# Old and new problems in the theory of the Aharonov–Bohm effect

G. N. Afanas'ev

*Joint Institute for Nuclear Research, Dubna*

Fiz. Elem. Chastits At. Yadra **21**, 172–250 (January–February 1990)

The current status of the theory of the Aharonov–Bohm effect is reviewed. Special attention is given to discussions of recent years on the existence of this effect. The experimental situation is analyzed, and the crucial experiments testing the AB effect are discussed.

## INTRODUCTION

In classical electrodynamics the force acting on a charged particle is determined by the electric and magnetic field strengths **E** and **H**:

$$\mathbf{F} = e\left(\mathbf{E} + \frac{1}{c}[\mathbf{V},\ \mathbf{H}]\right).$$

This implies the absence of scattering in regions of space in which $\mathbf{E} = \mathbf{H} = 0$. The electric scalar potential $\varphi$ and the magnetic vector potential **A** in the classical theory of electromagnetism play an auxiliary role: they mainly serve to simplify the field equations. Arbitrary gauge transformations leaving the electromagnetic field strength unchanged are allowed. Therefore, the field strengths **E** and **H** are the fundamental quantities in classical mechanics.

The situation is completely different in quantum mechanics. The reason is that the Schrödinger equation involves the electromagnetic potentials rather than the field strengths. Ehrenberg and Siday[1] were the first to point out the special role of the electromagnetic potentials in quantum mechanics. Ten years passed before Aharonov and Bohm worked on this problem.[2] They studied the scattering of charged particles on the magnetic field of a cylindrical solenoid. It was found that scattering occurs, even if the region where the field strength **E**, $\mathbf{H} \neq 0$ is inaccessible to the incident particles.[1)] This phenomenon, i.e., the physical effects of the fields inaccessible to the particles, is referred to as the Aharonov–Bohm effect (the AB effect). In their study Aharonov and Bohm used single-valued wave functions. However, in a multiply connected space (for example, the regions of space outside an impenetrable cylindrical solenoid) it is possible to have inequivalent representations of the angular momentum which correspond to non-single-valued (more accurately, discontinuous) wave functions. The equivalent status of these representations and the lack of any reliable criterion for singling out any one of them have, in recent years, led to lively discussion of the existence of the AB effect.

On the other hand, experiments involving electron scattering on a toroidal solenoid show that the diffraction pattern shifts when a current is switched on in the solenoid. In the so-called alternative (or hydrodynamical) interpretation of the AB effect, the positive results of such experiments are attributed to the penetration of the incident particles into the region where $\mathbf{H} \neq 0$. There are several reasons for this. Since a real cylindrical solenoid has a finite length, the magnetic field leaks out near the ends of the solenoid. This allows the results of experiments with cylindrical solenoids to be attributed to scattering on the tails of the magnetic field. The second reason is the finite value of a real potential, which prevents the penetration of particles into the region with $\mathbf{H} \neq 0$. An additional complication for an infinite cylindrical solenoid is the slow falloff of the vector potential, leading to significant distortion of the incident wave, which ceases to be single-valued. This leads to numerous paradoxes with angular-momentum nonconservation.

The plan of our discussion is as follows. In Sec. 1 we study the scattering of charged particles on an infinite cylindrical solenoid. We show that the discrepancy between the results of Ref. 3 and the Aharonov–Bohm results is due to a defect in the Born approximation used in Ref. 3 (as has been pointed out in many studies; see, for example, the review of Ref. 4), and to the difference between the physical situations considered in Refs. 2 and 3. Here we give the asymptote of the AB wave function valid for all values of the scattering angle. In this same section we discuss the physical meaning and the range of applicability of the Dirac phase factor.

In Sec. 2 we discuss questions of the single-valuedness of wave functions in multiply connected spaces. We show that multivalued wave functions become allowed. This is related to the fact that (in contrast to a simply connected space) the well-known Pauli proof of the single-valuedness of the wave functions does not apply. In particular, the angular-momentum operators on such wave functions do not cease to be Hermitian.

In Sec. 3 we analyze the attempts to prove the nonexistence of the AB effect found in the recent literature. All these attempts involve the use of multivalued (rather than single-valued) wave functions.

In Sec. 4 we show that by studying the switching-on of the magnetic field in the solenoid we can partially eliminate the above-mentioned uncertainties in the choice of the wave functions.

In Sec. 5 we study the scattering of charged particles on two infinite cylindrical solenoids with oppositely directed magnetic fluxes. We are forced to consider this situation by the difficulties in the choice of the incident wave for a single solenoid.

In Sec. 6 we study the scattering of charged particles on a toroidal solenoid. This eliminates the edge effects arising from leakage of the magnetic field near the ends of a cylindrical solenoid. The scattering amplitude does not contain any singularities (in contrast to the amplitude for scattering on a cylindrical solenoid), and the integrated scattering cross section proves to be finite.

In Sec. 7 we study scattering in multiply connected spaces. We show that it is possible for the AB effect to exist or not to exist in a given multiply connected space with nontrivial vector potentials and single-valued wave functions.

 0090-4759/90/010074-32$04.00  

The conditions for the existence of the effect are formulated.

In Sec. 8 we discuss the question of what gauge transformations are allowed. It turns out that any gauge transformations are allowed (including those which completely eliminate the vector potential outside the solenoid) if we are willing to work with multivalued (after the gauge transformation) wave functions. Otherwise, we must restrict ourselves to vector potentials which satisfy the Stokes theorem.

In Sec. 9 we consider the alternative interpretation of the AB effect. It is based on relating the results of experiments on detecting the AB effect to the penetration of the incident particles into the region with $H \neq 0$. For the examples of a single cylindrical solenoid, two cylindrical solenoids with $\Phi_1 = -\Phi_2$, and a toroidal solenoid we show that this interpretation is groundless.

In Sec. 10 we discuss the AB effect for bound states. The presence of a nonzero vector potential in a region of space accessible to the particles leads to additional transitions between energy levels.

In Sec. 11 we discuss the current experimental status of the AB effect. We show that there is almost unanimous indication that the AB effect exists. We analyze and propose critical experiments which can eliminate the doubts about the existence of this effect expressed in the literature. Here we also discuss the thought experiment showing that the effect of fields inaccessible to particles can be observed also in a simply connected space.

## 1. AN ELEMENTARY EXPOSITION OF THE AHARONOV–BOHM EFFECT

### 1.1. Scattering on a single finite cylindrical solenoid

Let us consider scattering on an impenetrable cylinder $C_0$ of radius $R$, whose axis coincides with the $z$ axis. A beam of charged spinless particles is incident along the $x$ axis. The wave function and scattering amplitude are

$$\left.\begin{aligned}\Psi_0 &= \sum_m i^{|m|}\left[J_{|m|}(k\rho) - H_m^{(1)}(k\rho)\,\frac{J_{|m|}(kR)}{H^{(1)}_{|m|}(kR)}\right]\exp(im\varphi);\\ f_0(\varphi) &= -\sqrt{\frac{2}{\pi ik}}\sum_m \frac{J_{|m|}(kR)}{H^{(1)}_{|m|}(kR)}\exp(im\varphi).\end{aligned}\right\} \tag{1}$$

Here, $\rho$, $z$, and $\varphi$ are the cylindrical coordinates, $J_m$ and $H_m^{(1,2)}$ are cylindrical functions, and $k^2 = 2\mu E/\hbar^2$.

We now place inside $C_0$ a cylindrical solenoid of radius $a < R$ with its axis coinciding with that of $C_0$. Outside the solenoid only one component of the magnetic vector potential is nonzero: $A_\rho = A_z = 0, A_\varphi = \Phi/2\pi\rho$ ($H$ is the magnetic field and $\Phi$ is its flux). The wave function and scattering amplitude are

$$\left.\begin{aligned}\Psi &= \sum_m \exp\left[i\pi\left(|m| - \frac{1}{2}|m-\gamma|\right)\right]\left[J_{|m-\gamma|}(k\rho)\right.\\ &\left. - H^{(1)}_{|m-\gamma|}(k\rho)\,\frac{J_{|m-\gamma|}(kR)}{H^{(1)}_{|m-\gamma|}(kR)}\right]\exp(im\varphi),\quad \gamma = \frac{e\Phi}{hc};\\ f(\varphi) &= -\frac{1}{\sqrt{2\pi ik}}\sum_m\left\{1+\exp[i\pi(|m|\right.\\ &\left. - |m-\gamma|)]\,\frac{H^{(2)}_{|m-\gamma|}(kR)}{H^{(1)}_{|m-\gamma|}(kR)}\right\}\exp(im\varphi).\end{aligned}\right\} \tag{2}$$

Obviously, $f \neq f_0$, and it is this difference which is the essence of the AB effect.

Subtracting from the total amplitude $f$ [see (2)] the amplitude $f_0$ for scattering on an impenetrable cylinder in the absence of a magnetic field [see (1)], we find the scattering amplitude $f_1$ due to the presence of the magnetic field:

$$f_1 = f - f_0 = \frac{1}{\sqrt{2\pi ik}}\sum f_{1m}\exp(im\varphi); \tag{3}$$

$$f_{1m} = \frac{H^{(2)}_{|m|}}{H^{(1)}_{|m|}} - \exp[i\pi(|m| - |m-\gamma|)]\,\frac{H^{(2)}_{|m-\gamma|}}{H^{(1)}_{|m-\gamma|}}. \tag{4}$$

In (4) the Hankel functions depend on the argument $x = KR$. Let us consider the case of a thin solenoid ($KR \ll 1$). Then from (4) we find

$$f_{1m} = \begin{cases} -1+\exp(i\pi\gamma) & \text{for } m > \gamma;\\ -1+\exp(-i\pi\gamma) & \text{for } m < \gamma.\end{cases} \tag{5}$$

Here the series for the scattering amplitude can be summed to give

$$f_1(\varphi) = -\frac{1}{\sqrt{2\pi ik}}\,\frac{\sin\pi\gamma}{\sin\frac{\varphi}{2}}\exp\left[i\left(m_\gamma + \frac{1}{2}\right)\varphi\right]. \tag{6}$$

Here $m$ is the closest integer less than $\gamma$. If in addition $0 < \gamma \ll 1$, then

$$f_{1m} = \begin{cases} i\pi\gamma, & m \geqslant 1;\\ -i\pi\gamma, & m \leqslant 0;\end{cases}$$

$$f_1(\varphi) = -\frac{1}{\sqrt{2\pi ik}}\,\frac{\pi\gamma}{\sin\frac{\varphi}{2}}\exp\left(\frac{i\varphi}{2}\right),$$

$$\sigma_{AB} = \frac{\pi\gamma^2}{2k}\,\frac{1}{\sin^2\frac{\varphi}{2}}. \tag{7}$$

Expressions (7) were obtained by Aharonov and Bohm in 1959 (Ref. 2).

Alternatively, in the Schrödinger equation we can neglect the terms quadratic in the vector potential and move the linear terms to the right-hand side, treating them as a perturbation:

$$\Delta\Psi + k^2\Psi = \frac{2ie}{\hbar c}\mathbf{A}\nabla\Psi = \frac{2i\gamma}{\rho^2}\,\frac{\partial\Psi}{\partial\varphi}. \tag{8}$$

We seek $\Psi$ in the form $\Psi_0 + \Psi_1$. Here $\Psi_0$ is the wave function in the absence of the magnetic field [see (1)]. For $\Psi_1$ we find the equation

$$\Delta\Psi_1 + k^2\Psi_1 = \frac{2i\gamma}{\rho^2}\,\frac{\partial\Psi_0}{\partial\varphi};$$

we write $\Psi_1$ as a series in $\exp(im\varphi)$:

$$\Psi_1 = \Sigma\Psi_{1m}\exp(i\,m\varphi).$$

Here $\Psi_{1m}$ satisfies the equation

$$\left.\begin{aligned}&\frac{d^2\Psi_{1m}}{d\rho^2} + \frac{1}{\rho}\,\frac{d\Psi_{1m}}{d\rho} - \frac{m^2}{\rho^2}\,\Psi_{1m} = -\frac{2m\gamma}{\rho^2}\,\Psi_{0m};\\ &\Psi_{0m} = i^{|m|}\left[J_{|m|}(k\rho) - H^{(1)}_{|m|}(k\rho)\,\frac{J_{|m|}(kR)}{H^{(1)}_{|m|}(kR)}\right];\end{aligned}\right\} \tag{9}$$

$\Psi_{1m}$ can be expressed in terms of the Green's function $G_{0m}$ $(\rho,\rho')$ and $\Psi_{0m}$:

$$\left.\begin{aligned}&\Psi_{1m} = -2m\gamma\int\frac{d\rho'}{\rho'}\,G_{0m}(\rho,\rho')\,\Psi_{0m}(\rho');\\ &G_{0m}(\rho,\rho') = \frac{\pi}{2i}H^{(1)}_{|m|}(k\rho_>)\\ &\quad\times\left[J_{|m|}(k\rho_<) - H^{(1)}_{|m|}(k\rho_<)\,\frac{J_{|m|}(kR)}{H^{(1)}_{|m|}(kR)}\right];\\ &\Delta G_{0m} + k^2 G_{0m} = \frac{1}{\rho}\,\delta(\rho-\rho').\end{aligned}\right\} \tag{10}$$

Finally, taking $\rho \to \infty$ in (10), we find the partial components of the scattering amplitude in the Born approximation:

$$f^{\mathrm{B}}_{1m} = 2\pi i m\gamma \int_R^\infty \frac{d\rho}{\rho}\left[J_{|m|}(k\rho) - H^{(1)}_{|m|}(k\rho)\frac{J_{|m|}(kR)}{H^{(1)}_{|m|}(kR)}\right]^2$$

This expression can be integrated by quadratures:

$$f^{\mathrm{B}}_{1m} = 2\pi i\gamma \frac{m}{|m|}\sum_{s=0}^{|m|-1}\varepsilon_s\left[J_s(kR) - \frac{J_{|m|}(kR)}{H^{(1)}_{|m|}(kR)}H^{(1)}_s(kR)\right]^2$$
$$\varepsilon_s = (1+\delta_{0s})^{-1}. \qquad (11)$$

We note that for $m = 0$ the right-hand side of (11) vanishes. Therefore, in this approximation $f^{\mathrm{B}}_{10}$ does not contain terms linear in $\gamma$. On the other hand, the exact amplitude $f_{10}$ is nonzero both for finite $R$ [see (4)] and for $R \to 0$ [see (7)]. Let us see what the amplitudes $f^{\mathrm{B}}_{1m}$ become for $R \to 0$. From (11) for $m \neq 0$ we find

$$f^{\mathrm{B}}_{1m} = i\pi\gamma\frac{m}{|m|}. \qquad (12)$$

Comparing (7) and (12), we conclude that

$$f^{\mathrm{B}}_{1m} = f_{1m} \quad (m \neq 0);$$
$$f^{\mathrm{B}}_{10} = 0, \quad f_{10} = i\pi\gamma. \qquad (13)$$

Let us now calculate the Born scattering amplitude:

$$f_{1\mathrm{B}} = -\frac{1}{\sqrt{2\pi ik}}\pi\gamma\operatorname{ctg}\frac{\varphi}{2}. \qquad (14)$$

The amplitudes $f_1$ and $f_{1\mathrm{B}}$ lead to different differential cross sections:

$$\sigma^{\mathrm{B}}(\varphi) = \frac{\pi\gamma^2}{2k}\operatorname{ctg}^2\frac{\varphi}{2} \neq \sigma_{\mathrm{AB}}(\varphi) = \frac{\pi\gamma^2}{2k}\frac{1}{\sin^2\frac{\varphi}{2}}. \qquad (15)$$

Expressions (14) and (15) were obtained in the well-known study of Ref. 3. The author of that study attributed the difference from the corresponding quantities in (7) to the use of a non-single-valued wave function by Aharonov and Bohm. Later, it was shown[6,7] that the AB wave function is single-valued (the non-single-valuedness of the incident wave is compensated by that of the scattered wave). In many subsequent studies (see, for example, Refs. 4, 6, and 8) the result (7) of Aharonov and Bohm was viewed as being correct, and the difference of (14) and (15) from (7) was viewed as a result of the inapplicability of perturbation theory owing to the slow falloff ($1/r$) of the vector potential.

Let us see what we get if the exact amplitude (4) is expanded in powers of $\gamma$. For $m \neq 0$ we have

$$f_{1m} = \frac{H^{(2)}_{|m|}}{H^{(1)}_{|m|}} - \left(1 + i\pi\gamma\frac{m}{|m|}\right)\frac{H^{(2)}_{|m|} - \gamma\frac{m}{|m|}\frac{\partial H^{(2)}_\nu}{\partial\nu}\Big|_{\nu=|m|}}{H^{(1)}_{|m|} - \gamma\frac{|m|}{m}\frac{\partial H^{(1)}_\nu}{\partial\nu}\Big|_{\nu=|m|}}$$
$$= \gamma\frac{H^{(2)}_{|m|}}{H^{(1)}_{|m|}}\frac{m}{|m|}$$
$$\times\left[-i\pi + \frac{1}{H^{(2)}_{|m|}}\frac{\partial H^{(2)}_\nu}{\partial\nu}\Big|_{\nu=|m|} - \frac{1}{H^{(1)}_{|m|}}\frac{\partial H^{(1)}_\nu}{\partial\nu}\Big|_{\nu=|m|}\right]. \qquad (16)$$

The derivative of the Hankel functions with respect to the indices is again expressed in terms of linear combinations of Hankel functions:[9]

$$\frac{\partial H^{(1,2)}_\nu}{\partial\nu}\Big|_{\nu=|m|} = \mp\frac{i\pi}{2}H^{(1,2)}_{|m|}$$
$$+\frac{m!}{2}\sum_{k=0}^{m-1}\left(\frac{x}{2}\right)^{k-m}\frac{1}{k!}\frac{1}{m-k}H^{(1,2)}_k. \qquad (17)$$

Subtracting these expressions from each other, we find

$$\left[\frac{1}{H^{(2)}_{|m|}}\frac{\partial H^{(2)}_\nu}{\partial\nu} - \frac{1}{H^{(1)}_{|m|}}\frac{\partial H^{(1)}_\nu}{\partial\nu}\right]_{\nu=|m|} = i\pi$$
$$+\frac{i\,|m|!}{J^2_{|m|}+N^2_{|m|}}\sum_{k=0}^{|m|-1}\left(\frac{x}{2}\right)^{k-|m|}\frac{J_kN_{|m|}-J_{|m|}N_k}{k!\,(|m|-k)}.$$

Substituting this into (16), we obtain

$$f_{1m} = \gamma\frac{m}{|m|}\frac{H^{(2)}_{|m|}}{H^{(1)}_{|m|}}\frac{im!}{J^2_{|m|}+N^2_{|m|}}$$
$$\times\sum_{k=0}^{|m|-1}\left(\frac{x}{2}\right)^{k-|m|}\frac{J_kN_{|m|}-J_{|m|}N_k}{k!\,(m-k)}. \qquad (18)$$

In order to make a comparison with (5), we must let the argument of the Bessel functions $x = kR$ tend to zero in (18). Here we keep in mind the fact that the difference $J_k(x)N_{|m|}(x) - J_{|m|}(x)N_k(x)$ is expressed as a finite series in inverse powers of $x$ (Ref. 9). The maximum exponent encountered in this series is $|m| - k$ ($0 \leqslant k \leqslant |m| - 1$). This implies that the largest contribution to the sum (18) for $x \to 0$ comes from the term with $k = 0$:

$$J_0N_{|m|} - N_0J_{|m|} \approx -\frac{1}{\pi}\left(\frac{2}{x}\right)^{|m|}(|m|-1)!$$

Taking into account the behavior of $J_{|m|}(x)$ and $N_{|m|}(x)$ at small $x$ $J_m \approx (x/2)^m(1/m!)$, $N_m \approx -((m-1)!/\pi)(2/x)^m$, we find

$$\frac{im!}{J^2_m+N^2_m}\sum_{k=0}^{m-1}\left(\frac{x}{2}\right)^{k-m}\frac{J_kN_m-J_mN_k}{k!\,(m-k)} \to -i\pi \quad (x \to 0) \qquad (19)$$

Therefore, for $m \neq 0$,

$$f_{1m} = i\pi\gamma\frac{m}{|m|},$$

which coincides with (7). The case $m = 0$ requires separate treatment. We have[9]

$$\frac{\partial H^{(1)}_\nu}{\partial\nu}\Big|_{\nu=0} = -\frac{i\pi}{2}H^{(1)}_0, \quad \frac{\partial H^{(2)}_\nu}{\partial\nu}\Big|_{\nu=0} = \frac{i\pi}{2}H^{(2)}_0, \qquad (20)$$

i.e., the sum in (17) does not contribute. For $f_{10}$ we find

$$f_{10} = \frac{H^{(2)}_0}{H^{(1)}_0} - \exp(-i\pi\gamma)\frac{H^{(2)}_\nu}{H^{(1)}_\nu}$$
$$= \frac{H^{(2)}_0}{H^{(1)}_0} - (1-i\pi\gamma)\frac{H^{(2)}_0 + \frac{i\pi\gamma}{2}H^{(2)}_0}{H^{(1)}_0 - \frac{i\pi\gamma}{2}H^{(1)}_0} \approx 0 \qquad (21)$$

(we neglect terms of order $\gamma^2$ and higher).

Therefore, the limit $\gamma \to 0$, $kR \to 0$ of the original expression (2) coincides with the limit $kR \to 0$ of the Born amplitude. This implies that the difference between the cross sections $\sigma^{\mathrm{B}}$ and $\sigma_{\mathrm{AB}}$ [see (15)] is not a consequence of the inapplicability of the Born approximation, as stated, for example, in the review of Ref. 4. Let us consider this question in more detail.[10] In going from (4) to (15) we assumed that for small $kR$ the relation $H^{(2)}_{|m-\gamma|}/H^{(1)}_{|m-\gamma|} \approx -1$ holds with good accuracy or, equivalently,

$$|J_{|m-\gamma|}(x)| \ll |N_{|m-\gamma|}(x)|. \qquad (22)$$

From the behavior of $J_\nu(x)$ at small $x$ and arbitrary $\nu$, $J_\nu(x) \approx (x/2)^\nu(1/\Gamma(1+\nu))$, and the definition of $N_\nu$ $((J_\nu\cos\nu\pi - J_{-\nu})/\sin\pi\nu)$, it follows that the inequality (22) is certainly satisfied if $m \neq 0$ (we recall that $0 \leqslant \gamma \leqslant \frac{1}{2}$). We still need to consider the case $m = 0$. The condition (22) is satis-

fied for $x \ll 1$ if $\gamma$ is not too small. To be more accurate, we write $J_\gamma(x)$ as

$$J_\gamma(x) \approx \exp\left(\gamma \ln \frac{x}{2}\right) \frac{1}{\Gamma(1+\gamma)} \quad (x \ll 1).$$

If the condition

$$\left|\gamma \ln \frac{x}{2}\right| \gg 1$$

is satisfied, then

$$|J_\gamma(x)| \ll |J_{-\gamma}(x)|, \quad |J_\gamma(x)| \ll |N_\gamma(x)|$$

and we find (5). However, if

$$\left|\gamma \ln \frac{x}{2}\right| \ll 1,$$

an expansion in $\gamma$ is valid,

$$J_\gamma(x) = J_0 + \frac{\pi\gamma}{2} N_0, \quad N_\gamma(x) = N_0(x) - \frac{\pi\gamma}{2} J_0,$$

and we arrive at (21) or, equivalently, the Born amplitude.

Therefore, the difference between the cross sections (7) and (15) is due to the different physical situations. The cross section (7) corresponds to the conditions

$$\gamma \ll 1, \; kR \ll 1, \; \left|\gamma \ln \frac{kR}{2}\right| \gg 1. \tag{23}$$

The cross section (8) corresponds to the conditions

$$\gamma \ll 1, \; kR \ll 1, \; \left|\gamma \ln \frac{kR}{2}\right| \ll 1. \tag{24}$$

We note that the conditions (24) are much easier to satisfy in an experiment.

Let us again return to the wave function describing the scattering on an infinite screened solenoid of finite radius $R$, $\Psi = \Psi_{AB} + \Psi_R$:

$$\left.\begin{aligned} \Psi_{AB} &= \sum \exp\left[i\pi\left(|m| - \frac{1}{2}|m-\gamma|\right)\right] \\ &\quad \times J_{|m-\gamma|}(k\rho)\exp(im\varphi); \\ \Psi_R &= -\sum \exp\left[i\pi\left(|m| - \frac{1}{2}|m-\gamma|\right)\right] \\ &\quad \times H^{(1)}_{|m-\gamma|}(k\rho) \\ &\quad \times \frac{J_{|m-\gamma|}(kR)}{H^{(1)}_{|m-\gamma|}(kR)} \exp(im\varphi). \end{aligned}\right\} \tag{25}$$

Everything is clear with the asymptotic behavior of the second term: it describes a scattered wave, $\Psi_R \approx (\exp(ik\rho)/\sqrt{\rho}) f_\gamma(\varphi)$, with

$$f_\gamma(\varphi) = -\left(\frac{2}{\pi i k}\right)^{1/2} \sum \frac{J_{|m-\gamma|}(kR)}{H^{(1)}_{|m-\gamma|}(kR)} \times \exp[im\varphi + i\pi(|m| - |m-\gamma|)].$$

The asymptotic behavior of $\Psi_{AB}$ is more complicated. At large distances $\Psi_{AB}$ reduces to a distorted plane wave

$$\Psi_{in} = \exp(ikx)\exp[i\gamma(\varphi - \pi)] \tag{26}$$

and a scattered wave $\Psi_s$. For this expressions were obtained in Ref. 7, which were valid for a particular range of variation of the scattering angle $\varphi$. The expression for $\Psi_s$ found later in Ref. 11 is valid for all $\varphi$:

$$\Psi_s \approx i \sin\pi\gamma \exp\left(\frac{i\varphi}{2}\right) \frac{\exp(ik\rho)}{\left(1 - 2\pi i k\rho \sin^2\frac{\varphi}{2}\right)^{1/2}}. \tag{27}$$

For $k\rho \sin^2(\varphi/2) \gg 1$ we arrive at the expressions found in Ref. 7:

$$\Psi_s \approx \sin\pi\gamma \frac{\exp(i\varphi/2)}{\sqrt{2\pi i k\rho}} \frac{\exp(ik\rho)}{\sin\frac{\varphi}{2}}. \tag{28}$$

Then on the positive $x$ semiaxis (above and below)

$$\left.\begin{aligned} \Psi_s(\varphi = 0) &= i \sin\pi\gamma \exp(ik\rho); \\ \Psi_s(\varphi = 2\pi) &= -i \sin\pi\gamma \exp(ik\rho). \end{aligned}\right\} \tag{29}$$

Therefore, $\Psi_{in}$ and $\Psi_s$ separately have a discontinuity on the positive $x$ semiaxis. Nevertheless, their sum (i.e., $\Psi_{AB}$) is everywhere continuous. In particular,

$$\Psi_{in}(\varphi = 0) + \Psi_s(\varphi = 0) = \Psi_{in}(\varphi = 2\pi) + \Psi_s(\varphi = 2\pi) \tag{30}$$

We note that the discontinuity of $\Psi_{in}$ and $\Psi_s$ leads to the appearance of a number of paradoxical situations. We shall consider them in Sec. 3.

### 1.2. On the physical nature of the Dirac phase factor

The Schrödinger equation satisfied by the function $\Psi$,

$$-\frac{\hbar^2}{2\mu}\left[\frac{\partial^2\Psi}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial\Psi}{\partial\rho} + \frac{1}{\rho^2}\left(\frac{\partial}{\partial\varphi} - i\gamma\right)^2\Psi\right] = E\Psi, \tag{31}$$

is formally satisfied by the solution

$$\Psi = \Psi_0 \exp(i\gamma\varphi). \tag{32}$$

Here $\Psi_0$ is the solution of the Schrödinger equation in the absence of a magnetic field. If as $\Psi_0$ we take the single-valued solution (1) satisfying the condition $\Psi_0(2\pi) = \Psi_0(0)$, the unitarity of the transformation (32) causes the probability density and current to be identical for $\Psi$ and $\Psi_0$. This means that the magnetic field does not contribute to the scattering.

On the other hand, if $\Psi$ in (32) is a single-valued function, then $\Psi_0$ is non-single-valued:

$$\Psi_0(2\pi) = \Psi_0(0)\exp(-2i\pi\gamma). \tag{33}$$

We shall show that in this case [i.e., when the condition (33) is satisfied] the wave functions $\Psi$ in (2) and (32) are identical. In fact, the solution of the free Schrödinger equation with the boundary condition (33) has the form

$$\Psi_0 = \sum A_m\left[J_{|m-\gamma|}(k\rho) - H^{(1)}_{|m-\gamma|}(k\rho)\frac{J_{|m-\gamma|}(kR)}{H^{(1)}_{|m-\gamma|}(kR)}\right] \times \exp[i(m-\gamma)\varphi)]. \tag{34}$$

Substituting (34) into (32) and fixing $A_m$ from the condition that the ingoing cylindrical wave vanish for $\rho \to \infty$, we arrive at (2).

The representation of the wave function in the form (32) or the more general form

$$\Psi_D(\mathbf{r}) = \Psi_0(\mathbf{r})\exp\left(\frac{ie}{\hbar c}\int_{r_0}^{r} \mathbf{A}(\mathbf{r})\,d\mathbf{r}\right) \tag{35}$$

was first introduced by Dirac[12] in 1931. Here $\Psi_D$ is the wave function in the presence of a magnetic field, and $\Psi_0$ is the solution of the free Schrödinger equation. The exponential factor in (35) is often referred to as the magnetic phase factor. A defect of the Dirac prescription is that $\Psi_D$ is not a single-valued function (if $\Psi_0$ is a single-valued solution of the free Schrödinger equation), so that it incorrectly describes the quantum state when a magnetic field is present. However, if $\Psi$ in (31) is a single-valued wave function, Eq.

(35), although correct, is practically useless, since the problem of finding the solution of the free Schrödinger equation with a boundary condition of the type (33) is no simpler than that of solving the "unabridged" equation with single-valued wave functions. In spite of this difficulty, the Dirac prescription (35) is often used to predict various effects of the magnetic field, and these effects have been observed experimentally. The arguments generally run as follows. The function $\Psi_0$ is written as a sum

$$\Psi_0(\mathbf{r}) = \Psi_0^{(1)}(\mathbf{r}) + \Psi_0^{(2)}(\mathbf{r}).$$

Each term corresponds to waves arriving at $\mathbf{r}$ by different paths 1 and 2. The effect of the magnetic field is included by the addition of a magnetic phase factor to each term separately:

$$\Psi_D(\mathbf{r}) = \Psi_0^{(1)}(\mathbf{r}) \exp\left(\frac{ie}{\hbar c}\int\limits_{r_0,(C_1)}^{r} \mathbf{A}\,d\mathbf{r}\right) + \Psi_0^{(2)}(\mathbf{r}) \exp\left(\frac{ie}{\hbar c}\int\limits_{r_0,(C_2)}^{r} \mathbf{A}\,d\mathbf{r}\right)$$

Factoring out one exponential, we find

$$\Psi_D(\mathbf{r}) = \exp\left(\frac{ie}{\hbar c}\int\limits_{r_0}^{r} \mathbf{A}\,d\mathbf{r}\right)\left[\Psi_0^{(1)}(\mathbf{r}) + \exp\left(\frac{ie\Phi}{\hbar c}\right)\Psi_0^{(2)}(\mathbf{r})\right] \tag{36}$$

Here $\Phi$ is the magnetic flux through the closed contour constructed from $C_1$ and $C_2$. The factor $\exp(ie\Phi/\hbar c)$ leads to interference between $\Psi_0^{(1)}$ and $\Psi_0^{(2)}$. The overall factor in front of the square brackets in (36) makes $\Psi_D$ multivalued.

The physical meaning of the Dirac magnetic factor was explained in the interesting study of Ref. 13, which we shall now discuss. Let us first consider an infinitely thin cylindrical solenoid. In the absence of a magnetic field, we have the standard expansion for a plane wave:

$$\Psi_0 = \exp(ikx) = \sum_{l=-\infty}^{\infty} i^l J_{|l|}(k\rho)\exp(il\varphi). \tag{37}$$

Now we replace the summation over $l$ by the Poisson summation formula:

$$\sum_{l=-\infty}^{\infty} F(l) = \sum_{m=-\infty}^{\infty}\int\limits_{-\infty}^{\infty} d\lambda F(\lambda)\exp(2i\pi m\lambda). \tag{38}$$

Here $F(\lambda)$ is an arbitrary interpolation of $F(l)$ to noninteger values of $l$. Applying this to (37), we obtain

$$\Psi_0(\rho,\varphi) = \sum_{m=-\infty}^{\infty} T_m(\rho,\varphi);$$

$$T_m(\rho,\varphi) = \int\limits_{-\infty}^{\infty} d\lambda \exp\left(\frac{i\pi}{2}|\lambda|\right) J_{|\lambda|}(k\rho)\exp[i\lambda(\varphi+2\pi m)]. \tag{39}$$

The individual terms in (39) are not single-valued:

$$T_m(r,\varphi+2\pi) = T_{m+1}(r,\varphi). \tag{40}$$

The condition (40) ensures that the entire sum (39) is single-valued, in spite of the non-single-valuedness of its individual terms. We also note that $T_m$ contains $\varphi$ in the combination $\varphi + 2\pi m$. Therefore, $T_m$ can be interpreted as a wave taking the value $\varphi + 2\pi m$ after $m$ complete (counterclockwise) circuits of the origin. The author of Ref. 13 refers to this as the $m$th "whirling wave" corresponding to the wave function $\Psi_0$. Each such wave is a multivalued solution of the Schrödinger equation in the absence of a magnetic field. Let us now apply the Dirac prescription (35) to each whirling wave:

$$T_m^D(\rho,\varphi) = T_m(\rho,\varphi)\exp[i\gamma(\varphi+2\pi m)]. \tag{41}$$

We note that the phase of the Dirac magnetic factor is taken to be $\varphi + 2\pi m$, since this is the value that the total angle takes after $m$ circuits. Now we substitute expression (41) for $T_m$ into (39). Then,

$$\Psi(\rho,\varphi) = \sum_{m=-\infty}^{\infty} T_m^D(\rho,\varphi)$$
$$= \sum_{m=-\infty}^{\infty}\int\limits_{-\infty}^{\infty} d\lambda \exp\left(\frac{i\pi}{2}|\lambda|\right) J_{|\lambda|}(k\rho)$$
$$\times\exp[i(\lambda+\gamma)(\varphi+2\pi m)]. \tag{42}$$

We treat $\lambda + \gamma$ as a new variable and apply to (42) the inverse of (38):

$$\Psi(\rho,\varphi) = \sum_{l=-\infty}^{\infty} i^{|l-\gamma|} J_{|l-\gamma|}(k\rho)\exp(il\varphi). \tag{43}$$

The single-valued function (43) satisfies the Schrödinger equation and coincides with the AB wave function.

Let us now consider a solenoid of finite radius. In order to keep the particles from penetrating inside the solenoid, we put an infinite repulsive potential inside the solenoid. Now for $\Psi_0$ we must use not a plane wave, but the solution corresponding to scattering on an impenetrable cylinder in the absence of a magnetic field. Let the expansion of $\Psi_0$ in partial waves have the form

$$\Psi_0(\rho,\varphi) = \sum_{l=-\infty}^{\infty} R_{|l|}(\rho)\exp(il\varphi). \tag{44}$$

Here $R_{|l|}(\rho)$ is the solution of the radial equation corresponding to angular momentum $l$ and infinite repulsion inside the cylinder. We apply to (44) the Poisson summation formula

$$\Psi_0 = \sum_{m=-\infty}^{\infty} T_m(\rho,\varphi),\quad T_m(\rho,\varphi)$$
$$= \int\limits_{-\infty}^{\infty} d\lambda R_{|\lambda|}(\rho)\exp[i\lambda(\varphi+2\pi m)]. \tag{45}$$

As in (41), we use the Dirac magnetic factor to shift the phase in (45):

$$\Psi(\rho,\varphi) = \sum_{m=-\infty}^{\infty} T_m(\rho,\varphi)\exp[i\gamma(\varphi+2\pi m)]. \tag{46}$$

Altogether, after the inverse Poisson transformation we obtain

$$\Psi(\rho,\varphi) = \sum_{l=-\infty}^{\infty} R_{|l-\gamma|}(k\rho)\exp(il\varphi),$$

which is an exact solution.

It can be concluded from all this that it is possible to obtain single-valued wave functions if the Dirac magnetic factor is applied to a suitable representation of the wave functions in the absence of a magnetic field. For the AB effect such a representation is the expansion in whirling waves $T_m$. This is related to the fact that the impenetrability of the cylinder makes the space multiply connected. Paths corresponding to different numbers of circuits of the origin cannot be transformed into each other by a continuous de-

formation. Therefore, the magnetic phase shifts must be different for different numbers of circuits.

Now we can understand why the "incorrect" representation of the wave function in the form (36) was used successfully earlier for analyzing the experimental data. It can happen that in an angular range under consideration only two of the whirling waves $T_m(\rho,\varphi)$ are important. For example, they might correspond to $m=0$ and $m=-1$. Then in this range of $\varphi$ the wave function can be approximated as

$$\Psi_0(\rho,\varphi) = T_0(\rho,\varphi) + T_{-1}(\rho,\varphi). \tag{47}$$

Applying the Dirac prescription (41) to each whirling wave, we obtain the analog of the representation (36):

$$\Psi(\rho,\varphi) = \exp(i\gamma\varphi)\,[T_0(\rho,\varphi) + \exp(-2i\pi\gamma)\,T_{-1}(\rho,\varphi)]. \tag{48}$$

We have already shown that this expression is not single valued. Now the reason for this is clear. Namely, the discarded whirling waves ($m\neq 0,\ -1$), even though they are small in the angular range under consideration, owing to (40) become large upon a change by $2\pi$. Therefore, they must be included in order to guarantee that $\Psi$ is single-valued after a circuit of the cylinder.

Let us return to Eq. (32). We invert it:

$$\Psi_0 = \Psi\exp(-i\gamma\varphi). \tag{49}$$

We recall that $\Psi$ is a single-valued solution of the Schrödinger equation in the presence of a magnetic field, and $\Psi_0$ is a multivalued solution of the free Schrödinger equation with a specific boundary condition:

$$\Psi_0(2\pi) = \Psi_0(0)\exp(-2i\pi\gamma).$$

The unitarity of the transformation (49) guarantees that all observables for $\Psi$ and $\Psi_0$ coincide. Now the usual arguments can be applied to $\Psi_0$, and the AB effect can be treated as the result of the interference of waves passing around the cylinder from different sides.

Therefore, by itself the representation of the solution in the form (32) is correct when $\Psi$ is single-valued. Since the very existence of the AB effect depends on the conditions for the single-valuedness of the wave function, we shall briefly discuss the question of the single-valuedness of wave functions in quantum mechanics.

## 2. THE SINGLE-VALUEDNESS OF WAVE FUNCTIONS IN QUANTUM MECHANICS

Many of the studies dealing with these questions are indebted to the following remark of Blatt and Weisskopf[14]: "... multiple-valued wave functions cannot be excluded *a priori*. Only physically measurable quantities, such as probability densities, and expectation values of operators must be single-valued."

We shall restrict ourselves to the consideration of a spinless particle in nonrelativistic quantum mechanics in a simply connected space. The wave function $\Psi$ is a function of the particle coordinates. Then the question of the single-valuedness or multivaluedness of the wave function can be posed as follows. Let us have $\Psi(P)$ at the point $P$. Will we obtain the same value $\Psi(P)$ if we move around some closed path and again return to $P$? The angular-momentum operators satisfy the relations $[L_i, \mathbf{L}^2]=0, [L_j, L_k]=i\hbar\varepsilon_{jks}L_s$. In spherical coordinates the eigenfunctions of $L_z$ and $\mathbf{L}^2$ satisfy the differential equations

$$L_zY_l^m = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}Y_l^m = m\hbar Y_l^m; \tag{50}$$

$$\mathbf{L}^2Y_l^m = \hbar^2 l(l+1)Y_l^m. \tag{51}$$

From the commutation relations it follows that the eigenvalues of $L_z$ lie in the ranges $-l\leqslant m\leqslant l$, where $l$ can be an integer or half-integer. For integer $l$ the regular solutions are the ordinary spherical harmonics $P_l^m(\cos\theta)\exp(im\varphi)$, which are single-valued functions of the particle position. Typical examples of wave functions with half-integer indices are

$$Y_{1/2}^{1/2} \sim (\sin\theta)^{1/2}\exp(i\varphi/2)\quad (l=1/2,\ m=1/2);$$
$$Y_{3/2}^{1/2} \sim (\sin\theta)^{1/2}\cos\theta\exp(i\varphi/2)\quad (l=3/2,\ m=1/2).$$

These functions are everywhere finite, but they do not transform into each other under the action of the angular-momentum operator. For example, application of the lowering operator $L_- = L_x - iL$ to $Y_{1/2}^{1/2}$ does not give the expected $Y_{1/2}^{-1/2}$. In fact,

$$L_-Y_{1/2}^{1/2} = h\exp(-j\varphi)\left(-\frac{\partial}{\partial\theta} + i\operatorname{ctg}\theta\frac{\partial}{\partial\varphi}\right)Y_{1/2}^{1/2}$$
$$\sim \frac{\cos\theta}{\sqrt{\sin\theta}}\exp(-i\varphi/2),$$

i.e., we obtain a function which, even though it has a singularity at $\theta=0,\pi$, is still square-integrable. It also turns out that $\mathbf{L}^2Y_{1/2}^{1/2}\neq 0$. Moreover, this function is no longer square-integrable. In addition, $\langle\frac{3}{2},\ -\frac{1}{2}|L_x|\frac{1}{2}\rangle\ \frac{1}{2}\neq 0$, although from the commutation relation $\mathbf{L}^2L_x - L_x\mathbf{L}^2 = 0$ it follows that any matrix element $\langle lm|L_x|l'm'\rangle$ vanishes for $l\neq l'$. The reason for this is that $\langle\mathbf{L}^2Y_{3/2}^{-1/2}, L_xY_{1/2}^{1/2}\rangle \neq \langle Y_{3/2}^{-1/2}\mathbf{L}^2L_xY_{1/2}^{1/2}\rangle$, i.e., the operator $\mathbf{L}^2$ is not Hermitian on functions generated from two-valued functions by successive application of $L_-$ (or $L_+$).

Pauli[15] realized that the "ladder" algorithm (i.e., the generation of wave functions by the successive application of lowering or raising operators) leads to singular functions which no longer lie in the region where $\mathbf{L}$ is defined [for example, $\mathbf{L}^2$ ceases to be self-adjoint (or Hermitian)]. Because of this, Pauli suggested that the condition for the single-valuedness of the wave functions should be replaced by the following: admissible wave functions are those on which the successive application of the operators generating the eigenvalues characteristic of a given problem realizes a representation of the corresponding transformation group (for example, the rotation group). This is the celebrated Pauli criterion.

Let us now turn to a discussion of the single-valuedness of wave functions in a multiply connected space.[16] We shall consider the AB Hamiltonian describing a particle of charge $e$ and mass $\mu$ in the field of a magnetic vector potential $\mathbf{A}$ (we assume that curl $\mathbf{A}=0$ everywhere, except on the $z$ axis):

$$H = (-i\nabla - e\mathbf{A})^2/2\mu,\quad A_r = A_\theta = 0,\quad A_\varphi = \frac{\Phi}{2\pi r\sin\theta},$$

where $\oint A_l dl = \Phi$ for any contour enclosing the $z$ axis.

It is easily verified that the differential operators

$$\mathbf{J} = \mathbf{r} \times (-i\nabla - e\mathbf{A})$$

$$\left(J_z = -\mathrm{i}\frac{\partial}{\partial\varphi} - \frac{e\Phi}{2\pi},\ J_\pm = J_x \pm \mathrm{i}J_y\right.$$

$$\left. = \exp(\pm\mathrm{i}\varphi)\left(\pm\frac{\partial}{\partial\theta} + \mathrm{i}\operatorname{ctg}\theta\frac{\partial}{\partial\varphi} + \frac{e\Phi}{2\pi}\operatorname{ctg}\theta\right)\right)$$

satisfy the relations

$$[\mathbf{J},\ H] = 0,\ [J_k,\ J_m] = i\epsilon_{kms}J_s \quad (\theta \neq 0,\ \pi). \tag{52}$$

Using $\mathbf{J}$, the Hamiltonian $H$ can be written in the form

$$H = \frac{1}{2\mu}\left(P_r^2 + \frac{1}{r^2}\mathbf{J}^2\right), \tag{53}$$

where $P_r\Psi = -(i/r)(\partial/\partial r)(r\Psi)$ for an arbitrary function $\Psi$. The only difference from the free case ($\mathbf{A} = 0$) is in the definition of $\mathbf{J}$. We must therefore consider the question of the self-adjointness of the operator

$$\mathbf{J}^2 = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\left(\frac{\partial}{\partial\varphi} - \frac{ie\Phi}{2\pi}\right)^2.$$

The self-adjointness condition is written as

$$\langle\Psi',\ \bar{\mathbf{J}}^2\Psi\rangle - \langle\bar{\mathbf{J}}^2\Psi',\ \Psi\rangle = 0. \tag{54}$$

Let us see what functions $\Psi$, $\Psi'$ satisfy this condition. The expanded form of (54) is

$$\int_0^{2\pi} d\varphi\left[(1-x^2)\left(\Psi\frac{\partial\bar{\Psi}'}{\partial x} - \bar{\Psi}'\frac{\partial\Psi}{\partial x}\right)\right]_{-1}^{1}$$

$$+\int_{-1}^{1}\frac{dx}{1-x^2}\left\{\Psi\exp\left(-\frac{ie\Phi\varphi}{2\pi}\right)\frac{\partial}{\partial\varphi}\left[\bar{\Psi}'\exp\left(\frac{ie\Phi\varphi}{2\pi}\right)\right]\right.$$

$$\left. - \bar{\Psi}'\exp\left(\frac{ie\Phi\varphi}{2\pi}\right)\frac{\partial}{\partial\varphi}\left[\Psi\exp\left(-\frac{ie\Phi\varphi}{2\pi}\right)\right]\right\}_0^{2\pi}. \tag{55}$$

The conditions for this expression to vanish have the form

$$\left.\begin{aligned}&\Psi(\theta,\ \varphi = 2\pi) = \exp(2i\pi\alpha)\,\Psi(\theta,\ \varphi = 0);\\ &\frac{\partial\Psi(\theta,\ \varphi)}{\partial\varphi}\Big|_{\varphi=2\pi} = \exp(2i\pi\alpha)\frac{\partial\Psi(\theta,\ \varphi)}{\partial\varphi}\Big|_{\varphi=0},\end{aligned}\right\} \tag{56}$$

where $\alpha$ is an arbitrary real parameter. Moreover, $\Psi$ must be finite for $\cos\theta = \pm 1$. For noninteger $\alpha$ and $0 < \varphi < 2\pi$ the conditions (56) for $\Psi$ to be multivalued cannot be physically excluded, since both the charge density and the current density remain single-valued.

We seek eigenfunctions of $\mathbf{J}^2$ satisfying (56) in the form

$$\left.\begin{aligned}&Y(\theta,\ \varphi) = \exp[\mathrm{i}(m+\alpha)\varphi]\,P(\theta);\\ &\mathbf{J}^2Y = l(l+1)\,Y,\quad m = 0, \pm 1, \pm 2\ldots\end{aligned}\right. \tag{57}$$

where $l$ is still arbitrary. For $P(\theta)$ we find the following differential equation:

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dP}{d\theta}\right) + \left[l(l+1) - \frac{m'^2}{\sin^2\theta}\right]P = 0. \tag{58}$$

Here $m' = m + \alpha - e\Phi/2\pi$.

It is easily verified that the eigenfunctions $Y(\theta,\varphi)$ realize a representation of the rotation group with generators $J_i$ only when $\alpha - e\Phi/2\pi$ is an integer. Here $m'$ is also an integer, and the only functions satisfying the self-adjointness condition for $\mathbf{J}^2$ have the form

$$Y(\theta,\ \varphi) \sim \exp\left(\frac{ie\Phi\varphi}{2\pi}\right)Y_{lm'}(\theta,\ \varphi),$$

$$l = |m'| + n,\ n = 0,\ 1,\ 2\ldots \tag{59}$$

where $Y_{lm'}$ are the ordinary spherical harmonics. These functions lead to the same physical consequences as in the free case. Therefore, acceptance of the Pauli criterion leads to the same dynamics as for a free particle, i.e., to the absence of the AB effect.

For noninteger $\alpha - e\Phi/2\pi$ the general solution of (58) can be written as the sum of two terms:

$$P(\theta) = (\sin\theta)^{m'}\left[aF\left(m'+l+1,\ m'-l;\ m'+1;\ \frac{1-\cos\theta}{2}\right)\right.$$

$$\left. + b(1-\cos\theta)^{-m'}F\left(l+1,\ -l;\ 1-m';\ \frac{1-\cos\theta}{2}\right)\right],$$

where $a$ and $b$ are arbitrary constants and $F$ is the hypergeometric function. From the condition of finiteness at $\cos\theta = 1$ we obtain $b = 0$ for $m' > 0$ and $a = 0$ for $m' < 0$. Let $m' > 0$; then for $\cos\theta \to -1$ we find

$$P(\theta) \to a\frac{2^{m'}\Gamma(m'+1)\,\Gamma(m')}{\Gamma(m'+l+1)\,\Gamma(m'-l)}\left(\frac{1+\cos\theta}{2}\right)^{-m'/2}.$$

This expression satisfies the requirement of finiteness at $\theta = \pi$ for $l = |m'| + n$. The same conclusion follows from the case $m' < 0$. Altogether, we obtain the following eigenfunctions:

$$Y(\theta,\ \varphi) = \text{const}\exp\left[\mathrm{i}\left(m' + \frac{e\Phi}{2\pi}\right)\varphi\right]$$

$$\times(\sin\theta)^{|m'|}P_{l-|m'|}^{|m'|,\,|m|}(\cos\theta), \tag{60}$$

where $P_n^{\alpha,\beta}(x)$ is a Jacobi polynomial. Since these polynomials are finite for $x = \pm 1$, the eigenfunctions (60) satisfy the condition of self-adjointness of $\mathbf{J}^2$. Furthermore, for noninteger $\alpha - e\Phi/2\pi$, $|m'| \neq 0$, these eigenfunctions vanish on the $z$ axis. They are therefore also applicable when the $z$ axis is inaccessible to the incident particles. However, the eigenfunctions (60) do not realize a representation of the rotation group for noninteger $m'$ and $l$. It is easily verified that operation on $Y(\theta,\rho)$ with $J_\pm$ leads to singular functions, i.e., functions different from the AB functions.

Therefore, the Pauli criterion requires that $\alpha - e\Phi/2\pi$ be an integer and leads to the free-particle dynamics. The criterion for the wave function to be single-valued requires integer $\alpha$ and leads to the usual description of the AB effect. Since self-adjointness occurs for any real $\alpha$, it can be expected that these $\alpha$ will correspond to physical situations different from the two just mentioned.

Let us consider the time evolution of an eigenfunction of $H$:

$$\Psi = \sum\exp[\mathrm{i}(m+\alpha)\varphi]\,\Psi_m(r,\ \theta,\ t),$$

where

$$\Psi_m(r,\ \theta,\ t)$$

$$= \int_0^\infty dk\exp\left(-\frac{\mathrm{i}k^2t}{2\mu}\right)$$

$$\times\sum_{n>0}a_{mn}P_{m+\alpha-\frac{e\Phi}{2\pi}}(\theta)\frac{1}{\sqrt{r}}J_{\left|m+\alpha-\frac{e\Phi}{2\pi}\right|+n+\frac{1}{2}}(kr).$$

Here $J_\beta(x)$ is a Bessel function;

$$P_{m+\alpha-\frac{e\Phi}{2\pi}}(\theta) = (\sin\theta)^{\left|m+\alpha-\frac{e\Phi}{2\pi}\right|}C_n^{\left|m+\alpha-\frac{e\Phi}{2\pi}\right|+\frac{1}{2}}(\cos\theta);$$

$a_{mn}(k)$ are arbitrary constants; $C_n^\nu(x)$ are Gegenbauer polynomials proportional to $P_n^{\nu-1/2,\nu-1/2}(x)$; and

$$C_n^\nu(x) = \frac{\Gamma(h+2\nu)\,\Gamma\left(\nu+\frac{1}{2}\right)}{\Gamma(2\nu)\,\Gamma\left(n+\nu+\frac{1}{2}\right)}P_n^{\nu-\frac{1}{2},\,\nu-\frac{1}{2}}(x).$$

We introduce the velocity operator

$$\mathbf{v}=\frac{\mathbf{j}}{\rho}=\frac{1}{\mu}\frac{1}{|\Psi|^2}\operatorname{Re}[\overline{\Psi}(-i\nabla-l\mathbf{A})\Psi].$$

The azimuthal component of $\mathbf{v}$ is

$$v_\varphi=\frac{1}{\mu r\sin\theta}\left[\alpha-\frac{e\Phi}{2\pi}+\operatorname{Re}\frac{\sum m\exp(im\varphi)\Psi_m(r,t)}{\sum\exp(im\varphi)\Psi_m(r,t)}\right].$$

It is easily seen that $v_\varphi$ depends only on $\alpha-e\Phi/2\pi$, but not on $\alpha$ and $\Phi$ separately. This is also true of the other components of $\mathbf{v}$ and also $|\Psi|^2$. Furthermore, the authors of Ref. 16 note that, although the wave function $\Psi$ depends both on $\alpha$ and $\Phi$, all physical quantities depend only on the difference $\alpha-e\Phi/2\pi$. The initial values ($t=0$) of $|\Psi|^2$ and $\mathbf{v}$ are fully determined if $a_{mn}(k)$ and $\alpha-e\Phi/2\pi$ are specified. The values of $|\Psi|^2$ and $\mathbf{v}$ at later times are determined uniquely and are independent of $\Phi$. Alternatively, let us take the initial conditions to be fixed. Then $a_{mn}(k)$ and $\alpha-e\Phi/2\pi$ are also fixed. Now for the given initial conditions (i.e., for given $\rho$ and $\mathbf{v}$) we shall study the cases corresponding to different fluxes of the magnetic field $\Phi$. They will correspond to different $\alpha$, but the same difference $\alpha-e\Phi/2\pi$, and, consequently, to the same $\mathbf{v}$ and $|\Psi|^2$. In this sense the magnetic flux along the $z$ axis does not lead to observable physical consequences.

These arguments become clearer if we use cylindrical coordinates $r$ and $\varphi$ (Ref. 17). At the time $t$ we have

$$\Psi(r,\varphi,t)=\Sigma\exp[i(m+\alpha)\varphi]\Psi_m(r,t), \tag{61}$$

where $\Psi_m(r,t)$ is a wave packet constructed from solutions of the Schrödinger equation:

$$\Psi_m(r,t)=\int_0^\infty dk\exp\left(-\frac{ik^2t}{2\mu}\right)C_m(k)J_{|m+\alpha-\gamma|}(kr). \tag{62}$$

For $t=0$, $\Psi_m(r,0)=\int_0^\infty dkC_m(k)J_{|m+\alpha-\gamma|}(kr)$. From this we find

$$C_m(k)=k\int_0^\infty r\,drJ_{|m+\alpha-\gamma|}(kr)\Psi_m(r,0).$$

We substitute $C_m$ into (62):

$$\Psi_m(r,t)=\int_0^\infty r'\,dr'\Psi_m(r',0)g_m(r,r',t), \tag{63}$$

where

$$g_m(r,r',t)=\int_0^\infty k\,dk\exp\left(-\frac{ik^2t}{2\mu}\right)J_{|m+\alpha-\gamma|}(kr)J_{|m+\alpha+\gamma|}(kr').$$

Now we write $\rho=|\Psi|^2$ and $v_\varphi$:

$$\rho(r,\varphi,t)=\sum_{m,m'}\exp[i(m-m')\varphi]\Psi_m(r,t)\overline{\Psi}_{m'}(r,t);$$

$$v_\varphi=\frac{\hbar}{2\mu r}\left[\alpha-\gamma+\operatorname{Re}\frac{\sum m\exp(im\varphi)\Psi_m(r,t)}{\sum\exp(im\varphi)\Psi_m(r,t)}\right]. \tag{64}$$

Let us mention two types of initial condition. Let the wave function be given at the initial time $t=0$. Then at later times the wave function, and also $\rho=|\Psi|^2$ and $\mathbf{j}$, depend explicitly on the magnetic flux [see (61)], i.e., we have the AB effect. Now let $\rho$ and $\mathbf{j}$ be given at the initial time ($\rho=\rho_0$, $\mathbf{j}=\mathbf{j}_0$ at $t=t_0$). From (64) it follows that $\rho$ and $\mathbf{j}$ depend only on the difference $\alpha-\gamma$. Therefore, for fixed $\rho_0$ and $\mathbf{j}_0$ the values of $\rho$ and $\mathbf{j}$ at later times are independent of $\gamma$, i.e., of the magnetic flux. In fact, the values $\gamma'=\gamma$ correspond to the same initial conditions only when the following condition is satisfied: $\alpha'-\gamma'=\alpha-\gamma$. Owing to this equality, $\rho$ and $\mathbf{j}$ for different $\gamma$ and fixed $\rho_0$ and $\mathbf{j}_0$ are the same.

## 3. ATTEMPTS TO PROVE THE NONEXISTENCE OF THE AB EFFECT

We shall illustrate a typical "proof" for the example of a plane rotator in a magnetic field:

$$H=\frac{1}{2\mu a^2}P_\varphi^2,\quad P_\varphi=\frac{\hbar}{i}\frac{\partial}{\partial\varphi}+\frac{\partial F}{\partial\varphi}. \tag{65}$$

A single-valued eigenfunction of $H$ has the form

$$\Psi_m=\exp(im\varphi)\exp\left[i\varphi\frac{F(2\pi)-F(0)}{h}-iF(\varphi)/\hbar\right]. \tag{66}$$

It corresponds to the eigenvalue

$$E_m=\frac{\hbar^2}{2\mu a^2}\left[m+\frac{F(2\pi)-F(0)}{h}\right]^2. \tag{67}$$

In the absence of a magnetic field we would have

$$H_0=\frac{1}{2\mu a^2}(P_\varphi^0)^2,\quad P_\varphi^0=\frac{\hbar}{i}\frac{\partial}{\partial\varphi},\quad \Psi_m^0=\exp(im\varphi),\quad E_m^0=\frac{\hbar^2m^2}{2\mu a^2}. \tag{68}$$

Therefore, $E_m\neq E_m^0$ for $F(0)\neq F(2\pi)$. On the other hand, the transformation from (65) to (68) is a simple change of the angular-momentum representation:

$$P_\varphi^0=SP_\varphi S^{-1},\quad S=\exp\left[\frac{i}{\hbar}F(\varphi)\right].$$

A transformation from one representation to another cannot lead to experimentally observable consequences. Since the magnetic field ($=\partial F/\partial\varphi$) can be eliminated by a simple change of the representation for $P_\varphi$, the AB effect was declared in Ref. 18 to be a mathematical function.

It is easy to see where the error lies in these arguments. Namely, it is impossible to consider operator transformations separately from the transformation of the basis in which these operators act. Let us write down the operators and wave functions before and after changing the representation:

$$\Psi_m=\exp(im\varphi)\exp\left[i\varphi\frac{F(2\pi)-F(0)}{h}-i\frac{F(\varphi)}{\hbar}\right],\qquad P_\varphi=\frac{\hbar}{i}\frac{\partial}{\partial\varphi}+F(\varphi), \tag{69}$$

$$\tilde{\Psi}_m^0=S\Psi_m=\exp(im\varphi)\exp\left[i\varphi\frac{F(2\pi)-F(0)}{h}\right],\qquad P_\varphi^0=\frac{\hbar}{i}\frac{\partial}{\partial\varphi}. \tag{70}$$

Therefore, $P_\varphi$ and $P_\varphi^0$ are equivalent only if they act in the space of functions $\tilde{\Psi}_m$ and $\Psi_0$, respectively. In this case $\Psi_m$ and $\tilde{\Psi}_m^0$ correspond to the same angular-momentum eigenvalue

$$\hbar m+F(2\pi)-F(0).$$

On the other hand, the representation

$$P_\varphi^0,\ \Psi_m^0=\exp(im\varphi) \tag{71}$$

is inequivalent to (69) and (70), since, although the relation $P_\varphi^0=SP_\varphi S^{-1}$ holds, $\Psi_m^0\neq S\Psi_m$. We considered a similar example in Sec. 1, where for the case of the scattering of charged particles by the magnetic field of a cylindrical solenoid we showed that it is fully possible for the magnetic field to be expelled from the region of space accessible to the incident particles. Here the solutions of the free Schrödinger equation must satisfy multivalued boundary conditions.

Misunderstandings also arise owing to incorrect identification of the angular-momentum operator.[19] Let us consider a particle of charge $e$ which moves around the $z$ axis, but does not penetrate inside a cylindrical solenoid with its axis coincident with the $z$ axis. When current does not pass through the solenoid the angular momentum is

$$L_z^0 = \frac{1}{i}\frac{\partial}{\partial\varphi}.$$

Its eigenvalues are integers. If now a current is gradually switched on, the charged particle experiences an electric field

$$\mathbf{E} = -\frac{2\pi}{x^2+y^2}\,\mathbf{e}_z \times \mathbf{r}\dot{\Phi}.$$

This leads to a change of the angular momentum

$$L_z' = [\mathbf{r} \times e\mathbf{E}]_z = -\frac{e\dot{\Phi}}{2\pi}.$$

The total increment of the angular momentum is

$$\Delta L_z = -e\Phi/2\pi.$$

After the current is switched on,

$$L_z = \frac{1}{i}\frac{\partial}{\partial\varphi} - \frac{e\Phi}{2\pi}.$$

The eigenvalues of $L_z$ are $m - e\Phi/2\pi$, i.e., they are not integers.

In order to find the error in these arguments, let us consider[20] the nonrelativistic Lagrangian

$$L = \frac{1}{2}\mu v^2 + \frac{e}{c}\mathbf{v}\cdot\mathbf{A}.$$

For the cylindrical solenoid under consideration

$$A_\rho = A_z = 0, \quad A_\varphi = \frac{\Phi}{2\pi\rho}$$

The canonical momentum

$$\mathbf{p} = \frac{\delta L}{\delta \mathbf{v}} = \mu\mathbf{v} + \frac{e}{c}\mathbf{A}$$

differs from the kinematical momentum $\mu\mathbf{v}$ by an amount $(e/c)\mathbf{A}$. For the angular momentum we find

$$\mathbf{L} = \mathbf{r}\times\mathbf{p} = \mu\mathbf{r}\times\mathbf{v} + \frac{e}{c}\mathbf{r}\times\mathbf{A}.$$

In quantum mechanics the operator $\mathbf{p}$, but not $\mu\mathbf{v}$, is equal to $(\hbar/i)\nabla$. Therefore the angular-momentum operator is the same, independently of whether or not a magnetic field is present, and is equal to

$$\frac{\hbar}{i}\,\mathbf{r}\times\nabla = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}.$$

Completeness of its eigenvalues and Hermiticity require that its eigenvalues be integers (in units of $\hbar$).

In Ref. 21 it was pointed out that both $\mathbf{H}$ and $\mathbf{E}$ (the electric field of the moving charge) are nonzero inside the solenoid. In crossed electric and magnetic fields the angular momentum

$$\mathbf{F} = \frac{1}{4\pi c}\int \mathbf{r}\times(\mathbf{E}\times\mathbf{H})\,dV = \frac{e}{4\pi c}\int \mathbf{r}\times(\mathbf{r}\times\mathbf{H})\,\frac{d^3\mathbf{r}}{r^3} \qquad (72)$$

arises. Altogether, we obtain

$$F_z = -\frac{e\Phi}{2\pi c}.$$

It is noteworthy that the angular momentum of the field depends only on the total flux $\Phi$ and is independent of its spatial distribution.

Let us now consider a deuteron located at the origin and an infinite solenoid which is sufficiently far away (for example, on the moon). The total angular momentum of such a system is

$$\mathbf{J} = \mathbf{L} + \mathbf{S} + \mathbf{F}, \qquad (73)$$

where $\mathbf{S}$ is the total spin of the neutron and the proton, and $\mathbf{L}$ is the mechanical orbital angular momentum in the deuteron (the deuteron center of mass is fixed at the origin). Since the flux $\Phi$ can take arbitrary values, $F_z$ can also take arbitrary values (not only integer or half-integer multiples of $\hbar$). This leads to the following paradoxical situation[21]: a) if the total angular momentum of the system is limited to integral or half-integral multiples of $\hbar$, the allowed values of the relative angular momentum, the centrifugal barrier, and the deuteron binding energy depend on the field of the solenoid located on the moon; if for $L_z = \hbar/2$ the centrifugal barrier of the deuteron is so large that no bound states exist, the deuteron located on the earth can be dissociated, leading to the appearance of a small magnetic field in the solenoid on the moon; b) if we require that the properties of the deuteron are independent of what happens to the solenoid on the moon, this can be achieved by integer eigenvalues of $L_z$ (the total angular momentum will have eigenvalues which are not necessarily integer or half-integer multiples of $\hbar$); c) the usual treatment using the Schrödinger equation with a vector potential which is finite at infinity shows that the magnetic field of the distant solenoid has no effect on the deuteron. However, this treatment does not take into account the contribution of the crossed fields to the angular momentum. This contribution is a physical effect which must be included. The paradox is resolved if the magnetic flux at infinity (the return flux) is taken into account. Let us treat the infinite solenoid as the limit of a solenoid of finite length $L$ for $L\to\infty$. The return flux outside the solenoid moves to infinity as the length of the solenoid is increased. From (72) it follows that when the return flux is included the total angular momentum of the field is equal to zero. Since the contribution of the return flux remains finite for an arbitrarily long solenoid, paradoxes arise if the return flux is ignored for the infinite solenoid. For a long but finite solenoid the usual Schrödinger equation with the vector potential includes the crossed-field contribution to the angular momentum. This is easily seen in the usual gauge $\operatorname{div}\mathbf{A} = 0$ and $\mathbf{A}(\infty) = 0$. In this gauge the angular momentum (73) is the generator of rotations, and the additional angular momentum introduced by the vector potential is exactly equal to the crossed-field contribution:

$$\mathbf{r}\times\frac{e}{c}\mathbf{A} = \mathbf{F} \quad (\operatorname{div}\mathbf{A} = 0,\ \mathbf{A}(\infty) = 0). \qquad (74)$$

Expression (74) is valid for all solutions of Maxwell's equations which fall off sufficiently rapidly at infinity. Therefore, this formalism implicitly includes the contribution to the angular momentum from distant return fluxes. We arrive at the following conclusions: a) the angular momentum of the field is physically justified and must be included in the total angular momentum; b) the contribution of the return fluxes to the angular momentum is finite for a solenoid of arbitrary length and therefore cannot be discarded for the infinite solenoid; c) the usual formalism (the Schrödinger equation with the vector potential) correctly includes the crossed-

field contribution to the angular momentum if the field falls off sufficiently rapidly at infinity.

In summary: the ordinary Schrödinger equation with the vector potential correctly (although implicitly) includes both the angular momentum of the field and the contribution of the return fluxes.

Let us now discuss the very interesting study of Ref. 22, which also puts the existence of the AB effect in doubt. Let $z_\mu(x, \xi)$ ($\mu = 1,...,4$; $z_4 = iz_0$) be single-valued differentiable functions of the coordinates $x_\mu$ and the real parameter $\xi$ ($-\infty < \xi < \infty$), satisfying the conditions $z_\mu(x, 0) = \xi_\mu$, $\lim_{\xi \to -\infty} |z_\mu(x, \xi)| \to \infty$.

Typical examples of $z_\mu$ satisfying these conditions are: a) $z_0 = x_0$, $\mathbf{z} = \mathbf{x} + \xi\mathbf{n}$ (where $\mathbf{n}^2 = 1$ and $\mathbf{n}$ is independent of $x_\mu$); b) $z_0 = x_0$, $\mathbf{z} = \mathbf{x}(1 - \xi)$.

Let $A_\mu(x)$ be single-valued functions obeying the conditions

$$\int_{-\infty}^{0} A_\mu(z) \frac{\partial z_\mu}{\partial \xi} d\xi < \infty, \quad \lim_{(\xi \to -\infty)} A_\nu(z) \frac{\partial z_\nu}{\partial x_\mu} = 0,$$

$$\int_{-\infty}^{0} d\xi F_{\rho\nu}(\xi) \frac{\partial z_\rho}{\partial \xi} \frac{\partial z_\nu}{\partial x_\mu} < \infty, \quad \left(F_{\nu\rho} = \frac{\partial A_\nu}{\partial x_\rho} - \frac{\partial A_\rho}{\partial x_\nu}\right), \tag{75}$$

where the integrand containing $F_{\rho\nu}$ is a continuous function of $\xi$ and $x_\mu$. Then

$$\frac{\partial}{\partial x_\mu} \int_{-\infty}^{\infty} d\xi A_\nu(\xi) \frac{\partial z_\nu}{\partial \xi} = A_\mu(x) + \int_{-\infty}^{0} d\xi F_{\rho\nu}(z) \frac{\partial z_\nu}{\partial \xi} \frac{\partial z_\rho}{\partial x_\mu}. \tag{76}$$

This relation is obtained by differentiation inside the integral and subsequent integration by parts. The condition (75) is necessary for these operations to be valid.

Let a charged particle have access to a region $R$ (which can be simply or multiply connected), where it interacts with the single-valued potential $A_\mu$. Then, if there exists a unique and differentiable path $z_\mu(x, \xi)$ lying in $R$ for any $x$ such that the conditions (75) and (76) are satisfied, the physical effects acting on the particle are determined exclusively by the field strength in $R$. This follows from (76) if we use the fact that the integral on the left-hand side of (76) is a single-valued function. Therefore, owing to gauge invariance, $A_\mu(x)$ is equivalent to the potential

$$A'_\mu = \int_{-\infty}^{0} d\xi F_{\nu\rho} \frac{\partial z_\nu}{\partial \xi} \frac{\partial z_\rho}{\partial x_\mu},$$

which depends only on the field strength in $R$.

The potential of an infinite cylindrical solenoid ($A_\rho = A_z = 0, A_\psi = \Phi/2\pi\rho$) does not satisfy the conditions (75). Let us now take a solenoid of finite length $2L$. Then $A_0 = A_\rho = A_z = 0$,

$$A_\varphi = \frac{\Phi}{4\pi\rho} \left\{ \frac{\operatorname{sgn}(L-z)}{\left[1 + \left(\frac{\rho}{L-z}\right)^2\right]^{1/2}} + \frac{\operatorname{sgn}(L+z)}{\left[1 + \left(\frac{\rho}{L+z}\right)^2\right]^{1/2}} \right\}. \tag{77}$$

Owing to the rapid ($\sim r^{-2}$) falloff for $r \to \infty$, this potential satisfies the conditions (75). Therefore, for a finite solenoid of arbitrary length the potential is determined by the strength of the electromagnetic field in $R$, i.e., by the leakage of the magnetic field outside the solenoid. In other words, for any real cylindrical solenoid the positive results of experiments seeking the AB effect can be attributed to leakage of the magnetic field. In a typical experiment on the detection of the AB effect, the solenoid length is less than 2 cm and the radius is several micrometers. The distance from the source to the screen is about 50 cm. We conclude that all the experiments with cylindrical solenoids have been carried out in a real simply connected space. The space is simply connected, owing to the finite dimensions of the solenoid and to the finite height of the potential barrier preventing the penetration of charged particles inside the solenoid.

In spite of the fact that there were no doubts about the mathematical aspects of Ref. 22, attempts were made[23] to find a different interpretation of its results. It was noted that for a finite solenoid the leakage of the magnetic field is a maximum near the ends ($z = \pm L$) and decreases rapidly towards the center ($z = 0$). If the charged-particle beam is focused near the plane $z = 0$, for these particles the conditions are the same as for an ideal infinite solenoid.

Let us now turn to the description of the AB effect in the hydrodynamical interpretation of quantum mechanics. We write the wave function as

$$\Psi = \exp(R + iS).$$

We introduce the probability density $\rho$ and the current $\mathbf{j}$:

$$\rho = |\Psi|^2 = \exp(2R),$$
$$\mathbf{j} = \rho\mathbf{v} = \frac{i\hbar}{2m}(\psi\nabla\bar{\Psi} - \bar{\psi}\nabla\Psi) - \frac{e}{mc}\rho\mathbf{A}, \tag{78}$$

where $\rho$ and $\mathbf{v}$ satisfy the following system of equations:

$$\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho\mathbf{v}) = 0, \quad m\frac{d\mathbf{v}}{dt} = -\operatorname{grad} Q + \mathbf{F}. \tag{79}$$

Here $Q = -(\hbar^2/2m)(\Delta\sqrt{\rho}/\sqrt{\rho})$, $\mathbf{F}$ is the Lorentz force [$= e\mathbf{E} + (e/c)(\mathbf{v}\times\mathbf{H})$], and

$$\frac{d\mathbf{v}}{dt} = \frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\nabla)\mathbf{v}.$$

Outside the solenoid $\mathbf{E} = \mathbf{H} = 0$, so that

$$m\frac{d\mathbf{v}}{dt} = -\operatorname{grad} Q.$$

Alternatively, integrating around any closed contour outside the solenoid,

$$\frac{d}{dt}\int(\mathbf{v}\,d\mathbf{s}) = 0.$$

This implies that the integral of the velocity field along any closed contour is an integral of the motion. If this contour encloses the solenoid, we have

$$\int \mathbf{v}\,d\mathbf{s} = \frac{i\hbar}{2m}\int \frac{\partial}{\partial\varphi}\ln\frac{\bar{\Psi}}{\Psi}d\varphi - \frac{e\Phi}{mc} = \frac{i\hbar}{2m}\ln\frac{\bar{\Psi}}{\Psi}\Big|_0^{2\pi} - \frac{e\Phi}{mc}. \tag{80}$$

For $t \to -\infty$ the asymptote of the wave function coincides with the distorted plane wave (26):

$$\Psi(t = -\infty) = \Psi_{\text{inc}} = \exp[i(kx + \gamma(\varphi - \pi)], \quad e = \frac{e\Phi}{hc}. \tag{81}$$

Substituting (81) into (80), we find

$$\int \mathbf{v}\,d\mathbf{s} = 0. \tag{82}$$

For $t \to \infty$ the asymptote of the wave function has the form

$$\Psi(t = \infty) = \Psi_{\text{inc}} + \frac{1}{\sqrt{\rho}}\exp(ik\rho)f(\varphi). \tag{83}$$

Above [see (30)] we saw that this function is single-valued. Therefore,

$$\int \mathbf{v}\, ds = -\frac{e\Phi}{mc}. \tag{84}$$

Expressions (82) and (83) coincide if $f(\varphi) = 0$, i.e., if there is no scattering on the magnetic field surrounding the solenoid. From this the authors of Ref. 24 concluded that the AB effect is absent. The error in these arguments lies in the identification of the multivalued wave function (26) with the incident wave. The true wave function must be single-valued [see, for example, Ref. 25 or the articles by Yang (pp. 5–9) and Aharonov (pp. 10–14) in Ref. 26].

Let us now consider the paradox associated with angular-momentum nonconservation in the field of an infinite cylindrical solenoid.[24] According to the Ehrenfest theorem,

$$\frac{d\mathbf{L}}{dt} = \langle \Psi \mid \mathbf{r} \times \mathbf{F} \mid \Psi \rangle.$$

Here $\mathbf{F}$ is the Lorentz force [in this case $= (e/c)(\mathbf{v}\times\mathbf{H})$], $m\mathbf{v} = \mathbf{p} - (e/c)\mathbf{A}$, and $\mathbf{L} = \langle \Psi | \mathbf{r} \times m\mathbf{v} | \Psi \rangle$. Since outside the solenoid $F = 0$, $\mathbf{L}$ is independent of time. Constructing a wave packet at $t = -\infty$ from the functions (81) and one at $t = \infty$ from the functions (82), we verify that in the first case $L_z = \hbar\gamma$, while in the second $L_z = 0$. Angular momentum is not conserved. The reason for this is that a multivalued wave function was used at $t = -\infty$. These paradoxes are studied in detail in Ref. 27.

#### 4. THE TIME-DEPENDENT AB FORMALISM

Here we briefly describe the contents of Refs. 16 and 28. We consider a plane rotator, a charged particle moving around the $z$ axis in a circle of radius $a$. The magnetic potential of the infinite solenoid is time-dependent:

$$A_r = A_\theta = 0, \quad A_\varphi = \frac{\Phi(t)}{2\pi a}.$$

As before, the magnetic flux is located in the region inaccessible to the charged particle. The electric field at the location of the charge is nonzero:

$$E_r = E_\theta = 0, \quad E_\varphi = -\Phi'(t)/2\pi.$$

The time-dependent AB Hamiltonian is

$$H = \frac{1}{2\mu a^2}\left(\frac{1}{i}\frac{\partial}{\partial\varphi} - \frac{e\Phi}{2\pi}\right)^2.$$

We seek a solution of the time-dependent Schrödinger equation in the form

$$\Psi(\varphi, t) = \sum \exp[i(m+\alpha)\varphi]\, \Psi_m(t).$$

Here $\alpha$ is an arbitrary real, time-independent parameter (otherwise, differentiation of $\Psi$ with respect to $t$ would lead to a term proportional to $\varphi$). The equation for $\Psi_m(t)$ can be solved exactly:

$$\Psi_m(t) = \exp\left[-i\int_0^t \frac{dt'}{\mu a^2}\left(m+\alpha-\frac{e\Phi(t)}{2\pi}\right)^2\right].$$

The "hydrodynamical" velocity ($\mathbf{v} = \mathbf{j}/\rho$, $\rho = |\Psi^2|$) is

$$v_\varphi = \frac{1}{\mu a}\left[\alpha - \frac{e\Phi(t)}{2\pi} + \mathrm{Re}\,\frac{\sum m\Psi_m(t)\exp(im\varphi)}{\sum \Psi_m(t)\exp(im\varphi)}\right].$$

What can be found from these expressions? The criterion for $\alpha - e\Phi/2\pi$ to be an integer is no longer applicable. In fact, if at some time $t_1$, $(\alpha - e\Phi(t_1))/2\pi$ is an integer, it will not be one at a different time $t_2$. Let us now consider a special form of the dependence $\Phi(t)$:

$$\Phi(t) = \Phi_0\theta(t),$$

$\Phi_0 = \text{const}$; $\theta(t) = 0$ for $t < 0$ and 1 for $t > 0$. Then

$$\Psi_m(t) = \begin{cases} a_m \exp\left[-\dfrac{it}{2\mu a^2}\left(m+\alpha-\dfrac{e\Phi}{2\pi}\right)^2\right], & t > 0; \\ a_m \exp\left[-\dfrac{it}{2\mu a^2}(m+\alpha)^2\right], & t < 0. \end{cases}$$

The eigenvalues for $t < 0$ and $t > 0$ are different.

What we interpreted earlier as the result of the presence of an irrotational vector potential can now be understood as the results of the action of the electric field at $t = 0$.

Let us now consider the scattering of charged particles on an infinite cylindrical solenoid with a time-dependent magnetic field. Then instead of (61) we have

$$\Psi = \int_{-\infty}^{\infty} dk_z \sum \exp[i(m+\alpha)\varphi] \exp(ik_z z)\, \Psi_{m,k_z}(\rho, t),$$

where $\Psi_{m,k_z}(\rho, t)$ satisfies the differential equation

$$\left[2i\mu\frac{\partial}{\partial t} - k_z^2 + \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right) - \frac{1}{\rho^2}\left(m+\alpha-\frac{e\Phi(t)}{2\pi}\right)^2\right]\Psi_{m,k_z}(\rho t) = 0$$

From this it follows that $\psi_{m,k_z}$ and, accordingly, $\mathbf{v}$ and $|\Psi|^2$ depend only on the difference $\alpha - e\Phi(t)/2\pi$. In the special case $\Phi(t) = \Phi_0\theta(t)$ the equation for $\Psi_{mk_z}$ can be solved exactly:

$$\Psi_{m,k_z}(\rho, t) = \int_0^\infty dk \exp\left(-it\frac{k^2+k_z^2}{2\mu}\right) a_{m,k_z}(k)\, J_{\left|m+\alpha-\frac{e\Phi_0}{2\pi}\right|}(k\rho); \quad (t > 0)$$

$$\Psi_{m,k_z}(\rho, t) = \int_0^\infty dk \exp\left(-it\frac{k^2+k_z^2}{2\mu}\right) b_{m,k_z}(k)\, J_{|m+\alpha|}(k\rho), \quad (t < 0), \tag{85}$$

where $b_{m,k_z}$ and $a_{m,k_z}$ are determined from the condition that the solutions match at $t = 0$:

$$b_{m,k_z}(k) = \int_0^\infty dk' a_{mk_z}(k') \int_0^\infty d\rho k\rho J_{|m+\alpha|}(k\rho)\, J_{\left|m+\alpha-\frac{e\Phi}{2\pi}\right|}(k'\rho).$$

These expressions show that the solution (60) satisfying the Pauli criterion is inapplicable. In fact, for the solution (85) to be consistent with the Pauli criterion it is necessary that $\alpha$ be an integer for $t < 0$ and $\alpha - e\Phi/2\pi$ be an integer for $t > 0$. It also follows from (85) that the scattering amplitudes for $t > 0$ ($\Phi \neq 0$) and $t < 0$ ($\Phi = 0$) are different. Above, this difference was associated with the magnetic field outside the solenoid. Now we have the alternative of associating this change with the presence of an electric field.

We note that the entire technique developed in Sec. 2 and in the present section was for the purpose of proving the mathematical consistency of noninteger values of $\alpha$. From the viewpoint of the results obtained in Sec. 2, the following two experimental situations must be equivalent: a) the case in which the magnetic field strength is equal to zero ($\Phi = 0$); b) the case in which first a magnetic field is switched on ($\Phi \neq 0$), and then the electron beam is emitted. This is related to the fact that all physical quantities depend

on the difference $\alpha - e\Phi/2\pi$. Therefore, for identical initial conditions [$\bar{\Psi}\Psi$ and $\mathbf{j} = \rho\mathbf{v}$ coincide at $t = 0$ for cases (a) and (b)] the difference $\alpha - e\Phi/2\pi$ is the same for cases (a) and (b). If the electron beam is emitted first and then the magnetic field is changed [we shall refer to this as case (c)], $\alpha$ remains constant. The induced electric field leads to a change of the wave function and a shift of the interference pattern.

Therefore, if the constant $\alpha$ is an integer, the AB effect is observed when the magnetic field and the electron beam are switched on in any order. However, if $\alpha$ is not an integer, the AB effect exists only in case (c).

We conclude that a correct description of the AB effect is obtained only when single-valued wave functions are used. The use of multivalued wave functions, although mathematically consistent, leads to serious ambiguities in the interpretation of the experimental results.

## 5. SCATTERING OF CHARGED PARTICLES ON TWO INFINITE CYLINDRICAL SOLENOIDS

Above, we mentioned that the slow ($\sim\rho,^{-1}$) falloff of the vector potential for a single solenoid leads to a change of the asymptote of the wave function. The splitting of the total wave function into an incident and a scattered wave becomes ambiguous, which leads to the appearance of various paradoxes. Here we attempt to avoid this problem by considering the scattering of charged particles on two infinite impenetrable parallel solenoids with magnetic fluxes of equal magnitude but opposite sign.[29-31]

### 5.1. Necessary mathematical information

In what follows we shall make extensive use of bicylindrical coordinates $\mu$, $\theta$, $z$:

$$x = \frac{a\sin\theta}{\operatorname{ch}\mu - \cos\theta}, \quad y = a\frac{\operatorname{sh}\mu}{\operatorname{ch}\mu - \cos\theta}, \quad z = z$$
$$(-\infty < \mu < \infty, \ -\pi < \theta < \pi, \ -\infty < z < \infty). \tag{86}$$

For fixed $\mu$ and varying $\theta, z$ the points (86) lie on the surface of a cylinder of radius $R = a/|\sinh\mu|$. Its axis, parallel to the $z$ axis, passes through the point $x = 0$, $y = d = a\coth\mu$. Since the scattering process is the same in any plane perpendicular to the $z$ axis, it is sufficient to restrict ourselves to the plane $z = 0$. In Fig. 1 we show the cross sections of the two solenoids corresponding to the values $\mu = \pm\mu_0$. For definiteness we assign to the upper solenoid ($\mu = \mu_0$) the index 1, and to the lower one ($\mu = -\mu_0$) the index 2. Then for $\mu_0 < \mu < \infty$ the point (86) lies inside the first solenoid, while for $-\infty < \mu < -\mu_0$ it lies inside the second. For $-\mu_0 < \mu < \mu_0$ this point lies outside the solenoid. For $\mu = \pm\mu_0$ the variation of the angle $\theta$ causes the points (86) to lie along one of the two circles shown in Fig. 1. There we also show the values of the angles $\theta$ on these circles. We note that points on adjacent segments $x = \pm\varepsilon$, $-a < y < a$ ($0 < \varepsilon \ll 1$) correspond to the values $\theta = \pm\pi$, respectively. The bicylindrical coordinates are expressed in terms of the polar coordinates $\rho, \varphi$ as

$$\mu = \frac{1}{2}\ln\frac{\rho^2 + 2a\rho\sin\varphi + a^2}{\rho^2 - 2a\rho\sin\varphi + a^2}, \quad \sin\theta = \frac{2a\rho\cos\varphi}{[(\rho^2 - a^2)^2 + 4a^2\rho^2\cos^2\varphi]^{1/2}}.$$

From this it follows that for $\rho \to \infty$

$$\mu \approx \frac{2a}{\rho}\cos\varphi, \quad \theta \approx \frac{2a}{\rho}\sin\varphi.$$

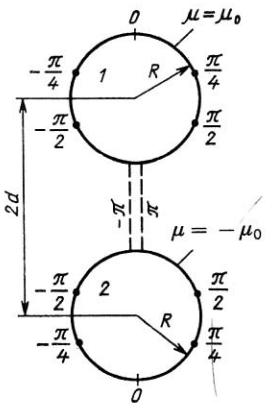


FIG. 1. The circles 1 and 2 correspond to two cylindrical solenoids with $\mu = \pm\mu_0$. The values of the angle $\theta$ along the circles are given.

Let us express the vector potential in bicylindrical coordinates. It is simplest to start from the Poisson equations:

$$(\partial_\mu^2 + \partial_\theta^2) A_x = -a\frac{1 - \operatorname{ch}\mu_0\cos\theta}{|\operatorname{ch}\mu_0 - \cos\theta|^2}$$
$$\times [H_1\delta(\mu - \mu_0) - H_2\delta(\mu + \mu_0)];$$

$$(\partial_\mu^2 + \partial_\theta^2) A_y = -\frac{\operatorname{sh}\mu_0\sin\theta}{(\operatorname{ch}\mu_0 - \cos\theta)^2}[H_1\delta(\mu - \mu_0) + H_2\delta(\mu + \mu_0)].$$

Here $H_1$ and $H_2$ are the magnetic field strengths inside the solenoids (we note that for an infinite cylindrical solenoid $H_z = \text{const}$, $H_x = H_y = 0$ inside the solenoid and $\mathbf{H} = 0$ on the outside). The solutions of these equations which are continuous and decrease for $\rho \to \infty$ are

$$A_x = a\sum_{n=1}^{\infty} R_n^x(\mu)\cos n\theta + a\frac{H_1 - H_2}{\exp(2\mu_0) - 1};$$
$$A_y = a\sum_{n=1}^{\infty} R_n^y(\mu)\sin n\theta, \tag{87}$$

where

$$R_n^x = (H_2 e^{-2\mu_0 n} - H_1)\, e^{-\mu n}, \quad R_n^y = (H_2 e^{-2\mu_0 n} + H_1)\, e^{-\mu n}$$

inside the first solenoid ($\mu > \mu_0$),

$$R_n^x = (H_2 - H_1 e^{-2\mu_0 n})\, e^{\mu n}, \quad R_n^y = (H_2 + H_1 e^{-2\mu_0 n})\, e^{\mu n}$$

inside the second ($\mu < \mu_0$), and

$$R_n^x = (H_2 e^{-\mu n} - H_1 e^{\mu n})\, e^{-2\mu_0 n},$$

$$R_n^y = (H_1 e^{\mu n} + H_2 e^{-\mu n})\, e^{-2\mu_0 n}$$

outside the solenoids ($-\mu_0 < \mu < \mu_0$).

We shall restrict ourselves to the most interesting case ($H_1 = -H_2 = H$). Then from (87) for $\rho \to \infty$ we find

$$A_x \approx \frac{\Phi d}{\pi\rho^2}\cos 2\varphi, \quad A_y \approx \frac{\Phi d}{\pi\rho^2}\sin 2\varphi;$$
$$\left(\Phi = \pi R^2 H, \ R = \frac{a}{\operatorname{sh}\mu_0}, \ d = a\operatorname{cth}\mu_0\right).$$

### 5.2. The generating function

Since outside the solenoids $\mathbf{H} = \operatorname{curl}\mathbf{A} = 0$, $\mathbf{A}$ can be written as the gradient of some function $\chi$ (Refs. 32 and 33). This function is discontinuous, since for certain closed con-

tours (more accurately, contours which enclose one of the solenoids) $\oint A_l dl \neq 0$. The function $\chi$ is

$$\chi = -\frac{\Phi}{2\pi}\left(\theta + 2\sum_{n=1}^{\infty}\frac{1}{h}\sin n\theta e^{-2\mu_0 n}\operatorname{ch}\mu n\right). \quad (88)$$

The following properties of the function $\chi$ will be important for what follows: 1) at large distances $\chi$ falls off as $\rho^{-1}$; 2) $\chi$ jumps from $-\frac{1}{2} = \Phi$ to $\frac{1}{2}\Phi$ in passing through the segment $(-d + R, d - R)$ of the $y$ axis; 3) $\chi$ is an odd function of $x$.

The generating-function formalism will play a key role in obtaining the scattering amplitude. Owing to its unfamiliarity and for illustrative purposes, we shall show how it works for the well-known example of a single pointlike cylindrical solenoid. In this case the generating function $\chi$ is $\Phi\varphi/2\pi$. The region of discontinuity of $\chi$ coincides with the positive $x$ semiaxis. For the scattering amplitude in first-order perturbation theory we have

$$f(\varphi) = \frac{e}{\hbar c}\frac{1}{\sqrt{2\pi ik}}\int e^{-ik(x'\cos\varphi + y'\sin\varphi)}\mathbf{A}\nabla\Psi_0\, dx'\, dy'. \quad (89)$$

Here $\Psi_0$ is the wave function in the absence of the magnetic field [in the present case it is $\exp(ikx)$]. Taking, as above, $\mathbf{A} = \operatorname{grad}\chi$ outside the solenoid and integrating by parts twice, we bring (89) to the form

$$f = \frac{e}{2\hbar c}\frac{1}{\sqrt{2\pi ik}}\int dx'\, dy' \operatorname{div}\left[e^{-ik(x'\cos\varphi + y'\sin\varphi)}\operatorname{grad}\chi\Psi_0 - \chi\Psi_0 \operatorname{grad} e^{-ik(x'\cos\varphi + y'\sin\varphi)}\right].$$

Owing to Gauss's theorem (taking into account the discontinuity of the function $\chi$), this integral reduces to an integral over a circle of large radius

$$I_1 = \frac{1}{2}\frac{\gamma ikR_0}{\sqrt{2\pi ik}}\int_0^{2\pi} e^{ikR_0[\cos\varphi' + \cos(\varphi - \varphi')]}\varphi'$$

$$\times\,[\cos\varphi' + \cos(\varphi - \varphi')]\, d\varphi' \quad (90)$$

and to an integral along the line where the function $\chi$ is discontinuous

$$I_2 = \frac{\pi ik\gamma}{\sqrt{2\pi ik}}\sin\varphi\int_0^{R_0} dx e^{ikx(1-\cos\varphi)}$$

$$= \frac{\pi\gamma}{\sqrt{2\pi ik}}\frac{\sin\varphi}{1-\cos\varphi}\left[e^{ikR_0(1-\cos\varphi)} - 1\right]. \quad (91)$$

Using the fact that the integrand in (90) is a rapidly oscillating function, for $R_0 \to \infty$ we obtain

$$I_1 \approx -\frac{\pi\gamma}{\sqrt{2\pi ik}}\frac{\sin\varphi}{1-\cos\varphi}e^{ikR_0(1-\cos\varphi)}. \quad (92)$$

Adding $I_1$ and $I_2$, we find

$$f_{AB} = -\frac{\pi\gamma}{\sqrt{2\pi ik}}\frac{\sin\varphi}{1-\cos\varphi},$$

i.e., the AB amplitude (14) for a pointlike solenoid. We note that in this case, since the function $\chi$ does not fall off for $\rho \to \infty$, the integral $I_1$ around a circle of infinitely large radius does not vanish.

### 5.3. The Born approximation

Let a beam of incident particles be directed along the $x$ axis (Fig. 2). In the first Born approximation we have

$$\Psi = \Psi_0 + \frac{2ie}{\hbar c}\int G_0(\mathbf{r}, \mathbf{r}')\, V_1\Psi_0(\mathbf{r}')\, dV'. \quad (93)$$

Here $\Psi_0$ and $G_0$ are the wave function and the Green's function corresponding to scattering on two impenetrable infinite cylinders coinciding with the solenoids in the absence of a magnetic field. They vanish at the surface (and also inside) of each of the cylinders, $V_1 = \mathbf{A}\nabla$. Since $\mathbf{V}_1$ and $\Psi_0$ in (93) are independent of $z$, the three-dimensional Green's function $G_0(\mathbf{r}, \mathbf{r}')$ can be integrated over $z$ and written as a two-dimensional Green's function $G_0(\boldsymbol{\rho}, \boldsymbol{\rho}')$ depending only on $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ [for example, integrating the three-dimensional plane-wave Green's function

$$-\frac{\exp(ikR)}{4\pi R}, \quad R = [(x-x')^2 + (y-y')^2 + (z-z')^2]^{1/2}$$

over $z'$, we obtain the two-dimensional Green's function $(1/4i)H_0^{(1)}(k|\boldsymbol{\rho} - \boldsymbol{\rho}'|)$]. Altogether, instead of (93) we obtain

$$\Psi = \Psi_0 + \frac{2ie}{\hbar c}\int G_0(\boldsymbol{\rho}, \boldsymbol{\rho}')\, V_1\Psi_0(\boldsymbol{\rho}')\, d^2(\boldsymbol{\rho}'), \quad (94)$$

with the integration carried out in the region outside the solenoids, where $\mathbf{A} = \operatorname{grad}\chi$. Taking this into account, $\mathbf{A}\operatorname{grad}\Psi_0 = \operatorname{grad}\chi\operatorname{grad}\Psi_0 = \frac{1}{2}\Delta\chi\Psi_0 \frac{1}{2}\chi\Delta\Psi_0 = \frac{1}{2}(\Delta + k^2)\chi\Psi_0$. Therefore,

$$\Psi = \Psi_0 + \frac{ie}{\hbar c}\int G_0(\boldsymbol{\rho}, \boldsymbol{\rho}')(\Delta + k^2)\chi\Psi_0 d^2\boldsymbol{\rho}'.$$

After double integration by parts we obtain

$$\Psi = \Psi_0 + \frac{ie}{\hbar c}\chi\Psi_0$$

$$+\frac{ie}{\hbar c}\int \operatorname{div}(G_0\operatorname{grad}\chi\Psi_0 - \chi\Psi_0\operatorname{grad}G_0)\, d^2\boldsymbol{\rho}'. \quad (95)$$

In order to obtain the scattering amplitude we must find the limit of this expression for $\rho \to \infty$. Let us consider each of the terms on the right-hand side of (95) separately. For $\rho \to \infty$ the first term reduces to an incoming plane wave and the scattering amplitude on two impenetrable cylinders in the absence of a magnetic field. The second term in (95) decreases as $\rho^{-1}$ for $\rho \to \infty$ and therefore does not contribute to the scattering amplitudes [we use the fact that for the two-dimensional case under consideration, the scattering amplitude is the coefficient of $\exp(ik\rho)/\sqrt{\rho}$. If the operator div in (95) were operating on a continuous function, by Gauss's theorem the integral in (95) could be reduced to integrals over the surface of the solenoids and the surface of a cylinder of sufficiently large radius (more accurately, over the projections of these surfaces on the $z = 0$ plane). The integral over the surface of the solenoids is equal to zero ($\Psi_0 = G_0 = 0$ on them). Owing to the asymptotic behavior of the function $\chi$, the integral along $C_0$ vanishes. Now let us use the fact that the divergence in (95) operates on a discontinuous

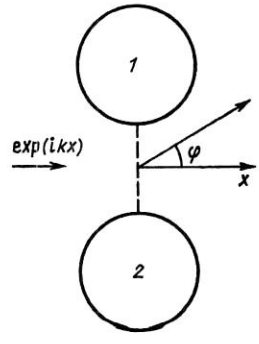


FIG. 2. An incident wave propagates perpendicular to the plane passing through the axis of the two solenoids. The scattering angle $\varphi$ is indicated. The dashed line shows the region where the generating function $\chi$ is discontinuous.

function. Detailed consideration shows that an additional integral over the region of discontinuity of the function $\chi$ arises:

$$\int_{-d+R}^{d-R} dy' G_0 \left( \frac{\partial \Psi_0}{\partial x'} - \Psi_0 \frac{\partial G_0}{\partial x'} \right) \Big|_{x'=0}, \quad \gamma = e\Phi/hc. \tag{96}$$

The wave equation is not separable in bicylindrical coordinates. Following the Kirchhoff method (see, for example, Ref. 34), we shall assume that in the region between the solenoids $\Psi_0$ and $G_0$ can be approximated by their plane-wave analogs:

$$\Psi_0 \approx \exp(ikx), \quad G_0 \approx \frac{1}{4i} H_0^{(1)} [k \sqrt{(x-x')^2 + (y-y')^2}].$$

Taking the limit $\rho \to \infty$ in (96), we find the following for the amplitude of scattering on the magnetic field of two solenoids with $\Phi_1 = -\Phi_2$ (which for brevity will be referred to as the magnetic scattering amplitude) (Ref. 31):

$$f_1(\varphi) = \gamma \sqrt{\frac{2\pi i}{k}} \frac{1+\cos\varphi}{\sin\varphi} \sin[k(d-R)\sin\varphi]. \tag{97}$$

The total scattering amplitude is equal to the sum of the amplitudes for scattering on the two impenetrable cylinders in the absence of magnetic fields $f_0$ and $f_1$. If the cylinder radii are very small, $f_0$ (for angles not too close to the zeros of $f_1$) can be neglected in comparison with $f_1$. Altogether, for the cross section for scattering on two pointlike solenoids we obtain

$$\sigma_1(\varphi) = \frac{2\pi\gamma^2}{k} \sin^2[k(d-R)\sin\varphi] \operatorname{ctg}^2 \frac{\varphi}{2}.$$

In contrast to the cross section for scattering on a single pointlike solenoid, the cross section for scattering on two pointlike solenoids is finite for any angles $\varphi$. The integrated cross section $\sigma = \int \sigma_1(\varphi) d\varphi$ can be calculated in the limiting cases of small and large wavelengths:

$$\sigma = \begin{cases} 16\pi^2 (d-R)\gamma^2, & k(d-R) \gg 1; \quad (98) \\ 6k\pi^2 (d-R)^2 \gamma^2, & k(d-R) \ll 1. \quad (99) \end{cases}$$

The limits of applicability of these expressions are the same as for the Kirchhoff method in optics: the wavelength $k^{-1}$ must be small compared with the "slit" width $d - R$. Therefore, the validity of expression (99) is doubtful, although it does vanish for $d = R$ (as it should).

For arbitrary orientation $(k_x, k_y)$ of the wave vector of the incident wave, the magnetic scattering amplitude is equal to

$$f_1 = \gamma \sqrt{\frac{2\pi i}{k}} \frac{k_x + k\cos\varphi}{k_y - k\sin\varphi} \sin[(d-R)(k_y - k\sin\varphi)]. \tag{100}$$

The approximations for which this expression was obtained should always be borne in mind. For example, let us consider the case in which the incident wave propagates along the $y$ axis. In this case (100) takes the form

$$f_1 = \gamma \sqrt{\frac{2\pi i}{k}} \frac{\cos\varphi}{1-\sin\varphi} \sin[k(d-R)(1-\sin\varphi)].$$

To obtain this expression, in the region where the function $\chi$ is discontinuous ($-d + R \leqslant y \leqslant d - R$) we approximated the exact wave function $\Psi_0$ by the incident wave $\exp(iky)$. Here the fact that in this region $\Psi_0 \simeq 0$, owing to the presence of a shadow, is not taken into account. In summary, expression (100) is valid for incident waves with wave vector close to the $x$ axis.

### 5.4. The high-energy approximation

Let us now calculate the scattering amplitude in the high-energy approximation. For this, in the Lippmann–Schwinger equation

$$\Psi = \Psi_0 + \int G_0 \left( \frac{2ie}{\hbar c} \mathbf{A}\nabla + \frac{e^2}{\hbar^2 c^2} \mathbf{A}^2 \right) \Psi d^2\boldsymbol{\rho}' \tag{101}$$

we make the following approximations, typical of this method:[35] the exact wave function $\Psi$ under the integral sign in (101) is replaced by its high-energy approximation:

$$\Psi \approx \exp\left( ikx + \frac{ie}{\hbar c} \int_{-\infty}^{x} A_x\, dx \right).$$

Then

$$\Psi = \Psi_0 - \frac{2ke}{\hbar c} \int G_0 e^{ikx'} \left( A_x + \frac{e}{2\hbar ck} \mathbf{A}^2 \right) \times \exp\left( \frac{ie}{\hbar c} \int_{-\infty}^{x'} A_x\, dx \right) dx'\, dy'.$$

For sufficiently high energies the term quadratic in the vector potential can be neglected:

$$\Psi = \Psi_0 + 2ik \int G_0 e^{ikx'} \frac{\partial}{\partial x'} \exp\left( \frac{ie}{\hbar c} \int_{-\infty}^{x'} A_x\, dx \right) dx'\, dy'$$

Replacing $G_0$ by its plane-wave analog $(1/4i)\, H_0^{(1)}(k|\boldsymbol{\rho} - \boldsymbol{\rho}'|)$ and taking the limit $\rho \to \infty$, for the magnetic scattering amplitude we obtain

$$f_1 = \sqrt{\frac{k}{2\pi i}} \int e^{ikx'(1-\cos\varphi)} e^{-iky'\sin\varphi} \frac{\partial}{\partial x'} \times \exp\left( \frac{ie}{\hbar c} \int_{-\infty}^{x'} A_x\, dx \right) dx'\, dy'.$$

Since at high energies scattering at small angles dominates, we can set $\exp[ikx'(1-\cos\varphi)] \approx 1$. Then

$$\left.\begin{aligned} f_{\text{H.E}} &= -\sqrt{\frac{2}{\pi i k}} [1 - \exp(2i\pi\gamma)] \frac{\sin[k(d-R)\sin\varphi]}{\sin\varphi}; \\ \sigma_{\text{H.E}} &= \frac{8}{\pi k} \left\{ \sin\pi\gamma \frac{\sin[k(d-R)\sin\varphi]}{\sin\varphi} \right\}^2; \\ \int \sigma_{\text{H.E}}\, d\varphi &= 16(d-R)\sin^2\pi\gamma. \end{aligned}\right\} \tag{102}$$

These expressions coincide with (97)–(99) for small $\gamma$.

The generalization to the case of cylinders of different radii is trivial. Let us consider two cylinders (Fig. 3) corresponding to $\mu = \mu_1$ and $\mu = -\mu_2$ ($\mu_2 > 0$). We set $R_1 = a/\sinh\mu_1$, $R_2 = a/\sinh\mu_2$, $d_1 = a\coth\mu_1$, and $d_2 = a\coth\mu_2$. Then if the magnetic fluxes in the cylinders are equal in magnitude and opposite in sign ($\Phi_1 = -\Phi_2 = \Phi$), the expressions (97) and (102) for the cross sections obtained earlier are valid, with $d$ and $R$ taken to be $d = \frac{1}{2}(d_1 + d_2)$ and $R = \frac{1}{2}(R_1 + R_2)$.

### 5.5. Scattering for an unusual orientation of the wave vector

Let us consider the case in which the initial wave vector is directed along the $y$ axis (Fig. 4). We place an impenetrable screen in the plane $x = 0$ between the impenetrable solenoids. Then the wave function vanishes on the screen and inside the cylinders. The following obvious relation holds:

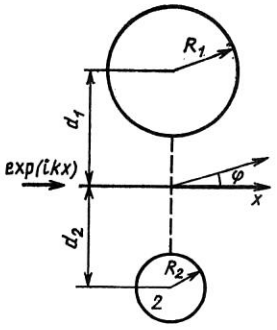


FIG. 3. The same as in Fig. 2 but for cylinders of different radii.

$$\Psi_W^{\gamma} = \Psi_W^0 \exp\left(\frac{ie\chi}{\hbar c}\right). \tag{103}$$

Here $\Psi_W^{\gamma}$ and $\Psi_W^0$ are the wave functions describing the scattering of charged particles on two impenetrable cylinders joined by an impenetrable screen, respectively, with and without a current in the solenoid. They are discontinuous and, as mentioned above, vanish at $x = 0$, $|y| < d - R$. Obviously, $\Psi_W^0$ is an even function of $x$. Let us now consider the case $\gamma = \frac{1}{2}$. From the properties of the generating function $\chi$ for the two solenoids and from (103), it follows that $\Psi_W^{1/2}$ changes sign for $x = 0$, $|y| < d - R$ and, therefore, has a node exactly at the location of the impenetrable screen. Therefore, the physical situation is not changed if the impenetrable screen is removed. As a result, we have

$$\Psi^{1/2} = \Psi_W^{1/2} = \Psi_W^0 \exp\left(\frac{ie\chi}{\hbar c}\right)\Big|_{\gamma=1/2}. \tag{104}$$

For infinitely thin solenoids ($\chi = -\Phi\theta/2\pi$) the preceding expression takes the form[29]

$$\Psi^{1/2} = \Psi_W^{1/2} = \Psi_W^0 \exp(-i\,\theta/2). \tag{105}$$

Here $\Psi_W^0$ is the wave function describing the scattering on the region of infinite length along the $z$ axis, width $2(d - R)$ along the $y$ axis, and zero width along the $x$ axis. A closed expression for this function is known.[36] Relation (104) implies that for $\gamma = \frac{1}{2}$ and initial wave vector along the $y$ axis the effect of the magnetic field is equivalent to the introduction of an impenetrable screen between the solenoids.[29,30] We note that for $\gamma = \frac{1}{2}$ the wave function vanishes on the segment of the $y$ axis lying between the solenoids even in the absence of the impenetrable screen. This fact will be used below.

## 6. SCATTERING OF CHARGED PARTICLES ON A TOROIDAL SOLENOID

In the preceding section we saw that the case of scattering on two cylindrical solenoids with $\Phi_1 = -\Phi_2$ has certain advantages over the case of scattering on a single solenoid. However, the finiteness of the solenoid length leads to leakage of the magnetic field and the possibility of interpreting the AB effect as the scattering of particles on these fields. From this point of view it is interesting to consider scattering on a solenoid of finite dimensions, the simplest of which is a toroidal solenoid. For what follows we shall need the basic quantities pertaining to the toroidal solenoid: the vector potential, the generating function, and so on.

### 6.1. The magnetic vector potential of a toroidal solenoid (Ref. 37)

The magnetic field of a toroidal solenoid $(\rho - d)^2 + z^2 = R^2$ is equal to zero outside the solenoid and $H_\varphi = g/\rho$, $H_\rho$, $H_z = 0$ inside it. Here $\rho$ is the distance of a point inside the solenoid from the solenoid axis (we assume that the solenoid axis coincides with the $z$ axis and that its equatorial plane coincides with the plane $z = 0$). The constant $g$ has the following dependence on the total number of windings $n$ and current strength $J$: $g = 2nJ/c$ ($c$ is the speed of light). The flux of the magnetic field through a cross section of the solenoid is $\Phi = 2\pi g(d - \sqrt{d^2 - R^2})$. In what follows we shall make extensive use of toroidal coordinates. Therefore, here we give the necessary mathematical details. Cylindrical coordinates are related to toroidal coordinates in the following manner:

$$\rho = a\frac{\operatorname{sh}\mu}{\operatorname{ch}\mu - \cos\theta}, \quad z = a\frac{\sin\theta}{\operatorname{ch}\mu - \cos\theta}, \quad \varphi = \varphi$$
$$(0 < \mu < \infty, \quad -\pi < \theta < \pi, \quad 0 < \varphi < 2\pi),$$

with $\mu$ and $\theta$ suitably expressed in terms of the spherical coordinates $r$ and $\theta_s$:

$$\operatorname{cth}\mu = \frac{r^2 + a^2}{2ra\sin\theta_s}, \quad \operatorname{ctg}\theta = \frac{r^2 - a^2}{2ra\cos\theta_s}. \tag{106}$$

From this it follows that $\mu$ and $\theta$ tend to zero for $r \to \infty$:

$$\mu \approx \frac{2a}{r}\sin\theta_s, \quad \theta \approx \frac{2a}{r}\cos\theta_s. \tag{107}$$

From (106) we see that neighboring points lying on opposite sides of the plane $r = 0$ for $\rho < a$ correspond to values of $\theta$ differing by $2\pi$:

$$\theta = \pm\left(\pi - \frac{2a\,|\delta z|}{|\rho^2 - a^2|}\right).$$

For fixed $\mu$ and varying $\theta$ and $\varphi$, the points $x, y, z$ form the surface of a torus $(\rho - a\coth\mu)^2 + z^2 = a^2/\sinh^2\mu$. Let the value $\mu = \mu_0$ correspond to the surface of the solenoid $T_0$. Then for $\mu < \mu_0$ ($> \mu_0$) the points $x, y, z$ lie outside (inside) the solenoid. The radii of the cross section $R$ and the axial line $d$ of the solenoid are expressed in terms of $a$ and $\mu_0$ as $R = d/\sinh\mu_0$, $d = a\coth\mu_0$. For $u_0 \to \infty$ the solenoid degenerates into a filament of radius $a$ lying in the plane $z = 0$.

The Cartesian components of the vector potential in the Coulomb gauge (div $\mathbf{A} = 0$) satisfy the following system of differential equations:

$$\Delta A_i = \frac{g}{a^2}\frac{(\operatorname{ch}\mu_0 - \cos\theta)^2}{\operatorname{sh}\mu_0}\,\delta(\mu - \mu_0)\,(\mathbf{n}_\theta)_i.$$

Here $\mathbf{n}_\theta$ is a vector tangent to the surface of the solenoid:

$$n_\theta = -\frac{1}{\operatorname{ch}\mu_0 - \cos\theta}(\sin\theta\operatorname{sh}\mu_0\cos\varphi,$$
$$\sin\theta\operatorname{sh}\mu_0\sin\varphi, \quad 1 - \operatorname{ch}\mu_0\cos\theta).$$

Owing to the axial symmetry, $A_z$ is independent of the angle $\varphi$, and $A_x$ and $A_y$ depend trivially on it: $A_x = A_\rho\cos\varphi$, $A_y = A_\rho\sin\varphi$. Obviously, $A_\varphi = A_x\sin\varphi - A_y\cos\varphi = 0$. Setting

$$A_\rho = \sqrt{\operatorname{ch}\mu - \cos\theta}\,\tilde{A}_\rho, \quad A_z = \tilde{A}_z\sqrt{\operatorname{ch}\mu - \cos\theta}, \tag{108}$$

we obtain the following equations for $\tilde{A}_\rho$ and $\tilde{A}_z$:

$$\left(\partial_\mu^2+\operatorname{cth}\mu\partial_\mu+\partial_\theta^2+\frac{1}{4}-\frac{1}{\operatorname{sh}^2\mu}\right)\tilde{A}_\rho=-\frac{g\sin\theta\delta(\mu-\mu_0)}{(\operatorname{ch}\mu_0-\cos\theta)^{3/2}};$$

$$\left(\partial_\mu^2+\operatorname{cth}\mu\partial_\mu+\partial_\theta^2+\frac{1}{4}\right)\tilde{A}_z=-\frac{g\delta(\mu-\mu_0)}{\operatorname{sh}\mu_0}$$

$$\times\frac{1-\operatorname{ch}\mu_0\cos\theta}{(\operatorname{ch}\mu_0-\cos\theta)^{3/2}}. \qquad (109)$$

We seek solutions of these equations in the form of series in $\sin n\theta$ and $\cos n\theta$: $A_z=\Sigma R_n^0(\mu)\cos n\theta$, $A_\rho=\Sigma R_n^1(\mu)\sin n\theta$. Expanding the right-hand sides of (109) in the same way, we find second-order ordinary differential equations for $R_n$. The solutions of these equations which are continuous and finite throughout space (including on the solenoid boundary) have the form

$$R_n^0=\frac{2\sqrt{2}g}{\pi}A_n^0\begin{cases}P_{n-\frac{1}{2}}(\operatorname{ch}\mu)\,Q_{n-\frac{1}{2}}(\operatorname{ch}\mu_0) & (\mu<\mu_0);\\ Q_{n-\frac{1}{2}}(\operatorname{ch}\mu)\,P_{n-\frac{1}{2}}(\operatorname{ch}\mu_0) & (\mu>\mu_0);\end{cases}$$

$$R_n^1=-\frac{2\sqrt{2}g}{\pi}Q_{n-\frac{1}{2}}(\operatorname{ch}\mu_0)$$

$$\times\begin{cases}P^1_{n-\frac{1}{2}}(\operatorname{ch}\mu)\,[Q_{n+\frac{1}{2}}(\operatorname{ch}\mu_0)-Q_{n-\frac{3}{2}}(\operatorname{ch}\mu_0)] & (\mu<\mu_0);\\ Q^1_{n-\frac{1}{2}}(\operatorname{ch}\mu)\,[P_{n+\frac{1}{2}}(\operatorname{ch}\mu_0)-P_{n-\frac{3}{2}}(\operatorname{ch}\mu_0)] & (\mu>\mu_0).\end{cases} \qquad (110)$$

Here we have taken

$$A_0^0=\frac{1}{2}Q_{\frac{1}{2}}(\operatorname{ch}\mu_0),$$

$$A_n^0=\frac{2\sqrt{2}g}{\pi}\left[\left(n+\frac{1}{2}\right)Q_{n+\frac{1}{2}}(\operatorname{ch}\mu_0)\right.$$

$$\left.-\left(n-\frac{1}{2}\right)Q_{n-\frac{3}{2}}(\operatorname{ch}\mu_0)\right];\quad P_\nu^\mu(z),\ Q_\nu^\mu(z)$$

$P_\nu^\mu(z)$ and $Q_\nu^\mu(z)$ are the Legendre functions of the first and second kind, respectively. Henceforth we shall understand the argument of the Legendre functions to be $\cosh\mu$ unless explicitly states otherwise. It is easily verified that $\operatorname{div}\mathbf{A}=0$. We note that $A_z$ and $A_\rho$ are even and odd functions of $z$, respectively. Furthermore, $A_\rho=0$ in the plane $z=0$ and on the $z$ axis. At large distances $A_\rho$ and $A_z$ fall off as $r^{-3}$:

$$A_z\approx\frac{4ga^3}{\pi r^3}C(1+3\cos 2\theta_s),\quad A_\rho\approx\frac{12ga^3}{\pi r^3}C\sin 2\theta_s. \qquad (111)$$

Here the constant is $C=(\pi^2/32)(\operatorname{ch}\mu_0/\operatorname{sh}^3\mu_0)$. Let us now consider the behavior of $A_z$ in more detail. On the $z$ axis,

$$A_z(\rho=0,z)=\frac{\sqrt{R}gd}{(d^2+z^2)^{3/4}}Q_{1/2}\left(\frac{d^2+z^2+R^2}{2R\sqrt{d^2+z^2}}\right), \qquad (112)$$

which is equal to $(16ga^3/|z|^3)C$ for large values of $z$. Now let us consider the behavior of $A_z$ in the plane $z=0$. At the origin

$$A_z=\frac{g}{\sqrt{\operatorname{ch}\mu_0}}Q_{1/2}\left(\frac{1+\operatorname{ch}^2\mu_0}{2\operatorname{ch}\mu_0}\right).$$

Inside the solenoid (rather, at $\rho=a=\sqrt{d^2-R^2}$) we have $A_z=gQ_{1/2}(\cosh\mu_0)P_{-1/2}(\cosh\mu_0)$. For large values of $\rho$ the potential $A_z$ is negative: $A_z=-(8g/\pi)(a^3/r^3)C$. It can be shown that $A_z$ changes sign inside the solenoid. These expressions become more transparent for the case of a thin solenoid ($R/d\ll 1$ or $\mu_0\gg 1$). In this case the vector potential is

$$A_\rho\approx\sqrt{2}\pi g\exp(-2\mu_0)\sin\theta\sqrt{\operatorname{ch}\mu-\cos\theta}\,P^1_{1/2};$$

$$A_z\approx\frac{\pi g}{\sqrt{2}}\exp(-2\mu_0)\sqrt{\operatorname{ch}\mu-\cos\theta}\,(P_{-\frac{1}{2}}-\cos\theta P_{1/2}) \qquad (113)$$

outside the solenoid and $A_\rho\approx g\exp(-\mu)\sin\theta$, $A_z\approx g\exp\times(-\mu)\cos\theta$ inside it. In the plane $z=0$, $A_z$ increases from $2\pi g\exp(-2\mu_0)=\pi gR^2/(2d^2)$ at the origin to $g\exp(-2\mu_0)=gR/(2d)$ on the inner boundary ($\rho=d-R$) of the solenoid. Inside the solenoid $A_z$ vanishes at $\rho=d$ and is negative, $-gR/2d$, on the outer wall ($\rho=d+R$). For large values of $\rho$, $A_z$ remains negative, decreasing for $\rho\to\infty$: $A_z\approx-(\pi ga^3/r^3)\exp(-2\mu_0)$. Let us now consider $A_z(\rho,z)$ as a function of $z$ for fixed $\rho$. On the $z$ axis $A_z(0,z)=(\pi g/2)\,[dR^2/(z^2+d^2)^{3/2}]$. For fixed $\rho<d$ the function $A_z(\rho,z)$ is positive for all $z$. For $\rho>d$, $A_z$ is negative for small $z$ and positive for large $z$. The zeros of $A_z$ in the $(\rho,z)$ plane lie on a line which begins at the point $(d,0)$ and has straight-line asymptotes $z=\pm\rho/\sqrt{2}$. Because of this, $\oint A_l dl$ along a closed contour is equal to the magnetic-field flux if the integration contour passes through the hole in the solenoid, and zero otherwise.

### 6.2. The generating function for a toroidal solenoid (Refs. 37 and 38)

Since outside the solenoid $\operatorname{curl}\mathbf{A}=0$, $\mathbf{A}$ can be represented as the gradient of some function $\chi$. Since $\oint A_l dl\neq 0$ for a closed contour passing through the hole in the solenoid, $\chi$ is a multivalued function of the Cartesian coordinates. For convenience let us first consider a thin solenoid ($R/d\ll 1$). Using the vector potential found above [see (113)], we easily find the expressions determining the function $\chi_0$:

$$\frac{1}{a}\frac{\partial\chi_0}{\partial\theta}=\frac{A_\theta}{\operatorname{ch}\mu-\cos\theta}=\frac{g\pi}{\sqrt{2}}\exp(-2\mu_0)\frac{\cos\theta P_{-1/2}-P_{1/2}}{(\operatorname{ch}\mu-\cos\theta)^{1/2}},$$

$$\frac{1}{a}\frac{\partial\chi_0}{\partial\mu}=\frac{A_\mu}{\operatorname{ch}\mu-\cos\theta}$$

$$=\sqrt{2}\pi g\exp(-2\mu_0)P^1_{-1/2}\frac{1}{(\operatorname{ch}\mu-\cos\theta)^{1/2}}. \qquad (114)$$

We adopt the convention that the argument of the Legendre functions is $\cosh\mu$ unless it states otherwise. Integrating (114) over $\theta$ and $\mu$ and equating them, we find

$$\chi_0(\theta,\mu)=\frac{\pi ga}{\sqrt{2}}\exp(-2\mu_0)$$

$$\times\left(P_{-\frac{1}{2}}\int_0^\theta\frac{\cos\theta\,d\theta}{\sqrt{\operatorname{ch}\mu-\cos\theta}}-P_{\frac{1}{2}}\int_0^\theta\frac{d\theta}{\sqrt{\operatorname{ch}\mu-\cos\theta}}\right)$$

$$=\sqrt{2}\pi ga\exp(-2\mu_0)\sin\theta\int_0^\mu\frac{d\mu}{\sqrt{\operatorname{ch}\mu-\cos\theta}}P^1_{-\frac{1}{2}}$$

$$-2\pi ga\exp(-2\mu_0)\left(1-\cos\frac{\theta}{2}\right)\operatorname{sgn}\theta. \qquad (115)$$

The integrals over the angle $\theta$ are expressed in terms of the elliptic integrals $E$ and $F$:

$$\chi_0(\theta,\mu)=\sqrt{2}\pi ga\exp(-2\mu_0)$$

$$\times\left\{\frac{\operatorname{ch}\mu P_{-\frac{1}{2}}-P_{\frac{1}{2}}}{\sqrt{1+\operatorname{ch}\mu}}F\left(\delta,\sqrt{\frac{2}{1+\operatorname{ch}\mu}}\right)\right.$$

$$-P_{-\frac{1}{2}}\left[\sqrt{1+\operatorname{ch}\mu}\,E\left(\delta,\sqrt{\frac{2}{1+\operatorname{ch}\mu}}\right)\right.$$

$$\left.\left.-\frac{\sin\theta}{\sqrt{\operatorname{ch}\mu-\cos\theta}}\right]\right\}.$$

From this expression and from the asymptotic behavior of the Legendre functions for $\mu \to 0$ we find

$$\chi_0 \approx -\pi g \exp(-2\mu_0) \frac{a^3}{r^2} \cos\theta_s \ (r \to \infty) \qquad (116)$$

($r$, $\theta_s$, and $\varphi$ are the usual spherical coordinates). Therefore, at large distances $\chi_0$ falls off as $r^{-2}$. We expand the integrand in (115) in $\cos n\theta$ and $\sin n\theta$ and integrate over $\theta$;

$$\chi_0 = ga \exp(-2\mu_0) \cdot \left\{-2\theta + \sum \frac{\sin n\theta}{n} \left[P_{-\frac{1}{2}}(Q_{n+\frac{1}{2}} + Q_{n-\frac{3}{2}}) - 2P_{\frac{1}{2}} Q_{n-\frac{1}{2}}\right]\right\}. \qquad (117)$$

It is easily verified that the integral $\oint A_l dl$ is equal to ($\mu = \text{const} < \mu_0$)

$$a \int_{-\pi}^{\pi} A_\theta \frac{d\theta}{\operatorname{ch}\mu - \cos\theta} = -\int_{-\pi}^{\pi} \frac{\partial \chi_0}{\partial \theta} d\theta = -4\pi g a \exp(-2\mu_0),$$

i.e., up to a sign it is the flux $\Phi$ of the magnetic field if the integration contour passes through the hole in the solenoid, and zero otherwise.

For a finite ($R \approx d$) solenoid the direct integration of the vector-potential components $A_\rho$, $A_z$ is very difficult. Here we give the final result (see Ref. 38 for details). The generating function for a finite solenoid has the form

$$\chi(\theta, \mu) = \frac{\chi_0(\theta, \mu)}{1 - \exp(-2\mu_0)} + \frac{4\sqrt{2}\, g}{\pi} \sqrt{\operatorname{ch}\mu - \cos\theta} \sum_{n=1}^{\infty} \beta_n P_{n-\frac{1}{2}} \sin n\theta, \qquad (118)$$

where $\chi_0$ (the generating function for a thin solenoid) is defined in (117) and

$$\beta_n = -\sum_{k=n}^{\infty} Q_{k-\frac{1}{2}}(\operatorname{ch}\mu_0) Q_{k+\frac{1}{2}}(\operatorname{ch}\mu_0).$$

Using the asymptotic behavior of the Legendre functions, from (118) we obtain

$$\chi(\theta, \mu) \approx -\frac{\pi g}{4} \frac{\operatorname{ch}\mu_0}{\operatorname{sh}^3 \mu_0} \frac{a^3}{r^2} \cos\theta_s. \qquad (119)$$

The following properties of the function $\chi$ will be needed below. First, $\chi$ falls off as $r^{-2}$ at large distances from the solenoid. Second, $\chi$ undergoes a finite jump (equal to $-\Phi = -4\pi g a/(\exp(2\mu_0) - 1)$ in crossing the circle of radius $d - R$ lying in the equatorial plane of the solenoid ($z = 0$). In other words, the region of discontinuity of the generating function of a toroidal solenoid fills a circle of radius $d - R$ lying in the equatorial plane.

### 6.3. The amplitude for scattering of charged particles by the magnetic field of a toroidal solenoid

In this section we shall consider the scattering of charged spinless particles by the irrotational vector potential of a toroidal solenoid. To prevent the penetration of particles into the region with $H \neq 0$, this region is screened by an infinite repulsive potential of suitable geometric form. This causes the wave function to vanish at the boundary of the region of infinite repulsion (as it does also inside this region).

*The first Born approximation.* We therefore must find the solution of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar c}\mathbf{A}\right)^2 \Psi + V\Psi = E\Psi, \qquad (120)$$

with $V = \infty$, $\Psi = 0$ inside the screen and on its boundary and $V = 0$ at other points in space.

Let us first consider the case in which the solenoid $T_0$ is enclosed in a screen of toroidal shape. We assume that it coincides with $T_0$. Then in the first Born approximation in the vector potential we have

$$\Psi = \Psi_0 + \frac{2ie}{\hbar c} \int G_0(\mathbf{r}, \mathbf{r}') \mathbf{A} \nabla \Psi_0 \, dV'. \qquad (121)$$

Here $\Psi_0$ and $G_0$ are the wave function and the Green's function corresponding to scattering on an impenetrable torus in the absence of a magnetic field. They obviously vanish at the boundary of $T_0$. The integration in (121) is carried out outside the solenoid, where $\mathbf{A} = \operatorname{grad}\chi$. Furthermore, we have the identity

$$\mathbf{A}\operatorname{grad}\Psi_0 = \frac{\partial \chi}{\partial x_k}\frac{\partial \Psi_0}{\partial x_k} = \frac{1}{2}\Delta(\chi\Psi_0) - \frac{1}{2}\chi\Delta\Psi_0 = \frac{1}{2}(\Delta + k^2)\chi\Psi_0$$

($\Delta$ and $k^2$ are the Laplacian and the wave number, respectively). Then we obtain

$$\Psi = \Psi_0 + \frac{ie}{\hbar c} \int G_0(\Delta + k^2)\chi\Psi_0 \, dV'. \qquad (122)$$

We integrate (122) twice by parts, using the equation for $G_0$:

$$\Psi = \Psi_0 + \frac{ie}{\hbar c}\chi\Psi_0 + \frac{ie}{\hbar c}\int \operatorname{div}[G_0 \operatorname{grad}\chi\Psi_0 - \chi\Psi_0 \operatorname{grad} G_0]\, dV'. \qquad (123)$$

To calculate the scattering amplitude we must find the limit of (123) for $\rho \to \infty$. The first term on the right-hand side gives the incident plane wave and the amplitude for scattering on an impenetrable torus in the absence of a magnetic field. The second term can be neglected, since $\chi$ falls off as $r^{-2}$ for $r \to \infty$. Let us consider the third term in (123). For the moment we neglect the fact that the vector function on which the divergence operates is discontinuous (owing to the $\chi$ function). Using Gauss's theorem, we replace the volume integral by an integral over the surface containing this volume. This surface consists of the surface of the torus $T_0$ and the surface of a sphere $C_R$ of sufficiently large radius $R$. The integral over the surface of the torus vanishes, since both $\Psi_0$ and $G_0$ vanish on it. We are therefore left with the integral over $C_R$:

$$R^2 \int \left[G_0 \frac{\partial(\chi\Psi_0)}{\partial R} - \chi\Psi_0 \frac{\partial G_0}{\partial R}\right] d\Omega' \quad (d\Omega' = \sin\theta_s' \, d\theta_s' \, d\varphi').$$

Since at large distances $\Psi_0 \approx \exp(ikz)$,

$$\chi \approx -\frac{\pi g}{4}\frac{\operatorname{ch}\mu_0}{\operatorname{sh}^3\mu_0} r^{-2} \cos\theta_s,$$

$$G_0 \approx -\frac{1}{4\pi r}\exp[ik(r - \mathbf{rn}')],$$

the integral over $C_R$ reduces to

$$-\frac{g}{16}\frac{kea^3}{\hbar c}\frac{\operatorname{ch}\mu_0}{\operatorname{sh}^3\mu_0}\frac{\exp(ikR)}{R}\int \cos\theta_s'(1 + \cos\theta_s') \times \exp[ik(R\cos\theta_s' - \mathbf{rn}')]\, d\Omega'.$$

It vanishes for $R \to \infty$. Now we take into account the fact that the divergence in (123) operates on a discontinuous function. Detailed considerations shows that Gauss's theorem must be modified. The surface integral must be augmented by an integral over the region of discontinuity of the $\chi$ function:

$$\frac{ie\Phi}{\hbar c}\int\left(G_0\frac{\partial\Psi_0}{\partial z'}-\Psi_0\frac{\partial G_0}{\partial z'}\right)\rho'\,d\rho'\,d\varphi\quad(\rho'\leqslant d-R,\ z'=0). \tag{124}$$

Let us now calculate the scattering amplitude. For this we take the limit $r\to\infty$ in (124), replacing $G_0$ by its asymptote:

$$f=-\frac{e\Phi}{4\pi\hbar c}\int\left[\exp(-ik\mathbf{n}\boldsymbol{\rho}')\frac{\partial\Psi_0}{\partial z'}-\Psi_0\frac{\partial}{\partial z'}\exp(-ik\mathbf{n}\mathbf{r}')\right]\rho'\,d\rho'\,d\varphi\quad(\rho'\leqslant d-R,\ z'=0). \tag{125}$$

We see that the scattering amplitude depends on the behavior of the wave function near the solenoid. The explicit form of the latter is unknown, since the wave equation is not separable in toroidal coordinates. We replace $\Psi_0$ by a plane wave $\exp(ikz)$. Then[38]

$$f=\frac{\gamma k}{2}(1+\cos\theta_s)\int\exp(-ik\rho'\sin\theta_s\cos\varphi)\,\rho'\,d\rho'\,d\varphi$$
$$=\pi\gamma(1+\cos\theta_s)(d-R)\frac{J_1[k(d-R)\sin\theta_s]}{\sin\theta_s},\quad\gamma=e\Phi/hc. \tag{126}$$

For a thin solenoid this expression reduces to

$$\pi\gamma(1+\cos\theta_s)\,d\frac{J_1(kd\sin\theta_s)}{\sin\theta_s}. \tag{127}$$

This expression was obtained earlier in the interesting study of Ref. 39. However, the procedure for obtaining (127) was somewhat dubious, because the vector potential used in that study had a $\delta$-type singularity in the plane $z=0$.

As a result, the discarded term $-(e^2/\hbar^2c^2)\mathbf{A}$ is more singular than the term $-(ie/\hbar c)(2\mathbf{A}\,\mathrm{grad}\,\Psi_0+\Psi_0\,\mathrm{div}\,\mathbf{A})$ used in the calculation ($\mathrm{div}\,\mathbf{A}\neq0$ for the potentials of Ref. 39).

*The high-energy approximation.* Let us now calculate the amplitude for magnetic scattering on a toroidal solenoid in the high-energy approximation. For this, in the Lippmann–Schwinger equation

$$\Psi=\Psi_0+\int G_0(\mathbf{r},\ \mathbf{r}')\,V_1\Psi\ (\mathbf{r}')\,dV'\left(V_1=\frac{2ie}{\hbar c}\mathbf{A}\nabla\right) \tag{128}$$

we make the following simplifications typical of this method:[35] we replace the Green's function $G_0$ by a plane-wave function $-(1/4\pi)(\exp(ik|\mathbf{r}-\mathbf{r}'|)/|\mathbf{r}-\mathbf{r}'|)$, and inside the integral for $\Psi$ we use its high-energy approximation

$$\exp\left(ikz+\frac{ie}{\hbar c}\int_{-\infty}^{z}A_z\,dz\right).$$

This expression is correct if

$$\left|\mathbf{A}^2\frac{e}{\hbar ck}\right|\ll|A_z|.$$

Otherwise, the terms quadratic in the potentials must be added to the integrand. For the vector potentials used in Ref. 39 this leads to an infinite scattering amplitude, as in the first Born approximation. On the other hand, the potentials used in that study are everywhere continuous and finite functions of the coordinates. Therefore, the condition cited above is satisfied at sufficiently high energies. For the magnetic scattering amplitude we find

$$f(\mathbf{n})=\frac{ek}{hc}\int\exp(i\mathbf{q}\mathbf{r}')\,A_z\exp\left(\frac{ie}{\hbar c}\int_{-\infty}^{z'}A_z dz\right)dV'. \tag{129}$$

Here $\mathbf{q}=\mathbf{k}-\mathbf{k}'$ is the momentum transfer, $\mathbf{k}'=\mathbf{n}\cdot\mathbf{k}$, and $\mathbf{k}=\mathbf{n}_z k$. Since scattering at small angles dominates at high energies, the vector $\mathbf{q}$ can be assumed to be perpendicular to the initial vector $\mathbf{k}$; therefore, $\mathbf{q}$ lies in the plane $z=0$. Then[38]

$$f(\mathbf{n})=\frac{ek}{hc}\int d^2\boldsymbol{\rho}'\exp(i\mathbf{q}\boldsymbol{\rho})\int_{-\infty}^{\infty}dzA_z\exp\left(\frac{ie}{\hbar c}\int_{-\infty}^{z}A_z\,dz\right)$$
$$=-\frac{ik}{2\pi}\int d^2\boldsymbol{\rho}\exp(i\mathbf{q}\boldsymbol{\rho})\left[\exp\left(\frac{ie}{\hbar c}\int_{-\infty}^{\infty}A_z\,dz\right)-1\right]$$
$$=-ik[\exp(2i\pi\gamma)-1]\int_0^{d-R}J_0(q\rho)\,\rho\,d\rho$$
$$=-ik(d-R)[\exp(2i\pi\gamma)-1]\,J_1[q(d-R)]/q. \tag{130}$$

This amplitude becomes the Born amplitude for small $\gamma$ and $\theta$. Expressions (126) and (130) are easily generalized to the case in which the toroidal solenoid is completely enclosed in an impenetrable torus (Fig. 5):

$$f_B=\pi\gamma\frac{1+\cos\theta}{\sin\theta}J_1(ka\sin\theta); \tag{131}$$

$$f_{H.E}=-ika[\exp(2i\pi\gamma)-1]\,J_1(qa)/q. \tag{132}$$

*Scattering for unusual configurations of the wave vector.* Let us again consider scattering on a toroidal solenoid surrounded by an impenetrable torus $(\rho-d)^2+z^2=R^2$. Now let the initial wave vector be perpendicular to the $z$ axis, for example, directed along the $x$ axis (Fig. 6). We place an impenetrable barrier at $z=0$, $\rho\leqslant d-R$. We then have the following relation:

$$\Psi_W^\gamma=\exp\left(\frac{ie\chi}{\hbar c}\right)\Psi_W^0\left(\gamma=\frac{e\Phi}{hc}\right). \tag{133}$$

Here $\Psi_W^\gamma$ and $\Psi_W^0$ are the wave functions corresponding to scattering on an impenetrable torus with a hole covered by an impenetrable barrier, respectively, with and without a current present in the solenoid windings. The subscript on the wave functions denotes the presence of the impenetrable barrier, and the superscript shows the value of the constant $\gamma$ at which they are calculated. The functions $\Psi_W^0$ and $\Psi_W^\gamma$ are everywhere continuous and vanish both inside the impenetrable cylinders and on the surface of the impenetrable barrier. The function $\chi$ is the generating function of the toroidal solenoid. From its explicit form it follows that it changes from $-\Phi/2$ to $\Phi/2$ in passing through the impenetrable barrier and is an odd function of $z$. Obviously, $\Psi_W^0$ is an even function of $z$. Now let $\gamma=\frac{1}{2}$. From (133) we find that $\Psi_W^{1/2}$ differs by a sign on opposite sides of the impenetrable barrier, and vanishes on it. It therefore has a node there. Thus, the impenetrable barrier can be removed without changing the physical conditions. Finally, we have

$$\Psi_W^{1/2}=\Psi_0^{1/2}=\exp\left(\frac{ie\chi}{\hbar c}\right)\Big|_{\gamma=\frac{1}{2}}\Psi_W^0. \tag{134}$$

For an infinitely thin solenoid $\Psi_W^0$ corresponds to scattering on a thin disc of radius $d$ lying in the plane $z=0$ (Ref. 40). A closed expression is known for this wave function.[36] For $\gamma=\frac{1}{2}$ and the initial wave vector in the plane $z=0$ the presence of the magnetic field is equivalent to covering the hole of the torus by an impenetrable barrier. From (134) it follows that the wave function for $\gamma=\frac{1}{2}$ vanishes for $z=0$, $\rho<d-R$ even in the absence of the impenetrable barrier. We shall make use of this fact below.

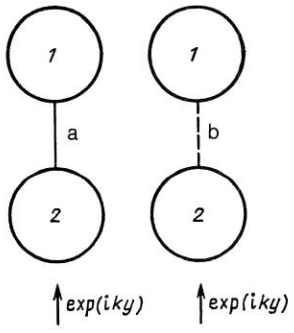


FIG. 4. An incident wave scatters on two impenetrable solenoids in the presence (a) and in the absence (b) of an impenetrable partition between them.

## 7. SCATTERING OF CHARGED PARTICLES IN AN IDEAL MULTIPLY CONNECTED SPACE

Now we surround one or several solenoids by topologically inequivalent screens and study the scattering on the magnetic fields accessible to the particles. Outside the solenoids the Schrödinger equation is satisfied by the following wave function:

$$\Psi = \Psi_0 \exp\left(\frac{ie\chi}{\hbar c}\right), \qquad (135)$$

where $\chi$ is the generating function defined above and $\Psi_0$ is the wave function in the absence of the magnetic field. The transformation (135) is unitary, and this guarantees that all the observables for $\Psi$ and $\Psi_0$ coincide. We always choose the wave function $\Psi$ (the solution of the initial full Schrödinger equation) to be single-valued. Therefore (owing to the discontinuity of the $\chi$ function) $\Psi_0$ is not single-valued. The absence of a unitary transformation relating the continuous function $\Psi$ to the continuous function $\Psi_0$ is the reason for the appearance of the AB effect. However, there exists an important exception when the two functions appearing in (135) are continuous. This occurs when the region of discontinuity of the function $\chi$ is inaccessible to the incident particles. Then at the discontinuity $\Psi = \Psi_0 = 0$ and the relation (135), which in the region of discontinuity is satisfied trivially, $0 = 0$, is a unitary transformation between the continuous functions with $\mathbf{A} \neq 0$ and $\mathbf{A} = 0$. In this situation the AB effect is absent. Typical examples are shown in Figs. 7–10: two cylindrical solenoids ($\Phi_1 = -\Phi_2$) in an impenetrable cylinder (Fig. 7), a toroidal solenoid in an impenetrable cylinder (Fig. 8) and in a sphere (Fig. 9), and in one section of an impenetrable torus (Fig. 10). We note that in the multiply connected space common to these cases (for example, the space outside the impenetrable torus) the AB effect can occur (Fig. 6), or it also can be absent (Fig. 10). This is also true for the impenetrable cylinder: the AB effect occurs if one solenoid is located inside the cylinder, and it does not occur if two solenoids with $\Phi_1 = -\Phi_2$ (Fig. 7) or a toroidal solenoid (Fig. 8) are located inside it. We conclude that the existence of the AB effect is determined not only by the fact that the space accessible to the particles is multiply connected, the nontriviality of the vector potentials in this space, and the single-valuedness of the wave functions which are used.[38,41] The condition for this effect to exist can be expressed either by means of the generating function (the AB effect occurs if the region of discontinuity of the $\chi$ function is accessible to the incident particles) or by means of a nonintegrable phase factor (there exist closed paths accessible to the particles along which $\oint A_l dl \neq 0$; Ref. 42).

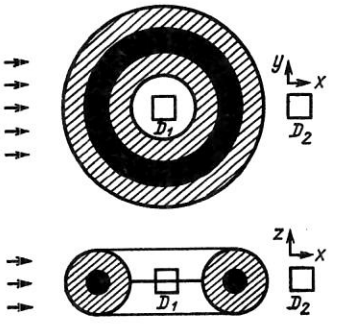


FIG. 6. A toroidal solenoid (in black) surrounded by an impenetrable torus (shaded region). For the initial wave vector in the plane $z = 0$ and $\gamma = \frac{1}{2}$ the wave function vanishes on the portion of the plane $z = 0$ coinciding with the hole in the torus.

## 8. GAUGE TRANSFORMATIONS OF THE VECTOR POTENTIALS

Since outside the solenoid $\mathbf{A} = \operatorname{grad} \chi$, we can attempt to eliminate the vector potential in this region (completely or partially) by means of a suitable gauge transformation. It is known[32,43] that for a single infinite cylindrical solenoid the complete elimination of the vector potential outside the solenoid leads to the appearance of a singular magnetic field on the solenoid axis with direction opposite to the original field. For two solenoids a singular magnetic field appears on the axis of each solenoid. Finally, in the case of a toroidal solenoid this gauge transformation leads to a magnetic field $H = -\Phi\delta(z)\delta(\rho - d)/2\pi$. In each of these cases the problem becomes inequivalent to the original one after the gauge transformation. However, there are some provisos. The situations before and after the gauge transformation are physi-

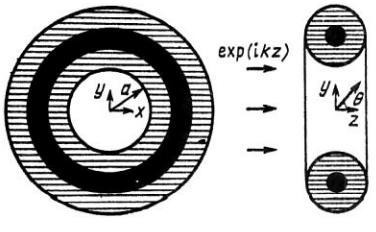


FIG. 5. A toroidal solenoid (in black) surrounded by a potential barrier of height $V_0$ having the form of a torus (shaded region). For $V_0 \to \infty$ the cross section for scattering on the magnetic field tends to a finite value [see (131) or (132)].

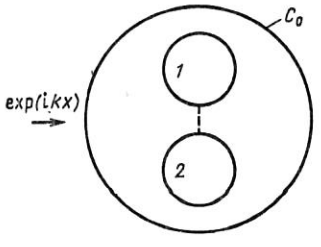


FIG. 7. An illustration of the absence of the AB effect in a multiply connected space. Two cylindrical solenoids with $\Phi_1 = -\Phi_2$ are surrounded by an impenetrable barrier of cylindrical shape.

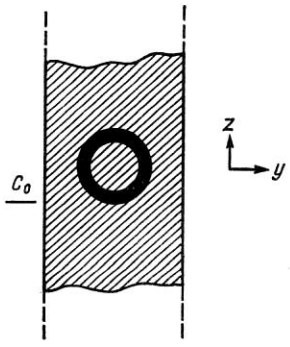


FIG. 8. An illustration of the absence of the AB effect in a multiply connected space. A toroidal solenoid is surrounded by an impenetrable cylindrical barrier.

cally inequivalent if single-valued wave functions are used in both cases. In fact, in the preceding section we saw that the vector potential outside the solenoid can always be made to vanish by means of a unitary transformation. The price for this is a discontinuity in the functions obtained after the gauge transformation. If this price is too high for us, we must restrict ourselves to gauge transformations which do not change the transformation properties of the wave functions.

### 8.1. An infinite cylindrical solenoid

As a first example let us consider an infinite cylindrical solenoid of radius $R$ (Ref. 38). In this case, along with the usual vector potential ($A_x = -HR^2y/2\rho^2, A_y = -HR^2x/2\rho^2$, $A_z = 0$ outside the solenoid and $A_x = -\frac{1}{2}Hy$, $A_y = \frac{1}{2}Hx, A_z = 0$ inside it) we can work with a vector potential $\mathbf{A}'$ giving the same magnetic field as $\mathbf{A}$ (i.e., $\mathbf{H} = H\mathbf{e}_z$ inside the solenoid and $\mathbf{H} = 0$ outside it). The only nonzero component of $\mathbf{A}'$ is $A'_y$, which is equal to $H(x + \sqrt{R^2 - y^2})$ inside the solenoid. Outside solenoid $A'_y \neq 0$ in the region $|y| \leqslant R$, $x > 0$ (Fig. 11): $A'_y = 2H\sqrt{R^2 - y^2}$. The vector potentials $\mathbf{A}$ and $\mathbf{A}'$ are related by a nonsingular gauge transformation:

$$\mathbf{A} = \mathbf{A}' - \operatorname{grad} \frac{\partial \alpha}{\partial y}.$$

The function $\alpha$ is found from the equation

$$\frac{\partial^2 \alpha}{\partial x^2} + \frac{\partial^2 \alpha}{\partial y^2} = Hf(x, y).$$

We have the function $f = x + \sqrt{R^2 - y^2}$ inside the solenoid. Outside the solenoid it is $2\sqrt{R^2 - y^2}$ in the shaded region of Fig. 11 and zero outside this region. For an infinitely thin solenoid $A'_y$ degenerates to

$$A'_y = \Phi\delta(y)\,\theta(x),$$

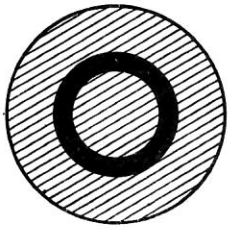

FIG. 9. An illustration of the absence of the AB effect owing to the fact that the space accessible to the particles is simply connected. A toroidal solenoid is surrounded by an impenetrable spherical barrier.

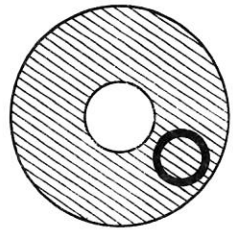

FIG. 10. An illustration of the absence of the AB effect in a multiply connected space. A toroidal solenoid is located in one section of an impenetrable torus.

i.e., the vector potential is nonzero on the positive $x$ semiaxis.

More general gauge transformations are discussed in Ref. 44.

### 8.2. Two solenoids with opposite fluxes

For two solenoids with $\Phi_1 = -\Phi_2$ the following vector potential $\mathbf{A}'$ gives the same magnetic field ($\pm H\mathbf{e}_z$) inside the solenoids as the field obtained earlier in Ref. 31: $A'_y = A'_z = 0$, $A'_x = f\Phi/\pi R^2$. The function $f$ is (Fig. 12): $d + \sqrt{R^2 - x^2} \mp y$ inside the first and second solenoids, respectively; $f = 2\sqrt{R^2 - x^2}$ between the solenoids ($-d + \sqrt{R^2 - x^2} \leqslant y \leqslant d - \sqrt{R^2 - x^2}$, $|x| < R$). Finally, $f = 0$ in the other regions of space. The vector potentials $\mathbf{A}$ and $\mathbf{A}'$ are related by a nonsingular gauge transformation:

$$\mathbf{A} = \mathbf{A}' + \operatorname{grad} \frac{\partial \Psi}{\partial x}.$$

The function $\Psi$ is found from the equation $\Delta\Psi = -fH$. For pointlike solenoids ($R \to 0$, but $\Phi = \text{const}$) the vector potential is nonzero only on the segment $|y| < d$ of the $y$ axis:

$$\mathbf{A}' = \mathbf{e}_x \Phi\delta(x).$$

### 8.3. A toroidal solenoid

For a toroidal solenoid $(\rho - d)^2 + z^2 = R^2$ the following vector potential $\mathbf{A}'$ gives the same magnetic field, $\mathbf{H} = \mathbf{e}_\varphi g/\rho, g = (\Phi/2\pi)(d - \sqrt{d^2 - R^2})^{-1}$ inside the solenoid and $\mathbf{H} = 0$ outside it, as the vector potential found in Ref. 38: $\mathbf{A}' = \mathbf{A}'_z\mathbf{e}_z$, where $A'_z = g \ln((d + \sqrt{R^2 - z^2})/\rho)$ inside the solenoid and $A'_z = g \ln((d + \sqrt{R^2 - z^2})/(d - \sqrt{R^2 - z^2}))$ outside it in the shaded region (Fig. 13). The vector potentials $\mathbf{A}$ and $\mathbf{A}'$ are related by the following gauge transformation:

$$\mathbf{A} = \mathbf{A}' + \operatorname{grad} \frac{\partial \alpha}{\partial z}.$$

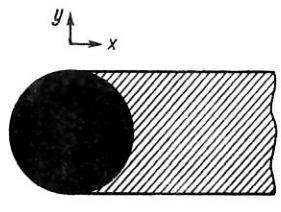


FIG. 11. The vector potential of a cylindrical solenoid in a nonstandard gauge. Outside the solenoid the vector potential is nonzero only in the shaded region.

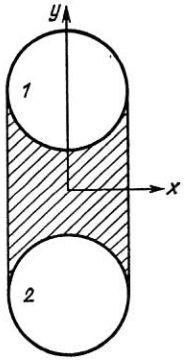


FIG. 12. The vector potential of two cylindrical solenoids with $\Phi_1 = -\Phi_2$ in a non-Coulomb gauge. Outside these solenoids the vector potential is nonzero only in the shaded region.

The functions are

$$\alpha = -\frac{g}{2\pi\sqrt{\rho}}\int_{-R}^{R} dz_1 \beta(z,\ \rho,\ z_1);$$

$$\beta(\rho,\ z,\ z_1) = \ln(d-\sqrt{R^2-z_1^2}) \int_0^{d-\sqrt{R^2-z_1^2}} \sqrt{\rho_1}\, d\rho_1 Q_{-\frac{1}{2}}(x)$$

$$-\ln(d+\sqrt{R^2-z_1^2}) \int_0^{d+\sqrt{R^2-z_1^2}} d\rho_1 \sqrt{\rho_1}\, Q_{-\frac{1}{2}}(x)$$

$$+\int_{d-\sqrt{R^2-z_1^2}}^{d+\sqrt{R^2-z_1^2}} \sqrt{\rho_1}\, d\rho_1 \ln \rho_1 Q_{-\frac{1}{2}}(x) \quad \left(x = \frac{\rho^2+\rho_1^2+(z-z_1)^2}{2\rho\rho_1}\right).$$

For an infinitely thin toroidal solenoid $R \to 0$ we obtain (for fixed $\Phi$):

$$\mathbf{A}' = \mathbf{e}_z \Phi \delta(z)\, \theta(d-\rho),$$

$$\mathbf{A} = \mathbf{A}' + \operatorname{grad} \frac{\partial \alpha}{\partial z},$$

$$\alpha = \frac{R^2 g}{4\pi}\left[2\pi(\sqrt{z^2+d^2} - |z|)\right.$$

$$\left. -2\sqrt{d}\int_0^{\rho} \frac{d\rho}{\sqrt{\rho}}\, Q_{1/2}\left(\frac{z^2+\rho^2+d^2}{2\,d\rho}\right)\right].$$

*We conclude that*: if we agree to work with discontinuous wave functions, then by means of a suitable gauge transformation the vector potential outside the solenoid can always be completely eliminated. Otherwise, we must restrict ourselves to only nonsingular gauge transformations which do not change the transformation properties of the wave functions.

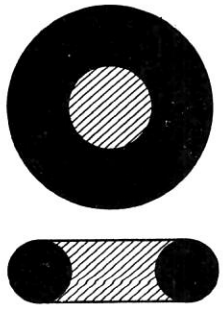

FIG. 13. The vector potential of a toroidal solenoid in a non-Coulomb gauge. Outside the solenoids the vector potential is nonzero only in the shaded region.

## 9. AN ALTERNATIVE INTERPRETATION OF THE AHARONOV–BOHM EFFECT

In the preceding sections the solenoids were surrounded by various absolutely impenetrable topologically inequivalent screens, and we studied the scattering on the magnetic field ($\mathbf{H} = 0$, $\mathbf{A} \neq 0$) surrounding these solenoids. On the other hand, there is an alternative to the AB interpretation of positive experimental results, in which the effect is attributed to leakages of the magnetic field (owing to the finite length of a real cylindrical solenoid) and to the nonvanishing probability that the incident particles scatter on regions of space with $\mathbf{H} \neq 0$ (owing to the finiteness of a real potential barrier). It is incorrect to go to the limit of an ideal solenoid, since the space arising after the transition to infinite barriers is multiply connected. We have seen that in such a multiply connected space inequivalent representations of the angular momentum are allowed which correspond to non-single-valued wave functions.

On the other hand, actual experiments are carried out in a simply connected space, where only single-valued wave functions are allowed. In the remainder of this section we shall assume that the solenoid is surrounded by potential barriers of finite height and shall study the scattering on them.

### 9.1. An infinite cylindrical solenoid

Let us consider a cylindrical solenoid of radius $R$ surrounded by a cylindrical potential barrier $C_0$ of height $V_0$ and radius $b \geqslant R$. Then in first-order perturbation theory in the dimensionless constant $\gamma = e\Phi/hc$ for the amplitude of scattering on the magnetic field we find

$$f_1 = \frac{1}{\sqrt{2\pi i k}} \sum f_{1m} \exp(im\varphi). \tag{136}$$

The partial amplitudes $f_{1m}$ $(m \neq 0)$ are[41]

$$f_{1m} = i\pi\gamma m \left\{-\frac{4}{\pi^2 h_m^2} I_{|m|}^2(k_1 R) - I_{|m|-1}(k_1 R)\, I_{|m|+1}(k_1 R)\right]$$

$$+\frac{1}{8\pi^2}\frac{1}{h_m^2}\frac{1}{|m|}\sum_{s=0}^{|m|}(-1)^{|m|-s}\Delta(0,\ s,\ m)\,[I_s^2 - I_s^2(k_1 R)]$$

$$+\frac{2}{|m|}\sum_{s=0}^{|m|}\Delta(0,\ s,\ m)\,[J_s - r_m H_s^{(1)}]^2\Big\}, \tag{137}$$

where $\Delta(0,\ s,\ m) = (1+\delta_{0s})^{-1}(1+\delta_{s,|m|})^{-1}$ and $\gamma = e\Phi/hc$.

Here and below, the arguments of the ordinary and modified Bessel functions are not indicated if they are $kb$ and $k_1 b$, respectively, furthermore, $k = \sqrt{2\mu E}/\hbar$, $k_1 = \sqrt{2\mu(V_0 - E)}/\hbar$, $r_m = j_m/h_m$, $h_m = kbI_{|m|}\dot{H}^{(1)}_{|m|} - k_1 bI_{|m|}H^{(1)}_{|m|}$, $j_m = kbI_{|m|}\dot{J}_{|m|} - k_1 bI_{|m|}J_{|m|}$.

The first line in this expression is equal to the contribution to the scattering amplitude from the magnetic field inside the solenoid, and the second line comes from the magnetic field outside the solenoid, but inside $C_0$. Finally, the third line is the contribution of the magnetic field outside $C_0$. For zero potential barrier ($V_0 = 0$) expression (137) becomes

$$f_{1m}^{0}=\mathrm{i}\pi\gamma m\Big[J_{|m|}^{2}(kR)-J_{|m|-1}(kR)\,J_{|m|+1}(kR)$$

$$+\frac{1}{|m|}\sum_{s=0}^{|m|}\Delta(0,\ s,\ m)\,J_{s}^{2}(kR)\Big].$$

The first and second lines in this expression come from the magnetic field inside and outside the solenoid. For an infinitely thin solenoid ($R\to 0$) the contribution of the magnetic field inside the solenoid vanishes, while the second line reduces to the usual AB amplitude:

$$f_{m}^{\mathrm{AB}}=-\mathrm{i}\pi\gamma\frac{m}{|m|};$$

$$f_{\mathrm{AB}}=\frac{1}{\sqrt{2\pi ik}}\sum f_{m}^{\mathrm{AB}}\exp(\mathrm{i}m\varphi)=-\frac{\pi\gamma}{\sqrt{2\pi ik}}\operatorname{ctg}\frac{\varphi}{2}.$$

For an infinitely high potential barrier ($V_0=\infty$) the incident wave does not penetrate inside $C_0$:

$$f_{1m}^{\infty}=2\pi\mathrm{i}\gamma\frac{m}{|m|}\sum_{l=0}^{|m|}[J_{l}(kb)-s_{m}H_{l}^{(1)}(kb)]^{2}\,\Delta(0,\ l,\ m)$$

$$(s_{m}=J_{|m|}(kb)/H_{|m|}^{(1)}(kb)). \qquad (138)$$

We conclude that for large values of the potential barrier the contribution of the regions with $H\neq 0$ to the scattering amplitude is negligible. This implies that the positive results of experiments to detect the AB effect cannot be attributed to the finite height of the potential barrier.

### 9.2. Two cylindrical solenoids and a toroidal solenoid

Now let there be located inside $C_0$ two solenoids with $\Phi_1=-\Phi_2$. Then expression (136) is valid with

$$f_{1m}=\frac{e}{\hbar c}\mathrm{i}^{-|m|}\int_{0}^{\infty}\rho\,d\rho\int_{0}^{2\pi}R_{0m}\exp(-\mathrm{i}m\varphi)\,\mathbf{A}\nabla\Psi_{0}\,d\varphi. \qquad (139)$$

Here $\mathbf{A}$ is the vector potential for the two solenoids with $\Phi_1=-\Phi_2$, $\Psi_0=\Sigma\mathrm{i}^{|m|}R_{0m}\exp(\mathrm{i}m\varphi)$; $R_{0m}=(2\mathrm{i}/\pi)I_{|m|}(k_1\rho)$ inside $C_0$, and $R_{0m}=J_{|m|}(k\rho)-r_{m}H_{|m|}^{(1)}(k\rho)$ outside.

We split the integral (139) into integrals inside the solenoids and outside them. Setting $\mathbf{A}=\operatorname{grad}\chi$ in the second integral and integrating by parts twice, we bring it to the form

$$\frac{e}{\hbar c}(-\mathrm{i})^{|m|}\int\rho\,d\rho\,d\varphi\,\operatorname{div}$$

$$\times(R_{0m}\mathrm{e}^{-\mathrm{i}m\varphi}\operatorname{grad}\chi\Psi_{0}-\chi\Psi_{0}\operatorname{grad}R_{0m}\mathrm{e}^{-\mathrm{i}m\varphi}). \qquad (140)$$

Owing to Gauss's theorem (taking into account the discontinuity of the $\chi$ function), this integral reduces to the sum of integrals over the surfaces of the solenoids, the surface of the cylinder of infinite radius, and the region of discontinuity of the $\chi$ function [which coincides with the segment $(-d+R,\ d-R)$ of the $y$ axis]. The integral over the surface of the cylinder of infinite radius vanishes, owing to the rapid falloff of the $\chi$ function. Then in (139) there remains the integration over the nearest neighborhood lying inside $C_0$. For $V_0\to\infty$ the functions $R_{0m}$ and $\Psi_0$ tend to zero inside $C_0$. Therefore, both $f_{1m}$ and $f_1$ vanish. Thus, for a given cylindrical barrier $C_0$ the amplitude for scattering on the magnetic field approaches the finite value (138) for $V_0\to\infty$ in one case (when there is one solenoid inside $C_0$), while in the second case (when there are two solenoids inside $C_0$) it approaches zero. This appears strange if the AB effect is interpreted as the result of scattering on regions of space with $H\neq 0$, since in both cases for finite $V_0$ the probability that the particles penetrate the region with $H\neq 0$ is also finite.

Now let us place a toroidal solenoid inside an infinite cylindrical barrier (Fig. 8). Following the same arguments as above, we conclude that the magnetic scattering amplitude falls off continuously as the height of the potential barrier is increased. This is also true for a toroidal solenoid located in a spherical potential barrier (Fig. 9) or in one section of a toroidal potential barrier (Fig. 10). On the other hand, the magnetic scattering amplitude tends to the finite value (131), (132) for a toroidal solenoid surrounded by a potential barrier in the form of a torus (Fig. 6). Thus, in some cases [a cylindrical solenoid in a cylindrical potential barrier and a toroidal solenoid in a toroidal potential barrier (Fig. 6)] the scattering amplitude tends to a finite value ($\neq 0$) as the height of the barrier is increased. However, in other cases (two cylindrical solenoids with $\Phi_1=-\Phi_2$ in a cylindrical potential barrier and a toroidal solenoid in cylindrical and spherical potential barriers or in one section of a toroidal potential barrier) the scattering amplitude tends to zero as the barrier height is increased. This difference is difficult to explain if the AB effect is interpreted as scattering on regions of space with $H\neq 0$. In fact, the probability that the particles penetrate the region with $H\neq 0$ is roughly the same for barriers of different geometrical shapes and tends to zero for $V_0\to\infty$. On the other hand, the fact that the magnetic amplitude is nonzero can naturally be interpreted either in terms of a nonintegrable phase factor (a finite contribution of the magnetic field to the scattering amplitude), if after the limit $V_0\to\infty$ there remain closed paths accessible to the particles along which $\oint A_l dl\neq 0$, or in terms of a generating function (the potential barrier does not completely screen the region of discontinuity of the generating function).

## 10. THE AB EFFECT FOR BOUND STATES

Let charged particles move between two impenetrable cylinders of radii $a$ and $b$. A solenoid with magnetic flux $\Phi$ is located inside the smaller cylinder. The Schrödinger equation has the form

$$-\frac{\hbar^{2}}{2\mu}\left(\nabla-\frac{ie}{\hbar c}\mathbf{A}\right)^{2}\Psi=E\Psi,\quad \mathbf{A}=\frac{\Phi}{2\pi\rho}\mathbf{e}_{\varphi}. \qquad (141)$$

The single-valued wave function corresponding to angular momentum $m$ and vanishing at $\rho=a$ and $\rho=b$ has the form

$$\Psi_{m}^{\gamma}=C_{m\gamma}[J_{|m-\gamma|}(k\rho)N_{|m-\gamma|}(ka)$$

$$-J_{|m-\gamma|}(ka)N_{|m-\gamma|}(k\rho)]e^{\mathrm{i}m\varphi}. \qquad (142)$$

Here $\gamma=e\Phi/hc$ and $C_{m\gamma}$ is an unimportant normalization factor. The eigenvalues are determined from the relation

$$J_{|m-\gamma|}(kb)N_{|m-\gamma|}(ka)-J_{|m-\gamma|}(ka)\,N_{|m-\gamma|}(kb)=0.$$

From this it follows that the eigenvalues depend explicitly on $\gamma$, i.e., on the magnetic flux $\Phi$. For $ka\gg 1$ and $kb\gg 1$ the eigenvalues can be found explicitly:[44]

$$k_{mn}(\gamma)=\frac{\pi n}{b-a}+\frac{4(m-\gamma)^{2}-1}{8\pi ab}\,\frac{b-a}{n}$$

The following non-single-valued wave function also satisfies the Schrödinger equation (141):

$$\Psi_m^0 = C_m[J_{|m|}(k\rho)N_{|m|}(ka) - N_{|m|}(k\rho)J_{|m|}(ka)]e^{i(m+\gamma)\varphi}. \quad (143)$$

In this case the eigenvalues are the same as in the absence of the magnetic field, i.e., for $\Phi = 0$:

$$J_{|m|}(kb)\, N_{|m|}(ka) - J_{|m|}(ka)\, N_{|m|}(kb) = 0.$$

Therefore, when non-single-valued wave functions are used the presence of the magnetic field does not lead to observable consequences. In fact, the choice of wave functions (141) or (143) has in recent years been the subject of lively discussion on the existence of the AB effect.

The following interesting problem was studied in Refs. 44–47. Let there be a constant magnetic field $H$ directed along the $z$ axis. It corresponds to the following vector potential:

$$\mathbf{A} = \frac{1}{2}(H \times \rho) = -\frac{1}{2} H\rho \mathbf{e}_\varphi.$$

The wave functions and eigenvalues of the Schrödinger equation with this vector potential are

$$\Psi^0_{N,m} \sim \rho^{|m|} L_N^{|m|}\left(\frac{\rho^2}{2}\right) \exp\left(-\frac{\rho^2}{4}\right) \exp(im\varphi),$$
$$E_0(N, m) = \frac{1}{2}(2N + 1 + |m| - m) \quad (144)$$

($E$ is in units of $\hbar\omega$ and $\rho$ is in units of $\hbar/\sqrt{m\omega}$; $\omega = eH/mc$). Now we place at the origin an infinitely thin cylindrical solenoid with symmetry axis along the $z$ axis. Then

$$\Psi^\gamma_{N,m} \sim \rho^{|m-\gamma|} L_N^{|m-\gamma|}\left(\frac{\rho^2}{2}\right) \exp\left(-\frac{\rho^2}{4}\right) \exp(im\varphi),$$
$$E_\gamma(N, m) \sim \frac{1}{2}(2N + 1 + |m - \gamma| + \gamma - m). \quad (145)$$

For definiteness let $0 < \gamma < 1$. Then for $m > 1$ the energy spectrum is the same in the presence or absence of the solenoid. For $m \leqslant 0$ the spectra are different:

$$E_0 = N - m + 1/2, \quad E_\gamma = N - m + \gamma + 1/2.$$

The essential difference between the wave functions (144) and (145) should be noted. For example, the dipole matrix element in the absence of the solenoid

$$\langle \Psi^0_{N',m+1} | x + iy | \Psi^0_{Nm} \rangle$$

is nonzero only for $N' = N, N - 1$. The same matrix element in the presence of the solenoid ($m = 0$)

$$\langle \Psi^\gamma_{N',1} | x + iy | \Psi^\gamma_{N0} \rangle$$

is nonzero for all $N'$. This implies that, in addition to the transitions related to the cyclotron resonances ($N \to N$, $N \to N + 1$) there arises an infinite number of transitions due to the difference of $\gamma$ from zero. This effect can be checked experimentally. In a time-varying electric field, transitions should be observed between states $\Psi^\gamma_{N,0}$, $\Psi^\gamma_{N',1}$.

## 11. EXPERIMENTAL VERIFICATION OF THE AHARONOV–BOHM EFFECT

We deliberately leave aside the early experiments on the observation of the AB effect carried out using a single cylindrical solenoid. The presence of magnetic-field leakages due to the finite length of the solenoid and to imperfections in its winding gave rise to justified criticism and led to uncertainties in the interpretation of the experimental results.[48]

At the present time the most complete experiments on the detection of the magnetic AB effect are those of Tonomura *et al.* with a magnetic field of a toroidal configuration.[49] In electron scattering a shift of the interference pattern has been observed as the magnetic field inside the toroid is changed. In order to prevent the penetration of the electrons into the region with $H \neq 0$, this region was screened by a superconducting layer and a layer of copper. The magnetic-field leakages were controlled by a magnetic interferometer. They turned out to be extremely small. Nevertheless, it has been suggested in the physics literature that they might simulate the AB effect. In Sec. 9 we showed that this is extremely unlikely. Another experiment proposed in Ref. 38 would essentially dispel these doubts about the existence of the AB effect. This experiment is based on the following idea. Let us return to expressions (131) and (132), which determine the contribution $f_m$ of the magnetic field to the total scattering amplitude $f$. The latter is the sum of the scattering amplitude in the absence of the magnetic field $f_0$ and $f_m$. From (131) and (132) it follows that $f_m$ depends on the radius of the hole in the torus $a$ and on the magnetic flux $\Phi$. The amplitude $f_0$ depends only on the parameters of the torus containing the solenoid (i.e., on $a$ and the radius of the torus cross section $R$). Then for fixed $a$, $R$, and $\Phi$ the parameters of the solenoid can be varied without changing $f$. This is illustrated in Fig. 14, where we show two configurations which are completely equivalent from the usual quantum-mechanical point of view [within the framework of which expressions (131) and (132) were obtained]. This shows that the AB effect arises from scattering on regions of space with $\mathbf{H} = 0$, $\mathbf{A} \neq 0$. On the other hand, these two configurations are different from the viewpoint of the alternative interpretation of the AB effect (see Sec. 9). In fact, the probability that the incident particles penetrate into the region with $H \neq 0$ is significantly larger for the configuration shown on the left-hand side of Fig. 14 than for the configuration on the right-hand side.

It is also interesting to study experimentally the scattering of charged particles on a toroidal solenoid located in an impenetrable sphere (Fig. 9) or in one section of an impenetrable torus (Fig. 10). The usual quantum-mechanical treatment predicts that the contribution of the magnetic field to the scattering amplitude is zero in both cases. On the other hand, this contribution is nonzero in the alternative interpretation of the AB effect.

Let a winding of wire of radius $b$ loop around a cylindrical solenoid of radius $R < b$. Then the resistance of the winding (in spite of the fact that it is entirely located in the region where $H = 0$) is a periodic function of the flux $\Phi$ of the magnetic field $H$ inside the solenoids. This effect has been observed experimentally.[50] Therefore, one can speak of a

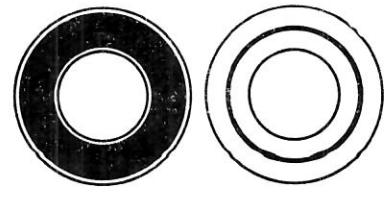

FIG. 14. Two configurations of a toroidal magnetic field inside a toroidal potential barrier, which are identical from the viewpoint of ordinary quantum mechanics, but differ in the alternative view of the AB effect.

force action of the vector potential. For example, for varying **A** we can make a switching device, incorporated in a circuit of the winding with a current, located in the region with $H = 0$. An excellent review of these questions can be found in Ref. 51.

Another experiment to verify the possibility of observing noninteger values of the angular momentum was proposed in Ref. 52. A cylindrical solenoid is surrounded by a superconducting cylinder. Then, allowing for the possibility of noninteger quantization of the angular momentum (i.e., for non-single-valued wave functions), in the external cylinder there should appear a magnetic flux

$$\Phi = \Phi_0 + \frac{1}{2}\frac{neH}{c}. \tag{146}$$

Here $\Phi_0$ is the magnetic flux inside the solenoid. The use of single-valued wave functions in the superconducting cylinder leads to a magnetic flux

$$\Phi' = \frac{1}{2}\frac{neH}{c}. \tag{147}$$

The difference between (146) and (147) can be verified experimentally.

*Crucial experiments to test the existence of the AB effect.* The following experiment would essentially unambiguously confirm or deny the existence of the AB effect.[53] It is based on the following theoretical prediction: under certain conditions the presence of fields inaccessible to particles (for example, the magnetic field $H$ inside an impenetrable solenoid) can lead to a decrease of the density and the probability current in the regions of space accessible to the particles. For example, let us consider an infinite cylindrical solenoid of radius $a$ surrounded by an impenetrable cylinder of radius $R$. In the absence of a current in the solenoid, the wave function $\Psi_0$ is determined by (1). At large distances $\Psi_0$ differs from the incident plane wave $\exp(ikx)$ by terms of order $(R/\rho)^{1/2}$. This is true of both the observables $\bar{\Psi}\Psi$ and $j = (i\hbar/2m)(\Psi\nabla\bar{\Psi} - \bar{\Psi}\nabla\Psi)$. Now we switch on a current in the solenoid. Then the wave function $\Psi$ is given by (2). Aharonov and Bohm showed that for a pointlike, unshielded solenoid located at the origin, with the incident wave propagating along the negative $x$ semiaxis, the wave function vanishes on the positive $x$ semiaxis for $\gamma = \frac{1}{2}$. This remarkable phenomenon also occurs for a finite screened solenoid. In fact, setting $\gamma = \frac{1}{2}$ in (2) and regrouping the terms of the sum (which is possible, owing to its absolute convergence), we find

$$\left.\begin{aligned}
&\Psi_{\frac{1}{2}} = \Psi_{AB}^{1/2} + \Psi_s^{1/2};\\
&\Psi_{AB}^{1/2} = -2\exp\left(\frac{i\varphi}{2} + \frac{i\pi}{4}\right)\\
&\qquad \times \sum_{m=0}^{\infty} i^m J_{m+\frac{1}{2}}(k\rho)\sin\left(m+\frac{1}{2}\right)\varphi;\\
&\Psi_s^{1/2} = 2\exp\left(\frac{i\varphi}{2} + \frac{i\pi}{4}\right)\sum_{m=0}^{\infty} i^m H^{(1)}_{m+\frac{1}{2}}(k\rho)\\
&\qquad \times \frac{J_{m+\frac{1}{2}}(kR)}{H^{(1)}_{m+\frac{1}{2}}(kR)}\sin\left(m+\frac{1}{2}\right)\varphi.
\end{aligned}\right\} \tag{148}$$

From this it follows that $\Psi_{1/2} = \Psi_{AB}^{1/2} = \Psi_s^{1/2} = 0$ on the positive $x$ semiaxis ($\varphi = 0$). For $\rho \to \infty$ we have

$$\left.\begin{aligned}
&\Psi_{AB}^{1/2} = -i\exp\left(ikx + \frac{i\varphi}{2}\right) + \frac{i\exp(i\varphi/2)\exp(ik\rho)}{\left(1 - 2\pi ik\rho\sin^2\frac{\varphi}{2}\right)^{1/2}};\\
&\Psi_s^{1/2} = \frac{\exp(ik\rho)}{\sqrt{\rho}} f_{\frac{1}{2}}(\varphi);\\
&f_{\frac{1}{2}}(\varphi) = 2\sqrt{\frac{2}{\pi ik}}\exp\left(\frac{i\varphi}{2}\right)\\
&\qquad \times \sum_{m=0}^{\infty} \frac{J_{m+\frac{1}{2}}(kR)}{H^{(1)}_{m+\frac{1}{2}}(kR)}\sin\left(m+\frac{1}{2}\right)\varphi.
\end{aligned}\right\} \tag{149}$$

Therefore, the wave function (and also the observables $|\Psi|^2$ and $j$) vanish for $\gamma = \frac{1}{2}$ on the positive $x$ semiaxis behind the solenoid. This suggests the following experimental setup (Fig. 15). We place a detector $D$ behind the solenoid. In the absence of a current in the solenoid, owing to diffraction of the incident particles there is a finite probability that the particles reach the detector $D$. For $\gamma = \frac{1}{2}$ the counting rate of an ideal detector $D$ must fall to zero. A real detector has a finite size. Let us discuss the resulting complications. First of all, we note that the largest distortion of the wave function by the magnetic field occurs for small $kR$. For this it is sufficient to compare the wave function and the scattering amplitude for zero magnetic field:

$$\Psi_s^0 \approx -H_0^{(1)}(k\rho)\left[1 + \frac{2i}{\pi}\left(\ln\frac{kR}{2} + C\right)\right]^{-1}, \quad C = 0{,}577\ldots;$$

$$f_0(\varphi) = -\sqrt{\frac{2}{\pi ik}}\left[1 + \frac{2i}{\pi}\left(\ln\frac{kR}{2} + C\right)\right]^{-1}$$

and for $\gamma (\equiv e\Phi/hc) = \frac{1}{2}$:

$$\Psi_s^{1/2} = -2\exp\left(\frac{i\varphi}{2} - \frac{i\pi}{4}\right)kRH_1^{(1)}(k\rho)\sin(\varphi/2);$$

$$f_{\frac{1}{2}}(\varphi) = 2\sqrt{\frac{2}{\pi ik}}\exp\left(\frac{i\varphi}{2} + \frac{i\pi}{4}\right)kR\sin(\varphi/2).$$

At high energies we cannot use the representation of the wave function in the form of series (1), (2), (148), (149), since they contain many terms of alternating sign. In this case the Kirchhoff approximation is suitable. For the part $\Psi_s^0$ of the wave function describing the scattered wave $[\Psi_0 = \exp(ikx) + \Psi_s^0]$ we find

$$\Psi_s^0(x, y) = -\frac{k}{4}\int_{-R}^{R} dy'\left[H_0^{(1)}(kt) + \frac{ix}{t}H_1^{(1)}(kt)\right],$$

$$t = \sqrt{x^2 + (y - y')^2}.$$

For $k\rho \to \infty$ we have

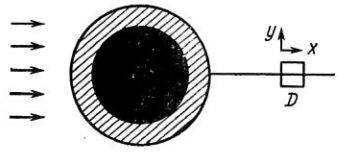


FIG. 15. A cylindrical solenoid (blackened region) surrounded by an impenetrable cylinder (shaded region). For the initial wave vector along the $x$ axis and $\gamma = \frac{1}{2}$ the counting rate of the detector $D$ located on the $x$ axis behind the cylinder is equal to zero.

$$\Psi_s^0 = -\left(\frac{k}{8\pi i}\right)^{1/2} \int_{-R}^{R} dy' \frac{\exp(ikt)}{\sqrt{t}}(1 + x/t). \qquad (150)$$

For $k \to \infty$ and fixed $R$, $\rho$ the integrand oscillates rapidly. For $|y| < R$ the value of the integral in (150) is determined by the contribution of the stationary point $y' = y$. Using the stationary-phase method, we obtain

$$\Psi_s^0 \approx -\exp(ikx) \quad (k \to \infty,\ |y| < R,\ x > 0).$$

Substituting this expression into $\Psi_0$, we find $\Psi_0 \approx 0$, i.e., a geometrical shadow appears behind the solenoid. For $\rho \gg R$, from (150) we find

$$\Psi_s^0 = -\sqrt{\frac{k}{8\pi i\rho}}(1 + \cos\varphi) \int_{-R}^{R} \exp(ikt)\, dy'. \qquad (151)$$

For $\rho \to \infty$ and fixed $k$ we have

$$\left.\begin{aligned} \Psi_s^0 &= \frac{\exp(ik\rho)}{\sqrt{\rho}} f_0(\varphi); \\ f_0(\varphi) &= -\frac{1}{\sqrt{2\pi i k}} \frac{\sin(kR\sin\varphi)}{\sin\varphi}(1 + \cos\varphi); \\ \sigma_0(\varphi) &= \frac{1}{2\pi k}\left[\frac{\sin(kR\sin\varphi)}{\sin\varphi}(1 + \cos\varphi)\right]^2. \end{aligned}\right\} \qquad (152)$$

The Kirchhoff approximation works (see, for example, Ref. 34) if the wavelengths of the incident particles are small compared with the size of the obstacles on which the scattering occurs (in the present case this corresponds to the condition $kR \gg 1$). The fields calculated in this approximation differ from the exact ones only very close to the scatterer. Numerical calculations[54] show that the Kirchhoff approximation is satisfactory for $kR \approx 2$–3. In typical experiments[55] on electron scattering on a cylindrical solenoid, the electron energy is $E \approx 20$ keV, and $R \approx 10^{-4}$ cm. This makes $k = 10^{10}$ cm$^{-1}$ and $kR \sim 10^6$. Therefore, the Kirchhoff approximation is certainly applicable.

For finite values of $kR$ and $k\rho \gg 1$ the wave function and the scattering cross section in the absence of a magnetic field oscillate with period $\pi/kR$ [see (152)]. The amplitude of the oscillations is particularly large for small values of $\varphi$, i.e., in the region behind the solenoid. The introduction of an impenetrable partition coinciding with the part of the plane $y = 0$ extending from $x = R$ to $x = \infty$ (which is equivalent to the presence inside an impenetrable cylinder of a magnetic flux $\Phi$ satisfying the condition $\gamma = e\Phi/hc = \frac{1}{2}$) drastically changes the diffraction theory. Owing to the oscillations, there is a sharp decrease of the wave function only very close to the plane $y = 0$. Let us now calculate $|\Psi|^2$ for finite values of $kR$ and $k\rho$. For this we use the wave functions $\Psi_0$ and $\Psi_{1/2}$ given by expressions (1) and (148). In Figs. 16 and 17 we show the results of the calculations of Ref. 56 for $kR = 1$ and $kR = 10$ for the values of $k\rho$ shown in the figures. For each value of $k\rho$ the region of the geometrical shadow ($0 \leqslant \varphi \leqslant \arcsin(R/\rho)$) was divided into 40 equal parts. Since the absolute values of the wave functions for $\gamma = 0$ and $\gamma = \frac{1}{2}$ do not change when the sign of $\varphi$ is changed, it is sufficient to calculate the wave functions for $\varphi > 0$ at each of the 21 points $\varphi_n = ((n-1)/21)\arcsin(R/\rho)$ $(1 \leqslant n \leqslant 21)$. Expression (148) was used to calculate the probability density $|\Psi_{1/2}(\rho, \varphi = \varphi_n)|^2$ for each $\varphi_n$. For convenience, it was related to the square of the absolute value of the wave function in the absence of a magnetic field, taken on the $x$ axis at the same value of $\rho$. Therefore, on the vertical axes in Figs. 16 and 17 we give the ratio $r = |\Psi_{1/2}(\rho, \varphi = \varphi_n)|^2 / |\Psi_0(\rho, \varphi = 0)|^2$, and on the horizontal axes we give the zone number $n$ for which the angle $\varphi_n$ is calculated. As expected, oscillations appear for increasing $R$ and $\rho$. In addition, the region where $|\Psi_{1/2}|^2$ is small narrows. For $kR \gg 1$ and $k\rho \gg 1$, which corresponds to the experimental conditions (see Ref. 55), the region behind the solenoid lies in the geometrical shadow. In it the wave function is very small. Therefore, the introduction of the aforementioned impenetrable partition, lying in the plane $y = 0$ and extending from $x = R$ to $x = \infty$ (which, as we have seen, is equivalent to the establishment inside the solenoid of a magnetic flux $\Phi$ with $\gamma = e\Phi/hc = \frac{1}{2}$), changes practically nothing in this region. In both cases, i.e., for $\gamma = 0$ and $\gamma = \frac{1}{2}$, the counting rate of detectors located behind the solenoid will be extremely small. At first glance it might seem that under these conditions it is senseless to carry out the experiment proposed in Ref. 53 to verify the existence of the AB effect. However, this is not so. The reason is that the experimental setups[55] make extensive use of negatively charged electrostatic systems (so-called Fresnel biprisms), which deflect the electron beam toward the $x$ axis. This effectively leads to a decrease of the wave number $k$ (Ref. 57). As a result, an interference pattern appears in the shadow region. For symmetric placement of the biprisms relative to the plane $y = 0$ the wave function again vanishes in the plane $y = 0$ $(R < x < \infty)$. This fact can be checked experimentally.

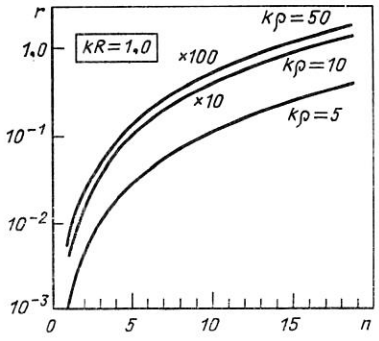


FIG. 16. A numerical illustration of the effect of the magnetic field on the particle detection probability. The shadow region was divided into 40 equal parts. At each of these points the square of the absolute value of the wave function for $\gamma = \frac{1}{2}$ is related to the square of the absolute value of the wave function in the absence of a magnetic field taken on the $x$ axis for the same value of $\rho$. Here $kR = 1$. The values of $k\rho$ are given for each curve.

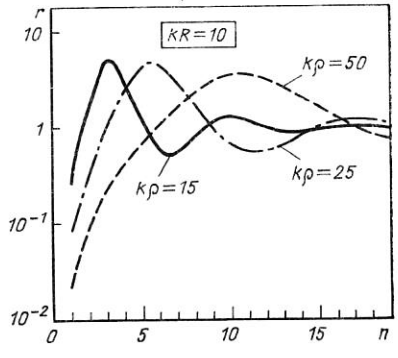


FIG. 17. The same as in Fig. 16 but for $kR = 10$.

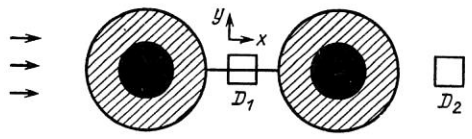


FIG. 18. Two cylindrical solenoids with $\Phi_1 = -\Phi_2 = \Phi$ surrounded by impenetrable cylindrical barriers (shaded regions). The incident wave propagates along the $x$ axis. For $\gamma = e\Phi/he = \frac{1}{2}$ the wave function vanishes on the segment of the $x$ axis lying between the cylinders. Therefore, the counting rate of the detector $D_1$ falls to zero.

Earlier we saw that for $\gamma = \frac{1}{2}$ the wave function vanishes between two solenoids with $\Phi_1 = -\Phi_2$ and in the equatorial plane ($z = 0$) of a toroidal solenoid $(\rho - d)^2 + z^2 = R^2$. From this fact it follows that particle detectors ($D_1$; see Figs. 6 and 18) located at the positions where the wave function vanishes must cease to detect particles for $\gamma = \frac{1}{2}$. On the other hand, the counting rate of detectors placed behind two solenoids with $\Phi_1 = -\Phi_2$ and behind a toroidal solenoid ($D_2$; see Figs. 6 and 18) is roughly the same for $\gamma = 0$ and $\gamma = \frac{1}{2}$.

*Fraunhofer diffraction on two cylindrical solenoids and a toroidal solenoid.* Above, we found the amplitudes for the scattering of charged particles on the magnetic field of two cylinders with $\Phi_1 = -\Phi_2$ (Fig. 2) and on the magnetic field of a toroidal solenoid (Fig. 5). However, it is not the cross section for scattering $\sigma_m = |f_m|^2$ on the magnetic field of the solenoid that is measured experimentally, but the shift of the diffraction pattern arising from the magnetic flux inside the solenoid. For this it is necessary to find the scattering cross section in the absence ($\sigma_0 = |f_0|^2$) and the presence of a magnetic field inside the solenoid ($\sigma_\gamma = |f_0 + f_m|^2$), and to compare them with each other and with the experimentally observed shift of the diffraction pattern.

In the Kirchhoff approximation the amplitude and the differential cross section for scattering on two impenetrable cylinders of radius $R$ [the cylinder axes pass through the points $\pm d$ ($d > R$) of the $y$ axis parallel to the $z$ axis, and the initial wave vector points along the $x$ axis is; Fig. 2] are[56]

$$\left.\begin{aligned} f_0^{2s} &= -\frac{2}{\sqrt{2\pi i k}}\,\frac{1+\cos\varphi}{\sin\varphi}\sin(kR\sin\varphi)\cos(kd\sin\varphi);\\ \sigma_0^{2s} &= \frac{2}{\pi k}\left[\frac{1+\cos\varphi}{\sin\varphi}\sin(kR\sin\varphi)\cos(kd\sin\varphi)\right]^2 \end{aligned}\right\} \tag{153}$$

Similarly, for the cross section for scattering on an impenetrable torus $(\rho - d)^2 + z^2 = R^2$ (with the initial wave vector along the $z$ axis; Fig. 5) we have

$$\begin{aligned} f_0^t &= \frac{i}{2}\,\frac{1+\cos\theta}{\sin\theta}\{(d+R)\,J_1[k(d+R)\sin\theta]\\ &\quad -(d-R)\,J_1[k(d-R)\sin\theta]\};\\ \sigma_0^t &= \frac{1}{4}\left(\frac{1+\cos\theta}{\sin\theta}\right)^2\{(d+R)\,J_1[k(d+R)\sin\theta]\\ &\quad -(d-R)\,J_1[k(d-R)\sin\theta]\}^2. \end{aligned} \tag{154}$$

The total amplitude for scattering on a solenoid surrounded by an impenetrable barrier is the sum of the scattering amplitude $f_0$ in the absence of a magnetic field and the amplitude $f_m$ for scattering on the magnetic field outside the impenetrable area. For $f_m$ we can use the scattering amplitudes obtained in Secs. 5 [see (97), (102)] and 6 [see (131), (132)], or the amplitudes found in Ref. 56 using the Dirac phase factor:[12,13]

$$f_m^D = -\frac{1}{\sqrt{2\pi i k}}[1-\exp(2i\pi\gamma)]\,\frac{1+\cos\varphi}{\sin\varphi}\sin[k(d-R)\sin\varphi]$$

for two cylinders with $\Phi_1 = -\Phi_2$ and

$$f_m^D = \frac{i}{2}[1-\exp(2i\pi\gamma)]\,(d-R)\,\frac{1+\cos\theta}{\sin\theta}\,J_1[k(d-R)\sin\theta]$$

for the toroidal solenoid.

For the total scattering amplitudes and cross sections we obtain

$$\begin{aligned} f_\gamma^{2s} &= -\frac{1}{\sqrt{2\pi i k}}\,\frac{1+\cos\varphi}{\sin\varphi}\{\sin[k(d+R)\sin\varphi]\\ &\quad -\exp(2i\pi\gamma)\sin[k(d-R)\sin\varphi]\},\\ \sigma_\gamma^{2s} &= \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2[\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\sin^2\pi\gamma\\ &\quad +\sin^2(kR\sin\varphi)\cos^2(kd\sin\varphi)\cos^2\pi\gamma] \end{aligned} \tag{155}$$

for two cylinders with $\Phi_1 = -\Phi_2$ and

$$\begin{aligned} f_\gamma^t &= \frac{i}{2}\,\frac{1+\cos\theta}{\sin\theta}\{(d+R)\,J_1[k(d+R)\sin\theta]\\ &\quad -\exp(2i\pi\gamma)\,(d-R)\,J_1[k(d-R)\sin\theta]\}, \end{aligned} \tag{156}$$

$$\begin{aligned} \sigma_\gamma^t &= \sigma_0^t + \left(\frac{1+\cos\theta}{\sin\theta}\right)^2(d^2-R^2)\sin^2\pi\gamma J_1[k(d+R)\sin\theta]\\ &\quad \times J_1[k(d-R)\sin\theta] \end{aligned}$$

for the toroidal solenoid. The cross section $\sigma_0^t$ is determined from (154). We shall be especially interested in the case $\gamma = \frac{1}{2}$. Then

$$\begin{aligned} \sigma_{\frac{1}{2}}^{2s} &= \frac{2}{\pi k}\left[\frac{1+\cos\varphi}{\sin\varphi}\sin(kd\sin\varphi)\cos(kR\sin\varphi)\right]^2,\\ \sigma_{\frac{1}{2}}^t &= \frac{1}{4}\left(\frac{1+\cos\theta}{\sin\theta}\right)^2\{(d+R)\,J_1[k(d+R)\sin\theta]\\ &\quad +(d-R)\,J_1[k(d-R)\sin\theta]\}^2. \end{aligned} \tag{157}$$

Let us now see what are the experimental consequences of electron scattering on two solenoids with $\Phi_1 = -\Phi_2$. From (153) it follows that in the absence of a magnetic field ($\gamma = 0$) the cross section for scattering on two cylinders (Fig. 2) has two families of zeros at angles given by the relations

$$\sin\varphi_n^{(R)} = \frac{h\pi}{kR}, \quad \sin\varphi_n^{(d)} = \frac{h+\frac{1}{2}}{kd}\,\pi,$$

or (for angles which are not too large)

$$\varphi_n^{(R)} = \frac{n\pi}{kR}, \quad \varphi_n^{(d)} = \frac{n+\frac{1}{2}}{kd}\,\pi. \tag{158}$$

On the other hand, for $\gamma = \frac{1}{2}$ [see (157)]

$$\varphi_n^{(R)} = \frac{n+\frac{1}{2}}{kR}\,\pi, \quad \varphi_n^{(d)} = \frac{n\pi}{kd}\,\pi. \tag{159}$$

Comparing (158) and (159), we conclude that the switching-on of the magnetic field leads (for $\gamma = \frac{1}{2}$) to a shift of the zeros of the first family by an amount $\pi/2kR$, and of those of the second by an amount $\pi/2kd$. However, it is the maxima of the cross sections, not the minima, which are observed experimentally. In Fig. 19 we show the cross sections calculated for the following parameters: $k = 2\times10^{10}$ cm$^{-1}$,

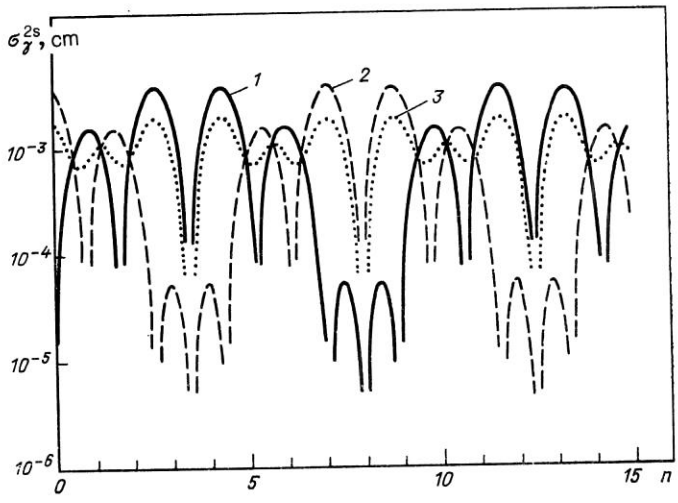


FIG. 19. Angular dependence of the cross section for electron scattering on two solenoids with $\Phi_1 = -\Phi_2$. In the Fraunhofer theory of diffraction the solid line corresponds to the value $\gamma = 0$, the broken line to $\gamma = 0.5$, and the dotted line to $\gamma = 0.25$. The situation changes drastically when the appropriate Fresnel theory is used. If the distance from the origin to the detector (see Fig. 1) is 106.1, the Fresnel diffraction theory exactly reproduces the results of the Fraunhofer theory with the same values of $\gamma$. For $\rho \approx 91$, curve 1 corresponds to $\gamma = 0.5$, curve 2 corresponds to $\gamma = 0$, and curve 3 corresponds to $\gamma = 0.25$. For $\rho = 98$ the values $\gamma = 0$ and $\gamma = 0{,}5$ correspond to the same curve 3.

$R = 10^{-4}$ cm, and $d = 5\times 10^{-4}$ cm. On the horizontal axis we plot the integers $n$ related to the scattering angle as $\varphi = (0.01 + 10^{-5}\cdot n)$ deg. For example, the number 5 correspond to the angle $\varphi = (0.01 + 5\times 10^{-5})$ deg. On the vertical axis we plot the cross section $\sigma_\gamma^{2s}$ [see (155)] in centimeters. The cross sections were calculated for $\gamma = 0$ (curve 1), $\gamma = \frac{1}{2}$ (curve 2), and $\gamma = \frac{1}{4}$ (curve 3). For greater clarity, in the upper part of Fig. 20 we show the locations of the maxima of the cross sections and their values. The maxima themselves are shown by vertical lines of length equal to the value of the cross section at the maximum. The solid lines correspond to $\gamma = 0$, and the broken ones to $\gamma = 0.5$.

Let us now turn to the toroidal solenoid. Since in real experiments (which will be discussed below) $kd \gg 1$ and $kR \gg 1$, for scattering angles which are not too small the Bessel functions in the cross sections (154) and (157) can be replaced by their asymptotic expressions:

$$\sigma_0' = \frac{1}{2\pi k}\frac{(1+\cos\theta)^2}{\sin^3\theta}\left\{\sqrt{d+R}\sin\left[k(d+R)\sin\theta - \frac{\pi}{4}\right] - \sqrt{d-R}\sin\left[k(d-R)\sin\theta - \frac{\pi}{4}\right]\right\}^2;$$

$$\sigma_1' = \frac{1}{2\pi k}\frac{(1+\cos\theta)^2}{\sin^3\theta}\left\{\sqrt{d+R}\sin\left[k(d+R)\sin\theta - \frac{\pi}{4}\right] + \sqrt{d-R}\sin\left[k(d-R)\sin\theta - \frac{\pi}{4}\right]\right\}^2.$$

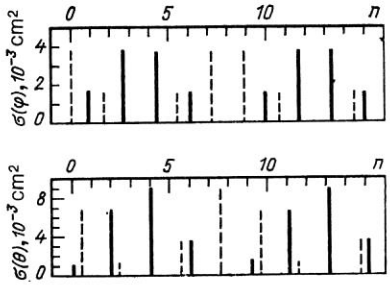


FIG. 20. Positions and values of the maxima in the cross sections for scattering on two cylindrical solenoids with $\Phi_1 = -\Phi_2$ (upper part of the figure) and on a toroidal solenoid (lower part): the solid vertical lines correspond to $\gamma = 0$ and the dashed ones to $\gamma = 0.5$. Only those maxima whose values exceed $10^{-3}$ are shown.

If $R \ll d$, we can neglect $R$ in the square root (but not in the argument of the sine, since $kR \gg 1$):

$$\sigma_0' = \frac{2d}{\pi k}\left[(1+\cos\theta)\sin(kR\sin\theta)\cos\left(kd\sin\theta - \frac{\pi}{4}\right)\right]^2 / \sin^3\theta;$$

$$\sigma_1' = \frac{2d}{\pi k}\left[(1+\cos\theta)\cos(kR\sin\theta)\sin\left(kd\sin\theta - \frac{\pi}{4}\right)\right]^2 / \sin^3\theta.$$

As for the two cylinders with $\Phi_1 = -\Phi_2$, we have two families of zeros of the scattering cross section:

$$\theta_n^{(R)} = \frac{n\pi}{kR},\quad \theta_n^{(d)} = \frac{n-\frac{1}{4}}{kd}\pi \quad (\gamma = 0); \tag{160}$$

$$\theta_n^{(R)} = \frac{n-\frac{1}{2}}{kR}\pi,\quad \theta_n^{(d)} = \frac{n+\frac{1}{4}}{kd}\pi \left(\gamma = \frac{1}{2}\right). \tag{161}$$

When the magnetic field is switched on, the zeros of the first family shift by $\Delta\theta_R = \pi/2kR$, and those of the second shift by $\Delta\theta_d = \pi/2kd$. The well-known experiments of Tonomura *et al.*[49] were carried out for the following parameters: electron energy $E \approx 150$ keV, $R = 10^{-4}$ cm, and $d = 4\times 10^{-4}$ cm. Then $k = 2\times 10^{10}$ cm$^{-1}$, $kR = 2\times 10^6$, and $kd = 8\times 10^6$. This gives $\Delta\theta_R = 8\times 10^{-7}$ and $\Delta\theta_d = 2\times 10^{-7}$. In these experiments a shift of the interference pattern is observed in the plane perpendicular to the wave vector of the incident wave (i.e., in the plane $z = \text{const}$). Unfortunately, the authors of Ref. 49 do not give the distances from the scatterer (the toroidal solenoid) to the detector. Judging from the sketch of the setup given in that study, we take it to be $r = 1$m. Then for the shift of the levels of the first family we have $\Delta z_R = r\Delta\theta_R \approx 0.8\,\mu$m, and for the second $\Delta z_d = r\Delta\theta_d \approx 0.2\,\mu$m. These are approximately the shifts observed in the interferograms given in Ref. 49. The typical angular dependence of the cross section is shown in Fig. 21. The cross sections were calculated for the parameters of the Tonomura experiment. The origin corresponds to the angle $\varphi = 0.01°$, and the step size along the horizontal axis is $10^{-5}$. On the vertical axis we give the cross sections calculated using Eqs. (154) and (157). The solid line corresponds to $\gamma = 0$ and the broken line to $\gamma = 0.5$. As for scat-

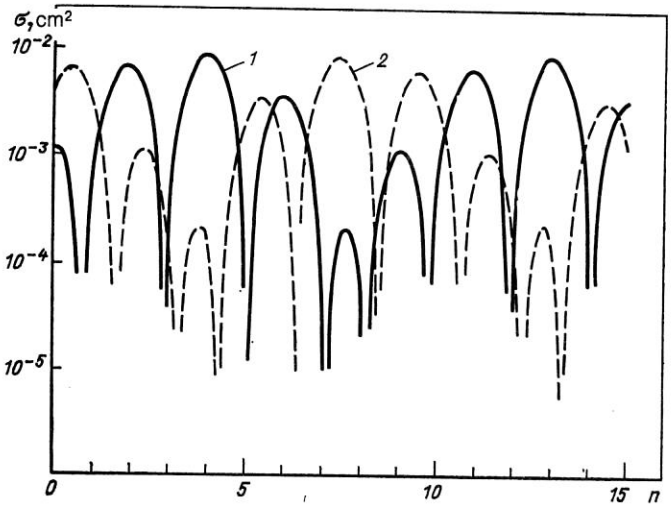


FIG. 21. Angular cross sections for electron scattering on a toroidal solenoid, calculated using the Fraunhofer diffraction theory: 1) $\gamma = 0$; 2) $\gamma = 0.5$. The horizontal scale is the same as in Fig. 19.

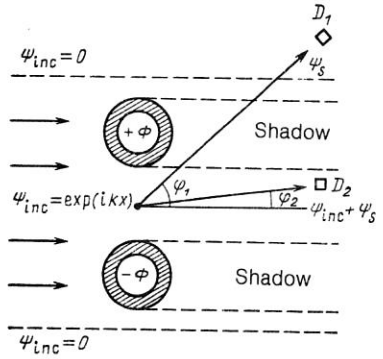


FIG. 22. Measurement of the differential cross sections using detectors under different physical conditions.

tering on two cylinders, the shift of the diffraction pattern when the magnetic field is switched on is clearly seen. For greater clarity, in the lower part of Fig. 20 we show the positions of the maxima and the values of the cross sections at them.

*Fresnel diffraction on two cylindrical solenoids.* The representation of the scattered wave as the product of an outgoing wave ($\exp(ik\rho)/\sqrt{\rho}$ for the two-dimensional case and $\exp(ikr)/r$ for the three-dimensional one) and a function depending only on angles is valid only at sufficiently large distances from the scatterer. Here along with the conditions $kr \gg 1$ and $R/r \ll 1$ (which we shall always assume to be valid) we must also have $kR^2/r \ll \pi$ (for a single cylindrical solenoid) and $kdR/r \ll \pi$ (for two cylindrical solenoids with $\Phi_1 = -\Phi_2$ and for a toroidal solenoid). In optics the scattering of light under such conditions is referred to as Fraunhofer diffraction.[34] In the experiments carried out on a single cylindrical solenoid[55] we have $kR^2/\rho \approx 5$. As we have already mentioned, the distance from the toroidal solenoid to the detector is not given in the studies by Tonomura *et al.*[49] For definiteness we take it to be 1m, as above. Then $kdR/r \approx 8$. In view of this, the conclusions drawn above on the shift of the interference pattern are qualitative in nature. Since $kdR/r$ is not small compared with $\pi$, the contribution of these terms must be included in the wave function. In optical language this corresponds to Fresnel diffraction.[34] The higher-order terms ($kR^3/\rho^2 \approx 3\times 10^{-5}$ for a single solenoid and $kdR^2/r^2 \approx 8\times 10^{-6}$ for two solenoids with $\Phi_1 = -\Phi_2$ and for a toroidal solenoid) are small and can be neglected. Let us show how the scattering cross sections vary for the example of two cylinders with $\Phi_1 = -\Phi_2$. When the conditions $kd \gg 1$, $p \gg d$, and $kdR^2/\rho^2 \ll 1$ are satisfied, the wave function corresponding to Fig. 2 is

$$\Psi_\gamma = \exp(ikx) + \Psi_s, \tag{162}$$

$$\Psi_s = -\frac{1}{2\sqrt{2i}}\frac{1+\cos\varphi}{\cos\varphi}\exp\left[ik\rho\left(1-\frac{1}{2}\operatorname{tg}^2\varphi\right)\right](A+iB);$$

$$A = C_1 + C_3 - \cos 2\pi\gamma\,(C_2+C_4) + \sin 2\pi\gamma\,(S_2+S_4);$$

$$B = S_1 + S_3 - \cos 2\pi\gamma\,(S_2+S_4) - \sin 2\pi\gamma\,(C_2+C_4). \tag{163}$$

Here $C_i = C(\rho_i)$ and $S_i = S(\rho_i)$; $C(x)$ and $S(x)$ are the usual Fresnel integrals ($C = \int_0^x \cos(\pi x^2/2)dx$, $S = \int_0^x \sin(\pi x^2/2)dx$). The arguments $\rho_i$ are $\rho_1 = \sqrt{k\rho/\pi}(((d+R)/\rho)\cos\varphi + \tan\varphi)$, $\rho_2 = \sqrt{k\rho/\pi}(((d-R)/\rho)\cos\varphi + \tan\varphi)$, $\rho_3 = \sqrt{k\rho/\pi}(((d+R)/\rho)\cos\varphi - \tan\varphi)$, and $\rho_4 = \sqrt{k\rho/\pi}(((d-R)/\rho)\cos\varphi - \tan\varphi)$.

For $|\tan\varphi| \gg (d+R)/\rho$ expression (163) simplifies:

$$\Psi_s = \frac{\exp(ik\rho)}{\sqrt{\rho}} f_\gamma(\rho, \varphi);$$

$$f_\gamma(\rho,\varphi) = -\frac{1}{\sqrt{2\pi ik}}\frac{1+\cos\varphi}{\sin\varphi}\left\{\exp\left[ik\frac{(d+R)^2}{2\rho}\cos^2\varphi\right]\right.$$

$$\times \sin[k(d+R)\sin\varphi] - \exp(2i\pi\gamma)\exp\left[ik\frac{(d-R)^2}{2\rho}\cos^2\varphi\right]$$

$$\left.\times \sin[k(d-R)\sin\varphi]\right\}. \tag{164}$$

If the detector is located outside the incident beam ($D_1$ in Fig. 22), then $|f_\gamma|^2$ (up to terms of order $[(d+R)/\rho]^2$) coincides with the differential flux through a cylindrical surface of radius $\rho$ and is the analog of the differential cross section for finite values of $\rho$:

$$\tilde{\sigma}_\gamma^{2s}(\rho,\varphi) = \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2$$

$$\times[\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\sin^2(\omega-\pi\gamma)$$

$$+\sin^2(kR\sin\varphi)\cos^2(kd\sin\varphi)\cos^2(\omega-\pi\gamma)],$$

$$\omega = \frac{kdR}{\rho}\cos^2\varphi. \tag{165}$$

Obviously, for $\rho \to \infty$ the Fresnel cross section (165) becomes the Fraunhofer cross section (155). We shall be especially interested in the cases $\gamma = 0$ and $\gamma = \frac{1}{2}$:

$$\tilde{\sigma}_0^{2s} = \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2[\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\sin^2\omega$$

$$+\sin^2(kR\sin\varphi)\cos^2(kd\sin\varphi)\cos^2\omega]; \tag{166}$$

$$\tilde{\sigma}_{\frac{1}{2}}^{2s} = \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2[\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\cos^2\omega$$

$$+\sin^2(kR\sin\varphi)\cos^2(kd\sin\varphi)\sin^2\omega]. \tag{167}$$

*Uncertainties and interpretation of the experimental data.* Let us briefly consider the complications which arise if experiments on electron scattering on solenoids are interpreted using expressions (153), (154), (156), and (157), in which the finiteness of the distance from the scatterer to the observation point is not taken into account. As an illustration we again consider electron scattering on two solenoids

with opposite magnetic fluxes. Usually, in experiments one studies a small number of diffraction maxima and observes their shift for different magnetic field strengths inside the solenoids. For definiteness, let the number of observed diffraction bands be 10. If the parameters $k$, $R$, and $d$ are the same as in the Tonomura experiments with the toroidal solenoid, then 10 maxima occur in an angular range [see (158) and (159)] approximately equal to $10(\pi/kd)\approx 2\cdot 10^{-6}$. Furthermore, let $\varphi_0$ be the angle near which the scattered particles are observed. Since $kdR/\rho$ (for $\rho = 1$m) is roughly equal to eight, in expression (155) the angle $\varphi$ in $\sin(\omega - \pi\gamma)$ and $\cos(\omega - \pi\gamma)$ can with good accuracy be replaced by $\varphi_0$. This cannot be done in $kR\sin\varphi$ and $kd\sin\varphi$, since (for $kR \gg 1$ and $kd \gg 1$) they vary significantly over the interval $10\pi/kd$. Then instead of (155) we have

$$\tilde{\sigma}_\gamma^{2s} = \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2 [\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\sin^2(\omega_0 - \pi\gamma) + \sin^2(kR\sin\omega)\cos^2(kd\sin\varphi)\cos^2(\omega_0 - \pi\gamma)],$$

$$\omega_0 = \frac{kdR}{\rho}\cos^2\varphi_0. \qquad (168)$$

Let us consider expression (168) for $\gamma = 0$ and $\gamma = \frac{1}{2}$:

$$\tilde{\sigma}_0^{2s} = \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2 [\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\sin^2\omega_0 + \sin^2(kR\sin\varphi)\cos^2(kd\sin\varphi)\cos^2\omega_0]; \qquad (169)$$

$$\tilde{\sigma}_{\frac{1}{2}}^{2s} = \frac{2}{\pi k}\left(\frac{1+\cos\varphi}{\sin\varphi}\right)^2 [\sin^2(kd\sin\varphi)\cos^2(kR\sin\varphi)\cos^2\omega_0 + \sin^2(kR\sin\varphi)\cos^2(kd\sin\varphi)\sin^2\omega_0]. \qquad (170)$$

Let $\omega_0$ be a multiple of $\pi$. Then $\tilde{\sigma}_0^{2s} = \sigma_0^{2s}$ and $\tilde{\sigma}_{1/2}^{2s} = \sigma_{1/2}^{2s}$. In this case an observer located at a finite distance, using the Fraunhofer expressions (153) and (157), correctly describes the positions of the maxima and minima of the cross section in the presence and absence of a magnetic field. Now let $\omega_0$ be a multiple of $\pi/2$ (rather, equal to $(n+\frac{1}{2})\pi$). Then $\tilde{\sigma}_0^{2s} = \sigma_{1/2}^{2s}$ and $\tilde{\sigma}_{1/2}^{2s} = \sigma_0^{2s}$.

Therefore, an observer who placed the detector at $\rho = \rho_0$, $\varphi = \varphi_0$, measures the cross section for $\gamma = \frac{1}{2}$, and compares it with the theoretical predictions (153) and (157) concludes that when a magnetic flux is present with $\gamma = \frac{1}{2}$ the cross section measured in the solenoids coincides with the theoretical cross section $\sigma_0^{as}$, corresponding to the absence of a magnetic field. From this one might conclude that the presence of a magnetic field in regions of space inaccessible to the particles has no effect on the scattering, so that there is no AB effect. It might be objected that the experimentalist will not use the theoretical cross sections (153) and (157), but will simply measure the diffraction pattern for $\gamma = 0$ and $\gamma = \frac{1}{2}$ and will decide about the existence of the AB effect on the basis of its shift. In fact, for integer or half-integer $\omega_0$ the shift of the diffraction pattern is correctly reproduced by the Fraunhofer theory. However, let $\rho = \rho_0$ and $\varphi_0$ in (169) and (170) be such that $\omega_0 = \pi(n+\frac{1}{4})$. Then $\tilde{\sigma}_0^{2s} = \tilde{\sigma}_{1/2}^{2s}$. This implies that an observer who placed the detector at $\rho = \rho_0$ and $\varphi = \varphi_0$ sees the same diffraction pattern for $\gamma = 0$ and $\gamma = \frac{1}{2}$. On the basis of this he concludes that the AB effect is absent. Let us again return to the cross sections in Fig. 19. We have already pointed out that they correspond to Fraunhofer diffraction, which describes the experimental situation only qualitatively. Using the Fresnel theory of diffraction and taking $\rho = 106.1$ cm and $\varphi_0 = 0.01°$ (with $\omega$ equal to an integer), we exactly reproduce the results of the Fraunhofer theory, with curves 1, 2, 3 corresponding to the values $\gamma = 0, 0.5, 0.25$. For $\rho = 90.95$ (half-integer $\omega$) curve 1 corresponds to $\gamma = 0.5$, curve 2 corresponds to $\gamma = 0$, and curve 3, as before, corresponds to $\gamma = 0.25$. For $\rho \approx 98$ the values $\gamma = 0$ and $\gamma = 0.5$ correspond to the same curve 3.

In summary, we urge that care be taken in the interpretation of the results of experiments on electron scattering on solenoids. It can happen that two different experimentalists, placing their detectors at different distances from the solenoid, arrive at contradictory conclusions about the existence of the AB effect (one fixes the shift of the diffraction pattern, and the other doesn't). In the case of scattering on two solenoids, the second observer must move in the radial direction by an amount $\Delta\rho = (4\omega_0/\pi - 1)^{-1}\rho_0$ in order to obtain the correct (in the sense of the Fraunhofer theory) value of the diffraction shift for $\gamma = 0$ and $\gamma = \frac{1}{2}$.

*Measurement of the intensity in the direct beam.* Let us discuss what the observer sees at small scattering angles. Since the detector ($D_2$ in Fig. 22) is located in the direct beam, the cross section is no longer equal to the square of the absolute value of the coefficient of the outgoing wave. According to the rules of quantum mechanics,[58] it is proportional to the radial component of the probability current:

$$\sigma = \frac{m\rho}{\hbar k} j_\rho, \quad \mathbf{j} = \frac{\hbar}{2im}(\psi\nabla\psi - \psi\nabla\bar{\psi}) - \frac{e}{mc}\mathbf{A}\bar{\psi}\psi. \qquad (171)$$

Substituting in this the wave function $\psi_\gamma$ given by (162) and discarding terms which are small compared with $kR$ (as long as they do not occur in an exponential) in the angular range of interest to us ($0 \leqslant \varphi \leqslant \varphi_0 = (d-R)/\rho$), we obtain

$$\sigma = \rho\left[\cos\varphi + \frac{1}{8}\left(\frac{1+\cos\varphi}{\cos\varphi}\right)^2(A^2+B^2) - \frac{(1+\cos\varphi)^2}{2\sqrt{2}\cos\varphi}(A\cos\Delta + B\sin\Delta)\right] - \frac{e\rho}{\hbar kc}\left[1 + \frac{1}{8}\left(\frac{1+\cos\varphi}{\cos\varphi}\right)^2(A^2+B^2) - \frac{1+\cos\varphi}{\sqrt{2}\cos\varphi}(A\cos\Delta + B\sin\Delta)\right]A_\rho. \qquad (172)$$

Here $A$ and $B$ are given by (163), and

$$\Delta = k\rho\left(\cos\varphi - 1 + \frac{1}{2}\operatorname{tg}^2\varphi\right) + \frac{\pi}{4}.$$

For small $\varphi$ the quantity $\cos\varphi - 1 + \frac{1}{2}\tan^2\varphi$ is roughly equal to $(9/24)\varphi^4$. Taking the maximum value for $\varphi$ ($\varphi_0 \approx (d-R)/\rho$, $\rho \approx 1$m), we obtain $k\rho(\cos - 1 + (1/2)\tan^2\varphi) \leqslant 3\cdot 10^{-11}$. We can therefore set $\Delta \approx \pi/4$. Let us estimate the contribution to the scattering cross section from the part of the probability current proportional to the vector potential $\mathbf{A}$. For two solenoids with $\Phi_1 = -\Phi_2$ and axes passing through the points $y = \pm d$ (Fig. 22), the radial component of the vector potential is

$$A_\rho = \frac{\Phi d\cos\varphi}{\pi(\rho^2+d^2)}\left[1 - \frac{4\rho^2d^2}{(\rho^2+d^2)^2}\sin^2\varphi\right]^{-1}.$$

Since $\rho \gg d$, $A_\rho \approx (\Phi d\cos\varphi)\pi\rho^2$ and $(e\rho/\hbar ck)A_\rho \approx \rho(e\Phi/hc)(2d/k\rho) \approx \rho(e\Phi/hc)\cdot 5\cdot 10^{-18}$. For $e\Phi/hc$ we used the same value as in the experiments of Tonomura *et al.* $e\Phi/hc \approx 5$. Then the contribution of the vector potential to the

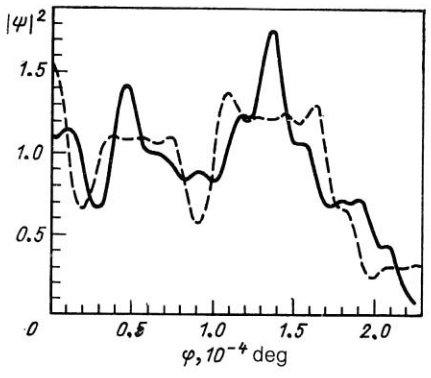


FIG. 23. Probability density of the wave function for electron scattering on two solenoids with $\Phi_1 = -\Phi_2$ in the Fresnel diffraction theory. The distance is $\rho = 106.1$. The solid line corresponds to $\gamma = 0$ and the broken line to $\gamma = 0.5$. The scattering angle in units of $10^{-4}$ deg is plotted along the horizontal axis.

scattering cross section can be neglected. As a result, $\sigma$ and $|\psi|^2$ are proportional ($0 \leqslant \varphi \leqslant (d-R)/\rho$):

$$\sigma = \rho\,|\,\psi\,|^2,$$

$$|\psi|^2 = 1 + \frac{1}{2}(A^2+B^2) - A - B. \qquad (173)$$

In Figs. 23–25 we show the dependence of $|\psi|^2$ on the scattering angle for angles lying between the solenoids ($-\varphi_0 < \varphi_0 < \varphi_0$, $\varphi_0 = (d-R)/\rho$) for three values of $\rho$ (106.1, 98, 90.95). We note that the above-mentioned complications related to the uncertainty in the interpretation of the shift of the diffraction pattern for $\gamma = 0$ and $\gamma = 0.5$ do not occur (in any case, for the values of $\rho$ considered). Furthermore, the deviations of $|\psi|^2$ from the plane-wave value (equal to 1) reach 70% and can be observed experimentally. The differences of $|\psi|^2$ for $\gamma = 0$ and $\gamma = 0.5$ are also fairly large. The decrease of $|\psi|^2$ for $\varphi > 2\cdot 10^{-4}$ is related to the approach to the shadow region. For clarity, in Fig. 26 we show the positions and values of the maxima of $|\psi|^2$ in the region between the solenoids. The solid lines correspond to $\gamma = 0$, and the broken ones to $\gamma = 0.5$. We give only those maxima whose value exceeds 1.

We conclude that the angular region lying between the solenoids is the most promising for the quantitative confirmation of the AB effect. In this region the differences of the cross sections for $\gamma = 0$ and $\gamma = 0.5$ are quite large and can be verified experimentally. In obtaining the cross section (172) we assumed that when the detector is located in the direct beam it is impossible to distinguish the scattered particles from the particles of the incident beam. However, if the

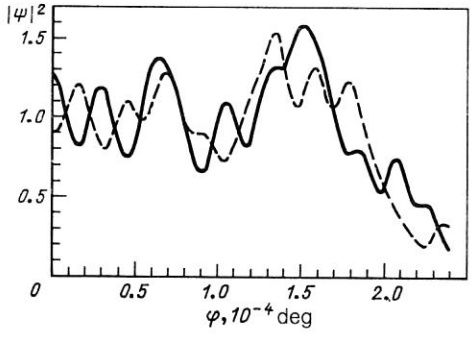


FIG. 24. The same as in Fig. 23, but for $\rho = 98$.

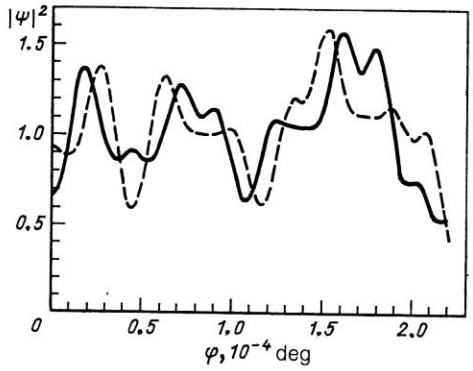


FIG. 25. The same as in Fig. 23, but for $\rho = 91$.

detector sees only the scattered particles, the total wave function $\psi$ in (171) should be replaced by $\psi_s$ [see (163)]. Then under the same experimental conditions for $|\varphi| < (d-R)/\rho$ we have

$$\sigma_s = \frac{1}{8}\rho\left(\frac{1+\cos\varphi}{\cos\varphi}\right)^2 (A^2+B^2)\left(1 - \frac{e}{\hbar kc}A_\rho\right).$$

*Concluding remarks about the experimental verification of the AB effect.* Here we would like to point out the present need for quantitative verification of the AB effect (which we understand to be an observable effect of fields inaccessible to particles). The reason for this is that there are alternative explanations of the shift of the interference pattern when the magnetic field is switched on [see, for example, Refs. 59–61 and, in particular, the animated discussion (pp. 307–312) after the report by Matteucci and Pozzi (pp. 297–306) in Ref. 62]. These explanations are qualitative in nature. Therefore, the experimental confirmation of possibly more

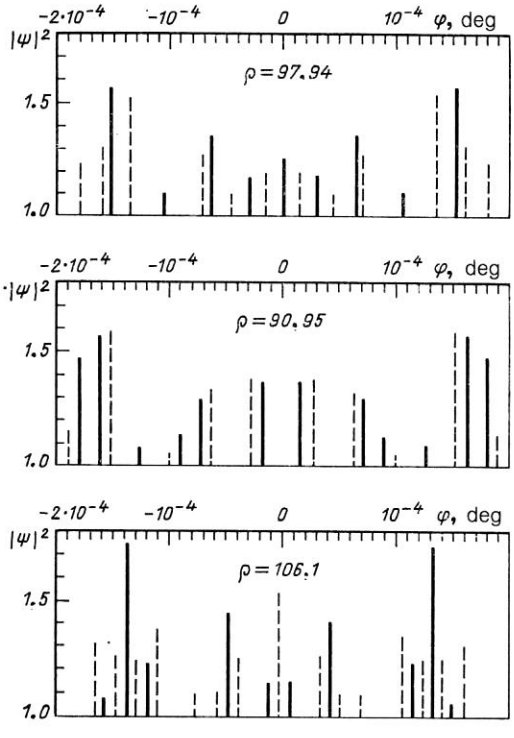


FIG. 26. Locations and maxima of $|\Psi|^2$. The conditions are the same as in Figs. 23–25. Only those maxima whose values exceed 1 are shown.

accurate calculations of the shifts of the diffraction pattern when a magnetic field is switched on would be a strong argument in favor of the physical nature of the vector potential (it is extremely unlikely that the different qualitative interpretations of the AB effect would lead to the same quantitative result). The future experiments which, in our opinion, can eliminate any doubts about the existence of the AB effect are the following:

1) Experiments on electron scattering on cylindrical and toroidal solenoids for the configurations of the initial wave vector, solenoids, and detectors shown in Figs. 6, 15, and 16. In the absence of a magnetic field the detectors $D$, $D_1$, and $D_2$ must see particles owing to the use of electrostatic biprisms even for very large $kR$ and $kd$ (which are experimentally realized). For $\gamma = e\Phi/hc = \frac{1}{2}$ the counting rate of the detectors $D$ and $D_1$ must decrease sharply, while the counting rate of the detectors $D_2$ must remain practically unchanged from its value in the absence of a magnetic field.

2) Electron scattering on two cylindrical solenoids with opposite magnetic fluxes in the configuration of the initial wave vector, solenoids, and detectors shown in Fig. 22. It is essential that there be sufficiently accurate calculated expressions for the cross sections (see Sec. 11), which allow the diffraction pattern to be described in the presence and absence of a magnetic field for realistic dimensions of the experimental setup.

3) Electron scattering on a toroidal solenoid. Unfortunately, the beautiful experiments of Tonomura *et al.*[49] can only be qualitatively described by the equations (154), (156), and (157) of the Fraunhofer diffraction theory. Let us estimate the distances at which the equations of the Fraunhofer diffraction theory begin to work. If we are satisfied with the value of the parameter $kdR/\rho$ equal to 0.1, then for $k$, $d$, and $R$ the same as in the experiments of Tonomura *et al.* ($k = 2\times 10^{10}$ cm$^{-1}$, $R = 10^{-4}$ cm, and $d = 4\times 10^{-4}$ cm) we obtain $\rho = 80$ m. These dimensions are unrealistic, owing both to the dimensions of the setup and to the weakening of the intensity of the scattered particles. As discussed above, the Fresnel theory of diffraction is adequate for describing these experiments.

*Can fields inaccessible to particles be observed in a simply connected space*? Let us again return to the toroidal solenoid surrounded by an impenetrable sphere (Fig. 9). The space accessible to the particles is simply connected. Is it possible to find out that a magnetic field exists by performing experiments outside the solenoid? Since there are no paths accessible to the particles along which $\oint A_l dl \neq 0$, then, according to Ref. 42, such experiments do not exist. The following thought experiment contradicts this statement. We note that the magnetic field inside the solenoid $[H = (\Phi/2\pi\rho)(d - \sqrt{d^2 - R^2})^{-1}]$ leads to the following increase of the solenoid mass:

$$\Delta m = \frac{\mathscr{E}_{\text{magn}}}{c^2}, \quad \mathscr{E}_{\text{magn}} = \frac{1}{8\pi}\int H^2\, dV = \frac{1}{8\pi}\,\frac{\Phi^2}{d - \sqrt{d^2 - R^2}}.$$

This changes the gravitational field outside the solenoid. Therefore, the scattering process for massive particles is different, depending on whether the current in the solenoid is present or absent. This is also true of the parameters of the bound-state orbits.

## CONCLUSIONS

The literature devoted to the AB effect amounts to thousands of titles. Owing to the restrictions on the length of this review, we could not discuss many interesting questions.[2)] Let us mention only a few of them.

1) The Aharonov–Carmi effect.[63,64] Let us describe it briefly. In a rest frame we have a particle of charge $e$ and mass $m$ moving with velocity $\mathbf{v}$. From the viewpoint of an observer located in a coordinate system rotating with constant angular velocity $\omega$, a Coriolis force and a centrifugal force act on the particle. Then electric and magnetic fields can be introduced which completely cancel these forces in a region bounded by two cylinders. From the viewpoint of our observer, no forces act on the particle. Nevertheless, the inertial vector potential $\mathbf{a}$ ($\boldsymbol{\omega} = \operatorname{curl}\mathbf{a}$) is nonzero, which leads to the same physical effects (for example, to a shift of the interference pattern) as in the case of the ordinary AB effect. Strangely enough, this effect was observed[64,65] before it was predicted. The existence of physical effects arising from a nonelectromagnetic vector potential was thereby demonstrated.

2) The analog of the AB effect for neutral particles. Aharonov and Casher[66] predicted this effect on the basis of the following arguments. They noted that the Lagrangian describing the interaction of a charged particle with a pointlike solenoid is invariant under the interchange of the coordinates of the particle and the solenoid. From this they concluded that an electrically charged filament must affect (as in the case of the ordinary AB effect, a magnetic solenoid affects a charged particle) a neutral particle possessing a magnetic moment. At the present time an experiment to verify this effect is being set up. Doubts that this effect exists have been expressed in the literature.[67]

3) The interpretation of the AB effect in terms of modular variables [see Ref. 68, and also the article of Aharonov (pp. 10–19) in Ref. 26]. We have already noted that in the scattering of charged particles on the magnetic field of a toroidal solenoid placed behind an impenetrable screen between two circular apertures the interference picture is different for $\mathbf{A} = 0$ and $\mathbf{A} \neq 0$. Meanwhile, the expectation values of arbitrary powers of the coordinates $x_i$ and momenta $P_i$ and their products are the same for $\mathbf{A} = 0$ and $\mathbf{A} \neq 0$. For example, in the case of the momenta this implies that the particle velocity is unchanged in AB scattering. In the studies just mentioned, quantities were constructed whose expectation values are different for $\mathbf{A} = 0$ and $\mathbf{A} \neq 0$. One of them has the form $f = \sin \boldsymbol{\alpha}\cdot\mathbf{p}/\alpha$. Here $\boldsymbol{\alpha}$ is the vector joining the centers of the holes in the screen and $\mathbf{p}$ is the momentum of the scattered particle. In spite of the fact that when the function $f$ is expanded in a series the expectation value of each term in the expansion $(\boldsymbol{\alpha}\cdot\mathbf{p})^k$ is equal to zero, the expectation value of the function $f$ is not equal to zero. This is also true of the limit of the function $f$ for $\alpha \to 0$.

4) The interpretation of the AB effect in terms of the Berry phase.[69]

5) The AB effect for coherent states. There is an excellent review on this subject in Russian.[44]

In conclusion, we quote F. M. Dostoevskiĭ ("Crime and Punishment"): "... That might be the subject of a new story, but our present story is ended."

[1)]This fact shocked and surprised most physicists. This brings to mind a line by W. Shakespeare ("Hamlet"): "There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy."
[2)]"That which I understood was beautiful, so I think that that which I have not understood is even more beautiful" (Socrates).


[1]W. Ehrenberg and R. Siday, Proc. R. Soc. London Ser. B **62**, 8 (1949).
[2]Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1985).
[3]E. L. Feĭnberg, Usp. Fiz. Nauk **78**, 53 (1962) [Sov. Phys. Usp. **5**, 753 (1962)].
[4]S. N. M. Ruijsenaars, Ann. Phys. (N.Y.) **146**, 1 (1983).
[5]S. Olariu and I. Popescu, Rev. Mod. Phys. **57**, 339 (1985).
[6]I. V. Tyutin, Preprint No. 27 [in Russian], P.N. Lebedev Physics Institute, Moscow (1974).
[7]V. D. Skarzhinskiĭ, Kratk. Soobshch. Fiz. No. 4, 8 (1984); M. V. Berry, R. G. Chambers, and M. D. Large, Eur. J. Phys. **1**, 154 (1980); K. Kawamura, Y. Zempo, and Y. Irie, Prog. Theor. Phys. **67**, 1263 (1982).
[8]E. Corinaldesi and F. Rafeli, Am. J. Phys. **46**, 1185 (1978); K. M. Purcell and W. C. Hennebèrger, Am. J. Phys. **46**, 1255 (1978); W. C. Henneberger, Phys. Rev. A **22**, 1383 (1980); Y. Aharonov, C. K. Au, E. C. Lerner, and J. Q. Liang, Phys. Rev. D **29**, 2396 (1984); B. Nagel, Phys. Rev. D **32**, 3328 (1985).
[9]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, Berlin, 1966).
[10]Y. Aharonov, C. K. Au, E. C. Lerner, and J. Q. Liang, Phys. Rev. D **29**, 2396 (1984); B. Nagel, Phys. Rev. D **32**, 3328 (1985); R. A. Brown, J. Phys. A **18**, 2497 (1985).
[11]T. Takabajasi, Hadr. J. Suppl. **1**, 219 (1985).
[12]P. A. M. Dirac, Proc. R. Soc. London Ser. A **133**, 60 (1931).
[13]M. V. Berry, Eur. J. Phys. **1**, 240 (1980).
[14]J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952) [Russ. transl., IL, Moscow, 1954, p. 607].
[15]W. Pauli, Helv. Phys. Acta **12**, 147 (1939) [Russ. transl., Nauka, Moscow, 1977].
[16]S. M. Roy and V. Singh, Nuovo Cimento **79A**, 391 (1984).
[17]W. C. Henneberger, J. Math. Phys. **22**, 116 (1981).
[18]P. Bocchieri and A. Loinger, Nuovo Cimento **47A**, 475 (1978); **59A**, 121; **66A**, 164 (1981); Lett. Nuovo Cimento **25**, 476 (1979); **35**, 469 (1982); **39**, 148 (1982); P. Bocchieri, A. Loinger, and G. Siragusa, Nuovo Cimento **51A**, 1 (1979); **56A**, 55 (1980).
[19]F. Wilczek, Phys. Rev. Lett. **48**, 1144 (1982).
[20]R. Jackiw and A. N. Redlich, Phys. Rev. Lett. **50**, 555 (1983).
[21]M. Peshkin, Phys. Rep. **80**, 375 (1981); H. J. Lipkin and M. Peshkin, Phys. Lett. **118B**, 385 (1982).
[22]S. M. Roy, Phys. Rev. Lett. **44**, 111 (1980).
[23]D. M. Greenberger, Phys. Rev. D **23**, 1460 (1981); U. Klein, Phys. Rev. D **23**, 1463 (1981); H. J. Lipkin, Phys. Rev. D **23**, 1466 (1981).
[24]D. H. Kobe and J. Q. Liang, Phys. Lett. **118A**, 475 (1986); J. Q. Liang and X. X. Ding, Phys. Lett. **119A**, 325 (1987).
[25]M. Bawin and A. Burnel, J. Phys. A **18**, 2123 (1985).
[26]*Proceedings of the International Symposium on Foundations of Quantum Mechanics in the Light of New Technology*, edited by S. Kamefuchi (Japan Physical Society, Tokyo, 1984).
[27]M. Bawin and A. Burnel, Phys. Lett. **122A**, 451 (1987).
[28]P. Frolov and V. D. Skarzhinsky, Nuovo Cimento **B76**, 32 (1983); D. Home and S. Sengupta, Am. J. Phys. **51**, 942 (1983).
[29]M. Peshkin, I. Talmi, and L. J. Tassie, Ann. Phys. (N.Y.) **12**, 426 (1961).
[30]B. Nagel, in *Fundamental Aspects of Quantum Theory*, edited by V. Gorini and A. Frigerio (Plenum Press, New York, 1986), p. 335.
[31]G. N. Afanas'ev, Preprint R2-87-154 [in Russian], JINR, Dubna (1987).
[32]D. Bohm and B. J. Hiley, Nuovo Cimento **A52**, 295 (1979); D. Bohm, R. D. Kaye, and C. Phillippidis, Nuovo Cimento **B71**, 75 (1982).
[33]S. Olariu and I. I. Popescu, Phys. Rev. D **33**, 1701 (1986); H. J. Rothe, Nuovo Cimento **A62**, 54 (1981); G. N. Afanasiev, JINR Rapid Commun. **6**, 17 (1985).
[34]M. Born and E. Wolf, *Principles of Optics*, 4th ed. (Pergamon Press, Oxford, 1969) [Russ. transl., Nauka, Moscow, 1970]; L. I. Mandel'shtam, *Lectures on Optics, the Theory of Relativity, and Quantum Mechanics* [in Russian] (Nauka, Moscow, 1972).
[35]R. Glauber, *Lectures in Theoretical Physics*, Vol. 1 (Interscience, New York, 1954); A. G. Sitenko, *Scattering Theory* [in Russian] (Vyshcha Shkola, Kiev, 1975).
[36]P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, Vol. II (McGraw-Hill, New York, 1953) [Russ. transl., Vol. 2, IL, Moscow, 1960].
[37]G. N. Afansiev, J. Comput. Phys. **69**, 196 (1987); G. N. Afanas'ev, Preprint R4-87-106 [in Russian], JINR, Dubna (1987).
[38]G. N. Afanas'ev, Preprint R4-87-107 [in Russian], JINR, Dubna (1987); G. N. Afanasiev, J. Phys. A **21**, 2095 (1988).
[39]V. L. Lyuboshits and Ya. A. Smorodinskiĭ, Zh. Eksp. Teor. Fiz. **75**, 40 (1978) [Sov. Phys. JETP **48**, 19 (1978)]; Preprint R2-1189 [in Russian], JINR, Dubna (1978).
[40]L. J. Tassie, Phys. Lett. **5**, 43 (1961).
[41]G. N. Afanas'ev, Preprint R4-87-569 [in Russian], JINR, Dubna (1987); *Questions of Atomic Science and Technology, in the Series: General and Nuclear Physics* [in Russian], No. 1 (1988), p. 49; G. N. Afanasiev, Nuovo Cimento **A99**, 647 (1988).
[42]C. N. Yang and T. T. Wu, Phys. Rev. D **12**, 3845 (1988).
[43]J. A. Mignaco and C. A. Novaes, Lett. Nuovo Cimento **26**, 453 (1979).
[44]V. D. Skarzhinskiĭ, Tr. Fiz. Inst. Akad. Nauk SSSR **167**, 139 (1986).
[45]V. G. Bagrov, D. M. Gitman, and V. D. Skarzhinskiĭ, Tr. Fiz. Inst. Akad. Nauk SSSR **176**, 151 (1986).
[46]R. R. Lewis, Phys. Rev. A **28**, 1228 (1983).
[47]E. Bakanas, Litov. Fiz. Sb. **28**, 15 (1988).
[48]A Loinger, Riv. Nuovo Cimento **10**, 1 (1987); D. H. Kobe and J. Q. Liang, Phys. Rev. A **37**, 1133 (1988); see also Bocchieri's and Loinger's papers in *Fundamental Aspects of Quantum Theory*, edited by V. Gorini and A. Frigerio (Plenum Press, New York, 1986).
[49]N. Osakabe, T. Matsuda, T. Kawasaki *et al.*, Phys. Rev. A **34**, 815 (1986); see also the same authors' paper in *Proceedings of the Second International Symposium "Foundations of Quantum Mechanics in the Light of New Technology,"* edited by M. Namiki, Y. Ohnuki, Y. Marajama, and S. Nomara (Japan Physics Society, Tokyo, 1987), p. 97.
[50]B. L. Al'tshuler, A. G. Aronov, B. Z. Spivak *et al.*, Pis'ma Zh. Eksp. Teor. Fiz. **35**, 476 (1982) [JETP Lett. **35**, 588 (1982)]; S. Washburn and R. A. Webb, Adv. Phys. **35**, 375 (1986); S. Datta, M. R. Mellock, S. Bandyopadhyay *et al.*, Phys. Rev. Lett. **55**, 2344 (1985).
[51]B. Schwarzschild, Physics Today, p. 17 (January 1986) [Russ. transl., Mir, Moscow, 1987].
[52]J. Q. Liang and P. Bocchieri, Phys. Lett. **127A**, 221 (1988); J. Q. Liang and X. X. Ding, Phys. Rev. Lett. **60**, 836 (1988).
[53]G. N. Afanas'ev, Preprint R2-88-282 [in Russian], JINR, Dubna (1988); G. N. Afanasiev, Nuovo Cimento **A100**, 967 (1988).
[54]S. Silver, J. Opt. Soc. Am. **52**, 131 (1962).
[55]G. F. Missiroli, G. Pozzi, and U. Valdre, J. Phys. E **14**, 649 (1981); see also references therein.
[56]G. N. Afanas'ev and V. M. Shilov, Preprint R4-88-841 [in Russian], JINR, Dubna (1988).
[57]J. Komrska, in *Advances in Electronics and Electron Physics*, edited by L. Marton (Academic Press, New York, 1971), p. 139.
[58]P. A. M. Dirac, *The Principles of Quantum Mechanics*, 3rd ed. (Clarendon Press, Oxford, 1947) [Russ. transl., Nauka, Moscow, 1979].
[59]T. M. Boyer, Nuovo Cimento **B100**, 685 (1987).
[60]E. Comay, Fizika **20**, 241 (1988).
[61]V. A. Namiot, Phys. Lett. **124A**, 9 (1988).
[62]*Quantum Uncertainties, Recent and Future Experiments and Interpretations*, edited by W. M. Honig, D. W. Kraft, and E. Panarella (Plenum Press, New York, 1987).
[63]Y. Aharonov and G. Carmi, Found. Phys. **3**, 493 (1973); **4**, 75 (1974); J. H. Harris and M. D. Semon, Found. Phys. **10**, 151 (1980).
[64]M. D. Semon, Found. Phys. **12**, 49 (1982).
[65]J. E. Zimmerman and M. Mercereau, Phys. Rev. Lett. **14**, 887 (1965).
[66]Y. Aharonov and A. Casher, Phys. Rev. Lett. **53**, 319 (1984).
[67]T. H. Boyer, Phys. Rev. A **36**, 5083 (1988).
[68]Y. Aharonov, H. Pendleton, and A. Petersen, Int. J. Theor. Phys. **2**, 213 (1969); M. D. Semon and J. R. Taylor, Nuovo Cimento **B97**, 25 (1987); **B100**, 389 (1987).
[69]M. V. Berry, Proc. R. Soc. London Ser. A **392**, 45 (1984); Y. Aharonov and J. Anandan, Phys. Rev. Lett. **58**, 1593 (1987).


Translated by Patricia Millard