Proc. R. Soc. Lond. A. **336**, 165–190 (1974)
Printed in Great Britain

Dislocations in wave trains

BY J. F. NYE AND M. V. BERRY
H. H. Wills Physics Laboratory, University of Bristol

(*Communicated by F. C. Frank, F.R.S. – Received* 17 *January* 1973)

When an ultrasonic pulse, containing, say, ten quasi-sinusoidal oscillations, is reflected in air from a rough surface, it is observed experimentally that the scattered wave train contains dislocations, which are closely analogous to those found in imperfect crystals. We show theoretically that such dislocations are to be expected whenever limited trains of waves, ultimately derived from the same oscillator, travel in different directions and interfere – for example in a scattering problem. Dispersion is not involved. Equations are given showing the detailed structure of edge, screw and mixed edge–screw dislocations, and also of parallel sets of such dislocations. Edge dislocations can glide relative to the wave train at any velocity; they can also climb, and screw dislocations can glide. Wavefront dislocations may be curved, and they may intersect; they may collide and rebound; they may annihilate each other or be created as loops or pairs. With dislocations in wave trains, unlike crystal dislocations, there is no breakdown of linearity near the centre. Mathematically they are lines along which the phase is indeterminate; this implies that the wave amplitude is zero.

1. OBSERVATION OF DISLOCATIONS

In this paper we introduce a new concept into wave theory: using elementary arguments, we show that wavefronts – that is, surfaces of constant phase – can contain dislocation lines, closely analogous to those found in crystals. The work originated in attempts to understand radio echoes from the bottom of the Antarctic ice sheet; the spatial fine structure, or 'fading pattern', of the echo may be used for precise determination of position (Nye, Kyte & Threlfall 1972; Walford 1972; Nye, Berry & Walford 1972). A laboratory analogue experiment was carried out, using ultrasound instead of radio waves. The relatively low frequency of ultrasound enabled the detailed phase structure of the echo to be studied, and wavefront dislocations were observed. We suspect that dislocations may often have been observed in phase sensitive experiments without their significance being appreciated (see, for example, figure 5 of a paper by Findlay (1951) showing dislocations in a radio wave field reflected from the ionosphere, and the remarks in §8 of the present paper about amphidromic points). It is possible that dislocations may find application in remote sensing as 'markers' in a wave field, because they are definite features recognizable even in the presence of noise. In this paper, however, our purpose is simply to demonstrate that wavefront dislocations exist, to deduce their detailed structure and to examine some of their properties.

We begin by describing the original observation. Ultrasonic pulses from a small source were incident, in air, on a rough surface, and the scattered pulses were received by a small moveable microphone and displayed on an oscilloscope. Each pulse from the source consisted of about 10 sinusoidal waves (frequency 100 kHz)

within a smoothly varying (approximately Gaussian) envelope (figure 1*a*). The echo from the rough surface was of longer duration (figure 1*b*), consisting of 50 or more approximately sinusoidal waves of fluctuating amplitude. By moving the microphone along a line it was possible to find points in space where the envelope of the echo had zero strength for some particular time delay. Figure 2*c* shows a

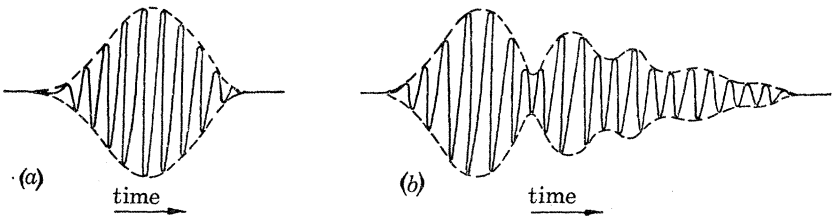

FIGURE 1. (*a*) Quasi-monochromatic pulse from a source.
(*b*) The echo from a rough surface.

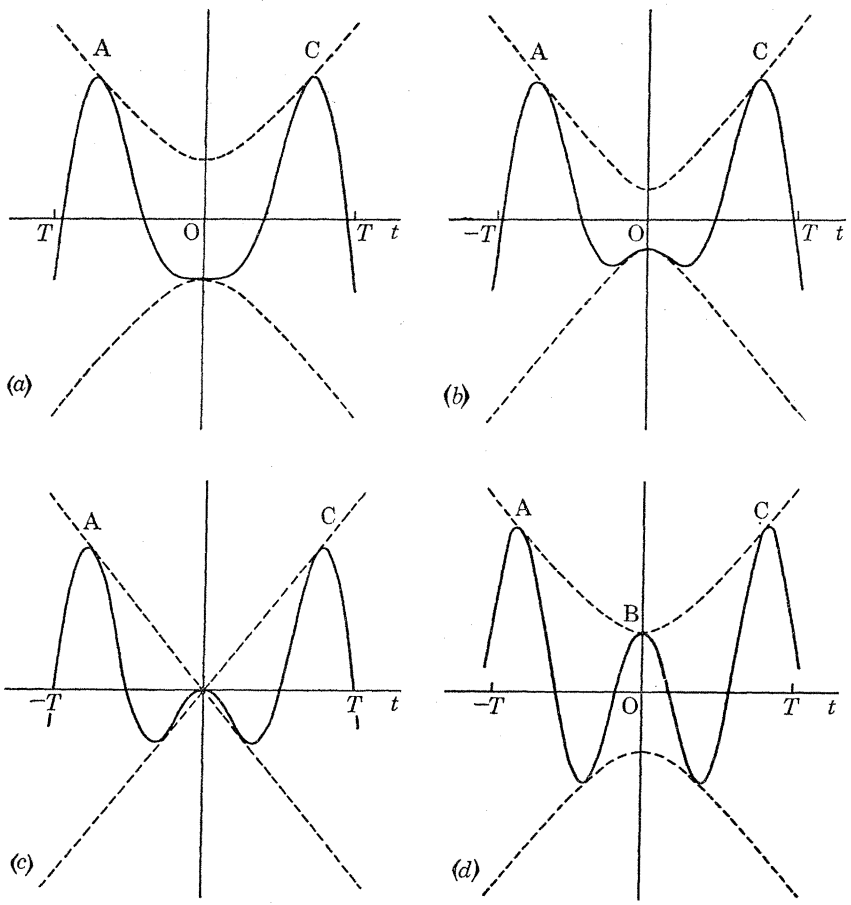

FIGURE 2. An edge dislocation – symmetrical case. Signal versus time t for four different positions of a receiver: (*a*) $x = -2\beta_e^*/k$, (*b*) $x = -\beta_e^*/k$, (*c*) $x = 0$, (*d*) $x = 2\beta_e^*/k$. $T = 2\pi/\omega$ = period.

symmetrical example of the time variation of the signal at such a point; the broken line is the envelope. By moving the microphone first to one side (figures 2*a*, *b*) and then the other (figure 2*d*) one could follow the crests A and C continuously through the transition and see that a new crest B was created between them. Figures 3*a*, *b*, *c*, *d* show an antisymmetrical example of the same thing; again a new crest B appears between crests A and C. In general, the phase of the carrier wave relative to the envelope is such that one observes an unsymmetrical curve which is a linear combination of the symmetrical and antisymmetrical cases.

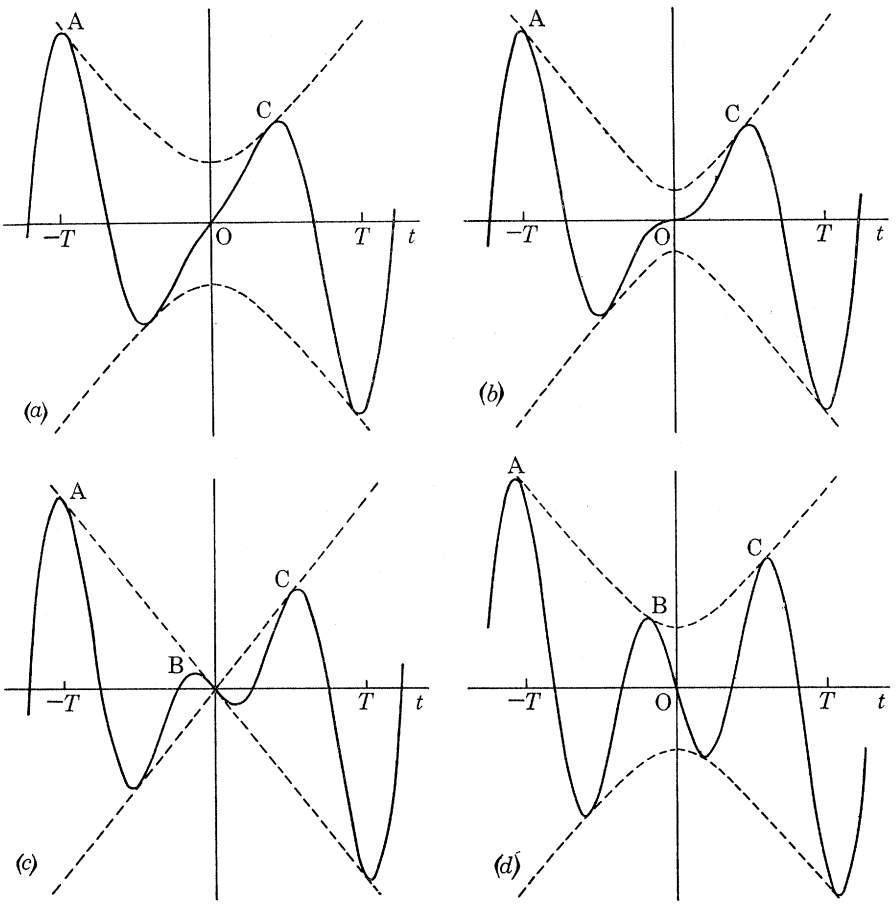

FIGURE 3. As figure 2 – antisymmetrical case. Figures 2 and 3 are calculated by taking the real and imaginary parts respectively of equation (13).

A possible spatial variation at a given time of the echo corresponding to figures 2*a* to *d* is shown in figure 4. As later examples will show, the wavefront dislocation may move backwards, forwards or sideways independently of the motion of the wavefronts; but for simplicity in this introductory section we consider the case in which the dislocation travels with the wavefronts, in a direction normal to their mean

plane. As the pattern of crests sweeps upwards a receiver at P notes the passage of crests A and C, and as the receiver is moved continuously to R a new crest B is found to appear as the receiver passes point Q. The spatial pattern has a crest wavefront that ends at N. (The precise location of N depends on what one defines as a crest; we

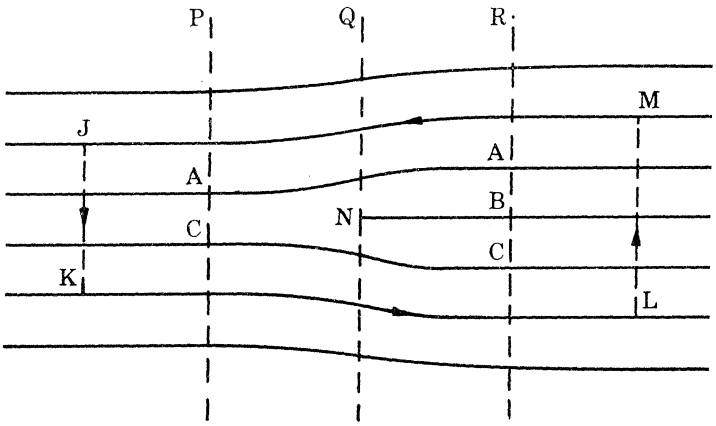

FIGURE 4. Spatial pattern of wave-crests at a fixed time. There is an edge dislocation at N.

shall deal with this presently.) In three dimensions the wavefront ends in a line. We shall call this linear structure an edge dislocation by analogy with the edge dislocations found in crystals (Read 1953; Nabarro 1967). In practice, since the amplitude falls to zero at N, the details of the transition may be submerged in the noise, but, by following a 'Burgers circuit' such as JKLM which is everywhere in places of strong signal and counting the number of wavefronts crossed, one can be quite sure that a dislocation has been enclosed. The analogy with crystal dislocations goes further: we shall see that it is possible to have both screw and mixed edge-screw dislocations, that wavefront dislocations are capable of movement by glide and climb, that they can intersect, that they can collide and rebound, or annihilate each other or be created as loops or pairs.

In general, pulses rather than monochromatic waves are essential for dislocations. In a monochromatic wave the time variation at a point must be strictly sinusoidal. This is clearly impossible at Q in figure 4, for the passage of N past Q is a unique event. At the same time there must be some periodicity if we are to identify a dislocation. Thus the dislocations are structures that disappear at both the monochromatic and the white noise limit; they need both the localization property of a pulse and the oscillation property of a continuous wave. Dispersion is not involved.

However, in certain degenerate cases dislocations can be produced by continuous waves. One is the non-localized interference fringe FF shown in figure 5, as produced, for example, by a Young's two-slit experiment. Monochromatic waves are moving upwards and their amplitude is zero on FF; there is a phase change of π across FF. This could be described as a row of edge dislocations of alternating sign or as a single

infinitely extended dislocation. Because all the crests behave identically the time variation at any point is still strictly sinusoidal. Other degenerate cases are the stationary pure screw dislocations and the localized interference fringe which we shall describe later.

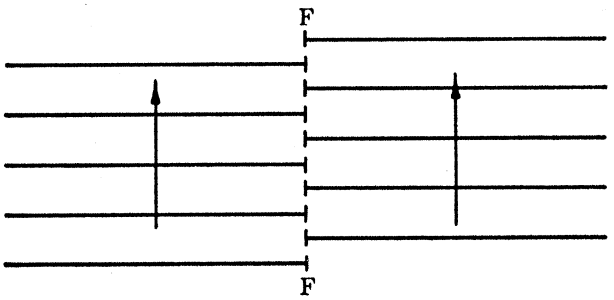

FIGURE 5. A non-localized interference fringe FF. The waves travel upwards.

We have introduced wavefront dislocations as structures observed when a pulse is diffracted from a rough surface and we shall deal more fully with this method of producing them in a separate paper. Here we show that they are a general feature to be expected whenever there is interference between pulses derived from a common oscillator; diffraction by a rough surface is but one way of making them.

2. THE FUNDAMENTAL MATHEMATICAL PROPERTY OF A DISLOCATION

What distinguishes a dislocation from other features of a wave field? To answer this, we must first define precisely the *amplitude* and *phase* of a wave. Let there be an oscillator with fixed angular frequency ω. After an arbitrary phase change the signal from it is amplitude-modulated and made to drive a source of scalar waves, the wave amplitude being proportional to the driving signal. We call such a source *quasi-monochromatic*. At the same time the original signal from the oscillator is changed in phase and amplitude-modulated in other fixed ways and made to drive other sources in a similar fashion. All the wave trains so produced travel together in an undispersive medium and their combined effect is observed at a certain point P. If the original oscillation is proportional to $\cos(-\omega t)$ we denote the resulting signal at P by $\psi_c(t)$. The phase of the original carrier wave is now changed by $\frac{1}{2}\pi$, while everything else is kept fixed. Thus, let the original oscillation now be proportional to $\sin(-\omega t)$, while the envelopes representing the various amplitude modulations and all the phase changes are held fixed, and denote the resulting signal at P by $\psi_s(t)$. We can say that, if the original oscillation were proportional to $e^{-i\omega t}$, the signal at P would be the complex wave function

$$\psi(t) = \psi_c(t) + i\psi_s(t) = \rho(t)\, e^{i\chi(t)}, \quad \text{say}, \tag{1}$$

(χ being real and ρ being real and positive) on the understanding that either the real or the imaginary parts represent the physical quantities. Then

$$\rho^2(t) = \psi_c^2(t) + \psi_s^2(t) \quad (\rho(t) \geqslant 0), \tag{2}$$

and

$$\tan\chi(t) = \psi_s(t)/\psi_c(t). \tag{3}$$

It is natural to define $\rho(t)$ and $\chi(t)$ as the amplitude and phase of the wave; they are both time-varying quantities deducible from the observed functions $\psi_c(t)$ and $\psi_s(t)$. It is important to notice that $\rho(t)$, for example, is not strictly deducible by observing, say, $\psi_c(t)$ alone; it is essential to change the phase of the carrier wave to obtain complete information about the amplitude and phase. With this definition $\rho(t)$ is the envelope of the observed oscillation as the phase of the original oscillation is varied. This is the strict meaning of the broken line in figures 2 and 3.

We now examine the behaviour of $\rho(t)$ and $\chi(t)$ in the neighbourhood of a dislocation. The argument applies to any dislocation, moving in any way; in particular it applies to figure 4, which represents a pure edge dislocation moving with the wave. The changing complex wave function at a fixed position P in space can be represented by a moving point on an Argand diagram whose axes are ψ_c and ψ_s. Because of the quasi-monochromaticity of the source, we expect the point to encircle the origin quasi-periodically, the average time for a circuit being $2\pi/\omega$. Figure 6*a* shows (schematically) the single loop traversed by the point on the Argand diagram between two successive crests A and C of the component ψ_c; this corresponds to the signal sketched in figure 2*a*. At an observation position R, which lies on the other side of the path of dislocation, an extra crest B has appeared between A and C, so that (figure 6*d*) *two* loops now lie between A and C on the Argand diagram, corresponding to the signal of figure 2*d*. (If time is plotted perpendicular to the diagrams of figure 6 the changing wavefunction at a fixed point in space is represented by a helix with varying cross-section and pitch. Viewed in this way the crests A and C remain identifiable, as distinct turns of the helix, throughout the transition.)

As the observation position moves from P to R, the curve representing $\psi(t)$ must change continuously between that of figure 6*a* and that of figure 6*d*, which has different topology. This can only happen if, for some intermediate observation position Q, *the curve passes through the origin*, where ψ_c and ψ_s, and hence the amplitude ρ, are zero. One possible form for the curve at Q is shown in figure 6*c*, while the curve for a position between P and Q is shown in figure 6*b*. We note, in passing, that in the configurations (not drawn) intermediate between figures 6*b* and *c* the phase χ would be retrograde for part of the curve. It is easily verified that the behaviour of $\psi_c(t)$ for the four curves in figure 6 reproduces the symmetrical signal versus time curves of figure 2, while the behaviour of $\psi_s(t)$ reproduces the antisymmetrical curves of figure 3.

Precisely where the dislocation lies is a matter of definition, and we choose the position where the amplitude ρ vanishes. This is the most fundamental definition from a theoretical point of view (see §6); it is preferable to choosing the cusp in figure 6*b* where a new loop is just about to appear; if we choose the position where

the pair of new zeros first appears, we should have a dislocation whose position depends on which 'projection' of the complex wave function represents the physical signal (that is, on the phase of the carrier wave). Since ρ is zero at a dislocation, it follows from the properties of polar coordinates in the Argand diagram that the phase χ is indeterminate. However, it is possible to have zeros of ρ, with indeterminate χ, that do not represent dislocations – the curve in the Argand diagram may

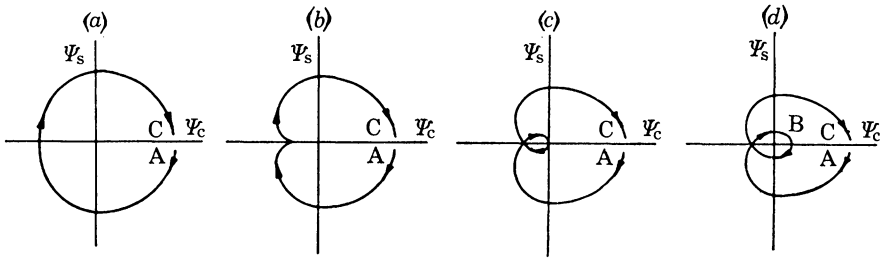

FIGURE 6. Curve in Argand diagram representing the complex wave function as a function of time for four positions near a dislocation. (*c*) is the configuration at a point through which the dislocation passes.

move up to the origin, and then move away again, without having swept through it; an example is equation (13) of §4 with $(kz-\omega t)$ replaced by $(kz-\omega t)^2$. The essential property of a dislocation is that χ changes by a multiple of 2π on a closed circuit around it. Thus, although the condition $\rho = 0$ is a useful indication of where to look for a dislocation, it is necessary to check that the structure located in this way does indeed have the topology that characterizes a dislocation. We shall examine this topology in a little more detail in §6.

The dislocations in a given wave field are situated at positions where $\psi_c(t) = 0$ and $\psi_s(t) = 0$. At a given instant t these two equations define surfaces in three-dimensional physical space, which intersect in a family of lines (except in some degenerate cases). As time proceeds, the surfaces move and hence the lines – the dislocations – also move. As they move the dislocation lines sweep out surfaces in space. Thus a receiver that displays the signal at a function of time need only explore a line in order to find a dislocation, which will be at the point where the line intersects the surface (with Dr M. E. R. Walford we have mapped some of these surfaces for the special case of the ultrasonic wave field described in §1).

3. THEORETICAL CONSTRUCTION OF DISLOCATIONS: QUALITATIVE TREATMENT

We now describe a simple way of constructing dislocations theoretically. Consider identical pulses of scalar waves emitted simultaneously from two point sources S_1 and S_2, and let the signal be observed at a point P. Figure 7 *a* shows one of the pulses, as observed at P, a monochromatic wave modulated in amplitude by an envelope, and the convention in figure 7 *b* will be used to denote the pulse envelope

and the positions of the crests and troughs within it. We use the words crests and troughs rather loosely here, and our argument is only qualitative; a strict treatment will follow later. The signal observed at P will depend on the path difference $S_1P - S_2P$. In figure 7 *c* the position of P is such that the path difference is two wavelengths, and so the carrier waves reinforce one another to give the resultant shown

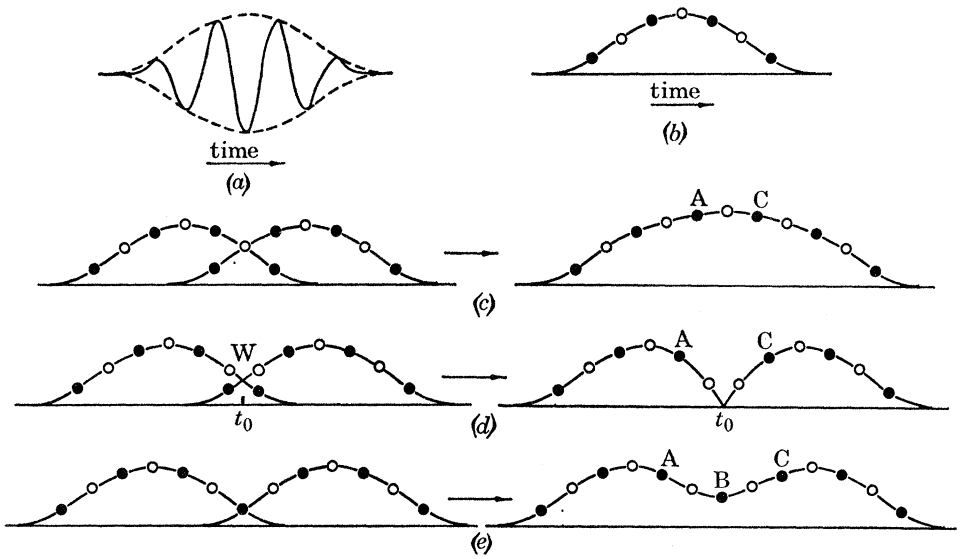

FIGURE 7. The construction of edge dislocations by the interference of pulses. Signal versus time. (*a*) the pulse form; (*b*) schematic representation of pulse: ●, crests, ○, troughs; (*c*) path difference is two wavelengths; (*d*) path difference is $2\frac{1}{2}$ wavelengths, giving destructive interference at time t; (*e*) path difference is three wavelengths, and a new crest B has appeared.

on the right (we ignore the fact that the two pulses will be of slightly different strengths). In figure 7 *d* P has moved so that the path difference is $2\frac{1}{2}$ wavelengths. There is destructive interference, but it is only complete at one time t_0. In figure 7 *e* the path difference is three wavelengths and the carrier waves reinforce again. There are six crests in figure 7 *c*, a new one is on the point of appearing in figure 7 *d*, and there are seven in figure 7 *e*. Clearly, when the path difference is zero there will be four crests, and when the path difference is large there will be two separate pulses with eight crests in all; crests appear as the path difference increases. Each time one appears there is a dislocation. The essential point is that, when the path difference is such that the carrier waves interfere destructively, the interference will only be complete at one time, namely the time when the two envelopes intersect; at this time the amplitude ρ of the combined wave is zero. This principle still holds when the two pulses are of unequal strength and duration, except that then the envelopes may intersect, for a given path difference, at more than one time. The spatial pattern of crests at fixed time shown in figure 4 corresponds with figure 7.

To proceed analytically we retain only those features of the theoretical model just

described which are essential for the production of an edge dislocation. The sources are placed at infinity so that there are now two pulses of plane waves crossing at a small angle, and instead of an approximately Gaussian envelope we choose a linear variation of amplitude through the pulse, one pulse rising as the other falls. This amounts to approximating the envelopes near the intersection point W in figure 7 *d* by their tangents. In this way we may calculate in detail the structure in time and space near a dislocation, but the results will not have significance outside the range where the envelopes may be approximated by their tangents.

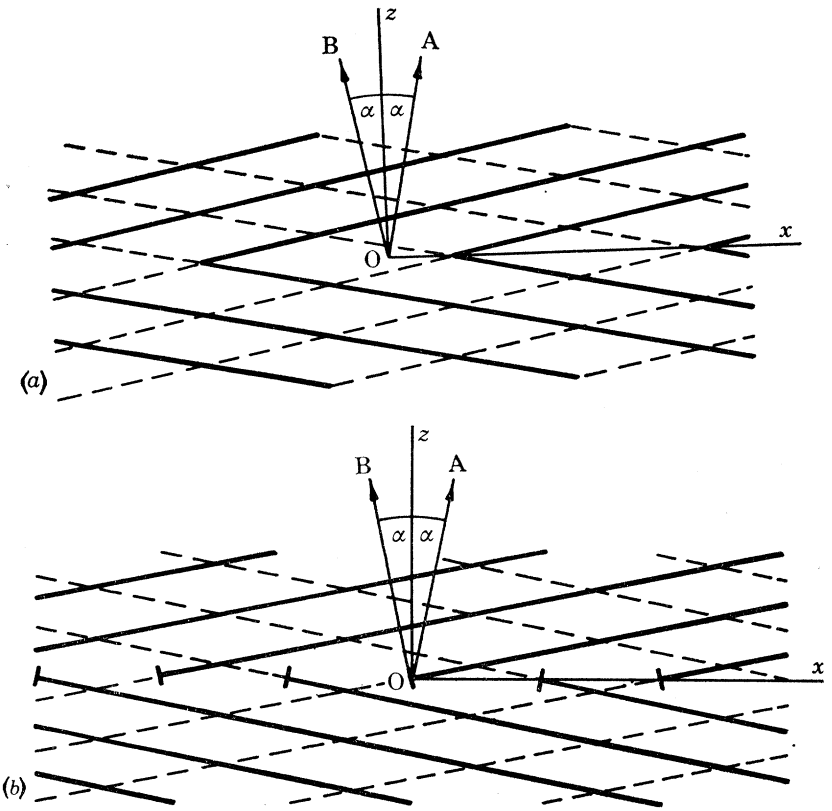

FIGURE 8. The spatial pattern of crests from two wave-trains A and B travelling at an angle to one another. A predominates in the lower half-space and B predominates in the upper half-space. (*a*) symmetrical case; (*b*) alternating case.

The general nature of the result to be expected is seen in figure 8 *a*, which shows the spatial pattern of wave crests at fixed time. The wave train A that travels upwards and to the right has an amplitude that rises towards the rear of the train; the wave train B that travels upwards and to the left has an amplitude that rises towards the front of the train. The amplitudes are arranged to be equal on Ox, so that in the upper half-space train B predominates and in the lower half-space A predominates. Thus, as we move along Ox, new wave crests appear in pairs. In

crystal dislocation language Ox is a tilt boundary. Note that figure 8*a* has been drawn so that the origin, where the two waves have equal amplitude, is a zero (approximately midway between a crest and a trough) for both waves; this is the same relationship as in figure 7*d*. Figure 8*b*, on the other hand, shows the pattern of crests that results if the phases are arranged so that at the origin, where the waves have equal amplitude, a crest of B falls on a trough of A. Edge dislocations of alternate character appear along Ox. Figures 8*a* and *b* merely show which of the two wave trains is dominant at any point; the true pattern of resultant wave crests will be similar but with local readjustments. The exact analysis which follows essentially calculates these local readjustments.

4. Exact analysis for edge dislocations

We shall consider complex wave functions ψ satisfying the scalar wave equation

$$c^2\nabla^2\psi = \frac{\partial^2\psi}{\partial t^2}, \tag{4}$$

the wave velocity c being constant.

Let the wave trains A and B, as in figures 8*a* and *b*, be the following solutions of (4):

$$\left.\begin{aligned} \psi_{\mathrm{A}} &= a_0\{1-\beta_{\mathrm{e}}(k_1x+k_3z-\omega t)\}\exp\left[\mathrm{i}(k_1x+k_3z-\omega t-\tfrac{1}{2}\pi)\right] \\ \text{and}\quad \psi_{\mathrm{B}} &= a_0\{1+\beta_{\mathrm{e}}(-k_1x+k_3z-\omega t)\}\exp\left[\mathrm{i}(-k_1x+k_3z-\omega t+\tfrac{1}{2}\pi)\right], \end{aligned}\right\} \tag{5}$$

with

$$\omega/k = c, \quad k_1 = k\sin\alpha, \quad k_3 = k\cos\alpha, \tag{6}$$

where a_0 and β_{e} are constants, ω is the fixed angular frequency at the original oscillator, k is the wave number corresponding to ω, and α is the angle between the wave normals and the z-axis. Put

$$\xi = k_1x, \quad \zeta = k_3z-\omega t. \tag{7}$$

The resultant complex disturbance ψ is given by

$$\psi = \psi_{\mathrm{A}}+\psi_{\mathrm{B}} \equiv \rho\,\mathrm{e}^{\mathrm{i}\chi}. \tag{8}$$

That is

$$\rho\,\mathrm{e}^{\mathrm{i}\chi} = 2a_0\{(1-\beta_{\mathrm{e}}\xi)\sin\xi+\mathrm{i}\beta_{\mathrm{e}}\zeta\cos\xi\}\,\mathrm{e}^{\mathrm{i}\xi} \tag{9}$$

Hence

$$\rho^2 = 4a_0^2\{(1-\beta_{\mathrm{e}}\xi)^2\sin^2\xi+(\beta_{\mathrm{e}}\zeta)^2\cos^2\xi\}, \tag{10}$$

and

$$\chi = \arctan\left\{\frac{\beta_{\mathrm{e}}\zeta}{(1-\beta_{\mathrm{e}}\xi)\tan\xi}\right\}+\zeta+2n\pi, \tag{11}$$

the ambiguity of an odd multiple of π in the value of arctan being resolved by making sure that the sine of the angle has the same sign as $\beta_{\mathrm{e}}\zeta\cos\xi$. If the integer n is chosen so that χ is between $-\pi$ and π, χ becomes what we shall call the reduced phase χ_0; otherwise n is zero and χ is a continuous, but multivalued, function. It is sufficient for our purpose to keep β_{e} small and restrict attention to a limited region of (ξ,ζ) space around the origin, so that the envelope expressions in (5) do not go negative.

It is easily seen from (9) or (10) that the dislocations, which occur where ρ is zero, lie along the lines parallel to the y-axis defined by $\zeta = 0$, $k_1 x = m\pi$, where m is any integer or zero, and that they alternate in character. Figures 9 *a* and *b* show lines of constant reduced phase χ_0 and constant amplitude ρ at time $t = 0$ for $\tan\alpha = 0.1$ and $\beta_e = 0.1$. If one traverses a closed circuit anticlockwise, χ increases by 2π for every dislocation enclosed. The lines of constant χ_0 may be read in different ways depending on the phase of the carrier wave relative to the pulse envelope in the

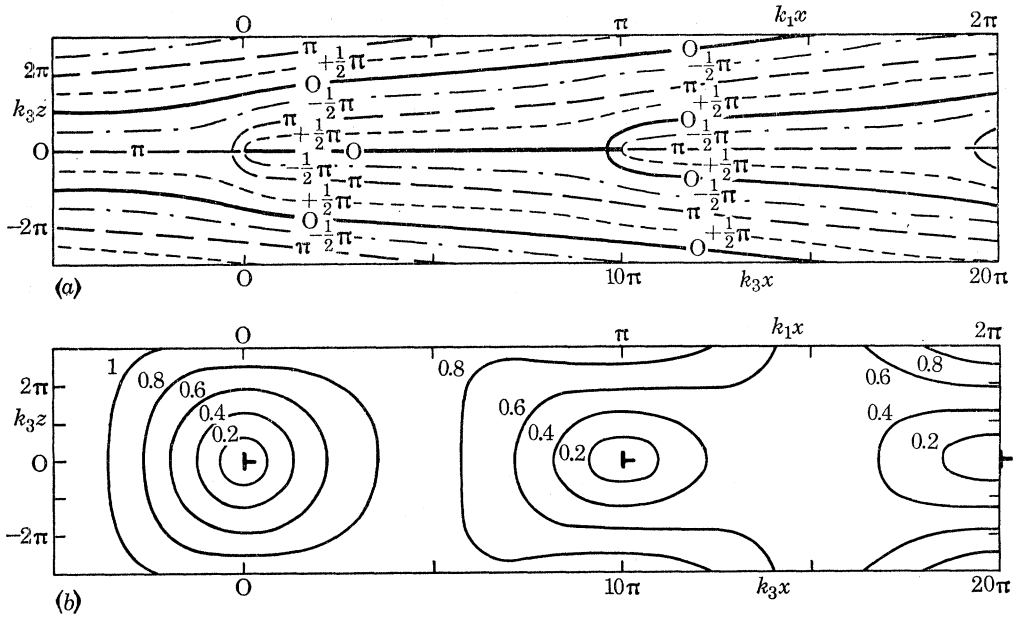

FIGURE 9. (*a*) A row of edge dislocations (tilt boundary) formed by the interference of two pulses of plane waves, each modulated linearly, $\tan\alpha = 0.1$ and $\beta_e = 0{\cdot}1$. The scales are arranged so that the pattern is a true-to-scale map in the (x, z) plane at $t = 0$. Lines of constant reduced phase χ_0 are shown; the dislocations are singularities of χ_0. Any arbitrary value of χ_0 may be chosen to represent crests. The whole pattern moves upwards with velocity $c \sec\alpha$, unchanged in form. (*b*) As (*a*) but showing contours of wave amplitude ρ. The numbers on the curves are values of $\rho/2a_0$. The dislocations are at places where $\rho = 0$.

original oscillation. For example, if the original oscillation is $\cos(-\omega t)$, the resultant is $\rho(x, y, z, t)\cos\chi(x, y, z, t)$ and the full lines where $\chi_0 = 0$ are to be read as crests; this corresponds to figure 8 *a*; at even values of m a new crest appears, while at odd values a crest splits into two. If, on the other hand, the original oscillation is $\sin(-\omega t)$ the crests are the lines where $\chi_0 = \frac{1}{2}\pi$. It can be seen that these run into the dislocations alternately from above and below, corresponding to figure 8 *b*. Alternatively, the lines of $\chi_0 = 0$ may be read as troughs and $\chi_0 = \pi$ as crests, and so on. Crests and troughs defined in this way are of course not necessarily exactly maxima and minima of the disturbance, but they are usually very close to them.

The disturbance described by (9) is not quite periodic in ξ, because both the

constituent wave trains grow stronger towards the left. The time dependence of the disturbance is contained entirely in ζ, and thus the whole pattern simply sweeps upwards parallel to Oz with velocity $\omega/k_3 = c\sec\alpha$. Thus the wavefronts move upwards at a speed greater than c, carrying the dislocations with them.

The fact that the dislocations are equally spaced is of course simply a consequence of using plane wave trains, and at the same time the details of the wave pattern far from the origin are not of primary interest because they depend on the unrealistic and artificial assumption of a linear envelope. We are really interested in the structure of a single dislocation. In order to study this structure undistorted by interaction with neighbouring dislocations in the row we use a limiting process in which the angle 2α between the two primary wave trains is decreased to zero, thus moving the dislocations apart. First define new constants A_0 and β_e^*:

$$A_0 = 2a_0 \sin\alpha \quad \text{and} \quad \beta_e^* = \beta_e \cot\alpha. \tag{12}$$

Then write equation (9) in terms of A_0, β_e^*, k, α, x and z, and let $\alpha \to 0$, keeping A_0, β_e^* and k fixed. In the limit

$$\psi = A_0\{kx + \mathrm{i}\beta_e^*(kz - \omega t)\}\exp \mathrm{i}(kz - \omega t), \tag{13}$$

so that

$$\left.\begin{aligned} \rho^2 &= A_0^2\{k^2x^2 + \beta_e{}^{*2}(kz-\omega t)^2\} \quad (\rho \geqslant 0), \\ \text{and} \qquad \chi &= \arctan\frac{\beta_e^*(kz-\omega t)}{kx} + kz - \omega t + 2n\pi. \end{aligned}\right\} \tag{14}$$

These equations describe a single edge dislocation parallel to Oy, passing through O at $t = 0$, and moving parallel to Oz at a speed $\omega/k = c$. Figure 10 shows ρ and reduced phase χ_0 at $t = 0$. The meaning of the parameter β_e^* (which determines the x scale) may be seen by noting that before the limiting process the change in amplitude of one of the primary wave trains over one period was $2\pi\beta_e a_0 = \pi\beta_e^* A_0$. Thus β_e^* can be regarded as a measure of the non-monochromaticity of the pulse, or of the bandwidth. In the monochromatic limit, $\beta_e^* = 0$, the dislocation becomes infinitely extended along Oz and the phase map turns into figure 5. The larger β_e^* the more compact is the dislocation in the z direction (in crystal dislocation language β_e^* is an inverse measure of the width of the dislocation; it determines the core structure). When $\beta_e^* = 1$ the ρ contours for $t = 0$ are circles and the inverse tangent in (14) is simply the polar coordinate angle θ measured from the x-axis. Thus, near the origin $\chi = \theta$.

At time $t = 0$, or with respect to coordinates carried along with the wave, we have, for $\beta_e^* = 1$,

$$\begin{aligned} \operatorname{grad}\chi &= \operatorname{grad}\theta + k\boldsymbol{n}_z, \\ &= \boldsymbol{n}_\theta/r + k\boldsymbol{n}_z, \end{aligned} \tag{15}$$

where $\boldsymbol{n}_\theta$ and $\boldsymbol{n}_z$ are unit vectors. Thus: if χ is regarded as a potential, we find that the field $\operatorname{grad}\chi$ is the sum of a uniform field along the z axis and a vortex (we were led to this interpretation by a suggestion of Professor F. C. Frank, F.R.S.). Since the uniform field is the undistorted wave, the disturbance produced by the dislocation

is the vortex. In this sense the dislocation may be pictured as a vortex that is carried along with the wave, rather as a vortex in a river is carried along by the main flow. When $\beta_e^* \neq 1$ the general result is the same but the vortex is distorted. There is a stagnation point to the left of the origin where the backward grad χ due to the vortex just cancels the grad χ due to the forward flow. It is readily found that this occurs at

$$x = -\beta_e^*/k. \tag{16}$$

This is the point S in figures 10 and 11, where two contours of $\chi_0 = \pi$ cross at right angles; it is a saddle-point for χ. Figure 11 shows the trajectories of grad χ at $t = 0$ when $\beta_e^* = 1$. Their equation is

$$(kx)^2 + (kz)^2 = Q_0 \exp\{-2(1+kx)\},$$

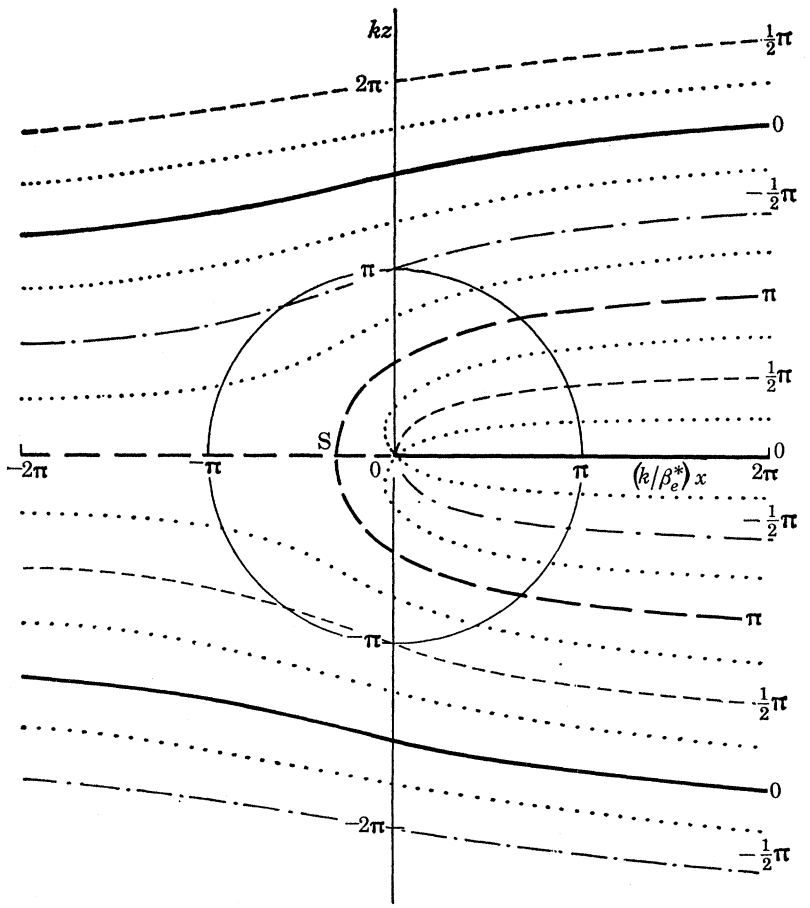

FIGURE 10. A single edge dislocation at the origin showing the lines of reduced phase χ_0 at $t = 0$. Any arbitrary value of χ_0 may be chosen to represent crests. The figure is a true-to-scale map when $\beta_e^* = 1$. ρ is proportional to the radial distance from the origin; the circle shows the contour for $\rho = \pi A_0 \beta_e^*$. S is the stagnation point.

where the constant Q_0 labels the different curves. We have tried without success to find a physical interpretation of this vortex motion (for example, as momentum or energy flux) valid for all types of nondispersive wave; in quantum mechanics $\rho^2 \operatorname{grad} \chi$ would represent a local expectation value of momentum.

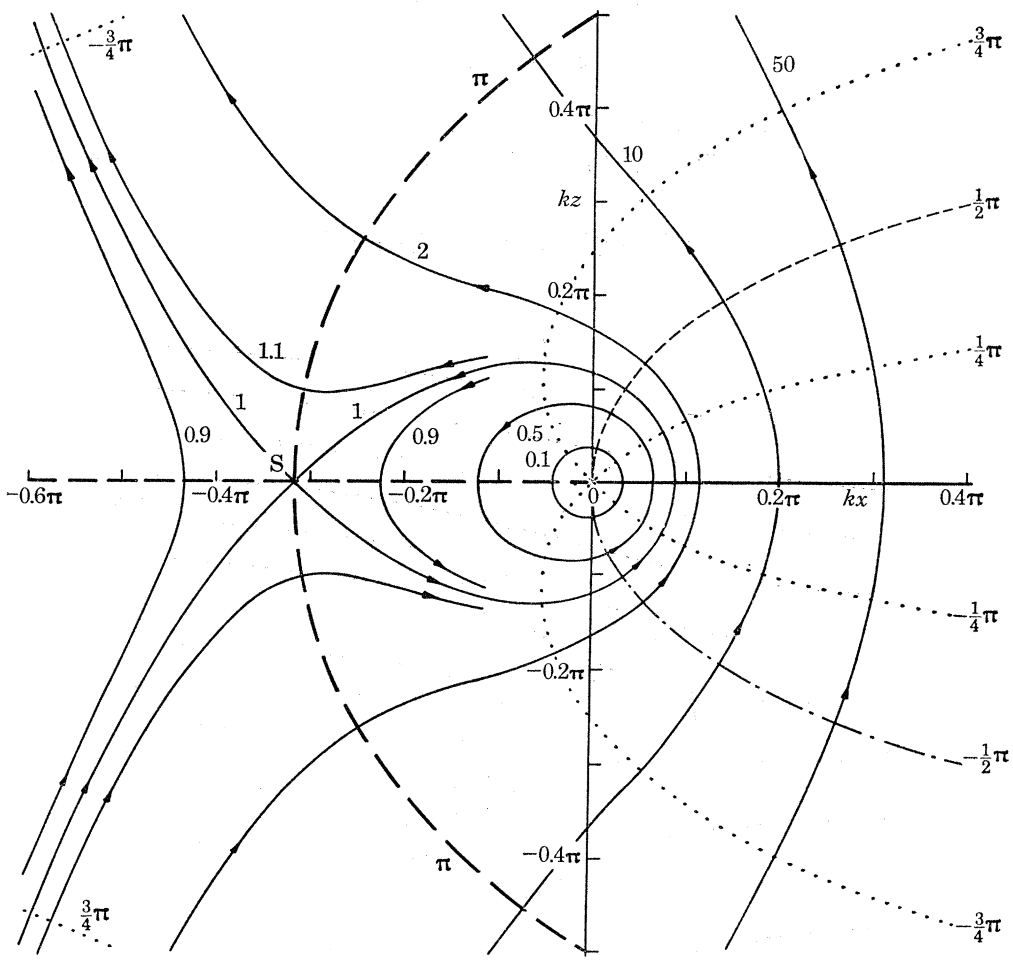

FIGURE 11. The same edge dislocation as in figure 10, but on a larger scale and showing the trajectories of grad χ at $t = 0$ for $\beta_e^* = 1$. The numbers on the curves are the values of Q_0. Curves of constant reduced phase χ_0 are also shown.

The details in figures 2 *a* to *d* may now be described more precisely. The figures show the disturbance at $z = 0$ corresponding to the real part $\psi_c(t)$ of $\psi(t)$ for four values of x. For $x < -2\beta_e^*/k$ the curve is concave upwards at the centre-point $t = 0$, but for $-2\beta_e^*/k < x < 0$ (figure 2 *b*) it is convex upwards and flanked by two minima (the downward curvature of the envelope outweighing the upward curvature of the cosine). Two new zeros appear precisely at the centre of the dislocation (figure 2 *c*), but they are preceded by the appearance of the new maximum slightly below

the zero level. In the antisymmetrical structure of figure 3, which corresponds to the imaginary part $\psi_s(t)$ of $\psi(t)$, the two new zeros appear at $x = -\beta_e^*/k$. As discussed in § 2, the unique property of the dislocation line $x = 0$, $\zeta = 0$ is that the disturbance is zero for *all* values of the phase of the carrier wave relative to the pulse envelope, or, in other words, for all linear combinations of the real and imaginary parts of the complex disturbance ψ.

5. Screw and mixed edge-screw dislocations

So far we have modulated the two primary wave trains linearly in their directions of propagation. If we try to modulate a train of plane waves parallel to the wavefronts the modulation will normally spread out by diffraction, but it may be verified that if the modulation is merely linear with position the wave propagates without change. Thus, for example, the wave

$$\psi = (a + bx)\exp[\mathrm{i}(kz - \omega t)], \tag{17}$$

where a and b are constants, satisfies the wave equation (4) (cf. equation (13)).

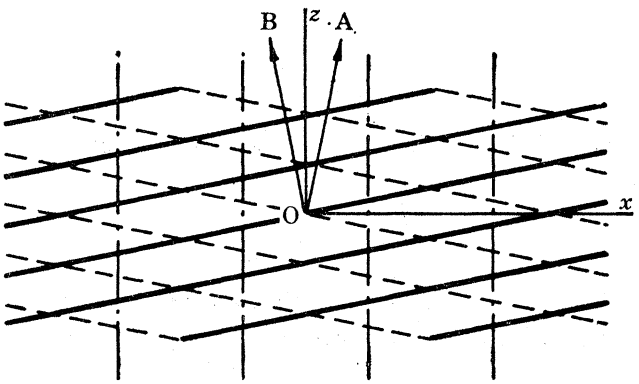

Figure 12. Interference between two monochromatic plane waves A and B making an angle with one another and modulated linearly in opposite directions along Oy, the direction that is common to both wavefronts. Oy points into the page. The result is a set of parallel screw dislocations (—·—·—) parallel to Oz and lying in a plane parallel to the page.

Accordingly, let the two primary waves A and B be now modulated linearly along Oy, the direction common to both wavefronts, so that as one rises the other falls, and let there be no modulation along their propagation directions so that the waves are monochromatic. The general effect is seen in figure 12. If the amplitudes are the same in the plane $y = 0$, one wave, B say, will dominate above the plane of the diagram while A will dominate below. The resultant is a set of screw dislocations in the plane $y = 0$, all parallel to Oz and equally spaced along Ox. It can be seen from figure 12 that they are right-handed. In crystallographic terms this is a twist boundary.

For mixed screw-edge dislocations we simply add modulation along the propagation directions. The analysis for screw, edge and mixed screw-edge dislocations is conveniently done all together. Let the two primary waves now be

$$\left.\begin{aligned}\psi_{\mathrm{A}} &= a_0\{1-\beta_{\mathrm{e}}(k_1x+k_3z-\omega t)+\beta_{\mathrm{s}}k_1y\}\exp\left[\mathrm{i}(k_1x+k_3z-\omega t-\tfrac{1}{2}\pi)\right],\\ \psi_{\mathrm{B}} &= a_0\{1+\beta_{\mathrm{e}}(-k_1x+k_3z-\omega t)-\beta_{\mathrm{s}}ky\}\exp\left[\mathrm{i}(-k_1x+k_3z-\omega t+\tfrac{1}{2}\pi)\right],\end{aligned}\right\} \tag{18}$$

where β_{s} is a constant. Writing $k_1y=\eta$, we find that the resultant disturbance $\psi=\psi_{\mathrm{A}}+\psi_{\mathrm{B}}$ is

$$\psi = 2a_0\{(1-\beta_{\mathrm{e}}\xi)\sin\xi+\mathrm{i}(\beta_{\mathrm{e}}\zeta-\beta_{\mathrm{s}}\eta)\cos\xi\}\exp(\mathrm{i}\zeta)=\rho\exp(\mathrm{i}\chi). \tag{19}$$

Hence
$$\rho^2 = 4a_0^2\{(1-\beta_{\mathrm{e}}\xi)^2\sin^2\xi+(\beta_{\mathrm{e}}\zeta-\beta_{\mathrm{s}}\eta)^2\cos^2\xi\} \quad (\rho\geqslant 0), \tag{20}$$

and
$$\chi = \arctan\left\{\frac{\beta_{\mathrm{e}}\zeta-\beta_{\mathrm{s}}\eta}{(1-\beta_{\mathrm{e}}\xi)\tan\xi}\right\}+\zeta+2n\pi, \tag{21}$$

where the conventions about the π are the same as before.

As always, the dislocations lie along the lines where $\psi_{\mathrm{c}}=\psi_{\mathrm{s}}=0$; from (19), they have the equations

$$\left.\begin{aligned}\xi &= m\pi,\\ \beta_{\mathrm{e}}\zeta-\beta_{\mathrm{s}}\eta &= 0,\end{aligned}\right\} \tag{22a}$$

or, at $t=0$,

$$\beta_{\mathrm{e}}^*z-\beta_{\mathrm{s}}y=0 \tag{22b}$$

(ignoring the solution $\beta_{\mathrm{e}}\xi=1$ as being outside the field of interest). Define constants β and δ so that

$$\left.\begin{aligned}\beta_{\mathrm{e}}^* &\equiv \beta\cos\delta,\\ \beta_{\mathrm{s}} &\equiv \beta\sin\delta.\end{aligned}\right\} \tag{23}$$

δ is then the inclination of the dislocation lines to Oy. $\delta=0$ corresponds to $\beta_{\mathrm{s}}=0$ and pure edge dislocations. $\delta=\frac{1}{2}\pi$ corresponds to $\beta_{\mathrm{e}}^*=\beta_{\mathrm{e}}=0$ and pure screw dislocations.

Proceeding to the limit $\alpha\to 0$ as before, we find

$$\psi = A_0[kx+\mathrm{i}\{\beta_{\mathrm{e}}^*(kz-\omega t)-\beta_{\mathrm{s}}ky\}]\exp\left[\mathrm{i}(kz-\omega t)\right], \tag{24}$$

so that

$$\left.\begin{aligned}\rho^2 &= A_0^2[k^2x^2+\{\beta_{\mathrm{e}}^*(kz-\omega t)-\beta_{\mathrm{s}}ky\}^2] \quad (\rho\geqslant 0),\\ \text{and}\quad \chi &= \arctan\frac{\beta_{\mathrm{e}}^*(kz-\omega t)-\beta_{\mathrm{s}}ky}{kx}+kz-\omega t+2n\pi.\end{aligned}\right\} \tag{25}$$

This describes a single mixed dislocation lying in the yz plane at an angle δ to Oy, passing through O at $t=0$ and moving parallel to Oz at velocity c. To obtain a pure screw dislocation put $\beta_{\mathrm{e}}^*=0$. Then

$$\chi = -\arctan\frac{\beta_{\mathrm{s}}y}{x}+kz-\omega t+2n\pi. \tag{26}$$

If one encircles the z-axis at constant z and fixed t, χ changes by 2π for each

revolution. If $\beta_s = 1$ the surfaces of constant phase at given time are helicoids and we obtain the simple helicoidal wave

$$\psi = A_0 kr \exp[\mathrm{i}(kz - \omega t - \phi)], \tag{27}$$

where r, ϕ, z are cylindrical polar coordinates. $\beta_e^* = 0$ represents the monochromatic limit; the screw dislocation is one of the special cases where a dislocation can exist in a pure monochromatic wave.

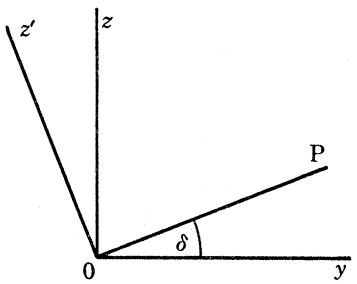

FIGURE 13. Axes for a mixed edge-screw dislocation. OP is the dislocation line; Ox points out of the plane of the paper.

For the mixed dislocation take a new axis Oz' in the yz plane and perpendicular to the line of the dislocation (figure 13); thus

$$z' = z\cos\delta - y\sin\delta. \tag{28}$$

Then equations (25) become, at $t = 0$,

$$\rho^2 = A_0^2 k^2(x^2 + \beta^2 z'^2), \quad (\rho \geqslant 0), \tag{29a}$$

$$\chi = \arctan(\beta z'/x) + kz + 2n\pi. \tag{29b}$$

This may be pictured as a vortex circulating about the dislocation line in a uniform flow parallel to Oz. If the inclination δ (which determines the edge-screw character) is varied while β is held fixed, equation (29a) shows that the dislocation always has the same field of ρ when it is viewed in its own coordinate system. β is a measure of the nonmonochromaticity of the pulses, by extension of the argument already given for a pure edge dislocation; β is also an inverse measure of how far the dislocation spreads out in the yz plane perpendicular to its length.

Now that we have derived equations (13) and (24) for isolated dislocations, by mixing two plane wave trains and using a limiting process, we can be wise after the event and see that these equations (but not those for rows of dislocations) could have been guessed *a priori*, by using the principle that a single plane wave train can be amplitude modulated linearly in any direction in its wavefront and the fact that this still holds if the coefficients are complex. From this point onwards, our method of constructing dislocations will be more direct. We shall simply present complex wave functions ψ which satisfy the wave equation (4) and shall then interpret them as representing various kinds of dislocations. ρ and χ, as always, are the amplitude and phase of ψ.

One way of realizing such dislocations physically would be to place on the plane $z = z_0$, where z_0 is large and negative, that distribution of sources which will produce the correct boundary condition $\psi = \psi(x, y, z_0, t)$. In each case it will be found that the required sources are quasi-monochromatic in the sense defined in §2, being derived by (complex) amplitude modulation from the primary oscillation $e^{-i\omega t}$.

Equation (27) is the simplest example from a class of helicoidal waves describing *multiple* pure screw dislocations, of 'strength' s; a more general expression which can easily be shown to satisfy the wave equation (4) is

$$\psi = (Ar^s + [B/r^s]) \exp[i(kz - \omega t \pm s\phi)], \tag{30}$$

where A and B are constants and s is a positive integer; the choice of sign in the exponent determines the 'handedness' of the dislocation. The solutions associated with the coefficient B are singular at the origin. Equation (30) is still not the most general expression for an isolated screw dislocation; for example, the wave reflected from a surface consisting of one turn of a helicoid near $z = 0$, whose pitch is s half-wavelengths, has the asymptotic form

$$\psi \underset{\substack{r\to\infty \\ z\to\infty}}{\longrightarrow} A \exp[i(kz - \omega t \pm s\phi)], \tag{31}$$

which is simpler than (30), but it can be shown that the core structure (r small) is very complicated.

Thus multiple screw dislocations exist. What about multiple edge and mixed dislocations? We conjecture that these cannot exist, although we are unable to give a general proof. The pattern of figure 8*a* is suggestive on this point. It shows 'extra half-planes' coming together in pairs as they would at edge dislocations of double strength. But when the local readjustments have taken place (figure 9*a*) each double dislocation has split into a pair of single-strength dislocations with phase-structures differing by π. A further argument is that the slightest perturbation splits a double pure screw dislocation into two single-strength mixed dislocations. To show this, we proceed by analogy with equation (24), which gives a mixed dislocation as the sum of a pure single-strength screw and a linearly modulated plane wave; thus we consider the wave

$$\psi = A_0\{r^2 e^{-2i\phi} + i\beta_e^*(kz - \omega t)\} \exp[i(kz - \omega t)]. \tag{32}$$

The dislocations are the lines where ψ_c and ψ_s are simultaneously zero, that is where

$$\left.\begin{aligned} r^2 \cos 2\phi &= 0, \\ r^2 \sin 2\phi &= \beta_e^* \zeta, \end{aligned}\right\} \tag{33}$$

ζ now denoting $kz - \omega t$. These describe two parabolas lying in orthogonal vertical planes, namely

$$\left.\begin{aligned} \zeta &= r^2/\beta_e^*, \quad & \phi &= \tfrac{1}{4}\pi, -\tfrac{3}{4}\pi, \\ \zeta &= -r^2/\beta_e^*, \quad & \phi &= -\tfrac{1}{4}\pi, \tfrac{3}{4}\pi. \end{aligned}\right\} \tag{34}$$

The dislocations are of single strength, and change their character with ζ, being pure

screw when $|\zeta| = \infty$ and pure edge when $\zeta = 0$. In the limit when $\beta_e^* = 0$ the parabolas degenerate into the original double pure screw along the z-axis, and in the limit $|\beta_e^*| = \infty$ they degenerate into two single-strength edge dislocations intersecting orthogonally in the plane $\zeta = 0$.

As well as suggesting that multiple dislocations do not exist unless they are pure screw, the wave (32) introduces two new properties of dislocation lines: they may be curved, and two edges may intersect. A simple model showing the intersection of a screw and an edge is

$$\psi = A_0\{kx + i\beta y(kz - \omega t)\}\exp[i(kz - \omega t)], \tag{35}$$

corresponding to dislocations where

$$x = 0 \quad \text{and} \quad y\zeta = 0, \tag{36}$$

that is, to a screw along Oz intersecting an edge parallel to Oy.

All the patterns of straight, curved, edge, screw and mixed dislocations discussed so far have arisen from waves which contain z and t only in the combination ζ; that is, they are of the form

$$\psi = \psi(x, y, \zeta). \tag{37}$$

It would be possible to generate a great variety of dislocation patterns, using the fact, derivable from (4), that waves of this type satisfy the two-dimensional Laplace equation

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = 0. \tag{38}$$

However, all such dislocations would be rigidly attached to the wavefronts, and in order to investigate the ways in which dislocations can move relative to the wavefronts, and to one another, it is necessary to widen the class of waves considered, to include solutions where z and t do not appear solely in the combination ζ. Some waves of this type will be examined in §7. In order to describe the motion of the resulting dislocations we first introduce some concepts from crystal dislocation theory.

6. Burgers circuit, glide and climb

The theorem, well-known for crystal dislocations, that the Burgers vector is conserved along the length of a single dislocation applies equally well for dislocations in wave trains. A closed circuit (e.g. JKLM in figure 4) is traversed at fixed time in the dislocated wave and the value of

$$(\lambda/2\pi)\oint d\chi = N\lambda, \quad \text{say}, \tag{39}$$

where λ is the wavelength $2\pi/k$, is the analogue of the Burgers vector. N is necessarily an integer or zero and equals the total strength of the dislocation lines encircled (with due regard to their signs). If the shape and position of the circuit is altered the value of the integral remains the same provided the circuit continues to

encircle the same dislocation lines. In crystal dislocation language the Burgers vector of each dislocation is $\lambda \boldsymbol{n}$, where $\boldsymbol{n}$ is a unit vector along the wave normal in an undislocated reference wave. The difference is that in the wave, as distinct from the crystal, components of the Burgers vector parallel to the reference wavefront are of no significance. Only the magnitude and sense of the Burgers vector is significant and these are given by $N\lambda$ (positive or negative).

The Burgers circuit just described was traversed at fixed time. If we now allow the circuit to travel with the wave it follows by continuity that the Burgers vector for the circuit remains the same provided it is not crossed by a dislocation line; the magnitude of the Burgers vector for the moving circuit remains equal to the total strength of the dislocation lines encircled multiplied by the wavelength.

Wave trains, unlike crystals, are oscillatory in time as well as space, and therefore one may consider a more general kind of 'Burgers circuit' where the line element can be wholly or partly time-like as well as space-like. If the wave disturbance is now thought of as varying in the four dimensions (x, y, z, t), the line integral of χ round any closed circuit drawn in this space must necessarily be a multiple of 2π; thus

$$\oint \mathrm{d}\chi = 2N\pi, \tag{40}$$

where N is an integer or zero. If the circuit lies in a plane for which t = constant, we have (39); on the other hand, the method of observing a dislocation described in § 1 and the argument about the Argand diagram in § 2 can be regarded as employing a circuit in the (x, t) plane. By shrinking the circuit around a dislocation continuously to zero one can see from (40) that χ must be singular on the dislocation itself – and it then follows from the properties of the Argand diagram, together with the physical requirements that the wave function must be continuous and single-valued, that ρ must be zero. Thus, by *defining* a dislocation by means of a Burgers circuit one is led automatically to the conclusion that the precise location of the line is the place where $\rho = 0$ (rather than, for example, various other alternative possibilities that might be suggested by figure 6).

With crystal dislocations a Burgers circuit always has to be chosen so as to avoid passing through 'bad crystal' where the lattice is so badly distorted that the crystallographic directions cannot be unambiguously identified (Frank 1951). It is interesting to notice that this particular complication does not appear with wave dislocations; χ can be well defined everywhere (provided the waves are ultimately derived from a source with a carrier frequency ω) and a dislocation has no disordered core. There is no breakdown of linearity in the wave equation near the singularities. Thus, however close together the dislocations may be, they do not lose their identities; they are always recognizable as singularities of χ where ρ is zero. With crystals, on the other hand, the dislocation concept itself has to be abandoned when the dislocations are very close together. Of course, in practice it may be impossible to resolve individual wave dislocations when they are very close together because of the presence of noise, but that is another matter.

A dislocation may move either by *glide*, defined as motion in the plane containing

the Burgers vector and the direction of the dislocation line (the *glide plane*), by *climb*, defined as motion perpendicular to this plane, or by a combination of glide and climb. Thus a pure edge dislocation, which is normal to its Burgers vector, glides by moving perpendicular to the wavefronts (more strictly, perpendicular to the reference wavefronts), and climbs by moving parallel to the wavefronts. For a pure screw dislocation, which is parallel to its Burgers vector, all vertical planes are glide planes, and climb has no meaning.

7. Motion, birth, annihilation and collision of dislocations

First we consider edge dislocations parallel to the y-axis. As we have already seen, to obtain moving dislocations, it is necessary that z and t should not appear only in the combination ζ. The simplest monochromatic wave of this new type involves z linearly, and can be seen by inspection of the wave equation to be

$$\psi_1 = (k^2x^2 + \mathrm{i}kz)\exp[\mathrm{i}(kz - \omega t)]. \tag{41}$$

The inevitable appearance of quadratic modulation in x means that moving dislocations obtained using ψ_1 will occur in pairs.

To obtain *glide*, we add ψ_1 to a multiple of the wave function (13) describing a single edge dislocation; this gives

$$\psi = \{Akx + k^2x^2 + \mathrm{i}B(kz - \omega t) + \mathrm{i}kz\}\exp[\mathrm{i}(kz - \omega t)]. \tag{42}$$

The dislocation lines satisfy

$$\left.\begin{aligned} Akx + k^2x^2 &= 0, \\ B(kz - \omega t) + kz &= 0, \end{aligned}\right\}$$

or

$$\left.\begin{aligned} x &= 0 \quad \text{or} \quad -A/k, \\ z &= \frac{Bc}{B+1}t. \end{aligned}\right\}$$

There are thus two dislocations, which glide parallel to Oz, and have velocity v_g given by

$$v_g = \frac{B}{B+1}c, \tag{43}$$

while the undistorted wavefronts move with velocity c. Thus in a frame of reference moving with the dislocations the lines of constant χ sweep through in sequence, backwards or forwards; glide is equivalent to a steady change of phase of the dislocation.

v_g can have all values, positive or negative, according to the value of B. For $-1 < B < 0$, v_g is negative: the dislocations move backwards. As $B \to -1$, the dislocations become infinitely spread out in the z direction and $|v_g| \to \infty$, rather as the point of coincidence of a vernier can be made to travel very fast even though the

scales themselves are only moving slowly. If $|B| \to \infty$, $v_g \to c$ and the disturbance consists of two non-gliding edge dislocations. Finally, if B is zero, v_g is zero; this is the monochromatic limit, and the two dislocations are stationary features of the wave, in fact two localized interference fringes. From this point of view our moving dislocations may be regarded as moving interference fringes.

It is also possible to obtain the 'glide' wave function (42) by a limiting process analogous to that used in § 4 to obtain the single edge dislocation: two primary plane wave trains, linearly modulated in their directions of propagation, are also linearly modulated parallel to their wavefronts and perpendicular to their common direction.

By replacing the linear factor $(kz - \omega t)$ in (42) by a quadratic factor, which corresponds to taking two interfering wave trains whose pulse envelope is not linear but quadratic, pairs of edge dislocations can be made to annihilate by glide, or to appear spontaneously. The wave function is

$$\psi = \{Akx + k^2x^2 + \mathrm{i}B(kz - \omega t)^2 + \mathrm{i}kz\} \exp[\mathrm{i}(kz - \omega t)], \tag{44}$$

whose dislocations lie at

$$\left.\begin{aligned} x &= 0 \quad \text{or} \quad -A/k, \\ z &= ct - \frac{1}{2Bk}\{1 \pm \sqrt{(1 - 4B\omega t)}\}. \end{aligned}\right\} \tag{45}$$

If $B > 0$ there are two pairs of approaching dislocations of opposite sign which glide together at $t = (4B\omega)^{-1}$ and mutually annihilate, while if $B < 0$ the pairs of dislocations suddenly appear at $t = -(4|B|\omega)^{-1}$ and then glide apart.

To obtain *climb* of edge dislocations we simply set $A = 0$ in (42), and replace B by $B - \mathrm{i}C$; thus the resulting dislocations satisfy

$$x^2 = \frac{Cz}{Bk}, \quad z = \frac{Bct}{B+1}. \tag{46}$$

There are two dislocations of opposite sign: both have the same z value which need not change with speed c, so that the dislocations can glide. In addition they may separate or approach one another, that is, they may climb. If $C/(B+1) > 0$, the two dislocations are born at $t = 0$ and separate by climb, initially with infinite speed; this would manifest itself as the sudden tearing of a wavefront (when $B+1 > 0$) or as the appearance of an expanding strip of wavefront (when $B+1 < 0$). (When a pulse is reflected from a scattering object the waves that have travelled the shortest path, and therefore arrive first, are usually travelling in the direction of the average reflected wave, whereas later arrivals tend to be oblique. It may be shown that this results *on the average* in the disappearance of wavefronts. Therefore, in this situation, we expect the appearance of a new piece of wavefront to be a comparatively rare event.) If $C/(B+1) < 0$, the two dislocations climb towards each other and annihilate at $t = 0$; this would manifest itself as the sudden disappearance of a strip of wavefront (when $B+1 < 0$), or, rarely, as the spontaneous healing of a tear (when $B+1 > 0$). In both cases the surface swept out by the moving

dislocations is a parabolic cylinder whose axis is parallel to Oy. These solutions may also be obtained by adding a third wave train, travelling along Oz, to the two considered previously (§4 and figure 8), and taking the limit $\alpha \to 0$.

To obtain a *circular climbing edge dislocation loop* (climbing prismatic dislocation) it is only necessary to notice that if x^2 is replaced by $\frac{1}{2}r^2$ in (41), where r is the radial coordinate of cylindrical coordinates, ψ_1 will satisfy the wave equation. Thus the wave

$$\psi = \{\tfrac{1}{2}k^2r^2 + \mathrm{i}kz + \mathrm{i}(B - \mathrm{i}C)(kz - \omega t)\} \exp[\mathrm{i}(kz - \omega t)] \tag{47}$$

describes a circular edge dislocation loop sweeping out a paraboloid of revolution whose axis lies along Oz. If $C/(B+1) > 0$ the loop is born at $t = 0$ (appearing as the sudden puncture of a wavefront, or, rarely, as the birth of an expanding circular island of wavefront), while if $C/(B+1) < 0$ the loop vanishes at $t = 0$ (disappearance of an island of wavefront, or, rarely, the spontaneous healing of a puncture). The solution (47) may also be obtained by a limiting process involving the interference of a plane wave with a toroidal wave (for example, from an annular source).

It is possible for edge dislocations to *collide and rebound* without annihilation. To construct this situation we require a slightly more complicated wave than (41), namely

$$\psi_2 = (\tfrac{1}{3}k^3x^3 + \mathrm{i}k^2xz) \exp[\mathrm{i}(kz - \omega t)]. \tag{48}$$

Combining ψ_2 with a plane wave with linear pulse envelope, we obtain

$$\psi = \{\tfrac{1}{3}k^3x^3 + B(kz - \omega t) + \mathrm{i}k^2xz\} \exp[\mathrm{i}(kz - \omega t)], \tag{49}$$

which has two edge dislocations of opposite sign, parallel to Oy, satisfying

$$\left.\begin{aligned} xz &= 0, \\ \tfrac{1}{3}k^3x^2 + B(kz - \omega t) &= 0. \end{aligned}\right\} \tag{50}$$

The dislocation for which $x = 0$ moves along Oz with speed c (i.e. it neither climbs nor glides), while the dislocation for which $z = 0$ moves along Ox with a speed which varies, becoming infinite at $x = 0$ (that is, it climbs along Ox while gliding backwards at speed c). At time $t = 0$ the two dislocations collide. Afterwards one moves off along Ox and the other along Oz. If we identify the dislocations by their signs it may be shown that each has been deflected through a right angle.

Finally we consider the *motion of pure screw dislocations*. In their simplest form, given by equation (27), these are monochromatic, and it is clear that the addition of another monochromatic wave such as ψ_1 (equation 41) cannot produce motion. But there is a 'complementary' solution to ψ_1, where t appears linearly instead of z; it is easily verified that this solution is

$$\psi_3 = (\omega t - \mathrm{i}k^2x^2) \exp[\mathrm{i}(kz - \omega t)]. \tag{51}$$

Adding ψ_3 to a real multiple of the 'screw' wave function (27), we obtain

$$\psi = \{\omega t - \mathrm{i}k^2x^2 + Ak(x - \mathrm{i}y)\} \exp[\mathrm{i}(kz - \omega t)], \tag{52}$$

which has a single dislocation, satisfying

$$\left.\begin{aligned} x &= -\omega t/Ak, \\ y &= -kx^2/A. \end{aligned}\right\} \tag{53}$$

Since it is parallel to Oz this is a screw dislocation; it glides along a parabola in the xy plane. Replacing A by iB, we find that the resulting wave has two screw dislocations of opposite sign, satisfying

$$\left.\begin{aligned} x = 0, \quad x &= B/k; \\ y &= -ct/B. \end{aligned}\right\} \tag{54}$$

These glide parallel to the y-axis while maintaining a constant separation. For *annihilation and creation of screw dislocations* we replace A in (52) by $A+iB$, obtaining the dislocation line equations

$$y = -\frac{ct+Ax}{B} = \frac{-kx^2+xB}{A}. \tag{55}$$

These describe two screw dislocations moving along a parabola in the xy plane; they are created and glide apart if $AB > 0$, and they glide together and annihilate if $AB < 0$.

8. Concluding remarks

We have shown that wavefront dislocation lines, along which the amplitude is zero, are perfectly analogous to crystal dislocations so far as their kinematic properties are concerned. On the other hand, there is nothing analogous to the forces between crystal dislocations or to their line tensions. We may also note that, while their topology is virtually identical, the superposition properties of the two sorts of dislocations are different. Crystal dislocations are lines where, mathematically, the stress and strain become infinite; so when two elastic strain-fields, both containing dislocations, are added together, the dislocations remain in place. In wave trains, on the other hand, $\rho = 0$ on the dislocations, and so when two dislocated wave fields are added each one tends to destroy the dislocations of the other, although new dislocations may be created elsewhere by interference. (If the second field is added gradually, the original dislocations move continuously, but they may be annihilated and new ones may be created.) Another difference is that, in crystals, climb tends to be a slow process compared with glide because it requires diffusion of lattice vacancies, whereas with wavefront dislocations there is no corresponding restriction on the speed of climb.

It is interesting to compare wavefront dislocations with caustic surfaces and focal lines. These are envelopes of the rays of geometrical optics, that is, loci of points where the density of rays is infinite. At caustics and foci the wave amplitude is generally large, becoming infinite in the limit when the wavelength becomes vanishingly small (in comparison with the radii of curvature of rays and the

dimensions of diffracting objects). Thus, when geometrical optics is a valid approximation, caustics and foci are the dominant features in the wave field. However, in this limit dislocations are unobservable, because they can only be unambiguously identified by measuring the phase change around a Burgers circuit, and phase is a quantity which varies infinitely fast when the wavelength is zero. In long waves, on the other hand, dislocations are readily observable, but caustics and foci lose their prominence because the amplitude is no longer especially large. Thus dislocations are entities complementary, in a certain sense, to foci and caustics.

For simplicity, and also to emphasize that dispersion is not necessary for the production of dislocations, we have restricted ourselves to the scalar wave equation (4). But dispersive wave trains can also contain dislocations; in particular, the time-dependent Schrödinger equation of quantum mechanics may be analysed by the methods of this paper, leading to the conclusion that matter waves contain a tangled 'cobweb' of dislocation lines on which the electron density is zero. These cobwebs are not superposable, even though the wave equation is linear.

Wavefront dislocations can also exist in two dimensions (for example in water waves), and take the form of points instead of lines. Only edge dislocation points may occur. Examples of the monochromatic case, where the dislocation points are localized interference fringes (see the paragraph following equation (43)) occur in a paper by Braunbek & Laukien (1952): their figure 2 shows the lines of constant phase for the field of a plane wave diffracted by a half-plane; several dislocation points may be seen. Other examples of this special case are the 'amphidromic points' for tides (Whewell 1833; Defant 1961) around which the lines of constant phase travel like the spokes of a wheel (the phase structure is identical with that very near the origins of figures 10 and 11). There are two amphidromic points in the North Sea.

It is clear that the use of other combinations of allowed modulation and the addition of further wave trains will produce more complex arrangements of dislocations moving in more complex ways. Moreover, the waves carrying the dislocations need not be plane. Reflexion of a pulse from a rough surface is one way of producing a large number of interacting pulses and may therefore be expected to result in a complex field of dislocations. Scattering of a pulse from any spatial array of scattering centres will produce a similar effect. These are subjects for further study, and some progress has been made on them. In this paper we have simply shown that dislocations exist and have tried to explore some of their elementary properties.

We are grateful to Dr M. E. R. Walford for his valuable help and advice. The ultrasonic apparatus used for the observations of § 1 was made by R. G. Kyte and D. C. Threlfall, to whom we also express our thanks, and an improved version was designed and made by Dr Walford. The experimental work was partly supported by a grant from the Natural Environment Research Council.

References

Braunbek, W. & Laukien, G. 1952 *Optik* **9**, 174–9.
Defant, A. 1961 *Physical oceanography*, volume 2. London: Pergamon Press.
Findlay, J. W. 1951 *J. Atmos. Terrest. Phys.* **1**, 353–366.
Frank, F. C. 1951 *Phil. Mag.* Series 7, **42**, 809–819.
Nabarro, F. R. N. 1967 *Theory of crystal dislocations*. Oxford: Clarendon Press.
Nye, J. F., Kyte, R. G. & Threlfall, D. C. 1972 *J. Glaciol.* **11**, 319–325.
Nye, J. F., Berry, M. V. & Walford, M. E. R. 1972 *Nature, Phys. Sci.* **240**, 7–9.
Read, W. T. 1953 *Dislocations in crystals*. New York: McGraw-Hill.
Walford, M. E. R. 1972 *Nature, Lond.* **239**, 93–95.
Whewell, W. 1833 *Phil. Trans. R. Soc. Lond.* **123**, 147–236.