JOURNAL OF MODERN OPTICS, 1998, VOL. 45, NO. 9, 1845–1858

Wave dislocation reactions in non-paraxial Gaussian beams

M. V. BERRY

H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, England

(*Received 21 October 1997*)

Abstract. For exact (i.e., non-paraxial) waves ψ representing freely propagating Gaussian beams in two and three dimensions, the patterns of phase singularities, that is zeros of ψ, are studied in detail. The zeros (points in two dimensions, and rings in three) are phase dislocations (optical vortices). The waves depend on a single parameter L, representing the radius of the waist of the beam. As L increases, pairs of dislocations interact and depart from the focal plane. Each such interaction comprises three events where the phase topology of ψ changes; each event is a reaction between the dislocations and associated phase saddles, conserving two topological quantum numbers. The same behaviour was predicted and observed by Karman *et al.* for beams truncated by apertures. The geometrical sensitivity of the wave to L is astonishing: changes in phase topology can occur when the waist expands by a few thousandths of a wavelength. The integral representing ψ is evaluated asymptotically, leading to a global explanation of the dislocations and topological events in terms of interference between complex and diffracted rays. Locally, all details of the topological changes can be captured by a local model, constructed using a classification devised by Nye.

1. Introduction

Several recent papers [1–3] reported the theoretical prediction and experimental discovery of remarkably rich phase dislocation structures in and near the focal plane of Gaussian beams truncated by apertures. The dislocations are rings: line zeros of the complex scalar wave representing the light. Around these lines, the phase changes by an integer multiple (typically ± 1) of 2π, so the wave fronts (surfaces of constant phase modulo 2π) are singular [4]. As the size of the aperture is varied, the dislocations change their topology, by splitting into three and annihilating in pairs. Nye [5] has devised local models of these events, in terms of an ingenious classification scheme reminiscent of catastrophe theory.

Here I shall show that the same patterns of singularities occur in a light field describing pure Gaussian beams (in two and three dimensions) that are not stopped by an aperture. These are exact—that is, not paraxial—solutions of the wave equation for light of wavelength $\lambda = 2\pi/k$. The Gaussian beams depend on a single dimensionless parameter L, namely the radius of the waist of the beam, measured in units of $1/k$. As L changes, the pattern of singularities can change very rapidly. In particular, the sequence of topological events corresponding to the dislocation reactions takes place when the beam radius changes by a few thousandths of a wavelength. Even for fixed L the waves can possess topologically interesting structure on scales much smaller than a wavelength. Karman *et al.* [2]

0950–0340/98 $12·00

pointed out that, since the beams are superpositions of real—that is propagating, rather than evanescent—waves, this fine detail is an example of superoscillation, that is of functions varying faster than their fastest Fourier component [6]. All the dislocations are of edge type and therefore different from the screw dislocations observed along the axes of Gauss–Laguerre beams [7–9]

Several levels of analytical description are possible for these waves. The integrals that represent them can be evaluated asymptotically in the focal plane, leading to a global picture of the dislocation patterns for all L in terms of the interference of complex geometrical rays and marginally evanescent diffracted rays. For small L, the dislocations form concentric rings in the focal plane, all with the same sign. As L increases, the dislocations split off from the focal plane in pairs, starting with the innermost. Locally, the dislocation reactions involved in each splitting can be analyzed in detail, in a slightly modified application of Nye's [5] scheme. There are three critical events:

(a) interaction of one of a pair of dislocations in the focal plane with a phase saddle that acts as a catalyst, enabling this dislocation to switch sign while spawning two satellites that move away from the focal plane;
(b) collision and separation of two saddles;
(c) annihilation of the sign-switched dislocation, along with two saddles, with the other dislocation in the original pair.

Earlier [6], I studied different representations of non-paraxial Gaussian beams. Those too were exact and non-singular solutions of the wave equation and also possessed dislocations (but confined to the focal plane, and not interacting). However, they were physically unsatisfactory because they included weak waves travelling backwards from infinity as well as the main beam travelling forwards. This feature is avoided in the new beams, which are formally similar to those devised by Karman *et al.* [2] and Beijersbergen [3], albeit differently interpreted.

2. Phase singularities and their two quantum numbers

Let the wave be represented by a complex scalar field

$$\psi(\mathbf{r}) = \rho(\mathbf{r}) \exp[\mathrm{i}\chi(\mathbf{r})] \tag{1}$$

with modulus ρ and phase χ. In two dimensions, the amplitude of a wave polarized perpendicular to the plane satisfies

$$\nabla^2\psi + k^2\psi = 0. \tag{2}$$

In three dimensions, solutions ψ of the same equation could represent a scalar approximation, in which polarization is neglected, or, exactly, a Cartesian component of the electromagnetic wave (whose state of polarization will depend on position, causing the dislocations to be decorated with disclinations [10–12] that will not be discussed further). Another possibility, discussed later, is a three-dimensional (3D) wave linearly polarized azimuthally with respect to the propagation direction, where ψ satisfies an equation similar to (2).

At a dislocation, $\psi = 0$. In local coordinates ξ, η, with the sense of the dislocation line defined to be along the positive ζ axis, the wave has the form

$$\psi = a\xi + b\eta + \cdots. \tag{3}$$

Here $a = a_1 + ia_2$ and $b = b_1 + ib_2$ are complex constants, and the dislocation is assumed nondegenerate. The sign of the dislocation, defined as ± 1 if the phase increases or decreases by 2π during a positive circuit in the ξ–η plane, is

$$\begin{aligned} S &= \frac{1}{2\pi}\oint \mathrm{d}\chi(\mathbf{r}) \\ &= \operatorname{sgn} \det \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \\ &= \operatorname{sgn} \operatorname{Im} a^* b. \end{aligned} \tag{4}$$

This can be written in invariant form as

$$S = \operatorname{sgn} \operatorname{Im} \nabla\psi^* \times \nabla\psi \cdot \mathbf{n}, \tag{5}$$

where $\mathbf{n}$ is the unit vector along the dislocation.

S is a topological quantum number, whose sum for all the dislocations involved in an interaction must be conserved. Independent of S is a second topological quantum number T that must also be conserved. This is the Poincaré index, defined on a surface as the signed number of times that the phase lines rotate in the direction in which the circuit is traversed. T is $+1$ for all non-degenerate dislocations (and also for phase extrema which, however, do not occur in the waves studied here), and -1 for phase saddles. As Nye pointed out [5], S depends on the phase labels associated with the wave fronts, while T depends only on the pattern of wave fronts and not on their labels. Conservation of S and T puts strong restrictions on dislocation reactions. For example [13], when two dislocations with opposite signs annihilate, they must be accompanied by two saddles that also disappear in the event. At a phase saddle,

$$\nabla\chi = \nabla \operatorname{Im} \log \psi = 0, \qquad \text{that is } \operatorname{Im} \psi^* \nabla\psi = 0. \tag{6}$$

This simply states that phase saddles are stagnation points of the local current (Poynting vector).

For our Gaussian beams, we choose the propagation direction as the positive z axis, and the focal plane as $z = 0$. In this plane, the above relations for dislocations and phase saddles simplify, because the waves possess symmetry:

$$\psi(x, y, -z) = \psi^*(x, y, z). \tag{7}$$

Therefore ψ is real in the focal plane. Moreover, ψ is symmetric about the z axis. From these facts, and defining the convenient functions

$$f(x) \equiv \psi(x, 0, 0), \quad g(x) \equiv \operatorname{Im} \partial_z \psi(x, 0, 0), \tag{8}$$

it is easy to find the following conditions from equations (6) and (5)

$$\text{dislocations in the focal plane} : f(x) = 0, \tag{9a}$$

$$\text{saddles in the focal plane} : g(x) = 0, \tag{9b}$$

$$\text{sign of a dislocation in the focal plane} : S = -\operatorname{sgn} f'(x) g(x). \tag{9c}$$

3. Non-paraxial Gaussian beams in two and three dimensions

It is convenient to work with waves in two dimensions x, z, but corresponding formulae for waves in three dimensions will also be given, and will later be shown to possess exactly the same structure of dislocations and dislocation reactions. Of course in two dimensions the dislocations are points in the plane, whereas in three dimensions they are rings centred on the z axis. Moreover, the analogue of the focal plane is the line $x = 0$, but for convenience we shall still refer to this as the focal plane. Henceforth we set $k = 1$, which is equivalent to measuring all lengths in units of $1/k = \lambda/2\pi$.

We seek a wave that is a smooth function of position in the whole unobstructed space, propagating in from $z = -\infty$ and out to $z = +\infty$, with its waist in the focal plane $z = 0$. Any such two-dimensional wave (2D) can be written as the superposition of non-evanescent plane waves:

$$\psi_2(x, z) = \int_{-1}^{1} \mathrm{d}K\, a(K) \exp\left[\mathrm{i}(Kx + z\sqrt{1 - K^2})\right] \tag{10}$$

(evanescent waves, with $|K| > 1$ and the square roots chosen so that these waves decay as $|z| \to \infty$, are excluded because they would make $\nabla\psi$ discontinuous across the focal plane). In defining non-paraxial Gaussian beams, we note first that it is not possible to satisfy the condition that ψ^2 is exactly Gaussian in the focal plane, because this is incompatible with the integration limits $|K| = 1$. Therefore we impose the weaker requirement that the amplitude $a(K)$ of the component with transverse wavevector K must be chosen so that ψ^2 reduces to the familiar Gaussian beam (which has no dislocations) in the paraxial approximation, where $\sqrt{1 - K^2}$ is replaced by $1 - K^2/2$ and the integration is extended to the whole real K axis. This defines the class of nonparaxial Gaussian beams.

We choose what seems to be the simplest member of the class: a Gaussian bundle of plane waves, truncated at $|K| = 1$. Thus

$$\psi_2(x, z) + \frac{L}{\sqrt{\pi}} \int_{-1}^{1} \mathrm{d}K \exp\left(-\tfrac{1}{2}K^2L^2\right) \exp\left[\mathrm{i}(Kx + z\sqrt{1 - K^2})\right]. \tag{11}$$

If the integration range were the whole real line, this choice would give an exactly Gaussian profile in the focal plane, with L corresponding to the exponential intensity 1/e radius of the beam waist. I have examined several related choices, in which, for example the bundle of plane waves is Gaussian in angle ($\sin^{-1}K$) rather than the transverse wave-vector component K, or where the sharp cut-off of the Gaussian at $K = 1$ is softened by replacing K^2L^2 by a function such as $K^2L^2(1 - K^2)$. All lead to dislocation patterns very similar to the one that we shall explore.

The simplest 3-D solution of equation (2) analogous to equation (11) is

$$\psi_{3,\mathrm{i}}(r, z) = L^2 \int_0^1 \mathrm{d}K\, K \exp\left(-\tfrac{1}{2}K^2L^2\right) J_0(Kr) \exp\left(\mathrm{i}z\sqrt{1 - K^2}\right), \tag{12}$$

where $r = \sqrt{x^2 + y^2}$). Paraxially, and after some rescaling, this is identical to an integral studied by Karman *et al.* [2], representing the Gaussian beam focused by a lens of finite aperture. It should be emphasized, however, that here equations (11) and (12) are regarded as freely propagating Gaussian beams, not truncated by any aperture—although in a formal sense the effects of the truncation at the limit of

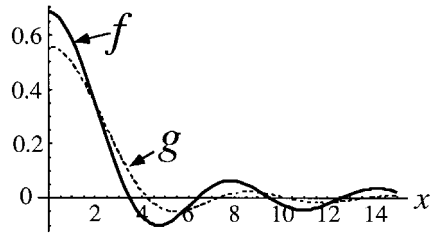

Figure 1. The functions f (——) and g (- - - -), defined in equations (14), for the 2D Gaussian beam ψ_2, for $L = 1$. Dislocations in the focal plane lie at the zeros of f, and phase saddles lie at the zeros of g.

evanescence $K = 1$ resemble those of an aperture. A related wave is the exact amplitude for light linearly polarized azimuthally; a little calculation shows that the azimuthal electric field is $r\psi_{3,\mathrm{ii}}(r, z)$, where

$$\psi_{3,\mathrm{ii}}(r, z) = \frac{L^4}{r}\int_0^1 \mathrm{d}K\, K^2 \exp\left(-\tfrac{1}{2}K^2L^2\right) J_1(Kr) \exp\left(\mathrm{i}z\sqrt{1-K^2}\right) \qquad (13)$$

(the Cartesian components of this field vanish smoothly on the axis $z = 0$, where there is a disclination). I shall not give detailed discussions of the waves $\psi_{3,\mathrm{i}}$ and $\psi_{3,\mathrm{ii}}$, or other 3D waves, because their dislocation structure is identical with that of equation (11).

4. Global singularity structure

For the 2D wave ψ_2 defined by equation (11), the focal plane functions f and g in equation (8) are

$$\begin{aligned} f_2(x) &= L\sqrt{\frac{2}{\pi}}\int_0^1 \mathrm{d}K \exp\left(-\tfrac{1}{2}K^2L^2\right)\cos(Kx), \\ g_2(x) &= L\sqrt{\frac{2}{\pi}}\int_0^1 \mathrm{d}K\, \sqrt{1-K^2} \exp\left(-\tfrac{1}{2}K^2L^2\right)\cos(Kx). \end{aligned} \qquad (14)$$

Figure 1 illustrates the behaviour of f and g. For small x they are approximately Gaussian, and for larger x they oscillate. The zeros of f are dislocations in the focal plane and are interlaced with the zeros of g, which are phase saddles. From equation (9c), the interlacing property implies that all the dislocations have the same sign ($S = +1$ for $x > 0$ and $S = -1$ for $x < 0$). As L increases, the region of x where f is Gaussian grows, and successive pairs of zeros disappear, as shown in figure 2.

As recognized by Karman *et al.* [2], the disappearances of zeros cannot represent a simple annihilation of dislocations, because this would violate the conservation of S. In fact, the dislocations move away from the focal plane in pairs. Figures 3 (a) and 3 (c) show the wave fronts and intensity contours before and after the first pair has departed. Figures 3 (b) and 3 (d) show the lines of local current (Poynting vector), parallel to $\nabla\psi$; this way of representing phase emphasizes the vortices at the dislocations and the stagnation points at the phase saddles.

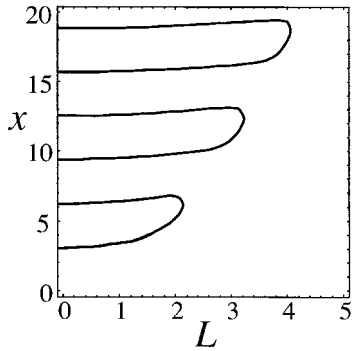

Figure 2. Dislocation tracks in the focal plane, as the beam waist L increases, for the 2D Gaussian beam ψ_2.

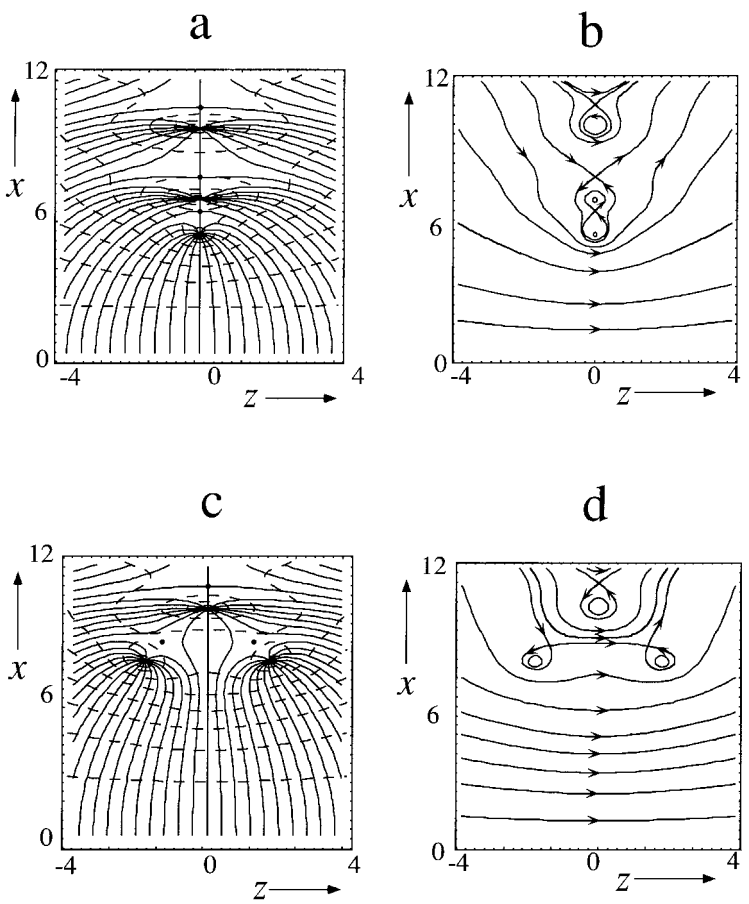

Figure 3. (*a*), (*c*) Wave fronts (contours of $\chi \bmod 2\pi$) of the 2D Gaussian beam ψ_2 (——) at intervals of $\Delta\chi = \pi/8$, and contours of modulus $|\psi|$ (- - -), with phase saddles marked with dots. (*b*), (*d*) lines of energy flow (proportional to $\nabla\chi$). In (*a*), (*b*), $L = 2$; in (*c*), (*d*), $L = 2.5$.

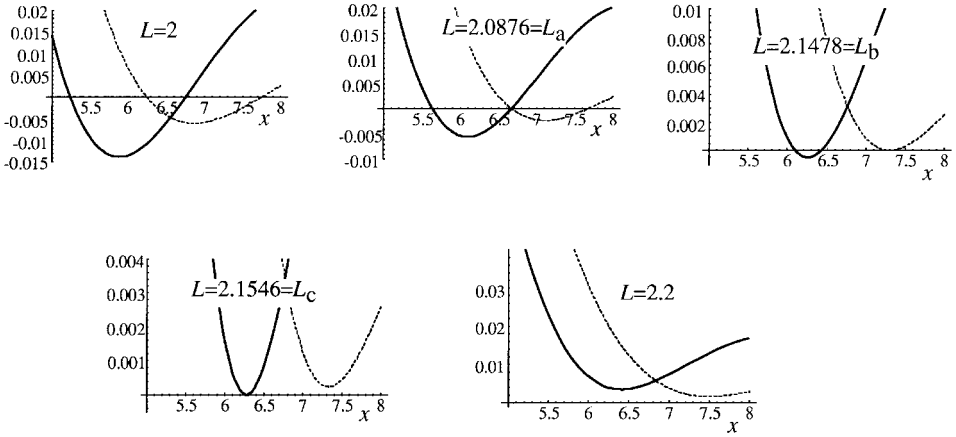

Figure 4. As figure 1, for a series of L values including the events L_a, L_b and L_c where the topology of the phase field changes as the first pair of dislocations leaves the focal plane.

Dislocations leave the focal plane in a process that can be analyzed into three events, illustrated in figure 4. We shall refer to the dislocation with smallest x, in the focal plane, as the 'lower' dislocation, and its neighbour, with the second smallest value of x, as the 'upper' dislocation. Similarly, the 'lower' saddle lies between the lower and upper dislocations, and the 'upper' saddle is above the upper dislocation. In event a, at $L = L_a$, the lower saddle passes through the upper dislocation, that is $f(x_a) = g(x_a) = 0$; from equation (9 *c*), this converts the sign of the dislocation to $S = -1$. Conservation of S requires the creation of two new dislocations, with $S = +1$, and these move away from the focal plane. Conservation of T requires that the newly born dislocations are accompanied by two saddles. In event b, at $L = L_b$, the lower saddle, that has catalysed the sign switch of the upper dislocation, collides with the upper saddle; the two zeros of g annihilate, that is $g(x_b) = g'(x_b) = 0$, but because of the conservation of T the saddles cannot disappear; they move away from the focal plane. In event c, at $L = L_c$, the two lowest zeros of f coalesce, that is $f(x_c) = f'(x_c) = 0$, in an annihilation event with the lower dislocation (with $S = +1$) and the upper dislocation (now with $S = -1$) disappearing along with the two saddles that were created in event a.

The first row of table 1 lists the values of L and x for these three events. Note how close the three L values are. The interval $L_c - L_a = 0.067$, over which all the changes in topology occur, corresponds to a change in the radius of the beam of only $(L_c - L_a)/2\pi = 0.011$ wavelengths. For the interval $L_c - L_b = 0.0068$ between the collision of the two saddles and the final annihilation, the change in radius is only 0.001 wavelengths.

The outcome of these events is that the upper and lower dislocations and saddles have, in effect, moved away from the focal plane, leaving the next pair of focal plane dislocations exposed to suffer the same fate for a series of larger values of L_a, L_b and L_c, listed in the fourth row of table 1. As L is further increased, the process repeats indefinitely.

Table 1. Beam widths L and positions x corresponding to topological transformations of 2D and 3D Gaussian beams, as the first ($m = 1$) and second ($m = 2$) pair of dislocations leave the focal plane.

	L_a	L_b	L_c	x_a	x_b	x_c
2D, $m = 1$, exact $\psi_2(11)$	2.0876	2.1478	2.1546	6.6776	7.3049	6.2832
2D, $m = 1$, approximation (15)	2.063	2.119	2.137	6.619	7.226	6.208
2D, $m = 1$, approximation (18)	2.081	2.142	2.140	6.768	7.461	6.283
2D, $m = 2$, exact ψ_2 (11)	3.1612	3.1985	3.2441	13.0684	13.6506	12.5668
2D, $m = 2$, approximation (15)	3.152	3.187	3.238	13.036	13.604	12.528
2D, $m = 2$, approximation (18)	3.154	3.190	3.237	13.144	13.745	12.566
3D, $m = 1$, $\psi_{3,i}$	2.2180	2.2700	2.2923	7.4395	8.0276	7.0156
3D, $m = 1$, $\psi_{3,ii}$	2.4507	2.4918	2.5367	8.8913	9.4206	8.4172

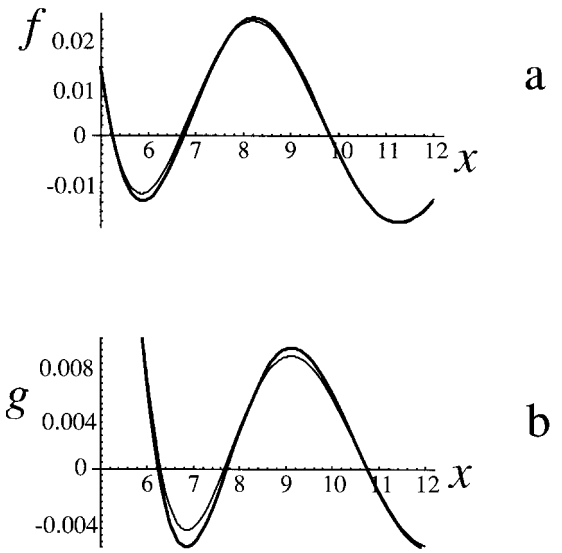

Figure 5. Comparison of (*a*) f and (*b*) g with the approximations (15), for $L = 2$.

To describe the events analytically, we evaluate the integrals (14) asymptotically, for large L. Contributions come from the saddle point at $x = \mathrm{i}/L^2$, representing a complex geometrical ray piercing the focal plane, and the end-point at $x = 1$, representing two marginally evanescent diffracted rays (one from each exponential in the cosine) travelling in the focal plane. The contributions can be separated by transforming the integration over the range $(0, 1)$ to two integrals, over the ranges $(0, \infty)$ and $(1, \infty)$. Standard techniques [14, 15] lead to

$$\begin{aligned} f_2(x) &\approx \exp\left(-\frac{x^2}{2L^2}\right) - \sqrt{\frac{2}{\pi}}\frac{L\exp(-L^2/2)}{\sqrt{L^4+x^2}}\cos\left[x+\tan^{-1}\left(\frac{x}{L^2}\right)\right], \\ \frac{g_2(x)}{\sqrt{1+x^2/L^4}} &\approx \exp\left(-\frac{x^2}{2L^2}\right) - \frac{L^3\exp(-L^2/2)}{(L^4+x^2)^{5/4}}\sin\left[x+\tfrac{3}{2}\tan^{-1}\left(\frac{x}{L^2}\right)\right]. \end{aligned} \tag{15}$$

In figure 5, these approximations are compared with the exact f_2 and g_2.

The first term in each of the formulae (15) is the Gaussian arising from the saddle point, and the second term is the oscillations from the end points. Comparison of the two exponentials shows that the oscillations dominate when $x > L^2$; this is the region where the dislocations and phase saddles lie. Asymptotically, the positions of these singularities are

$$\left.\begin{array}{c}\text{dislocations: } x \approx n\pi \\ \text{phase saddles: } x \approx (n+\frac{1}{4})\pi\end{array}\right\} \quad (x \gg L^2). \tag{16}$$

As the second and fifth rows of table 1 show, the approximations (15) reproduce, to an accuracy of a few per cent, the critical values L_a, L_b and L_c described above. A further simplification is possible, since the singularities involved in the corresponding events lie at positions x close to L, so in order to calculate the critical L and x values it is legitimate to write

$$x = L^2 + \xi, \quad \xi \ll L^2. \tag{17}$$

The approximations (15) can now be simplified:

$$\begin{aligned} f_2(x) &\approx \exp\left(-\frac{L^2}{2}\right)\left(\exp(-\xi) - \frac{1}{L\sqrt{\pi}}\cos(L^2+\xi+\tfrac{1}{4}\pi)\right), \\ g_2(x) &\approx \sqrt{2}\exp\left(-\frac{L^2}{2}\right)\left(\exp(-\xi) - \frac{1}{L^2 2^{5/4}}\sin(L^2+\xi+\tfrac{3}{8}\pi)\right). \end{aligned} \tag{18}$$

Algebra leads to

$$\begin{aligned} x_a &\approx (2m-\tfrac{1}{4})\pi + \tan^{-1}\left(\frac{2^{5/4}\sqrt{2m}}{\sin(\frac{3}{8}\pi)} - \cot(\tfrac{3}{8}\pi)\right), \\ L_a &\approx \sqrt{x_a} + \frac{1}{2\sqrt{x_a}}\log\left(\frac{\cos(x_a+\frac{1}{4}\pi)}{\pi\sqrt{2m}}\right), \\ x_b &\approx (2m+\tfrac{3}{8})\pi, \\ L_b &\approx \sqrt{(2m+\tfrac{3}{8})\pi} - \frac{1}{2\sqrt{(2m+\tfrac{3}{8})\pi}}\log[2^{7/4}(2m+\tfrac{3}{8})\pi], \\ x_c &\approx 2m\pi, \\ L_c &\approx \sqrt{2m\pi} - \frac{1}{2\sqrt{2\pi m}}\log(2\pi\sqrt{m}) \quad (m = 1, 2, \ldots). \end{aligned} \tag{19}$$

The values calculated from these formulae are shown in the third and sixth rows of table 1. For $m = 1$, the magnitudes are approximately correct but the events b and c occur in the wrong order. For $m = 2$, all the values are very accurate.

A simple summary of the results of this analytical theory is as follows. Dislocations arise from interference between marginally evanescent waves travelling towards and away from the beam axis, and the dislocation reactions result from the interference of these waves with the complex Gaussian ray. The set of events corresponding to the mth pair of dislocations leaving the focal plane occurs near $\xi = 2m\pi$, for a beam radius near $L = \sqrt{2m\pi}$, over a range given (for large m) by $\Delta L \approx [\log(m)]/(2\sqrt{2\pi m})$. The range is very small, as expected for a super-

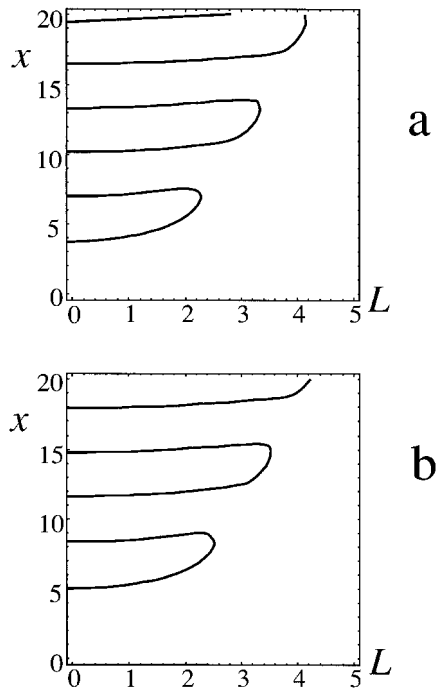

Figure 6. As figure 2, for the 3D Gaussian beams (*a*) $\psi_{3,\mathrm{i}}$ and (*b*) $\psi_{3,\mathrm{ii}}$.

oscillatory phenomenon [6, 16]. Consistent with this, the amplitude of the wave in the vicinity of these events is small too: $\exp(-L^2/2) \approx \exp(-m\pi)$.

For the 3D waves (12) and (13), the behaviour of the dislocations is very similar to that for the 2D wave (11) that we have been dicussing in detail. This is illustrated by figure 6 (to be compared with figure 2), which shows the tracks of dislocations in the focal plane, and the last two rows of table 1, which lists the critical values of the first set of events.

5. Local singularity structure

To obtain a local model reproducing the transition illustrated in figure 3, incorporating the three events described in the previous section and illustrated in figure 4, we employ Nye's [5] classification of exact solutions of equation (2). This is in terms of modulations $F(x, z)$ of a wave travelling in the z direction, that is (with $k = 1$)

$$\psi_{\mathrm{model}}(x, z) = F(x, z)\exp(\mathrm{i}z). \tag{20}$$

In the present application, x is a local coordinate in the focal plane. The functions F are polynomials, listed according to their highest power of x. The first few are

$$\begin{aligned} &F_0 = 1, \quad F_1 = x, \quad F_2 = x^2 + \mathrm{i}z, \\ &F_3 = \tfrac{1}{3}x^3 + \mathrm{i}zx, \quad F_4 = \tfrac{1}{6}x^4 - \tfrac{1}{2}z^2 + \mathrm{i}z(x^2 - \tfrac{1}{2}). \end{aligned} \tag{21}$$

Any superposition of these functions, with arbitrary complex constants, is a solution of equation (2); to guarantee the symmetry (7), the constants must be real.

The functions f and g defined by equation (8) are

$$f(x) = F(x, 0), \quad g(x) = F(x, 0) + \operatorname{Im} \partial_z F(x, 0). \tag{22}$$

To reproduce the sequence in figure 4, f and g must each have a single minimum, and the minima must be separated by an intersection where the curves have opposite slopes. This cannot be achieved with any superposition with the modulation F_2 or F_3 in equation (21) as its highest member; it is necessary to include at least F_4. It is not necessary to include F_3, since this can be eliminated by shifting the origin of ξ. Nye [4] examined the case where ψ is even in x, but this is unnecessary; indeed, the sequence that he studied in detail does not include the event b, where two phase saddles leave the focal plane. A simpler choice is

$$F = F_4 + BF_1 + AF_0 = \tfrac{1}{6}x^4 - \tfrac{1}{2}z^2 + Bx + A + \mathrm{i}z(x^2 - \tfrac{1}{2}), \tag{23}$$

where A and B are positive. Thus

$$f(x) = \tfrac{1}{6}x^4 + Bx + A, \quad g(x) = \tfrac{1}{6}x^4 + Bx + A + x^2 - \tfrac{1}{2}. \tag{24}$$

We wish f and g each to have one minimum rather than two; obviously these must have x negative. Each of the events a, b and c corresponds to a locus in A, B space. Consider first event a, where f and g are both zero. For f and g to coincide between their minima, $\xi = -1/\sqrt{2}$. For this coincidence to be a zero of f and g,

$$A = \frac{1}{\sqrt{2}}B - \frac{1}{24}. \tag{25}$$

Now consider event c, where $f = 0$ at its minimum. The minimum lies at $x = -(\frac{3}{2}B)^{1/3}$. It is also a zero if

$$A = \tfrac{1}{2}(\tfrac{3}{2}B)^{4/3}. \tag{26}$$

At event b, $g = 0$ at its minimum; the locus can be found as the solution of a cubic equation and is most easily determined numerically. Choosing $B = 1$, we obtain, for the A values of the three events,

$$\begin{aligned} A_\mathrm{a} &= \frac{1}{\sqrt{2}} - \frac{1}{24} = 0.665\,44, \\ A_\mathrm{b} &= 0.740\,99, \; A_\mathrm{c} = \tfrac{1}{2}(\tfrac{3}{2})^{4/3} = 0.858\,54. \end{aligned} \tag{27}$$

Figure 7 shows the wave fronts of ψ_model for a series of values of A. The three events can be seen at A_a (figure 7 (*b*)), A_b (figure 7 (*d*)) and A_c (figure 7 (*f*)). They occur exactly as described in the previous section and as observed by Karman *et al.* [1, 2]. I have performed similar calculations with the integrals for ψ_2, $\psi_{3,\mathrm{i}}$ and $\psi_{3,\mathrm{ii}}$, and confirmed that the same local transformations of wave fronts occur.

Abandoning the requirement that event b occurs between events a and c, we can choose

$$A = \tfrac{1}{8}, B = \frac{1}{3\sqrt{2}} \tag{28}$$

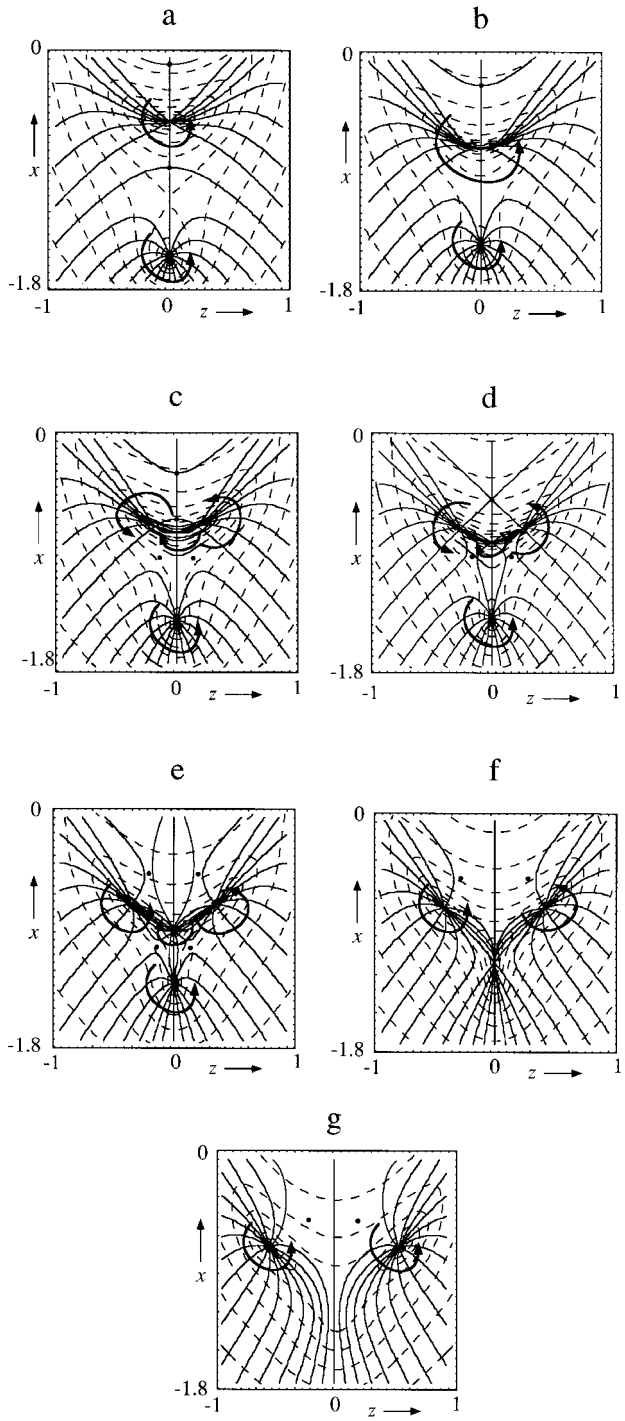

Figure 7. Wave fronts (——) and contours of modulus (- - -), with phase saddles marked by dots and the signs of dislocations (direction of increasing phase) marked with arrows, for the model wave ψ_{model} (equations (20) and (23) with $B = 1$), for (*a*) $A = 0.5$, (*b*) $A = A_a = 0.665\,44$, (*c*) $A = 0.7$, (*d*) $A = A_b = 0.740\,99$, (*e*) $A = 0.8$, (*f*) $A = A_c = 0.858\,54$ and (*g*) $A = 1$.

and make events a and c coincide in a 'supersingularity'. This is a dislocation of strength 2, also mentioned by Nye [5] (and, incidentally, disproving our old conjecture [4] that multiple edge dislocations do not exist).

6. Concluding remarks

It might be difficult to generate the non-paraxial, freely propagating beams experimentally. I cannot think of a simple way. Nevertheless, it seems worthwhile to present these calculations, for several reasons.

First, to demonstrate that the same dislocation topologies as are generated with an aperture [1–3] can also, in principle, be produced (non-paraxially) without an aperture.

Second, to emphasize that an essential feature in the production of phase dislocations is interference. For the non-paraxial Gaussian beams that I have described, the interference is between a complex ray and marginally evanescent waves. This interpretation holds in two and three dimensions, apparently independently of exactly what vector wave the complex scalar amplitude represents, and explains why the same pattern of dislocations occurs for ψ_2, $\psi_{3,\mathrm{i}}$ and $\psi_{3,\mathrm{ii}}$. The same asymptotics that led to equation (15) can easily be applied to integrals representing the truncated Gaussian (paraxial and non-paraxial) beams studied by Karman *et al.* [1, 2] and Beijersbergen [3]. The results are very similar, but the interpretation is slightly different: the marginally evanescent waves are replaced by diffracted rays from the edge of the truncating aperture.

Third, to develop Nye's [5] local classification in slightly greater detail, to bring out the role of the phase saddles in the sequence of events leading to the departure of pairs of dislocations from the focal plane.

Acknowledgement

I thank Professor J. F. Nye for helpful discussions and suggestions. Some of this work was carried out at the Isaac Newton Institute, for whose support I am grateful.

References

[1] Karman, G. P., Beijersbergen, M. W., van Duijl, A., and Woerdman, J. P., 1997, Creation and annihilation of phase singularities in a focal field, *Optics Lett.*, **22,** 1503–1505.

[2] Karman, G. P., Beijersbergen, M. W., van Duijl, A., Bouwmeester, D., and Woerdman, J. P., 1998, Airy pattern reorganization and sub-wavelength structure in a focus, *J. opt. Soc. Am.* A, **15,** 884–899.

[3] Beijersbergen, M., 1996, Phase singularities in optical beams, (Thesis, *Huygens Laboratory*, University of Leiden, p. 98.

[4] Nye, J. F., and Berry, M. V., 1974, Dislocations in wave trains, *Proc. R. Soc.* A, **336**, 165–190.

[5] Nye, J. F., 1998, Unfolding of higher-order wave dislocations, *J. Opt. Soc. Am.* A, **15,** (to be published).

[6] Berry, M. V., 1994, Evanescent and real waves in quantum billiards, and Gaussian Beams, *J. Phys.* A, **27,** L391–L398.

[7] Bazhenov, V. Y., Soskin, M. S., and Vasnetsov, M. V., 1990, Laser beams with screw wavefront dislocations. *JETP Lett.*, **52**, 429–431.

[8] Bazhenov, V. Y., Soskin, M. S., and Vasnetsov, M. V., 1992, Screw dislocations of wavefront, *J. mod. Optics.,* **39**, 985–990.
[9] Basistiy, I. V., Bazhenov, V. Y., Soskin, M. S., and Vasnetsov, M. V., 1993, Optics of light beams with screw dislocations, *Optics Commun.*, **103**, 422–428.
[10] Nye, J. F., 1983, Lines of circular polarization in electromagnetic wave fields, *Proc. R. Soc. Lond.* A, **389**, 279–290.
[11] Nye, J. F., and Hajnal, J. V., 1987, The wave structure of monochromatic electromagnetic radiation, *Proc. R. Soc.* A, **409**, 21–36.
[12] Nye, J. F., 1991, Phase gradient and crystal-like geometry in electromagnetic and elastic wavefields, *Sir Charles Frank, O.B.E.; an eightieth birthday tribute*, edited by R. G. Chambers, J. E. Enderby, A. Keller, A. R. Lang, and J. W. Steeds (Bristol: Adam Hilger), pp. 220–231.
[13] Nye, J. F., Hajnal, J. V., and Hannay, J. H., 1988, Phase saddles and dislocations in two-dimensional waves such as the tides, *Proc. R. Soc. Lond.* A, **417**, 7–20.
[14] Copson, E. T., 1965, *Asymptotic Expansions* (Cambridge University Press).
[15] Olver, F. W. J., 1974, *Asymptotics and Special Functions* (London: Academic Press).
[16] Berry, M. V., 1994, Faster than Fourier, *Quantum Coherence and Reality; in Celebration of the 60th Birthday of Yakir Aharonov*, edited by J. S. Anandan, and J. L. Safko (Singapore: World Scientific), pp. 55–65.