Pramāṇa – J. Phys., Vol. 35, No. 5, November 1990, pp. 421–428. © Printed in India.

Aharonov–Bohm effect from five dimensional space-time

S BISWAS
Department of Physics, University of Kalyani, Kalyani 741 235, India

MS received 3 October 1989; revised 13 August 1990

Abstract. We present an alternative non-quantum mechanical description of Aharonov–Bohm effect.

Keywords. Aharonov–Bohm effect; Kaluza–Klein; interference.

PACS No. 04·50

1. Introduction

Aharonov–Bohm (AB) effect is still now an area of active debate, not with the effect itself but with the interpretation. The usual accepted quantum mechanical explanation seems to imply a violation of the fundamental principle of quantum mechanics (Bocchieri and Loinger 1984). In Aharonov experiment electrons are passed around a confining magnetic flux between the slits in a two slit experiment (figure 1). The interference pattern observed on the screen is explained due to superposition of two parts of the wave function

$$\psi = \psi_1 + \psi_2 \tag{1}$$

where ψ_1 now denotes the wave function that describes the electron transversing the path 1, and ψ_2 the path appropriate to path 2. The usual substitution

$$\mathbf{P} \to \mathbf{P} + \frac{e}{c}\mathbf{A}(\mathbf{r}, t) \tag{2}$$

makes the Schrödinger equation violate the principle of gauge invariance as the Hamiltonian is changed according to (Gasirowicz 1974)

$$\frac{1}{2m}\left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A}\right)^2 \to \frac{1}{2m}\left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A} + \frac{e}{c}\nabla f\right)^2$$

under the gauge transformation

$$\mathbf{A} \to \mathbf{A} + \nabla f(r, t).$$

To restore the gauge invariance the transformation

$$\psi(\mathbf{r}, t) \to \exp[(-ie/\hbar c) f(r, t)]\psi(r, t), \tag{3}$$

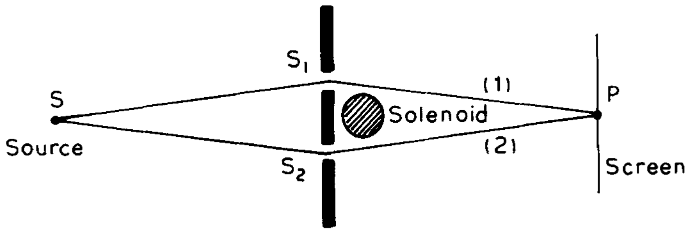

Figure 1.

is made in the Schrödinger equation and for the region $F_{\mu\nu}=0$, one takes $\mathbf{A}=\nabla f$. There results a change of wave function as

$$\psi_1 \to \psi_1 \exp\left[(-ie/\hbar c)\int_1 \mathbf{A}\cdot\mathbf{dr}\right], \tag{4}$$

$$\psi_2 \to \psi_2 \exp\left[(-ie/\hbar c)\int_2 \mathbf{A}\cdot\mathbf{dr}\right], \tag{5}$$

so that

$$\psi = (\psi_1 \exp(ie\Phi/\hbar c) + \psi_2)\exp\left[(-ie/\hbar c)\int_2 \mathbf{A}\cdot\mathbf{dr}\right], \tag{6}$$

where

$$\Phi = \int_1 \mathbf{dr}'\cdot\mathbf{A}(\mathbf{r}',t) - \int_2 \mathbf{dr}'\mathbf{A}(\mathbf{r}',t) \neq 0 \tag{7}$$

now introduces a relative phase resulting in a change in interference pattern (Gasiorowicz). On equation (6) we note that if ψ_1 and ψ_2 are single valued, it seems then, as if, ψ will be multi-valued. However, this is not so. The paths in the integral $\exp[-ie/\hbar c\int_2\ldots]$ are taken only one side (i.e. not encircling the flux containing region) so that the expressions like (7) will be zero i.e. ψ is single valued. For the paths encircling the region $F_{\mu\nu}\neq 0$,

$$\oint A_\mu \mathrm{d}x^\mu = \text{flux enclosed} \neq 0$$

and hence $A_\mu = \partial_\mu \theta$ with θ multiple valued function resulting in a multiply-connected space (Salam 1976). This was the suggestion of Aharonov and Bohm. One may then argue that the phase change in wave function as in (3) is an artefact of Schrödinger equation. As the substitution (2) as well as the gauge invariance is a well established fact of electrodynamics, one needs an explanation surpassing the argument mentioned above.

Now Bocchirie's argument is that if $\hat{F}$ be an operator corresponding to a physical observable, the expectation value in the superposed state is given by

$$(\psi,\hat{F}\psi) = (\psi_1, F\psi_1) + (\psi_2,\hat{F}\psi_2) + 2\mathrm{Re}\{\exp[-ie\Phi/\hbar](\psi_1,\hat{F}\psi_2)\}. \tag{8}$$

Dependence of the expectation value on Φ makes the thing worse. For example, if we take $\hat{F}=\hat{H}$, the Hamiltonian operator, then the mean value of the energy of the electrons be dependent on the confined flux, even though the external source does not perform any work (since $F_{\mu\nu}=0$). Hence the violation of fundamental principle. Attempts are then made (Ferrari and Griego 1986) to interpret the effect in a way so that the solution of Schrödinger equation is not required in the process. A suitable explanation might then arise from spacetime description where electromagnetic potentials arise as metrical phenomenon. Ferrari and Griego (1986) obtained a simple explanation of the AB effect using De Broglie's hypothesis plus a connection between the proper flight-time of the electrons and the electromagnetic potentials using five dimensional Kaluza–Klein description.

In this paper we adopt a similar approach. It is found that Ferrari and Griego's approach will not lead to the explanation of the effect. Redefinition of the De Broglie's wave vector in presence of e.m. field is the clue in obtaining the flux dependent phase difference resulting in the AB effect. This is a new result of this paper. We proceed with 5D Kaluza–Klein theory and construct the Lagrangian of a material particle that parallels the construction in four dimensional Minkowski space-time. We obtain solutions independent of Schrödinger equation showing a correct phase change of wave function thus surpassing Bocchirie's argument.

2. Five dimensional Kaluza–Klein space-time

Let us start with the interval in five dimension,

$$d\tau^2 = g_{ij}dx^i dx^j \tag{9}$$

$i, j = 0, 1, 2, 3, 5$. As our everyday experience is a four dimensional world, the metric tensor g_{ij} should satisfy the condition

$$g_{ij,5} = 0 \tag{10}$$

i.e., it has a vanishing derivative with respect to the newly introduced co-ordinate x_5. To work with (9) one must find the effective four dimension metric tensor so that the co-ordinate x^μ characterizes the usual four dimensional space-time. This requires that between different frames, x^μ transforms as

$$x^\mu = f^\mu(\hat{x}^\nu). \tag{11}$$

Now in order to satisfy (10), x^5 must transform as

$$x^5 = \hat{x}^5 + f^5(\hat{x}^\nu). \tag{12}$$

Thus the quantity

$$\gamma_{\mu\nu} = g_{\mu\nu} - \beta^2 g_{\mu 5} g_{\nu 5} \tag{13}$$

is independent of the x^5 co-ordinate. Taking $g_{55} = 1$ the interval (9) will now read as

$$\begin{aligned} d\tau^2 &= (\gamma_{\mu\nu} - \beta^2 g_{\mu 5} g_{\nu 5})dx^\mu dx^\nu + 2\beta g_{\mu 5} dx^\mu dx^5 + (dx^5)^2 \\ &= (dx^5 + \beta A_\mu dx^\mu)^2 + \gamma_{\mu\nu} dx^\mu dx^\nu, \end{aligned} \tag{14}$$

where we have taken as usual $g_{\mu 5} = g_{5\mu} = A_\mu$, the four vector potential.

The equations of motion of a material particle in the 5D space are obtained from the variation principle

$$\delta \int d\tau = 0. \tag{15}$$

Introducing a parameter s, we write

$$\mathrm{d}\tau = L\left(x^i, \frac{\mathrm{d}x^i}{\mathrm{d}s}\right)\mathrm{d}s, \tag{16}$$

where

$$L\left(x^i, \frac{\mathrm{d}x^i}{\mathrm{d}s}\right) = \left[\left(\frac{\mathrm{d}x^5}{\mathrm{d}s} + \beta A_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}s}\right)^2 + \gamma_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}x\right]^{\frac{1}{2}}. \tag{17}$$

Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\partial L}{\partial x^i/\mathrm{d}s}\right) - \frac{\partial L}{\partial x^i} = 0, \tag{18}$$

now give

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}x^5}{\mathrm{d}\tau} + \beta A_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) = 0, \tag{19}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\gamma_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right) - \frac{1}{2}\frac{\partial \gamma_{\nu\alpha}}{\partial x^\mu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau} = \beta\left(\frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}. \tag{20}$$

Equation (20) will represent the motion of a charge particle in e.m. field and gravitational field with the identification

$$F_{\mu\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}, \tag{21}$$

$$\beta = (q/m_0), \tag{22}$$

where m_0 is the rest energy of the particle. Equation (19) now gives

$$\frac{\mathrm{d}x^5}{\mathrm{d}\tau} + (q/m_0)A_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = k, \tag{23}$$

where K is constant along the trajectory. We will now show that equation (23) is a restatement of gauge invariance (dynamical outcome of our formulation). There are two ways of looking at (23). If $F_{\mu\nu}$ describes the e.m. field in four dimension then the gauge invariance

$$A_\mu \to A_\mu - \frac{\partial f}{\partial x^\mu}, \tag{24}$$

is inherent in the description. So to satisfy (23), x^5 must transform as

$$x^5 \to x^5 + f(x^\mu)(q/m_0). \tag{25}$$

This is explicitly the condition (12) mentioned earlier. In the other way, the transformation (25) is reflected in four space-time as gauge invariance. Thus a co-ordinate transformation in internal space

$$(x^\mu, x^5) \to (x^\mu, x^5 + (q/m_0)f), \tag{26}$$

keeps the photon massless i.e., gauge invariance is maintained in four dimension. Thus even if $F_{\mu\nu} = 0$ in a given region of space-time, the five co-ordinate might change, as in (25), to couple the charged particle at x^μ as if it is moving through a potential field $A^\mu = \partial_\mu f(x^\mu)$. The formulation thus provides the way the electron is affected, even if $F_{\mu\nu} = 0$, in the five dimensional space-time.

In Ferrari and Griego's work, (14) is written using (23) as

$$d\tau = K(dx^5 + \beta A_\mu dx^\mu) + \gamma_{\mu\nu}\frac{dx^\nu}{d\tau}dx^\nu. \tag{27}$$

The proper time between two paths in AB experiment is then

$$\tau_1 = \frac{q}{m_0}\int_1 A^\mu dx^\mu + \int_1 \gamma_{\mu\nu}\frac{dx^\mu}{d\tau}dx^\nu, \tag{28}$$

$$\tau_2 = \frac{q}{m_0}\int_2 A^\mu dx^\mu + \int_2 \gamma_{\mu\nu}\frac{dx^\mu}{d\tau}dx^\nu, \tag{29}$$

assuming that along the trajectories in four dimensional space-time $x^5 = \text{constant}$. In Ferrari and Griego's work the first term of (27) was neglected and using De Broglie's hypothesis

$$m_0\frac{dx^\mu}{d\tau} = \hbar K^\mu, \tag{30}$$

it was shown that a phase difference

$$\Delta\varphi = \frac{q}{\hbar}\Phi + \oint \mathbf{k}\cdot\mathbf{dr}, \tag{31}$$

where Φ is the flux. It is thus claimed by Ferrari and Griego that the (31) will explain the AB phase change. We will now show that the phase change observed by these authors is apparent and (31) cannot be obtained as demanded by them. The above authors while deriving (31) neglected the first term of (27) on the assumption that $x^5 = \text{constant}$ along the paths 1 and 2 as electron travels in four dimensional space-time. As a result they obtained (31) coming out of the first term of eqs (28) and (29). But in order to attach a physical significance to 5D space-time, we must accept that at each point of the electron path the space-time is five dimensional and the assumption $x^5 = \text{constant}$ can in no way be justified. Furthermore using (23) we can substitute $dx^5 + \beta A_\mu dx^\mu = k d\tau$ in (27) resulting in

$$d\tau^2 = \tilde{\gamma}_{\mu\nu} dx^\mu dx^\nu, \tag{32}$$

where

$$\gamma_{\mu\nu} = (1/1 - K^2)\gamma_{\mu\nu}$$

so that

$$\tau = \int \tilde{\gamma}_{\mu\nu}\frac{dx^\mu}{d\tau}dx^\nu. \tag{33}$$

Thus we see that the proper time τ_1 and τ_2 will not now show any contribution from e.m. field as is apparently shown in (28) and (29) and hence the phase change apparently shown in (31) will be absent. In the next section we show how the effect still persists in the Kaluza–Klein description provided one modifies the relation (30) taking into account the special theory of relativity requirement viz., eq. (38).

3. Phase change in Kaluza–Klein description

In the proper frame of the electron, the wave function of the electron is written as

$$\psi(x^\mu) = \psi_0 \exp(im_0/\hbar)\tau, \tag{34}$$

where

$$\tau = \int \tilde{\gamma}_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \mathrm{d}x^\nu. \tag{35}$$

To surpass Bocchieri's argument we note that wave particle dualism does not necessarily imply to accept the full quantum mechanical formulation. We will come to this point later. In view of the relation (32), we consider (33) as a wave associated with a particle moving freely in four dimensional space-time even in the presence of electromagnetic potentials $A_\mu = f(x^\mu), \mu$. This peculiar property arises due to (26), (23) i.e., due to gauge invariance of electromanetic field. So in our approach, using (34) and (35) we get

$$\psi_1(x^\mu) = \psi_{10} \exp\left[i\frac{m_0}{\hbar} \int \gamma_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \mathrm{d}x^\nu \right], \tag{36}$$

$$\psi_2(x^\mu) = \psi_{20} \exp\left[i\frac{m_0}{\hbar} \int \gamma_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \mathrm{d}x^\nu \right]. \tag{37}$$

The factor $(1 - k^2)$ is absorbed in m_0. Now we note the wave vector associated with the wave described by (34) is not given by the relation (30) but is written as

$$m_0 \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \hbar \left(k^\mu + \frac{q}{\hbar c} A^\mu \right). \tag{38}$$

Use of generalized wave vector in (37) is a necessary requirement of special theory of relativity. The substitutions (33) and (38) apparently show that the role of the fifth dimension be now absent in the description. However this is not so. The constraint (23) demands the transformation (26) which in turn introduces a potential $A^\mu = \partial_\mu f(x^\mu)$ even if $F_{\mu\nu} = 0$, the condition maintained in AB experiment. Thus the substitution (38) tacitly takes into account the role of the fifth dimension in the picture. The usual minimal substitution (2) then finds an explanation through (23). In four dimensional language the gauge transformation as well as the transformation (2) need a phase transformation in the wave function. It may be mentioned that starting from a four dimensional interval like (32) one can arrive at the AB phase shift with the substitution (38) obscuring the role of fifth dimension. In such a case A^μ has no physical connection with e.m. potential as well as the gauge transformation be then totally absent in the description. The situation is cured only through the Kaluza–Klein description via eqs (23) and (38). Equations (23) and (25) are totally absent in four dimensional description. These are essential in order to have gauge invariance as well as a potential like $\partial_\mu f$ as is suggested by Aharonov and Bohm.

Now assuming $f(x^\mu)$ time independent

$$(m_0/\hbar)\int_1 \gamma_{\mu\nu}\frac{dx^\mu}{d\tau}dx^\nu = \int_1 \gamma_{\mu\nu}\left(k^\mu + \frac{q}{\hbar c}A^\mu\right)dx^\nu$$
$$= \int \omega\, dt - \int \mathbf{k}\cdot\mathbf{dr} - \frac{q}{\hbar c}\int_1 \mathbf{A}\cdot\mathbf{dr}.$$

Similar result will hold for path 2. It should be pointed out that we have not solved the Schrödinger equation to arrive at the above equation. At the point P on the screen

$$\psi = \psi_1 + \psi_2 = \exp\left[-\frac{iq}{\hbar c}\int \mathbf{A}\cdot\mathbf{dr} - i\int \mathbf{k}\cdot\mathbf{dr} - i\int \omega dt\right]$$
$$\times\left\{\psi_{10} + \psi_{20}\exp\left[i\left(\frac{q}{\hbar c}\Phi + \oint \mathbf{k}\cdot\mathbf{dr}\right)\right]\right\}$$

where

$$\Phi = \int_{s,1}^{P} \mathbf{A}\cdot\mathbf{dr} - \int_{s,2}^{P} \mathbf{A}\cdot\mathbf{dr} \oint \mathbf{K}\cdot\mathbf{dr} = \int_1 \mathbf{K}\cdot\mathbf{dr} - \int_2 \mathbf{K}\cdot\mathbf{dr}.$$

The phase difference of the two waves at P is then

$$\Delta\varphi = \frac{q}{\hbar c}\Phi + \oint \mathbf{k}\cdot\mathbf{dr}. \tag{39}$$

Equation (39) now permits us to explain the AB effect. The occurrence of WKB-term $\oint \mathbf{k}\cdot\mathbf{dr}$ is quite natural. It is the phase picked up by electron due to additional path length. However, one may be tempted to indicate the presence of interference pattern due to this path difference between the two electron paths. But it is ruled out on account of the fact that in AB experiment a change in the strength of the confined flux within the solenoid changes the observed interference pattern and is usually measured in experiment.

The controversy raised by few people regarding the existence of the effect itself (Dewilt 1962; Roy 1980; Bocchieri *et al* 1978, 1979, 1980) is best answered in favour of the effect observed by electron holography experiment (Tonomura *et al* 1979, 1982). In Mollenstedt group experiment (Tonomura 1982) two electron waves from the same source travel around the solenoid and are overlapped coherently to cause interference fringes on the film. The magnetic flux around the solenoid is then changed varying the coil current '*i*' resulting in frings shift. Thus when one measures the fringe shift due to two values of magnetic field, the WKB term cancels out as the electron's path remains unaffected due to change of magnetic flux within the solenoid. Our explanation then coincides with the usual explanation that one obtains from the solution of Schrödinger equation. The five dimensional constraint (23) in one way has an inbuilt structure of gauge invariance, as well as on the other way, generate the AB phase without having solved the Schrödinger equation.

Now returning to Bocchierie's objection we note that the explanation that we have put in this paper allows us to measure any physical observable according to special theory of relativity. What we measure in AB experiment is the wave function itself i.e., the relative phase of the wave function. However, this does not violate the quantum postulate (Pradhan 1985). In order to surpass Bocchierie's argument one has to establish the existence of the effect without solving Schrödinger equation. This has been done in this paper from five dimensional Kaluza–Klein description.

Acknowledgements

The author thanks Profs P Dasgupta and P Rudra of the University of Kalyani for discussion and the referee for some useful suggestions.

References

Bocchieri *et al* 1978 *Nuovo Cimento* **A7** 475
Bocchieri *et al* 1979 *Nuovo Cimento* **51** 1
Bocchieri P and Loinger 1982 *Lett. Nuovo Cimento* **35** 469
Bocchieri P and Loinger 1984 *Lett. Nuovo Cimento* **39** 148
Bocchieri P *et al* 1964 *Nuovo Cimento* **56, 55, 51** 12
Dewilt D S 1962 *Phys. Rev.* **125** 2189
Ferrari J A and Griego J 1985 *Nuovo Cimento* **B96** 41
Gasiorowicz S 1974 *Quantum Physics* (New York: Wiley) Ch. 13
Pradhan T 1985 *in Proceedings of the National Seminar on sixty years of quantum mechanics* (Sàntiniketan)
Roy S M 1980 *Phys. Rev. Lett* **33** 111
Salam A 1976 in *Physics and Contemporary needs* (ed.) Rlazuddin, (New York: Plenum Press) p. 308
Tonomura A *et al* 1979 *J. Electron Microsc.* **28** 1
Tonomura A *et al* 1982a *Phys. Rev. Lett* **48** 1443
Tonomura A *et al* 1982b *Lecture Notes in Physics* (ed.) K Kikkawa *et al* (Springer Verlag) Vol. 176 p. 146