PROCEEDINGS OF THE ROYAL SOCIETY A

Proc. R. Soc. A (2007) **463**, 1697–1711
doi:10.1098/rspa.2007.1842
Published online 17 April 2007

Wave dislocation lines threading interferometers

By M. V. Berry

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

In interferometers where a wave is divided into two beams that propagate along separate branches before being recombined, the closed circuit formed by the two branches must be threaded by wave dislocation lines. For a large class of interferometers, it is shown that the (signed) dislocation number, defined in a suitable asymptotic sense, jumps by $+1$ as the phase difference between the beams increases by 2π. The argument is based on the single-valuedness of the wave function in the branches and leaking between them. In some cases, the jumps occur when the phase difference is an odd multiple of π. The same result holds for the Aharonov–Bohm wave function, where the waves passing above and below a flux line experience different phase shifts; in this case, where the wave is not concentrated onto branches, the threading dislocations coincide with the flux line.

Keywords: singularities; vortices; phase; interference; topology

1. Introduction

In simple interferometers, a beam of light or matter waves, here assumed to be scalar or linearly polarized, is split into two at the entrance port. Each beam accumulates a different phase shift while propagating, essentially one dimensionally, along a different branch of the interferometer. The beams are recombined at the exit port, where the phase difference Γ is detected, for example by the shift of an interference pattern beyond the exit port. My purpose here is to point out a curious consequence of the fact that the wave cannot be completely confined to the two branches, so there will always be some leakage between them, especially near the entrance and exit ports, and a wave Ψ, albeit of very low intensity, must exist throughout the three-dimensional space in which the interferometer is situated.

The wave Ψ—a complex scalar field—must be single-valued in the space between the branches. This will be true whatever the nature of the intervening medium (refracting, reflecting, absorbing, active or inactive), or the mechanism of propagation (e.g. diffraction, scattering, tunnelling), or the wave equation (linear or nonlinear). It follows that the total phase change around any circuit formed by the branches of the interferometer must be an integer multiple S of 2π. S is the total strength of wavefront dislocation lines (Nye & Berry 1974; Nye 1999) threading the interferometer. On a dislocation line (also called a nodal line or a wave vortex), the phase $\arg\Psi$ is singular.

Received 19 January 2007
Accepted 6 March 2007

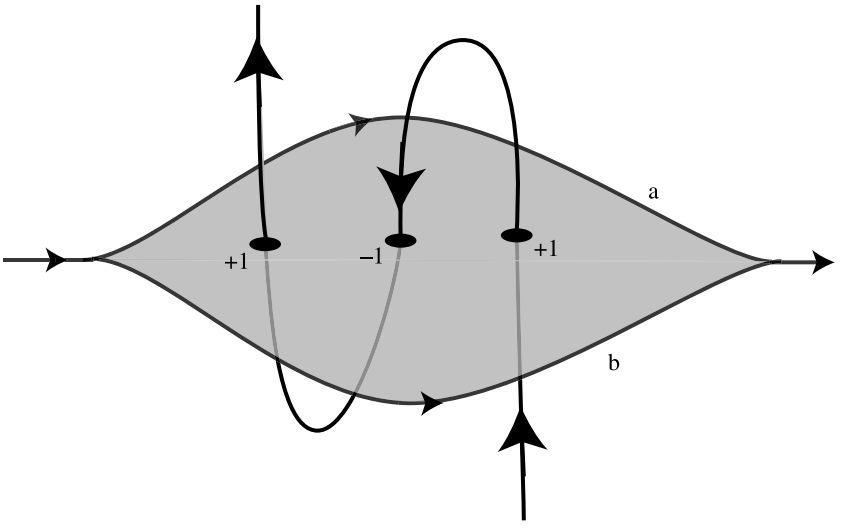

Figure 1. A dislocation line piercing an interferometer, with total strength $S=+1$.

The threading dislocations will pierce any surface spanning the branches (figure 1), at points P around which $\arg\Psi$ increases by $2\pi n$, in a sense defined by a normal to the surface; usually $n=\pm 1$. S is the sum of the (signed) strengths n over all P. In general, the phase difference Γ measured by the interferometer will not be an integer multiple of 2π, so $S\neq\Gamma/2\pi$. Therefore, it is interesting to ask: how does S depend on Γ?

The question is imprecise unless we specify more closely what is meant by the 'circuit formed by the branches'. In order for the beams in an interferometer not to spread by diffraction, they must be many wavelengths wide; hence, the branches cannot be regarded simply as lines without thickness. What follows will be an asymptotic theory, in which the ratio ε of the width of the beams to their total length is small. Thus, the overlap between the beams, necessary for Ψ to be single-valued, is confined to the regions ε close to the entrance and exit ports, and the circuit can be any closed curve (as in figure 1) that does not stray outside the beams.

As parameters describing the interferometer vary, S can change only by a dislocation entering or leaving the circuit while passing through one of the branches of the interferometer. After giving a general theory (§2), I will show that under reasonable assumptions, S jumps by $+1$ as Γ increases by 2π, i.e.

$$S(\Gamma+2\pi)-S(\Gamma)=1. \tag{1.1}$$

In some cases (§§3 and 5), Γ is the only relevant parameter determining S, and a model (§3) for a wide class of interferometers shows that the jumps in S occur as Γ passes half-integers. Mach–Zehnder interferometers conform to equation (1.1), but the details are slightly different: as a separate treatment (§4) shows, the jumps occur not at half-integers but at values of Γ that depend on the phases in the individual branches, as well as the reflection and transmission coefficients of the beam splitters.

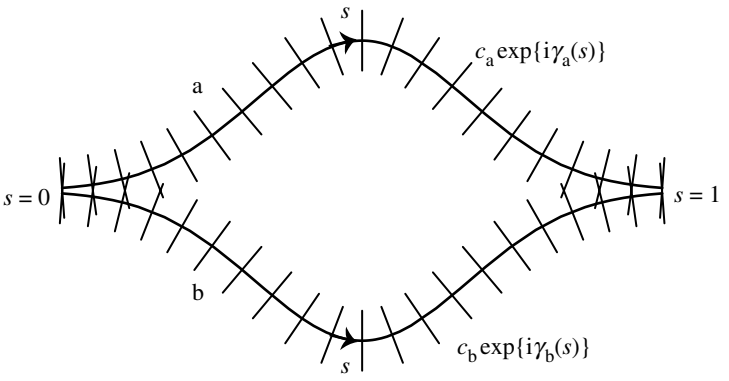

Figure 2. Schematic two-beam interferometer, showing the unperturbed waves with amplitudes c_a and c_b localized in each branch and propagating essentially one dimensionally. Near the entrance and exit ports, the wave in each branch wave must incorporate leakage between the branches, giving a single-valued total wave Ψ.

This phenomenon is not restricted to conventional interferometers. For example, the Aharonov–Bohm effect (AB; Aharonov & Bohm 1959), where a localized magnetic flux line scatters a beam of electrons incident on it, can be regarded as an interferometer inducing a phase difference between waves passing on either side of the flux line (Ehrenberg & Siday 1949; Feynman *et al.* 1965). In terms of the quantum flux $\alpha(\equiv \text{charge} \times \text{magnetic flux}/\hbar)$, the phase difference $2\pi\Gamma = 2\pi\alpha$. In the AB situation, the interfering waves are distributed throughout space, rather than being concentrated in branches. Therefore, we can choose as the 'interferometer circuit' any loop enclosing the flux line. Nevertheless (§5), the result (1.1) holds and, reinterpreting an old result (Berry *et al.* 1980), all the dislocation strengths are concentrated onto the flux line.

2. General theory

Referring to figure 2 and with the chosen circuit, let s be a parameter measuring the distance along the branches a and b of the interferometer, with entrance port $s=0$ and exit port $s=1$, and let the wave functions in the branches be $\Psi_a(s)$ and $\Psi_b(s)$, where $\Psi_a(0)=\Psi_b(0)$ and $\Psi_a(1)=\Psi_b(1)$, reflecting single-valuedness. Then, the dislocation strength around the circuit from $s=0$ to $s=1$ on branch b and from $s=1$ back to $s=0$ on branch a is

$$S = \frac{1}{2\pi}\int_0^1 \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}[\arg\Psi_b(s) - \arg\Psi_a(s)] = \frac{1}{2\pi}\int_0^1 \mathrm{d}s\, \mathrm{Im}\left[\frac{\partial_s\Psi_b(s)}{\Psi_b(s)} - \frac{\partial_s\Psi_a(s)}{\Psi_a(s)}\right], \tag{2.1}$$

where, here and henceforth, subscripts denote differentiation. We will use this formula for numerical calculations.

The only mechanism by which S can change as parameters vary is by Ψ becoming zero as a dislocation crosses either branch. Only one parameter needs to vary, and the parameter we are interested in (which as we will see is sometimes the only relevant one) is the phase difference Γ. Let a dislocation cross one of the branches at $s=s_c$,

$\Gamma=\Gamma_{\rm c}$. We seek the sign of the change ΔS as Γ increases through $\Gamma_{\rm c}$. $\Delta S=+1(-1)$ corresponds to an $n=+1$ dislocation entering (leaving) the interferometer, or an $n=-1$ leaving (entering). We will sometimes incorporate the sign of n and refer to $\Delta S=+1(-1)$ simply as a dislocation entering/leaving.

Consider the case where the crossing is on the b branch, on which s increases. In the space s, Γ, the change ΔS is $1/2\pi$ times the difference between the phase change accumulated passing just above $s_{\rm c}$ (i.e. $\Gamma>\Gamma_{\rm c}$) and that just below $s_{\rm c}$ (i.e. $\Gamma<\Gamma_{\rm c}$). Thus, ΔS is minus the dislocation strength, defined as $1/2\pi$ times the phase change in a positive circuit of the dislocation. If the dislocation crosses the branch a, on which s decreases, the sign is reversed. Thus, using a well-known formula for dislocation strength (Berry 1998, 2004),

$$\Delta S=\mathrm{sgn}\,\mathrm{Im}(\partial_s\Psi^*(s_{\rm c},\Gamma_{\rm c})\partial_\Gamma\Psi(s_{\rm c},\Gamma_{\rm c}))\times\begin{pmatrix}+1(\text{branch a})\\-1(\text{branch b})\end{pmatrix}. \tag{2.2}$$

We will use this formula in later sections.

For the unperturbed waves in the branches, ignoring the leakage between them, we write (figure 2)

$$\Psi_{\rm a0}(s)=c_{\rm a}\exp\{{\rm i}\gamma_{\rm a}(s)\},\quad \Psi_{\rm a0}(s)=c_{\rm b}\exp\{{\rm i}\gamma_{\rm b}(s)\}. \tag{2.3}$$

The picture envisaged here is that a single locally plane wave propagates in each branch; the amplitudes are $c_{\rm a}$ and $c_{\rm b}$, and the evolving phases are $\gamma_{\rm a}(s)$ and $\gamma_{\rm b}(s)$. These must satisfy (figure 3*a*)

$$\gamma_{\rm a}(0)=\gamma_{\rm b}(0)=0,\quad \gamma_{\rm a}(1)=\Gamma_{\rm a},\quad \gamma_{\rm b}(1)=\Gamma_{\rm b}, \tag{2.4}$$

where $\Gamma_{\rm a}$ and $\Gamma_{\rm b}$ are the phases accumulated along the branches, so the phase difference is

$$\Gamma_{\rm b}-\Gamma_{\rm a}\equiv\Gamma. \tag{2.5}$$

In numerical computations, we will use the simple model

$$\gamma_{\rm a}(s)=s\Gamma_{\rm a},\quad \gamma_{\rm b}(s)=s\Gamma_{\rm b}. \tag{2.6}$$

The expressions (2.3) approximate the waves everywhere, except near the entrance and exit ports. Near $s=0$ and $s=1$, the full waves $\Psi_{\rm a}(s)$ and $\Psi_{\rm b}(s)$ differ from $\Psi_{\rm a0}(s)$ and $\Psi_{\rm b0}(s)$ because of the leakage, which is necessary for the wave in the space containing the interferometer to be single-valued.

We model the leakage with a function $f(s,\varepsilon)$ (figure 3*b*), in which ε represents the ratio of the extent of the leakage region to the total length of the branches. We shall consider $\varepsilon\ll1$, corresponding to an interferometer with well-separated branches, for which $f\ll1$ except near $s=0$ and $s=1$, where $f=1$. In the numerical computations below, I will use

$$f(s,\varepsilon)=g(s,\varepsilon)+g(1-s,\varepsilon),\quad \text{where } g(s,\varepsilon)=\exp\left\{-\frac{s^2}{2\varepsilon^2}\right\}. \tag{2.7}$$

3. Simple interferometer model

In this section, we regard the leakage as a direct contamination of the wave in each branch from that in the other. This simple model applies, for example, to the Young and Rayleigh interferometers (Born & Wolf 2005), as well as single-

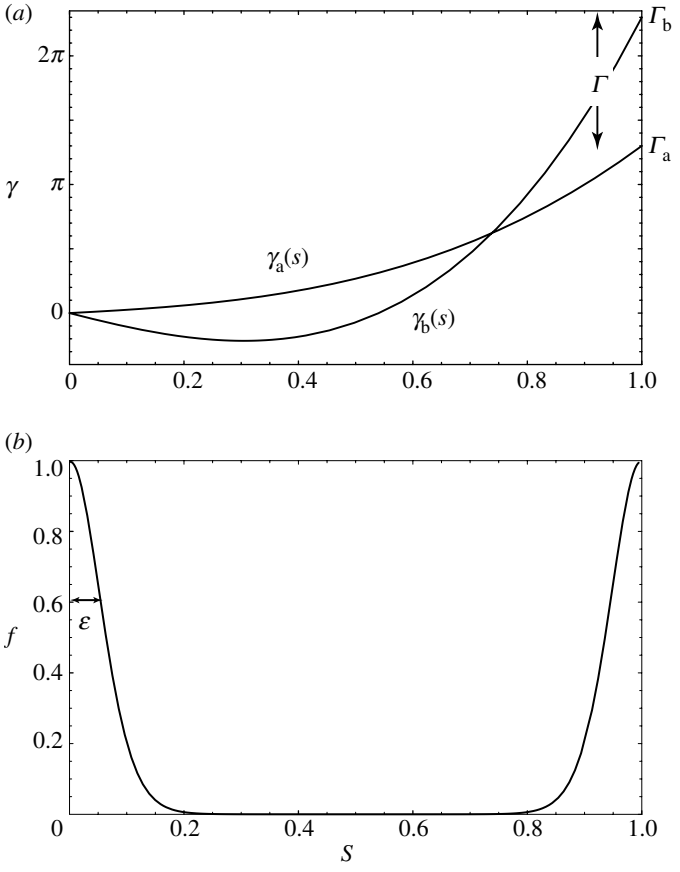

Figure 3. (*a*) Phases developing along the branches of the interferometer and (*b*) function describing leakage between branches near the entrance and exit ports.

crystal neutron (Rauch & Werner 2000) and X-ray (Hart 1975) interferometers. Thus, the total wave in each branch is given by the following modification of the unperturbed expression (2.3):

$$\begin{aligned}\Psi_{\mathrm{a}}(s) &= c_{\mathrm{a}}\exp\{\mathrm{i}\gamma_{\mathrm{a}}(s)\} + f(s,\varepsilon)c_{\mathrm{b}}\exp\{\mathrm{i}\gamma_{\mathrm{b}}(s)\}\\ \Psi_{\mathrm{b}}(s) &= c_{\mathrm{b}}\exp\{\mathrm{i}\gamma_{\mathrm{b}}(s)\} + f(s,\varepsilon)c_{\mathrm{a}}\exp\{\mathrm{i}\gamma_{\mathrm{a}}(s)\}.\end{aligned} \tag{3.1}$$

Ψ_{a} and Ψ_{b} agree at $s=0$ and $s=1$, as they must. Equation (3.1) can be simplified by defining the amplitude ratio and evolving phase difference

$$r \equiv \frac{c_{\mathrm{b}}}{c_{\mathrm{a}}}, \quad \gamma(s) \equiv \gamma_{\mathrm{b}}(s) - \gamma_{\mathrm{a}}(s), \tag{3.2}$$

giving

$$\begin{aligned}\Psi_{\mathrm{a}}(s) &= c_{\mathrm{a}}\exp\{\mathrm{i}\gamma_{\mathrm{a}}(s)\}[1 + rf(s,\varepsilon)\exp\{\mathrm{i}\gamma(s)\}]\\ \Psi_{\mathrm{b}}(s) &= c_{\mathrm{a}}\exp\{\mathrm{i}\gamma_{\mathrm{a}}(s)\}[r\exp\{\mathrm{i}\gamma(s)\} + f(s,\varepsilon)].\end{aligned} \tag{3.3}$$

Dislocations are determined by the vanishing of the factors in square brackets.

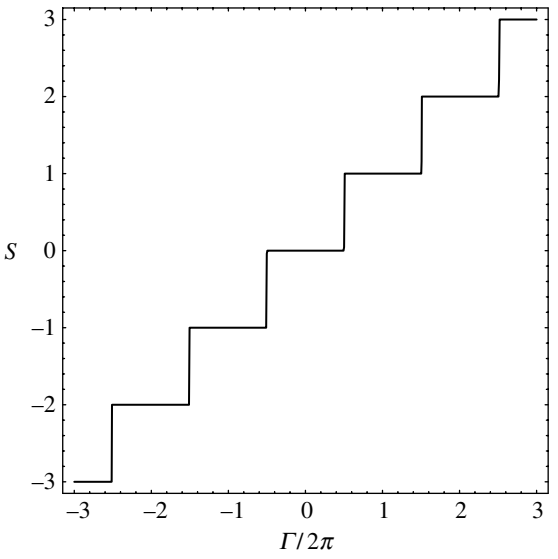

Figure 4. Dislocation strength S versus phase difference Γ, for amplitude ratio $r=0.6$, and leakage parameter $\varepsilon=0.01$, for the simple interferometer of §2.

Thus, the dislocation strength (2.1) becomes

$$S=\frac{1}{2\pi}\mathrm{Im}\int_0^1 \mathrm{d}s\left(\frac{[\mathrm{i}r\partial_s\gamma(s)\exp\{\mathrm{i}\gamma(s)\}+\partial_s f(s,\varepsilon)]}{r\exp\{\mathrm{i}\gamma(s)\}+f(s,\varepsilon)}-\frac{r[\partial_s f(s,\varepsilon)+\mathrm{i}f(s,\varepsilon)\partial_s\gamma(s)]}{\exp\{-\mathrm{i}\gamma(s)\}+rf(s,\varepsilon)\exp\{\mathrm{i}\gamma(s)\}}\right). \tag{3.4}$$

Figure 4 shows a numerical example using the models (2.6) and (2.7). Evidently, S increases by unity as Γ increases by 2π, as claimed in equation (1.1). Moreover, the jumps occur as Γ increases through $\Gamma_\mathrm{c}=(2n+1)\pi$.

To explain these observations, we must determine the zeros of Ψ_a and Ψ_b. From equation (3.3), destructive interference can occur only near the exit port, and there are two possibilities

$$\begin{aligned}&\text{zero on branch a, at } s_\mathrm{c}: f(s_\mathrm{c},\varepsilon)=\frac{1}{r}\quad(\text{if } r>1),\\ &\text{or zero on branch b, at } s_\mathrm{c}: f(s_\mathrm{c},\varepsilon)=r\quad(\text{if } r<1),\\ &\text{and } \gamma(s_\mathrm{c})=(2n+1)\pi.\end{aligned} \tag{3.5}$$

Now,

$$\text{as } \varepsilon\to 0,\quad s_\mathrm{c}\to 1\quad\text{and}\quad \gamma(s_\mathrm{c})\to\Gamma_\mathrm{c}=(2n+1)\pi, \tag{3.6}$$

hence, zero-crossing requires Γ_c to be a half-integer multiple of 2π, and the crossings occur near the exit port, as claimed. With the particular phase evolution (2.5) and leakage function (2.6), the approach to these asymptotic values is

$$s_\mathrm{c}=1-\varepsilon\sqrt{2\log(1/r)},\quad \Gamma_\mathrm{c}=\frac{n+1/2}{1-\varepsilon\sqrt{2\log(1/r)}}. \tag{3.7}$$

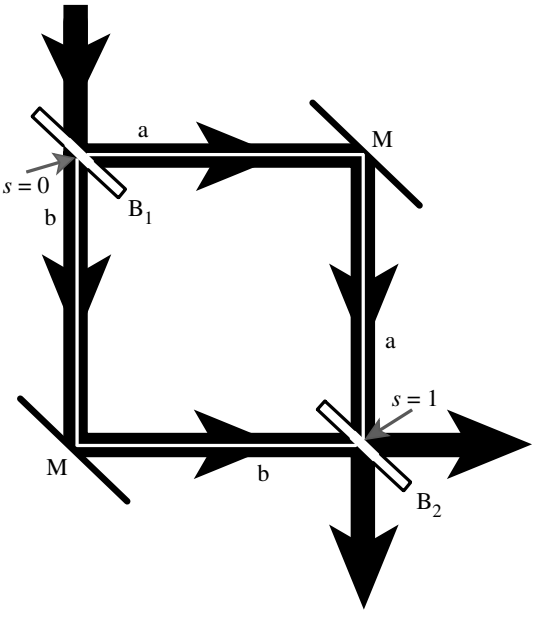

Figure 5. Schematic of Mach–Zehnder interferometer, with mirrors M and beam splitters B_1 and B_2. The white square denotes the notional circuit within which the threading dislocations are counted.

To determine whether the change of dislocation strength ΔS is $+1$ or -1 as Γ increases through Γ_c, we use the formula (2.2). Evaluating the derivatives of Ψ_a and Ψ_b, together with the information that near $s=1$ both $\gamma(s)$ and $f(s)$ are increasing functions of s, and $\gamma(s)$ is also an increasing function of Γ and $\exp\{i\gamma(s_c)\}$ is negative real, we find, for both the a and b branch crossings,

$$\Delta S = +1. \tag{3.8}$$

To determine the absolute value of S, as well as its increase as Γ passes Γ_c, we study the case $\Gamma=0$. Consider $r<1$ (an analogous argument works for $r>1$), then it follows from equation (3.3) that the only possibility of Ψ winding round the origin, i.e. of a non-zero S, is on the b branch. But such winding does not occur, because $\gamma(s)$ is close to zero near the entrance and exit ports $s=0$ and $s=1$, and for $\varepsilon \ll 0$ any possible large excursions in $\Gamma(s)$ must take place in regions where $f \ll 1$; it follows that $\mathrm{Im}\,\Psi \ll \mathrm{Re}\,\Psi$, thus excluding winding. We conclude that for this class of interferometers, the dislocation strength is independent of the amplitude ratio r and also independent of ε when $\varepsilon \ll 1$, and is given by

$$S = \text{closest integer to } \frac{\Gamma}{2\pi}, \tag{3.9}$$

as figure 4 illustrates.

4. Mach–Zehnder interferometer

With this type of interferometer (figure 5), and the similar Jamin and Sirks–Pringsheim interferometers (Born & Wolf 2005), the contamination is different from the case just considered, in which the a and b beams are contaminated with

the b and a beams. To understand what happens instead, consider the circuit indicated in figure 5, corresponding to the centre lines of the beams in the branches. At the entrance port, the 'a' beam is contaminated by the wave incident on the beam splitter B_1, and the 'b' beam is uncontaminated; at the exit port, the 'a' beam is contaminated by the 'a' wave reflected by B_2 and the 'b' wave transmitted by B_2, and the 'b' beam is contaminated by the 'a' wave transmitted by B_2 and the 'b' wave reflected by B_2.

We denote the complex reflection and transmission amplitudes of the beam splitters by ρ and τ. In this section, it is convenient to employ the leakage functions $g(s)$ and $g(1-s)$ for the entrance and exit ports individually, as in equation (2.7) (which corresponds to Gaussian beams). Thus, the waves in the branches can be written as

$$\begin{aligned}\Psi_{\mathrm{a}} &= \rho\exp\{\mathrm{i}\gamma_{\mathrm{a}}(s)\} + g(s,\varepsilon) + \left(\rho^2\exp\{\mathrm{i}\Gamma_{\mathrm{a}}\} + \tau^2\exp\{\mathrm{i}\Gamma_{\mathrm{b}}\}\right)g(1-s,\varepsilon)\\ \Psi_{\mathrm{b}} &= \tau\exp\{\mathrm{i}\gamma_{\mathrm{b}}(s)\} + \rho\tau(\exp\{\mathrm{i}\Gamma_{\mathrm{a}}\} + \exp\{\mathrm{i}\Gamma_{\mathrm{b}}\})g(1-s,\varepsilon).\end{aligned} \tag{4.1}$$

In the simplest model, the beam splitters can be regarded as infinitely thin, reversible and non-absorbing, so that the incoming and outgoing waves on the two sides are connected by a unitary matrix. Furthermore, single-valuedness demands that the total wave (incoming+outgoing) must be the same on both sides. These requirements imply the following relations between ρ and τ:

$$|\tau|^2 + |\rho|^2 = 1, \quad \mathrm{Re}\,\rho^*\tau = 0, \quad 1+\rho = \tau. \tag{4.2}$$

The unique solution, involving a single parameter θ, is

$$\rho = -\exp(\mathrm{i}\theta)\cos\theta, \quad \tau = -\mathrm{i}\exp(\mathrm{i}\theta)\sin\theta. \tag{4.3}$$

These expressions are periodic in θ with period π.

For simplicity, we henceforth consider the 'democratic' case $|c_{\mathrm{a}}|^2 = |c_{\mathrm{b}}|^2 = 1/2$, where $\theta = \pi/4$, so

$$\rho = -\frac{\exp(\mathrm{i}\pi/4)}{\sqrt{2}} = -\frac{1}{2}(1+\mathrm{i}), \quad \tau = \frac{\exp(-\mathrm{i}\pi/4)}{\sqrt{2}} = \frac{1}{2}(1-\mathrm{i}). \tag{4.4}$$

The qualitative behaviour of the dislocation strength is similar for other values of θ. In addition, we use the model phase and leakage functions (2.6) and (2.7). Then, the waves (4.1) become, after neglecting a common factor,

$$\begin{aligned}\Psi_{\mathrm{a}} &= -\exp\left\{\mathrm{i}\left(\frac{1}{4}\pi + \Gamma_{\mathrm{a}}(s-1) - \frac{1}{2}\Gamma\right)\right\} + \sqrt{2}\exp\left\{-\mathrm{i}\left(\Gamma_{\mathrm{a}} + \frac{1}{2}\Gamma\right)\right\}\exp\left\{-\frac{s^2}{2\varepsilon^2}\right\}\\ &\quad + \sqrt{2}\sin\left(\frac{1}{2}\Gamma\right)\exp\left\{-\frac{(1-s)^2}{2\varepsilon^2}\right\}\\ \Psi_{\mathrm{b}} &= \exp\left\{\mathrm{i}\left(-\frac{1}{4}\pi + \Gamma_{\mathrm{a}}(s-1) + \Gamma\left(s-\frac{1}{2}\right)\right)\right\} - \sqrt{2}\cos\left(\frac{1}{2}\Gamma\right)\exp\left\{-\frac{(1-s)^2}{2\varepsilon^2}\right\}.\end{aligned} \tag{4.5}$$

Now, there is a non-trivial dependence on the phase Γ_{a} as well as the phase difference Γ.

Near the entrance port, the branch b consists of just the transmitted wave, so there is no interference; therefore, Ψ_{b} cannot vanish and no dislocations cross this branch. For the a branch, equation (4.5) gives, for $\varepsilon \ll 1$ and again neglecting an overall factor,

$$\Psi_{\mathrm{a}} \approx -\exp\left\{\mathrm{i}\left(\frac{1}{4}\pi + \Gamma_{\mathrm{a}} s\right)\right\} + \sqrt{2}\,\exp\left\{-\frac{s^2}{2\varepsilon^2}\right\} \quad (s \ll 1). \tag{4.6}$$

It follows that dislocations can cross the branch a as Γ_{a} varies, but these events are independent of Γ, as is physically obvious. Moreover, the crossings can occur only when $s < \varepsilon\sqrt{(\log 2)}$, so the corresponding values of Γ_{a} are separated by phase differences of order $1/\varepsilon$. We do not consider such crossings further.

Near the exit port,

$$\left.\begin{aligned}
\Psi_{\mathrm{a}} &\approx -\exp\left\{\mathrm{i}\left(\frac{1}{4}\pi + \Gamma_{\mathrm{a}}(s-1) - \frac{1}{2}\Gamma\right)\right\} + \sqrt{2}\sin\left(\frac{1}{2}\Gamma\right)\exp\left\{-\frac{(1-s)^2}{2\varepsilon^2}\right\} \\
\Psi_{\mathrm{b}} &\approx \exp\left\{\mathrm{i}\left(-\frac{1}{4}\pi + \Gamma_{\mathrm{a}}(s-1) + \frac{1}{2}\Gamma\right)\right\} - \sqrt{2}\cos\left(\frac{1}{2}\Gamma\right)\exp\left\{-\frac{(1-s)^2}{2\varepsilon^2}\right\}
\end{aligned}\right\} \quad (1-s \ll 1). \tag{4.7}$$

It is easy to see that a dislocation crosses the exit port $s=1$ when $\Gamma_{\mathrm{c}} = (2n+1/2)\pi$, since then both Ψ_{a} and Ψ_{b} vanish. From equation (2.2) applied to Ψ_{a} or Ψ_{b}, the change in S when Γ increases through Γ_{c} is easily calculated as

$$\Delta S = \operatorname{sgn} \Gamma_{\mathrm{a}}. \tag{4.8}$$

When $\Gamma_{\mathrm{a}} < 0$, this corresponds to a dislocation leaving the interferometer as Γ increases, rather than entering, and so appears to violate equation (1.1). But we will see that this is not the case, because as Γ changes by 2π, other dislocations can enter and leave the interferometer, in a manner that maintains the validity of equation (1.1).

These additional dislocation events depend on the value of Γ_{a}. It might be thought that this dependence is negligible, because equation (4.7) involves $\Gamma_{\mathrm{a}}(1-s)$ and $1-s \sim \varepsilon \ll 1$. This would be wrong, because Γ_{a} can be so large that $\Gamma_{\mathrm{a}}(1-s)$ can be large too. For an interferometer whose arms are of length l and contain beams of wavenumber k and width w,

$$\Gamma_{\mathrm{a}} \sim kl, \quad \varepsilon \sim \frac{w}{l}, \quad \text{i.e.} \quad \Gamma_{\mathrm{a}}\varepsilon \sim kw \gg 1, \tag{4.9}$$

since the beams must be sufficiently wide to prevent diffraction spreading.

To determine the dislocation crossings Γ_{c}, s_{c}, it is convenient to define the exit-port-crossing variable

$$\sigma \equiv \frac{1-s}{\varepsilon} \quad (\text{branch a}), \quad \sigma \equiv \frac{s-1}{\varepsilon} \quad (\text{branch b}), \tag{4.10}$$

so that the exit port corresponds to $\sigma = 0$, with $\sigma < 0$ on branch b and $\sigma > 0$ on branch a. Then, the vanishing of $\mathrm{Im}\Psi_{\mathrm{a}}$ and $\mathrm{Im}\Psi_{\mathrm{b}}$ in equation (4.7) can be written in the common form

$$\Gamma_{\mathrm{c}} = \left(2n + \frac{1}{2}\right)\pi - 2\Gamma_{\mathrm{a}}\varepsilon\sigma_c, \tag{4.11}$$

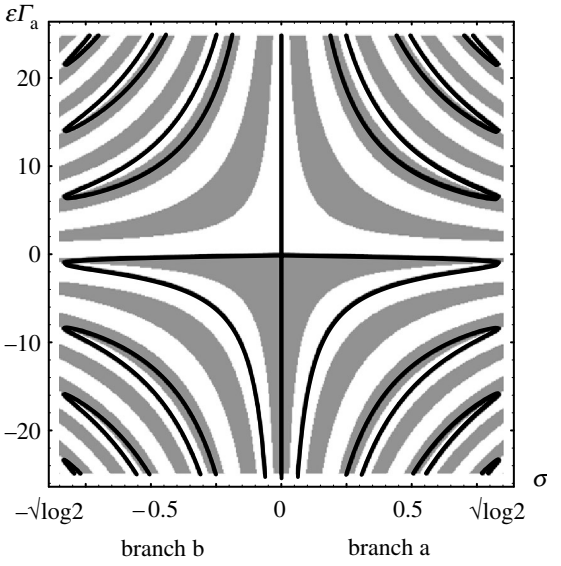

Figure 6. Curves are the loci (4.12) of dislocation-crossing events for the Mach–Zehnder interferometer; for each Γ_a, the values of s indicate where the crossings occur as Γ increases by 2π, and equation (4.11) gives the corresponding phase differences Γ. The white (grey) regions correspond to dislocations entering (leaving) the interferometer according to equation (4.13), i.e. $\Delta S = +1(-1)$.

with the values of σ_c determined by substitution into equation (4.7), which gives

$$\exp\left(\frac{1}{2}\sigma_c^2\right) = \sqrt{2}\sin\left(\frac{1}{4}\pi - \varepsilon\Gamma_a|\sigma_c|\right). \tag{4.12}$$

All solutions must have $\sigma < \sqrt{\log 2}$, i.e. $|1-s| < \varepsilon\sqrt{\log 2}$. Application of equation (2.2), followed by some algebra, now gives the sign of the dislocation jump at Γ_c, s_c as

$$\Delta S = \mathrm{sgn}\left[\varepsilon\Gamma_a \tan\left(\frac{1}{4}\pi - \varepsilon\Gamma_a|\sigma_c| + |\sigma_c|\right)\right]. \tag{4.13}$$

Figure 6 shows the locus (4.12), with the positive and negative regions (4.13) superimposed. It is clear that the number of crossings (as Γ increases by 2π) is odd for any value of Γ_a, and there is always one more dislocation entering than leaving, demonstrating that equation (1.1) holds for Mach–Zehnder interferometers. For small positive values of $\varepsilon\Gamma_a$, there is a single crossing, at the exit port $\sigma = 0$, $\Gamma = \pi/2$, and for small negative values there are one negative and two positive crossings. As $|\Gamma_a|$ increases, additional crossings appear in groups of four, two positive and two negative. The dependence on Γ_a is weak: in order to change the number of crossings as Γ increases by 2π, Γ_a must change by $O(2\pi/\varepsilon)$, which is large.

Figure 7*a–c* shows the locations and signs of dislocation-crossing events in the σ, Γ plane for three cases, and figure 7*d–f* shows how the corresponding dislocation strengths depend on Γ. In figure 7*b*, for example, three dislocations enter and two leave as Γ increases by 2π. In every case, the net number entering is $+1$, as claimed in equation (1.1).

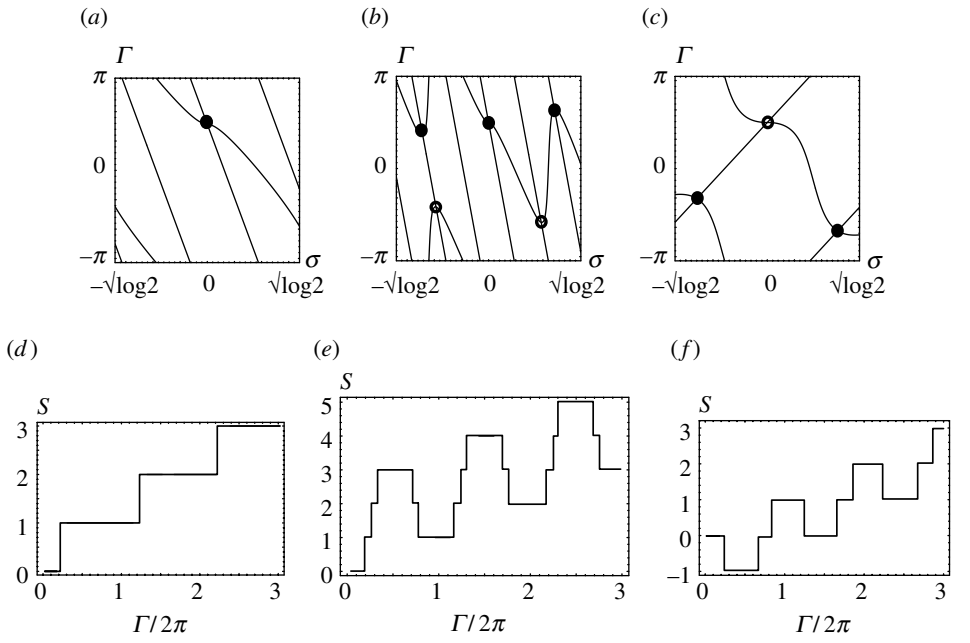

Figure 7. (*a–c*) Nodal lines of $\mathrm{Re}\Psi_{\mathrm{a}}$ and $\mathrm{Im}\Psi_{\mathrm{a}}$ ($s>0$) and $\mathrm{Re}\Psi_{\mathrm{b}}$ and $\mathrm{Im}\Psi_{\mathrm{b}}$ ($s<0$); intersections indicate dislocation-crossing events for Mach–Zehnder interferometer, with filled circles indicating dislocations entering the interferometer and open circles indicating dislocations leaving. (*d–f*) Signed number S of dislocations threading the interferometer, computed from equations (2.1) and (4.5). In this figure, $\varepsilon=0.01$; (*a*,*d*) $\Gamma_{\mathrm{a}}=500$, (*b*,*e*) $\Gamma_{\mathrm{a}}=1000$ and (*c*,*f*) $\Gamma_{\mathrm{a}}=-200$.

5. Aharonov–Bohm interferometer

The single-valued AB wave function in the plane, for a beam (plane wave) of charged quantum particles, incident from negative $x=r\cos\phi$ and diffracted by an isolated magnetic flux line with strength α is in units where the wavenumber is unity (Aharonov & Bohm 1959; Olariu & Popescu 1985),

$$\Psi_{\mathrm{AB}}(r,\phi,\alpha)=\sum_{l=-\infty}^{\infty}(-\mathrm{i})^{|l-\alpha|}(-1)^{l}J_{|l-\alpha|}(r)\exp(\mathrm{i}l\phi). \tag{5.1}$$

The leading-order large-r asymptotics, representing the incident plane wave modified by Dirac's magnetic phase factor, is (Ehrenberg & Siday 1949), for $0<|\phi|\leq\pi$,

$$\Psi_{\mathrm{AB}}(r,\phi,\alpha)\approx\exp\{\mathrm{i}(r\cos\phi+\alpha(\phi-\pi\,\mathrm{sgn}\,\phi))\}. \tag{5.2}$$

This is commonly interpreted interferometrically, as a wave propagating 'above' the flux line, that is for ϕ decreasing from $+\pi$ to zero through positive ϕ, and a wave 'below' the flux line, with ϕ increasing from $-\pi$ to zero through negative ϕ. The phase difference between these waves is Γ, where $\Gamma=2\pi\alpha$, and its observation as a shift of interference fringes is the celebrated AB effect (Olariu & Popescu 1985). Although equation (5.2) is continuous across $\phi=\pm\pi$ (incident direction), the formula possesses an unphysical discontinuity in the forward

direction $\phi=0$, where, as is well known (Olariu & Popescu 1985), the incident and diffracted waves cannot be separated (for the relationship between the single-valued equation (5.1) and the multi-valued equation (5.2); see Berry 1980).

As with other interferometers that we have considered, the phase difference implies that $\Psi_{\rm AB}$ is dislocated. But now there is no natural circuit, and we can define a strength that, in principle, could depend on radius, corresponding to dislocations piercing the plane at different points, namely, for circular loops with radius r,

$$S_{\rm AB}(r,\alpha)=\frac{1}{(2\pi)^2}\,{\rm Im}\int_{-\pi}^{\pi}{\rm d}\phi\,\frac{{\rm d}\Psi_{\rm AB}(r,\phi)/({\rm d}\phi)}{\Psi_{\rm AB}(r,\phi)}. \tag{5.3}$$

In fact, all the dislocation strength in $\Psi_{\rm AB}$ is concentrated on the flux line at $r=0$ and $S_{\rm AB}$, independent of r, is given by equation (1.1). For $r\ll 1$, this was shown directly from equation (5.1) (Berry *et al.* 1980), and a heuristic argument was given for the same result to hold for $r\gg 1$.

A better argument is based on the following modification (Berry & Shelankov 1999) of equation (4.12), uniformly valid in f for large r:

$$\begin{aligned}\Psi_{\rm ABapp}(r,\phi,\alpha)&=\exp\{{\rm i}(r\cos\phi+\alpha\phi)\}\\&\times\left[\cos(\pi\alpha)-{\rm i}\sin(\pi\alpha){\rm erf}\left\{\exp\left(-\frac{1}{4}{\rm i}\pi\right)\sin\left(\frac{1}{2}\phi\right)\sqrt{2r}\right\}\right]\quad(0\le|\phi|\le\pi).\end{aligned} \tag{5.4}$$

(This is the leading order of a series (Berry & Shelankov 1999) in powers of $1/r$, giving an intriguing representation of $\Psi_{\rm AB}$ in terms of the Cornu spiral.) As shown in figure 8, equation (5.4) is extremely accurate. Figure 9 shows a sample wavefront pattern with a dislocation of strength -3; the picture was computed from equation (5.4), but is indistinguishable from a computation based on the exact formula (5.1), showing that the uniform asymptotics is accurate for small r too.

Computations directly with equation (5.3), using either $\Psi_{\rm AB}$ or $\Psi_{\rm ABapp}$, generate graphs identical to figure 4, so we conclude that

$$S_{\rm AB}(r,\alpha)=\text{closest integer to }\alpha. \tag{5.5}$$

As with conventional interferometers, $S_{\rm AB}$ jumps as $\alpha=\Gamma/2\pi$ increases through a half-integer. But the mechanism is different: instead of a dislocation crossing the interferometer circuit, in the AB case there is a nodal line (in the plane) in the forward direction $\phi=0$ when α is a half-integer, enabling each wavefront $\arg\Psi_{\rm AB}=\text{constant}+2n\pi$ to unzip and reconnect with its neighbour $n+1$ as α increases through a half-integer. Figure 8, for $\alpha=-2.51$, shows some wavefronts that have just experienced the reconnection. The unzipping phenomenon was observed in an AB analogue experiment (Berry *et al.* 1980) based on water waves.

6. Concluding remarks

In the phenomenon identified here, the circuit formed by the branches of an interferometer is threaded by dislocation lines, whose signed number changes discontinuously as the phase difference between the arms increases continuously, with a net increase of $+1$ as the phase difference increases by 2π. It occurs in a

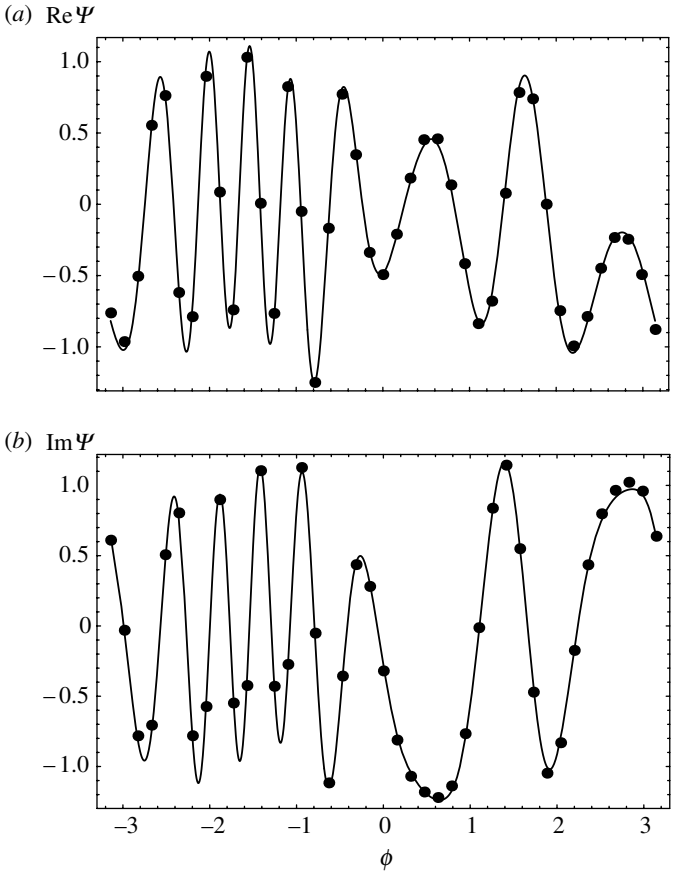

Figure 8. Comparison of Aharonov–Bohm wave function $\Psi_{\mathrm{AB}}(r, \phi, \alpha)$ (full curves) with approximation $\Psi_{\mathrm{ABapp}}(r, \phi, \alpha)$ (dots), for $r=5$, $\alpha=3.7$. (*a*) ReΨ and (*b*) ImΨ.

wide class of interferometers, including that corresponding to the AB effect, where there are no well-defined branches. The only requirement seems to be that there is a single-valued wave function throughout the space occupied by the interferometer.

Is the phenomenon universal? A natural conjecture is that it occurs for all interferometers for which a circuit can be defined. (The 'circuit' stipulation excludes interferometers of the Michelson and Twyman–Green types (Born & Wolf 2005), where, at least in the idealized form, each beam path is self-reciprocal, so the beams do not form a circuit that could be threaded by the dislocations.)

For phase differences Γ not close to the critical values Γ_{c} corresponding to the passage through the branches of the interferometer (e.g. $(2n+1)\pi$ in §§3 and 5), the dislocations can lie deep inside the circuit, where the wave Ψ is very weak, perhaps unobservably so. But when $\Gamma \approx \Gamma_{\mathrm{c}}$, it should be possible to detect the dislocations and follow them as they cross a branch close to the exit port; it would be interesting to do so.

The assumption throughout has been that the waves are scalar or linearly polarized. In optical interferometers where the state of polarization is changed in the branches, the vector nature of the light cannot be ignored, and the analysis

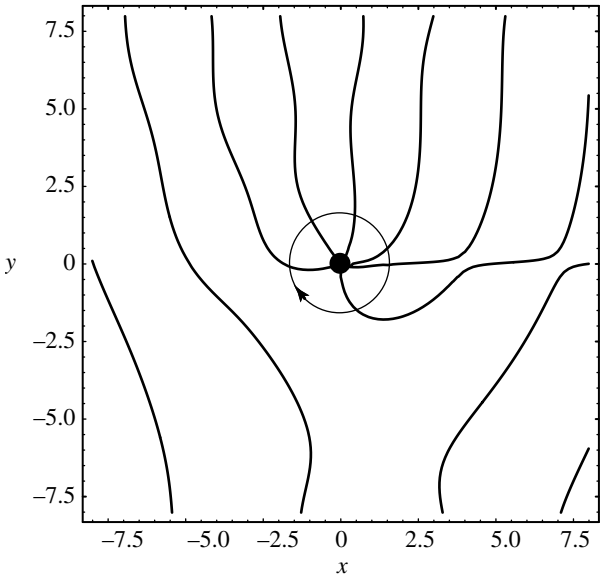

Figure 9. Aharonov–Bohm wavefronts $\mathrm{Im}\Psi_{\mathrm{AB}}=0$ (i.e. crests $\arg\Psi_{\mathrm{AB}}=2n\pi$ and troughs $\arg\Psi_{\mathrm{AB}}=(2n+1)\pi$) for $\alpha=-2.505$, showing a dislocation of strength -3 on the flux line at the origin; the wave is incident from $x=-\infty$, and the arrow denotes the direction of increasing phase.

must incorporate the spatially varying polarization in the region containing the circuit. With polarized light, there are typically no dislocations. Instead, the singularities threading the circuit are C lines, on which the polarization is purely circular, and L lines, on which the polarization is purely linear (Nye & Hajnal 1987; Nye 1999; Berry 2001; Berry & Dennis 2001). The numbers of C and L lines entering and leaving the interferometer during a 2π variation of the phase difference measured by the interferometer remain to be explored. Such a study should incorporate any geometric phase associated with polarization changes (Pancharatnam 1956; Berry 1987; Bhandari 1997).

I am grateful to late Professor Geoffrey Opat for a suggestion, many years ago, that inspired the research reported here, and I would like to acknowledge helpful discussions with Professors John Hannay, John Nye and Sandu Popescu, and Dr Mark Dennis and Mrs P. York.

References

Aharonov, Y. & Bohm, D. 1959 Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.* **115**, 485–491. (doi:10.1103/PhysRev.115.485)

Berry, M. V. 1980 Exact Aharonov–Bohm wave function obtained by applying Dirac's magnetic phase factor. *Eur. J. Phys.* **1**, 240–244. (doi:10.1088/0143-0807/1/4/011)

Berry, M. V. 1987 The adiabatic phase and Panacharatnam's phase for polarized light. *J. Mod. Optic.* **34**, 1401–1407.

Berry, M. V. 1998 Wave dislocations in nonparaxial Gaussian beams. *J. Mod. Optic.* **45**, 1845–1858. (doi:10.1080/095003498150745)

Berry, M. V. 2001 Geometry of phase and polarization singularities, illustrated by edge diffraction and the tides. In *Singular optics 2000*, vol. 4403 (ed. M. Soskin), pp. 1–12. Alushta, Ukraine: SPIE.

Berry, M. V. 2004 Index formulae for singular lines of polarization. *J. Opt. A* **6**, 675–678. (doi:10.1088/1464-4258/6/7/003)

Berry, M. V. & Dennis, M. R. 2001 Polarization singularities in isotropic random vector waves. *Proc. R. Soc. A* **457**, 141–155. (doi:10.1098/rspa.2000.0660)

Berry, M. V. & Shelankov, A. 1999 The Aharonov–Bohm wave and the Cornu spiral. *J. Phys. A* **32**, L447–L455. (doi:10.1088/0305-4470/32/42/101)

Berry, M. V., Chambers, R. G., Large, M. D., Upstill, C. & Walmsley, J. C. 1980 Wavefront dislocations in the Aharonov–Bohm effect and its water-wave analogue. *Eur. J. Phys.* **1**, 154–162. (doi:10.1088/0143-0807/1/3/008)

Bhandari, R. 1997 Polarization of light and topological phases. *Phys. Rep.* **281**, 1–64. (doi:10.1016/S0370-1573(96)00029-4)

Born, M. & Wolf, E. 2005 *Principles of optics*. London, UK: Pergamon.

Ehrenberg, W. & Siday, R. E. 1949 The refractive index in electron optics and the principles of dynamics. *Proc. Phys. Soc. B* **62**, 8–21. (doi:10.1088/0370-1301/62/1/303)

Feynman, R., Leighton, R. B. & Sands, M. 1965 *The Feynman lectures on physics*. Boston, MA: Addison-Wesley.

Hart, M. 1975 Ten years of X-ray interferometry. *Proc. R. Soc. A* **346**, 1–22. (doi:10.1098/rspa.1975.0163)

Nye, J. F. 1999 *Natural focusing and fine structure of light: caustics and wave dislocations*. Bristol, UK: Institute of Physics Publishing.

Nye, J. F. & Berry, M. V. 1974 Dislocations in wave trains. *Proc. R. Soc. A* **336**, 165–190. (doi:10.1098/rspa.1974.0012)

Nye, J. F. & Hajnal, J. V. 1987 The wave structure of monochromatic electromagnetic radiation. *Proc. R. Soc. A* **409**, 21–36. (doi:10.1098/rspa.1987.0002)

Olariu, S. & Popescu, I. I. 1985 The quantum effects of electromagnetic fluxes. *Rev. Mod. Phys.* **57**, 339–436. (doi:10.1103/RevModPhys.57.339)

Pancharatnam, S. 1956 Generalized theory of interference, and its applications. Part I. Coherent pencils. *Proc. Ind. Acad. Sci. A* **44**, 247–262.

Rauch, H. & Werner, S. A. 2000 *Neutron interferometry*. Oxford, UK: Clarendon Press.