IL NUOVO CIMENTO VOL. 47 A, N. 4 21 Ottobre 1978

Nonexistence of the Aharonov-Bohm Effect.

P. BOCCHIERI

Istituto di Fisica Teorica dell'Università - Pavia, Italia
Istituto Nazionale di Fisica Nucleare - Sezione di Pavia, Italia

A. LOINGER

Istituto di Scienze Fisiche dell'Università - Milano, Italia

(ricevuto il 2 Giugno 1978)

Summary. — In this paper the Aharonov-Bohm effect is investigated and it is shown that it has a purely mathematical origin. All the physical consequences of quantum mechanics turn out to be dependent on the field strengths and not on the potentials.

1. – As is well known (1), the canonical commutation rules $q_r p_s - p_s q_r = i\hbar\delta_{rs}$ $(r, s = 1, 2, ..., n)$ give, when the q's are diagonal, the following expressions for the p's:

$$p_s = -i\hbar \frac{\partial}{\partial q_s} + \frac{\partial F(q)}{\partial q_s}, \tag{1}$$

$F(q)$ being an arbitrary real function of $q_1, q_2, ..., q_n$. In going from a given F to another, the state functions are multiplied by a phase factor $\exp[i\gamma(q)]$. In particular, going to the representation in which $F = 0$, one has

$$\gamma(q) = \frac{F(q)}{\hbar} + \text{const}. \tag{2}$$

Generally, these changes have no observable consequences, except when the function F is discontinuous.

(1) Cf. J. VON NEUMANN: *Math. Ann.*, **104**, 570 (1931); P. A. M. DIRAC: *The Principles of Quantum Mechanics*, Fourth Edition (Oxford, 1958), p. 92.

Let us remark that the existence of different operator representations for the p's rests only on the structure of the commutation rules. A change of operator representation can always be performed by keeping the electromagnetic potentials *fixed.*

2. – Let us consider a free particle moving on a circumference of radius a (free plane rotator). Nowhere in the space an electromagnetic field is present. The customary Hamiltonian of the system is

$$H_0 = \frac{1}{2Ma^2}\left(-i\hbar\frac{\partial}{\partial\varphi}\right)^2, \tag{3}$$

M being the mass and φ $(0 \leqslant \varphi < 2\pi)$ the anomaly of the particle ([2]). In another representation of p_φ the Hamiltonian is

$$H = \frac{1}{2Ma^2}\left[-i\hbar\frac{\partial}{\partial\varphi} + \frac{\partial F(\varphi)}{\partial\varphi}\right]^2. \tag{4}$$

As we shall see, there are two cases. When $F(\varphi)$ is continuous, in particular if $\lim\limits_{\varphi\to 2\pi} F(\varphi) = F(0)$, H and H_0 have the same eigenvalues, and the corresponding eigenfunctions $u_m(\varphi)$, $u_{0,m}(\varphi)$ $(m = 0, \pm 1, \pm 2, \ldots)$ satisfy the relation

$$u_m(\varphi) = u_{0,m}(\varphi)\exp\left[-\frac{i}{\hbar}F(\varphi)\right]. \tag{5}$$

The eigenfunctions u_m, $u_{0,m}$ are obtained under continuity conditions. In particular, one has $\lim\limits_{\varphi\to 2\pi} u_m(\varphi) = u_m(0)$, $\lim\limits_{\varphi\to 2\pi} u_{0,m}(\varphi) = u_{0,m}(0)$. On the contrary, when $\lim\limits_{\varphi\to 2\pi} F(\varphi) \neq F(0)$, the eigenvalues of H and H_0 are different and eqs. (5) do not hold any longer (see sect. **3**). One has equivalence only among operator representations of p_φ for which the difference $\lim\limits_{\varphi\to 2\pi} F(\varphi) - F(0)$ has the same value.

The solutions $\psi(\varphi, t)$ and $\psi_0(\varphi, t)$ of the time-dependent Schrödinger equations $H\psi = i\hbar(\partial\psi/\partial t)$ and $H_0\psi_0 = i\hbar(\partial\psi_0/\partial t)$ are connected by

$$\psi(\varphi, t) = \psi_0(\varphi, t)\exp\left[-\frac{i}{\hbar}F(\varphi)\right]; \tag{5'}$$

([2]) « We remind the reader that in using nonrectangular co-ordinates one has an observable associated to a co-ordinate only when it is a *single*-valued *real*-valued function on the manifold of all configurations. Thus in using polar co-ordinates r, θ, φ for a single particle there will be no self-adjoint operator corresponding to θ or φ, as these are multiple-valued functions. On the other hand, there will be one corresponding to $\sin\theta$, $\cos\theta$, $\sin\varphi$, $\cos\varphi$, etc. Also there will be one corresponding to the *discontinuous* single-valued function obtained by insisting that φ be between 0 and 2π ». (G. W. MACKEY: *The Mathematical Foundations of Quantum Mechanics*, A Lecture-Note Volume (New York, N. Y., etc., 1963), p. 103.)

of course, if one of the two functions ψ, ψ_0 is continuous, the other is discontinuous, unless both of them belong to equivalent representations. The reverse is also true: if ψ, ψ_0 are continuous, eq. (5') cannot hold if ψ and ψ_0 belong to inequivalent representations.

3. – Let us define a function $f(\varphi)$ as follows:

$$\frac{\partial F(\varphi)}{\partial \varphi} = -\hbar\alpha f(\varphi)\,, \tag{6}$$

where α is an adimensional constant, and

$$\int\limits_0^{2\pi} f(\varphi)\,\mathrm{d}\varphi = 2\pi\,. \tag{7}$$

The solution $\chi(\varphi)$ of the equation

$$H\chi(\varphi) = E\chi(\varphi) \tag{8}$$

is

$$\chi(\varphi) = \text{const}\exp\left[im'\varphi + i\alpha\int\limits_0^{\varphi} f(\varphi')\,\mathrm{d}\varphi'\right] \tag{9}$$

with

$$E \equiv \frac{\hbar^2 m'^2}{2Ma^2}\,. \tag{10}$$

If the eigenvalue problem (9) is solved by imposing the continuity condition on the wave function χ,

$$\lim_{\varphi\to 2\pi}\chi(\varphi) = \chi(0)\,, \tag{11}$$

one obtains (cf. MERZBACHER [3])

$$E \equiv E_m = \frac{\hbar^2}{2Ma^2}(m-\alpha)^2 \qquad (m = 0, \pm 1, \pm 2, \ldots)\,, \tag{12}$$

$$\chi(\varphi) \equiv \chi_m(\varphi) = \text{const}\exp[im\varphi]\exp\left[i\alpha\left[\int\limits_0^{\varphi} f(\varphi')\,\mathrm{d}\varphi' - \varphi\right]\right]. \tag{13}$$

---

(3) W. EHRENBERG and R. E. SIDAY: *Proc. Phys. Soc.*, **62** B, 8 (1949); Y. AHARONOV and D. BOHM: *Phys. Rev.*, **115**, 485 (1959); **123**, 1511 (1961); **125**, 2192 (1962); W. H. FURRY and N. F. RAMSEY: *Phys. Rev.*, **118**, 623 (1960); M. PESHKIN, I. TALMI and L. J. TASSIE: *Ann. of Phys.*, **12**, 426 (1961); E. MERZBACHER: *Amer. Journ. Phys.*, **30**, 237 (1962); R. P. FEYNMAN, R. B. LEIGHTON and M. SANDS: *The Feynman Lectures on Physics* (Reading, Mass., 1964), sect. II-**15**-12; T. T. WU and C. N. YANG: *Phys. Rev. D*, **12**, 3845 (1975); C. MARTIN: *Lett. Math. Phys.*, **1**, 155 (1976); W. DRECHSLER and M. E. MAYER: *Fiber Bundle Techniques in Gauge Theories*, Lecture Notes in Physics (Berlin, etc., 1977) part two, p. 14; and many other papers.

Consequently, the choice of the representation of the operator p_φ affects the *differences* among the energy eigenvalues (unless α is an integer). This is due to the fact that only the differences among the eigenvalues of the angular momentum—and not their absolute magnitudes—are determined by the rotational symmetry of the system.

These conclusion are a logical consequence of the fact that one requires that the eigenfunctions are continuous in both the considered representations. As we shall see in sect. **5**, the eigenvalues can be made representation independent by means of slight modifications of the conditions imposed on the state functions.

Let us now remark that the change of operator representation $i\hbar(\partial/\partial\varphi) \to \to i\hbar(\partial/\partial\varphi) + \hbar\alpha f(\varphi)$ produces on the energy eigenvalues and eigenfunctions an effect which is identical with the so-called magnetic effect of the vector potential, first discovered by EHRENBERG and SIDAY ([3]) and then investigated by AHARONOV and BOHM ([3]).

One is easily convinced that this is the case if

i) one takes $\alpha = e\Phi/(2\pi\hbar c)$, where e is the electric charge of the particle and Φ the magnetic flux of an infinitely long solenoid, having a radius $r_0 < a$; the axis of the solenoid goes through the centre of the rotator and is perpendicular to its plane;

ii) one interprets $\Phi f(\varphi)/(2\pi a)$ as the A_φ-component of the Stokesian (in the whole space) vector potential $\boldsymbol{A}$ created by the solenoid on the circumference of radius a; the potential $\boldsymbol{A}$ is (for $r \geqslant r_0$) $\boldsymbol{A} = \{A_r = A_z = 0;\ A_\varphi = = \Phi f(\varphi)/(2\pi r)\}$.

Therefore the Aharonov-Bohm effect has a purely mathematical origin. This is consistent with the fact that in this problem there is no physical reason to prefer the above $\boldsymbol{A}$ to the following non-Stokesian one, which gives the same field strengths:

$$\boldsymbol{A} = 0 \qquad \text{for } r \geqslant r_0, \tag{14}$$

$$A_r = A_z = 0\,,\quad A_\varphi = (\Phi/2\pi r_0^2) f(\varphi) r - \Phi f(\varphi)/(2\pi r) \qquad \text{for } 0 < r \leqslant r_0\,. \tag{15}$$

Obviously, in this gauge there is no Aharonov-Bohm effect. This way of reasoning shows that the effect is gauge-dependent; this seems astonishing, at first, as in eq. (12) only Φ appears, which is a gauge-invariant concept.

4. – In effect, in their papers([3]) AHARONOV and BOHM have studied problems a little different from that of the free rotator. In particular, they investigated the diffraction of a Schrödinger electron wave by an infinite solenoid of magnetic flux Φ, shielded by an impenetrable cylindrical wall; in this way, the electron is never subject to a Lorentz force (cf. sect. **4** of their first paper). In their treatment, they use the Stokesian vector potential, having components

$A_r = A_z = 0$, $A_\varphi = \Phi f(\varphi)/(2\pi r)$, which we have already seen. In the limit in which the radius r_0 of the solenoid is vanishing, under the condition that Φ remains constant, AHARONOV and BOHM find the following expression for the cross-section $\mathrm{d}\sigma/\mathrm{d}\varphi$:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\varphi} = \frac{\sin^2 \pi\alpha}{2\pi k} \frac{1}{\cos^2(\varphi/2)}. \tag{16}$$

Equation (16) has been deduced integrating the Schrödinger equation:

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial}{\partial\varphi} + i\alpha\right)^2 + k^2\right]\eta(r, \varphi) = 0\,, \tag{17}$$

where $\boldsymbol{k}$ is the wave vector of the incident particle, and for simplicity's sake $f(\varphi) = 1$ has been chosen. Evidently, the considerations made for the plane rotator hold also in this case, and therefore the cross-section (16) describes a gauge-dependent effect. One concludes, in particular, that the interference experiments [4] performed with the aim of verifying the Aharonov-Bohm effect must in reality concern something else, *e.g.* effects due to leakage fields of solenoids, magnetic « whiskers », etc., as pointed out by PRYCE (cf. CHAMBERS [4]). The following sentence taken from a paper by MARTON *et al.* [4] is very significant: « The tilted fringes [...] can be fully explained by the assumption of a flux leaking normal to the axis through the surface of the whisker. No other types of magnetic field are required to explain this result. [...] The fringe tilt may be thought to be entirely due either to *a*) the changing internal flux, as AHARONOV and BOHM propose, or to *b*) the leaking field which skews the interferometer beams, as PRYCE [...] proposes. Flux closure demands a leakage field in the region of changing flux; hence, the two descriptions are equivalent ».

5. – This unsatisfactory situation can be completely avoided if one remarks that the continuity of the state functions is too restrictive a criterion, not necessary both from a physical and a mathematical point of view. A correct requirement is that probability density $\varrho = \bar{\psi}\psi$ and current density $\boldsymbol{j}$ are continuous in all the representations of $\boldsymbol{p}$ and for all the state functions ψ. For the plane rotator one has

$$j_\varphi = \frac{\hbar}{2Mai}\left(\bar{\psi}\frac{\partial\psi}{\partial\varphi} - \psi\frac{\partial\bar{\psi}}{\partial\varphi}\right) + \frac{1}{Ma}\bar{\psi}\psi\frac{\partial F}{\partial\varphi}\,. \tag{18}$$

---

(4) F. G. WERNER and D. R. BRILL: *Phys. Rev. Lett.*, **4**, 349 (1960); R. G CHAMBERS: *Phys. Rev. Lett.*, **5**, 3 (1960); H. BOERSCH, H. HAMISCH, K. GROHMANN and D. WOHLLEBEN: *Zeits. Phys.*, **165**, 79 (1961); **169**, 263 (1962); H. A. FOWLER, L. MARTON, J. AROL SIMPSON and J. A. SUDDETH: *Journ. Appl. Phys.*, **32**, 1153 (1961); W. BAYH: *Zeits. Phys.*, **169**, 492 (1962); B. LISCHKE: *Phys. Rev. Lett.*, **22**, 1366 (1969); G. MATTEUCCI, G. F. MISSIROLI and G. POZZI: *Giorn. di Fis.*, **18**, 264 (1977); and many other papers.

If one writes

$$\psi = \overset{+}{\sqrt{\varrho}} \exp [iS] , \tag{19}$$

a reasonable condition is that *ψ is single valued and S has the same discontinuities* (mod 2π) *as* $-F(\varphi)/\hbar$. It is straightforward to verify that in this case j_φ is continuous. Then the eigenvalues are independent of F, and the eigenfunctions ω_m are related to those, $\omega_{0,m}$, of the case $F=0$ by the equation

$$\omega_m = \omega_{0,m} \exp\left[-\frac{i}{\hbar} F\right] \equiv \text{const} \exp [im\varphi] \exp\left[-\frac{i}{\hbar} F\right] . \tag{20}$$

So all the operator representations of $\boldsymbol{p}$ are physically equivalent and the problem of Aharonov and Bohm vanishes.

6. – In 1949 EHRENBERG and SIDAY published a very interesting paper ([3]), that can be considered, from a conceptual point of view, as a natural continuation of the famous Schrödinger's second memoir of 1926 ([5]). In their paper, which is based on the *semi-classical* method typical of electron optics, they show, in particular, that the electron refractive index μ is proportional to the scalar product of the classical canonical momentum $\boldsymbol{p}^{(\text{cl})} = M\boldsymbol{v} + (e/c)\boldsymbol{A}$ by the unit vector tangent to the classical trajectory. Time-independent electromagnetic fields are assumed. The ratio of the de Broglie wave-length $\lambda_0 = h/(Mv)$, for $\boldsymbol{A}=0$, to the wave-length λ is equal to the inverse ratio μ/μ_0 of the corresponding refractive indices. Therefore one has

$$\lambda = \frac{h}{Mv + (e/cv)\,\boldsymbol{v}\cdot\boldsymbol{A}} . \tag{21}$$

The optical path of a given electron ray R is equal to $\int_R \mu \, ds$.

Now, we recall that even in classical dynamics the canonical momentum $\boldsymbol{p}^{(\text{cl})}$ is fixed within the gradient of an arbitrary function. Thus, so far as the problems investigated in the preceding sections are concerned, we are always allowed to perform the formal substitution

$$p_\varphi^{(\text{cl})} \to p_\varphi^{(\text{cl})} + \text{const}\,\frac{f(\varphi)}{r} , \tag{22}$$

that does not modify in any way the classical electron trajectories. Therefore, contrary to a claim made by EHRENBERG and SIDAY in the last page of their work, also in the semi-classical approach of electron optics the vector potential does not produce any physical effect, if the electron moves in a region in which the magnetic field is zero. Indeed, the effect of the Stokesian vector

([5]) E. SCHRÖDINGER: *Ann. der Phys.*, **79**, 489 (1926).

potential on the de Broglie wave-length λ can be exactly counterbalanced by the change (22) in the realization of the classical canonical momentum.

Finally, we point out that EHRENBERG and SIDAY require that the vector potential is Stokesian in the whole space, as only in this case the refractive index is everywhere continuous, even inside the solenoid. The following potential, which we have already mentioned, $\boldsymbol{A}=0$ for $r\geqslant r_0$, $A_r=A_z=0$, $A_\varphi=(\Phi/2\pi r_0^2)f(\varphi)r-\Phi f(\varphi)/(2\pi r)$ for $0<r\leqslant r_0$, is, on the contrary, not Stokesian everywhere and becomes infinite at the origin. However, if we isolate the solenoid by means of a cylindrical wall, this last potential can still be chosen also in the semi-classical approach of Ehrenberg and Siday, thus removing again the Aharonov-Bohm effect.

7. – It has been known since long that the Schrödinger equation

$$\left\{\frac{1}{2M}\left[-i\hbar\nabla-\frac{e}{c}\boldsymbol{A}(\boldsymbol{x},t)\right]^2+eV(\boldsymbol{x},t)\right\}\psi(\boldsymbol{x},t)=i\hbar\frac{\partial\psi(\boldsymbol{x},t)}{\partial t}\tag{23}$$

is equivalent to a system of nonlinear differential equations, of «hydrodynamical» type, *in which only the field strengths, and not the potentials, appear.* Here we shall follow the procedure of Nelson (⁶). Let us write

$$\left\{\begin{aligned}&\psi=\exp[R+iS] && (R,\,S\text{ real functions})\,,\\&\nabla R(\boldsymbol{x},t)=\frac{M}{\hbar}\boldsymbol{u}(\boldsymbol{x},t)\,,\\&\nabla S(\boldsymbol{x},t)=\frac{M}{\hbar}\left[\boldsymbol{w}(\boldsymbol{x},t)+\frac{e}{Mc}\boldsymbol{A}(\boldsymbol{x},t)\right].\end{aligned}\right.\tag{24}$$

Remark that $\boldsymbol{w}(\boldsymbol{x},t)$ is gauge-invariant even if one goes from a Stokesian to a non-Stokesian potential. In this case the discontinuity of $F(\varphi)$ at $\varphi=0$ is exactly neutralized by the discontinuity of S, and the expression $\nabla S-(e/\hbar c)\boldsymbol{A}$ turns out to be continuous and gauge-invariant. Of course, the same happens for the pure change of operator representation

$$-i\hbar\nabla\to-i\hbar\nabla+\nabla F(\boldsymbol{x})\,.$$

NELSON proves that eq. (23) is completely equivalent to the following system

(⁶) E. NELSON: *Dynamical Theories of Brownian Motion,* Mathematical Notes (Princeton, N. J., 1967), p. 130. Cf. also G. W. MACKEY: *The Mathematical Foundations of Quantum Mechanics,* A Lecture-Note Volume (New York, N. Y., etc., 1963), p. 96.

of differential equations:

$$\text{(25)}\qquad \begin{cases} \dfrac{\partial \boldsymbol{u}}{\partial t} = -\dfrac{\hbar}{2M}\nabla(\nabla\cdot\boldsymbol{w}) - \nabla(\boldsymbol{u}\cdot\boldsymbol{w}), \\ \dfrac{\partial \boldsymbol{w}}{\partial t} = \dfrac{\hbar}{2M}\nabla^2\boldsymbol{u} - (\boldsymbol{w}\cdot\nabla)\boldsymbol{w} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \dfrac{e}{M}\left(\boldsymbol{E} + \dfrac{\boldsymbol{w}}{c}\wedge\boldsymbol{B}\right), \end{cases}$$

under the conditions

$$\text{(25')}\qquad \nabla\wedge\boldsymbol{u} = 0\,, \qquad \nabla\wedge\boldsymbol{w} + \frac{e}{Mc}\boldsymbol{B} = 0\,.$$

Notice that the usual current density $\boldsymbol{j}$ is given by the product $\varrho\boldsymbol{w}$, where $\varrho = \bar{\psi}\psi = \exp[2R]$, and the standard statistical interpretation of quantum mechanics can be maintained also in this formalism.

The « hydrodynamical » equations show that, in their time evolution, ϱ and $\boldsymbol{j}$ depend only on the fields $\boldsymbol{E}$, $\boldsymbol{B}$ and not on the potentials V, $\boldsymbol{A}$: therefore they leave no room for effects of the kind of Aharonov and Bohm's.

* * *

We thank heartily our friend G. SIRAGUSA for useful discussions.

RIASSUNTO

In questo lavoro si studia l'effetto Aharonov-Bohm e si dimostra che esso ha un'origine puramente matematica. Tutte le conseguenze fisiche della meccanica quantistica dipendono solo dai campi elettromagnetici e non dai potenziali.

Об отсутствии эффекта Ааронова-Бома.

Резюме (*). — В этой статье исследуется эффект Ааронава-Бома. Показывается, что этот эффект имеет чисто математическое происхождение. Оказывается, что все физические следствия квантовой механики зависят от интенсивностей полей, а не от потенциалов.

(*) *Переведено редакцией.*