Gauge invariance of the Aharonov-Bohm effect

1984 J. Phys. A: Math. Gen. 17 2973

(http://iopscience.iop.org/0305-4470/17/15/015)

J. Phys. A: Math. Gen. 17 (1984) 2973–2982. Printed in Great Britain

Gauge invariance of the Aharonov–Bohm effect

M Babiker and R Loudon
Physics Department, Essex University, Colchester CO4 3SQ, UK

Received 25 May 1984

Abstract. It is argued that much of the recent controversy over the Aharonov–Bohm effect has been fuelled by the widespread consideration of solenoids of infinite length; in this limit, the field and the vector potential have somewhat anomalous properties and it appears superficially that the Aharonov–Bohm phase shift can be changed by a gauge transformation. This incorrect impression is here removed by considering a solenoid of finite length and by carefully distinguishing the longitudinal and transverse parts of the vector potential. It is shown that the phase shift depends only on the transverse part of the vector potential and it cannot be changed by a gauge transformation. The nature of the difficulty that occurs in the limit of a solenoid of infinite length is examined in detail. The finite-length theory given here also provides guidelines for the design of experiments on the Aharonov–Bohm effect.

1. Introduction

The Aharonov–Bohm effect (Aharonov and Bohm 1959, Ehrenberg and Siday 1949) concerns the phase difference, proportional to

$$\oint \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l},$$

between the paths from source to screen followed by the two beams in an electron interference experiment. Here $\boldsymbol{A}$ is the magnetic vector potential, and the line integral is taken round the closed loop obtained by subtracting one vector path from the other. The phase difference, and thus the positions of the interference fringes, depend upon the size of the magnetic flux Φ enclosed by the loop (see Feynman *et al* 1964 and Erlichson 1970 for readable accounts). Most of the interest in the effect derives from the circumstance that it appears to depend upon the values of $\boldsymbol{A}$ rather than the magnetic field $\boldsymbol{B}$ at the electron positions. It should be emphasised that the electron beams are usually arranged to pass only through regions of sufficiently small $\boldsymbol{B}$ that the direct influence of the Lorentz force on the electron motion can be neglected.

There has recently been some controversy over the extent to which the vector potential $\boldsymbol{A}$ experienced by the electron beams, and hence their phase difference, can be modified, or even completely transformed away, by means of gauge transformations. This possibility is supported by Bocchieri and Loinger (1978) and by Bocchieri *et al* (1979), and it has accordingly been claimed by Bocchieri and his coworkers that the Aharonov–Bohm effect does not exist. Less drastically, Roy (1980) and Home and Sengupta (1983) purport to show that the Aharonov–Bohm phase shift does exist, but that its magnitude can be entirely expressed in terms of the magnetic field $\boldsymbol{B}$ in regions of space accessible to the electron beams. The conventional view, that the Aharonov–Bohm effect does exist and is determined by the values of $\boldsymbol{A}$ rather than $\boldsymbol{B}$ sampled by the electron beams, has been reaffirmed by Klein (1979), Zeilinger (1979) and Bohm and Hiley (1979) using various electromagnetic and quantum mechanical arguments.

The discussions of the theory of the Aharonov–Bohm effect mentioned above almost all assume that the magnetic field is produced by a solenoid of infinite length placed between the two electron paths. We believe that this procedure is misleading and has in fact led to some of the misunderstandings concerning the nature of the effect. Not only must the solenoid be finite in any experimental realisation, but also, and more seriously, the vector potential of an infinite solenoid is somewhat pathological and needs careful treatment. The discussion that follows is based upon the potential of a finite-length solenoid, which is of course the quantity needed for comparison with observations, and which also allows analysis of the infinite-length limit. Our results support the conclusion of Zeilinger (1979) and Bohm and Hiley (1979) that the Aharonov–Bohm phase shift cannot be transformed away, and they give an unambiguous prescription for fixing the gauge of the vector potential.

2. Potential theory

In this section we briefly review some more or less standard results of electromagnetic theory. Maxwell's equations in the magnetostatic limit give

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} \tag{1}$$

and

$$\nabla \cdot \boldsymbol{B} = 0, \tag{2}$$

where the current $\boldsymbol{J}$ associated with a time-independent charge density satisfies

$$\nabla \cdot \boldsymbol{J} = 0. \tag{3}$$

The vector potential $\boldsymbol{A}$ is introduced in the usual manner by

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \tag{4}$$

where according to the Helmholtz theorem (see for example Morse and Feshbach 1953) $\boldsymbol{A}$ can be separated into a longitudinal part $\boldsymbol{A}^{\parallel}$ and a transverse part $\boldsymbol{A}^{\perp}$ that satisfy

$$\boldsymbol{A} = \boldsymbol{A}^{\parallel} + \boldsymbol{A}^{\perp}, \tag{5}$$

$$\nabla \times \boldsymbol{A}^{\parallel} = \boldsymbol{0} \tag{6}$$

and

$$\nabla \cdot \boldsymbol{A}^{\perp} = 0. \tag{7}$$

The procedure for making this separation is described by Power (1964). Clearly from (4) and (6),

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}^{\perp}, \tag{8}$$

and substitution into (1) gives

$$\nabla^2 \boldsymbol{A}^{\perp} = -\mu_0 \boldsymbol{J}. \tag{9}$$

According to Stokes theorem (e.g. Morse and Feshbach 1953), any vector field $\boldsymbol{F}$ satisfies

$$\oint \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{l} = \int \nabla\times\boldsymbol{F}\cdot \mathrm{d}\boldsymbol{S} \tag{10}$$

where the integral on the right is taken over any surface that is bordered by the closed-line path of the integral on the left. Thus for the two parts of the vector potential, Stokes theorem gives

$$\oint \boldsymbol{A}^{\parallel}\cdot \mathrm{d}\boldsymbol{l} = 0 \tag{11}$$

and

$$\oint \boldsymbol{A}^{\perp}\cdot \mathrm{d}\boldsymbol{l} = \int \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{S}. \tag{12}$$

A gauge transformation replaces an initial vector potential $\boldsymbol{A}$ by a new potential

$$\boldsymbol{A}' = \boldsymbol{A} - \nabla\chi, \tag{13}$$

where χ is some function of position. Since

$$\nabla\times\nabla\chi \equiv \boldsymbol{0}, \tag{14}$$

the gauge transformation changes only the longitudinal part of $\boldsymbol{A}$, and $\boldsymbol{A}^{\perp}$ is unchanged. Thus the magnetic field $\boldsymbol{B}$ is invariant under a gauge transformation. It is nevertheless sometimes useful to consider the same problem in a variety of gauges, as for example in the interaction of electromagnetic radiation with charged particles (Babiker and Loudon 1983). However, in the present problem it is convenient to take

$$\boldsymbol{A}^{\parallel} = \boldsymbol{0} \tag{15}$$

throughout.

3. Field of a finite solenoid

Consider a circular solenoid of radius a and length $2L$ that lies symmetrically on either side of the xy plane with its axis coincident with the z axis. The solenoid carries a current I and has n turns per unit length, n being sufficiently large that the separation between adjacent turns is much smaller than a. We use cylindrical polar coordinates (ρ, ϕ, z) and consider the fields only in the $z=0$ plane. Then the vector potential has only a ϕ component that depends only on ρ, and with the help of a result of Jackson (1975) we find

$$A_{\phi}^{\perp}(\rho) = \frac{2B_0 a}{\pi}\int_0^L \mathrm{d}z \frac{(2-k^2)K(k) - 2E(k)}{k^2[(a+\rho)^2+z^2]^{1/2}}, \tag{16}$$

where $K(k)$ and $E(k)$ are complete elliptic integrals (Gradshteyn and Ryzhik 1965), k is a dimensionless variable given by

$$k^2 = 4a\rho/[(a+\rho)^2+z^2], \tag{17}$$

and

$$B_0 = \mu_0 I n \tag{18}$$

is a convenient normalising field magnitude. This result holds both inside and outside the solenoid.

It is clear by symmetry that $\boldsymbol{B}$ has only a z component in the $z=0$ plane; its magnitude is obtained from (8) as

$$B_z(\rho)=\frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho A_\phi^\perp)=\frac{A_\phi^\perp}{\rho}+\frac{\partial A_\phi^\perp}{\partial\rho}. \tag{19}$$

With the vector potential given by (16), it is only possible to obtain the complete variation of magnetic field with position by numerical integration. Figure 1 shows an example of such a calculation for the case $L=2a$. It is seen that the magnetic field takes its largest value at the centre of the solenoid, falls off smoothly for increasing ρ with no particular feature at $\rho=a$, and changes sign at some radial distance, denoted by ρ_0, well outside the solenoid. The inset to the figure shows the detailed variation of the field in the vicinity of its sign change, which occurs at $\rho=\rho_0=2.6a$ for $L=2a$.

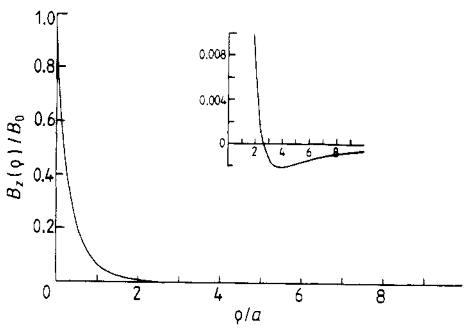

Figure 1. Variation of magnetic field with radial distance ρ in the $z=0$ plane of a solenoid of length $2L$ where $L=2a$. The inset shows an enlarged view of the variation in the vicinity of the sign change at ρ_0.

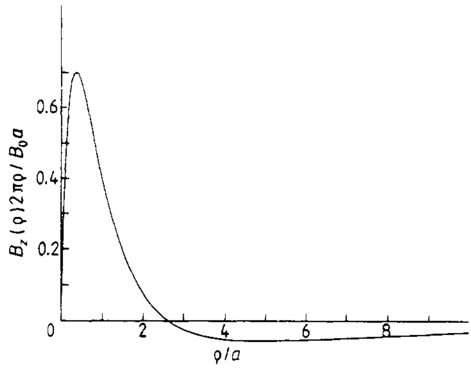

Figure 2. Variation of $B_z(\rho)2\pi\rho/B_0a$ with radial distance ρ in the $z=0$ plane of a solenoid of length $2L$ where $L=2a$. For a fixed value of ρ the total flux enclosed by a disc of radius ρ is the net area under the curve between the origin and that value of ρ. For equation (25) to be valid the positive and negative areas must be equal.

Although the integrand in (16) is complicated for general values of ρ, there are fortunately some simple analytical approximations that hold in the limits $\rho \ll a$ and $\rho \gg a$. For both cases $k \ll 1$, and the standard expansions of the complete elliptic integrals (Gradshteyn and Ryzhik 1965) enable the integration to be performed with the result

$$A_\phi^\perp(\rho)=\frac{B_0a^2}{2}\frac{\rho}{(a+\rho)^2}\frac{L}{[(a+\rho)^2+L^2]^{1/2}}\,. \tag{20}$$

The corresponding magnetic field is

$$B_z(\rho)=\frac{B_0a^2L}{2}\frac{2aL^2+(2a-\rho)(a+\rho)^2}{(a+\rho)^3[(a+\rho)^2+L^2]^{3/2}} \qquad (\rho \ll a \text{ and } \rho \gg a). \tag{21}$$

The fields at the centre of the solenoid are

$$A_\phi^\perp(0)=0 \qquad \text{and} \qquad B_z(0)=B_0L(a^2+L^2)^{-1/2}. \tag{22}$$

The radial distance at which the magnetic field vanishes can also be determined from the approximate expression (21) in the case where the solenoid length is much larger than its radius. The numerator of (21) then vanishes at

$$\rho_0\approx(2aL^2)^{1/3} \qquad (L\gg a); \tag{23}$$

this distance is much larger than a.

Finally, we note that in the limit of very large radial distances, (20) and (21) give

$$\left.\begin{aligned} A_\phi^\perp(\rho)&=B_0a^2L/2\rho^2 \\ B_z(\rho)&=-B_0a^2L/2\rho^3 \end{aligned}\right\} \quad \text{for } \rho\gg a \text{ and } \rho\gg L. \tag{24}$$

The integral of the vector potential round a circle of very large radius therefore vanishes as $1/\rho$, and it follows from Stokes theorem (12) that

$$\lim_{\rho\to\infty} 2\pi\rho A_\phi^\perp(\rho)=\int_0^\infty B_z(\rho)2\pi\rho\,\mathrm{d}\rho=0 \qquad \text{(correct)}. \tag{25}$$

The total magnetic flux through the $z=0$ plane therefore vanishes, a consequence of the Maxwell equation (2) which requires the magnetic field $\boldsymbol{B}$ to be solenoidal. The magnetic flux lines are closed, and they therefore pass twice through the plane, once in the vicinity of the solenoid and once in the opposite direction at some radial distance larger than ρ_0. Figure 2 shows the equal positive and negative areas spanned by the integrand of (25) for the case $L=2a$.

4. Aharonov–Bohm effect with semicircular paths

Consider an Aharonov–Bohm experiment in which the electron paths have the forms of semicircles concentric with the axis of a solenoid, such that the phase difference is proportional to an integral of $A_\phi^\perp$ around a circle of radius R. The required integral represents the magnetic flux passing through a disc of radius R and with the use of Stokes theorem (12), it can be written

$$\Phi(R)\equiv 2\pi R A_\phi^\perp(R)=\int_0^R B_z(\rho)2\pi\rho\,\mathrm{d}\rho. \tag{26}$$

Experiments with the circular configuration assumed here can in principle be realised with the use of Josephson-junction interferometers (Jaklevic *et al* 1965, Tonomura *et al* 1982).

Figure 3 shows the variation of the flux (26) with R for a variety of solenoid lengths. The maximum in each curve is seen by differentiation of (26) to occur at the radius ρ_0 for which $B_z=0$. For long solenoids ρ_0 is given to a good approximation by (23), and this radius of maximum flux has a magnitude that lies between the diameter and length of the solenoid. The phase shift in an Aharonov–Bohm experiment with semicircular paths is thus a maximum for semicircle radius ρ_0. We note that the magnetic field $\boldsymbol{B}$ is then zero along the entire electron paths. This configuration thus displays in its most acute form the essence of the Aharonov–Bohm situation in which the vector potential $\boldsymbol{A}^\perp$ has significant values but the magnetic field B_z is insignificant

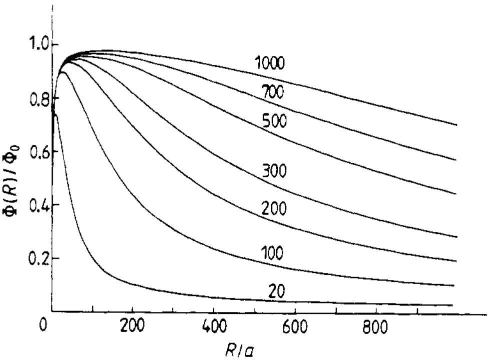

Figure 3. Magnetic flux through a disc of radius R as a function of R for various values of solenoid length $2L$. The graphs shown are for $L/a = 20$, 100, 200, 300, 500, 700 and 1000.

at the electron positions. The enclosed flux is of course entirely determined by the magnetic field, but only by its magnitudes at positions not sampled by the electrons. For very long solenoids the maximum enclosed flux tends to the value

$$\Phi_0 \equiv \Phi(\rho_0) = \pi\mu_0 Ina^2 = \pi a^2 B_0 \qquad (L/a \to \infty). \tag{27}$$

It is seen from figure 3 that the flux $\Phi(R)$ falls rapidly to zero for $R < \rho_0$. The fall-off for $R > \rho_0$ is much more gradual and the zero limiting value of Φ at infinite radial distance given by (25) is only approached when R is very much greater than the solenoid length.

Expressions used in earlier discussions of the Aharonov–Bohm effect are obtained from appropriate limits of (20) and (21). Thus Roy (1980) gives a vector potential for a solenoid of zero diameter but finite flux that agrees at $z = 0$ with the $a \to 0$ limit of (20) when Φ_0 is held at a constant value. We note from (21) that the magnetic field is negative in this case for all $\rho > 0$, and the vanishing of the total flux through the $z = 0$ plane in accordance with (25) is secured by the positive contribution of a singular magnetic field at $\rho = 0$.

Most discussions assume however that L is much larger than any other dimensions of the system, when (20) and (21) reduce to

$$A_\phi^\perp(\rho) = \tfrac{1}{2}B_0 a^2 \rho/(a+\rho)^2 \tag{28}$$

and

$$B_z(\rho) = B_0 a^3/(a+\rho)^3, \tag{29}$$

where both results are valid for $L \to \infty$ and $\rho \ll a$ or $\rho \gg a$. In the former case, the limiting results can be simplified still further to

$$\begin{aligned} A_\phi^\perp(\rho) &= \tfrac{1}{2}B_0\rho = \Phi_0\rho/2\pi a^2 \\ B_z(\rho) &= B_0 \end{aligned} \qquad (L \to \infty, \rho \ll a), \tag{30}$$

and in the latter case to

$$\begin{aligned} A_\phi^\perp(\rho) &= \tfrac{1}{2}B_0 a^2/\rho = \Phi_0/2\pi\rho \\ B_z(\rho) &= B_0 a^3/\rho^3 \end{aligned} \qquad (L \to \infty, \rho \gg a). \tag{31}$$

We note that these results are quite different from those obtained in (24) where the radial distance ρ is assumed to be very much larger than both the radius and length of the solenoid. In particular, the order in which the limits are taken in (31) produces a field and a vector potential that are no longer related by (19). The difficulties that can result from the use of the limiting results (28) to (31) are discussed in the following section.

Figure 4 compares the vector potentials calculated for a long solenoid with $L = 500a$ from the exact expression (16), from the approximation (28), and from the severe approximation (30) and (31). The last equations provide a poor match to the exact expression for the range of ρ plotted.

5. Field of an infinite solenoid

Most treatments of the Aharonov–Bohm effect, including the most accessible textbook account (Feynman *et al* 1964), work with the limiting forms of vector potential given in (30) and (31). We wish to point out that these forms can easily lead to confusion. Indeed, their use has led to a protracted and continuing controversy (Bocchieri and Loinger 1978, Bocchieri *et al* 1979, Home and Sengupta 1983) over the possibility of transforming away the vector potential at large radial distances ρ by means of a gauge transformation.

It is immediately clear from (29) and (31) that the limit $L \to \infty$ of an infinitely long solenoid removes the region of negative magnetic field that ordinarily occurs for sufficiently large ρ, as illustrated in figures 1 and 2. The magnetic field is everywhere positive in the limit and the integral of the magnetic field over the entire $z = 0$ plane

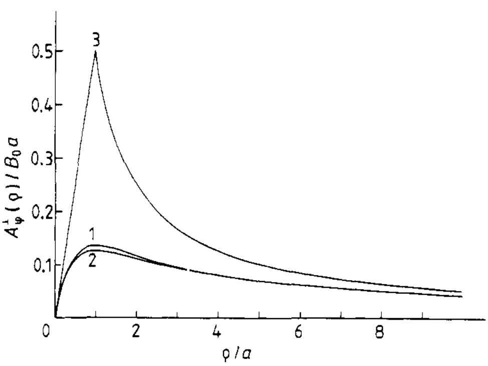

Figure 4. Comparison between the functional variations with ρ of the exact ($L = 500a$) vector potential (curve 1), the reasonable approximation (curve 2) and the extreme approximation (curve 3).

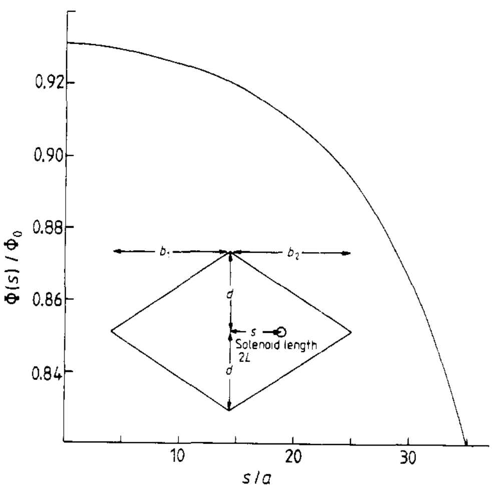

Figure 5. Variation of the flux enclosed by the electron paths in a typical electron interference experiment as a function of the position s of the solenoid. The inset shows a schematic arrangement of the experiment and defines the parameters. The parameter values used in the computation of the flux are $L = 500a$, $b_1 = b_2 = 45a$ and $d = 30a$.

on the right-hand side of Stokes theorem (12) cannot therefore vanish. Indeed, the limit (31) gives

$$\lim_{\rho\to\infty} 2\pi\rho A_\phi^\perp(\rho) = \int_0^\infty B_z(\rho) 2\pi\rho \, d\rho = \Phi_0 \qquad \text{(incorrect).} \tag{32}$$

This is in conflict with the correct result (25), which is itself a direct consequence of the solenoidal nature of the magnetic field expressed by the Maxwell equation (2). The incorrect result (32) follows from the ordering of limits in which L first tends to infinity and ρ is subsequently taken to infinity. In real experiments, L is of course always finite and the Stokes integral begins to acquire the necessary negative contributions when the radius of integration exceeds the value ρ_0 given approximately by (23).

The limiting vector potential (31) is derivable from a scalar potential

$$\chi = \Phi_0 \phi / 2\pi, \tag{33}$$

since

$$\nabla\chi = (0, \Phi_0/2\pi\rho, 0). \tag{34}$$

It follows from the identity (14) and the property (6) that the limiting $A_\phi^\perp(\rho)$ in (31) has the false appearance of a *longitudinal* vector. With this interpretation, it may then be tempting to remove $A_\phi^\perp(\rho)$ by means of a gauge transformation (13) with χ from (33) taken as the gauge function, as in the work of Bocchieri *et al* (1979) and Home and Sengupta (1983). Such a procedure is however improper. The potential (33) taken to be present over all of the $z = 0$ plane is that of a vortex line at $\rho = 0$ (see for example Morse and Feshbach 1953). Any gauge transformation that uses (33) as the gauge function has the effect of introducing a vortex line at $\rho = 0$, whereas no such vortex line is present in the field of a solenoid. It is therefore not permissible to remove $A_\phi^\perp(\rho)$ by means of a gauge transformation. The vector potential (31) only has the appearance of a longitudinal vector because of the manner in which the infinite solenoid limit has been taken. The correct potential given accurately by (16) and approximately by (20) is a transverse vector, and there is of course no question of removing it, or indeed of making any useful modifications to it, by means of gauge transformations.

The difficulties described in the present section make the use of the limiting forms (30) and (31) of the vector potential dangerous. It is, therefore, necessary to use instead the exact solution (16), some of whose numerical predictions have been illustrated in the figures, or where appropriate the approximate potential (20). Much of the interest in the Aharonov–Bohm effect is concerned with its basic principles and these can easily be obscured by the common restriction of treatments to a limit in which well established properties of magnetic fields appear to be violated.

6. Aharonov–Bohm effect with rectilinear paths

The semicircular-path Aharonov–Bohm experiment treated in § 4 has a useful configuration for discussions of the principles involved in the effect. However, practical realisations usually involve electron-beam paths that approximately make up a kite-shaped contour or the lozenge contour shown in the inset to figure 5 (see for example Möllenstedt and Bayh 1962), the detailed experimental paths being somewhat more complicated than these simple idealised shapes. Observation of the Aharonov–Bohm effect requires the detection of a shift, proportional to the enclosed flux, of a double-slit interference pattern within the unshifted single-slit diffraction envelope.

The flux Φ enclosed by the beam paths in figure 5 clearly depends on the lozenge dimensions and the coordinate s of the solenoid axis, in addition to the solenoid length $2L$. It can be evaluated by carrying out the integration on either side of Stokes theorem (12). Clearly the simplification of the left-hand side shown in (26) no longer occurs, and it is necessary to obtain the enclosed flux by numerical integration. The main part of figure 5 shows the variation of enclosed flux with the solenoid coordinate s for the lozenge dimensions shown in the inset and for a total solenoid length of $2L = 1000a$. The curve was obtained by numerical integration of the right-hand side of (12) with the magnetic field approximated by (21). The zero-field radius given by (23) for these solenoid dimensions is

$$\rho_0 \approx 80a,$$

and the magnetic flux through the contour is everywhere positive except when the solenoid approaches the end of the lozenge.

It should be noted that the total flux for $s = 0$ lies somewhat below the maximum possible value Φ_0. This is partly because the maximum value cannot be achieved for solenoids of finite length even for circular contours (see figure 3), and partly because the lozenge shaped contour excludes a further modest amount of positive flux. The percentage of total positive flux included in the contour falls as the solenoid is moved away from the centre of the lozenge, causing positive flux to move outside the contour and, for $s > 35a$, negative flux to move into the contour. However, the variation of flux with s is relatively slight for the range of s shown in figure 5. This absence of large variation occurs because, as can be seen from figure 3, a positive flux of about $0.8\Phi_0$ lies within a circular disc of radius $10a$ centred on the solenoid. Thus large changes in enclosed flux would be expected only when the solenoid lies within about $10a$ of the lozenge end, and this is confirmed by the numerical data. Nevertheless, it should be emphasised that the total flux in figure 5 always lies sufficiently below Φ_0 that the calculated values must be used in any detailed interpretation of Aharonov–Bohm experiments.

7. Conclusions

The calculations presented here show that it is straightforward to calculate the Aharonov–Bohm phase shift for a solenoid of finite length. The enclosed flux, and hence the phase shift, take maximum values at the radial distance ρ_0 for which the magnetic field vanishes. Thus an idealised Aharonov–Bohm experiment with semicircular electron paths of radius ρ_0 would show the maximum effect of the solenoid field for a case in which the magnetic field $\boldsymbol{B}$ at the electron positions is strictly zero. The principles of such an experiment display in pure form the dependence of the observed result on the magnitude of $\boldsymbol{A}$ rather than $\boldsymbol{B}$ at the electron locations. It would of course be difficult to perform the experiment in practice since the electron beams could not be exactly confined at a given radius even if semicircular trajectories could be achieved.

The apparent possibility of changing the calculated phase shift by means of gauge transformations has been shown to be a consequence of working with a limiting form of vector potential that applies to a solenoid of infinite length. The derivation of this limiting form from the finite-length theory shows that the required vector potential is entirely transverse, and the phase shift to which it gives rise cannot possibly be affected by any lawful gauge transformations, since these change only the longitudinal part of the vector potential. The transverse vector potential is fixed by the current in the solenoid and the $\boldsymbol{B}$ field to which it gives rise. The gauge-independent formulation of the Aharonov–Bohm effect has been discussed by Wódkiewicz (1984) from a quite different viewpoint.

Finally, we comment on suggestions by Roy (1980) and by Home and Sengupta (1983) that the Aharonov–Bohm phase shift can be expressed in terms of $\boldsymbol{B}$ fields accessible to the electron beams. The former paper has already been critically discussed by Greenberger (1981), Klein (1981), and Lipkin (1981). In the latter paper, a major contribution to the phase shift comes from magnetic fields close to one end of the long solenoid. However, this region of space is certainly not sampled by the electron beams in the usual experimental set-up, and their remark seems to add nothing of substance to the usual freedom in Stokes theorem (12) to choose on the right-hand side any area that is bordered by the contour of the line integral on the left-hand side. It is of course the case that the transverse vector potential $\boldsymbol{A}^{\perp}$ experienced by the electron beams is determined by the overall spatial distribution of magnetic field $\boldsymbol{B}$. Nevertheless, there is no way in which the phase shift can be expressed solely in terms of the magnetic fields actually sampled by the electron beams in their routes round opposite sides of the solenoid, and this remains the most interesting feature of the Aharonov–Bohm effect.

References

Aharonov Y and Bohm D 1959 *Phys. Rev.* **115** 485–91
Babiker M and Loudon R 1983 *Proc. R. Soc.* A **385** 439–60
Bocchieri P and Loinger A 1978 *Nuovo Cimento* A **47** 475–82
Bocchieri P, Loinger A and Siragusa G 1979 *Nuovo Cimento* A **51** 1–17
Bohm D and Hiley B J 1979 *Nuovo Cimento* A **52** 295–308
Ehrenberg W and Siday R E 1949 *Proc. Phys. Soc.* B **62** 8–21
Erlichson H 1970 *Am. J. Phys.* **38** 162–73
Feynman R P, Leighton R B and Sands M 1964 *The Feynman Lectures on Physics* (Reading, Mass.: Addison-Wesley)
Gradshteyn I S and Ryzhik I M 1965 *Table of Integrals, Series and Products* (New York: Academic)
Greenberger D M 1981 *Phys. Rev.* D **23** 1460–2
Home D and Sengupta S 1983 *Am. J. Phys.* **51** 942–7
Jackson J D 1975 *Classical Electrodynamics* (New York: Wiley)
Jaklevic R C, Lambe J, Mercereau J E and Silver A H 1965 *Phys. Rev.* **140** 1628–37
Klein U 1979 *Lett. Nuovo Cimento* **25** 33–7
—— 1981 *Phys. Rev.* D **23** 1463–5
Lipkin H J 1981 *Phys. Rev.* D **23** 1466–7
Möllenstedt G and Bayh W 1962 *Phys. Blätt.* **18** 299–305
Morse P M and Feshbach H 1953 *Methods of Theoretical Physics* (New York: McGraw-Hill)
Power E A 1964 *Introductory Quantum Electrodynamics* (London: Longmans)
Roy S M 1980 *Phys. Rev. Lett.* **44** 111–4
Tonomura A, Matsuda T, Suzuki R, Fukuhara A, Osakabe N, Umezaki H, Endo J, Shinagawa K, Sugita Y and Fujiwara H 1982 *Phys. Rev. Lett.* **48** 1443–6
Wódkiewicz K 1984 *Phys. Rev.* A **29** 1527–8
Zeilinger A 1979 *Lett. Nuovo Cimento* **25** 333–6