J. Phys. A: Math. Gen. **32** (1999) 5627–5641. Printed in the UK PII: S0305-4470(99)04457-1

Aharonov–Bohm beam deflection: Shelankov's formula, exact solution, asymptotics and an optical analogue

M V Berry
H H Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK

Received 21 May 1999

Abstract. Using a paraxial analysis, Shelankov (1998 *Europhys. Lett.* **43** 623) has shown that charged particles in a beam of small angular width $1/w$, aimed at a magnetic flux line with quantum flux α, are deflected through an angle D proportional to $\sin(2\pi\alpha)/w$, vanishing in the classical limit and also vanishing if the incident beam has zero intensity at the flux line. These properties are confirmed by numerical calculations based on an exact solution of the Schrödinger equation, and the paraxial wavefunction is obtained as an asymptotic approximation for large w. Paraxial theory suggests that the same deflection will occur for a light beam reflected by a mirror containing a step of height $\pi\alpha$. A theory of this optical phenomenon, based on an exact solution (for which D is not periodic in α), shows that convergence to the paraxial D is fast for $\alpha \ll 1$, and slow for α near $\frac{1}{2}$.

1. Introduction

In the Aharonov–Bohm (AB) effect (Aharonov and Bohm 1959, Olariu *et al* 1985), inaccessible magnetic flux diffracts electrons, producing interference fringes whose positions depend on the value of the flux. Classical particles would not be deflected because they traverse a region where there is no magnetic field, and hence there is no classical force. A natural question is: are quantum particles deflected? The usual AB theory cannot give an answer, because it is formulated in terms of an infinitely extended (plane) incident wave, for which deflections cannot be defined. One way to discuss deflection is to consider an incident wave that is a collimated beam rather than a plane wave. Recently, Shelankov (1998) has given a persuasive analysis (reviewed in section 2 for later reference) of the propagation of a beam in the presence of a flux line, based on a paraxial approximation to the Schrödinger equation, and concludes that there is a deflection, periodic in the flux as all AB effects must be, and given by a simple formula (equation (12)).

My aim here is twofold: to underpin Shelankov's theory by deriving his formula as an asymptotic approximation (section 4) from an exact analysis (section 3), and to explore an interesting optical analogue (section 5) that it suggests. The exact formulation is in terms of an incident beam that is monochromatic, constructed by superposing infinitely many AB waves incident from different directions with the same energy.

Other approaches are possible: the deflection of time-dependent wavepackets has been studied by Keating and Robbins (in preparation) in a careful analysis of the (Hermitian) quantum force operator; they show that although there is no classical force, there is a quantum force, giving rise to a deflection that in the appropriate limit is given by Shelankov's formula. A related problem is the determination of reaction forces on a vortex line giving rise to diffraction

0305-4470/99/305627+15$30.00

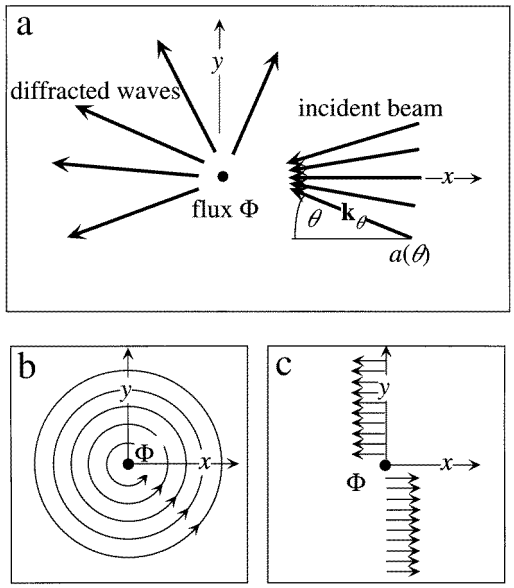

Figure 1. (*a*) Scattering geometry for incident beam with angular spectrum $a(\theta)$; (*b*) vector potential $\boldsymbol{A}_c(\boldsymbol{r})$ in circular gauge; (*c*) vector potential $\boldsymbol{A}_s(\boldsymbol{r})$ in Shelankov (sheet) gauge.

of quasiparticles; this was considered by Iordanskii (1966), and the associated (sometimes controversial) physics has been reviewed by Sonin (1997).

Figure 1(*a*) shows the (essentially two-dimensional) problem to be studied here. A monochromatic beam of electrons (charge $-e$) travelling in the $\boldsymbol{r} = (x, y)$ plane, with polar coordinates (r, ϕ), is incident from $x = +\infty$. The beam consists of plane waves travelling in directions θ (where $|\theta| < \pi/2$) with amplitudes $a(\theta)$. Piercing the plane at the origin is a magnetic flux line, of strength Φ.

Diffraction of the beam depends on the value of the flux in quantum units, that is on

$$\alpha = -\frac{e\Phi}{h} \tag{1}$$

and arises from the vector potential $\boldsymbol{A}(\boldsymbol{r})$ associated with the flux. Two such potentials will be useful in what follows: $\boldsymbol{A}_c(\boldsymbol{r})$, in the circular gauge (figure 1(*b*)), and $\boldsymbol{A}_s(\boldsymbol{r})$, in the sheet or Shelankov gauge (figure 1(*c*)); they are defined as

$$\boldsymbol{A}_c(\boldsymbol{r}) = \frac{\Phi}{2\pi r}\boldsymbol{e}_\phi \qquad \boldsymbol{A}_s(\boldsymbol{r}) = -\frac{\Phi}{2}\delta(x)[\Theta(y) - \Theta(-y)]\boldsymbol{e}_x \tag{2}$$

where Θ denotes the unit step. Both potentials satisfy $\nabla \times \boldsymbol{A}(\boldsymbol{r}) = \Phi\delta(\boldsymbol{r})$, and they are related by

$$\chi(\phi) \equiv -\frac{e}{h}\int_{+\infty \boldsymbol{e}_x}^{r} (\boldsymbol{A}_c(\boldsymbol{r}') - \boldsymbol{A}_s(\boldsymbol{r}'))\cdot \mathrm{d}\boldsymbol{r}' = \alpha\left[\phi - \pi\left\{\Theta\left(\phi - \frac{\pi}{2}\right) + \Theta\left(\phi - \frac{3\pi}{2}\right)\right\}\right] \tag{3}$$

$$(0 \leqslant \phi \leqslant 2\pi)$$

$$\chi(\phi + 2\pi) = \chi(\phi)$$

(of course, $\chi(\phi)$ is a single-valued function of position).

The only effect of the common wavenumber k of the plane waves in the beam is to set the scale of the diffraction pattern, so we immediately set $k = 1$; this amounts to measuring all lengths in units of wavelength/2π.

All formulae will be valid for general amplitudes $a(\theta)$ of the angular spectrum. For explicit calculations, we will use the Gaussian model

$$a(\theta) = \exp(-\tfrac{1}{2}\theta^2 w^2) \qquad (|\theta| < \tfrac{1}{2}\pi). \tag{4}$$

The r.m.s. angular width of the incident beam is $1/(w\sqrt{2})$; it is helpful to bear in mind that $w = 2$ corresponds to a width 27.01°, $w = 5$ to 8.1°, and $w = 20$ to 2.03°. In the paraxial regime, w is large.

2. Shelankov's paraxial theory

Paraxially, that is replacing $\cos\theta$ by $1 - \theta^2/2$ and ignoring the limits $|\theta| < \pi/2$, the incident wave, for $x > 0$, is the superposition

$$\psi_{\rm inc}(\boldsymbol{r}) = \frac{\exp\{-{\rm i}x\}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} {\rm d}\theta\, a(\theta) \exp\left\{{\rm i}\left(y\theta + \frac{1}{2}x\theta^2\right)\right\}. \tag{5}$$

(Hereafter, lower case ψ will denote waves in the paraxial approximation.) Immediately to the right of the plane $x = 0$ containing the flux, this wave is

$$\psi_0(y) \equiv \psi_{\rm inc}(0_+, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} {\rm d}\theta\, a(\theta) \exp\{{\rm i}y\theta\}. \tag{6}$$

For the model (4),

$$\psi_0(y) = \frac{1}{w} \exp\left(-\frac{y^2}{2w^2}\right) \tag{7}$$

describing the waist of a Gaussian beam with coordinate width w.

In Shelankov's gauge $\boldsymbol{A}_s$, the effect of the flux is confined to the sheet $x = 0$, where the vector potential acts as a phase-changing screen. Thus the transmitted wave immediately beyond the screen is

$$\psi(0_-, y) = \exp\{{\rm i}\pi\alpha\ {\rm sgn}\ y\}\psi_{\rm inc}(0_+, y) = \exp\{{\rm i}\pi\alpha\ {\rm sgn}\ y\}\psi_0(y). \tag{8}$$

The beam deflection $D_{\rm paraxial}$ can be calculated in terms of the transmitted angular distribution, namely (cf (6))

$$\begin{aligned} c_{\rm paraxial}(\alpha, \theta) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} {\rm d}y\ \psi(0_-, y) \exp(-{\rm i}\theta y) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} {\rm d}y\ \psi_0(y) \exp\{{\rm i}(\pi\alpha\ {\rm sgn}\ y - \theta y)\} \end{aligned} \tag{9}$$

as the average angle of the power spectrum $|c_{\rm paraxial}(\alpha, \theta)|^2$. Thus

$$D_{\rm paraxial}(\alpha) = \frac{\int_{-\infty}^{\infty} {\rm d}\theta\, \theta |c_{\rm paraxial}(\alpha, \theta)|^2}{\int_{-\infty}^{\infty} {\rm d}\theta\, |c_{\rm paraxial}(\alpha, \theta)|^2}. \tag{10}$$

Calculation of this average is subtle, because for non-zero flux $c_{\rm paraxial}(\alpha, \theta)$ decays as $|\theta| \to \infty$ as $1/\theta$ (the origin of this slow decay is the singularity in the integrands of (9) at $y = 0$), so the numerator integral in (10) diverges and must be regularized. One regularization was given by Shelankov (1998). Another regularization is based on two observations. First, the decay is the same at $\theta = \pm\infty$, so the odd part of $|c_{\rm paraxial}|^2$ decays faster (exponentially fast, for the Gaussian model (4)), and both integrals in (10) can be made to converge. Second, the phase screen transmission function can be written in terms of its even and odd parts:

$$\exp\{{\rm i}\pi\alpha\ {\rm sgn}\ y\} = \cos(\pi\alpha) + {\rm i}\ {\rm sgn}\ y \sin(\pi\alpha). \tag{11}$$

Then a short calculation leads to

$$D_{\text{paraxial}}(\alpha) = \sin(2\pi\alpha)\frac{|\psi_0(0)|^2}{\int_{-\infty}^{\infty} \mathrm{d}y\, |\psi_0(y)|^2}. \tag{12}$$

This is Shelankov's formula (his expression has the opposite sign because he considers waves incident from $x = -\infty$). When the wavenumber k is reinstated, it appears in the denominator, making the angle D_{paraxial} dimensionless and also showing that D_{paraxial} vanishes classically ($k \to \infty$) and is comparable to the diffraction spreading of the beam without flux. Evidently the deflection is an oscillatory function of the flux, and depends on the probability density of the paraxially incident wave where it strikes the flux line. It is possible to obtain the result (12) from the expectation value of the transverse current operator $-\mathrm{i}\partial/\partial y$, but we will see that the non-paraxial generalization is problematic.

For later comparison with the exact solution, we note that the paraxially transmitted wave, for $x < 0$, can be calculated by regarding $\psi(0_-, y)$ (equation (8)) as a source of diffracted waves. Thus

$$\psi(\boldsymbol{r}) = \frac{\exp\{-\mathrm{i}(x + \frac{1}{4}\pi)\}}{\sqrt{2\pi|x|}} \int_{-\infty}^{\infty} \mathrm{d}u\, \psi_0(u) \exp\left\{\mathrm{i}\left[\pi\alpha\, \mathrm{sgn}\, u - \frac{\mathrm{i}}{2x}(y-u)^2\right]\right\} \qquad (x < 0). \tag{13}$$

With the model (4), explicit formulae are, for the transmitted spectrum (9),

$$c_{\text{paraxial}}(\alpha, \theta) = \exp\left(-\frac{1}{2}\theta^2 w^2\right)\left[\cos(\pi\alpha) + \sin(\pi\alpha)\mathrm{erfi}\left(\frac{w\theta}{\sqrt{2}}\right)\right] \tag{14}$$

for the transmitted wave (13),

$$\psi(\boldsymbol{r}) = \frac{\exp(-\mathrm{i}x)}{\sqrt{w^2 + \mathrm{i}|x|}} \exp\left\{-\frac{y^2}{2(w^2 + \mathrm{i}|x|)}\right\}\left[\cos(\pi\alpha) + \mathrm{i}\sin(\pi\alpha)\,\mathrm{erf}\left\{\frac{\mathrm{i}wy}{\sqrt{2|x|(|x| - \mathrm{i}w^2)}}\right\}\right] \tag{15}$$

and, for the deflection (12),

$$D_{\text{paraxial}}(\alpha) = \frac{\sin(2\pi\alpha)}{w\sqrt{\pi}}. \tag{16}$$

Here erf is the error function (Abramowitz and Stegun 1972), and $\mathrm{erfi}(z) = -\mathrm{i}\,\mathrm{erf}(\mathrm{i}z)$. Note that when $\alpha = 0$ the wave (15) becomes the paraxially propagating Gaussian beam (with $|x|$ replaced by x the formula is then valid for positive as well as negative x).

3. Exact solution

In the circular gauge, a formal but multi-valued solution of Schrödinger's equation for a plane wave travelling in the direction $\boldsymbol{k}_\theta = (-\cos\theta, \sin\theta)$ (figure 1), in the presence of the flux, is

$$\exp\{\mathrm{i}(\boldsymbol{k}_\theta \cdot \boldsymbol{r} + \alpha\phi)\} = \exp\{\mathrm{i}(-r\cos(\phi + \theta) + \alpha(\phi + \theta))\}\exp\{-\mathrm{i}\alpha\theta\}. \tag{17}$$

The corresponding single-valued AB solution (Aharonov and Bohm 1959), describing the scattering of the plane wave by the flux, is

$$\Psi_\theta(\boldsymbol{r}) = \sum_{l=-\infty}^{\infty} J_{|l-\alpha|}(r)(-\mathrm{i})^{|l-\alpha|}\exp\{\mathrm{i}l(\phi + \theta)\}\exp\{-\mathrm{i}\alpha\theta\}. \tag{18}$$

Therefore, the wave describing the scattering of a beam with angular distribution $a(\theta)$ is

$$\Psi(\boldsymbol{r}) = \frac{1}{\sqrt{2\pi}} \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta\, a(\theta) \Psi_\theta(\boldsymbol{r}) = \sum_{l=-\infty}^{\infty} J_{|l-\alpha|}(r)(-\mathrm{i})^{|l-\alpha|} \exp\{\mathrm{i}l\phi\} b(l-\alpha) \tag{19}$$

where

$$b(l) = \frac{1}{\sqrt{2\pi}} \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta\, a(\theta) \exp\{\mathrm{i}l\theta\}. \tag{20}$$

Equations (19) and (20) constitute the exact AB wave for an incident beam. By restricting directions to the range $-\pi/2 < \theta < \pi/2$, we are declaring that the incident beam contains no waves travelling back towards $+x$, and no evanescent waves (which would have $\sin\theta > 1$ and hence θ complex). Absence of evanescence amounts to the physical requirement that the beam without flux is smooth and finite in intensity for all $\boldsymbol{r}$; if evanescent waves were permitted, different angular superpositions would be required in different regions (e.g. $x < 0$ and $x > 0$) to prevent the intensity blowing up in the growing direction, and it can be shown that then $\nabla\Psi$ would be discontinuous at the boundaries of the regions (some interesting properties of non-paraxial beams made of real plane waves are discussed by Berry (1994, 1998)).

The coefficients $b(l)$ describe a function very close to the paraxial unscattered waist wave amplitude (6), discretely sampled. For the model (4),

$$b(l) = \frac{1}{w} \exp\left(-\frac{l^2}{2w^2}\right) \left[1 - \mathrm{Re}\,\mathrm{erfc}\left\{\frac{1}{\sqrt{2}}\left(\frac{\pi w}{2} - \mathrm{i}\frac{l}{w}\right)\right\}\right] \tag{21}$$

and asymptotics of erfc gives

$$b(l) - \psi_0(l) \approx -\exp\left(-\frac{1}{8}\pi^2 w^2\right) \sqrt{\frac{2}{\pi}} \frac{[\frac{1}{2}\pi w^2 \cos(\frac{1}{2}\pi l) - l \sin(\frac{1}{2}\pi l)]}{l^2 + \frac{1}{4}\pi^2 w^4} \qquad (w \gg 1). \tag{22}$$

For large w (that is, paraxial beams) this difference is negligible if l is small; for large l, the difference decays slowly (as $1/l$) and dominates ψ_0 if $|l| > w^2$.

Because of the scattering from the flux, the angular spectrum is no longer $a(\theta)$ as in the incident wave. Moreover, the superposition of plane waves corresponding to $\Psi(\boldsymbol{r})$ need not be the same in all regions of the plane, because of the singularity at the flux line. To find the deflection, we need the superposition, with angular amplitude $c(\alpha, \theta)$, on the scattering side of the flux, that is $x < 0$ or $\pi/2 < \phi < 3\pi/2$. To calculate $c(\alpha, \theta)$, we use the sheet gauge, because $\boldsymbol{A}_s$ is zero for $x < 0$, to transform from the circular gauge using (3). The power spectrum $|c(\theta)|^2$ is the same for all $x < 0$. We choose the line 0_-, on which $\phi = \pi - \frac{1}{2}\pi \operatorname{sgn} y$, and

$$\begin{aligned} \Psi(\boldsymbol{r}) \exp\{-\mathrm{i}\chi(\boldsymbol{r})\} &= \exp\{\mathrm{i}\alpha\pi\} \Psi(0_-, y) \exp\{-\mathrm{i}\alpha\pi(1 - \tfrac{1}{2}\operatorname{sgn} y)\} \\ &= \frac{1}{\sqrt{2\pi}} \int \mathrm{d}\theta\, c(\alpha, \theta) \exp\{\mathrm{i}y\boldsymbol{e}_y \cdot \boldsymbol{k}_\theta\} \\ &= \frac{1}{\sqrt{2\pi}} \int \mathrm{d}\theta\, c(\alpha, \theta) \exp\{\mathrm{i}y \sin\theta\}. \end{aligned} \tag{23}$$

The integral can include evanescent waves (complex θ), but we are only interested in the real waves $-\pi/2 < \theta < \pi/2$, since these determine the deflection of the beam far from the flux. Now $c(\alpha, \theta)$ is obtained by Fourier transforming over y:

$$c(\alpha, \theta) = \frac{|\cos\theta|}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}y\, \exp\{-\mathrm{i}y \sin\theta\} \exp\{\mathrm{i}\alpha\pi\} \Psi(0_-, y) \exp\left\{-\mathrm{i}\alpha\pi\left(1 - \frac{1}{2}\operatorname{sgn} y\right)\right\}$$

$$= \sum_{l=-\infty}^{\infty} (-\mathrm{i})^{|l-\alpha|}(-1)^l b(l-\alpha) I(l-\alpha) \tag{24}$$

where

$$I(\lambda) \equiv \frac{|\cos\theta|}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}y \ \exp\{-\mathrm{i}y\sin\theta\} J_{|\lambda|}(|y|) \exp\left\{-\mathrm{i}\frac{1}{2}\pi\lambda \ \mathrm{sgn}\ y\right\}. \tag{25}$$

Evaluating the integral using formulae (11.4.37–8) of Abramowitz and Stegun (1972) gives the angular spectrum exactly as

$$c(\alpha,\theta) = \sqrt{\frac{2}{\pi}} \sum_{l=-\infty}^{\infty} (-1)^l (-\mathrm{i})^{|l-\alpha|} \cos\left\{\frac{1}{2}\pi(l-\alpha) + |l-\alpha|\theta\right\} b(1-\alpha). \tag{26}$$

Finally, the deflection can be defined as the average value of the transverse momentum $\sin\theta$ over this spectrum (the results are hardly altered by choosing θ instead), namely (cf (10))

$$D(\alpha) = \frac{\int_{-\pi/2}^{\pi/2} \mathrm{d}\theta \ \sin\theta |c(\alpha,\theta)|^2}{\int_{-\pi/2}^{\pi/2} \mathrm{d}\theta \ |c(\alpha,\theta)|^2}. \tag{27}$$

From equation (26) it follows that $c(\alpha,\theta)$ changes sign when α changes by unity. For the case where $b(l)$ is even, representing a transversely symmetric beam striking the flux centrally, $c(\alpha,\theta)$ possesses an additional symmetry, relating α,θ to $-\alpha,-\theta$, so that:

$$c(\alpha,\theta) = -c(\alpha+1,\theta) = c(-\alpha,-\theta). \tag{28}$$

These imply $D(\alpha) = -D(-\alpha) = -D(1-\alpha)$, and so $D(\frac{1}{2}+\alpha) = -D(\frac{1}{2}-\alpha)$, further implying $D(0) = D(\frac{1}{2}) = 0$, so that it is necessary to study $D(\alpha)$ only for $0 \leqslant \alpha \leqslant \frac{1}{2}$. The corresponding paraxial quantities (9) and (12) possess the additional symmetry $D_{\mathrm{paraxial}}(\frac{1}{4}-\alpha) = D_{\mathrm{paraxial}}(\frac{1}{4}+\alpha)$.

It might be thought that an alternative definition of D would be as the expectation value of the transverse momentum operator along any slice with constant $x < 0$, for example

$$D(\alpha) \equiv \frac{\int_{-\infty}^{\infty} \mathrm{d}y \ \Psi^*(0_-,y)(-\mathrm{i}\partial_y)\Psi(0_-,y)}{\int_{-\infty}^{\infty} \mathrm{d}y \ |\Psi(0_-,y)|^2} \qquad (x<0) \tag{29}$$

(when k is reinstated in the denominator of this expression, it is dimensionless, as an angle must be). This expectation is however frustrated, because the normalization integral almost always diverges, for a curious reason explained in the appendix.

Now we present some numerical calculations based on the model (4), for different values of the beam width w and flux α. Figure 2 shows how accurately the AB deflection D is described by Shelankov's paraxial formula (16), even for small w. The maxima of the $D(\alpha)$ curves are shifted slightly to the right of $\alpha = \frac{1}{4}$, indicating that the paraxial symmetry $D_{\mathrm{paraxial}}(\frac{1}{4}-\alpha) = D_{\mathrm{paraxial}}(\frac{1}{4}+\alpha)$ is (weakly) broken for small w. A further comparison (figure 3) is provided by the spectrum $|c(\theta)|^2$.

Figure 4 shows density plots of the wave intensities, indicating that the excellent agreement between paraxial and exact solutions extends to the wavefunction itself everywhere in the scattering region $x < 0$. In the region $x > 0$, the exact solutions show interference between the scattered waves and the incident beam. More details are shown in figure 5, as slices of the wave intensity for different α. This shows the rightward shift of the beam from zero at $\alpha = 0$, through its maximum near $\alpha = \frac{1}{4}$, and back to zero at $\alpha = \frac{1}{2}$ and also the fact that in the incidence region $x > 0$ the beam (figure 5(*f*)) is hardly shifted at all, although it is affected by waves reflected from the flux.

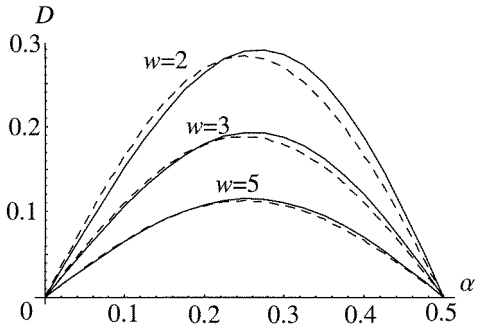

Figure 2. Deflection D versus flux α, for the indicated beam widths w in the model (4). Full curves: the exact AB deflection (27); dashed curves: deflection from Shelankov's formula (16).

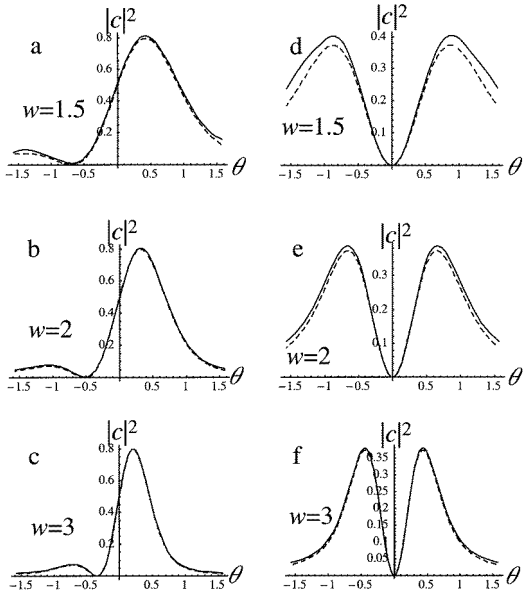

Figure 3. Angular spectrum $|c(\theta)|^2$ for AB scattered beams with the indicated widths w. Full curves: exact spectrum (26); dashed curves: paraxial approximation (14). In (*a*)–(*c*), $\alpha = \frac{1}{4}$; in (*d*)–(*f*), $\alpha = \frac{1}{2}$.

Shelankov's formula (12) predicts that if the incident beam strikes the flux line non-centrally, by being shifted sideways by y_0, the deflection is smaller. In the computations, this shift can be exactly accommodated simply by changing the argument of $b(l)$ in formula (26) for $c(\theta)$ from $l-\alpha$ to $l-\alpha-y_0$. For the Gaussian model (7), the deflection (equation (16)) should be reduced by a factor of $\exp(-y_0^2/2w^2)$. Numerical calculations confirm this, although the approach to paraxiality as $w \to \infty$ is slower than for symmetric incidence. Since $\psi_0(y)$, and therefore $b(l)$, are no longer even functions, the second symmetry in (28) is no longer exact. In particular, the exact D does not vanish when $\alpha = \frac{1}{2}$, however, this symmetry-breaking is hardly detectable if w exceeds about 2.

A more striking prediction is that if the incident wave (without flux) vanishes at the origin, $D(\alpha)$ should be zero for all α. We can test this by choosing the model (cf (7))

$$\psi_0(y) = \frac{y^2}{w}\exp\left(-\frac{y^2}{2w^2}\right). \tag{30}$$

The corresponding angular spectrum is, from (6),

$$a(\theta) = w^2(1-w^2\theta^2)\exp\{-\tfrac{1}{2}w^2\theta^2\}. \tag{31}$$

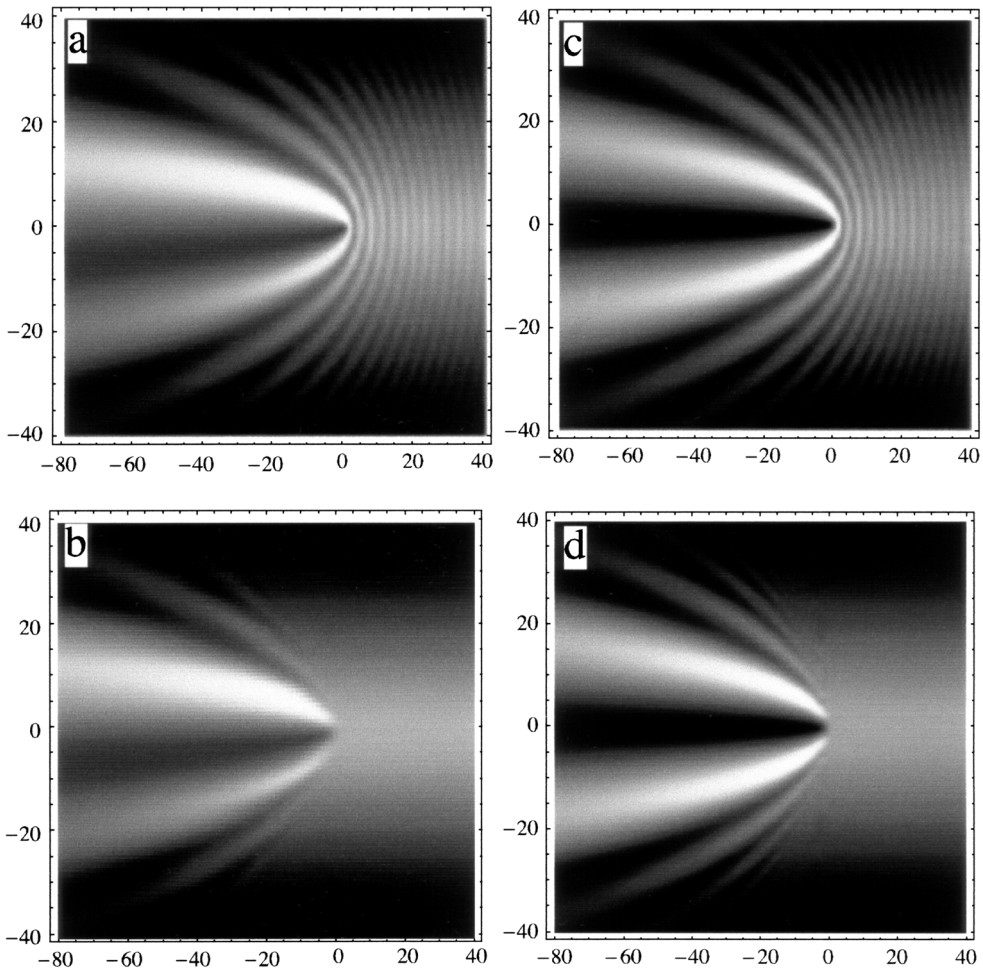

Figure 4. Density plots of the exact AB wave intensity $|\Psi(r)|^2$ (equation (19)) (*a*) and (*c*), and the paraxial wave intensity $|\psi(r)|^2$ (equation (15), with $\alpha = 0$ for the incident region $x > 0$), for $w = 20$. In (*a*), (*b*): $\alpha = \frac{1}{4}$; in (*b*), (*d*): $\alpha = \frac{1}{2}$. x horizontal, y vertical.

Using (20) and (26), the angular spectrum $c(\theta)$ of the scattered wave can be computed from the exact theory. Figure 6 shows an example of $|c(\theta)|^2$. The exact and paraxial curves agree rather well, even for such a small w (wide-angle incident beam). Although the curves are not symmetric about $\theta = 0$, D_{paraxial} is exactly zero, from Shelankov's theory, and D calculated numerically from the exact formula (27) is 0.006, which is zero within the accuracy of these computations.

This sensitivity to screening of the flux by the incident beam is a remarkable phenomenon, because the exact AB wave (19) vanishes at $r = 0$ for α non-zero whatever the form of the incident beam, so there is a sense in which the flux is always hidden from the electrons. Nevertheless, it is possible to detect, by the deflection of a beam, whether the incident beam intensity in the absence of flux vanishes at the place where the flux line will be inserted. In the familiar AB scattering, the situation is different: the fractional shift of the far-field fringes is α whether the flux line is screened by the incident beam, as in the original biprism experiment

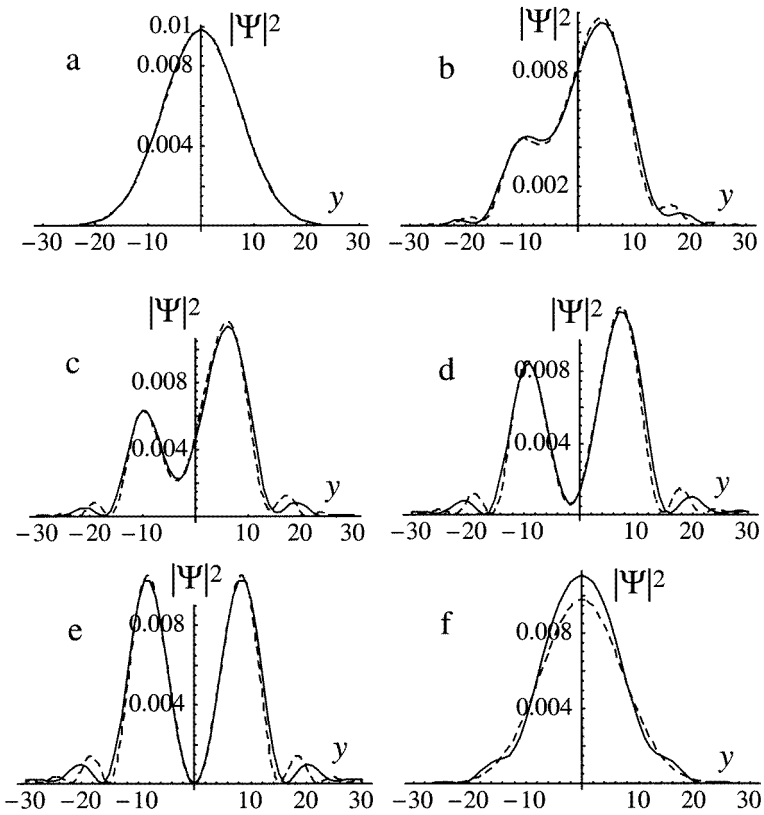

Figure 5. Transverse slices of the exact AB wave intensities (full curves) and paraxial approximations (dashed curves), for $w = 10$ and (*a*)–(*e*) $x = -20$ (scattering region), (*f*) $x = +20$ (incidence region). In (*a*), $\alpha = 0$; (*b*), $\alpha = \frac{1}{8}$; (*c*), $\alpha = \frac{1}{4}$; (*d*), $\alpha = \frac{3}{8}$; (*e*), $\alpha = \frac{1}{2}$; (*f*), $\alpha = \frac{1}{4}$.

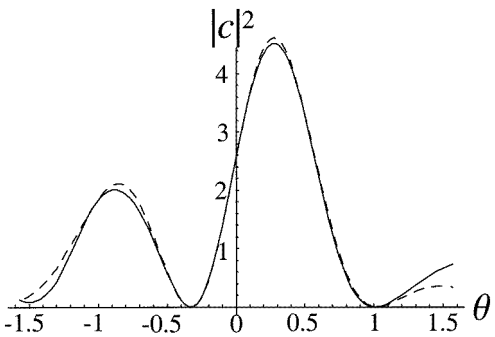

Figure 6. Angular spectrum $|c(\theta)|^2$ for AB scattered beam with width $w = 1.5$ in the model (31), for an incident wave vanishing on the flux line, which here has $\alpha = \frac{1}{4}$. Full curve: exact spectrum (26); dashed curve: paraxial approximation (9). (For larger w, the exact and paraxial curves are hard to distinguish on this scale.)

of Chambers (1960), or an unobstructed plane wave, as in the original solution of Aharonov and Bohm (1959).

Numerical computation leaves several questions unanswered. Does D vanish exactly when the incident intensity would be zero on the flux line, or is D merely very small? In the latter case, how does D depend on the details of the wavefunction near the flux? How does D vanish as $w \to \infty$? These are not just mathematical niceties. In any practical experiment, the flux will be contained in a tube of finite radius, which of course must be made impenetrable

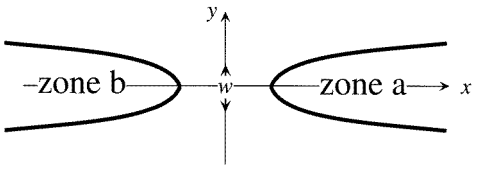

Figure 7. Zones for the calculation of the paraxial asymptotics of $\Psi(\boldsymbol{r})$ in terms of the integrals (36). In zone a only $I_0^{\mathrm{II}}(\boldsymbol{r})$ contributes significantly, in zone b $I_0^{\mathrm{I}}(\boldsymbol{r})$ contributes if $y > 0$, and $I_{-1}^{\mathrm{I}}(\boldsymbol{r})$ contributes if $y < 0$.

by a scalar potential (representing its walls), in order to eliminate any non-AB deflection from the electrons' encounter with the real magnetic field. But walls cause the wave to vanish even when $\alpha = 0$, and then a strict interpretation of Shelankov's formula threatens to make D zero when $\alpha \neq 0$. On physical grounds, however, I expect that if the flux tube is much thinner than an electron wavelength (a non-trivial practical constraint) the preceding theory can be applied as though the thickness were zero, and there will be a deflection for $\alpha \neq 0$; this deserves further study.

4. Paraxial asymptotics

The paraxial regime is $w \gg 1$ (incident beam narrow in angle), with $\boldsymbol{r}$ located in two regions (figure 7): zone a, where $r \gg w$ (far from the flux) and $|\phi| \ll 1$ (near-backward direction), and zone b, where $r \gg w$ and $|\pi - \phi| \ll 1$ (near-forward direction).

There are four steps in the derivation of the paraxial wave (13) from the exact AB wave (19). The first, which is exact, is to transform the sum (19) to a series of integrals by the Poisson sum formula (Lighthill 1958):

$$\Psi(\boldsymbol{r}) = \exp(\mathrm{i}\alpha\phi) \sum_{m=-\infty}^{\infty} \exp(2\pi \mathrm{i} m\alpha) I_m(\boldsymbol{r}) \tag{32}$$

where

$$I_m(\boldsymbol{r}) = \int_{-\infty}^{\infty} \mathrm{d}\lambda \, J_{|\lambda|}(r) b(\lambda) \exp(-\tfrac{1}{2}\mathrm{i}\pi|\lambda|) \exp\{\mathrm{i}\lambda(\phi + 2\pi m)\}. \tag{33}$$

This is the 'whirling-wave' decomposition of the wave, discussed elsewhere (Berry 1980).

The second step is to replace $b(\lambda)$ by its close approximation $\psi_0(\lambda)$, which is valid when $w \gg 1$ (cf equations (6) and (20)). The decay of $\psi_0(\lambda)$ cuts off the integrals in (33) when $|\lambda| > w$. Then, since $r \gg w$, all the Bessel functions contributing significantly to the sum have argument $\gg$ order. This makes possible the third step, which is to replace $J_{|\lambda|}(r)$ by the appropriate asymptotic expression, namely the following variant of the Debye formula (Abramowitz and Stegun 1972):

$$J_{|\lambda|}(r) \approx \sqrt{\frac{2}{\pi r}} \cos\left(r + \frac{|\lambda|^2}{2r} - \frac{1}{2}|\lambda|\pi - \frac{1}{4}\pi\right) \tag{34}$$

(note that we are not assuming $r \gg |\lambda|^2$, when this formula reduces to the simplest large-argument approximation). Thus

$$I_m(\boldsymbol{r}) \approx I_m^{\mathrm{I}}(\boldsymbol{r}) + I_m^{\mathrm{II}}(\boldsymbol{r}) \tag{35}$$

where

$$I_m^{\mathrm{I}}(\boldsymbol{r}) = \frac{1}{\sqrt{2\pi r}} \exp\left\{\mathrm{i}\left(r - \frac{1}{4}\pi\right)\right\} \int_{-\infty}^{\infty} \mathrm{d}\lambda\, \psi_0(\lambda) \exp\left\{\mathrm{i}\left(-\pi|\lambda| + \frac{\lambda^2}{2r} + \lambda(\phi + 2\pi m)\right)\right\}$$
$$I_m^{\mathrm{II}}(\boldsymbol{r}) = \frac{1}{\sqrt{2\pi r}} \exp\left\{-\mathrm{i}\left(r - \frac{1}{4}\pi\right)\right\} \int_{-\infty}^{\infty} \mathrm{d}\lambda\, \psi_0(\lambda) \exp\left\{\mathrm{i}\left(-\frac{\lambda^2}{2r} + \lambda(\phi + 2\pi m)\right)\right\}. \quad (36)$$

The final step is to retain only those integrals whose integrands possess stationary points on the real λ axis with $|\lambda| \ll w$; in comparison with these, the other integrals are exponentially small. The stationary points are located at:

$$\text{integral } I^{\mathrm{I}} : \lambda = r(\pi \operatorname{sgn}\lambda - \phi - 2\pi m) \qquad \text{integral } I^{\mathrm{II}} : \lambda = r(\phi + 2\pi m). \quad (37)$$

In zone a, $|\phi|$ is small and $r \gg w$, and the only stationary point with $|\lambda| \ll w$ comes from the integral I^{II}, with $m = 0$. We obtain, from (32), (35) and (37),

$$\Psi(\boldsymbol{r}) \approx \frac{\exp\{\mathrm{i}(\alpha\phi + \frac{1}{4}\pi - r)\}}{\sqrt{2\pi r}} \int_{-\infty}^{\infty} \mathrm{d}\lambda\, \psi_0(\lambda) \exp\left\{\mathrm{i}\left(-\frac{\lambda^2}{2r} + \lambda\phi\right)\right\} \qquad \text{(zone a)}. \quad (38)$$

Then using $r \approx x + y^2/2x$ and $\phi \approx y/x$ gives precisely the paraxial incident beam, together with the α-dependent phase factor associated with the flux in the circular gauge. Strictly, the paraxial approximation fails to incorporate waves scattered back towards $x = +\infty$; these inevitable reflected waves originate in the asymptotics associated with the discontinuity at $\lambda = 0$ in the integrals I^{I}.

In zone b, only the integrals I^{I} have stationary points with $|\lambda| \ll w$, from $m = 0$ when $\pi - \phi > 0$ (i.e. $y > 0$ and $\lambda > 0$), and from $m = -1$ when $\pi - \phi < 0$ (i.e. $y < 0$ and $\lambda < 0$). Equations (32), (35) and (37) now lead to

$$\Psi(\boldsymbol{r}) \approx \frac{\exp\{\mathrm{i}(\alpha(\phi - \pi) - \frac{1}{4}\pi + r)\}}{\sqrt{2\pi r}} \int_{-\infty}^{\infty} \mathrm{d}\lambda\, \psi_0(\lambda) \exp\left\{\mathrm{i}\left(\pi\alpha \operatorname{sgn}\lambda + \frac{\lambda^2}{2r} - \lambda(\pi - \phi)\right)\right\}$$
(zone b). (39)

In this zone, $r \approx -x - y^2/2x$, and $\phi \approx \pi + y/x$, giving exactly Shelankov's paraxial scattered wave (13), together with the α-dependent phase factor associated with the transformation (3) to $\boldsymbol{A}_{\mathrm{c}}$ from $\boldsymbol{A}_{\mathrm{s}}$.

5. Optical analogue: diffraction by a reflecting step

In the paraxial theory, and using the sheet gauge $\boldsymbol{A}_{\mathrm{s}}(\boldsymbol{r})$, the vector potential acts as a phase-changing screen (equation (8)). This suggests that AB beam deflection can be simulated with other kinds of waves, for example light. The screen could be a slab of transparent material whose refracting surface is stepped, or, as in the case to be examined here, a mirror whose surface contains a step with straight edges (figure 8). To implement an AB flux α, the phase change between waves reflected by the two sides of the step must be $2\pi\alpha$. Therefore the height of the step should be $\pi\alpha$ in units where the wavenumber k is unity, that is $\alpha/2$ wavelengths. Focusing a beam with spatial width $w \gg 1$—to ensure paraxiality—onto the step ought to generate a reflected wave whose deflection is given by Shelankov's formula (12); for a Gaussian beam, the deflection should be (16).

This is an interesting optical effect, worth studying for its own sake, and seeking to detect experimentally. In any such experiment, it is important that the step be perpendicular to the surface and sharp on the wavelength scale. Any deviation from these conditions will spoil the AB analogy by giving rise to a geometrical-optics deflection, corresponding to electrons penetrating into the flux and therefore experiencing a classical force.

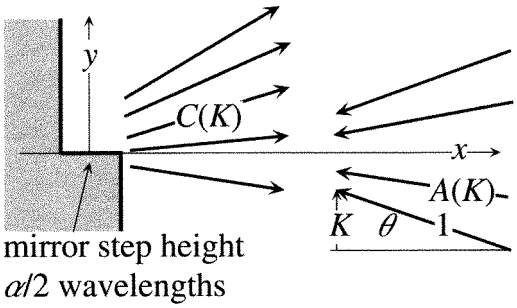

Figure 8. Diffraction by a mirror step of height $\pi\alpha$, that is, $\alpha/2$ wavelengths; the incident beam has angular spectrum $A(K)$, and the reflected waves have angular spectrum $C(K)$.

However, paraxiality is an approximation, and it is necessary to compare its predictions with those of an exact solution for waves reflected by a step. This is not the same as the exact AB solution (section 3) in the presence of a flux line. One important difference is that reflection from a step is not periodic in the step height: periodicity should emerge in the paraxial limit. Another is that there is no exact analytical solution for step reflection, making it necessary to formulate the problem in a way that can be implemented computationally, automatically incorporating the complicated diffracted (and multiply diffracted) waves from the edges of the step.

The step will be a perfectly reflecting surface whose profile (figure 8) is

$$x_{\text{step}}(y) = -\tfrac{1}{2}\pi\alpha \operatorname{sgn} y \tag{40}$$

(again we take $k = 1$). Light polarized in the z direction can be represented by a scalar wave $\Psi(\boldsymbol{r})$ vanishing on the step. It is convenient to label plane waves in the incident and reflected superpositions by their amplitudes as a function of the transverse wavenumber $K = \sin\theta$, rather than the direction θ. The amplitudes will be denoted $A(K)$ and $C(K)$, related to those defined previously by

$$\begin{aligned} a(\theta) &= \cos\theta A(\sin\theta) \qquad \text{i.e. } A(K) = \frac{a(\sin^{-1}K)}{\sqrt{1-K^2}} \\ c(\theta) &= \cos\theta C(\sin\theta) \qquad \text{i.e. } C(K) = \frac{c(\sin^{-1}K)}{\sqrt{1-K^2}}. \end{aligned} \tag{41}$$

The incident and reflected waves can now be written as

$$\begin{aligned} \Psi_{\text{inc}}(\boldsymbol{r}) &= \frac{1}{\sqrt{2\pi}}\int_{-1}^{1} \mathrm{d}K\, A(K) \exp\left\{\mathrm{i}\left[Ky - x\sqrt{1-K^2}\right]\right\} \\ \Psi_{\text{refl}}(\boldsymbol{r}) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}K\, C(K) \exp\left\{\mathrm{i}\left[Ky + x\sqrt{1-K^2}\right]\right\}. \end{aligned} \tag{42}$$

The limits of integration indicate that Ψ_{inc} does not contain evanescent waves ($|K| > 1$), but Ψ_{refl} does; then the square root is positive imaginary.

Now the problem is to determine $C(K)$ given $A(K)$, using the fact that the total wavefunction must vanish on the mirror (40). Thus

$$\begin{aligned} &\int_{-\infty}^{\infty} \mathrm{d}K\, C(K) \exp\left\{\mathrm{i}\left[Ky - \tfrac{1}{2}\pi\alpha \operatorname{sgn} y\sqrt{1-K^2}\right]\right\} \\ &\qquad = -\int_{-1}^{1} \mathrm{d}K\, A(K) \exp\left\{\mathrm{i}\left[Ky + \tfrac{1}{2}\pi\alpha \operatorname{sgn} y\sqrt{1-K^2}\right]\right\}. \end{aligned} \tag{43}$$

This must hold for all y. Multiplying by $\exp\{\mathrm{i}(-Qy + \pi\alpha \operatorname{sgn} y/2)\}$, and integrating over y, using the trick (11) and also

$$\int_{-\infty}^{\infty} \mathrm{d}y \ \operatorname{sgn} y \exp\{\mathrm{i}(K-Q)y\} = \frac{2\mathrm{i}}{K-Q} \tag{44}$$

leads to

$$\begin{aligned} C(Q)\cos\left\{\tfrac{1}{2}\pi\alpha\left(\sqrt{1-Q^2}-1\right)\right\} &+ \frac{1}{\pi}\int_{-\infty}^{\infty}\mathrm{d}K\, C(K)\frac{\sin\left\{\tfrac{1}{2}\pi\alpha\left(\sqrt{1-K^2}-1\right)\right\}}{K-Q} \\ &= -A(Q)\cos\left\{\tfrac{1}{2}\pi\alpha\left(\sqrt{1-Q^2}+1\right)\right\} \\ &\quad + \frac{1}{\pi}\int_{-1}^{1}\mathrm{d}K\, A(K)\frac{\sin\left\{\tfrac{1}{2}\pi\alpha\left(\sqrt{1-K^2}+1\right)\right\}}{K-Q} \end{aligned} \tag{45}$$

where the singularities at $K = Q$ are interpreted as principal values.

From $C(K)$, the deflection can be calculated from (27) and (41) as

$$D = \frac{\int_{-1}^{1}\mathrm{d}K\, K\sqrt{1-K^2}|C(K)|^2}{\int_{-1}^{1}\mathrm{d}K\,\sqrt{1-K^2}|C(K)|^2}. \tag{46}$$

In this formalism, the paraxial approximation is obtained by neglecting K^2 and Q^2, leading to the explicit formula

$$C(Q) \approx C_{\text{paraxial}}(Q) = -A(Q)\cos(\pi\alpha) + \frac{\sin(\pi\alpha)}{\pi}\int_{-\infty}^{\infty}\mathrm{d}K\,\frac{A(K)}{K-Q}. \tag{47}$$

It is not hard to show that this is equivalent to (9), apart from an overall minus sign arising from reflection.

By discretizing the integral operators in the exact equations (45) and representing them as finite matrices (setting the diagonal matrix element $(K-Q)^{-1}$ to zero when $K = Q$), an effective numerical scheme can be implemented, and $C(K)$ determined by matrix inversion. We employ the model (4), that is

$$A(K) = \frac{\exp\{-\frac{1}{2}(w\sin^{-1}K)^2\}}{\sqrt{1-K^2}}. \tag{48}$$

Evanescent waves can be incorporated by including transverse wavenumbers $|K| \leqslant K_{\max}$, $|Q| \leqslant K_{\max}$, where $K_{\max} > 1$. Of course, evanescent waves are necessary to make the wave vanish on the step (figure 9). However, when w is large the influence of the evanescent waves on the real waves (that is, on $C(K)$ for $|K| < 1$, and $c(\theta)$ for $|\theta| < \pi/2$) which determine the distant reflected field and the beam deflection, is negligible (figure 10(*a*)).

For small steps ($\alpha < \frac{1}{4}$), $c(\theta)$ is well approximated by the paraxial formula (14) (figure 10(*b*)). As α approaches $\frac{1}{2}$ (step height of $\frac{1}{4}$ wavelength), the agreement gets worse, and the expected symmetry does not materialize at $\alpha = \frac{1}{2}$ (figure 10(*c*)). As w increases, the curves for $\alpha = \frac{1}{2}$ do get more symmetrical, albeit slowly. Therefore convergence of the deflection D to D_{paraxial} is slow when α is not small (figure 11). It would be useful to have an asymptotic estimate of the rate at which $D \to 0$ (if indeed it does) as $w \to \infty$ when $\alpha = \frac{1}{2}$.

Acknowledgments

I thank Andrei Shelankov, Jonathan Keating and Jonathan Robbins for telling me about their unpublished researches on AB deflection, and for their kind and helpful comments on this paper.

Appendix. Failure of slice normalization

Consider any wave $\Psi(\boldsymbol{r})$ that can be expressed as an angular superposition of non-evanescent waves in directions $(\cos\theta, \sin\theta)$, where θ is real:

$$\Psi(\boldsymbol{r}) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}\theta\, a(\theta) \exp\{\mathrm{i}(x\cos\theta + y\sin\theta)\}. \tag{A1}$$

We slice Ψ on the line

$$x = x_0 + s\cos\gamma \qquad y = s\sin\gamma \tag{A2}$$

and attempt to normalize along the slice by requiring that

$$N(\gamma) \equiv \int_{-\infty}^{\infty} \mathrm{d}s\, |\Psi(x_0 + s\cos\gamma, s\sin\gamma)|^2 \tag{A3}$$

is finite.

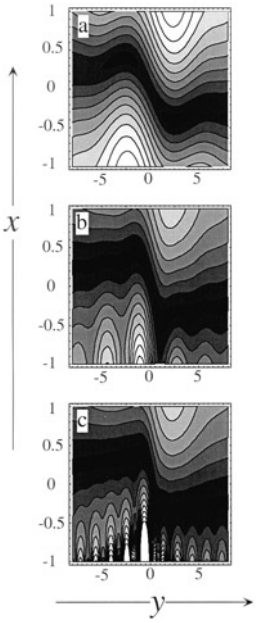

Figure 9. Contour plots of intensity $|\Psi(\boldsymbol{r})|^2$ for wave diffracted by step $\alpha = \frac{1}{4}$, with incident beam width $w = 10$, including transverse wavenumbers (*a*): $|K| < 1$ (no evanescent waves); (*b*): $|K| < 2$; (*c*): $|K| < 4$. As more evanescent waves are included, the vanishing of Ψ on the step is better approximated, the wave in the unphysical region behind the mirror ($x < x_{\mathrm{step}}$) gets wilder, but the contours in the far field hardly change.

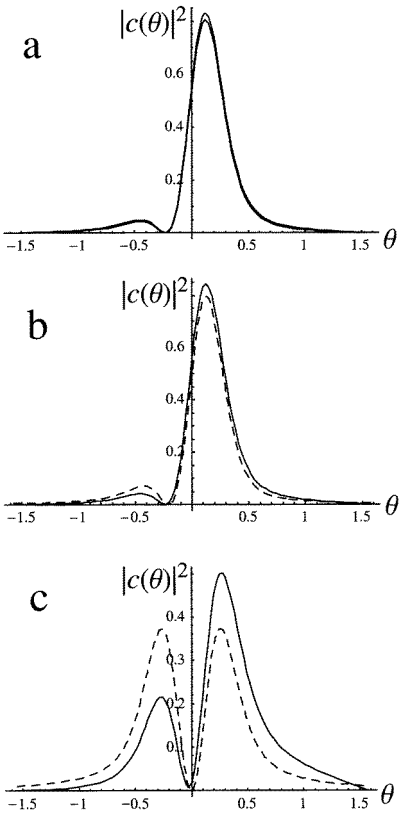

Figure 10. Power spectrum $|c(\theta)|^2$ of light beam ($w = 5$) reflected from a mirror step, for (*a*), (*b*): $\alpha = \frac{1}{4}$; (*c*) $\alpha = \frac{1}{2}$. In (*a*), the upper curve is calculated including evanescent waves with $|K| < 2$, and the lower curve with no evanescent waves ($|K| < 1$). In (*b*) and (*c*) the dashed curves show the paraxial approximation (14).

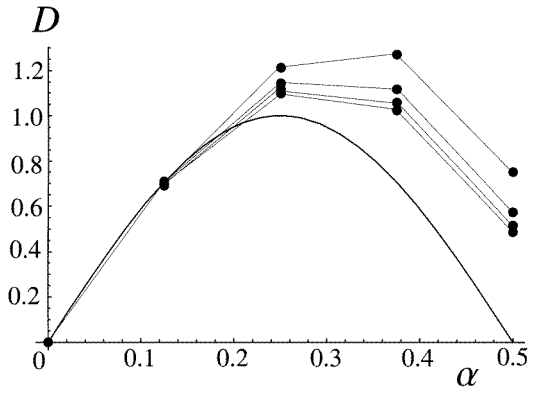

Figure 11. Beam deflection $w\sqrt{\pi}D$ from mirror steps with height $\pi\alpha$. Filled circles joined by straight lines, for (top to bottom): $w = 5$, 10, 20, 60; smooth curve: $w\sqrt{\pi}D_{\text{paraxial}} = \sin(2\pi\alpha)$.

However, a simple calculation gives

$$N(\gamma) = \int \mathrm{d}\theta \, \frac{|a(\theta)|^2 + a^*(\theta)a(2\gamma - \theta)\exp\{\mathrm{i}x_0(\cos(2\gamma - \theta) - \cos\theta)\}}{|\sin(\theta - \gamma)|} \tag{A4}$$

showing that, because of the singularity of the denominator, $N(\gamma)$ diverges unless $a(\gamma)$ vanishes. This has the simple interpretation that Ψ is not normalizable on any slice for which there is a plane wave in the superposition. Imagine now that the incident wave is such that an interval of angles $\Delta\theta$ is absent from the superposition, so that normalization is possible for slices in directions γ within $\Delta\theta$. Then scattering will almost surely send waves into this interval, destroying the normalization.

It is therefore not surprising that the AB wave (19) and (20) is not normalizable in any direction. The same is true for the paraxial wave (13) and (15), except for the slice $\gamma = \pi/2$, where the transmitted wave (8), immediately beyond the phase-changing sheet at $x = 0$, is normalizable.

Without normalization, expectation values of operators, for example the transverse current in the slice at x (expectation value of $\delta(\hat{x} - x)\hat{p}_y$, where carats denote operators and p is momentum) cannot be calculated.

References

Abramowitz M and Stegun I A 1972 *Handbook of Mathematical Functions* (Washington, DC: National Bureau of Standards)

Aharonov Y and Bohm D 1959 Significance of electromagnetic potentials in the quantum theory *Phys. Rev.* **115** 485–91

Berry M V 1980 Exact Aharonov–Bohm wavefunction obtained by applying Dirac's magnetic phase factor *Eur. J. Phys.* **1** 240–4

——1994 Evanescent and real waves in quantum billiards, and Gaussian beams *J. Phys. A: Math. Gen.* **27** L391–8

——1998 Wave dislocations in nonparaxial Gaussian beams *J. Mod. Opt.* **45** 1845–58

Chambers R G 1960 *Phys. Rev. Lett.* **5** 3–5

Iordanskii S V 1966 Mutual friction force in a rotating Bose gas *Sov. Phys.–JETP* **22** 160–7

Lighthill M J 1958 *Introduction to Fourier Analysis and Generalized Functions* (Cambridge: Cambridge University Press)

Olariu S and Popescu I I 1985 The Aharonov–Bohm effect *Rev. Mod. Phys.* **57** 339–436

Shalenkov A 1998 Magnetic force exerted by the Aharonov–Bohm flux line *Europhys. Lett.* **43** 623–8

Sonin E B 1997 Magnus force in superfluids and superconductors *Phys. Rev.* B **55** 485–501