8}$$

with α given by

$$\alpha = q \int_{-\infty}^{+\infty} \phi(t)\, dt. \tag{2.9}$$

The interference pattern on the screen is then described by

$$\psi_\alpha^* \psi_\alpha = |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1||\psi_2| \cos(\delta + \alpha). \tag{2.10}$$

Thus, the interference pattern is predicted to shift even though the particles creating it are never acted upon by any force. The shift is gauge invariant, as we would expect, but shows that some type of interaction has occurred, which, from a classical point of view, is a surprise.

The approach that Aharonov *et al.* take to this is as follows: Normally we describe an interaction by finding some variable that it changes. In classical physics, for example, it is the change in a particle's momentum that characterizes its interaction with a source. In light of this we ask: is it possible to describe the interaction in the potential effect in terms of the change of some variable belonging to a particle contributing to the interference pattern? For example, we can ask if $\langle p \rangle$, where p is the linear momentum of a particle along the screen, depends on α. If not, then the interaction leaves the momentum unchanged, and hence is nonclassical.

We have proved what Aharonov *et al.* suggested to be the case, that the interaction leaves $\langle p^n \rangle$ unchanged, for n a positive integer. That is, we have proved that the interaction does not affect the linear momentum, kinetic energy, or any higher moment of the particle's momentum. Our proof depends on the fact that while the particle is in the cylinders ψ_1 and ψ_2 are nonoverlapping, and requires the cylinders to be short enough and wide enough so that during the time the particle is inside the spreading of ψ_1 and ψ_2 is unimportant.

Having proved that $\langle p^n \rangle$ is unaffected for $n = 1, 2, 3, \ldots$, we can then use this result to prove that $\langle x^n \rangle$ is also unaffected. In this way we have proved that, although the interference pattern shifts when the potential is applied, the interaction that causes the shift does not affect any moment of the position or momentum of any particle contributing to the pattern.

III. MODULAR MOMENTUM

Aharonov *et al.* indicated that there is a variable whose expectation value does change in the potential effect. If we assume that the two slits are centered at $x = 0$ and $x = l$, and take p to be the x component of the momentum operator, then the variable they ask us to consider is

$$A = (p_0/2\pi)\sin 2\pi p/p_0 \tag{3.1}$$

with $p_0 = h/l$. Knowledge of A determines p within a multiple of p_0, so A is related to the momentum modulo p_0. For convenience, we simply call A the modular momentum.

We can give two heuristic arguments that a variable like A should be important in the potential effect.[16] From a physical standpoint, we know that a particle which lands at a maximum in the interference pattern will be in one of the directions θ for which

$$l\sin\theta = n\lambda \tag{3.2}$$

where n is an integer and λ the wavelength of the incident particle. This equation is easily rewritten using the de Broglie relation to read

$$p = h\sin\theta/\lambda = np_0. \tag{3.3}$$

Thus, a particle landing in a maximum has had its x component of momentum shifted to some multiple of p_0 by the two-slit barrier. If a potential difference is applied to the cylinders and the experiment run again, then the maxima are created by particles with momentum

$$p = np_0 + \bar{p} \tag{3.4}$$

where $0 < \bar{p} < p_0$. Thus, the interaction between a charged particle and the source of potential causes a shift of momentum in *fractions* of p_0. This suggests that the interaction will be described by a variable simply related to the x momentum *modulo* p_0.

From a mathematical standpoint, we note that if we express the modulo variable as a sine or cosine function, then its expectation value in the momentum representation will involve quantities like

$$\int \tilde{\psi}(p,t)^* e^{-ilp}\tilde{\psi}(p,t)\,dp \tag{3.5}$$

since $2\pi p/p_0 = lp$. If we let l be a variable, rather than fix its value as the slit separation, then the integral (3.5) is simply the Fourier transform of the momentum probability density $|\tilde{\psi}|^2$. This means that knowledge of $\sin lp$ or $\cos lp$ for all values of l will be equivalent to knowledge of the momentum distribution itself, and if the interaction changes $|\tilde{\psi}|^2$ at all, it should also change its Fourier transform, and hence $\langle A\rangle$. These considerations lead us to believe that A, as defined in Eq. (3.1), will be important in describing the interaction.

We have been able to prove the following results about A:

(1) In the Heisenberg picture, A evolves according to

$$\frac{d}{dt} A(t) = -\frac{1}{2l}\{[V(x+l,t) - V(x,t)]e^{ilp} + [V(x,t) - V(x-l,t)]e^{-ilp}\} \tag{3.6}$$

where $V(x,t)$ is the potential energy of the particle and all operators are in the Heisenberg picture. Thus, the time evolution of $A(t)$ depends upon the potential energy at several different points in space at the same time. This means that the change in $A(t)$ is determined by a nonlocal equation, even though the Hamiltonian

$$H = p^2/2m + V(x,t)$$

is local.

(2) If we multiply Eq. (3.6) on the left by $\psi(x,0)^*$, on the right by $\psi(x,0)$, and integrate, we find

$$\frac{d}{dt}\langle A\rangle = -\int_{-\infty}^{+\infty}\left[\frac{V(x+l,t) - V(x,t)}{l}\right] \times \left[\frac{\psi(x,0)^*\psi(x+l,0) + \psi(x+l,0)^*\psi(x,0)}{2}\right]dx. \tag{3.7}$$

This equation was derived by Aharonov *et al.* in the Schrödinger picture, is true for any wavefunction, and gives the time evolution of $\langle A\rangle$. As Aharonov *et al.* note, in the limit as l goes to zero, Eq. (3.7) reduces to Ehrenfest's theorem

$$\frac{d}{dt}\langle p\rangle = \left\langle -\frac{\partial V}{\partial x}\right\rangle. \tag{3.8}$$

This means that Eq. (3.7) is a generalization of Ehrenfest's theorem, and that we can identify the potential difference inside the integral as a generalized, nonlocal force. If we confine ourselves to two-slit interference experiments, then $\psi(x,0)$ is very well peaked about the slits at $x = 0$ and $x = 1$, and Eq. (3.7) becomes

$$\frac{d}{dt}\langle A\rangle = -\left[\frac{V(l,t) - V(0,t)}{l}\right]\langle\cos lp\rangle. \tag{3.9}$$

This shows clearly that it is finite differences of the potential energy, and not forces, that lead to a change in $\langle A \rangle$.

(3) If we evaluate $\langle A \rangle$ in the Schrödinger picture,

$$\langle A \rangle = \frac{1}{2i} \int_{-\infty}^{+\infty} \psi(x,t)^* [(e^{ilp} - e^{-ilp})] \psi(x,t)\, dx \tag{3.10}$$

we find that when the particle hits the screen

$$\langle A \rangle = (\hbar/2l) \sin \alpha. \tag{3.11}$$

Thus, $\langle A \rangle$ *is* changed by the interaction.

(4) The results of (2) and (3) together lead to an interesting mathematical question. Suppose that, rather than expressing the sine function in Eq. (3.10) as the sum of two exponentials, we had used instead its power series representation. In this case we might have argued that, since none of the moments of the momentum is affected by the interaction, the modular momentum should also remain unaffected.

The problem with this argument is that the power series representation of the sine function cannot be applied to all wavefunctions. At the very least it can only be used on wavefunctions that are infinitely differentiable, and for which the resulting power series is uniformly convergent. On the other hand, a wavefunction need only be $\mathscr{L}_2(\mathbf{R})$ to be a member of the Hilbert Space. In our case, even though our two-slit wavefunctions may be infinitely differentiable, they are not analytic. This is because the physical conditions require them to be identically zero in an interval along the x-axis, and the only analytic function for which this is true is the trivial function. Since our two-slit wavefunctions are not analytic, sin lp acting on them does not produce the same result as its power series, and the power series argument is not valid.

The same objection does not apply when we represent the sine function as the difference of two exponentials. The exponentials can be defined by their power series on a set of vectors dense in $\mathscr{L}_2(\mathbf{R})$, called analytic vectors, and their properties established in the usual way.[17,18] Then, because the exponentials are continuous unitary operators their extension to the rest of the Hilbert Space is unique, and their properties apply to any vector in the space, including two-slit wavefunctions. The interesting point here is that because of the nonanalyticity of the two-slit wavefunctions, which is required by the physical situation, the interaction in the potential effect can change the modular momentum while leaving all the moments of the linear momentum unchanged.

IV. SUMMARY

Equations (3.9) and (3.11) show that modular momentum gives a new formalism and language for describing the potential effect. According to Eq. (3.9) the modular momentum of the particle is conserved unless a generalized nonlocal force acts. This generalized nonlocal force can cause a change in the modular momentum even when none of the moments of the linear momentum are affected. This is in fact what happens since the generalized nonlocal force is nonzero even though the local force vanishes wherever the particle can be found. We note that this description of the potential effect, in terms of the change in modular momentum, is quite similar to the classical description of interactions. The work of Aharonov *et al.* indicates that such an approach can be generalized to other variables, such as energy, angular momentum, and position, and that not only will it lead to a new understanding of known quantum effects, but it also could provide a way to make new predictions of effects peculiar to quantum mechanics which have no classical analog.

Acknowledgements

We would like to thank Professor Aharonov for making available to us some lecture notes on this subject, and for several helpful conversations. We also would like to thank Professor K. Jagannathan and Professor R. H. Romer for many helpful discussions.

Note Added in Proof

For proofs of the results described here, and their extension to the magnetic A–B effect, see our paper, 'Expectation Values in the Aharonov-Bohm Effect', submitted to *Phys. Rev.* *D***15**.

References

1. Y. Aharonov and D. Bohm, *Phys. Rev.* **115**, 485 (1959).
2. D. Wisnivesky and Y. Aharonov, *Ann. Phys.* (*NY*) **45**, 479 (1967).
3. T. T. Wu and C. N. Yang, *Phys. Rev.* *D***12**, 3845 (1975).
4. J. Stachel, *Phys. Rev.* *D***26**, 1281 (1982).
5. Y. Aharonov and G. Carmi, *Found. Phys.* **3**, 493 (1973).
6. J. H. Harris and M. D. Semon, *Found. Phys.* **10**, 151 (1980).
7. M. D. Semon, *Found. Phys.* **12**, 49 (1982).
8. Y. Aharonov and D. Bohm, *Phys. Rev.* **123**, 1511 (1962); **125**, 2192 (1963); **130**, 1625 (1963).
9. B. S. DeWitt, *Phys. Rev.* **125**, 2189 (1962).
10. P. D. Noerdlinger, *Nuovo Cim.* **23**, 158 (1962).
11. F. J. Belinfante, *Phys. Rev.* **128**, 2832 (1962).
12. E. J. Konopinski, *Am. J. Phys.* **46**, 499 (1978).
13. H. Erlichson, *Am. J. Phys.* **38**, 162 (1970).
14. D. Home and S. Sengupta, *Am. J. Phys.* **51**, 942 (1983).
15. P. Frolov and V. D. Skarzhinsky, *Nuovo Cim.* **76B**, 35 (1983).
16. Y. Aharonov, H. Pendleton and A. Petersen, *Int. J. Theor. Phys.* **2**, 213 (1969).
17. T. F. Jordan, *Linear Operators for Quantum Mechanics* (Wiley, New York, 1969), p. 112.
18. M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. II (Academic Press, New York, 1975), pp. 200ff.