PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Dislocations in Diffraction Patterns: Continuous Waves and Pulses

F. J. Wright and J. F. Nye

Phil. Trans. R. Soc. Lond. A 1982 **305**, 339-382
doi: 10.1098/rsta.1982.0041

Phil. Trans. R. Soc. Lond. A **305**, 339–382 (1982) [339]
Printed in Great Britain

DISLOCATIONS IN DIFFRACTION PATTERNS: CONTINUOUS WAVES AND PULSES

By F. J. WRIGHT†‡ and J. F. NYE, F.R.S.
H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, U.K.
† *and Institut für Informationsverarbeitung, Universität Tübingen, F.R.G.*

(*Received* 28 *July* 1981)

CONTENTS

‡ Permanent address: Department of Applied Mathematics, Queen Mary College, London University, Mile End Road, London E1 4NS, U.K.

Vol. 305. A 1490. 22 [Published 16 June 1982

When a monochromatic wave is scattered by fixed objects it produces a stationary three-dimensional diffraction pattern, but if the driving oscillation is a quasi-monochromatic pulse the diffraction pattern naturally changes with time. Both the fixed and the moving patterns may be usefully characterized by their dislocation lines, where the amplitude is zero and the phase is indeterminate. The paper studies the relation between the fixed dislocation lines (continuous wave (c.w.) null lines, interference fringes) of the monochromatic pattern and the moving lines of the pulse pattern. The latter sweep out surfaces, called the dislocation trajectories.

A simple model system with two interfering beams illustrates how dislocation lines can appear and disappear from the head or tail of the pulse and how pairs of dislocations can be created in two different ways. It also shows how the shape and bandwidth of the pulse affect not so much the positions of the trajectories, but their lengths and the way they are traversed in time by the dislocations.

In a general theory the pulse is regarded as a slightly perturbed continuous wave and its behaviour is deduced from the continuous wave response of the (linear) system within the bandwidth of the pulse. The method is based on obtaining the fastest initial convergence of a functional series. For very small bandwidth it leads to a generalization of the concept of group velocity. In this case dislocations by interference appear only very close to c.w. nulls. For larger, but still small, bandwidth, moving dislocations appear whose trajectories are approximately parts of *frequency minimum surfaces*, that is, surfaces in the continuous-wave diffraction pattern where the amplitude is a minimum with respect to changes in frequency. These surfaces contain the c.w. null lines. In the next approximation it is found that the trajectory surfaces nearly, but not quite, contain these lines. In certain two-dimensional cases the trajectories follow valleys in the landscape defined by the continuous-wave amplitude pattern. The prediction of the arrival times of the dislocations is delicate. General formulae, suitable for numerical computation, are given in terms of frequency derivatives of the continuous-wave diffraction pattern. They give successful results in three applications to pulsed wave fields: the diffraction pattern near a cusped caustic, the two-beam model, and a piston radiator.

1. Introduction

Conventional remote sensing systems, such as radar or ultrasonics, customarily use pulses of quasi-monochromatic waves. When interference effects are encountered a useful, and usual, approximation is to consider the waves to be monochromatic. If an oscillator produces a continuous harmonic scalar wave, which is scattered by a system of fixed objects, the resulting three-dimensional diffraction pattern is fixed in space. If, however, the original oscillation is modulated in amplitude so as to produce a pulse of finite length, rather than an infinite wave train, the resulting diffraction pattern will naturally change with time. In this paper we study the relation between the moving diffraction pattern produced by the pulse and the corresponding stationary diffraction pattern produced by the continuous wave.

It is useful to regard both diffraction patterns as based on their dislocations (Nye & Berry 1974; Nye 1981). These are structurally stable features, being singular lines in space where the wave amplitude vanishes and the phase is indeterminate. For the continuous wave the dislocation lines are stationary and are simply localized interference fringes; for the pulse they are moving lines which trace out surfaces, called the dislocation trajectories. Our main concern will be to explore the relation between these basic structural elements of the two diffraction patterns: the fixed dislocation lines of the monochromatic pattern and the moving dislocation lines of the pulse pattern, with their associated trajectories.

2. A TWO-BEAM MODEL

In §3 we give a general theory for pulses of small bandwidth; but it is helpful to begin with a two-dimensional model system which, although simple, nevertheless displays some of the most important features of the general case.

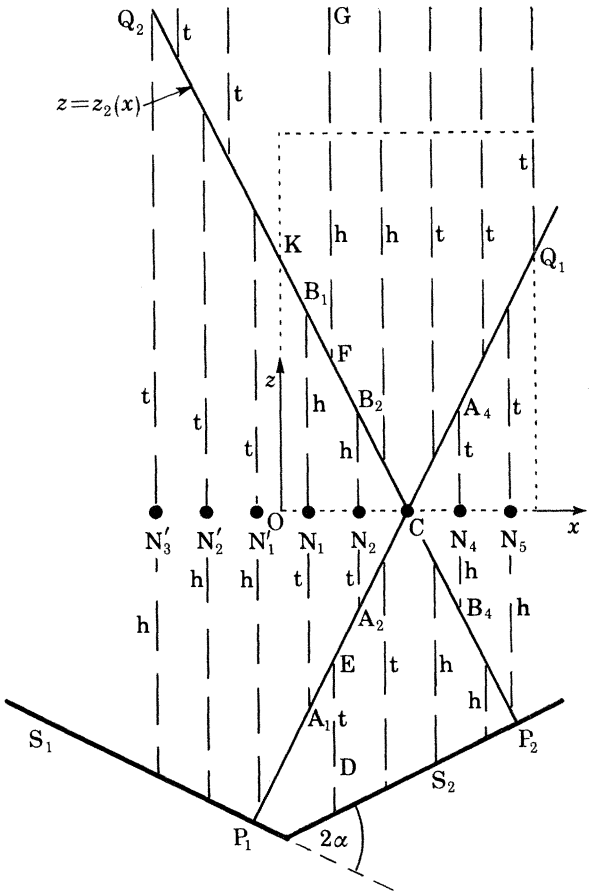

FIGURE 1. Two-beam model. The sources S_1, S_2 emit plane-wave pulses whose strength varies linearly parallel to the wavefronts. The amplitude is zero on the 'null normals' P_1Q_1, P_2Q_2. Dislocations move along the vertical trajectories (dashed lines) through the c.w. nulls ..., N_2', N_1', N_1, N_2, ... The dislocations may be in the head (h) or the tail (t) of the pulse.

Two infinite sources of plane waves S_1 and S_2 are set at an angle 2α to one another (figure 1) in a homogeneous non-dispersive medium. To start with they emit continuous waves of angular frequency ω_0, in phase with one another. If the plane waves were uniform in amplitude and the two sources were equal in strength, they would produce a set of vertical interference fringes through the points $\ldots N_2', N_1', N_1, N_2, \ldots$, the amplitude along the fringes being exactly zero. In a general two-dimensional diffraction pattern the zeros are not lines but points. This is because both the real and the imaginary parts of the complex wave amplitude have to vanish; each of these conditions separately defines a line, and the lines intersect in points. To remove this degeneracy in our model we impose spatial modulation on S_1 and S_2: the amplitude of each plane wave is made to vary linearly parallel to its wavefronts. Modulation parallel to the wavefronts

would normally spread out by diffraction, but if it is merely linear the wave propagates without change. Accordingly, we arrange the amplitude to be zero at P_1 and P_2 and to *increase* linearly to the left at the same rate for each source. The wave normals through P_1 and P_2 intersect at C, and it can be seen that the locus of points where the two interfering continuous waves have the same amplitude is the horizontal line through C. Destructive interference will only be complete at points on this line, and the result is that the vertical interference fringes collapse into point nulls at $\ldots N_2', N_1', N_1, N_2 \ldots$. The degeneracy has been removed.

The sources S_1 and S_2, amplitude modulated parallel to their lengths as just described, are now made to emit simultaneously the pulse represented by the complex function

$$\psi_0(t) = f(t)\, e^{-i\omega_0 t}, \tag{1}$$

where $f(t)$ is an envelope function. In this section we shall assume, for simplicity, that $f(t)$ is real and positive definite but in later sections we shall allow $f(t)$ to be more general. The actual oscillation is either the real or the imaginary part of $\psi_0(t)$, or some linear combination of the two. We know that there will now be a field of moving dislocation points (Nye & Berry 1974). Take axes Ox, Oz (figure 1). At a given instant t the dislocations are at the points $\boldsymbol{r} \equiv (x, z)$ where both the real and imaginary parts of the received signal $\psi(\boldsymbol{r}, t)$ are zero. Alternatively stated, they are at the points $\boldsymbol{r}$ where, at a given instant t, the wave amplitude $|\psi(\boldsymbol{r}, t)|$ vanishes and hence the phase is indeterminate. The moving dislocation points trace out lines in the two-dimensional observation space of $\boldsymbol{r}$, which are the dislocation trajectories. In the limit as $f(t)$ becomes infinitely long, so that $\psi_0(t)$ is a continuous oscillation, these linear trajectories must in some way condense on to the continuous-wave null points (called hereafter c.w. nulls). One of our tasks is to examine precisely how this happens.

First we find the trajectories. The complex wave function at an arbitrary point (x, z) may be written as

$$\psi(x, z, t) = \frac{1}{2}\left(1 - \frac{x}{x_1} + \frac{z}{z_1}\right)\psi_0(t - \tau_1) + \frac{1}{2}\left(1 - \frac{x}{x_1} - \frac{z}{z_1}\right)\psi_0(t - \tau_2), \tag{2}$$

where $x_1 = OC$, $z_1 = OK$, and τ_1, τ_2 are the travel times from S_1 and S_2 to (x, z). The linear modulating factors are each taken as $\frac{1}{2}$ at O. Writing

$$\bar{\tau} = \tfrac{1}{2}(\tau_1 + \tau_2), \quad \Delta\tau = \tfrac{1}{2}(\tau_1 - \tau_2), \tag{3}$$

and using (1) we have

$$\psi(x, z, t) = \tfrac{1}{2} e^{-i\omega_0(t-\bar{\tau})}\left\{\left(1 - \frac{x}{x_1} + \frac{z}{z_1}\right) f(t - \bar{\tau} - \Delta\tau)\, e^{i\omega_0 \Delta\tau} + \left(1 - \frac{x}{x_1} - \frac{z}{z_1}\right) f(t - \bar{\tau} + \Delta\tau)\, e^{-i\omega_0 \Delta\tau}\right\}. \tag{4}$$

In terms of position in the field $\Delta\tau$ is given by

$$\Delta\tau = (x \sin\alpha)/c, \tag{5}$$

and the mean time delay or travel time $\bar{\tau}$ of the two separate pulses is

$$\bar{\tau} \equiv \bar{\tau}(z) = \bar{\tau}_0 + (z \cos\alpha)/c, \tag{6}$$

where $\bar{\tau}_0$ is the mean time delay at O and c is the wave velocity. The 'null normals' P_1Q_1, P_2Q_2 divide the space into four sectors according to the signs of the two linear modulations. For the moment we look only at the sectors to the right and left of C where the two signs are the same.

Then the condition for a dislocation, namely $\psi(x, z, t) = 0$, gives

$$\omega_0 \Delta\tau = (n + \tfrac{1}{2})\pi \quad (n = \text{an integer}) \tag{7}$$

and

$$\left(1 - \frac{x}{x_1} + \frac{z}{z_1}\right) f(t - \bar{\tau} - \Delta\tau) = \left(1 - \frac{x}{x_1} - \frac{z}{z_1}\right) f(t - \bar{\tau} + \Delta\tau), \tag{8}$$

assuming $f(t) > 0$. Equation (7) shows that, for a dislocation, the two carrier waves must interfere destructively, while (8) shows that, in addition, the envelopes of the two pulses must be of equal height (figure 2*a*) at the time and place in question. The actual signal at a point where (7) is satisfied is shown in the upper diagram of figure 2*b*. Equation (7), which is independent of t

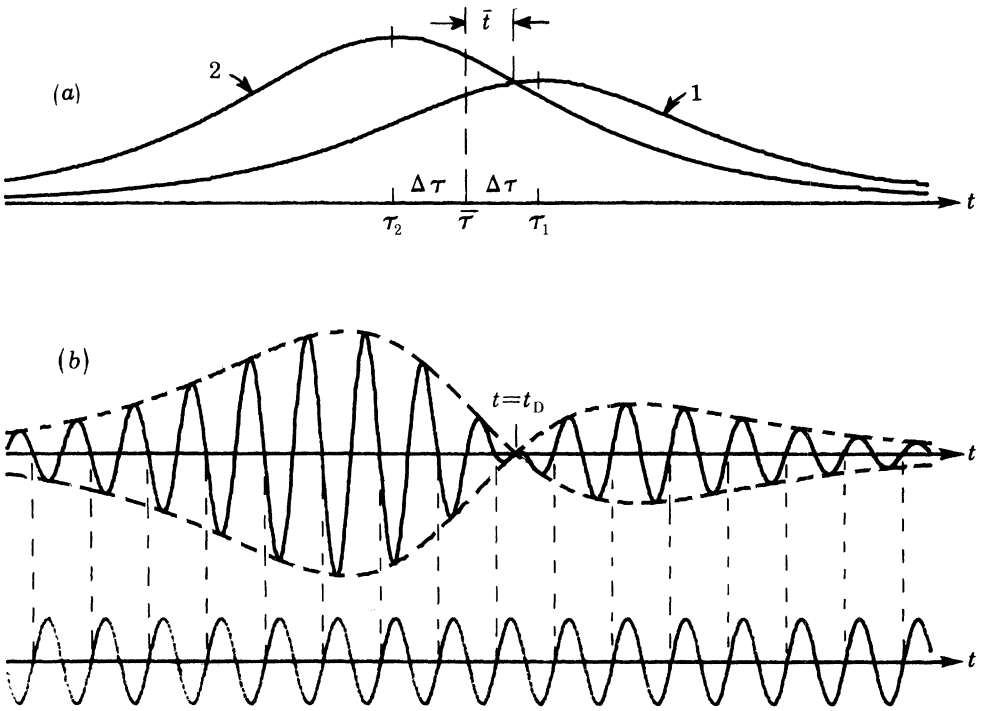

FIGURE 2. Interference of two pulses. (*a*) The dislocation occurs at the time when the envelopes of the two pulses are of equal height. (*b*) The combined signal, Re $\psi(t)$, contains a dislocation at $t = \bar{\tau} + \bar{t} = t_D$. The lower diagram shows the wave that would be received if the sources were unmodulated. The phase difference between the modulated and unmodulated waves is 0 or π.

(and also of the spatial modulations, apart from their relative signs), shows that the dislocation trajectories are simply vertical lines through the c.w. nulls $\ldots N_2', N_1', N_1, N_2 \ldots$, as indicated by the dashed lines in figure 1 terminating on P_1Q_1 and P_2Q_2. Equation (8) then determines the time t of the dislocation. In fact, a point that satisfies (7) lies on a trajectory only if (8) gives a real finite time t for the dislocation. We shall later find as a consequence that, depending on the pulse shape, only parts of the dashed lines shown in figure 1 are actually used as trajectories.

It follows from (4) and (7) that on a trajectory the signal can be expressed as

$$\psi(x, z, t) = \mathscr{R}(x, z, t) \exp \mathrm{i}\{\omega_0 \bar{\tau}(z) - \omega_0 t + \tfrac{1}{2}\pi\}, \tag{9}$$

where $\mathscr{R}(x, z, t)$ is a real envelope function. Thus $\omega_0 \bar{\tau}(z)$, which varies linearly with z by (6), measures the spatial variation of the phase of the carrier wave within the envelope. The time

$t = \bar{\tau}(z)$ can thus be thought of as marking the time of a particular phase of the carrier wave as the receiver moves upwards.

It is convenient to measure the dislocation time t relative to $\bar{\tau}$, defining

$$\bar{t} = t - \bar{\tau}, \tag{10}$$

as the relative delay of the dislocation. Now $\bar{t}$ varies as the receiver moves upwards, and therefore the dislocation moves relative to the carrier wave; in the nomenclature of Nye & Berry (1974) and Nye (1981), which is taken from crystal dislocation theory, the dislocation glides.

Equation (8) for the dislocation time can be expressed more conveniently as

$$(z_2+z)\, f(\bar{t}-\Delta\tau) = (z_2-z)\, f(\bar{t}+\Delta\tau) \quad (|z| < |z_2|), \tag{11}$$

where $z_2 \equiv z_2(x) = z_1(x_1-x)/x_1$ is simply the z-coordinate of the end of the trajectory that lies on P_2CQ_2: for instance B_1 for the trajectory A_1B_1, and B_4 for A_4B_4. Thus the value of z_2 depends on the particular trajectory being considered. Solving for z gives

$$\frac{z}{z_2} = \frac{f(\bar{t}+\Delta\tau) - f(\bar{t}-\Delta\tau)}{f(\bar{t}+\Delta\tau) + f(\bar{t}-\Delta\tau)} \quad (|z| < |z_2|). \tag{12}$$

The rate of sideways linear modulation of the two waves, provided it is non-zero, does not affect the position or time of the dislocation; it does affect the steepness of the amplitude well in which the dislocation sits.

We now derive an expression for the velocity of the gliding dislocation as it passes through a c.w. null point in this model, assuming the pulse envelope $f(t)$ to be analytic and peaked (that is decreasing monotonically from its non-degenerate maximum at $t = 0$). At a c.w. null the two constituent pulses are of equal strength and figure 2*a* shows that, provided $\Delta\tau$ is small compared with the pulse length, the relevant parts of the pulse envelope are close to the pulse maximum. We therefore approximate $f(t)$ by

$$f(t) = f(0) + \tfrac{1}{2} f''(0)\, t^2 + \ldots, \tag{13}$$

substitute this form in (12) and solve for $\bar{t}$. Thus

$$\bar{t} \approx \frac{z f(0)}{z_2 f''(0)\, \Delta\tau} \quad (\Delta\tau \text{ small}). \tag{14}$$

Then differentiating with respect to z gives

$$\left(\frac{\mathrm{d}\bar{t}}{\mathrm{d}z}\right)_{z=0} \approx \frac{f(0)}{z_2 f''(0)\, \Delta\tau} \quad (\Delta\tau \text{ small}). \tag{15}$$

We see that the glide reverses sign with z_2 and with $\Delta\tau$; $f(0)$ is positive, while $f''(0)$ is negative. Thus between O and C in figure 1, where $\Delta\tau$ and z_2 are both positive, the dislocation glides from tail to head, while to the left of O and to the right of C it glides from head to tail. The physical reason is (see figure 2*a*) that when the stronger of the two signals arrives first the intersection of the envelopes, and therefore the dislocation, is in the tail. Figure 1 is labelled accordingly.

The actual velocity v of the dislocation is given by

$$\frac{1}{v} = \frac{\mathrm{d}t}{\mathrm{d}z} = \frac{\mathrm{d}\bar{\tau}}{\mathrm{d}z} + \frac{\mathrm{d}\bar{t}}{\mathrm{d}z} = \frac{\cos\alpha}{c} + \frac{f(0)}{z_2 f''(0)\, \Delta\tau} \quad (\Delta\tau \text{ small}) \tag{16}$$

with the use of (6), where the first of the two terms comes from the motion of the carrier wave. For displaying the relative motion of the carrier wave and the dislocation, slowness (reciprocal velocity) is the natural quantity rather than velocity (Nye 1981).

In describing dislocation glide one must be aware of a trap. If the last term in (16) is sufficiently negative, v can also be negative. Figure 3 shows the track in (z, t) of a dislocation with negative v. We can think of this dislocation in two ways. In a series of time displays like the upper diagram of figure 2*b*, such as one might have on an oscilloscope, as z increases the dislocation moves from the tail of the pulse to the head. However, in space, a series of snapshots with increasing t would show the dislocation passing from head to tail. In describing glide we must take care to specify whether we are using the time or the space picture.

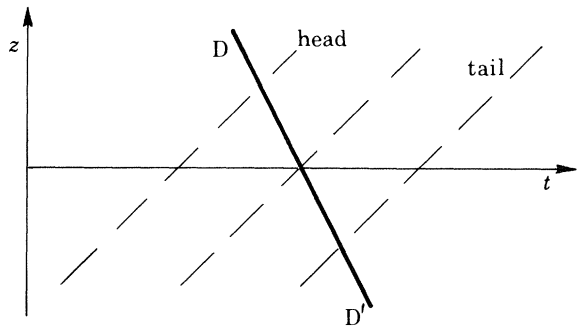

FIGURE 3. The line DD′ is the track in (z, t) of a dislocation moving with negative velocity v.

We have not yet considered possible dislocation trajectories in the two sectors above and below C, where the signs of the two linear modulations are opposite. Equations (7) and (11) are now replaced by

$$\omega_0 \Delta\tau = n\pi \quad (n = \text{an integer}) \tag{17}$$

and

$$(z_2+z)f(\bar{t}-\Delta\tau)+(z_2-z)f(\bar{t}+\Delta\tau) = 0 \quad (|z| > |z_2|). \tag{18}$$

The trajectories defined by (17) interleave the previous ones, as shown in figure 1. Equation (18) can be solved for z to give

$$\frac{z}{z_2} = \frac{f(\bar{t}+\Delta\tau)+f(\bar{t}-\Delta\tau)}{f(\bar{t}+\Delta\tau)-f(\bar{t}-\Delta\tau)} \quad (|z| > |z_2|). \tag{19}$$

Note that (18) has no solutions for $\bar{t}$ when $\Delta\tau = 0$. Therefore the z-axis ($n = 0$) is never used as a trajectory *for any positive definite pulse envelope* (but see the discussion of intrinsic dislocations in §3.2).

2.1. *Hyperbolic secant pulse envelope*

Some results turn out to be sensitive to the form of the envelope function $f(t)$, particularly its rates of rise and decay, and so we need to make a realistic choice. A Gaussian pulse, for example, would not be a good choice, because an experimental pulse is unlikely to have a tail that decays as fast as $\exp(-t^2)$; it is more likely to decay as $\exp(-t)$. A simple symmetrical pulse with the latter behaviour is the hyperbolic secant

$$f(t) = \operatorname{sech} \sigma t \quad (\sigma > 0), \tag{20}$$

where σ^{-1} is a measure of the pulse length, and we shall adopt this as our principal example.

With this pulse shape, formula (15) becomes

$$\mathrm{d}\bar{t}/\mathrm{d}z = -1/z_2\sigma^2\Delta\tau. \tag{21}$$

The limit $\sigma \to 0$ is the continuous wave and we see that this gives $\mathrm{d}\bar{t}/\mathrm{d}z \to \infty$ and hence $\mathrm{d}z/\mathrm{d}t = 0$. The track in (z, t) (figure 3) becomes horizontal and the dislocation is permanently at $z = 0$.

We can now calculate in more detail the behaviour of the dislocations produced by a hyperbolic secant pulse in this two-beam model. The trajectories are parts of the sets of vertical lines shown in figure 1, and we must now use equations (12) and (19) to find the (z, t) relation. In the right and left sectors equation (12) with (20) gives

$$z/z_2 = -\tanh(\sigma\Delta\tau)\tanh(\sigma\bar{t}), \tag{22}$$

or equivalently

$$\bar{t} = -\frac{1}{\sigma}\operatorname{artanh}\left(\frac{z}{z_2\tanh\sigma\Delta\tau}\right), \tag{23}$$

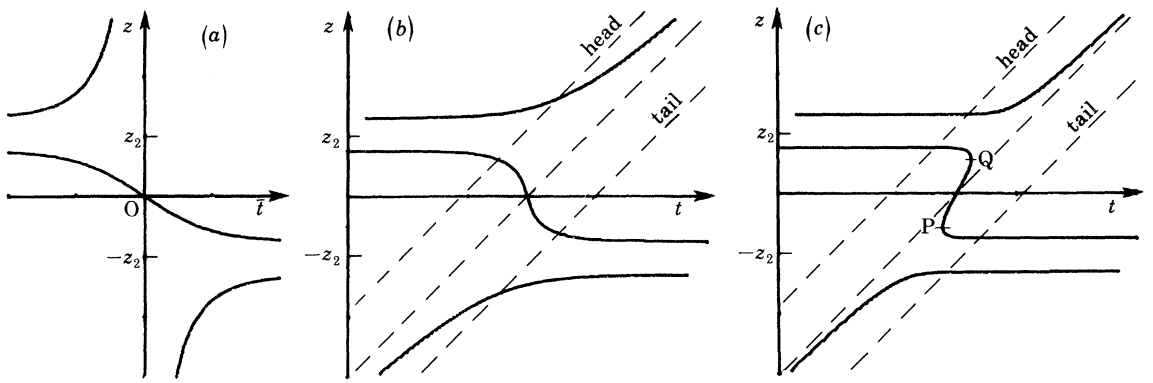

FIGURE 4. Two-beam model with a hyperbolic secant pulse envelope. (*a*) The $z, \bar{t}$-relation, for $\sigma\Delta\tau = 1$, $z_2 > 0$. (*b*) Worldlines of the dislocations. (*c*) The same as (*b*) but with $\mathrm{d}t/\mathrm{d}z$ positive at $z = 0$.

with $|z| < |z_2\tanh\sigma\Delta\tau|$. The last restriction is important; it means there are no solutions for $\bar{t}$ when z is outside the indicated range and that this range is less than the full range $|z| < |z_2|$ of the broken lines shown in figure 1. In other words, as we mentioned before, only a part of each of these lines, centred on the c.w. null, is actually used as a trajectory. The range of z is different for each trajectory because both z_2 and $\Delta\tau$ vary (linearly) with x. For large $|x|$, that is large $|\Delta\tau|$, the trajectories approach the full range $|z| < |z_2|$.

To study the interleaving trajectories in the upper and lower sectors we use equation (19) with (20) to give

$$z/z_2 = -\coth\sigma\Delta\tau\coth\sigma\bar{t} \tag{24}$$

or, solving for $\bar{t}$,

$$\bar{t} = -\frac{1}{\sigma}\operatorname{artanh}\left(\frac{z_2}{z\tanh\sigma\Delta\tau}\right), \tag{25}$$

with the restriction $|z| > |z_2\coth\sigma\Delta\tau|$. Again, only part of the full range $|z| > |z_2|$ is actually used by the dislocation as a trajectory.

For clarity of description we now choose z_2 and $\Delta\tau$ to be both positive (the range of x from O to C in figure 1). Figure 4*a* shows z as a function of $\bar{t}$ given by equations (22) and (24), and figures 4*b*,*c* show this converted to a relation between z and t according to the transformation, from (6) and (10),

$$t = \bar{\tau}_0 + (z\cos\alpha)/c + \bar{t}. \tag{26}$$

Figures 4*b*, *c* are obtained from figure 4*a* by shifting to the right and shearing parallel to the $\bar{t}$ axis. They differ in the sign of $\mathrm{d}t/\mathrm{d}z$ at $z = 0$; for using (21) we have

$$\left(\frac{\mathrm{d}t}{\mathrm{d}z}\right)_{z=0} = \frac{\cos\alpha}{c} - \frac{1}{z_2\,\sigma^2\Delta\tau}, \tag{27}$$

which can clearly be positive or negative.

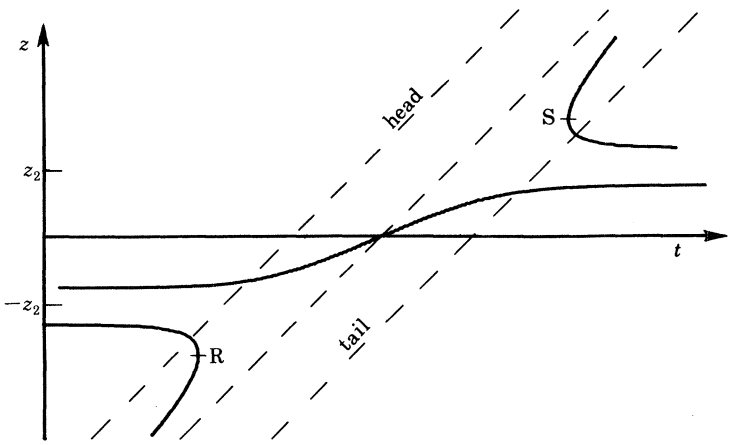

FIGURE 5. As for figure 4*b*, but with $\sigma\Delta\tau = -1$, $z_2 > 0$.

In figures 4*b*, *c* the central dashed line is the worldline of the reference time $t = \bar{\tau}(z)$. When $z \to 0$ or $z \to \pm\infty$ the two interfering pulses are of equal amplitude, and in this case, if the pulse envelopes are symmetric, $\bar{\tau}(z)$ clearly defines the centre of the resulting (symmetric) composite pulse envelope. Therefore *in these limits* we can regard $\bar{\tau}(z)$ as the pulse centre time. (In the general case, however, it is not clear how to define the 'pulse centre', but see § 3.7, and note that $T_1 = \bar{\tau}$ in the above limits.) In relation to typical trajectories such as the relevant portions of DE, A_2B_2, FG in figure 1 we could read the dislocation worldlines in figure 4*b* as follows. As a receiver moves from infinity to E a dislocation seen on a time display (oscilloscope screen) in the centre of the pulse moves towards the tail and leaves at a point below E. As the receiver moves from A_2 to B_2 a dislocation enters the tail at a point above A_2, passes through the centre of the pulse when $z = 0$ and leaves through the head at a point below B_2. On FG a dislocation is seen to enter the head at a point above F and move asymptotically to the pulse centre.

If, on the other hand, we ask how the spatial pattern of the ascending composite pulse changes with time we read from figure 4*b* that there is a dislocation on DE near the pulse centre which lags behind and leaves the tail at a point below E. Meanwhile, at a point below B_2 a dislocation enters the head of the pulse, moves downwards and is left behind at a point above A_2. A further dislocation appears above F in the head of the pulse, and is ultimately caught up by the pulse centre.

The reading of figure 4*c* as a series of time displays is very similar to figure 4*b*, but the succession of spatial patterns represented is rather different. At P a pair of dislocations is born; one moves backwards to the tail while the other moves rapidly forwards to annihilate at Q another dislocation that entered through the head and moved backwards. One way of achieving figure 4*c*, and hence pair creation and annihilation, is to decrease the wave angle α. Then z_2 increases without limit and the first term in (27) dominates. The critical value of α is arctan $(M^{\frac{1}{2}}/L)$,

where M is the relative delay ($2\Delta\tau$) between the two pulses and L is the pulse length $2\sigma^{-1}$, both measured in carrier wave periods.

When z_2 and $\Delta\tau$ are of opposite sign, that is, the range of x in figure 1 is to the left of O and to the right of C, $\bar{t}$ changes sign and the dislocation worldlines in (z, t) are as in figure 5. In the changing spatial pattern there is now pair annihilation at R in the head and pair creation at S in the tail. These types of dislocation behaviour also occur in the more realistic piston radiator model (Wright 1977; Wright & Berry in preparation). As $\sigma \to 0$ (continuous wave limit) the central worldlines in figures 4*b* and 5 approach the t-axis and the upper and lower ones recede to infinity, leaving a single stationary dislocation at $z = 0$.

2.2. *Gaussian pulse envelope*

A similar analysis for the Gaussian pulse envelope

$$f(t) = \exp(-\tfrac{1}{2}\sigma^2 t^2) \tag{28}$$

(chosen to have the same-shaped peak up to quadratic terms as the hyperbolic secant in (20)) gives, in place of (22),

$$\bar{t} = -\frac{1}{\sigma^2\Delta\tau}\operatorname{artanh}\frac{z}{z_2} \quad (|z| < |z_2|) \tag{29}$$

and, in place of (25),

$$\bar{t} = -\frac{1}{\sigma^2\Delta\tau}\operatorname{artanh}\frac{z_2}{z} \quad (|z| > |z_2|). \tag{30}$$

Qualitatively the corresponding dislocation worldlines are the same as shown in figures 4*b*, *c* and 5 for the hyperbolic secant pulse envelope, but with the important difference that the asymptotes are exactly at the levels $z = \pm z_2$, so that the full ranges of the broken lines in figure 1 are used as trajectories. Note that the last statement is true for all non-zero σ, but that for $\sigma = 0$ the used parts of these lines suddenly collapse on to the c.w. nulls. We have already remarked that the hyperbolic secant pulse is more realistic than the Gaussian, whose rise and fall is unphysically rapid; we now see that the dislocations produced by the hyperbolic secant pulse display a smoother approach to the continuous wave limit.

2.3. *Lorentzian pulse envelope*

Whereas the Gaussian envelope decays faster than the hyperbolic secant, the Lorentzian envelope

$$f(t) = (1 + \tfrac{1}{2}\sigma^2 t^2)^{-1}, \tag{31}$$

again chosen to give the same shape of peak up to quadratic terms, decays more slowly. Its dislocations exhibit qualitatively new features, as we now show.

Equations (12) and (19), with (31), result in

$$\zeta = -2(\sigma\Delta\tau)(\sigma\bar{t})/\{2 + (\sigma\Delta\tau)^2 + (\sigma\bar{t})^2\}, \tag{32}$$

where

$$\zeta = \begin{cases} z/z_2 & \text{if } |z| < |z_2| \\ z_2/z & \text{if } |z| > |z_2|, \end{cases} \tag{33}$$

as shown in figure 6*a*. Shearing these curves parallel to the $\bar{t}$ axis gives the dislocation worldlines of figure 6*b*. As before, $\mathrm{d}t/\mathrm{d}z$ at $z = 0$ can be positive or negative. Figure 6*c* shows the worldlines for the regions where $\Delta\tau$ and z_2 are opposite in sign (to the left of O and to the right of C in figure 1).

There are two new features in figures 6*b*, *c*. One is that the worldline for the central dislocation now begins and ends at $z = 0$. The other is the presence of points where the tangent is horizontal, $\mathrm{d}z/\mathrm{d}t = 0$ (previously we found points where it was vertical). This means that pair creation and annihilation is now seen in the time (oscilloscope) display as the receiver is moved vertically, as well as in the changing spatial pattern of the upward-moving composite wave train. The reason is simple. Figure 7 shows two Lorentzian pulse envelopes of equal length with their delays and relative strengths arranged so that they touch. This would represent the moment of pair creation. It is impossible to arrange either Gaussian or hyperbolic secant pulse envelopes (of equal pulse length) in this way.

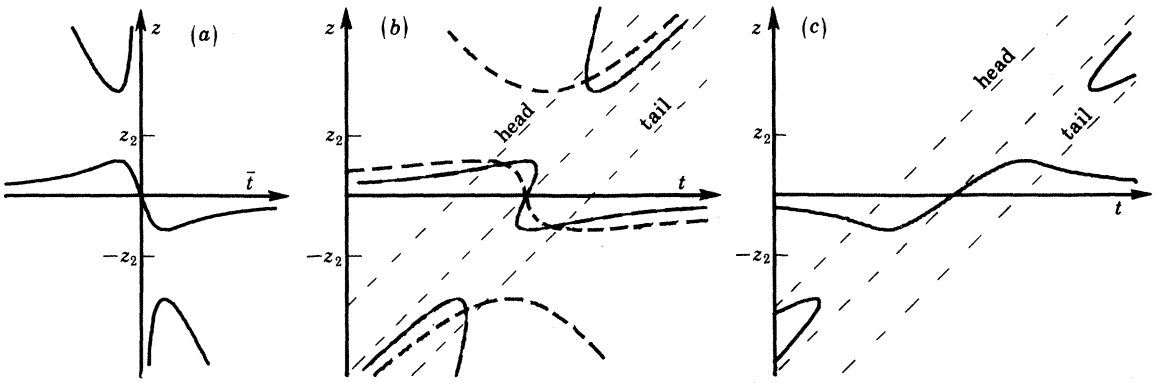

FIGURE 6. Two-beam model with a Lorentzian pulse envelope. (*a*) The $z,\bar{t}$-relation, for $\sigma\Delta\tau = 1$, $z_2 > 0$. (*b*) World-lines of the dislocations: $\mathrm{d}t/\mathrm{d}z$ at $z = 0$ can be positive (solid curves) or negative (dashed curves). (*c*) World-lines for $\sigma\Delta\tau = -1$, $z_2 > 0$.

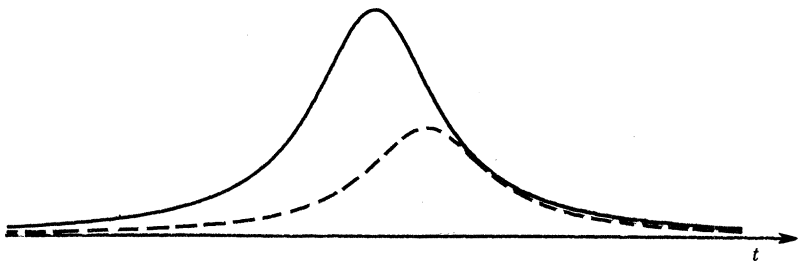

FIGURE 7. Two Lorentzian pulse envelopes of equal length arranged so that they touch.

Thus two distinct types of pair creation may be recognized during the glide of dislocations: the first corresponds to the worldline being vertical (the type described by Nye & Berry (1974)) and the second, which is new, to the worldline being horizontal.

2.4. *Summary of results for pulses*

We learn from these examples that in the two-beam model the pulse shape does not affect the positions of the dislocation trajectories but it does affect their length and the way they are traversed in time. With the realistic hyperbolic secant pulse the trajectories are traversed monotonically with time, but with the less realistic Lorentzian pulse the dislocations come momentarily to rest at the end of a trajectory and reverse direction. An important result is that, as $\sigma \to 0$, the dislocations in hyperbolic secant and Lorentzian pulses become progressively more localized in the neighbourhoods of the c.w. nulls. Dislocations in Gaussian pulses do not: the

length of the trajectory remains the same right up to the limit $\sigma = 0$, when it suddenly collapses to zero; but this behaviour would be hard to observe because the dislocations far from a c.w. null would be in the far head or tail of the pulse and liable to be lost in the noise.

2.5. *Pair creation in general*

How general are the two types of pair creation shown by the Lorentzian pulse? Stepping aside for the moment from the two-beam model, if we consider the general behaviour of point dislocations in the two space dimensions (x, z) we note first that the worldline in (x, z, t) is a continuous smooth line. Generically there would be special points on the line where it touches a plane $t =$ constant; these points represent pair creation of the first kind. But generically there will be no point on the line where the tangent points exactly along the t axis, and therefore no point on the line where, even locally, the behaviour is that corresponding to the second kind of pair creation. The reason it occurs in the two-beam model is that here the dislocation worldlines in (x, z, t) lie entirely in (z, t) planes (glide), with no component of motion parallel to Ox (climb). Climb would remove the second kind of pair creation.

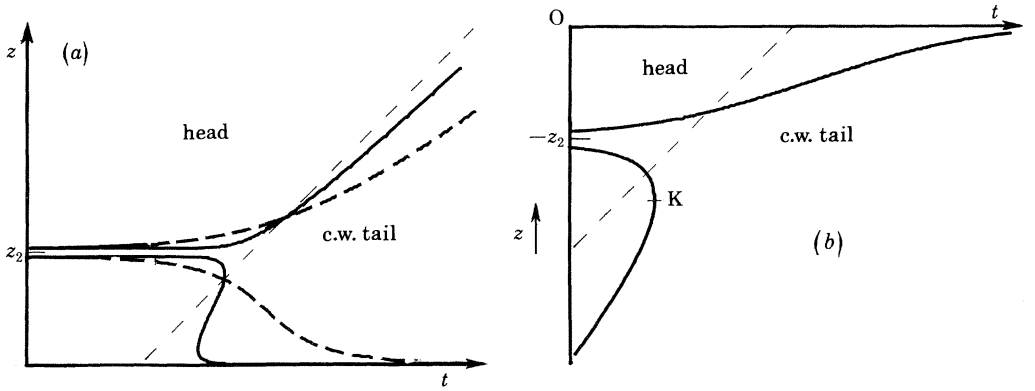

FIGURE 8. Two-beam model with a 'smooth-step' envelope function, showing how the dislocations reach their final positions at the c.w. nulls ($z = 0$). (*a*) Worldlines for $\Delta\tau > 0$, $z_2 > 0$: with pair creation and annihilation (solid curve) and without (dashed curve). (*b*) Worldlines for $\Delta\tau < 0$, $z_2 > 0$.

2.6. *Establishment of the continuous-wave nulls*

If a continuous wave is 'switched on', where do the nulls that are eventually stationary come from? We can study an example in the two-beam model by taking as the envelope the 'smooth-step' function

$$f(t) = \tfrac{1}{2}(1 + \tanh \sigma t), \tag{34}$$

which rises from near 0 to near 1 in a time of order σ^{-1}. Then equations (12) and (19) give

$$\zeta = \frac{\tanh \sigma(\bar{t} + \Delta\tau) - \tanh \sigma(\bar{t} - \Delta\tau)}{2 + \tanh \sigma(\bar{t} + \Delta\tau) + \tanh \sigma(\bar{t} - \Delta\tau)}, \tag{35}$$

where, as before,

$$\zeta = \begin{cases} z/z_2 & \text{if } |z| < |z_2| \\ z_2/z & \text{if } |z| > |z_2|, \end{cases}$$

and when this is combined with equation (26) we find the dislocation worldlines shown in figures 8*a*, *b*.

Only part of each dashed line in figure 1 is used as a trajectory: qualitatively, those parts labelled 'head'. In all cases, with increasing t, dislocations that finally become deposited at the c.w. nulls enter through the head of the step. But the process may be complicated by pair creation (solid curve in figure 8*a*). In figure 8*a* it is accompanied by another dislocation which enters through the head and moves to infinity, and in figure 8*b* by two dislocations in the lower part of the field which mutually annihilate at K. The details must depend on the shape of the step but because the hyperbolic tangent approaches its asymptote in a simple exponential way this example is thought to be quite realistic.

3. General theory

3.1. *The physical system*

We now present a more general theory. As before, we represent the original oscillation by the complex scalar $\psi_0(t)$. Thus we suppose that a signal generator S_0 produces an oscillation that is either the real or the imaginary part of $\psi_0(t)$, or some linear combination of the two. The scalar $\psi_0(t)$ can represent either a continuous oscillation or a pulse. We suppose that all processes are linear. After various time delays and linear filterings, signals from S_0 are radiated from sources $S_1, S_2, \ldots, S_n$ into a general medium. The waves are scattered at stationary scattering objects, and the diffraction pattern is observed by placing a movable receiver at the point P given by the position vector $\boldsymbol{r}$. After a further delay and linear filtering the 'received signal' $\psi(\boldsymbol{r}, t)$ is displayed on an oscilloscope. The receiver at P need not be perfect: in general, it will have a polar diagram for reception, a certain size and a certain bandwidth. But we make sure that these characteristics depend only on the position of the receiver: for example, when at P it must point in a fixed direction, but this direction can vary as P is changed.

This system, which is linear throughout, is general enough to represent either a scattering experiment with a single source or an interference experiment that uses several separate transmitting transducers all supplied by a single pulse generator. It corresponds closely to what can actually be done in practice, for example in an experiment with ultrasonic pulses.

3.2. *Pulses as perturbed continuous waves: dislocations*

It is always possible, in principle, to deduce the response of a linear system when driven by a pulse from a knowledge of its continuous wave response, via a Fourier transform. To examine particular features of the pulse response, such as its dislocation structure, one needs to evaluate the integral representation derived from the Fourier transform, which is generally not possible in a simple closed form. Further, it requires knowledge of the continuous wave response for all frequencies, which in practice is not usually available.

However, if we restrict ourselves to pulses with spectra concentrated about their centre (carrier) frequency, it should be possible to deduce the pulse behaviour approximately from the continuous wave response within the pulse bandwidth. We now do this by developing a perturbation expansion in σ, the pulse bandwidth, and then solving for the phenomenon of interest, here the dislocations, for small σ. (Real pulses, produced by resonant systems, tend to have small values of σ; for example $\sigma/\omega_0 \approx 10^{-2}$ would not be unusual.) The specific objective in §§ 3.2–3.6 is to evolve a workable numerical method of computing the dislocation trajectories and times from information about the continuous wave diffraction pattern at frequencies near the centre frequency. However, we keep the formulation as general as possible with a view to its application

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

to other pulse phenomena, such as tracing the course of a pulse maximum (which might have more immediate relevance to current remote sensing methods).

The Fourier transforms of the original oscillation $\psi_0(t)$ and of the received signal $\psi(\boldsymbol{r}, t)$, namely

$$\overline{\psi}_0(\omega) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \psi_0(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t, \qquad \overline{\psi}(\boldsymbol{r}, \omega) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \psi(\boldsymbol{r}, t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t, \tag{36}$$

are related by the transfer function for the whole system $a(\boldsymbol{r}, \omega)$:

$$\overline{\psi}(\boldsymbol{r}, \omega) = a(\boldsymbol{r}, \omega)\, \overline{\psi}_0(\omega). \tag{37}$$

The transfer function $a(\boldsymbol{r}, \omega)$ defined in this way evidently includes not only the characteristics of the diffraction pattern in space, but also those of the receiver. If, however, the receiver were perfect, the function $a(\boldsymbol{r}, \omega)$ would constitute complete information about the diffraction pattern, and we shall think of it in that way. For each frequency ω, $a(\boldsymbol{r}, \omega)$ then represents a spatial variation of complex amplitude, which is, in fact, the monochromatic diffraction pattern for that particular frequency of radiation. Thus $a(\boldsymbol{r}, \omega)$ represents the family of monochromatic diffraction patterns for all frequencies.

A very general class of linear systems can be represented by a transfer function of the form $a(\boldsymbol{r}, \omega)$, which can include inhomogeneity and dispersion. Only singularities of the transfer function in the finite complex frequency plane affect the present general theory: we assume these singularities to be poles. Causality requires the poles to lie only in the lower half ω-plane (see for example Pippard 1978). If a system is excited transiently, such poles will produce damped sinusoidal responses which are largely independent of the form of the drive. A pole at complex frequency produces a maximum, or resonance, in the real-frequency response, associated with a natural frequency of the system. (However, the converse is not true; a frequency maximum of the transfer function need not be associated with any singularity in the finite ω-plane, as shown by the two-beam model, and by the even simpler example $a(\omega) \equiv \cos \omega\tau$, where τ is a constant.) We shall see that the assumption that $a(\boldsymbol{r}, \omega)$ is an entire function of ω (that is, has no singularities in the entire finite complex ω-plane) leads to a convergent, rather than asymptotic, perturbation expansion.

Resonances in a medium usually lead to absorption, that is, minima in the transfer function. Such resonances are perhaps more aptly called 'anti-resonances' from our point of view. They may be associated with complex zeros, but will not affect the analyticity of $a(\boldsymbol{r}, \omega)$. However, practical transducers are usually designed to resonate near the carrier frequency, and these resonances will normally introduce poles into $a(\boldsymbol{r}, \omega)$. We may still be able to model such a system by an entire transfer function driven by a suitably modified pulse, or restrict attention to that part of the system (for example, just the medium) which has an entire transfer function, and in what follows we shall mainly assume this to be the case.

Let the original oscillation now be of the form (1), namely

$$\psi_0(t) = f(t)\, \mathrm{e}^{-\mathrm{i}\omega_0 t}. \tag{1}$$

Subject only to the existence of its Fourier transform, we initially allow a completely general complex $f(t)$, which describes both amplitude and phase modulation. Then

$$\overline{\psi}_0(\omega) = \bar{f}(\omega - \omega_0), \tag{38}$$

where $\bar{f}(\omega)$ is the Fourier transform of $f(t)$. Hence the received signal is

$$\psi(\boldsymbol{r}, t) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \bar{\psi}(\boldsymbol{r}, \omega)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}\omega$$

$$= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} a(\boldsymbol{r}, \omega)\, \bar{f}(\omega - \omega_0)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}\omega \tag{39}$$

with the use of (37) and (38).

We shall assume that the envelope function $f(t)$ is slowly varying; so $\bar{f}(\omega - \omega_0)$ will be concentrated around $\omega = \omega_0$ and only values of the integrand for ω near ω_0 will contribute significantly to this integral. Then, if we make a Taylor expansion of $a(\boldsymbol{r}, \omega)$ about ω_0, the resulting series representation for $\psi(\boldsymbol{r}, t)$ should converge rapidly. In this way we can expect to deduce approximately the received pulse pattern and, in particular, its dislocation structure, from knowledge of the monochromatic diffraction patterns for frequencies at and near ω_0.

The Fourier transforms (36) are invariant under shifts in time; so we are free to choose a *local time origin* $T(\boldsymbol{r})$ when we invert $\bar{\psi}(\boldsymbol{r}, \omega)$. Accordingly we define a local time $\tilde{t}(\boldsymbol{r}) \equiv t - T(\boldsymbol{r})$, and write (39) as

$$\psi(\boldsymbol{r}, t) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \{a(\boldsymbol{r}, \omega)\, \mathrm{e}^{-\mathrm{i}\omega T(\boldsymbol{r})}\}\, \bar{f}(\omega - \omega_0)\, \mathrm{e}^{-\mathrm{i}\omega \tilde{t}(\boldsymbol{r})}\, \mathrm{d}\omega. \tag{40}$$

In effect, we have replaced the original transfer function $a(\boldsymbol{r}, \omega)$ by the delayed transfer function

$$A_T(\boldsymbol{r}, \omega) \equiv a(\boldsymbol{r}, \omega)\, \mathrm{e}^{-\mathrm{i}\omega T(\boldsymbol{r})}, \tag{41}$$

where the subscript T is to remind us of the functional dependence of A on $T(\boldsymbol{r})$.

At first sight we appear to have made the problem more complicated, but the use of this local time origin turns out to be the crux of the theory, and the choice of $T(\boldsymbol{r})$ its most subtle facet. It is a central difficulty that there is no absolute way of determining $T(\boldsymbol{r})$. No *exact* results can depend on $T(\boldsymbol{r})$, but its choice affects the initial rate of convergence of the perturbation series.

We shall find that the best choice for $T(\boldsymbol{r})$ depends on where we decide to truncate the perturbation series. In the lowest-order approximation, appropriate only for very long pulses, where we keep only the first term of the series, the optimal choice for $T(\boldsymbol{r})$, denoted by $T_0(\boldsymbol{r})$, has a simple physical significance which we now describe.

First consider the trivial example where a pulse is emitted and then received without change at $\boldsymbol{r}$, after a delay $\tau(\boldsymbol{r})$. We have

$$\psi(\boldsymbol{r}, t) = \psi_0\{t - \tau(\boldsymbol{r})\}, \quad \bar{\psi}(\boldsymbol{r}, \omega) = \mathrm{e}^{\mathrm{i}\omega\tau(\boldsymbol{r})}\, \bar{\psi}_0(\omega) \tag{42}$$

so that

$$a(\boldsymbol{r}, \omega) = \mathrm{e}^{\mathrm{i}\omega\tau(\boldsymbol{r})}.$$

Then, if $T_0(\boldsymbol{r})$ is chosen to be $\tau(\boldsymbol{r})$, we note that $A_T(\boldsymbol{r}, \omega) \equiv 1$.

More generally, in equation (40), if $\bar{f}$ is sufficiently concentrated, that is, if the pulse is long enough, a good approximation should be to replace $A_T(\boldsymbol{r}, \omega)$ by $A_T(\boldsymbol{r}, \omega_0)$, provided also that $A_T(\boldsymbol{r}, \omega)$ varies sufficiently slowly with ω. The function $A_T(\boldsymbol{r}, \omega)$ will vary with ω both in amplitude and in phase. We cannot alter the rate of change of its amplitude by adjusting $T(\boldsymbol{r})$, but we can at least reduce to zero the rate of change of its phase at $\omega = \omega_0$. If $\gamma(\boldsymbol{r}, \omega)$ denotes the phase of $A_T(\boldsymbol{r}, \omega)$, and $\varphi(\boldsymbol{r}, \omega)$ denotes the phase of $a(\boldsymbol{r}, \omega)$, we have from (41)

$$\gamma = \varphi - \omega T \tag{43}$$

and hence

$$\left(\frac{\partial \gamma}{\partial \omega}\right)_{\omega=\omega_0} = \left(\frac{\partial \varphi}{\partial \omega}\right)_{\omega=\omega_0} - T. \tag{44}$$

Thus, by choosing $T_0 = (\partial\varphi/\partial\omega)_{\omega=\omega_0}$, we secure $(\partial\gamma/\partial\omega)_{\omega=\omega_0} = 0$, as we wanted, and we show later (§3.5.1) that this choice for T_0 is indeed optimal. In the simple example above it would give $T_0(\boldsymbol{r}) = \tau(\boldsymbol{r})$.

Having thus chosen $T_0(\boldsymbol{r})$ we put $A_T(\boldsymbol{r}, \omega) = A_T(\boldsymbol{r}, \omega_0)$ in (40) and invert the Fourier transform to obtain

$$\psi(\boldsymbol{r}, t) \approx A_T(\boldsymbol{r}, \omega_0)\, \mathrm{e}^{-\mathrm{i}\omega_0 \tilde{t}} f(\tilde{t}) = a(\boldsymbol{r}, \omega_0)\, \mathrm{e}^{-\mathrm{i}\omega_0 t} f(t - T_0(\boldsymbol{r})). \tag{45}$$

In this approximation the pulse waveform is simply the carrier wave $\mathrm{e}^{-\mathrm{i}\omega_0 t}$ propagated to the point $\boldsymbol{r}$ by the transfer function $a(\boldsymbol{r}, \omega_0)$ and modulated by the envelope function $f(t)$ delayed by $T_0(\boldsymbol{r})$. Therefore we shall call $T_0(\boldsymbol{r})$ the *pulse delay time*. Since the phase of $a(\boldsymbol{r}, \omega_0)$ is $\varphi(\boldsymbol{r}, \omega_0)$, the phase of the carrier wave is evidently

$$\varphi(\boldsymbol{r}, \omega_0) - \omega_0 t = -\omega_0\{t - \varphi(\boldsymbol{r}, \omega_0)/\omega_0\},$$

from which we recognize that *its* delay time is $\varphi(\boldsymbol{r}, \omega_0)/\omega_0$, in contrast to $T_0 = (\partial\varphi/\partial\omega)_{\omega=\omega_0}$ for the envelope.

In a general system, driven by a short pulse, the signal received at $\boldsymbol{r}$ will consist of a sequence of differently delayed pulses arriving from the different sources (which may be continuously distributed) and via different routes. As the pulse length is increased, the sequence of pulses will begin to merge. For a pulse consisting of a modulated carrier, as the pulse length tends to infinity, the envelope of the received signal will look like a single scaled copy of the envelope of the driving pulse, from (45), delayed by the time $T_0(\boldsymbol{r})$. Note that we let the pulse length tend to infinity, but not become infinite, so that we can still discern the shape of the envelope.

To illustrate the connection of this discussion with the group velocity phenomenon consider a plane wave travelling a distance x through a uniform dispersive medium with wave number $k(\omega)$. Then the transfer function is

$$a(x, \omega) \equiv \mathrm{e}^{\mathrm{i}k(\omega)x}.$$

From standard wave theory the propagation of quasimonochromatic plane wave pulses is described by the group velocity

$$v_{\mathrm{g}} \equiv (\mathrm{d}\omega/\mathrm{d}k)_{\omega=\omega_0}.$$

Then, if the driving pulse $\psi_0(t) \equiv f(t)\, \mathrm{e}^{-\mathrm{i}\omega_0 t}$ is quasimonochromatic, the signal received at x is

$$\begin{aligned} \psi(x, t) &= \mathrm{e}^{\mathrm{i}\{k(\omega_0)x - \omega_0 t\}} f(t - x/v_{\mathrm{g}}) \\ &= a(x, \omega_0)\, \mathrm{e}^{-\mathrm{i}\omega_0 t} f\{t - T_0(x)\}, \end{aligned} \tag{46}$$

where

$$T_0(x) = (\mathrm{d}k/\mathrm{d}\omega)_{\omega=\omega_0} x.$$

Equation (46) has precisely the form of our general result (45), and $T_0(x)$ is indeed the rate of change of the phase $\varphi = k(\omega)x$ of $a(x, \omega)$ at $\omega = \omega_0$. This prescription for T_0 clearly depends on ω_0 in general. Of course, if the medium is non-dispersive $T_0(x)$ reduces to x/c, where c is the wave velocity.

For non-planar waves which exhibit diffraction this simple prescription for the pulse delay time in terms of the group velocity ceases to be possible. The alternative prescription

$$T_0 = (\partial\varphi/\partial\omega)_{\omega=\omega_0}$$

constitutes a generalization of the idea of group velocity to much more general (and hence realistic) systems. It is well known in linear systems theory (see for example Faulkner 1969) as the 'group delay time', but the derivation is different. Let us note that we have yet to confirm our crude derivation of equation (45) and to justify the prescription for $T_0(\boldsymbol{r})$.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

However, the main purpose of this paper is to find wavefront dislocations, and it is helpful to notice that they can arise in two different ways. If the driving pulse envelope $f(t)$ has any zeros, these will propagate through the wavefield; we call them *intrinsic* dislocations since they are inherent in the driving pulse. Our primary interest in this paper is dislocations produced by the diffraction in the wavefield; we call these *extrinsic* dislocations. They occur even when the driving pulse itself has no zeros, and are what Nye & Berry (1974) had in mind when they coined the name 'wavefront dislocation'.

Intrinsic dislocations have not yet been studied. As a simple example, suppose the two-beam system of § 2 is driven by a pulse whose envelope $f(t)$ has a single zero at $t = a$. This produces a single intrinsic dislocation satisfying $\Delta\tau = 0$, $t - \bar{\tau} = a$,

$$\text{i.e. trajectory:} \quad x = 0,$$

$$\text{time:} \quad t = a + \bar{\tau}_0 + (z/c)\cos\alpha.$$

It is interesting that this trajectory along $x = 0$ is completely forbidden to extrinsic dislocations.

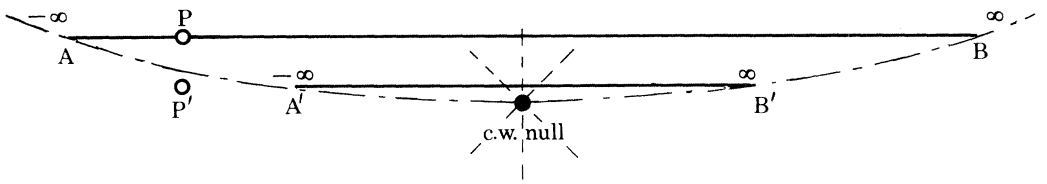

FIGURE 9. With a hyperbolic secant pulse shape, in the limit $\sigma \to 0$ a dislocation trajectory collapses on to a c.w. null. While the length of the trajectory is proportional to σ, its distance from the c.w. null is proportional to σ^2. The dislocation times run from $-\infty$ to $+\infty$. In the degenerate two-beam model the trajectories for different σ fall on top of one another, and always pass through the c.w. null.

Equation (45) predicts no extrinsic dislocations (other than c.w. nulls). To understand this let us consider, anticipating later results, what happens physically as the bandwidth σ of the driving pulse is decreased to zero (figure 9). For small non-zero σ and a symmetrical pulse envelope a typical trajectory surface AB avoids a c.w. null line by a distance proportional to σ^2. For a hyperbolic secant pulse the dislocation line sweeps out a strip of length AB proportional to σ, the time of arrival of the dislocation passing from $-\infty$ to $+\infty$. Thus, as σ decreases the trajectory surface migrates (vertically in figure 9) and simultaneously contracts in length (to A′B′); finally, as $\sigma \to 0$, it condenses on to the c.w. null line. (The two-beam model is degenerate in that the trajectories with $|z| < |z_2|$ always pass through the c.w. nulls.) Suppose then that we have fixed σ and have found the trajectory AB, labelled along its course with the arrival times t. If σ is now decreased a typical point P will no longer show a dislocation, because the trajectory will have shifted, but we could regain the dislocation, at a different time, by moving slightly perpendicular to the trajectory. Eventually, however, the dislocation will be lost in the far head or tail of the pulse ($t = \pm\infty$), and this will happen *before* σ reaches zero (in figure 9, P would have reached P′ when AB reached A′B′). Thus, for any point not on a c.w. null line, as σ is decreased the signal will eventually not contain a dislocation. This is why the lowest-order approximation (45), valid for very small σ, does not contain any extrinsic dislocations (other than the c.w. nulls). To find extrinsic dislocations we have to fix σ as non-zero and then seek an approximation to $\psi(\boldsymbol{r}, t)$ that is good enough to reveal the positions and times of the zeros.

This discussion illuminates another puzzling point. The spatial pattern of $\varphi(\boldsymbol{r}, \omega_0)$, the phase of $a(\boldsymbol{r}, \omega_0)$, near the c.w. null is indicated in figure 9 by the radial equiphase lines, the phase changing by 2π in one revolution about the null (Nye & Berry 1974; Nye 1981). As ω is changed the null point and the radial pattern move, and it is evident that, for a point $\boldsymbol{r}$ near the c.w. null, the rate of change $(\partial\varphi/\partial\omega)_{\omega=\omega_0}$, and hence T_0, can become very large. Since T_0 is the pulse delay time this is a strange conclusion, but it is nevertheless correct. For example, in the two-beam model it may be verified from first principles that on a dislocation trajectory, for fixed small z, as σ is decreased, first the dislocation disappears into the far head or tail of the pulse as described above, and then the signal envelope does indeed become a much attenuated replica of the original pulse envelope but displaced (backwards or forwards) to very large times. The time displacement can be made indefinitely large by choosing a smaller fixed z, provided σ is then allowed to be sufficiently small. (This means that it is quite possible for the peak of the received envelope to be observed *before* the peak of the source envelope has been transmitted. However, there is no reason to think that causality is violated. To discuss causality one would need a signal with a sharp onset, but in the present case the small bandwidth, on which the result depends, severely restricts how fast the source envelope can vary.) Of course, at the c.w. null itself, as $\sigma \to 0$, the whole signal becomes zero.

3.3. *The perturbation series*

All our analysis is in the time and frequency domains at fixed $\boldsymbol{r}$, so we shall usually drop all explicit $\boldsymbol{r}$ dependences. Defining a new frequency variable $W \equiv \omega - \omega_0$, and using the delayed transfer function defined in (41), our integral representation (40) for the received signal becomes

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega\tilde{t}}\,(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} A_T(\omega_0 + W)\, \bar{f}(W)\, \mathrm{e}^{-\mathrm{i}W\tilde{t}}\, \mathrm{d}W \tag{47}$$

where, as before, $t = T + \tilde{t}$, and T denotes simply an arbitrary local time origin. Expanding $A_T(\omega_0 + W)$ about $W = 0$ and integrating term by term gives formally

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\,(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \left\{ \sum_{n=0}^{\infty} \frac{W^n}{n!} A_{T,0}^{(n)} \right\} \bar{f}(W)\, \mathrm{e}^{-\mathrm{i}W\tilde{t}}\, \mathrm{d}W \tag{48}$$

$$= \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}} \sum_{n=0}^{\infty} \frac{\mathrm{i}^n}{n!} A_{T,0}^{(n)} f^{(n)}(\tilde{t}), \tag{49}$$

where we have used inverse Fourier transforms. Here

$$A_{T,0}^{(n)} \equiv \left. \frac{\partial^n A_T(\omega)}{\partial \omega^n} \right|_{\omega=\omega_0} \quad \text{and} \quad f^{(n)}(t) \equiv \frac{\mathrm{d}^n f(t)}{\mathrm{d}t^n}.$$

Subscript zero will always denote that functions of ω are evaluated at $\omega = \omega_0$.

We have already argued that, if the system has no 'singular resonances', $a(\boldsymbol{r}, \omega)$ is an entire function of ω, and hence so is $A_T(\boldsymbol{r}, \omega)$. Therefore, the Taylor expansion of $A_T(\boldsymbol{r}, \omega)$, as used in (48), is uniformly convergent in the whole complex ω-plane. On the real ω-axis $|\bar{f}(\omega)|$ is bounded; hence the expansion of the whole integrand of (48) is uniformly convergent for all real W. Therefore, integration term by term is justified, and (49) converges.

If, on the other hand, $a(\boldsymbol{r}, \omega)$ has any poles, the integrand of (48) will have a finite radius of convergence and (49) will be an asymptotic series, because we are integrating outside the radius of convergence (Dingle 1973, p. 3). The significance of this is that, typically, the terms of a

convergent series initially increase in magnitude and then ultimately decrease (converge) to zero, whereas the terms of an asymptotic series initially converge in magnitude to some minimum value and then ultimately diverge. (As an excellent introduction to asymptotic series the first chapter of Dingle (1973) is recommended.) In both cases a truncated series may accurately represent the infinite series only if it is truncated in the converging region. Therefore we must truncate a convergent series late enough, but an asymptotic series early enough. Since (49) is not a power series its behaviour may be more complicated, and it is difficult to check that it is converging at the point of truncation. We shall assume that it is. Then subsequent analysis is formally the same whether or not the series ultimately converges or diverges.

The perturbation series in equation (49) expresses the received wavefunction as an oscillatory factor $e^{-i\omega_0 \tilde{t}}$ multiplied by a weighted sum of the original pulse envelope f and its derivatives, all delayed by the time T. The time T is arbitrary: nevertheless equation (49) is an *exact* formal representation of the received wavefunction, which from its derivation is independent of T. This is possible because T has become intricately interwoven into the functional form of the infinite series, appearing in both the delayed transfer function and the envelope function,

$$A^{(n)}_{T,0} \equiv \frac{\partial^n}{\partial \omega^n}\{a(\omega)\, e^{-i\omega T}\}\Big|_{\omega_0}, \quad f^{(n)}(\tilde{t}) \equiv \frac{d^n}{dt^n} f(t-T)$$

respectively, and also in the oscillatory factor $e^{-i\omega_0(t-T)}$. (In Appendix A we show explicitly how these T variations within (49) conspire to cancel out.) The time T cannot affect the *value* of the infinite series, but it may affect the *rate* of convergence of the early terms, and hence the accuracy of a particular truncation.

For reference, we introduce the bandwidth σ explicitly by defining

$$f(t) \equiv F(\sigma t), \tag{50}$$

so that

$$f^{(n)}(t) \equiv \sigma^n F^{(n)}(\sigma t), \tag{51}$$

where $F^{(n)}$ denotes the nth derivative of F. This does not turn the perturbation series into a power series in σ, because of the σ-dependence remaining in F; but it does make explicit the sense in which successive derivatives of f become rapidly smaller in the quasimonochromatic limit $\sigma \to 0$. The $F^{(n)}$ facilitate comparison of derivatives of different orders, by virtue of being dimensionless. For notational simplicity we shall continue to work in terms of f, referring back to (51) when necessary.

There are two important parameters affecting the rate of convergence of (49): the bandwidth σ of the driving pulse and $T(\boldsymbol{r})$. They are completely independent, but while σ is fixed by the system we are free to choose T. For a particular σ, we choose T to maximize the rate of initial convergence of (49). The smaller is σ, the more rapid is this rate.

This raises the obvious question: on what scale must σ be small? For the series (49) representing $\psi(t)$ to converge rapidly, the ratio of successive terms must be small. From (51), this requires that

$$\frac{1}{n+1}\left|\frac{A^{(n+1)}_{T,0} f^{(n+1)}(\sigma\tilde{t})}{A^{(n)}_{T,0} f^{(n)}(\sigma\tilde{t})}\right| \equiv \frac{\sigma}{n+1}\left|\frac{A^{(n+1)}_{T,0}}{A^{(n)}_{T,0}}\right|\left|\frac{F^{(n+1)}(\sigma\tilde{t})}{F^{(n)}(\sigma\tilde{t})}\right| \ll 1$$

for all n. The term $|A^{(n+1)}_{T,0}/A^{(n)}_{T,0}|$ has the dimensions of time and for the two-beam model, taking $T \equiv \bar{\tau}$, we find (for example from (90)) that it is proportional to $\Delta\tau$. In general, we expect that $|A^{(n+1)}_{T,0}/A^{(n)}_{T,0}| \propto \Delta\tau_1$, where $\Delta\tau_1$ is some characteristic time analogous to $\Delta\tau$ in the two-beam model. We may call $\Delta\tau_1$ the 'elongation time' of the pulse, and generally it will depend upon position (cf. equation (5) for $\Delta\tau$) and ω_0.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

We expect rapid convergence if $\sigma\Delta\tau_1 \ll 1$. This sets the (position- and frequency-dependent) scale on which the pulse bandwidth σ must be small, in order that our subsequent perturbation formulae be valid. The only relation between σ and ω_0 is through the ω_0-dependence of $\Delta\tau_1$: there is no direct constraint of the form $\sigma \ll \omega_0$ in this theory, and in fact the carrier frequency ω_0 plays little role in the analysis. We return to the elongation time $\Delta\tau_1$ in §3.8.

To use (49) in practice we want to truncate it at the smallest possible number of terms. Let us define $\Psi_N(t;\, T)$ to be the partial sum of (49) up to the Nth derivatives. Then

$$\psi(t) \equiv \Psi_N(t;\, T) + R_{N+1}(t;\, T)$$

defines the remainder $R_{N+1}(t;\, T)$. The T dependences of Ψ and R, which are crucial, have been made explicit; the sum ψ is, of course, independent of T. For given N, to approximate $\psi(t)$ by $\Psi_N(t;\, T)$ we require the remainder to be as small as possible; therefore we choose T to minimize $|R_{N+1}(t;\, T)|$.

3.4. *Minimizing the truncation error*

For given $\boldsymbol{r}$ we must expect the optimum value of the local time origin T to differ according to the value of N that is chosen; for given N let us denote it by T_N. In this section we shall derive the approximate *necessary* conditions for $|R_{N+1}(t;\, T)|$ to be a minimum, valid when $\sigma\Delta\tau_1 \ll 1$. They are independent of t and of the pulse shape, as we would hope, and may be expressed purely in terms of the 'coefficients' $A^{(n)}_{T,0}$. Unfortunately the conditions are not sufficient, so that our results are not completely reliable. Nevertheless, they seem to work well in practice, and our ultimate justification of the theory appeals to this fact. Note that T_N can never be determined from the truncated series Ψ_N alone.

Denoting temporarily $R_{N+1}(t;\, T)$ by just R, we require $\partial|R|/\partial T = 0$, $\partial^2|R|/\partial T^2 > 0$ for $|R|$ to have a local minimum with respect to T. If we assume regularity in T, this is a necessary, but not sufficient, condition for an unconstrained global minimum at finite T. Denoting a complex conjugate by an asterisk, we have

$$\frac{\partial|R|^2}{\partial T} = 2|R|\frac{\partial|R|}{\partial T}$$

$$\equiv \frac{\partial(RR^*)}{\partial T} = 2\,\mathrm{Re}\left(R\frac{\partial R^*}{\partial T}\right)$$

and

$$\frac{\partial^2|R|^2}{\partial T^2} = 2|R|\frac{\partial^2|R|}{\partial T^2} + 2\left(\frac{\partial|R|}{\partial T}\right)^2$$

$$\equiv \frac{\partial^2(RR^*)}{\partial T^2} = 2\,\mathrm{Re}\left(R\frac{\partial^2 R^*}{\partial T^2}\right) + 2\left|\frac{\partial R}{\partial T}\right|^2.$$

Then $\partial|R|/\partial T = 0$, $\partial^2|R|/\partial T^2 > 0$ implies

$$\left.\begin{aligned} \mathrm{Re}\left(R\frac{\partial R^*}{\partial T}\right) &= 0, && (52a)\\ \mathrm{Re}\left(R\frac{\partial^2 R^*}{\partial T^2}\right) + \left|\frac{\partial R}{\partial T}\right|^2 &> 0. && (52b)\end{aligned}\right\}$$

Our next task is to differentiate

$$R_{N+1}(t;\, T) \equiv \mathrm{e}^{-\mathrm{i}\omega_0(t-T)} \sum_{n=N+1}^{\infty} \frac{\mathrm{i}^n}{n!} A^{(n)}_{T,0} f^{(n)}(t-T) \qquad (53)$$

with respect to T. First we evaluate

$$\frac{\partial A^{(n)}_{T,0}}{\partial T} \equiv \frac{\partial^n}{\partial\omega^n}\frac{\partial}{\partial T}A_T(\omega)\bigg|_{\omega_0}.$$

From (41), without the explicit $\boldsymbol{r}$ dependences,

$$A_T(\omega) \equiv a(\omega)\,\mathrm{e}^{-\mathrm{i}\omega T} \quad \text{so} \quad \partial A_T(\omega)/\partial T = -\mathrm{i}\omega A_T(\omega).$$

Differentiating n times with respect to ω, using Leibnitz' formula, and setting $\omega = \omega_0$ gives

$$\partial A^{(n)}_{T,0}/\partial T = -\mathrm{i}(\omega_0 A^{(n)}_{T,0} + nA^{(n-1)}_{T,0}). \tag{54}$$

By differentiating (53), and using (54), the derivative of the exponential cancels the first term in (54) leaving

$$\frac{\partial R_{N+1}(t;T)}{\partial T} = \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\left\{\sum_{n=N+1}^{\infty}\frac{\mathrm{i}^{n-1}}{(n-1)!}A^{(n-1)}_{T,0}f^{(n)}(\tilde{t}) - \sum_{n=N+1}^{\infty}\frac{\mathrm{i}^{n}}{n!}A^{(n)}_{T,0}f^{(n+1)}(\tilde{t})\right\}, \tag{55}$$

where again $\tilde{t} = t - T$. All terms but one now cancel, leaving simply

$$\frac{\partial R_{N+1}(t;T)}{\partial T} = \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\frac{\mathrm{i}^{N}}{N!}A^{(N)}_{T,0}f^{(N+1)}(\tilde{t}). \tag{56}$$

This exact result is astonishing: the T-derivative of the infinite sum for $R_{N+1}(t;T)$ reduces to a single term, involving the last 'coefficient' $A^{(N)}_{T,0}$ occurring in $\Psi_N(t;T)$, which does not appear in $R_{N+1}(t;T)$ at all.

Differentiating again and using (54) produces a similar cancellation leading to

$$\frac{\partial^2 R_{N+1}(t;T)}{\partial T^2} = \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\frac{(-\mathrm{i})^N}{N!}\{\mathrm{i}NA^{(N-1)}_{T,0}f^{(N+1)}(\tilde{t}) + A^{(N)}_{T,0}f^{(N+2)}(\tilde{t})\}. \tag{57}$$

At this stage we must introduce some approximations to get tractable results. First, conditions (52) involve not only derivatives, but also $R_{N+1}(t;T)$ itself. The best we can do is to approximate $R_{N+1}(t;T)$ by its first term. This should be valid if ψ (and hence R_{N+1}) has rapid initial convergence, that is, if $\sigma\Delta\tau_1 \ll 1$ and T_N is correctly chosen, unless the first term of $R_{N+1}(t;T)$ happens to be zero or very small. With this approximation, conditions (52) cease to be sufficient even for a local minimum, although with approximate equality in (52a) they remain necessary. Second, we assume that $f(t)$ is varying sufficiently slowly that the second term in (57) is negligible, which, from the discussion at the end of §3.3, will again be true if $\sigma\Delta\tau_1 \ll 1$. Small errors in the second derivative should not matter, since we only constrain the sign of $\partial^2|R|/\partial T^2$, not its value.

Summarizing our results for the truncation error, using once again $\tilde{t} \equiv t - T$, we have

$$R_{N+1}(t;T) \approx \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\mathrm{i}^{N+1}\left\{\frac{A^{(N+1)}_{T,0}}{(N+1)!}\right\}f^{(N+1)}(\tilde{t}), \tag{58a}$$

$$\frac{\partial}{\partial T}R_{N+1}(t;T) = \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\mathrm{i}^{N+1}\left\{-\frac{\mathrm{i}A^{(N)}_{T,0}}{N!}\right\}f^{(N+1)}(\tilde{t}), \tag{58b}$$

$$\frac{\partial^2}{\partial T^2}R_{N+1}(t;T) \approx \mathrm{e}^{-\mathrm{i}\omega_0\tilde{t}}\mathrm{i}^{N+1}\left\{-\frac{A^{(N-1)}_{T,0}}{(N-1)!}\right\}f^{(N+1)}(\tilde{t}). \tag{58c}$$

The formal term-by-term cancellation in (55) appears to be independent of the ultimate convergence of the infinite series. Therefore, as long as ψ is truncated in its convergent region, equations (58) should hold even if (49) is asymptotic.

One essential outcome of the approximations leading to (58) is that all three equations in (58) have identical time dependences, so that on inserting (58) into (52) the time dependences cancel out, and the conditions for minimal truncation error reduce to

$$\left.\begin{aligned} \operatorname{Im}\left(A_{T,0}^{(N+1)}\, A_{T,0}^{(N)*}\right) &\approx 0 \qquad (59a)\\ N \operatorname{Re}\left(A_{T,0}^{(N+1)}\, A_{T,0}^{(N-1)*}\right) &< (N+1)\,|A_{T,0}^{(N)}|^2. \qquad (59b)\end{aligned}\right\}$$

These conditions, which are independent of t and valid for $\sigma\Delta\tau_1 \ll 1$, apply even if $f(t)$ is complex. For given N they determine T_N (if they have a solution) in terms of the transfer function $a(\omega)$ and the carrier frequency ω_0 only. Conditions (59) appear to be satisfied trivially by $A_{T,0}^{(N+1)} = 0$. However, this makes the approximation (58a) of R_{N+1} by its first term invalid, so this solution is invalid.

To proceed, we must separate real and imaginary parts. This works out most conveniently if we write the transfer function in amplitude-phase representation as

$$a(\boldsymbol{r},\omega) \equiv M(\boldsymbol{r},\omega)\, \mathrm{e}^{\mathrm{i}\varphi(\boldsymbol{r},\omega)} \quad (M > 0). \tag{60}$$

This is valid everywhere except where $a(\boldsymbol{r},\omega) = 0$ (because then φ is indeterminate), so $M = 0$ is excluded. To handle these c.w. nulls we must use the real–imaginary representation

$$a(\boldsymbol{r},\omega) \equiv p(\boldsymbol{r},\omega) + \mathrm{i}q(\boldsymbol{r},\omega). \tag{61}$$

This form is also more appropriate for numerical computations. We shall develop the theory in (M,φ)-representation, and then convert to (p,q)-representation as necessary, using the formulae in Appendix B.

In amplitude-phase representation, from (41),

$$A_T(\boldsymbol{r},\omega) \equiv M(\boldsymbol{r},\omega)\, \mathrm{e}^{\mathrm{i}\{\varphi(\boldsymbol{r},\omega)-\omega T(\boldsymbol{r})\}} \equiv M(\boldsymbol{r},\omega)\, \mathrm{e}^{\mathrm{i}\gamma(\boldsymbol{r},\omega)}, \tag{62}$$

where, as in (43), $\gamma \equiv \varphi - \omega T$. We shall use primes instead of superscripts to denote differentiation, drop explicit dependences, and use subscript zero as before to denote evaluation at $\omega = \omega_0$. Then, noting that $\gamma^{(n)} \equiv \varphi^{(n)}$ for $n \geqslant 2$, we have, for later use in conditions (59),

$$\left.\begin{aligned} A_{T,0} &\equiv M_0\, \mathrm{e}^{\mathrm{i}\gamma_0},\\ A'_{T,0} &\equiv (M'_0 + \mathrm{i}M_0\gamma'_0)\, \mathrm{e}^{\mathrm{i}\gamma_0},\\ A''_{T,0} &\equiv \{(M''_0 - M_0\gamma'^2_0) + \mathrm{i}(2M'_0\gamma'_0 + M_0\varphi''_0)\}\, \mathrm{e}^{\mathrm{i}\gamma_0},\\ A'''_{T,0} &\equiv [(M'''_0 - 3M'_0\gamma'^2_0 - 3M_0\varphi''_0\gamma'_0) + \mathrm{i}\{3M''_0\gamma'_0 + 3M'_0\varphi''_0 + M_0(\varphi'''_0 - \gamma'^3_0)\}]\, \mathrm{e}^{\mathrm{i}\gamma_0}. \end{aligned}\right\} \tag{63}$$

It may be shown, with the aid of equations (54), that conditions (59) for a local minimum of $|R_{N+1}(t;T)|$ are in fact *exactly* the conditions for a local minimum of $|A_{T,0}^{(N+1)}|$. The terms $R_{N+1}(t;T)$ and $A_{T,0}^{(N+1)}$ are approximately related by (58a), which approximates $R_{N+1}(t;T)$ by its first term. Therefore we see that, when this approximation is valid, it is also valid to ignore the T-dependence of $f^{(N+1)}(t-T)$, the last factor in (58a). (This follows from (58a) *alone* only under the additional assumption that $\sigma \ll \omega_0$.) Since, from (63), $|A_{T,0}^{(N+1)}|^2$ is a polynomial (of degree $2N+2$) in γ'_0, and hence in T (because $\gamma'_0 \equiv \varphi'_0 - T$), its global minimum is guaranteed to be one of the local minima given by (59). It can be shown that the prescriptions given below do give the global minimum of $|A_{T,0}^{(N+1)}|$, and hence, within our approximations, that of $|R_{N+1}(t;T)|$.

3.5. *Solution by successive approximation*

Having completed the mathematical groundwork, let us return to our main problem, to use equation (49), namely

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega_0 \tilde{t}} \sum_{n=0}^{\infty} \frac{\mathrm{i}^n}{n!} A^{(n)}_{T,0} f^{(n)}(\tilde{t}), \tag{49}$$

to predict the behaviour of a pulse of small bandwidth. Conditions (59) specify T, and initially we shall use representation (63) for $A^{(n)}_{T,0}$. We proceed by truncating (49) at $n = N$ for successively higher values of N until we produce a satisfactory formulation for the phenomenon we seek. At each new value of N the prescription T_N for T changes (slightly).

3.5.1. *Zero order*: $N = 0$

From equation (49):

$$\psi(t) \approx \Psi_0(t; T) = \mathrm{e}^{-\mathrm{i}\omega_0 \tilde{t}} A_{T,0} f(\tilde{t}) = M_0 \mathrm{e}^{\mathrm{i}(\varphi_0 - \omega_0 t)} f(t - T). \tag{64}$$

Condition (59*a*) for T_0 is $\mathrm{Im}\,(A'_{T,0} A^*_{T,0}) = 0$. With the use of (63) this becomes $M_0^2 \gamma'_0 = 0$. Since $M_0 = 0$ is excluded, we must have (using the definition (43) of γ in terms of φ)

$$\gamma'_0 \equiv \varphi'_0 - T_0 = 0. \tag{65}$$

For $N = 0$, the first term on the right of (57) is identically zero; so the approximation (58*c*) must be replaced by the second term in (57), involving $f^{(N+2)}(\tilde{t})$. For $\sigma \Delta\tau_1 \ll 1$ this term is negligible compared with terms involving $f^{(N+1)}(\tilde{t})$, from (51), so condition (52*b*), from which (59*b*) is derived, is trivially satisfied. Therefore $T_0 = \varphi'_0$ is optimal, and the zero-order approximation is

$$\psi(\boldsymbol{r}, t) \approx M(\boldsymbol{r}, \omega_0) \mathrm{e}^{\mathrm{i}\{\varphi(\boldsymbol{r}, \omega_0) - \omega_0 t\}} f\{t - \varphi'(\boldsymbol{r}, \omega_0)\}. \tag{66}$$

This achieves one of our stated aims by completely verifying approximation (45) and finding $T_0(\boldsymbol{r})$. The time delay of the carrier is given by $\varphi(\boldsymbol{r}, \omega_0)/\omega_0$, while the time delay of the envelope is given by $\varphi'(\boldsymbol{r}, \omega_0)$. The zero-order approximation gives only a crude description of a pulsed wavefield, and its main use is to provide a generalization of group velocity. It describes roughly the propagation of gross features of the driving pulse, such as its peaks. However, to study the more subtle dislocation lines we must proceed to higher order. (Zero order gives no extrinsic dislocations, and degenerate intrinsic dislocation surfaces; both become lines in first order.)

3.5.2. *First order*: $N = 1$

From equation (49):

$$\psi(t) \approx \Psi_1(t; T) = \mathrm{e}^{-\mathrm{i}\omega_0 \tilde{t}} \{A_{T,0} f(\tilde{t}) + \mathrm{i} A'_{T,0} f'(\tilde{t})\}. \tag{67}$$

General analysis becomes rather involved, so we shall from now on specifically seek extrinsic dislocations and shall assume f to be real. The point $\boldsymbol{r}$ is on a dislocation trajectory if $\psi(\boldsymbol{r}, t) = 0$ for some real t. Consequently we are on a dislocation trajectory to order N if $\Psi_N(t; T) = 0$ has real solutions for t. To first order, equating (67) to zero, inserting representations (63) and separating real and imaginary parts gives

$$\left.\begin{aligned} M_0 \{f(\tilde{t}) - \gamma'_0 f'(\tilde{t})\} &= 0, \\ M'_0 f'(\tilde{t}) &= 0. \end{aligned}\right\} \tag{68}$$

Since we are excluding $M_0 = 0$, and generically f and f' do not have simultaneous zeros, (68) implies

$$f(\tilde{t}) - \gamma_0' f'(\tilde{t}) = 0, \tag{69a}$$

$$M_0' = 0. \tag{69b}$$

Equation (69b) is independent of time, and represents one real equation in $\boldsymbol{r}$; so its solution is a surface in three-dimensional space. This is the dislocation trajectory; for any point on it (69a) gives the dislocation time t (via $t = T + \tilde{t}$, once we have prescribed T).

When $f(t)$ is at most linear in t equations (69) are exact, and *independent of* T, since then all the derivatives of f higher than f' in (49) vanish. But, as we shall see, (69) can still be a good approximation, provided T is correctly specified, for more general pulses.

From now on we consider solutions only on the surfaces in space on which $M_0' = 0$. Condition (59a) for T_1 is

$$\operatorname{Im}(A''_{T,0} A'^{*}_{T,0}) = 0.$$

If we use (63), and set $M_0' = 0$, this becomes

$$-M_0 \gamma_0' (M_0'' - M_0 \gamma_0'^2) = 0. \tag{70}$$

The condition obtained in the $N = 0$ approximation, $\gamma_0' = 0$, is still a solution, but then the dislocation equation (69a) reduces to $f(\tilde{t}) = 0$, which only gives intrinsic dislocations. To find extrinsic dislocations we must take

$$M_0'' - M_0 \gamma_0'^2 = 0 \quad \text{which implies} \quad \gamma_0' = \pm\sqrt{(M_0''/M_0)}. \tag{71}$$

This gives real values of

$$T_1 \equiv \varphi_0' - \gamma_0' = \varphi_0' \mp \sqrt{(M_0''/M_0)} \tag{72}$$

only when $M_0'' > 0$, thereby restricting the dislocation trajectories to *frequency minima* or 'anti-resonances' of the c.w. amplitude. Condition (59b) for T_1 is

$$\operatorname{Re}(A''_{T,0} A^{*}_{T,0}) < 2|A'_{T,0}|^2.$$

If we use (63) and set $M_0' = 0$, this becomes

$$M_0(M_0'' - M_0 \gamma_0'^2) < 2(M_0 \gamma_0')^2,$$

which is trivially satisfied by (71). We appear to have two alternative optimal prescriptions (72) for T_1 (whose mean exactly equals the zero-order prescription $T_0 = \varphi_0'$).

However, this prescription is suspect because, when $M_0' = 0$ and $M_0'' - M_0 \gamma_0'^2 = 0$, from (63)

$$A''_{T,0} = \mathrm{i} M_0 \varphi_0'' \, \mathrm{e}^{\mathrm{i}\gamma_0}.$$

It turns out that for the two-beam model φ_0'' is identically zero when $M_0' = 0$, and for a more realistic model such as an acoustic piston radiator (Wright 1977, Wright & Berry in preparation) φ_0'' is certainly very small when $M_0' = 0$. (This is one of a number of apparent coincidences that play an important role in the present theory, and are under investigation.) Therefore $A''_{T,0}$ is zero or very small, and approximation (58a) is invalid. To apply first-order theory consistently to extrinsic dislocations we would have to approximate $R_2(t; T)$ by its first two terms. This would involve $A'''_{T,0}$, and also the time t and envelope $f(t)$. Rather than do this, it is better to proceed to second-order perturbation theory, which involves $A'''_{T,0}$ but not the time or envelope. This will confirm one of the prescriptions (72).

3.5.3. *Second order*: $N = 2$

From equation (49)

$$\psi(t) \approx \Psi_2(t; T) = \mathrm{e}^{-\mathrm{i}\omega_0 \tilde{t}}\{A_{T,0} f(\tilde{t}) + \mathrm{i}A'_{T,0} f'(\tilde{t}) - \tfrac{1}{2}A''_{T,0} f''(\tilde{t})\}.$$

Equating $\Psi_2(t; T)$ to zero, inserting representations (63) and separating real and imaginary parts, gives

$$\left.\begin{aligned} M_0\{f(\tilde{t}) - \gamma'_0 f'(\tilde{t})\} - \tfrac{1}{2}(M''_0 - M_0\gamma'^2_0) f''(\tilde{t}) &= 0, \\ M'_0 f'(\tilde{t}) - \tfrac{1}{2}(2M'_0\gamma'_0 + M_0\varphi''_0) f''(\tilde{t}) &= 0 \end{aligned}\right\} \quad \begin{aligned} &(73a) \\ &(73b) \end{aligned}$$

(since $\gamma''_0 = \varphi''_0$). We shall assume that $\varphi''_0 = 0$ when $M'_0 = 0$. Then our $N = 1$ equations, namely

$$\left.\begin{aligned} M'_0 = 0 \quad \text{with} \quad \varphi''_0 &= 0, \\ f(\tilde{t}) - \gamma'_0 f'(\tilde{t}) &= 0, \\ M''_0 - M_0\gamma'^2_0 &= 0, \end{aligned}\right\} \quad \begin{aligned} &(74a) \\ &(74b) \\ &(74c) \end{aligned}$$

satisfy equations (73) exactly. (If, in fact, $\varphi''_0 \neq 0$, there would be a small shift in the dislocation trajectories and times, as one would expect in the next approximation; and the time and space dependences would be coupled, making the full second-order solution much more complicated. Since our purpose is mainly to confirm the first-order solution, we shall take $\varphi''_0 = 0$.)

Condition (59*a*) for T_2 is

$$\mathrm{Im}\,(A'''_{T,0} A''^*_{T,0}) = 0.$$

This is satisfied by equations (74), since they make $A''_{T,0}$ vanish, as we have already discussed. Condition (59*b*) for T_2 is

$$2\,\mathrm{Re}\,(A'''_{T,0} A'^*_{T,0}) < 3|A''_{T,0}|^2.$$

Since $A''_{T,0} = 0$, this reduces, with the use of (63) and (74), to

$$\{3M''_0\gamma'_0 + M_0(\varphi'''_0 - \gamma'^3_0)\} M_0\gamma'_0 < 0.$$

Replacing $M_0\gamma'^2_0$ by M''_0, from (74*c*), gives

$$2M''^2_0 + M^2_0\gamma'_0\varphi'''_0 < 0.$$

A necessary, but not sufficient, condition for this to be satisfied is $\gamma'_0\varphi'''_0 < 0$, that is

$$\mathrm{sgn}\,(\gamma'_0) = -\mathrm{sgn}\,(\varphi'''_0).$$

This equation enables us to discriminate between the two alternative signs in (71). Generically $A'''_{T,0} \neq 0$; so our approximation of $R_3(t; T)$ by its first term should be valid (unlike the situation to first order).

To summarize, we have confirmed one of the first-order solutions, that which may persist with only small shifts to second order. The trajectory is that part of the surface $\mathscr{S}$ given by:

$$M'_0 = 0, \quad M''_0 > 0 \quad \text{(trajectory)}, \tag{75a}$$

on which the time t is real, where

$$\left.\begin{aligned} t = T_1 + \tilde{t}, \quad T_1 &= \varphi'_0 - \gamma'_0, \\ \gamma'_0 = -\mathrm{sgn}\,(\varphi'''_0)\sqrt{(M''_0/M_0)}, \\ f(\tilde{t}) - \gamma'_0 f'(\tilde{t}) &= 0. \end{aligned}\right\} \quad \text{(time)} \tag{75b}$$

and $\tilde{t}$ satisfies

This prescription for T_1 gives minimal truncation error. (We can now see that the alternative sign in (72) actually gives *maximal* truncation error.) We shall take equations (75) as our *first-order perturbation formulae* for extrinsic dislocations. Although we had to work to second order to fix the sign of γ_0', (75) is strictly only a first-order solution of $\psi(\boldsymbol{r}, t) = 0$, unless it happens that $\varphi_0'' = 0$ when $M_0' = 0$, when fortuitously (75) may become exact to second order. Generally, φ_0'' will produce a small shift in formulae (75), embodied in equations (73).

The only ultimate verification of these formulae comes from comparison with exact results for suitable models, as we shall discuss later.

3.6. *The dislocation trajectories*

Where the bandwidth σ is sufficiently small that $\sigma\Delta\tau_1(\boldsymbol{r}, \omega_0) \ll 1$, in the first-order approximation extrinsic dislocations travel along the surfaces $\mathscr{S}$ in space on which $M'(\boldsymbol{r}, \omega_0) = 0$, $M_0''(\boldsymbol{r}, \omega_0) > 0$: that is, surfaces in the continuous wave diffraction pattern for frequency ω_0 where the amplitude is a minimum with respect to changes in frequency. This simple and unexpected conclusion constitutes one of the main results of this paper.

In the next approximation there will be a correction to this expression for the dislocation trajectories, depending on σ. However, it is important to remark that, provided the original pulse envelope is symmetrical in time, the correction is of at least second order in σ. The reason is that if, with a symmetrical pulse, the sign of σ is changed, the physical system remains unaltered. Therefore, the exact expression for the trajectories must be even in σ. This means that the error in taking $M_0' = 0$ for the trajectories is of order σ^2 or higher. The expression for the time must also be even in σ.

Generically, $M_0'' = 0$ defines lines in the surface $M_0' = 0$ where the minima change into maxima, beyond which our perturbation solution cannot extend. We have found no valid perturbation solutions for extrinsic dislocations (those not inherent in the driving pulse) where $M_0'' < 0$, but we have not proved that there are none. We conjecture that for small σ dislocations are only produced by diffraction at frequency minima of M. However, as the pulse length decreases, dislocations might appear where not predicted, before they cease to appear approximately where they are predicted!

The dislocation line may not actually reach all parts of the surface $\mathscr{S}$. This depends on the pulse shape, through (75*b*), which gives

$$f(\tilde{t})/f'(\tilde{t}) = -\operatorname{sgn}(\varphi_0''')\sqrt{(M_0''/M_0)}. \tag{76}$$

Physically $f(t) \to 0$ as $t \to \pm\infty$; so $f(t)$ has at least one stationary point and $|f/f'|$ is unbounded above. Of course $|f/f'|$ is always bounded below by zero, but for some pulse shapes it may have a lower bound $B > 0$. Then the extent of the trajectory is limited by

$$M_0''/M_0 \geqslant B^2.$$

For example, our most realistic model pulse shape

$$f(t) \equiv F(\sigma t) = \operatorname{sech}\sigma t, \tag{77}$$

gives

$$f/f' = -\sigma^{-1}\coth\sigma t,$$

and hence $B = \sigma^{-1}$. Therefore the trajectory is restricted to

$$M_0''/M_0 \geqslant 1/\sigma^2.$$

The Lorentzian envelope also gives $B > 0$, but for the Gaussian $B = 0$ and the full extent of $\mathscr{S}$ is used.

Figure 10*a* shows in an Argand diagram how the complex delayed transfer function $A_T(\boldsymbol{r}, \omega)$ varies with frequency ω at a particular point $\boldsymbol{r}$ (not a c.w. null) on $\mathscr{S}$. For a point $\boldsymbol{r}$ on $\mathscr{S}$ close to a c.w. null for $\omega = \omega_0$ the representative point passes close to the origin, and $M(\boldsymbol{r}, \omega)$ is a minimum at $\omega = \omega_0$. For a point $\boldsymbol{r}$ exactly at a c.w. null the locus in figure 10*a* must pass through the origin at $\omega = \omega_0$. The minimum in M then becomes a corner, with discontinuous slope and infinite curvature (figure 10*b*) and M' passes discontinuously through zero at $\omega = \omega_0$. Therefore the surfaces $\mathscr{S}$ contain, in this limiting sense, the lines $M_0 = 0$. In fact, differentiating M continues the lines $M_0 = 0$ into surfaces $M_0' = 0$, by replacing the two real equations implicit in $M_0 = 0$ by just one real equation $M_0' = 0$.

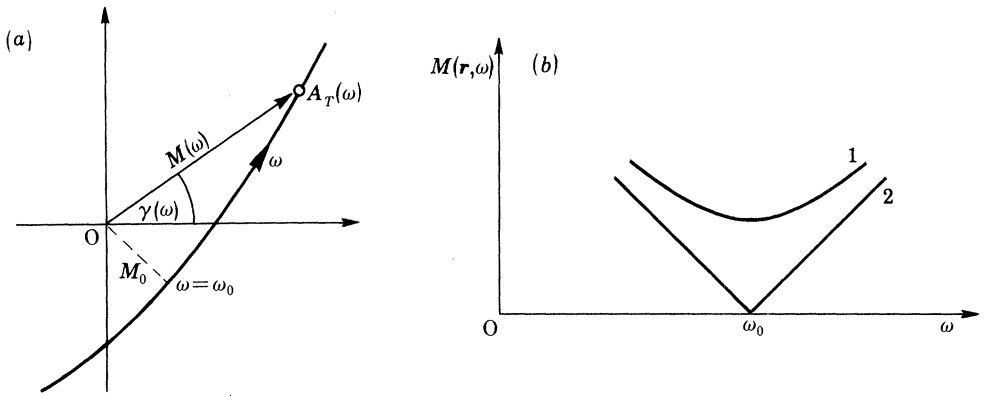

FIGURE 10. (*a*) Argand diagram for the delayed transfer function $A_T(\boldsymbol{r}, \omega)$ as ω varies for a fixed point $\boldsymbol{r}$ on $\mathscr{S}$. The surface $\mathscr{S}$ is defined by $M_0' = 0$, $M_0'' > 0$ and, to first order, contains the dislocation trajectory. (*b*) The modulus $M(\boldsymbol{r}, \omega)$ of $A_T(\boldsymbol{r}, \omega)$ as a function of ω: curve 1 is for a general point $\boldsymbol{r}$ on $\mathscr{S}$; curve 2 is for $\boldsymbol{r}$ at a c.w. null for $\omega = \omega_0$.

The fact that the lines on which $M_0 = 0$ lie in the surfaces $\mathscr{S}$ which give the approximate trajectories produces a problem, because the amplitude-phase representation (60) for the transfer function breaks down when $M = 0$, since then M and φ are singular. However, we show in Appendix B that the spatial behaviour of the frequency derivatives of M and φ may be 'regularized' by multiplying by appropriate powers of M. These regularized quantities may then be interpolated through a c.w. null in space. In general, we must evaluate third derivatives of the transfer function solely to fix the sign of the square root in (75*b*). However, the analysis in Appendix C shows that for trajectories that pass (approximately) through c.w. nulls it is sufficient to choose the sign of the square root so that T_1 is regular through the nulls, thus saving some computation.

3.7. *The meanings of* $T_0(\boldsymbol{r})$ *and* $T_1(\boldsymbol{r})$

Let us now review the meanings of $T_0(\boldsymbol{r})$ and $T_1(\boldsymbol{r})$. The time $T_0(\boldsymbol{r})$ is the optimum local time origin when we retain only the zero-order term in the perturbation series. It is given by $T_0(\boldsymbol{r}) = \varphi_0'(\boldsymbol{r})$ and has the physical significance of the pulse delay time for such small bandwidth that the pulse shape is undistorted. It is singular at a c.w. null (§3.2).

The time $T_1(\boldsymbol{r})$ is the optimum local time origin when we retain the first two terms in the series, and we have only defined it at points on the surfaces $\mathscr{S}$, on which $M_0' = 0$, $M_0'' > 0$. It is regular within this domain, and in particular it is regular at a c.w. null (Appendix C). When the pulse

shape is distorted, as it must be to contain a dislocation (figure 2*b*), the 'delay time of the pulse' can no longer be uniquely specified. One way of dealing with this problem is to define a *reference pulse*, namely the original pulse envelope undistorted and delayed by $T_1(\boldsymbol{r})$, and to use this as a standard against which to compare the actual distorted pulse envelope. We can also, as a matter of convenience, use $t = T_1(\boldsymbol{r})$ as a reference time from which to measure the position of the dislocation. (Note that $T_0(\boldsymbol{r})$ would not serve this purpose well because it becomes indefinitely large near a c.w. null.) Accordingly we call $T_1(\boldsymbol{r})$ the *dislocation reference time*.

In general, of course, the carrier wave moves relative to the reference pulse and we may recall that glide is defined (Nye 1981) as motion of the dislocation relative to the carrier wave rather than relative to the pulse. It is interesting that in the two-beam model (where the medium is non-dispersive) $\bar{\tau}$, which is related to the delay of the carrier wave, rather than the delay of the pulse, appears as the natural time origin in the exact analysis.

The first-order perturbation theory gives an interesting result for the signal received at a c.w. null. If we write (76) as

$$|f'(\tilde{t})/f(\tilde{t})| = \sqrt{(M_0/M_0'')},$$

then regularize and insert expressions (B 2) and (B 6*a*) from Appendix B, we find that

$$|f'(\tilde{t})/f(\tilde{t})| \sim |(p^2+q^2)/(pq'-qp')| \to 0,$$

as $M^2 = p^2+q^2 \to 0$. We deduce that, in the first-order approximation, $f'(\tilde{t}) = 0$ at a c.w. null.

The reference pulse envelope is $f(t-T_1) \equiv f(\tilde{t})$. Therefore our result means that, for a receiver at a c.w. null, the dislocation occurs in the signal at a time when the slope of the reference pulse envelope is zero. For example, if $f(t)$ has a single peak at $t = 0$, as in § 2, so that $f'(0) = 0$, and has no other stationary points, it follows from $f'(\tilde{t}) = 0$ that $\tilde{t} = 0$; hence $t = T_1$. At a c.w. null, the dislocation in the actual signal appears simultaneously with the peak of the reference pulse.

We now look more closely at the dislocation behaviour near a c.w. null.

3.8. *Behaviour near a continuous-wave null*

We have seen that the approximate dislocation trajectories are parts of the surfaces $\mathscr{S}$ and that these surfaces contain the c.w. null lines for $\omega = \omega_0$. Do the exact trajectories also contain the c.w. null lines? They do so in the two-beam model, but we shall see in this section that, in general, the c.w. null lines lie near but not on the exact trajectory surfaces, intersecting them in isolated points. By restricting attention to the immediate neighbourhood of a c.w. null we can express the dislocation conditions as power series in the bandwidth σ and this enables us to proceed in a more rigorous fashion than before.

We use a real–imaginary representation for $a(\boldsymbol{r}, \omega)$,

$$a(\boldsymbol{r}, \omega) \equiv \mathrm{e}^{\mathrm{i}\beta_0}\{P(\boldsymbol{r}, \omega) + \mathrm{i}Q(\boldsymbol{r}, \omega)\}, \tag{78}$$

where β_0 is an arbitrary phase to be chosen for convenience. Take as origin $\boldsymbol{r} \equiv (x, y, z) = 0$ a point O on a c.w. null line for $\omega = \omega_0$. Putting $W = \omega - \omega_0$ as before, the local approximation for P and Q in powers of x, y, z, W may be written

$$P = a_1x + a_2y + a_3z + a_4W + a_5xW + a_6yW + a_7zW + a_8W^2 + \ldots, \tag{79a}$$

$$Q = b_1x + b_2y + b_3z + b_4W + b_8W^2 + \ldots, \tag{79b}$$

where the a_r and b_r are constants. Out of the 16 terms quadratic in x, y, z, W we have explicitly written only five; the rest are to be understood. For $\boldsymbol{r} = 0$ the point in an Argand diagram representing $P+\mathrm{i}Q$ passes through the origin of the diagram as ω passes through ω_0. We choose β_0 so that the locus in fact traverses the origin vertically. Then, since

$$P = a_4 W + a_8 W^2 + \ldots, \quad Q = b_4 W + b_8 W^2 + \ldots,$$

we have $a_4 = 0$.

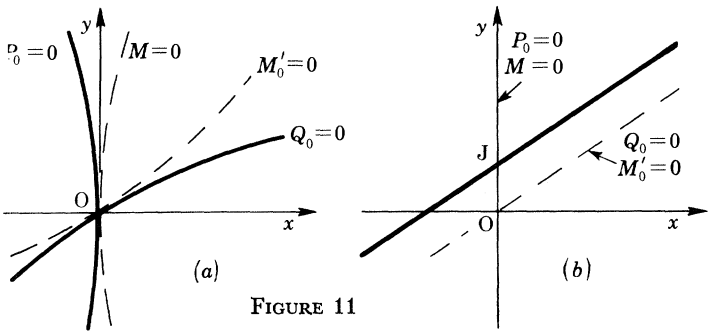

FIGURE 11

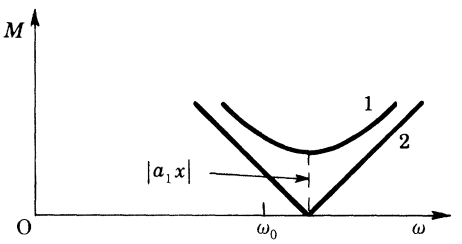

FIGURE 12

FIGURE 11. (*a*) The c.w. null line for $\omega = \omega_0$ is the intersection of the surfaces $P_0 = 0$ and $Q_0 = 0$. At O, an arbitrary point on this line, $P_0 = 0$ is tangential to the null surface $M = 0$, and $Q_0 = 0$ is tangential to the surface $M_0' = 0$. The axes are chosen so that $a_2 = a_3 = b_3 = 0$. (*b*) For small σ, the dislocation trajectory (solid line) is locally a plane parallel to $Q_0 = 0$ and makes intercepts on the axes proportional to σ^2.

FIGURE 12. Curve 1 shows $M(\omega)$ for a general point on the trajectory in figure 11*b*. $M(\omega)$ has a minimum of height $|a_1 x|$ which is displaced from ω_0 by a frequency difference proportional to σ^2. Curve 2 shows the same but for point J in figure 11*b*; the height of the minimum is zero.

As before we shall use the subscript zero to denote quantities evaluated at $W = 0$. The surface $P_0 = 0$ has a special meaning. As the continuous-wave frequency is changed the null line will trace out a 'null surface' in (x, y, z) which we can label $M = 0$ (figure 11*a*). For a given W the null line is the intersection of the two surfaces $P = 0$ and $Q = 0$. Each of these equations describes, in general, a family of surfaces in (x, y, z), parametrized by W; in our case they are, locally,

$$a_1 x + a_2 y + a_3 z = 0, \tag{80a}$$

$$b_1 x + b_2 y + b_3 z + b_4 W = 0, \tag{80b}$$

with W appearing only in $Q = 0$. Hence, projecting out W, the null surface $M = 0$ is represented locally simply by the first equation, which is $P_0 = 0$. The surface $P_0 = 0$ is therefore tangential at the origin to the null surface $M = 0$ (figure 11*a*).

Likewise the surface $Q_0 = 0$ has a special meaning. In general, for given $(\boldsymbol{r}, W)$,

$$M^2 = P^2 + Q^2.$$

Differentiating with respect to W and putting $W = 0$ we have

$$M_0 M_0' = P_0 P_0' + Q_0 Q_0'.$$

Equations (79*a*, *b*) with $a_4 = 0$ show that $P_0 P_0'$ is $O(r^2)$ while $Q_0 Q_0'$ is $O(r)$. Hence, to first order $M_0 M_0' = Q_0 Q_0'$. Now consider a point, not the origin, where $M_0' = 0$. Then, to first order $Q_0 Q_0' = 0$. But $Q_0' \neq 0$ near the origin (to first order it equals b_4). Hence $Q_0 = 0$. This shows that the surface $M_0' = 0$ is tangential to the surface $Q_0 = 0$ at the origin (figure 11*a*).

Now choose Oz to be tangential to the c.w. null line and Ox to be perpendicular to the surface $P_0 = 0$, thereby making $a_2 = a_3 = b_3 = 0$.

We must now find the condition for a dislocation. In the approach of this section we do not need to introduce the local time origin $T(\boldsymbol{r})$ because we can ensure sufficiently rapid convergence of the perturbation series by choosing σ and $\boldsymbol{r}$ small enough; so we simply use equation (49) with $T \equiv 0$ and $A_T(\boldsymbol{r}, \omega) = a(\boldsymbol{r}, \omega)$; that is, for given $\boldsymbol{r}$,

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega_0 t} \sum_{n=0}^{\infty} \mathrm{i}^n (n!)^{-1} a_0^{(n)} f^{(n)}(t). \tag{81}$$

If the point $\boldsymbol{r}$ lies on a dislocation trajectory, $\psi(t) = 0$ for some real t. Using this condition, dividing by $\mathrm{e}^{\mathrm{i}(\beta_0 - \omega_0 t)}$, and separating real and imaginary parts (taking f as real) we have

$$\left.\begin{aligned} P_0 f(t) - Q_0' f'(t) - \tfrac{1}{2} P_0'' f''(t) + \ldots = 0, \\ Q_0 f(t) + P_0' f'(t) - \tfrac{1}{2} Q_0'' f''(t) + \ldots = 0. \end{aligned}\right\} \tag{82a, 82b}$$

Now introduce the bandwidth σ explicitly as in (50) and use the expressions (79) (with $a_2 = a_3 = b_3 = a_4 = 0$):

$$\left.\begin{aligned} (a_1 x + \ldots) F(\sigma t) - \sigma (b_4 + \ldots) F'(\sigma t) - \sigma^2 (a_8 + \ldots) F''(\sigma t) + \ldots = 0, \\ (b_1 x + b_2 y + \ldots) F(\sigma t) + \sigma (a_5 x + a_6 y + a_7 z + \ldots) F'(\sigma t) - \sigma^2 (b_8 + \ldots) F''(\sigma t) + \ldots = 0, \end{aligned}\right\} \tag{83a, 83b}$$

where dots denote higher-order terms in x, y, z and σ. With small but non-zero σ these two simultaneous equations in x, y, z, t should determine, by elimination of t, a surface in (x, y, z) which is the trajectory, and also an expression for t as a function of position on the trajectory.

In the two-beam model σ is 'small' when it is much less than $|\Delta\tau|^{-1}$, or equivalently $2|\Delta\tau| \ll 2\sigma^{-1}$; that is, when the extra spreading of the received pulse is much less than the original pulse length $2\sigma^{-1}$. The dimensionless small quantity is $|\sigma\Delta\tau|$. In the general case the quantity analogous to $\Delta\tau$ is $\Delta\tau_1$, the elongation time introduced in § 3.3, which measures the spread of the time delays of the constituent pulses that make up the received pulse, and which is governed by the transfer function $a(\boldsymbol{r}, \omega)$. We define $\epsilon = |\sigma\Delta\tau_1|$ and, thinking of $\Delta\tau_1$ as locally constant, we call ϵ the (dimensionless) delay bandwidth of our pulse.

We anticipate from the principle stated at the beginning of § 3.6 that, if the trajectory does not pass through O, it will miss it by a distance proportional to σ^2, at any rate for a symmetric pulse. Therefore we look for a solution in the small region around O where x, y, z (suitably scaled to be dimensionless) are $O(\epsilon^2)$.

On the part of the trajectory within the small region near O we guess (and verify later) that t will be such that σt is $O(\epsilon)$. We assume $F'(0) = 0$, which is the case if F is peaked at zero, and since σt is small, we can then write

$$\begin{aligned} F(\sigma t) &= F(0) + O(\epsilon^2), \\ F'(\sigma t) &= \sigma t F''(0) + O(\epsilon^2), \\ F''(\sigma t) &= F''(0) + \sigma t F'''(0) + O(\epsilon^2). \end{aligned}$$

Knowing now the orders, in terms of ϵ, of all the terms in equations (83) we keep only the lowest order, which is ϵ^2, that is terms in σ^2, or x, y, z; thus

$$\left.\begin{aligned} a_1 x F(0) - \sigma^2 b_4 t F''(0) - \sigma^2 a_8 F''(0) = 0, \\ (b_1 x + b_2 y) F(0) - \sigma^2 b_8 F''(0) = 0. \end{aligned}\right\} \tag{84a, 84b}$$

It is important to remark that the term in (84b) involving t has dropped out, being $O(\epsilon^3)$.

Equation (84a) then gives for the dislocation time

$$t = -\frac{a_8}{b_4} + \frac{a_1}{\sigma^2 b_4}\frac{F(0)}{F''(0)}x, \tag{85}$$

and (84b) gives for the trajectory

$$b_1 x + b_2 y = \sigma^2 b_8 \frac{F''(0)}{F(0)}. \tag{86}$$

The factor $F''(0)/F(0) = -1$ for all three of our illustrative pulse envelopes (Gaussian, hyperbolic secant, Lorentzian).

Equation (86) defines a plane, locally parallel to $Q_0 = 0$ and therefore parallel to $M_0' = 0$, but making intercepts on the axes of length $O(\epsilon^2)$ (figure 11b). This justifies the original guess that a solution exists at a distance from O of order ϵ^2. Then (85) shows that, in this region, σt is indeed of order ϵ, as anticipated. Equation (85) shows that the time of the dislocation is $-a_8/b_4$ at point J, where $x = 0$, and varies linearly with x along the trajectory. In terms of derivatives with respect to frequency, taken at O not J, this time is $-\frac{1}{2}P_0''/Q_0'$.

This result is compatible with our general perturbation theory. In first-order theory the approximate dislocation trajectory (part of the surface $M_0' = 0$) passes through O and not through J. The best we can do in first-order theory is to find the time at which the dislocation passes through the c.w. null at O. The analysis in § 3.7 shows that for a peaked pulse this time is $t = T_1$.

The quantities p and q of the general theory are related to P and Q of the present local theory by

$$p + \mathrm{i}q \equiv \mathrm{e}^{\mathrm{i}\beta_0}(P + \mathrm{i}Q).$$

With this relation equation (C 4) of Appendix C gives, at a c.w. null,

$$T_1 = \frac{1}{2}\frac{p_0' q_0'' - q_0' p_0''}{p_0'^2 + q_0'^2} = \frac{1}{2}\frac{P_0' Q_0'' - Q_0' P_0''}{P_0'^2 + Q_0'^2}.$$

However, we have chosen β_0 so that $P_0' = 0$ at the origin, which is the c.w. null. Therefore

$$t = T_1 = -\tfrac{1}{2}P_0''/Q_0'.$$

This shows that the time from the perturbation theory at which the dislocation passes through the null is identical with the time at which the dislocation passes (more accurately) through J in the local theory.

It is of interest to note the local behaviour of M; we have

$$M^2 = P^2 + Q^2 = a_1^2 x^2 + (b_1 x + b_2 y + b_4 W)^2, \tag{87}$$

and at a point on the trajectory, by (86) with $F''(0)/F(0) = -1$,

$$M^2 = a_1^2 x^2 + (b_4 W - \sigma^2 b_8)^2.$$

For given x, $M(W)$ has a minimum of $|a_1 x|$ at $W = \sigma^2 b_8/b_4$ (figure 12), and at $x = 0$, that is J in figure 11b, the height of the minimum is zero. As a check of consistency note that J lies on the locus $M = 0$, the surface swept out by the c.w. null as W is changed; thus J is on the c.w. null line for frequency $\omega = \omega_0 + \sigma^2 b_8/b_4$.

For a general choice of O on the c.w. null line, b_8 in (86) will not be zero, and so the trajectory will not pass through O. But there is, in general, a surface in space on which $b_8 = 0$, and very near where the c.w. null line cuts this surface there will be special points where the trajectory

intersects the null line. To examine these special points in detail one would need to look at higher-order terms.

The three-dimensional picture is therefore as follows. For small bandwidth, the dislocation trajectory surfaces are very close to the surfaces $M_0' = 0$, $M_0'' > 0$, which contain the c.w. null lines. However, the trajectories themselves do not contain the c.w. null lines; they intersect them in isolated points.

We have now justified the assertion made in connection with figure 9 that, for a hyperbolic secant pulse envelope, the distance of the trajectory from the c.w. null is proportional to σ^2 as $\sigma \to 0$. The other assertion, that the length of the trajectory becomes proportional to σ as $\sigma \to 0$ is justified as follows.

In first-order perturbation theory the ends of the trajectory (where $t = \pm\infty$) are given by the condition derived in §3.6, which may be written

$$M_0/M_0'' = \sigma^2. \tag{88}$$

From (87) the behaviour of M on $M_0' = 0$ (whose equation is $b_1 x + b_2 y = 0$: see figure 11*b*) for small x and W is given by

$$M^2 = a_1^2 x^2 + b_4^2 W^2.$$

Differentiating twice with respect to W and putting $M_0' = 0$ we have $M_0 M_0'' = b_4^2$, and hence $M_0/M_0'' = M_0^2/b_4^2 = a_1^2 x^2/b_4^2$. Combining this with (88) shows that $|x|$ for the end points of the trajectory, and thus also the length of the trajectory, are proportional to σ.

4. Applications

We now apply the result of the perturbation theory of §3 to three examples.

4.1. *The cusp diffraction pattern*

As the wavelength is reduced a diffraction pattern approaches the intensity distribution given by a simple ray theory; for light the theory is geometrical optics. The envelope of a family of rays is a caustic surface, and this may be thought of as the most primitive type of focusing. The next most primitive type of focusing results in a crease in a caustic surface, called a cusp line, and the diffraction pattern we now describe is that associated with this cusp line. It is the second in a hierarchy of diffraction patterns (Berry 1976) which follows the list of Thom's (1972) and Arnol'd's (1975) elementary catastrophes.

In the general case, if all we know is the c.w. diffraction pattern at frequency ω_0, it is obviously impossible in principle to determine the surfaces $M_0' = 0$, simply because a derivative with respect to frequency is involved. On the other hand, diffraction patterns such as this one tend to vary with frequency in a highly systematic way (changing in size anisotropically); so in practice it may indeed be possible to infer the approximate position of the surfaces $M_0' = 0$ from the pattern at ω_0, given some form of additional information, such as the scaling law used below.

With the z-axis taken along the cusped edge, the modulus $M(x, y; \omega)$ of the complex cusp 'diffraction catastrophe', which is essentially two-dimensional (figure 13), is given by

$$M(x, y; \omega) \propto \omega^{\frac{1}{4}} \mathscr{M}(X, Y),$$

where $X = \omega^{\frac{1}{2}} x$, $Y = \omega^{\frac{3}{4}} y$ and $\mathscr{M}(X, Y)$ is a function first computed by Pearcey (1946). We assume here that the conditions are such that the position of the caustic itself is unaffected by change of frequency. The frequency dependence of the diffraction pattern is then described completely by the various powers of ω that appear in the above scalings.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

A short calculation shows that the condition $\partial M/\partial\omega = 0$ is equivalent to

$$\mathscr{M} + 2X\frac{\partial \mathscr{M}}{\partial X} + 3Y\frac{\partial \mathscr{M}}{\partial Y} = 0. \tag{89}$$

It is convenient to work in (X, Y) space. Then, by numerically differentiating the values of $\mathscr{M}(X, Y)$ already computed (Appendix C of Berry *et al.* 1979) we can easily plot the solutions of (89), as drawn in figure 13. (The numerical differentiation becomes unstable where $\mathscr{M}(X, Y)$ is slowly varying but this could be avoided by differentiating M (or $\mathscr{M}$) analytically and then

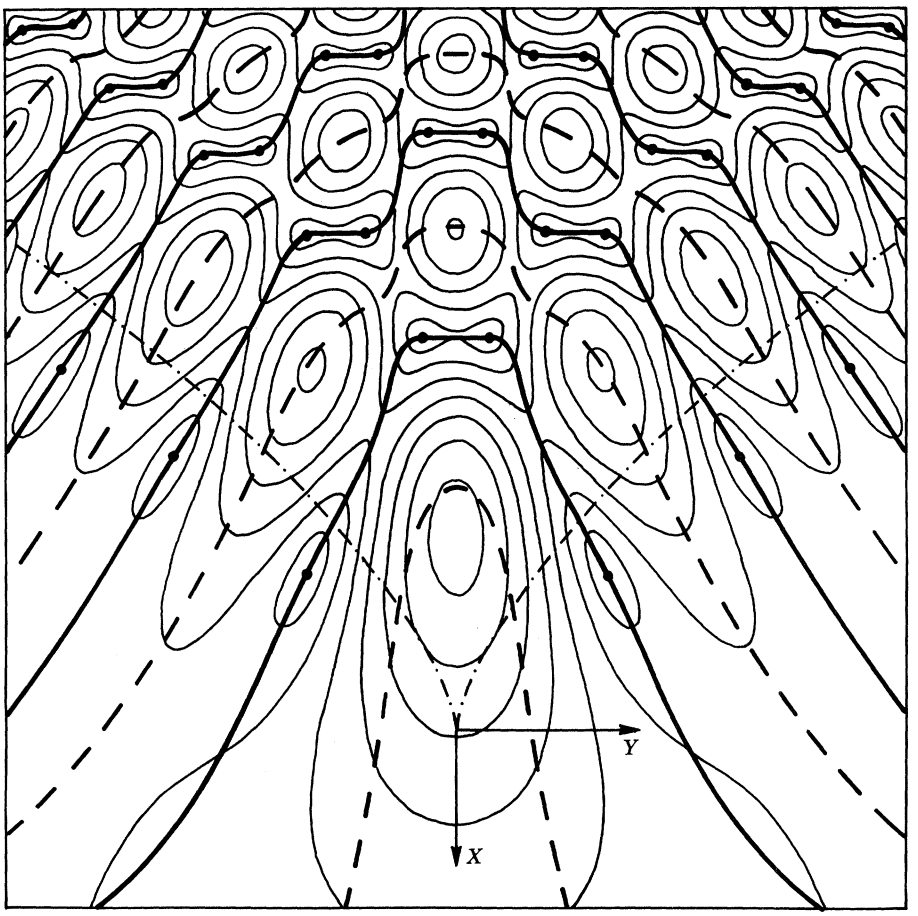

FIGURE 13. Amplitude contours of the cusp diffraction catastrophe. The loci of frequency minima $\mathscr{S}$ (thick solid lines) link the c.w. nulls (filled circles) and contain the approximate pulse dislocation trajectories for small bandwidth. The loci of frequency maxima (thick dashed lines) approximately link the spatial maxima. The cusp caustic is shown as a dot–dash line.

evaluating it numerically.) The solutions form two classes. Those shown by full lines link up all the c.w. null points by running up the valleys and over the saddles; these are the frequency minima. Those shown by dashed lines link up all the continuous wave maxima, and run down the ridges and over the saddles not used by the other family of curves; these are the frequency maxima. The lines do not necessarily pass exactly through the saddles and the maxima, but they are nearly always very close. It was a surprise to find the very close correspondence between the lines where M is a minimum with respect to frequency and the valleys in the M landscape, that

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

is, the landscape represented by M as a function of position, for fixed frequency. The correspondence arises because the main effect of a frequency change is to move the valleys sideways without much change in the height of their bottoms.

The perturbation theory predicts that, for a pulse of small bandwidth, the dislocation trajectories will be close to these valley lines. Experiments using ultrasonic waves (Humphrey & Nye to be published) show agreement with the prediction, but the directional characteristics of the receiver, which modify both the continuous wave diffraction pattern and the dislocation trajectories, must be taken into account in making a close comparison.

4.2. *The two-beam model*

The two-beam model of § 2 provides an exactly soluble problem against which to test realistically the accuracy of the perturbation formulae for the dislocation trajectories and times (and also to test the general numerical implementation described in Appendix D, although the general method was not used to obtain the results presented here).

First we find the transfer function. Write equation (2) in terms of the variable $z_2(x) = z_1(x_1-x)/x_1$:

$$\psi(x,z,t) = \tfrac{1}{2}z_1^{-1}\{(z_2+z)\,\psi_0(t-\tau_1) + (z_2-z)\,\psi_0(t-\tau_2)\};$$

then, with the help of (42), find $\overline{\psi}(\boldsymbol{r},\omega)$; hence, using (3), the transfer function is

$$a(x,z,\omega) \equiv M\,\mathrm{e}^{\mathrm{i}\varphi} = (z_2/z_1)\,\mathrm{e}^{\mathrm{i}\omega\bar{\tau}}\,(\cos\omega\Delta\tau + \mathrm{i}\,(z/z_2)\sin\omega\Delta\tau). \tag{90}$$

Derivatives of M and φ may be evaluated by applying the formulae in Appendix B to (90), taking $p \equiv (z_2/z_1)\cos\omega\Delta\tau$, $q \equiv (z/z_1)\sin\omega\Delta\tau$ and adding $\bar{\tau}$ to the resulting value of φ_0'. We notice that

$$pq'' - qp'' \equiv 0. \tag{91}$$

In this sense the two-beam model is degenerate. The first term in the expression for $M^4\varphi''$, (B 6*b*), is identically zero, and consequently $M_0^4\varphi_0'' = 0$ when $M_0 M_0' = 0$. Then, as remarked at the end of § 3.5, the result (75), which in general is correct to first order, satisfies $\psi = 0$ to second order, (73), as well (although T_1 is not *necessarily* optimal to second order).

In fact the prediction of (75*a*) that the trajectories satisfy $M_0' = 0$, $M_0'' > 0$ is not only correct to second order, but is exact. This can be seen analytically, or quite simply from the following physical argument. Consider the sources S_1, S_2 in figure 1 linearly modulated in space, as described, and emitting continuous waves of frequency $\omega = \omega_0$ to produce an interference pattern. At a fixed point on one of the dashed lines in figure 1 (but not on the horizontal line through C) we have unequal amplitudes but opposite phases, giving partial destructive interference. As ω is changed the amplitudes at this point are fixed but the relative phase changes. Clearly, then, the modulus of the resultant will be a minimum at $\omega = \omega_0$. If the fixed point considered were one of the nulls $\ldots \mathrm{N}_2', \mathrm{N}_1', \mathrm{N}_1, \mathrm{N}_2 \ldots$ the minimum of the modulus would be zero (as in figure 10*b*). We already know that the dislocation trajectories are parts of the broken lines; this proves that the broken lines are exactly the surfaces $\mathscr{S}$, where $M_0' = 0$, $M_0'' > 0$.

A further question (in view of § 4.1) is: do the trajectories follow the valleys in the landscape of the continuous wave amplitude? Figure 14*a* shows contours of c.w. amplitude at the pulse centre frequency for part of the field of figure 1. The c.w. amplitude and the dislocation trajectories are symmetric in z, so we only consider $z \geqslant 0$. The vertical $\mathscr{S}$ lines (full lines) appear to lie along the

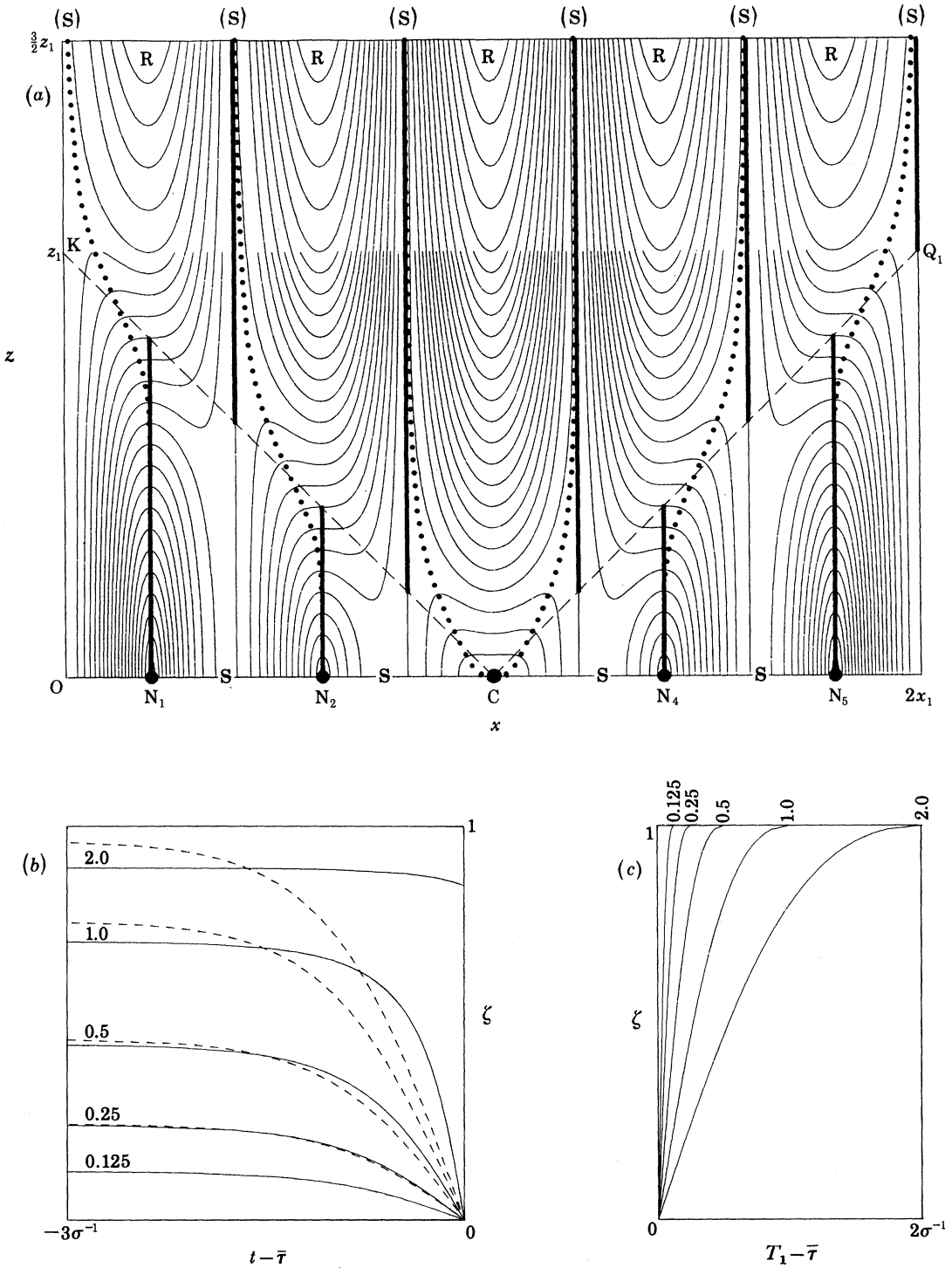

FIGURE 14. The two-beam model (region shown boxed in figure 1). (*a*) Amplitude contours showing the 'null normals' (dashed lines) and the dislocation trajectories (thick solid lines). The contour spacing is halved for $z < z_1$. N_i indicates a c.w. null (black dot), R a ridge, S a saddle and (S) an asymptotic saddle, from which the line of steepest descent is shown as a dotted line. (*b*) Accuracy of the perturbation theory. The solid curves are plotted from equation (93) for the perturbation theory; the dashed curves from equation (94) for the exact theory. The pairs of curves are labelled on the left by their values of $\sigma\Delta\tau$. (*c*) Relation of the carrier reference time $\bar{\tau}$ to the dislocation reference time T_1, for the same set of $\sigma\Delta\tau$ values (shown at the top) as used in (*b*).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

valley bottoms of the c.w. amplitude, at least far from the null normals CK, CQ_1. But there is a problem of definition in the term 'valley bottom', just as there is for any landscape. We shall adopt the definition 'line of steepest descent from a saddle point'. There are two sets of saddle points of the amplitude, both quite hard to recognize: one set lies along the x-axis, between (but not midway between) the nulls, each one being highly compressed in the x-direction; the other set is a row of degenerate saddle points at infinity on the $\mathscr{S}$ lines with each one infinitely compressed in the x-direction. The lines of steepest descent from the first set run along the x-axis and do not correspond to any dislocation trajectories. The dotted lines show the lines of steepest descent from the second degenerate set, and it will be seen that they do roughly follow the $\mathscr{S}$ lines. (But note that as the wave angle α is changed the pattern becomes more or less elongated in the z-direction, and the lines of steepest descent change their relative positions.)

Since the valleys clearly run approximately parallel to z we could try the alternative definition 'loci of points where the amplitude is a minimum with respect to x at fixed z'. The $\mathscr{S}$ lines passing through c.w. nulls lie exactly along such valley bottoms, but the other $\mathscr{S}$ lines do not. It is also interesting to see that the $\mathscr{S}$ lines through the c.w. nulls end on horizontal inflexions of the amplitude contours.

We now consider the dislocation times. From (7) and (17) the trajectories are given by

$$p_0 = 0 \quad \text{for} \quad |z| < |z_2| \quad \text{and} \quad q_0 = 0 \quad \text{for} \quad |z| > |z_2|,$$

where the subscript zero indicates, as usual, evaluation at $\omega = \omega_0$. In both cases, the formulae of Appendix B show that $M_0^6 \varphi_0'''$ is negative on the trajectories, so (75b) leads to

$$T_1 = \bar{\tau} + \Delta\tau/\zeta - (\Delta\tau/\zeta)\,(1-\zeta^2)^{\frac{1}{2}} \tag{92}$$

for the dislocation reference time, where

$$\zeta = \begin{cases} z/z_2 & \text{if} \quad |z| < |z_2| \\ z_2/z & \text{if} \quad |z| > |z_2|. \end{cases}$$

We recall that $z = \pm z_2$ represents the end points of each vertical dashed line in figure 1; thus, for the trajectories that contain the c.w. nulls, ζ is a dimensionless measure of the distance from the null. At the c.w. nulls T_1 has the finite limit $T_1 = \bar{\tau}$, where $z = 0$ and hence $\zeta = 0$.

If we choose the hyperbolic secant pulse shape (77), formula (75b) gives for the dislocation time

$$t = \bar{\tau} + \frac{\Delta\tau}{\zeta} - \frac{\Delta\tau(1-\zeta^2)^{\frac{1}{2}}}{\zeta} - \frac{1}{\sigma}\operatorname{artanh}\left\{\frac{\zeta}{\sigma\Delta\tau(1-\zeta^2)^{\frac{1}{2}}}\right\}, \tag{93}$$

with the restriction $|\zeta(1-\zeta^2)^{-\frac{1}{2}}| < |\sigma\Delta\tau|$. This is to be compared with the exact formula given by (23) and (25), with (10), namely

$$t = \bar{\tau} - \frac{1}{\sigma}\operatorname{artanh}\left(\frac{\zeta}{\tanh\sigma\Delta\tau}\right) \quad (|\zeta| < |\tanh\sigma\Delta\tau|). \tag{94}$$

Let us first consider the régime $|\zeta| \ll |\sigma\Delta\tau| \ll 1$, which is essentially the same as that treated for the general case in §3.8. Here both (93) and (94) reduce to the same approximate form,

$$t \approx \bar{\tau} - \zeta/\sigma^2\Delta\tau. \tag{95}$$

Physically this means that the perturbation theory gives the correct result if the difference between the time delays of the two pulses, $2|\Delta\tau|$, is much less than the pulse length, $2\sigma^{-1}$, and if the point examined is either very near a c.w. null or, for the trajectories with $|z| > |z_2|$, is such that $|z| \gg |z_2|$.

The appearance of σ^2 in the denominator of (95) shows once again why one has to approach the limit $\sigma \to 0$ with due care. As $\sigma \to 0$ the slope of the central $z(t)$ curve in figure 4*b* approaches zero, and in the limit all values of t are present at $z = 0$; that is, the dislocation is stationary at $z = 0$. To retain a dislocation ζ has to tend to zero at least as fast as $\sigma\Delta\tau$ (as is also illustrated by the contracting trajectories in figure 9). If, in contrast, we had taken $|\sigma\Delta\tau| \ll \zeta < 1$ there would be no dislocation, for the point would be beyond the end of the trajectory, and the received signal would be simply an attenuated and much advanced or retarded copy of the original pulse, as discussed at the end of § 3.2.

The comparison between (93) and (94) near to the limit $|\zeta| \ll |\sigma\Delta\tau| \ll 1$ can be carried a little further by expanding in powers of the small quantities ζ, $\sigma\Delta\tau$ and $\zeta/\sigma\Delta\tau$. Equation (93) gives

$$t = \bar{\tau} - \frac{\zeta}{\sigma^2\Delta\tau}\left\{1 + \frac{1}{3}\left(\frac{\zeta}{\sigma\Delta\tau}\right)^2 - \tfrac{1}{2}(\sigma\Delta\tau)^2 + \tfrac{1}{2}\zeta^2 + \ldots\right\},$$

while (94) gives

$$t = \bar{\tau} - \frac{\zeta}{\sigma^2\Delta\tau}\left\{1 + \frac{1}{3}\left(\frac{\zeta}{\sigma\Delta\tau}\right)^2 + \tfrac{1}{3}(\sigma\Delta\tau)^2 + \ldots\right\}.$$

This illustrates the general principle given at the beginning of § 3.6 that the expression for t must be even in σ. It is a vindication of our choice of T that the leading term is correct. The correction term in $(\zeta/\sigma\Delta\tau)^2$ is correctly reproduced and the remaining error is of order $(\sigma\Delta\tau)^2$ and ζ^2.

We should emphasize that, although the perturbation theory assumes $|\sigma\Delta\tau|$ to be small, it is not restricted to that very small neighbourhood of the c.w. nulls where $|\zeta| \ll |\sigma\Delta\tau|$. However, for fixed small σ, as ζ increases, the dislocation time becomes large and then infinite. Therefore, for fixed small σ, one cannot expect to predict very large dislocation times with high absolute accuracy. A fairer test is to choose a (large) time and ask how accurately ζ is predicted. For example, we might choose $t = \infty$, which represents an end point of the trajectory, and calculate the corresponding value of ζ (call it ζ_∞). The result given by the perturbation theory is, from (93),

$$|\zeta_\infty| = |\sigma\Delta\tau\{1 + (\sigma\Delta\tau)^2\}^{-\frac{1}{2}}|, \tag{96}$$

which is to be compared with the exact value,

$$|\zeta_\infty| = |\tanh \sigma\Delta\tau|, \tag{97}$$

given by (94). The two formulae are very similar, both being approximately $|\sigma\Delta\tau|$ for small $|\sigma\Delta\tau|$ and approaching 1 as $|\sigma\Delta\tau| \to \infty$. The maximum difference between them is 8.1% (when $|\sigma\Delta\tau| = 1.43$); when $|\sigma\Delta\tau| = 0.1$ they agree to 0.16% and when $|\sigma\Delta\tau| = 0.01$ they agree to 0.0017%. Thus these are the accuracies with which the perturbation theory predicts the length $|2\zeta_\infty|$ of the dislocation trajectories. Note that we only *expect* good results when $|\sigma\Delta\tau| \ll 1$.

Numerical comparison of (93) with (94) is shown in figure 14*b* for *all* allowed values of ζ and a wide range of values of $\sigma\Delta\tau$, confirming our previous analysis. Since both equations are anti-symmetric in ζ, it is sufficient to consider the range $0 \leqslant \zeta \leqslant 1$. Agreement is best near the origin and (coincidentally) near $t - \bar{\tau} \approx -1.5\sigma^{-1}$. When plotted in terms of the physical variables z and t the results appear as in figures 4 and 5 ($\bar{t}$ in figure 4*a* is the same as $t - \bar{\tau}$ in figure 14*b*). For comparison, note that the relative signs of ζ and z depend on $x - x_1$ through z_2, and that the sign of $\sigma\Delta\tau$ depends on x. If $\bar{\tau}$, given by (6), is fairly large, the relative error in t will be much less than that in $t - \bar{\tau}$ shown in figure 14*b*. This figure shows clearly how changing $\sigma\Delta\tau$ changes the

length of the trajectory, or equivalently that decreasing $\sigma\Delta\tau$ pushes the dislocation further into the head or tail of the pulse.

To put the values of $\sigma\Delta\tau$ into perspective, consider a pulse containing L carrier periods $(2\pi/\omega_0)$ within its pulse length $(2\sigma^{-1})$; that is $\sigma/\omega_0 = 1/L\pi$. It must be remembered that $\Delta\tau$ depends on x. Taking a very short pulse with $L = 3$ gives $\sigma\Delta\tau = \frac{1}{6} = 0.17$ on the trajectory at smallest x in figure 14*a*, that through N_1, and $\sigma\Delta\tau = \frac{5}{3} = 1.7$ on the trajectory at largest x in figure 14*a*, that through Q_1. A reasonably long pulse with $L = 30$ gives $\sigma\Delta\tau = \frac{1}{60}$ and $\frac{1}{6}$ respectively. For the short pulse we can only expect good results from the perturbation theory for small $|x|$, whereas for the longer pulse we can expect (very) good results for $|x| \lesssim 2x_1$. This discussion illustrates the fact that typically the convergence of a quasi-monochromatic perturbation theory, such as ours, will be non-uniform in space.

To complement our discussion in §3.7 of the meaning of T_1, the dislocation reference time which emerges from the perturbation theory, we relate T_1 to the carrier reference time $\bar{\tau}$ in figure 14*c*.

4.3. *The piston radiator*

Wright (1977) has computed from first principles the acoustic wave dislocations produced when a rigid plane circular piston, moving in an infinite plane baffle, is driven with a quasi-monochromatic pulse having a Gaussian envelope. This 'exact' computation provides a further test of the perturbation theory given in the present paper. It shows that the dislocation trajectories follow very closely the surfaces defined by $M_0' = 0$, $M_0'' > 0$, as expected. At the single c.w. null the dislocation position and time are predicted exactly. Elsewhere on the trajectories preliminary computations (Appendix D) show a maximum error in the dislocation times of about 10% for a bandwidth $\sigma/\omega_0 = 1/(3\pi) \approx 0.1$. The exact and the perturbation results naturally take different forms, and much of the error probably arises from difficulties of comparison, as well as from numerical problems. A full description of these results and comparisons will be given by Wright & Berry (in preparation).

5. Time dependence of the signal received on a dislocation trajectory

We have concentrated so far on calculating the trajectory and time of the dislocation. We now derive a simple result about the time dependence of the received signal.

A point receiver placed on a dislocation trajectory (not at a c.w. null) will find a signal $\psi(t)$ similar to that illustrated in figure 2*b*, and given by equation (49). Substituting the amplitude-phase representation (62), by using equations (63), we obtain

$$\psi(t) = \{\mathscr{R}(t) + \mathrm{i}\,\mathscr{I}(t)\}\,\mathrm{e}^{\mathrm{i}(\varphi_0 - \omega_0 t)}, \tag{98}$$

where the real functions $\mathscr{R}(t)$ and $\mathscr{I}(t)$ are given to second order $(N = 2)$ by the left-hand sides of (73*a*) and (73*b*): that is

$$\left.\begin{aligned} \mathscr{R}(t) &\equiv M_0\{f(\tilde{t}) - \gamma_0' f'(\tilde{t})\} - \tfrac{1}{2}(M_0'' - M_0\gamma_0'^2) f''(\tilde{t}), \\ \mathscr{I}(t) &\equiv M_0' f'(\tilde{t}) - \tfrac{1}{2}(2M_0'\gamma_0' + M_0\varphi_0'') f''(\tilde{t}), \end{aligned}\right\} \tag{99}$$

where $\tilde{t} = t - T_2$. Because the amplitude of $\psi(t)$ falls to zero at the time of arrival of the dislocation, t_D say, both $\mathscr{R}(t)$ and $\mathscr{I}(t)$ will, in general, pass linearly through zero at $t = t_D$. But, for a point on the trajectory, to the approximation of (73*b*), $M_0' = 0$ since we assume $\varphi_0'' = 0$.

Thus $\mathscr{I}(t)$ is so weakly dependent on t that, to this approximation, t drops out altogether and $\mathscr{I}(t) \equiv 0$. Thus

$$\psi(t) \approx \mathscr{R}(t)\, \mathrm{e}^{\mathrm{i}(\varphi_0 - \omega_0 t)}, \tag{100}$$

and near the dislocation

$$\psi(t) \propto (t - t_{\mathrm{D}})\, \mathrm{e}^{\mathrm{i}(\varphi_0 - \omega_0 t)} \tag{101}$$

approximately, with a real (positive or negative) constant of proportionality. (In the two-beam model (100) is exact, as shown by (9).)

Now φ_0 is the phase of the Fourier component of frequency ω_0 in the response function $a(\omega)$. So the phase relation between the received signal in a pulse experiment and in the corresponding continuous wave experiment, where the carrier wave is unmodulated, is very simple in this approximation: the signals are in phase on one side of the dislocation point $t = t_{\mathrm{D}}$ and in anti-phase on the other, as illustrated in figure 2*b* (although this figure is drawn for the two-beam model, the phase relation shown is general). It might be thought that this surprising general relation with the continuous wave experiment, which holds *for all t*, would have to break down if there were, for example, a deep dip in the envelope at some value of t, signalling that another dislocation was nearby. But it has to be remembered that, when σ is small enough to make the result a good approximation, the envelope of $\psi(t)$ will in fact be comparatively featureless (recall that in the $N = 0$ approximation (66) it has the same shape as the original pulse, with no dislocation).

The result does not hold for a general point in the field, but because it holds wherever $M_0' = 0$ and $\varphi_0'' = 0$ it does apply even on those parts of the surfaces $\mathscr{S}$ that are not actually used as trajectories for dislocations.

We may note that the physical signal represented by any chosen linear combination of real and imaginary parts in (100) has equally spaced zeros. The dislocation time ($t = t_{\mathrm{D}}$) can lie at any point of the cycle. In fact this attribute of a dislocation, its position in the cycle, can be made the basis for a classification of dislocations, their movement and their structure. The classification is not subject to the same small bandwidth limitation as that inherent in the perturbation theory, and is more appropriately treated in a separate paper (Nye 1981)†.

6. Discussion

In this paper §§2 and 3 embody two different approaches to the central problem of relating pulse dislocations to continuous wave diffraction patterns. In the two-beam model of §2 the results are exact and there is no restriction on bandwidth. The model shows in complete detail some of the variety of behaviour that pulse dislocations can exhibit, particularly pair creation

† *Note by J.F.N.* It is stated in Nye (1981, p. 236) that, for small bandwidth, in the neighbourhood of a point on the dislocation trajectory the plane-wave approximation to the continuous wave is identical in phase with the plane carrier wave for the dislocation (wave vector $\boldsymbol{k}_m$) as it passes through that point, except for a possible phase reversal. In fact this is only true for a neighbourhood confined entirely to the trajectory. The reason is that the phase relation derived in the present §5 holds only for points on the trajectory: therefore it gives no information on the rate of change of phase normal to the trajectory. This means that it gives no information on the normal component of $\boldsymbol{k}_m$, but only on the two components that are tangential to the trajectory (and that constitute the vector $\boldsymbol{k}_t$). The correct argument proceeds as follows. We have shown above that on the trajectory, for small bandwidth, the phase of the dislocated pulse wave ψ is the same (modulo π) as the phase of the continuous wave $\psi_{\mathrm{c.w.}}$, say. It is shown in Nye (1981, p. 233) that on the trajectory, near an isolated dislocation; the phase of ψ is the same (modulo π) as the phase of the carrier wave of the pulse dislocation, ψ_{carrier}. It follows that, on the trajectory, for small bandwidth near an isolated dislocation, the phases of all three waves ψ, $\psi_{\mathrm{c.w.}}$ and ψ_{carrier} are the same (modulo π).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

and how dislocations can appear and disappear in the head and tail of the pulse. It also shows the important effect of pulse shape and bandwidth, not upon the positions of the dislocation trajectories, but upon their length and upon the arrival times of the dislocations.

In the two-beam model the diffraction pattern is very simple, the medium is non-dispersive but the pulse can have any shape. In contrast, the perturbation theory of § 3 deals with the most general three-dimensional diffraction pattern but is restricted to small bandwidth; the greater the path, or time delay, differences in the diffraction pattern, the smaller must be the bandwidth. The medium is general (for example, inhomogeneous and dispersive) but all processes must be linear. Section 3 is essentially an attempt to set up a workable computational scheme for deducing the pulse dislocation behaviour from knowledge of the continuous wave diffraction pattern at and near the centre frequency; it is based on the subtle problem of producing (via the choice of $T(\boldsymbol{r})$), the fastest initial convergence of a functional series. We do not give precise conditions for its validity; it is to be judged by comparing its results with those of exact computation. In the two-beam model and the model of the piston radiator it serves very well to predict the dislocation trajectories, and it is also successful in the much more delicate matter of predicting the dislocation times. The results are expressed in terms of the c.w. diffraction pattern at the centre frequency ω_0 and its first and second derivatives with respect to frequency (the third derivative being used to decide a choice of sign). As the bandwidth increases it is quite natural that the method will need information about successively higher frequency derivatives if it is to give accurate results. A particularly simple result is the prediction that the dislocation lines travel approximately along the surfaces $\mathscr{S}$: that is the surfaces in the c.w. diffraction pattern for ω_0 where the c.w. amplitude is a minimum with respect to changes in frequency.

In § 3.8 we show how the perturbation series can be given a strict justification in a restricted domain, near the c.w. null lines. In particular, we show how, for small bandwidth, the dislocation trajectory surfaces do not, in general, actually contain the c.w. null lines but come very near to doing so; they intersect the null lines in isolated points.

This study was carried out at the H. H. Wills Physics Laboratory, University of Bristol, at the Department of Applied Mathematics, Queen Mary College, University of London and at the Institut für Informationsverarbeitung, Universität Tübingen. We are grateful to Professor M. V. Berry for his interest and encouragement. F.J.W. thanks the Science Research Council for financial support in Bristol, and in connection with the work in Tübingen thanks the Alexander von Humboldt Foundation for a research fellowship, Professor W. Güttinger and his staff for their hospitality, and Mr R. Brause for his assistance with the computer.

Appendix A. The T-independence of $\psi(t)$

From its derivation, the infinite series (49) for $\psi(t)$ must be independent of T, and therefore the T-dependences of the terms of the infinite sum must cancel. It is clear that this *must* happen, and the purpose of this Appendix is to make clear *how* it happens. We shall show that, via a resummation, the delay T of the transfer function A_T generates a Taylor expansion of $f(t)$ about $t - T$, thereby making explicit the mechanism by which T cancels.

The infinite-series representation for $\psi(t)$ is

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega_0(t-T)} \sum_{n=0}^{\infty} \frac{\mathrm{i}^n}{n!} A^{(n)}_{T,0} f^{(n)}(t-T). \tag{A 1}$$

Here $$A^{(n)}_{T,0} \equiv \frac{\partial^n}{\partial \omega^n}\{a(\omega)\,\mathrm{e}^{-\mathrm{i}\omega T}\}\Big|_{\omega_0},$$

which, using Leibnitz's formula, we may expand as

$$A^{(n)}_{T,0} \equiv \sum_{m=0}^{n} \binom{n}{m} a_0^{(m)} (-\mathrm{i}T)^{n-m}\,\mathrm{e}^{-\mathrm{i}\omega_0 T}.$$

Then (A 1) becomes

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega_0 t} \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{\mathrm{i}^n}{m!(n-m)!} a_0^{(m)} (-\mathrm{i}T)^{n-m} f^{(n)}(t-T).$$

Changing the order of summation to $\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}$ and writing $n-m \equiv n'$ gives

$$\begin{aligned}\psi(t) &= \mathrm{e}^{-\mathrm{i}\omega_0 t} \sum_{m=0}^{\infty} \sum_{n'=0}^{\infty} \frac{\mathrm{i}^{m+n'}}{m!\,n'!} a_0^{(m)} (-\mathrm{i}T)^{n'} f^{(m+n')}(t-T) \\ &= \mathrm{e}^{-\mathrm{i}\omega_0 t} \sum_{m=0}^{\infty} \frac{\mathrm{i}^m}{m!} a_0^{(m)} \sum_{n'=0}^{\infty} \frac{T^{n'}}{n'!} f^{(m+n')}(t-T). \qquad \text{(A 2)}\end{aligned}$$

The resummation has generated the Taylor expansion of $f^{(m)}(t)$ about $t-T$ recognized in the second sum in (A 2). Performing the second summation gives

$$\psi(t) = \mathrm{e}^{-\mathrm{i}\omega_0 t} \sum_{m=0}^{\infty} \frac{\mathrm{i}^m}{m!} a_0^{(m)} f^{(m)}(t),$$

which is precisely the result of setting $T = 0$ in (A 1). Thus, whatever value of T is used in (A 1) it drops out when the resummation is made.

The effect of delaying the transfer function by any time T is to advance by T, via the above Taylor resummation, the time at which the envelope function f and its derivatives are evaluated.

Appendix B. The real–imaginary representation of the transfer function

In numerical applications of the perturbation theory it is most convenient to represent the transfer function in terms of its real and imaginary parts as

$$a(\boldsymbol{r}, \omega) \equiv p(\boldsymbol{r}, \omega) + \mathrm{i}q(\boldsymbol{r}, \omega),$$

where p and q are real functions. This representation is also useful in analysing the neighbourhood of a c.w. null, where $a(\boldsymbol{r}, \omega) = 0$. Generically, at a c.w. null, p and q pass linearly through zero with respect to both $\boldsymbol{r}$ and ω, whereas the amplitude $M(\boldsymbol{r}, \omega)$ and phase $\varphi(\boldsymbol{r}, \omega)$ are both singular. However, frequency derivatives of M and φ may be 'regularized' in the neighbourhood of a c.w. null by multiplication by a suitable minimal power of M. It is essential, both numerically and analytically, to work in terms of the regularized derivatives near c.w. nulls, and the purpose of this Appendix is to find them in terms of p and q and their frequency derivatives.

We start from the identity

$$M(\boldsymbol{r}, \omega)\,\mathrm{e}^{\mathrm{i}\varphi(\boldsymbol{r}, \omega)} \equiv p(\boldsymbol{r}, \omega) + \mathrm{i}q(\boldsymbol{r}, \omega), \qquad \text{(B 1)}$$

and subsequently drop the explicit dependences on $\boldsymbol{r}$ and ω. Then

$$M^2 \equiv p^2 + q^2 \qquad \text{(B 2)}$$

is a regular function of p and q, but M alone is not. Since p and q are regular (generically linear) functions of $\boldsymbol{r}$ and ω near the null, M^2 is a regular function of $\boldsymbol{r}$ and ω, but M is not.

We now differentiate (B 1) successively with respect to frequency using primes to denote frequency derivatives, multiply by the complex conjugate of (B 1) and equate real and imaginary parts to give, up to third order,

$$MM' \equiv pp' + qq', \tag{B 3a}$$

$$M^2\varphi' \equiv pq' - qp', \tag{B 3b}$$

$$M(M'' - M\varphi'^2) \equiv pp'' + qq'', \tag{B 4a}$$

$$M(2M'\varphi' + M\varphi'') \equiv pq'' - qp'', \tag{B 4b}$$

$$M(M''' - 3M'\varphi'^2 - 3M\varphi'\varphi'') \equiv pp''' + qq''', \tag{B 5a}$$

$$M(3M''\varphi' + 3M'\varphi'' + M\varphi''' - M\varphi'^3) \equiv pq''' - qp'''. \tag{B 5b}$$

The right-hand side of each of these identities is regular at $p = q = 0$, and generically tends to zero linearly, since derivatives of p and q will not have zeros coincident with those of p and q themselves. Multiplying by M^2 preserves this regularity. The terms MM' and $M^2\varphi'$ are already in regularized form in (B 3a, b). To find the higher regularized derivatives we solve the above equations, multiplying by M^2 as necessary, to give

$$M^3M'' \equiv (p^2+q^2)(pp''+qq'') + (pq'-qp')^2, \tag{B 6a}$$

$$M^4\varphi'' \equiv (p^2+q^2)(pq''-qp'') - 2(pp'+qq')(pq'-qp'), \tag{B 6b}$$

$$M^5M''' \equiv (p^2+q^2)^2(pp'''+qq''') + 3(MM')(M^2\varphi')^2 + 3(M^2\varphi')(M^4\varphi''), \tag{B 7a}$$

$$M^6\varphi''' \equiv (p^2+q^2)^2(pq'''-qp''') - 3(M^3M'')(M^2\varphi') - 3(MM')(M^4\varphi'') + (M^2\varphi')^3. \tag{B 7b}$$

(We have left formulae (B 7) in a mixed form for simplicity.)

The regularity of these expressions is important and quite subtle. Some of them could be divided by a power of M^2 to leave a result which tended to a finite limit as p and $q \to 0$, but this limit would depend on the direction in which it was taken. Therefore the result would be singular, and the above regularized derivatives are in fact multiplied by the smallest possible powers of M. Note the uniform increase of this power thus: $M^{2n-1}M^{(n)}$ and $M^{2n}\varphi^{(n)}$.

For numerical purposes involving interpolation, it is obviously essential that the function whose values are being interpolated is regular, otherwise wildly erroneous results may be produced.

Appendix C. Regularity of the dislocation reference time $T_1(\boldsymbol{r})$

The quantity $T_1(\boldsymbol{r})$ which we have called the 'dislocation reference time' is given by (75b) as

$$T_1 = \varphi_0' + \operatorname{sgn}(\varphi_0''')\sqrt{(M_0''/M_0)}, \tag{C 1}$$

being defined only on the surfaces $\mathscr{S}$ satisfying $M_0' = 0$, $M_0'' > 0$. We have seen that these surfaces contain lines on which $M_0 = 0$, and there the individual terms on the right of (C 1) are singular. However, we now show that their sum is regular within $\mathscr{S}$, so that our first-order perturbation formulae (75) remain usable near a c.w. null, when properly expressed.

To do so we re-express (C 1) in terms of the real–imaginary (p, q) representation using Appendix B. First consider sgn (φ_0'''). Setting $M_0' = 0$ in (B 7b) gives

$$M_0^6\varphi_0''' = (p_0^2+q_0^2)^2\,(p_0q_0''' - q_0p_0''') - 3(M_0^3M_0'')\,(M_0^2\varphi_0') + (M_0^2\varphi_0')^3. \tag{C 2}$$

Let δ represent distance in $\mathscr{S}$ from a c.w. null. Then generically $p_0 = O(\delta)$, $q_0 = O(\delta)$, and hence $M_0 = O(\delta)$. From (B 6a) and (B 3b)

$$M_0^3M_0'' = (p_0q_0' - q_0p_0')^2 + O(\delta^3) = (M_0^2\varphi_0')^2 + O(\delta^3);$$

so that from (C 2)

$$M_0^6\varphi_0''' = -2(M_0^2\varphi_0')^3 + O(\delta^4),$$

since from (B 3b) $M_0^2\varphi_0' = O(\delta)$. Therefore

$$\text{sgn}\,(\varphi_0''') \sim -\text{sgn}\,(\varphi_0') = -\text{sgn}\,(p_0q_0' - q_0p_0')$$

as $M_0 \to 0$.

Using this result in (C 1) gives, as $M_0 \to 0$,

$$\begin{aligned} T_1 &\sim \frac{M_0^2\varphi_0'}{M_0^2} - \text{sgn}\,(\varphi_0')\frac{\sqrt{(M_0^3M_0'')}}{M_0^2} \\ &= \frac{p_0q_0' - q_0p_0'}{p_0^2+q_0^2} - \frac{\text{sgn}\,(p_0q_0' - q_0p_0')}{p_0^2+q_0^2}\{(p_0q_0' - q_0p_0')^2 + (p_0^2+q_0^2)\,(p_0p_0'' + q_0q_0'')\}^{\frac{1}{2}} \\ &= \frac{p_0q_0' - q_0p_0'}{p_0^2+q_0^2} - \frac{p_0q_0' - q_0p_0'}{p_0^2+q_0^2}\left\{1 + \frac{(p_0^2+q_0^2)\,(p_0p_0'' + q_0q_0'')}{(p_0q_0' - q_0p_0')^2}\right\}^{\frac{1}{2}}. \end{aligned} \tag{C 3}$$

The orders of the terms are given by

$$T_1 = O(\delta^{-1}) - O(\delta^{-1})\{1 + O(\delta)\}^{\frac{1}{2}}$$

and the two divergences exactly cancel, leaving the finite result $T_1 = O(1)$ as $M_0 \to 0$.

Specifically, since $p_0p_0' + q_0q_0' = 0$ in $\mathscr{S}$,

$$T_1 \sim -\frac{1}{2}\frac{p_0p_0'' + q_0q_0''}{p_0q_0' - q_0p_0'} = \frac{1}{2}\frac{p_0'q_0'' - q_0'p_0''}{p_0'^2 + q_0'^2} \tag{C 4}$$

as $M_0 \to 0$. Higher-order terms in the expansion of (C 3) are clearly regular in p_0 and q_0, and hence T_1 is regular in $\boldsymbol{r}$ near a c.w. null, as required.

We now see that we can avoid using third derivatives for any trajectory which, to lowest order, passes through a c.w. null, simply by choosing the sign of the square root so as to make T_1 regular within $\mathscr{S}$ through the null.

Finally we remark that, if φ_0''' were to change sign other than when $M_0 = 0$ or $M_0'' = 0$, T_1 would suffer a discontinuity, but this does not happen in any of the models so far studied. It appears that φ_0''' is related to M_0'', just as φ_0'' is apparently related to M_0'. These connections are under investigation.

Appendix D. Computational details

So far, computations have only been performed for problems in two space dimensions. These arise from three-dimensional problems which have translational or rotational symmetry, or simply from taking a plane section through a general system. A rectangular grid is set up over the region of interest. At each grid point the transfer function $a(\boldsymbol{r}, \omega_0)$ and its first three frequency

derivatives are evaluated at the pulse carrier frequency. Most numerical processes involve interpolation, and only regular quantities can be reliably interpolated by polynomials in the standard way. In practice, best results are obtained by interpolating $a(\boldsymbol{r}, \omega_0)$ and its derivatives directly where possible. The best way to display the behaviour of M_0, φ_0 and their derivatives is to plot contours, by using the regularized derivatives of Appendix B.

The first task is to find the dislocation trajectories. An array of values of $M_0 M_0'$ evaluated at each grid point is set up, and its zero-level contour traced. An excellent contouring routine is that written by Heap (1974); this searches for a contour and then outputs sequentially the points where it crosses the sides (and diagonals, in the more sophisticated version, which we recommend) of grid rectangles. Each such point is a potential trajectory point. The contouring algorithm uses linear interpolation, and it is easy to modify the code to interpolate simultaneously one or more subsidiary arrays. In this way values of p_0, q_0, p_0'... are interpolated at each potential trajectory point. The value of $M_0^3 M_0''$ is calculated, and if it is positive the algorithm proceeds; otherwise it seeks the next potential trajectory point. The values of $M_0^2 \varphi_0'$ and $M_0^2 + E$ are evaluated, where E is a very small constant used to avoid division by zero at c.w. nulls, and sgn (φ_0''') is evaluated from $M_0^6 \varphi_0'''$ or by the method discussed in Appendix C. From these we evaluate

$$T_1 = \{M_0^2 \phi_0' + \mathrm{sgn}\,(\varphi_0''') \sqrt{(M_0^3 M_0'')}\}/(M_0^2 + E).$$

It is essential that the computer performs the addition before the division, so that the limiting value of T_1 at (or near) a c.w. null is found correctly. Finally the dislocation time t is computed from (75b).

The trajectories, which generally fall into a number of disconnected curves, are plotted as they are traced, and the sequence of values of T_1 and t stored. Later separate graphs for each disconnected trajectory are plotted showing times as functions of distance along the trajectory.

References

Arnol'd, V. I. 1975 *Russ. math. Surv.* **30** (5), 1–75.
Berry, M. V. 1976 *Adv. Phys.* **25**, 1–26.
Berry, M. V., Nye, J. F. & Wright, F. J. 1979 *Phil. Trans. R. Soc. Lond.* A **291**, 453–484.
Dingle, R. B. 1973 *Asymptotic expansions: their derivation and interpretation*. London: Academic Press.
Faulkner, E. A. 1969 *Introduction to the theory of linear systems*. London: Chapman and Hall.
Heap, B. R. 1974 *National Physical Laboratory Report* NAC 47.
Nye, J. F. 1981 *Proc. R. Soc. Lond.* A **378**, 219–239.
Nye, J. F. & Berry, M. V. 1974 *Proc. R. Soc. Lond.* A **336**, 165–190.
Pearcey, T. 1946 *Phil. Mag.* **37**, 311–317.
Pippard, A. B. 1978 *The physics of vibration*, vol. 1. Cambridge University Press.
Thom, R. 1972 *Stabilité structurelle et morphogénèse*. Reading, Massachusetts: Benjamin. (English translation: *Structural stability and morphogenesis*. Reading, Massachusetts: Benjamin (1975).)
Wright, F. J. 1977 Wavefield singularities. Ph.D. thesis, Bristol University.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES