Significance of Electromagnetic Potentials in the Quantum Theory

Y. AHARONOV AND D. BOHM

H. H. Wills Physics Laboratory, University of Bristol, Bristol, England

(Received May 28, 1959; revised manuscript received June 16, 1959)

In this paper, we discuss some interesting properties of the electromagnetic potentials in the quantum domain. We shall show that, contrary to the conclusions of classical mechanics, there exist effects of potentials on charged particles, even in the region where all the fields (and therefore the forces on the particles) vanish. We shall then discuss possible experiments to test these conclusions; and, finally, we shall suggest further possible developments in the interpretation of the potentials.

1. INTRODUCTION

IN classical electrodynamics, the vector and scalar potentials were first introduced as a convenient mathematical aid for calculating the fields. It is true that in order to obtain a classical canonical formalism, the potentials are needed. Nevertheless, the fundamental equations of motion can always be expressed directly in terms of the fields alone.

In the quantum mechanics, however, the canonical formalism is necessary, and as a result, the potentials cannot be eliminated from the basic equations. Nevertheless, these equations, as well as the physical quantities, are all gauge invariant; so that it may seem that even in quantum mechanics, the potentials themselves have no independent significance.

In this paper, we shall show that the above conclusions are not correct and that a further interpretation of the potentials is needed in the quantum mechanics.

2. POSSIBLE EXPERIMENTS DEMONSTRATING THE ROLE OF POTENTIALS IN THE QUANTUM THEORY

In this section, we shall discuss several possible experiments which demonstrate the significance of potentials in the quantum theory. We shall begin with a simple example.

Suppose we have a charged particle inside a "Faraday cage" connected to an external generator which causes the potential on the cage to alternate in time. This will add to the Hamiltonian of the particle a term $V(x,t)$ which is, for the region inside the cage, a function of time only. In the nonrelativistic limit (and we shall

assume this almost everywhere in the following discussions) we have, for the region inside the cage, $H = H_0 + V(t)$ where H_0 is the Hamiltonian when the generator is not functioning, and $V(t) = e\phi(t)$. If $\psi_0(x,t)$ is a solution of the Hamiltonian H_0, then the solution for H will be

$$\psi = \psi_0 e^{-iS/\hbar}, \quad S = \int V(t) dt,$$

which follows from

$$i\hbar \frac{\partial \psi}{\partial t} = \left(i\hbar \frac{\partial \psi_0}{\partial t} + \psi_0 \frac{\partial S}{\partial t} \right) e^{-iS/\hbar} = [H_0 + V(t)]\psi = H\psi.$$

The new solution differs from the old one just by a phase factor and this corresponds, of course, to no change in any physical result.

Now consider a more complex experiment in which a single coherent electron beam is split into two parts and each part is then allowed to enter a long cylindrical metal tube, as shown in Fig. 1.

After the beams pass through the tubes, they are combined to interfere coherently at F. By means of time-determining electrical "shutters" the beam is chopped into wave packets that are long compared with the wavelength λ, but short compared with the length of the tubes. The potential in each tube is determined by a time delay mechanism in such a way that the potential is zero in region I (until each packet is well inside its tube). The potential then grows as a function of time, but differently in each tube. Finally, it falls back to zero, before the electron comes near the
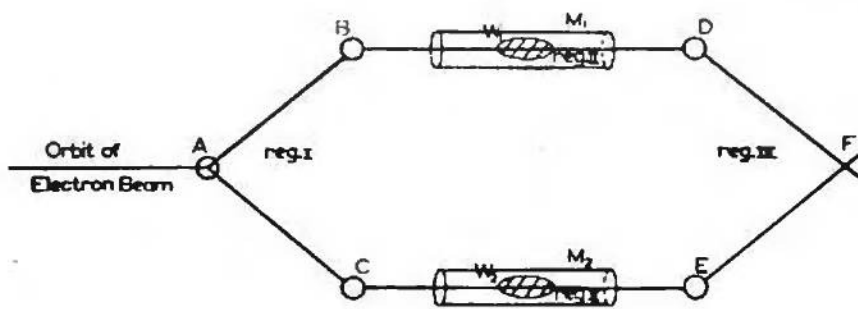
FIG. 1. Schematic experiment to demonstrate interference with time-dependent scalar potential. A, B, C, D, E: suitable devices to separate and divert beams. W_1, W_2: wave packets. M_1, M_2: cylindrical metal tubes. F: interference region.

FIG. 2. Schematic experiment to demonstrate interference with time-independent vector potential.

other edge of the tube. Thus the potential is nonzero only while the electrons are well inside the tube (region II). When the electron is in region III, there is again no potential. The purpose of this arrangement is to ensure that the electron is in a time-varying potential without ever being in a field (because the field does not penetrate far from the edges of the tubes, and is nonzero only at times when the electron is far from these edges).

Now let $\psi(x,t)=\psi_1^0(x,t)+\psi_2^0(x,t)$ be the wave function when the potential is absent (ψ_1^0 and ψ_2^0 representing the parts that pass through tubes 1 and 2, respectively). But since V is a function only of t wherever ψ is appreciable, the problem for each tube is essentially the same as that of the Faraday cage. The solution is then

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

where

$$S_1=e\int \varphi_1 dt, \quad S_2=e\int \varphi_2 dt.$$

It is evident that the interference of the two parts at F will depend on the phase difference $(S_1-S_2)/\hbar$. Thus, there is a physical effect of the potentials even though no force is ever actually exerted on the electron. The effect is evidently essentially quantum-mechanical in nature because it comes in the phenomenon of interference. We are therefore not surprised that it does not appear in classical mechanics.

From relativistic considerations, it is easily seen that the covariance of the above conclusion demands that there should be similar results involving the vector potential, A.

The phase difference, $(S_1-S_2)/\hbar$, can also be expressed as the integral $(e/\hbar)\oint \varphi dt$ around a closed circuit in space-time, where φ is evaluated at the place of the center of the wave packet. The relativistic generalization of the above integral is

$$\frac{e}{\hbar}\oint \left(\varphi dt - \frac{A}{c}\cdot dx \right),$$

where the path of integration now goes over any closed circuit in space-time.

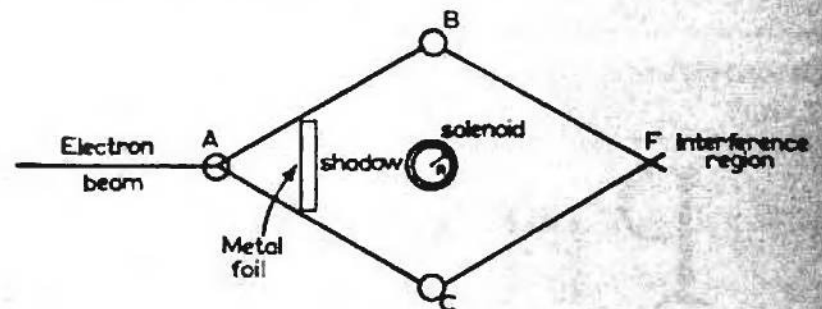As another special case, let us now consider a path in space only ($t=$constant). The above argument suggests that the associated phase shift of the electron wave function ought to be

$$\Delta S/\hbar=-\frac{e}{ch} \oint A\cdot dx,$$

where $\oint A\cdot dx= \int H\cdot ds=\phi$ (the total magnetic flux inside the circuit).

This corresponds to another experimental situation. By means of a current flowing through a very closely wound cylindrical solenoid of radius R, center at the origin and axis in the z direction, we create a magnetic field, H, which is essentially confined within the solenoid. However, the vector potential, A, evidently, cannot be zero everywhere outside the solenoid, because the total flux through every circuit containing the origin is equal to a constant

$$\phi_0= \int H\cdot ds= \int A\cdot dx.$$

To demonstrate the effects of the total flux, we begin, as before, with a coherent beam of electrons. (But now there is no need to make wave packets.) The beam is split into two parts, each going on opposite sides of the solenoid, but avoiding it. (The solenoid can be shielded from the electron beam by a thin plate which casts a shadow.) As in the former example, the beams are brought together at F (Fig. 2).

The Hamiltonian for this case is

$$H=\frac{[P-(e/c)A]^2}{2m}.$$

In singly connected regions, where $H=\nabla\times A=0$, we can always obtain a solution for the above Hamiltonian by taking $\psi=\psi_0 e^{-iS/\hbar}$, where ψ_0 is the solution when $A=0$ and where $\nabla S/\hbar=(e/c)A$. But, in the experiment discussed above, in which we have a multiply connected region (the region outside the solenoid), $\psi_0 e^{-iS/\hbar}$ is a non-single-valued function[1] and therefore, in general, not a permissible solution of Schrödinger's equation. Nevertheless, in our problem it is still possible to use such solutions because the wave function splits into two parts $\psi=\psi_1+\psi_2$, where ψ_1 represents the beam on

---

[1] Unless $\phi_0=nhc/e$, where n is an integer.

one side of the solenoid and ψ_2 the beam on the opposite side. Each of these beams stays in a simply connected region. We therefore can write

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

where S_1 and S_2 are equal to $(e/c)\int \mathbf{A} \cdot d\mathbf{x}$ along the paths of the first and second beams, respectively. (In Sec. 4, an exact solution for this Hamiltonian will be given, and it will confirm the above results.)

The interference between the two beams will evidently depend on the phase difference,

$$(S_1 - S_2)/h = (e/hc)\int \mathbf{A} \cdot d\mathbf{x} = (e/hc)\phi_0.$$

This effect will exist, even though there are no magnetic forces acting in the places where the electron beam passes.

In order to avoid fully any possible question of contact of the electron with the magnetic field we note that our result would not be changed if we surrounded the solenoid by a potential barrier that reflects the electrons perfectly. (This, too, is confirmed in Sec. 4.)

It is easy to devise hypothetical experiments in which the vector potential may influence not only the interference pattern but also the momentum. To see this, consider a periodic array of solenoids, each of which is shielded from direct contact with the beam by a small plate. This will be essentially a grating. Consider first the diffraction pattern without the magnetic field, which will have a discrete set of directions of strong constructive interference. The effect of the vector potential will be to produce a shift of the relative phase of the wave function in different elements of the gratings. A corresponding shift will take place in the directions, and therefore the momentum of the diffracted beam.

3. A PRACTICABLE EXPERIMENT TO TEST FOR THE EFFECTS OF A POTENTIAL WHERE THERE ARE NO FIELDS

As yet no direct experiments have been carried out which confirm the effect of potentials where there is no field. It would be interesting therefore to test whether such effects actually exist. Such a test is, in fact, within the range of present possibilities.[2] Recent experiments[3,4] have succeeded in obtaining interference from electron beams that have been separated in one case by as much as 0.8 mm.[3] It is quite possible to wind solenoids which are smaller than this, and therefore to place them between the separate beams. Alternatively, we may obtain localized lines of flux of the right magnitude (the

[2] Dr. Chambers is now making a preliminary experimental study of this question at Bristol.
[3] L. Marton, Phys. Rev. 85, 1057 (1952); 90, 490 (1953). Marton, Simpson, and Suddeth, Rev. Sci. Instr. 25, 1099 (1954).
[4] G. Mollenstedt, Naturwissenschaften 42, 41 (1955); G. Mollenstedt and H. Düker, Z. Physik 145, 377 (1956).

magnitude has to be of the order of $\phi_0 = 2\pi c\hbar/e \sim 4 \times 10^{-7}$ gauss cm²) by means of fine permanently magnetized "whiskers".[5] The solenoid can be used in Marton's device,[3] while the whisker is suitable for another experimental setup[4] where the separation is of the order of microns and the whiskers are even smaller than this.

In principle, we could do the experiment by observing the interference pattern with and without the magnetic flux. But since the main effect of the flux is only to displace the line pattern without changing the interval structure, this would not be a convenient experiment to do. Instead, it would be easier to vary the magnetic flux within the same exposure for the detection of the interference patterns. Such a variation would, according to our previous discussion, alter the sharpness and the general form of the interference bands. This alteration would then constitute a verification of the predicted phenomena.

When the magnetic flux is altered, there will, of course, be an induced electric field outside the solenoid, but the effects of this field can be made negligible. For example, suppose the magnetic flux were suddenly altered in the middle of an exposure. The electric field would then exist only for a very short time, so that only a small part of the beam would be affected by it.

4. EXACT SOLUTION FOR SCATTERING PROBLEMS

We shall now obtain an exact solution for the problem of the scattering of an electron beam by a magnetic field in the limit where the magnetic field region tends to a zero radius, while the total flux remains fixed. This corresponds to the setup described in Sec. 2 and shown in Fig. 2. Only this time we do not split the plane wave into two parts. The wave equation outside the magnetic field region is, in cylindrical coordinates,

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial}{\partial \theta} + i\alpha \right)^2 + k^2 \right]\psi = 0, \quad (1)$$

where $\mathbf{k}$ is the wave vector of the incident particle and $\alpha = -e\phi/ch$. We have again chosen the gauge in which $A_r = 0$ and $A_\theta = \phi/2\pi r$.

The general solution of the above equation is

$$\psi = \sum_{m=-\infty}^{\infty} e^{im\theta}[a_m J_{m+\alpha}(kr) + b_m J_{-(m+\alpha)}(kr)], \quad (2)$$

where a_m and b_m are arbitrary constants and $J_{m+\alpha}(kr)$ is a Bessel function, in general of fractional order (dependent on ϕ). The above solution holds only for $r > R$. For $r < R$ (inside the magnetic field) the solution has been worked out.[6] By matching the solutions at $r = R$ it is easily shown that only Bessel functions of positive order will remain, when R approaches zero.

[5] See, for example, Sidney S. Brenner, Acta Met. 4, 62 (1956).
[6] L. Page, Phys. Rev. 36, 444 (1930).

This means that the probability of finding the particle inside the magnetic field region approaches zero with R. It follows that the wave function would not be changed if the electron were kept away from the field by a barrier whose radius also went to zero with R.

The general solution in the limit of R tending to zero is therefore

$$\psi = \sum_{m=-\infty}^{\infty} a_m J_{|m+\alpha|} e^{im\theta}. \qquad (3)$$

We must then choose a_m so that ψ represents a beam of electrons that is incident from the right ($\theta=0$). It is important, however, to satisfy the initial condition that the current density,

$$j = \frac{h(\psi^*\nabla\psi - \psi\nabla\psi^*)}{2im} - \frac{e}{mc}A\psi^*\psi, \qquad (4)$$

shall be constant and in the x direction. In the gauge that we are using, we easily see that the correct incident wave is $\psi_{\text{inc}} = e^{-ikx}e^{-i\alpha\theta}$. Of course, this wave function holds only to the right of the origin, so that no problem of multiple-valuedness arises.

We shall show in the course of this calculation that the above conditions will be satisfied by choosing $a_m = (-i)^{|m+\alpha|}$, in which case, we shall have

$$\psi = \sum_{m=-\infty}^{\infty} (-i)^{|m+\alpha|} J_{|m+\alpha|} e^{im\theta}.$$

It is convenient to split ψ into the following three parts: $\psi = \psi_1 + \psi_2 + \psi_3$, where

$$\psi_1 = \sum_{m=1}^{\infty} (-i)^{m+\alpha} J_{m+\alpha} e^{im\theta},$$

$$\psi_2 = \sum_{m=-\infty}^{-1} (-i)^{m+\alpha} J_{m+\alpha} e^{im\theta},$$

$$= \sum_{m=1}^{\infty} (-i)^{m-\alpha} J_{m-\alpha} e^{-im\theta}, \qquad (5)$$

$$\psi_3 = (-i)^{|\alpha|} J_{|\alpha|}.$$

Now ψ_1 satisfies the simple differential equation

$$\frac{\partial\psi_1}{\partial r'} = \sum_{m=1}^{\infty} (-i)^{m+\alpha} J_{m+\alpha}' e^{im\theta}$$

$$= \sum_{m=1}^{\infty} (-i)^{m+\alpha} \frac{J_{m+\alpha-1} - J_{m+\alpha+1}}{2} e^{im\theta}, \quad r'=kr \qquad (6)$$

where we have used the well-known formula for Bessel functions:

$$dJ_\gamma(r)/dr = \tfrac{1}{2}(J_{\gamma-1} - J_{\gamma+1}).$$

As a result, we obtain

$$\frac{\partial\psi_1}{\partial r'} = \frac{1}{2}\sum_{m'=0}^{\infty} (-i)^{m'+\alpha+1} J_{m'+\alpha} e^{i(m'+1)\theta}$$

$$- \frac{1}{2}\sum_{m'=2}^{\infty} (-i)^{m'+\alpha-1} J_{m'+\alpha} e^{i(m'-1)\theta} \qquad (7)$$

$$= \frac{1}{2}\sum_{m'=1}^{\infty} (-i)^{m'+\alpha} J_{m'+\alpha} e^{im'\theta}(-ie^{i\theta} + i^{-1}e^{-i\theta})$$

$$+ \tfrac{1}{2}(-i)^\alpha [J_{\alpha+1} - ie^{i\theta}J_\alpha].$$

So

$$\partial\psi_1/\partial r' = -i\cos\theta\,\psi_1 + \tfrac{1}{2}(-i)^\alpha(J_{\alpha+1} - iJ_\alpha e^{i\theta}).$$

This differential equation can be easily integrated to give

$$\psi_1 = A\int_0^{r'} e^{ir'\cos\theta}[J_{\alpha+1} - iJ_\alpha e^{i\theta}]dr', \qquad (8)$$

where

$$A = \tfrac{1}{2}(-i)^\alpha e^{-ir'\cos\theta}.$$

The lower limit of the integration is determined by the requirement that when r' goes to zero, ψ_1 also goes to to zero because, as we have seen, ψ_1 includes Bessel functions of positive order only.

In order to discuss the asymptotic behavior of ψ_1, let us write it as $\psi_1 = A[I_1 - I_2]$, where

$$I_1 = \int_0^\infty e^{ir'\cos\theta}[J_{\alpha+1} - ie^{i\theta}J_\alpha]dr',$$

$$I_2 = \int_r^\infty e^{ir'\cos\theta}[J_{\alpha+1} - ie^{i\theta}J_\alpha]dr'. \qquad (9)$$

The first of these integrals is known[7]:

$$\int_0^\infty e^{i\beta r}J_\alpha(kr) = \frac{e^{i[\alpha\arcsin(\beta/k)]}}{(k^2-\beta^2)^{\frac{1}{2}}}, \quad 0<\beta<k, \quad -2<\alpha.$$

In our cases, $\beta=\cos\theta$, $k=1$, so that

$$I_1 = \left[\frac{e^{i\alpha(\frac{1}{2}\pi-|\theta|)}}{|\sin\theta|} - ie^{i\theta}\frac{e^{i(\alpha+1)(\frac{1}{2}\pi-|\theta|)}}{|\sin\theta|}\right]. \qquad (10)$$

Because the integrand is even in θ, we have written the final expression for the above integral as a function of $|\theta|$ and of $|\sin\theta|$. Hence

$$I_1 = e^{i\alpha(\frac{1}{2}\pi-|\theta|)}\left[\frac{ie^{-i|\theta|} - ie^{i\theta}}{|\sin\theta|}\right]$$

$$= 0 \quad \text{for } \theta<0,$$

$$= e^{-i\alpha\theta}2i^\alpha \quad \text{for } \theta>0, \qquad (11)$$

where we have taken θ as going from $-\pi$ to π.

[7] See, for example, W. Gröbner and N. Hofreiter, *Integraltafel* (Springer-Verlag, Berlin, 1949).

We shall see presently that I_1 represents the largest term in the asymptotic expansion of ψ_1. The fact that it is zero for $\theta < 0$ shows that this part of ψ_1 passes (asymptotically) only on the upper side of the singularity. To explain this, we note that ψ_1 contains only positive values of m, and therefore of the angular momentum. It is quite natural then that this part of ψ_1 goes on the upper side of the singularity. Similarly, since according to (5)

$$\psi_2(r',\theta,\alpha)=\psi_1(r',-\theta,-\alpha),$$

it follows that ψ_2 will behave oppositely to ψ_1 in this regard, so that together they will make up the correct incident wave.

Now, in the limit of $r' \to \infty$ we are allowed to take in the integrand of I_2 the first asymptotic term of J_α,[8] namely $J_\alpha \to (2/\pi r)^{\frac{1}{2}} \cos(r'-\frac{1}{2}\alpha-\frac{1}{4}\pi)$. We obtain

$$I_2=\int_r^\infty e^{ir'\cos\theta}(J_{\alpha+1}-ie^{i\theta}J_\alpha)dr' \to C+D, \quad (12)$$

where

$$C=\int_r^\infty e^{ir'\cos\theta}[\cos(r'-\tfrac{1}{2}(\alpha+1)\pi-\tfrac{1}{4}\pi)]\frac{dr'}{(r')^{\frac{1}{2}}}\left(\frac{2}{\pi}\right)^{\frac{1}{2}},$$

$$D=\int_r^\infty e^{ir'\cos\theta}[\cos(r'-\tfrac{1}{2}\alpha-\tfrac{1}{4}\pi)]\frac{dr'}{(r')^{\frac{1}{2}}}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}(-i)e^{i\theta}.$$

$$(13)$$

Then

$$C=\int_r^\infty e^{ir'\cos\theta}[e^{i[r'-\frac{1}{2}(\alpha+1)\pi-\frac{1}{4}\pi]}$$
$$+e^{-i[r'-\frac{1}{2}(\alpha+1)\pi-\frac{1}{4}\pi]}]\frac{dr'}{(2\pi r')^{\frac{1}{2}}}$$
$$=\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{(-i)^{\alpha+\frac{1}{2}}}{(1+\cos\theta)^{\frac{1}{2}}}\int_{[r'(1+\cos\theta)]^{\frac{1}{2}}}^\infty \exp(+iz^2)dz$$
$$+\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{i^{\alpha+\frac{1}{2}}}{(1-\cos\theta)^{\frac{1}{2}}}\int_{[r'(1-\cos\theta)]^{\frac{1}{2}}}^\infty \exp(-iz^2)dz, \quad (14)$$

where we have put

$$z=[r'(1+\cos\theta)]^{\frac{1}{2}} \quad \text{and} \quad z=[r'(1-\cos\theta)]^{\frac{1}{2}},$$

respectively.

Using now the well-known asymptotic behavior of the error function,[9]

$$\int_a^\infty \exp(iz^2)dz \to \frac{i}{2}\frac{\exp(ia^2)}{a},$$
$$\int_a^\infty \exp(-iz^2)dz \to \frac{-i}{2}\frac{\exp(-ia^2)}{a}, \quad (15)$$

[8] E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, Inc., New York, 1943), fourth edition, p. 138.
[9] Reference 8, p. 24.

we finally obtain

$$C=\left[\frac{(-i)^{\alpha+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{ir'}}{[r'(1+\cos\theta)^2]^{\frac{1}{2}}}\right.$$
$$\left.+\frac{i^{\alpha+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{-ir'}}{[r'(1-\cos\theta)^2]^{\frac{1}{2}}}\right]e^{ir'\cos\theta}, \quad (16)$$

$$D=\left[\frac{(-i)^{\alpha-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{ir'}}{[r'(1+\cos\theta)^2]^{\frac{1}{2}}}\right.$$
$$\left.+\frac{i^{\alpha-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{-ir'}}{[r'(1-\cos\theta)^2]^{\frac{1}{2}}}\right]e^{ir'\cos\theta}(-i)e^{i\theta}. \quad (17)$$

Now adding (16) and (17) together and using (13) and (9), we find that the term of $1/(r')^{\frac{1}{2}}$ in the asymptotic expansion of ψ_1 is

$$\frac{(-i)^{\frac{1}{2}}}{2(2\pi)^{\frac{1}{2}}}\left[(-1)^\alpha\frac{e^{ir'}}{(r')^{\frac{1}{2}}}\frac{1+e^{i\theta}}{1+\cos\theta}+i\frac{e^{-ir'}}{(r')^{\frac{1}{2}}}\frac{1-e^{i\theta}}{1-\cos\theta}\right]. \quad (18)$$

Using again the relation between ψ_1 and ψ_2 we obtain for the corresponding term in ψ_2

$$\frac{(-i)^{\frac{1}{2}}}{2(2\pi)^{\frac{1}{2}}}\left[(-1)^{-\alpha}\frac{e^{ir'}}{(r')^{\frac{1}{2}}}\frac{1+e^{-i\theta}}{1+\cos\theta}+i\frac{e^{-ir'}}{(r')^{\frac{1}{2}}}\frac{1-e^{-i\theta}}{1-\cos\theta}\right]. \quad (19)$$

Adding (18) and (19) and using (11), we finally get

$$\psi_1+\psi_2 \to \frac{(-i)^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\left[\frac{ie^{-ir'}}{(r')^{\frac{1}{2}}}+\frac{e^{ir'}}{(r')^{\frac{1}{2}}}\frac{\cos(\pi\alpha-\frac{1}{2}\theta)}{\cos(\frac{1}{2}\theta)}\right]$$
$$+e^{-i(r'\cos\theta+\alpha\theta)}. \quad (20)$$

There remains the contribution of ψ_3, whose asymptotic behavior is [see Eq. (12)]

$$(-i)^{|\alpha|}J_{|\alpha|}(r') \to (-i)^{|\alpha|}\left(\frac{2}{\pi r'}\right)^{\frac{1}{2}}\cos(r'-\tfrac{1}{4}\pi-\tfrac{1}{2}|\alpha|\pi).$$

Collecting all terms, we find

$$\psi=\psi_1+\psi_2+\psi_3 \to e^{-i(\alpha\theta+r'\cos\theta)}+\frac{e^{ir'}}{(2\pi i r')^{\frac{1}{2}}}\sin\pi\alpha\frac{e^{-i\theta/2}}{\cos(\theta/2)}, \quad (21)$$

where the $\pm$ sign is chosen according to the sign of α.

The first term in equation (21) represents the incident wave, and the second the scattered wave.[10] The scattering cross section is therefore

$$\sigma=\frac{\sin^2\pi\alpha}{2\pi}\frac{1}{\cos^2(\theta/2)}. \quad (22)$$

[10] In this way, we verify, of course, that our choice of the a_m for Eq. (3) satisfies the correct boundary conditions.

When $\alpha = n$, where n is an integer, then σ vanishes. This is analogous to the Ramsauer effect.[11] σ has a maximum when $\alpha = n + \frac{1}{2}$.

The asymptotic formula (21) holds only when we are not on the line $\theta = \pi$. The exact solution, which is needed on this line, would show that the second term will combine with the first to make a single-valued wave function, despite the non-single-valued character of the two parts, in the neighborhood of $\theta = \pi$. We shall see this in more detail presently for the special case $\alpha = n + \frac{1}{2}$.

In the interference experiment discussed in Sec. 2, diffraction effects, represented in Eq. (21) by the scattered wave, have been neglected. Therefore, in this problem, it is adequate to use the first term of Eq. (21). Here, we see that the phase of the wave function has a different value depending on whether we approach the line $\theta = \pm \pi$ from positive or negative angles, i.e., from the upper or lower side. This confirms the conclusions obtained in the approximate treatment of Sec. 2.

We shall discuss now the two special cases that can be solved exactly. The first is the case where $\alpha = n$. Here, the wave function is $\psi = e^{-ikx}e^{-i\alpha\theta}$, which is evidently single-valued when α is an integer. (It can be seen by direct differentiation that this is a solution.)

The second case is that of $\alpha = n + \frac{1}{2}$. Because $J_{(n+\frac{1}{2})}(r)$ is a closed trigonometric function, the integrals for ψ can be carried out exactly.

The result is

$$\psi = \frac{i^{\frac{1}{2}}}{\sqrt{2}}e^{-i(\frac{1}{2}\theta + r'\cos\theta)}\int_0^{[r'(1+\cos\theta)]^{\frac{1}{2}}} \exp(iz^2)dz. \quad (23)$$

This function vanishes on the line $\theta = \pi$. It can be seen that its asymptotic behavior is the same as that of Eq. (2) with α set equal to $n + \frac{1}{2}$. In this case, the single-valuedness of ψ is evident. In general, however, the behavior of ψ is not so simple, since ψ does not become zero on the line $\theta = \pi$.

5. DISCUSSION OF SIGNIFICANCE OF RESULTS

The essential result of the previous discussion is that in quantum theory, an electron (for example) can be influenced by the potentials even if all the field regions are excluded from it. In other words, in a field-free multiply-connected region of space, the physical properties of the system still depend on the potentials.

It is true that all these effects of the potentials depend only on the gauge-invariant quantity $\oint \mathbf{A} \cdot d\mathbf{x} = \int \mathbf{H} \cdot d\mathbf{s}$, so that in reality they can be expressed in terms of the fields inside the circuit. However, according to current relativistic notions, all fields must interact only locally. And since the electrons cannot reach the regions where the fields are, we cannot interpret such effects as due to the fields themselves.

In classical mechanics, we recall that potentials cannot have such significance because the equation of motion involves only the field quantities themselves. For this reason, the potentials have been regarded as purely mathematical auxiliaries, while only the field quantities were thought to have a direct physical meaning.

In quantum mechanics, the essential difference is that the equations of motion of a particle are replaced by the Schrödinger equation for a wave. This Schrödinger equation is obtained from a canonical formalism, which cannot be expressed in terms of the fields alone, but which also requires the potentials. Indeed, the potentials play a role, in Schrödinger's equation, which is analogous to that of the index of refraction in optics, The Lorentz force $[e\mathbf{E} + (e/c)\mathbf{v}\times\mathbf{H}]$ does not appear anywhere in the fundamental theory, but appears only as an approximation holding in the classical limit. It would therefore seem natural at this point to propose that, in quantum mechanics, the fundamental physical entities are the potentials, while the fields are derived from them by differentiations.

The main objection that could be raised against the above suggestion is grounded in the gauge invariance of the theory. In other words, if the potentials are subject to the transformation $\mathbf{A}_\mu \to A_\mu' = A_\mu + \partial\psi/\partial x_\mu$, where ψ is a continuous scalar function, then all the known physical quantities are left unchanged. As a result, the same physical behavior is obtained from any two potentials, $A_\mu(x)$ and $A_\mu'(x)$, related by the above transformation. This means that insofar as the potentials are richer in properties than the fields, there is no way to reveal this additional richness. It was therefore concluded that the potentials cannot have any meaning, except insofar as they are used mathematically, to calculate the fields.

We have seen from the examples described in this paper that the above point of view cannot be maintained for the general case. Of course, our discussion does not bring into question the gauge invariance of the theory. But it does show that in a theory involving only local interactions (e.g., Schrödinger's or Dirac's equation, and current quantum-mechanical field theories), the potentials must, in certain cases, be considered as physically effective, even when there are no fields acting on the charged particles.

The above discussion suggests that some further development of the theory is needed. Two possible directions are clear. First, we may try to formulate a nonlocal theory in which, for example, the electron could interact with a field that was a finite distance away. Then there would be no trouble in interpreting these results, but, as is well known, there are severe difficulties in the way of doing this. Secondly, we may retain the present local theory and, instead, we may try to give a further new interpretation to the poten-

[11] See, for example, D. Bohm, *Quantum Theory* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1951).

tials. In other words, we are led to regard $A_\mu(x)$ as a physical variable. This means that we must be able to define the physical difference between two quantum states which differ only by gauge transformation. It will be shown in a future paper that in a system containing an undefined number of charged particles (i.e., a superposition of states of different total charge), a new Hermitian operator, essentially an angle variable, can be introduced, which is conjugate to the charge density and which may give a meaning to the gauge. Such states have actually been used in connection with recent theories of superconductivity and superfluidity[12] and we shall show their relation to this problem in more detail.

ACKNOWLEDGMENTS

[12] See, for example, C. G. Kuper, *Advances in Physics*, edited by N. F. Mott (Taylor and Francis, Ltd., London, 1959), Vol. 8, p. 25, Sec. 3, Par. 3.