ADVANCES IN PHYSICS, 1976, VOL. 25, NO. 1, 1–26

Waves and Thom's theorem

By M. V. BERRY
H. H. Wills Physics Laboratory, Bristol BS8 1TL, U.K.

[Received 30 December 1975]

ABSTRACT

Short-wave fields can be well approximated by families of trajectories. These families are dominated by their singularities, i.e. by caustics, where the density of trajectories is infinite. Thom's theorem on singularities of mappings can be rigorously applied and shows that structurally stable caustics—that is those whose topology is unaltered by 'generic' perturbation—can be classified as 'elementary catastrophes'. Accurate asymptotic approximations to wave functions can be built up using the catastrophes as skeletons: to each catastrophe there corresponds a canonical diffraction function. Structurally unstable caustics can be produced by special symmetries, and the detailed form of the caustic that results from symmetry-breaking can often be determined by identifying the structurally unstable caustic as the special section of a higher-dimensional catastrophe. Sometimes it is clear that the unstable caustic is a special section of a catastrophe of infinite co-dimension; these fall outside the scope of Thom's theorem and suggest new directions for mathematical investigation. The discussion is illustrated with numerous examples from optics and quantum mechanics.

CONTENTS

§ 1. INTRODUCTION

Underlying wave theories are their short-wave limits in the form of Hamiltonian descriptions involving families of trajectories. In the case of quantum mechanics these trajectories are the paths of classical particles, while for electromagnetism they are the rays of geometrical optics. A wave field corresponds to a *family* of trajectories rather than a single trajectory. The family is defined by an action function S from which its individual trajectories can be derived; in isotropic media (where the Hamiltonian function depends on the canonical momentum $\mathbf{p}$ only through its length $|\mathbf{p}|$) the trajectories are normal to the contour surfaces of S, i.e. to the wavefronts. General reviews of Hamiltonian theories have been given by Synge (1954, 1960) and Goldstein (1951).

A family can, and usually does, exhibit a property that does not reside in any of its individual trajectories, namely *focusing*. This occurs on the contact of neighbouring trajectories, i.e. on their *envelope*. The general term for a focal region is a *caustic*, and it is caustics that form the subject of this paper.

Caustics are doubly important in wave theory. In the first place, the intensity, that is the density of trajectories, is infinite on caustics. Therefore under short-wave conditions, where the trajectory picture is a good approximation, wave fields will be dominated by caustics, which will form the significant structure of images formed on screens or photographic plates placed in the field. This is well expressed by Stavroudis (1972):

"The caustic is one of the few things in geometrical optics that has any physical reality. Wavefronts and rays are not realizable; they are just convenient symbols on which we can hang our ideas. The caustic on the other hand is real and becomes visible by blowing a cloud of smoke in the region of the focus of a lens."

In the second place, it is precisely on caustics that the elementary trajectory picture must break down because solutions of wave equations always have finite intensity provided the wavelength is not exactly zero. Therefore caustics are 'windows' through which wave effects can still be discerned when the wavelength has become so short that interference fringes, etc. are too fine to be seen elsewhere in the wave field. It has been known for some time (Berry 1966, 1969 a, b, Berry and Mount 1972) that refined trajectory pictures can be devised for various types of caustic giving short-wave asymptotic approximations to wave fields, with the property that they are valid *uniformly* through the caustic regions.

Until recently, however, it was not at all clear what forms caustics could take. Now, however, the celebrated theorem of Thom (Wassermann 1974, Bröcker 1975) enables certain caustics—those which are 'structurally stable'—to be classified according to their topology. Thom's theorem is a piece of pure mathematics; it also forms the basis of a more speculative system of ideas called 'catastrophe theory' (Thom 1975). The applicability of the theorem to certain caustics of families of trajectories has been proved by Jänich (1974), and the derivation of uniform asymptotic approximations to wave fields in the presence of these caustics has been put on a rigorous basis by Maslov (1972) (see also Kravtsov (1968) and Duistermaat (1974)). There is no need to repeat these derivations here.

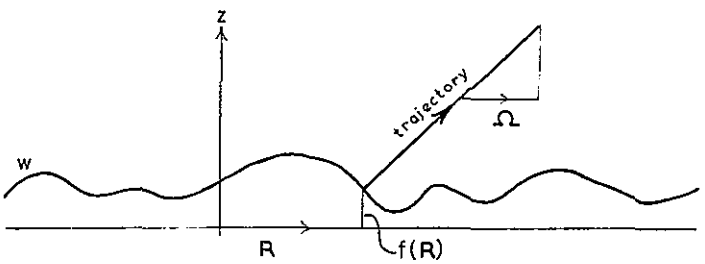
The purpose of this article is twofold. First, we explain how Thom's theorem can be used descriptively and predictively to solve problems in wave physics. Secondly, we point out some important classes of 'non-generic' caustic which lie outside the scope of Thom's theorem. A fully general treatment of all these subjects does not yet exist. Therefore throughout most of this paper we employ for illustrative purposes a model system which is very simple but which nevertheless comprehends a variety of non-trivial wave phenomena; this model is introduced in § 2. In § 3 we state Thom's theorem. In § 4 we give a series of examples of caustics that are 'elementary catastrophes', and also show how the theorem can predict the form of caustics resulting from certain kinds of symmetry-breaking. Section 5 is devoted to wave functions: we show how uniform asymptotic approximations can be

obtained in terms of canonical ' diffraction functions ' which can be derived in simple generic cases from Thom's catastrophes. Finally, in § 6 we investigate some caustics that are catastrophes of infinite co-dimension, resulting from high symmetries whose mathematical implications are not yet understood.

§ 2. Model system : the far field from a perturbed plane wavefront

Consider a monochromatic plane wave travelling along the direction $+z$ in a homogeneous isotropic medium. Let the wavefront at $z=0$ be deformed into a surface W by lifting the part at $\mathbf{R}\equiv(x, y)$ from $z=0$ to $z=f(\mathbf{R})$ (fig. 1). We assume that all radii of curvature of W are large in comparison with the wavelength λ, and that all slopes are small (i.e. $|\nabla f|\ll 1$) ; this does not mean that $f(\mathbf{R})$ is small, and indeed we are particularly interested in cases where f/λ can be large so that a trajectory description is valid. The restriction on slopes is not necessary but greatly simplifies the algebra.

Fig. 1



Notation for initial wavefront W and trajectories.

A variety of smooth phase objects will produce such wavefronts, and we give a few examples : (i) Undulating refracting surfaces such as the panes of glass used for bathroom windows. If n is the refractive index, then to produce W by transmission the profile of the undulating surface must be $f(\mathbf{R})n/n-1$. (ii) Irregular water droplet ' lenses ' on inhomogeneously dirty flat glass surfaces. Again the profile is $f(\mathbf{R})n/n-1$, where n is now the refractive index of water. For small droplets the internal pressure p is constant, and $f(\mathbf{R})$ is restricted by the surface tension γ to obey the equation

$$\nabla^2 f(\mathbf{R}) = -\frac{(n-1)p}{n\gamma}. \tag{1}$$

This has the general solution

$$f(\mathbf{R}) = \mathrm{Re}\, g(x+iy) - \frac{(n-1)p}{4n\gamma} R^2, \tag{2}$$

where g is any analytic function and $R\equiv|\mathbf{R}|$. The optics of such droplets have been briefly described by Minnaert (1954) and Larmor (1891) (see also the night scene in Fellini 1963). The caustics can be clearly seen by viewing a distant street lamp with the eye close behind small raindrops on a window pane. (iii) Undulating reflecting surfaces ($n=-1$), such as crystal surfaces

illuminated by beams of atoms (Berry 1975). (iv) Scattering of a beam of particles of mass m and (high) energy E by a spherically symmetric field of force with potential $V(r)$, where the wavefront beyond the scatterer is given by

$$f(R) = \int_R^\infty dr \frac{r}{\sqrt{(r^2 - R^2)}} \frac{V(r)}{E} \tag{3}$$

which, being cylindrically symmetric, involves only R. This type of scattering is discussed by Glauber (1958).

We discuss only the *far field* of the waves from W, to which only the *directions* of the trajectories and waves contribute. We specify directions by projections $\boldsymbol{\Omega}$ on the xy plane (fig. 1), i.e. by $\Omega_x = \sin\theta\cos\phi$, $\Omega_y = \sin\theta\sin\phi$, where θ and ϕ are the usual polar angles with z as axis. The intensity $I(\boldsymbol{\Omega})$ is defined by the flux through $d\boldsymbol{\Omega}$ far from W when there is unit flux through unit area of W itself. (Flux is energy flow in optics, particle flow in mechanics.)

Each point $\mathbf{R}$ on W gives rise to a trajectory normal to W, whose direction is

$$\boldsymbol{\Omega}(\mathbf{R}) = -\nabla f(\mathbf{R}). \tag{4}$$

For given $\boldsymbol{\Omega}$ there may be several points $\mathbf{R}$ satisfying this equation ; we label them $\mathbf{R}_i(\boldsymbol{\Omega})$. Then the intensity I on the trajectory picture is

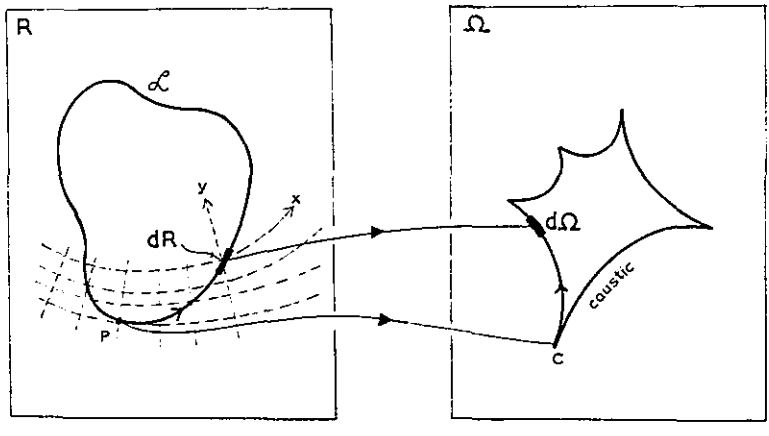
$$I(\boldsymbol{\Omega}) = \sum_i \left| \frac{d\boldsymbol{\Omega}}{d\mathbf{R}_i(\boldsymbol{\Omega})} \right|^{-1}, \tag{5}$$

where $|d\boldsymbol{\Omega}/d\mathbf{R}|$ is the Jacobian of the *mapping* from $\mathbf{R}$ to $\boldsymbol{\Omega}$ that each trajectory defines in accordance with eqn. (4). Explicitly,

$$\left| \frac{d\boldsymbol{\Omega}}{d\mathbf{R}} \right| = \left| \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 \right| \equiv |K(\mathbf{R})|, \tag{6}$$

where $K(\mathbf{R})$ is the *Gaussian curvature* of W at $\mathbf{R}$.

Fig. 2



Caustic as image of line $\mathscr{L}$ where Gaussian curvature of W vanishes.

The caustics of the family of trajectories travelling to infinity from W are the *singularities* of the mapping $\mathbf{R}\rightarrow\boldsymbol{\Omega}$; they occur where $|d\boldsymbol{\Omega}/d\mathbf{R}|$ vanishes and I is infinite. Therefore caustics are the images in $\boldsymbol{\Omega}$ of the lines $\mathscr{L}$ on W where K vanishes (fig. 2). These directional caustics at infinity are of course the asymptotes of the more complicated spatial caustics beyond W. Thom's theorem, to be discussed in § 3, tells us about the forms the caustics may take.

Crucial to the application of Thom's theorem is the fact that the trajectories (4) can be derived from a *generating function* $\Phi(\mathbf{R}, \boldsymbol{\Omega})$, of the form

$$\Phi(\mathbf{R}, \boldsymbol{\Omega}) = \boldsymbol{\Omega} \cdot \mathbf{R} + f(\mathbf{R}), \tag{7}$$

by the gradient conditions

$$\nabla_{\mathbf{R}}\phi = 0. \tag{8}$$

The generating function Φ is a single-valued non-singular function of its arguments $\mathbf{R}$ and $\boldsymbol{\Omega}$, whereas the *action function* of the family of trajectories, namely

$$S_i(\boldsymbol{\Omega}) = \Phi(\mathbf{R}_i(\boldsymbol{\Omega}), \boldsymbol{\Omega}), \tag{9}$$

is a many-valued function of $\boldsymbol{\Omega}$ because there may be many points $\mathbf{R}_i$ sending trajectories to $\boldsymbol{\Omega}$. The different 'branches' S_i join on the caustic. That (9) is indeed the 'directional action' can easily be verified by calculating the optical distance along a trajectory from W to some finite point and then letting the point recede to infinity in direction $\boldsymbol{\Omega}$.

The existence of a generating function is not confined to our simple example; Jänich (1974) and Maslov (1972) (see also Duistermaat (1974)) show that all Hamiltonian families of trajectories can be derived from generating functions. Introducing the language of catastrophe theory, we say that $\boldsymbol{\Omega}$ are the *control parameters*, on which the observable intensity I depends. In the general case $\boldsymbol{\Omega}$ need not represent direction; it could include position in space, time, parameters affecting the shape of scattering potentials or diffracting objects, the refractive index of a medium, etc. The auxiliary quantities $\mathbf{R}$, on which I does not depend, are called the *state variables*. The vectors $\mathbf{R}$ and $\boldsymbol{\Omega}$ may have any number of components. It is usually not difficult to find the generating function for a particular situation and there are general techniques (described by the authors cited) involving the introduction of variables canonically conjugate to those in which the caustics appear, the generating function being the solution of an appropriate Hamilton–Jacobi equation in mixed coordinate-momentum representation (see also Kravtsov 1968).

On the *wave theory* the intensity for the simple model system can be found to a close approximation by methods described by Berry (1975). The result (which is almost obvious) is the Fraunhofer diffraction integral

$$I(\boldsymbol{\Omega}) = |\psi(\boldsymbol{\Omega})|^2 = \left| \frac{k}{2\pi} \iint d\mathbf{R} \exp\left[ik\Phi(\mathbf{R}, \boldsymbol{\Omega})\right] \right|^2, \tag{10}$$

where ψ is the wave amplitude, Φ is given by eqn. (7), and $k \equiv 2\pi/\lambda$.

In § 5 we shall show how to evaluate diffraction integrals of this type. At this point we note that the exponent in the integrand is proportional to precisely the function Φ just introduced to generate the trajectories. Maslov (1972) shows that this is a general feature of integrals that are asymptotic approximations (valid as $k \to \infty$) to wave functions. (For an example of such an integral representation in scattering theory see § 6 of Berry and Mount (1972).)

§ 3. Thom's theorem

Caustics, then, are singularities of gradient maps of the form (8) derived from generating functions $\Phi(\mathbf{R}, \boldsymbol{\Omega})$. Thom's theorem concerns caustics with the property of *structural stability*. This means that a perturbation (e.g. of the Hamiltonian or of the initial wavefront W) leaves the local structure of the caustic unchanged in the sense that the perturbed and unperturbed caustics are related by a smooth reversible transformation (a 'diffeomorphism'). A situation in which the caustics are structurally stable is 'generic'—there is 'nothing special' about it. A structurally unstable caustic, on the other hand, whose topological type is changed by perturbation, results from a 'non-generic' situation where some special symmetry exists.

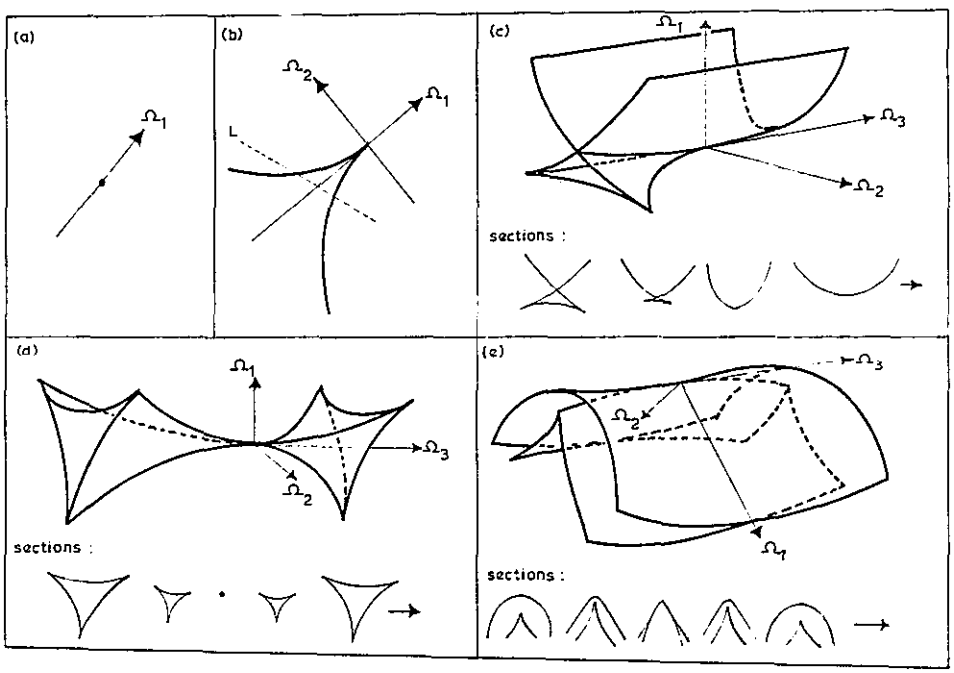
The theorem states (Wassermann 1974, Bröcker 1975) that there exists only a finite number of types of structurally stable caustic for each value of the dimensionality n of the control parameter space $\boldsymbol{\Omega}$, and it gives explicit standard forms for the generating functions when n is less than 7. n is called the *co-dimension* of the caustic. Thom calls the standard caustics the 'elementary catastrophes'. It is remarkable that for fixed n the structurally stable caustics are not affected by adding extra state variables $\mathbf{R}$; depending on the caustic type, only one component x, or two components x, y, of $\mathbf{R}$, determine the caustic, and any other components can always be transformed so as to enter $\Phi(\mathbf{R}, \boldsymbol{\Omega})$ quadratically, and then (cf. eqn. (8)) they cannot produce singularities.

Structurally stable caustics—'elementary catastrophes'—with co-dimensions 1, 2 and 3. x and y are components of the state vector $\mathbf{R}$ and Ω_1, Ω_2 and Ω_3 are control parameters Ω.

Co-dimension	Name	Generating function $\Phi(\mathbf{R}, \boldsymbol{\Omega})$	Singularity index σ
1	Fold	$x^3/3 + \Omega_1 x$	$\frac{1}{6}$
2	Cusp	$x^4/4 + \Omega_1 x^2/2 + \Omega_2 x$	$\frac{1}{4}$
3	Swallow-tail	$x^5/5 + \Omega_3 x^3/3 + \Omega_2 x^2/2 + \Omega_1 x$	$\frac{3}{10}$
3	Elliptic umbilic	$x^3 - 3xy^2 - \Omega_3(x^2+y^2) - \Omega_1 x - \Omega_2 y$	$\frac{1}{3}$
3	Hyperbolic umbilic	$x^3 + y^3 + \Omega_3 xy - \Omega_1 x - \Omega_2 y$	$\frac{1}{3}$

The table gives the names and generating functions of the structurally stable caustics for $n \leqslant 3$, and on fig. 3 are sketched the forms of these caustics in Ω space. What is peculiar to each catastrophe is the local nature of its highest singularity. Consider, for example, the cusp (fig. 3 *b*); a generic section such as the randomly chosen line $\mathscr{L}$ will have at most the 'fold' point catastrophes characteristic of one control dimension, and this will be structurally stable. However, a non-generic section, passing through the

Fig. 3

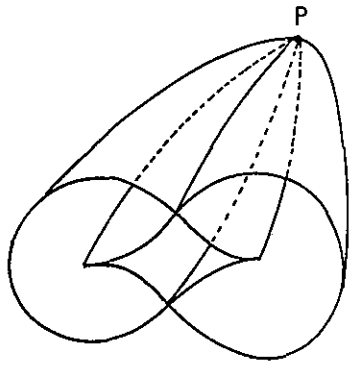


Structurally stable caustics with co-dimension $n \leqslant 3$: (*a*) fold, (*b*) cusp, (*c*) swallow-tail, (*d*) elliptic umbilic, (*e*) hyperbolic umbilic. These are the 'elementary catastrophes'.

cusp point $\Omega_1 = \Omega_2 = 0$, is unstable because if we move it to the right the caustic point vanishes, and if we move it to the left the caustic point splits into two. In the three-dimensional catastrophes the non-generic sections $\Omega_3 = 0$ are two-dimensional and contain not the structurally stable cusps and folds but either an isolated point (elliptic umbilic), a finite-angled corner (hyperbolic umbilic), or a point where the caustic curvature diverges smoothly (swallow-tail), all three of these points being structurally unstable. The pattern of development from instability to stability as the section moves from non-generic to generic is called the *smooth unfolding of the singularity*. As n increases, so does the number of catastrophes, a fact appreciated as early as 1873 by Maxwell (Campbell and Garnett 1882).

We emphasize that the standard forms of the table describe only the *local* caustic structure, and leave the *global* topology completely unspecified. This is illustrated in fig. 4, which shows a hyperbolic umbilic (singularity at P) globally different from fig. 3 (*e*).

Fig. 4



Hyperbolic umbilic with global topology different from fig. 3 (*e*) (after Zeeman).

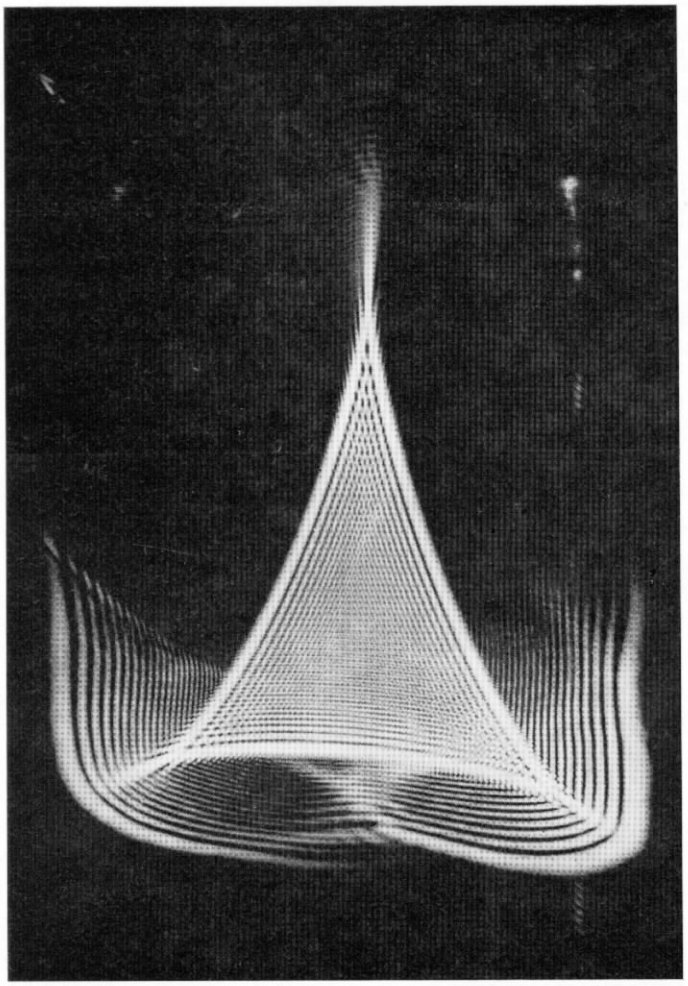
Armed with Thom's theorem, one can quite easily recognize non-generic caustics. Their appearance always indicates the presence of some symmetry. If the symmetry is broken generically by parameters ϵ, these may be added to the control parameters $\boldsymbol{\Omega}$. If $(\boldsymbol{\Omega}, \epsilon)$ is of sufficiently small dimensionality, the now structurally stable unfolding of the caustic is an elementary catastrophe. This ability to predict the details of symmetry-breaking is a very powerful feature of Thom's theorem ; we shall see an application of it in § 4. Often, however, ϵ is of infinite dimension—there are infinitely many topologically different ways of breaking the symmetry—and Thom's theorem is inapplicable ; we shall give several examples of this in § 6.

§ 4. Caustics of finite co-dimension

For the simple model system of § 2, the caustics appear in the two-dimensional space of directions $\boldsymbol{\Omega}$, as lines on a distant screen whose Cartesian coordinates are proportional to the control parameters $\Omega_1 = \Omega_x$ and $\Omega_2 = \Omega_y$. Reference to the table shows that the only structurally stable caustics are smooth fold lines which may be interrupted only at cusp points where the caustic curve reverses direction and crosses its tangent (i.e. these caustics may not come to an end or have finite-angled corners or be isolated points). To verify this experimentally, simply shine a laser beam through a generic object (e.g. examples (i) or (ii) of § 2—the irregular glass called ' Atlantic ' made by Pilkington's is excellent for this purpose as fig. 5 shows) and let the image fall on a light-coloured wall in a darkened room ; the size of the irregularities of the objects should be about the same as the width of the beam. It is difficult to lay down conditions guaranteeing genericity ; usually it is sufficient for no special circumstances to have attended the manufacture of the refracting objects. We shall amplify this point later.

How do folds and cusps arise in this model system ? Recall that caustics in $\boldsymbol{\Omega}$ are images under the mapping (4) of the line $\mathscr{L}$ on W where the Gaussian curvature K vanishes (fig. 2). $K=0$ implies that at least one of the two principal curvatures of W vanishes at each point on $\mathscr{L}$. Set up local Cartesian coordinates $\mathbf{R} = \mathbf{i}x + \mathbf{j}y$ whose origin is at 0 on $\mathscr{L}$, and let $\mathbf{i}$ be locally parallel to the principal direction corresponding to that curvature which vanishes. Then $\partial^2 f/\partial x^2$ and $\partial^2 f/\partial x \partial y$ both vanish at 0, and the line element $d\boldsymbol{\Omega}$ of caustic

Fig. 5



Far-field caustic from irregular window pane of Pilkington's 'Atlantic' glass, showing partially unfolded elliptic and hyperbolic umbilics.

resulting from the line element $d\mathbf{R} = \mathbf{i}\,dx + \mathbf{j}\,dy$ on $\mathscr{L}$ near 0 is easily calculated from (4) to be

$$d\boldsymbol{\Omega} = -\mathbf{i}\,\frac{\partial^2 f}{\partial y^2}\,dy. \tag{11}$$

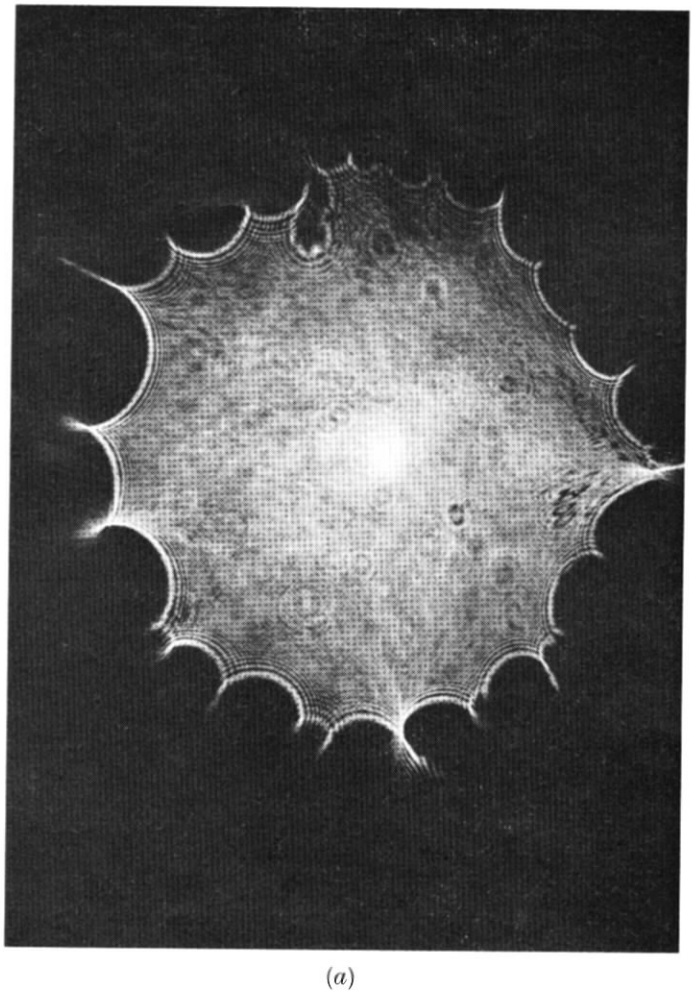
Thus the caustic is locally perpendicular to the direction on W of the principal curvature whose vanishing generates it; this is a special case of a result proved in a text by Eisenhart (1960).

As 0 moves round $\mathscr{L}$, the caustic is generated by adding all the $d\boldsymbol{\Omega}$ given by (11). The caustic will be smooth (i.e. a fold catastrophe) at almost all points. The exceptions are the cusps (C on fig. 2) which arise from points P (fig. 2) where $\mathscr{L}$ touches the direction $\mathbf{i}$ of vanishing curvature, since there $dy = 0$ and (11) shows that $d\boldsymbol{\Omega}$ reverses direction.

It is quite easy to produce caustics with no cusps. For example, choose the shape of W to be a cylindrically symmetric Gaussian hill. Then $\mathscr{L}$ is the circle corresponding to the radius R at which the profile inflects. The direction of vanishing curvature is radial, and therefore never touches $\mathscr{L}$, so that there are no cusps. The caustic is a circle in $\boldsymbol{\Omega}$. For case (iv) of § 2 this example corresponds precisely to potential scattering of 'rainbow' type (Ford and Wheeler 1959). The optical rainbow is produced by reflections as well as refractions (Descartes 1637) but it too is a circular directional caustic, and as such is the simplest natural example of an elementary catastrophe. (An end-point on a smooth curve is not a two-dimensional catastrophe, so that if the crock of gold exists it must be structurally unstable.)

For irregular droplet lenses (case (ii) of § 2), on the other hand, we can show that the caustics (fig. 6) *must* have cusps. First, recall that the profiles

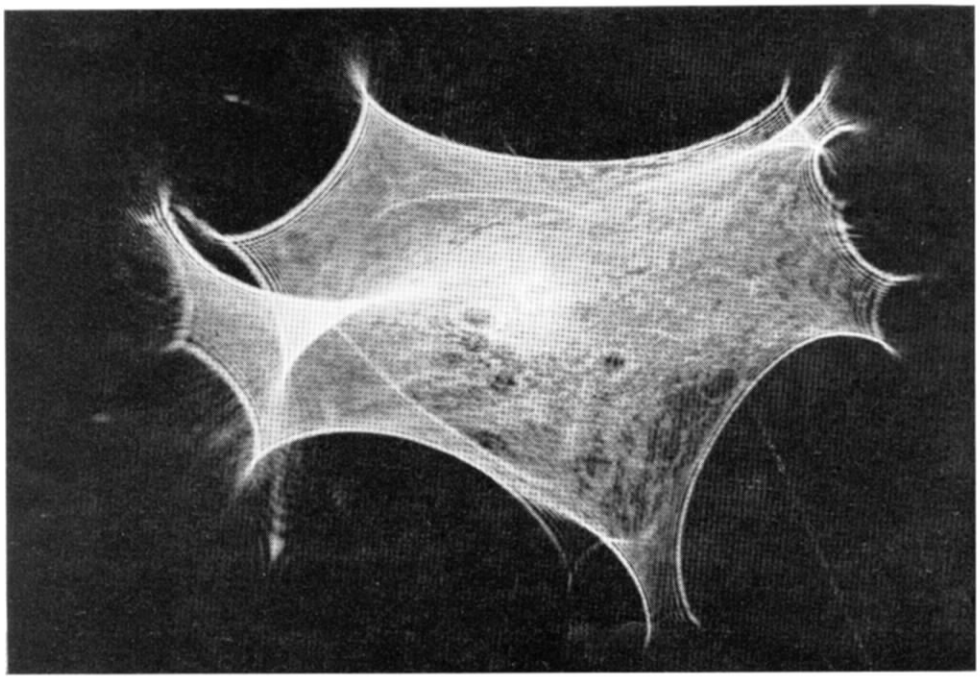
Fig. 6



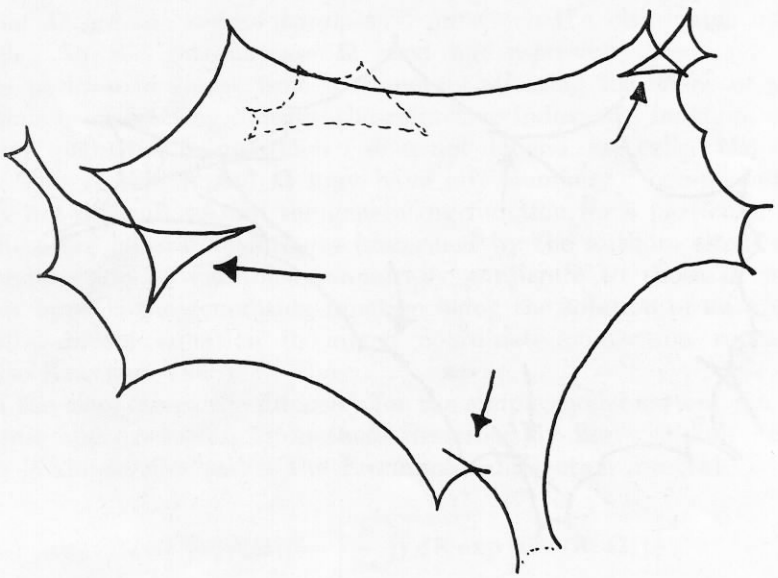
(*a*)

Far-field caustics from irregular water droplet on flat glass plate. (*c*) and (*e*) are tracings of (*b*) and (*d*). Swallow-tails are marked with arrows.

Fig. 6 (*continued*)

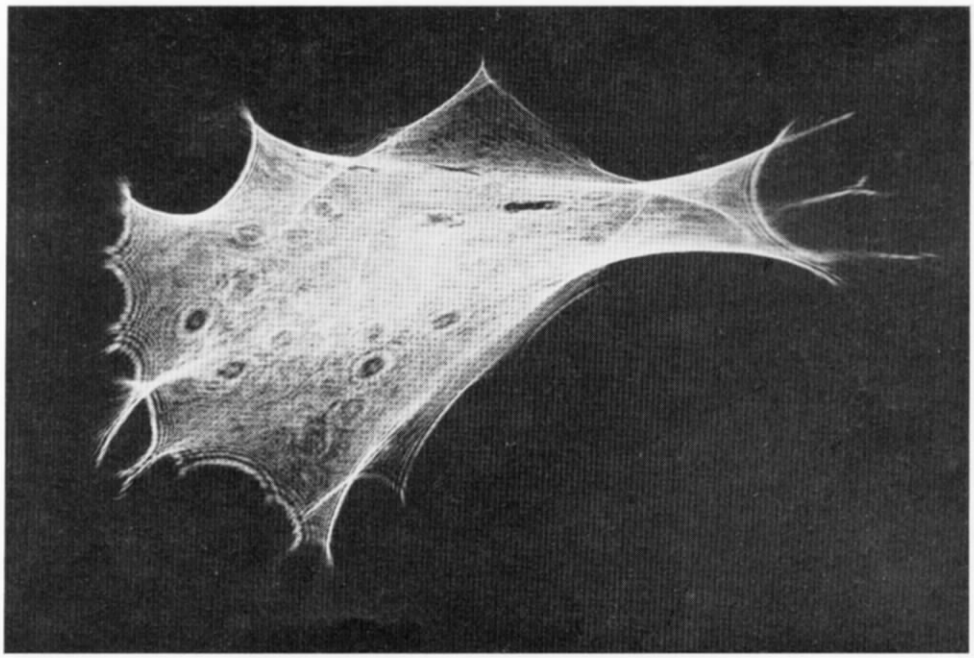


(*b*)

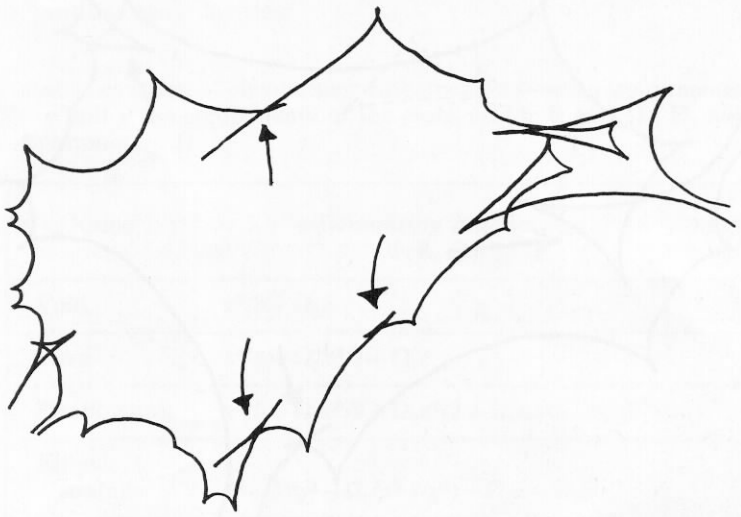


(*c*)

Fig. 6 (*continued*)



(*d*)



(*e*)

$f(\mathbf{R})$ are limited by surface tension to have the form of eqn. (2). Combined with (6) this gives, as the equation of the line $\mathscr{L}$,

$$|g''(\zeta)| = \frac{(n-1)p}{2n\gamma}, \tag{12}$$

where $\zeta = x + iy$, x and y being now global and not local Cartesians. Therefore we may write g'' on $\mathscr{L}$ as

$$g''(\zeta) = \frac{(n-1)p}{2n\gamma} \exp\,[i\chi(\zeta)]. \tag{13}$$

During one circuit of $\mathscr{L}$ the phase χ increases monotonically (Titchmarsh 1949), by an amount

$$\Delta\chi \equiv \oint_{\mathscr{L}} d\chi = \oint_{\mathscr{L}} d(\mathrm{Im}\,\ln g'') = \oint_{\mathscr{L}} \frac{d\zeta g'''}{g''} = 2m\pi, \tag{14}$$

where m is the number of zeros of g'' inside $\mathscr{L}$, non-simple zeros being counted according to their multiplicity.

Next we require the angle ψ made with the x-axis by the direction of the vanishing curvature. From (2), the curvature C_ψ of the line in W with direction ψ is easily found to be

$$C_\psi = |g''| \cos\,(2\psi + \chi) - \frac{(n-1)p}{2n\gamma}. \tag{15}$$

The two principal directions are the values of ψ for which this expression is extremal, and the direction of vanishing curvature (for which (12) holds) is

$$\psi = -\frac{\chi}{2}. \tag{16}$$

The difference between this angle and the angle ϕ made by $\mathscr{L}$ with the x-axis is

$$\phi - \psi = \phi + \frac{\chi}{2} \tag{17}$$

and cusps occur whenever $\phi - \psi$ passes through a multiple of π. Now, in one circuit of $\mathscr{L}$, ϕ changes by 2π and $\chi/2$ by $m\pi$. Therefore if $\phi + \chi/2$ is monotonic round $\mathscr{L}$ there are $m+2$ cusps. If $\phi + \chi/2$ were not monotonic, a rotation of coordinates in $\mathbf{R}$ would change ϕ but not χ and hence the number of times $\phi + \chi/2$ passed through a multiple of π would be altered. But the number of cusps is invariant under rotation of coordinates, and so, by contradiction, $\phi + \chi/2$ must be monotonic.

Finally, we observe that m must exceed zero for non-trivial cases, since any analytic function $g''(\zeta)$ with no zeros inside $\mathscr{L}$ must be a constant and this corresponds to droplets which are caps of ellipsoids (cf. eqn. (2)) which give no far-field caustics. Therefore we have shown that the caustic of the droplet defined by $g(\zeta)$ has two more cusps than $g''(\zeta)$ has zeros in $\mathscr{L}$. Thus the smallest possible number of cusps is three. It is possible to prove that

the caustics from these surface tension dominated droplets must always be concave outwards (fig. 6); occasionally apparent exceptions to this are observed in the form of caustics with points of inflexion, but the corresponding drops are always larger and visibly affected by gravity so that the theory based on eqn. (1) does not apply.

A family of model droplets illustrating these properties is that generated by

$$g(\zeta) = \zeta^n \tag{18}$$

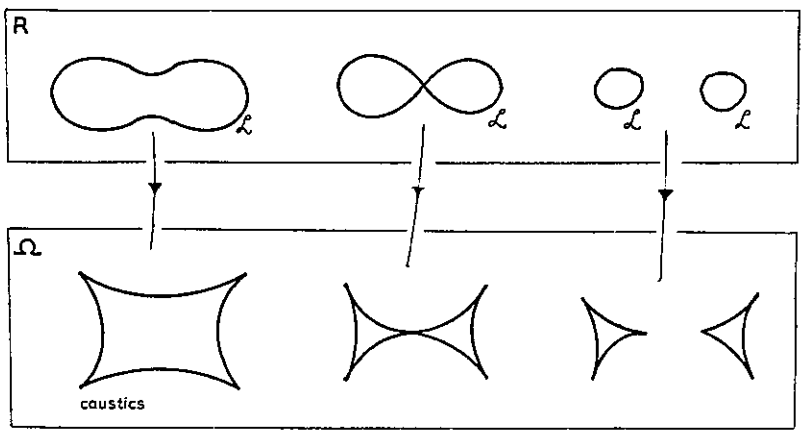
where $n = 3, 4, 5 \ldots$, $\mathscr{L}$ is a circle, and the caustic is a hypocycloid of n cusps (Lawrence 1972). By comparing an observed caustic with n cusps with one of these 'standard caustics', it might be possible to estimate the variation of the contact angle around the droplet perimeter, and hence obtain information about inhomogeneity of wetting (Dr. E. Zichy of I.C.I. is investigating this). The droplets $f(\mathbf{R})$ corresponding to (18) have fluted maxima. Their caustics should form the basis of a new class of 'hyperelliptic' catastrophes (the elliptic umbilic corresponds to $n = 3$).

An interesting possibility not covered by eqn. (18) occurs when the m zeros of g'' are isolated instead of coincident. Then, if g'' is a polynomial function of ζ, $\mathscr{L}$ will enclose all the zeros for sufficiently large p (eqn. (12)) and there will be $m+2$ cusps on the caustic. For very small p, on the other hand, $\mathscr{L}$ consists of m disconnected small closed curves, one round each zero, and there will be m closed caustics each with three cusps. Therefore there must in general be $m-1$ intermediate values of p at which $\mathscr{L}$ separates and two new cusps are born. The cusps appear in the manner shown in fig. 7, which is drawn for the droplet generated by

$$g''(\zeta) = 1 - \zeta^2; \tag{19}$$

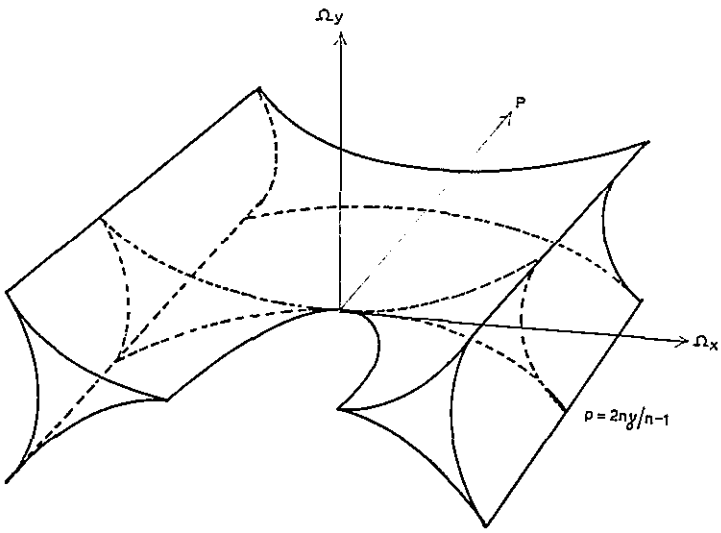
this can have a maximum at $\mathbf{R} = 0$ that splits into two as p diminishes through $2n\gamma/n - 1$. We can regard p as a control parameter additional to Ω_x and Ω_y. However, the birth of two cusps in this fashion is not a higher catastrophe, since, as fig. 8 shows, $p = 2n\gamma/n - 1$ is simply a particular section touching a cusp edge that forms a smooth curve.

Fig. 7



Birth of two cusps for droplet generated by (19) as p passes through $2n\gamma/n - 1$.

Fig. 8



Caustic in $\mathbf{\Omega}$, p control space from droplet generated by (19).

For a given $f(\mathbf{R})$ defining a wavefront W, the control parameter space $\mathbf{\Omega}$ is two-dimensional. However, as the last example showed, we can think of $f(\mathbf{R})$ as embedded in a family of functions generated by one or more parameters ϵ (e.g. pressure in a droplet). In this higher-dimensional control space the catastrophes are not restricted to folds and cusps, and changing ϵ may produce caustics in $\mathbf{\Omega}$ which are singular two-dimensional sections of umbilics, swallow-tails and higher catastrophes.

Often there is no possibility of varying ϵ; then, however, it may be the case that in 'function space' $f(\mathbf{R})$ is 'close to' a higher catastrophe, whose slightly unfolded section will be recognizable in $\mathbf{\Omega}$. For hyperbolic and elliptic umbilics, the table and eqn. (7) show that it is natural to take $\Omega_1 = \Omega_x$ and $\Omega_2 = \Omega_y$, so that Ω_3 is the fixed parameter in $f(\mathbf{R})$. Such sections with Ω_3 fixed (figs. 3 (*d*) and 3 (*e*)) are often seen in caustics from irregular glass; fig. 5 is an example showing both umbilics. For water droplet caustics the elliptic umbilic is generated by eqn. (18) with $n = 3$; hyperbolic umbilics cannot occur because the outer branch would have to be convex outwards thus violating the theorem that the droplet caustics must always be concave outwards.

Swallow-tails are less common than umbilics in caustics generated by our simple model. The state space is one-dimensional and there is no natural identification of Ω_1, Ω_2 or Ω_3 with Ω_x and Ω_y; in addition a partially unfolded swallow-tail would correspond to two close tangencies of $\mathscr{L}$ with the direction of vanishing curvature, a circumstance unlikely often to occur. The author has scanned several hundred undulations of irregular glass with a laser beam without seeing a single unmistakeable swallow-tail. Thom (1975, fig. 4) gives a rather unconvincing example. However, it is possible to produce swallow-tails in the far field of droplet lenses whose perimeters are highly convoluted. Figures 6 (*b*) and 6 (*d*) show examples of this; the swallow-tails are marked

on the tracings of the caustics on figs. 6 (*c*) and 6 (*e*). Figure 6 (*e*) also shows three 'specimens' of sections through the four-dimensional 'butterfly' catastrophe (Thom 1975). (On fig. 6 (*b*) there appears the faint image of a second caustic superimposed on the principal one; it arises from a second droplet dimly illuminated by the periphery of the laser beam.) This sort of search for catastrophes resembles not so much physics as specimen-collecting in botany. Individual specimens of a species of plant can vary enormously in size but share topological features (in animal species, by contrast, individuals are more nearly congruent); similarly, individual water droplet lenses vary enormously, and until the droplet hits the dusty glass surface one never knows exactly what its caustic will look like, only that some kind of catastrophe (species) will be involved. Liesegang (1953) shows beautiful swallowtails formed in the beam of an electron microscope.

By deliberately considering wavefronts W whose form $f(\mathbf{R})$ is non-generic it is possible to produce a structurally unstable caustic. If this can be identified as the singular section of a higher catastrophe then the topology of the caustic that would be produced by a generic perturbation of W can be deduced simply by inspecting the unfolding of the catastrophe. Thus can Thom's theorem be used predictively.

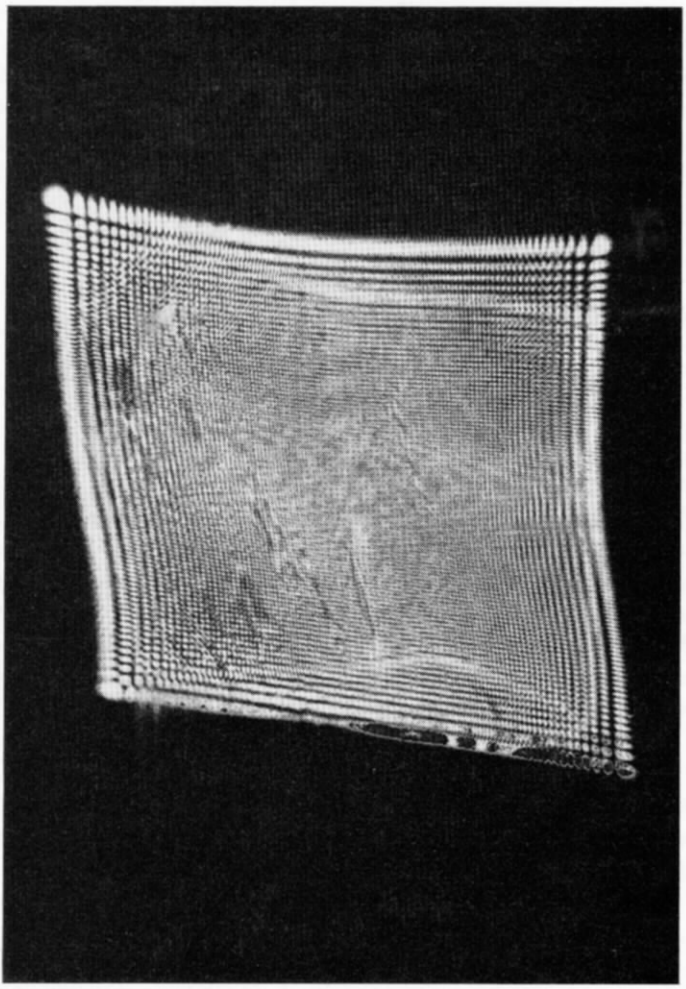
For example, consider the non-generic case where $f(\mathbf{R})$ is *separable*, that is

$$f(\mathbf{R}) = g(\mathbf{a} \cdot \mathbf{R}) + h(\mathbf{b} \cdot \mathbf{R}), \tag{20}$$

where $\mathbf{a}$ and $\mathbf{b}$ are arbitrary non-parallel unit vectors on $\mathbf{R}$. It is not hard to show from the ray condition (4) and the caustic condition $K=0$ that the caustics in $\boldsymbol{\Omega}$ are straight lines parallel to $\mathbf{a}$ and $\mathbf{b}$, meeting in corners but not crossing. In particular, if g and h are periodic functions whose topology is that of a sinusoid (i.e. two inflexions per period), then the caustic will be a parallelogram. The author tested this by sending a laser beam through two panes of glass on each of which a one-dimensional pattern of ridges had been impressed when the glass was softened by heating. The two panes were placed with the ridged faces together and the two sets of ridges not parallel. Figure 9 shows the result. Each non-generic corner is identified by inspection of fig. 3 as the section $\Omega_3=0$ of a hyperbolic umbilic, and this is confirmed by expanding $f(\mathbf{R})$ near the points on W that generate the corners, whereupon the functional form given in the table is obtained.

If now we imagine $f(\mathbf{R})$ (given by (20) with periodic f and g) perturbed by adding a non-separable term $\epsilon f_1(\mathbf{R})$, then each corner will unfold according to fig. 3 (*e*) so that the caustic will now be a distorted four-cusped astroid surrounded by a smooth convex closed curve. This prediction was first made during a study of the reflection of atomic beams from crystal surfaces; for a full description of this problem, and a photograph of an optical analogue confirming the prediction, see Berry (1975). Another application of catastrophe theory explains a curious feature of the near-field caustic lines painted by the sun on the sand beneath the sea (Minnaert 1954, p. 33): they frequently meet in triple junctions. These are non-generic and moreover do not correspond to sections of any three-dimensional catastrophe (fig. 3). This suggests that the triple junctions are really much more complicated generic structures imperfectly resolved or blurred by the $\frac{1}{2}^\circ$ divergence caused by the finite

Fig. 9



Far-field caustic from two superposed panes of ridged glass, showing singular sections of four hyperbolic umbilics.

disc of the sun. And, indeed, experiment confirms this (see Berry and Nye 1976).

§ 5. Caustics and wave functions

Near caustics the wave intensity shows dramatic diffraction effects. These are obvious on figs. 5, 6 and 9. In this section we give an analytical description of these phenomena, based on the generalization of the diffraction integral (10), namely

$$\psi(\mathbf{\Omega}) \equiv \left(\frac{k}{2\pi}\right)^{m/2} \iint d^m\mathbf{R} \exp\left[ik\Phi(\mathbf{R}, \mathbf{\Omega})\right] a(\mathbf{R}, \mathbf{\Omega}), \tag{21}$$

where a is a slowly varying function of its arguments (see Duistermaat 1974), and m is the dimensionality of the state space $\mathbf{R}$.

For short wavelengths k is large so that the exponential oscillates rapidly with $\mathbf{R}$ and the principal contributions come from those points where the phase $k\Phi$ is stationary. The stationary phase condition is eqn. (8), and this simply tells us that the stationary points $\mathbf{R}_i(\boldsymbol{\Omega})$ contributing to $\psi(\boldsymbol{\Omega})$ are those sending *trajectories* to $\boldsymbol{\Omega}$. For a general choice of control parameters $\boldsymbol{\Omega}$ it will usually be the case that the different $\mathbf{R}_i(\boldsymbol{\Omega})$ are well separated, so that the method of stationary phase can be used to evaluate (21) asymptotically for large k. This involves the $m \times m$ matrix

$$M_{\Phi}(\mathbf{R}) \equiv \left\{ \frac{\partial^2 \Phi}{\partial R_i \partial R_j} \right\} ; \tag{22}$$

the method yields ψ as

$$\psi(\boldsymbol{\Omega}) \approx \sum_i \frac{\exp\left(i[kS_i(\boldsymbol{\Omega}) + \alpha_i \pi/4]\right)}{|\det M_{\Phi}(\mathbf{R}_i(\boldsymbol{\Omega}))|^{1/2}} a(\mathbf{R}_i, \boldsymbol{\Omega}), \tag{23}$$

where α_i is the signature of M at $\mathbf{R}_i$ and S_i is the action (9).

The determinant in (23) is simply the Jacobian of the mapping from $\mathbf{R}$ to $\boldsymbol{\Omega}$. Taking the absolute square of (23), we obtain the intensity $I(\boldsymbol{\Omega})$ given by the pure trajectory picture (eqn. (5)), plus interference terms which for large k oscillate rapidly in $\boldsymbol{\Omega}$ (see § 6.2 of Berry and Mount (1972)).

Variation of $\boldsymbol{\Omega}$ will usually result in caustics being encountered ; then the mapping $\mathbf{R} \rightarrow \boldsymbol{\Omega}$ is singular, $\det M$ vanishes, two or more stationary points $\mathbf{R}_i$ coalesce, the formula (23) diverges and the simple stationary-phase method fails to give correctly the contribution of the trajectories to the asymptotic form of ψ. The nature of the divergence depends on the manner in which the stationary points coalesce, that is, on the topological type of the caustic.

In these circumstances a uniform approximation to (21), valid whether or not $\boldsymbol{\Omega}$ is near a caustic, can be found by comparison with a simpler integral that we may write as

$$j(\boldsymbol{\omega}) \equiv \left(\frac{k}{2\pi}\right)^{m/2} \iint d^m \mathbf{r} \exp\left[ik\phi(\mathbf{r}, \boldsymbol{\omega})\right]. \tag{24}$$

The function ϕ is chosen to be simpler in form than Φ but identical so far as the topology of coalescence of its stationary points is concerned. If the caustic produced by Φ is generic, then it is obvious that a sensible choice for ϕ is the standard form (table) given by Thom for the corresponding elementary catastrophe, with $\mathbf{r}$ as state variables and $\boldsymbol{\omega}$ as control variables.

Therefore Thom's theorem does far more than classify caustics : it also classifies the diffraction functions $j(\boldsymbol{\omega})$ that clothe the skeleton of structurally stable caustics when k is large but not infinite. The simplest of these functions, clothing the fold catastrophe, was introduced by Airy (1838) (see also Abramowitz and Stegun (1964)), while that clothing the cusp was introduced by Pearcey (1946) (see also Berry (1975)). Airy and Pearcey diffraction patterns are clearly visible on figs. 5 and 6. The pattern across the singular section of the hyperbolic umbilic (caustic corner) can be seen on fig. 9 ; analytically the wave function for this case is simply the product of two Airy functions. Mr. F. J. Wright is currently computing plane sections of the diffraction patterns for the swallow-tail and the elliptic and hyperbolic umbilics.

Before showing how to express (21) in terms of the comparison integrals (24), we introduce the *singularity index* σ, which gives a gross indication of the strength of the caustic. The greatest amplitude ψ occurs for control parameters $\mathbf{\Omega}_0$ at which the highest-order coalescence of stationary points occurs, and σ is defined by

$$|\psi(\mathbf{\Omega}_0)| = 0(k^\sigma). \tag{25}$$

For the elementary catastrophes (table) the singularity occurs at $\mathbf{\Omega}=0$ and it is not hard to find σ; the values are given in the table. It is clear that σ increases with the co-dimension n. This is not surprising since the more control parameters there are the greater is the number of stationary points that can be made to coalesce. Arnol'd (1973) has calculated σ for some non-elementary castastrophes with finite co-dimension; the general problem is extremely difficult. For caustics of infinite co-dimension (§ 6) only a few results exist: glory scattering has $\sigma=\frac{1}{2}$ and the forward diffraction peak formed by scattering from a potential whose long range tail decays as (distance)$^{-\mu}$ has $\sigma=(\mu+1)/(\mu-1)$ (Berry 1969 a). Finally, we note that σ does not always exist; scattering from an exponentially decaying potential, for example, gives a forward amplitude proportional to $k\ln^2 k$ (Keller and Levy 1963).

Now we return to the problem of evaluating (21). Change the variables from $\mathbf{R}$ to $\mathbf{r}$ by the mapping

$$\Phi(\mathbf{R}, \mathbf{\Omega}) = C(\mathbf{\Omega}) + \phi(\mathbf{r}, \boldsymbol{\omega}(\mathbf{\Omega})), \tag{26}$$

to make (21) as similar as possible to (24). Choose $C(\mathbf{\Omega})$ and the comparison parameters $\boldsymbol{\omega}(\mathbf{\Omega})$ in such a way that the mapping is one-to-one. This can be achieved by making the stationary points $\mathbf{R}_i(\mathbf{\Omega})$ of Φ map onto the stationary points $\mathbf{r}_i(\boldsymbol{\omega})$ of ϕ, as can be seen from

$$\nabla_{\mathbf{R}}\Phi \,.\, d\mathbf{R} = \nabla_{\mathbf{r}}\phi \,.\, d\mathbf{r}. \tag{27}$$

For co-dimension n there can be $n+1$ stationary points involved in the coalescence (because each new parameter Ω enables another stationary point to be moved into coincidence with any pre-existing cluster). Therefore the $n+1$ equations

$$\Phi(\mathbf{R}_i(\mathbf{\Omega}), \mathbf{\Omega}) \equiv S_i(\mathbf{\Omega}) = \phi(\mathbf{r}_i(\boldsymbol{\omega}(\mathbf{\Omega})), \boldsymbol{\omega}(\mathbf{\Omega})) + C(\mathbf{\Omega}) \tag{28}$$

are just sufficient to determine the $n+1$ quantities $C, \boldsymbol{\omega}$. For some values of $\mathbf{\Omega}$ (on the 'dark side' of caustics) there will be fewer than $n+1$ real points $\mathbf{R}_i$, but by analytic continuation it is always possible to find *complex trajectories* $\mathbf{R}_i$, so that we always have $n+1$ equations (see, e.g., Berry 1966).

In terms of the new variables $\mathbf{r}$, (21) becomes, exactly,

$$\psi(\mathbf{\Omega}) = \left(\frac{k}{2\pi}\right)^m \exp\left[ikC(\mathbf{\Omega})\right] \iint d^m\mathbf{r}\, g(\mathbf{r}, \mathbf{\Omega}) \exp\left[ik\phi(\mathbf{r}, \boldsymbol{\omega}(\mathbf{\Omega}))\right], \tag{29}$$

where

$$g(\mathbf{r}, \mathbf{\Omega}) \equiv \left|\frac{d\mathbf{R}}{d\mathbf{r}}\right| a(\mathbf{R}(\mathbf{r}), \mathbf{\Omega}). \tag{30}$$

Next we approximate by replacing g by the first term of an expansion about the stationary points $\mathbf{r}_i$, on the grounds that it is the stationary points that dominate the behaviour of (29). The expansion can be generated by writing

$$g(\mathbf{r}, \boldsymbol{\Omega}) = g_0(\boldsymbol{\omega}) + \mathbf{g}_1(\boldsymbol{\omega}) \cdot \nabla_{\boldsymbol{\omega}} \phi(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{h}(\mathbf{r}, \boldsymbol{\omega}) \cdot \nabla_{\mathbf{r}} \phi(\mathbf{r}, \boldsymbol{\omega}), \tag{31}$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{\Omega})$, $\mathbf{g}_1$ is an n-component vector and $\mathbf{h}$ an m-component vector. At the stationary points, $\mathbf{h} \cdot \nabla_{\mathbf{r}} \phi$ vanishes and so we neglect this term altogether. (If we included it, and expanded each component of $\mathbf{h}$ in a similar fashion to (31), the result would be an asymptotic series for (29) involving inverse powers of k. The derivation of the complete asymptotic series, and their interpretation as exact representations of wave functions ψ (Dingle 1973) is a challenging task for the future.) Neglecting $\mathbf{h}$, we can write (29) in terms of the $n+1$ quantities g_0, $\mathbf{g}_1$, and the comparison integrals $j(\boldsymbol{\omega})$ and their derivatives :

$$\psi(\boldsymbol{\Omega}) = \exp\left[ikC(\boldsymbol{\Omega})\right] \left[g_0(\boldsymbol{\omega}(\boldsymbol{\Omega})) j(\boldsymbol{\omega}(\boldsymbol{\Omega})) + \frac{\mathbf{g}_1(\boldsymbol{\omega}(\boldsymbol{\Omega}))}{ik} \cdot \nabla_{\boldsymbol{\omega}} j(\boldsymbol{\omega}(\boldsymbol{\Omega})) \right]. \tag{32}$$

To find g_0 and $\mathbf{g}_1$ we use the $n+1$ equations obtained by evaluating (31) at the stationary points. The left-hand side requires a knowledge of $|d\mathbf{R}/d\mathbf{r}|$ at each stationary point (cf. (30)), and differentiation of (27) leads to

$$\left| \frac{d\mathbf{R}}{d\mathbf{r}} \right|_i = \left| \frac{\det M_{\phi}(\mathbf{r}_i(\boldsymbol{\omega}(\boldsymbol{\Omega})))}{\det M_{\Phi}(\mathbf{R}_i(\boldsymbol{\Omega}))} \right|^{1/2}, \tag{33}$$

where M is defined by (22). Therefore g_0 and $\mathbf{g}_1$ are determined by solving

$$\left| \frac{\det M_{\phi}(\mathbf{r}_i(\boldsymbol{\omega}(\boldsymbol{\Omega})))}{\det M_{\Phi}(\mathbf{R}_i(\boldsymbol{\Omega}))} \right|^{1/2} a(\mathbf{R}_i(\boldsymbol{\Omega}), \boldsymbol{\Omega}) = g_0 + \mathbf{g}_1 \cdot \nabla_{\boldsymbol{\omega}} \phi(\mathbf{r}_i(\boldsymbol{\omega}(\boldsymbol{\Omega})), \boldsymbol{\omega}(\boldsymbol{\Omega})). \tag{34}$$

The wave function is now given by (32) in terms of $j(\boldsymbol{\omega})$ defined by (24). This technique gives a *uniform approximation* : for $\boldsymbol{\Omega}$ far from caustics the comparison integrals themselves may be expanded by the stationary phase method and we quickly recover (23). When $\boldsymbol{\Omega}$ is very close to the highest singularity $\boldsymbol{\Omega}_0$, on the other hand, we can neglect the second term in (32) and set $\boldsymbol{\omega}$ proportional to $\boldsymbol{\Omega}$ in the first term, thus obtaining a *transitional approximation* valid very close to $\boldsymbol{\Omega}_0$ (for examples see Airy (1838) and Ford and Wheeler (1959)). The uniform approximation for the fold catastrophe was first obtained by Chester, Friedman and Ursell (1957), and introduced into wave theory by Berry (1966), and Ludwig (1966, 1967). This latter author also devised uniform approximations for some other elementary catastrophes, as did Connor (1973, 1974, 1976). The whole subject has been put on a rigorous basis, so far as structurally stable caustics are concerned, by Duistermaat (1974). However, we emphasize that the result (32) is not restricted to these cases ; in fact Berry (1969 a) obtained uniform approximations for two structurally unstable caustics of infinite co-dimension (see § 6). Numerically, uniform approximations are very accurate (Mount 1973, Berry and Mount 1972), and they have been employed in the inversion of scattering data to determine interaction potentials (Buck 1971, 1974, Mullen and Thomas 1973).

§ 6. Caustics of infinite co-dimension

It sometimes happens that a wave system is strongly constrained to have a high degree of symmetry, so that the caustic is structurally unstable and extra parameters ϵ are necessary to restore genericity, that is, to produce all the different structurally stable caustics 'near' to the original one. We are concerned here with cases where the number of components of ϵ is infinite, so that the caustic is of infinite co-dimension. There is as yet no mathematical scheme classifying caustics of this type; Thom's theorem is inapplicable because it is restricted to co-dimension less than 7. We discuss three examples.

The first occurs when the wavefront W (§ 2) has cylindrical symmetry and its profile $f(\mathbf{R})$ has a maximum not at the origin. This maximum forms a circle ($R = R_0$ say) in $\mathbf{R}$, near which W is locally part of a torus. This situation corresponds to *glory scattering* in quantum mechanics (Berry 1969 a) or meteorology (Van de Hulst 1957, Nussenzveig 1969) where a trajectory with non-zero impact parameter interacts with a spherical object and emerges in the forward or backward direction. It can also be produced by an annular droplet on a glass surface, $f(R)$ being given by (2) with $g = \ln(x+iy)$.

The Gaussian curvature K vanishes on $R = R_0$, so that (§ 2) this circle generates the caustic. When $R = R_0$ the slope ∇f vanishes, and so according to eqn. (4) the corresponding trajectories all have $\mathbf{\Omega} = 0$: the far-field caustic is a single point in the forward direction. A point is structurally unstable in $\mathbf{\Omega}$, and the possibility suggests itself that we might be seeing the singularity $\Omega_3 = 0$ of the elliptic umbilic (fig. 3 (*d*)).

That this is not the case can be seen in two ways. First, the generating function $\Phi(\mathbf{R}, \mathbf{\Omega})$ (eqn. (7)), when written in simple standard form obtained by expansion about R_0, is

$$\Phi(\mathbf{R}, \mathbf{\Omega}) = -\frac{(\mathbf{R}-\mathbf{R}_0)^2}{2a} + \Omega R \cos\theta \tag{35}$$

where θ is the polar angle in $\mathbf{R}$, measured with respect to the direction of $\mathbf{\Omega}$. This function has different symmetry from the elliptic umbilic (table), and cannot be transformed into it. Second, the comparison integral (24) for this case is, for large k and small Ω,

$$j(\mathbf{\Omega}) = R_0(2\pi a)^{1/2}k^{1/2}\exp(-i\pi/4)J_0(kR_0\Omega), \tag{36}$$

where J_0 is the zero-order Bessel function of the first kind. Therefore the singularity index, defined by eqn. (25) and found by setting $\mathbf{\Omega} = 0$, is $\sigma = \frac{1}{2}$, so that the glory is a stronger singularity than the elliptic umbilic or indeed any other elementary catastrophe.

The focusing into $\mathbf{\Omega} = 0$ arises from the circular symmetry of W. If this is broken by perturbations $f_1(\mathbf{R}, \epsilon)$ with parameters ϵ of the profile $f(\mathbf{R})$, then the point caustic should break up into closed cusped figures, the number of cusps depending on the topology of f_1. But there is an infinity of topologically different ways in which the (continuous group) symmetry of a circle can be broken, so that to find all the structurally stable caustics 'near' the glory requires infinitely many parameters $\epsilon_1, \epsilon_2 \ldots$. We conjecture that such

a generic unfolding of the function (35) is obtained by a series of circular harmonic perturbations, namely

$$f_1(\mathbf{R}, \epsilon) = \sum_{n=2}^{\infty} \epsilon_{n-1} R^n \cos(n\theta + \alpha_n), \tag{37}$$

involving a countable infinity of parameters.

Our second caustic of infinite co-dimension is the *forward diffraction peak* from localized objects such as a wavefront deformation $f(\mathbf{R})$ vanishing smoothly as $\mathbf{R} \to \infty$. This caustic is also the isolated point $\mathbf{\Omega} = 0$, produced by the flat 'region at infinity' for which $\nabla f = 0$. K vanishes over an area rather than on a line, so the forward peak is an even stronger structurally unstable caustic than the glory.

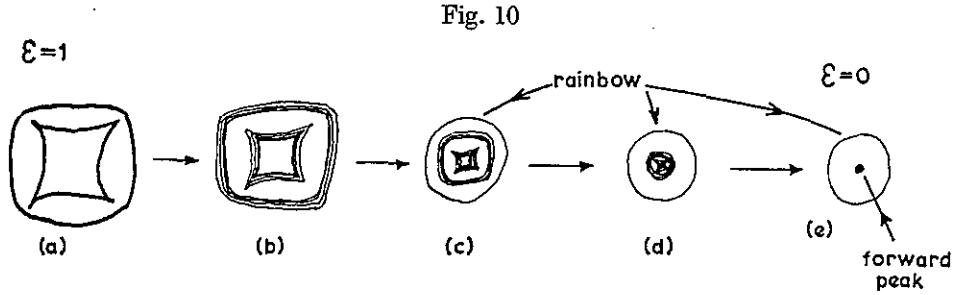
To evaluate the singularity index σ for this case it is first necessary to make the diffraction integral (10) converge by subtracting the forward delta function that would occur if W were flat ($f = 0$). The resulting differential scattering cross section is obtained from (10) by replacing $\exp(ik\Phi)$ by $\exp(ik\Phi) - 1$. A detailed analysis (Berry 1969 a) then shows that σ depends on the manner in which $f(\mathbf{R})$ goes to zero at infinity. For example. in the case of potential scattering (eqn. (3)) when $V(r)$ has the Van der Waals' tail, $-r^{-6}$, we found $\sigma = 1{\cdot}4$. This is powerful focusing indeed, and the forward peak is the strongest directional caustic yet discovered.

As we might expect, the unfolding of the forward peak is more complicated than that of the glory. The 'symmetry' to be broken is the asymptotic vanishing of $f(\mathbf{R})$, so that a typical symmetry-breaking perturbation $\epsilon f_1(\mathbf{R})$ must have maxima, minima and saddles right out to $\mathbf{R} = \infty$. Each undulation gives a caustic (perhaps resembling fig. 5 or 6) and in $\mathbf{\Omega}$ space these will all be superposed. As $\epsilon \to 0$ this mass of caustics shrinks down to $\mathbf{\Omega} = 0$ giving the forward peak. To examine this process in detail, let the forward peak be produced by a cylindrically symmetrical function $f_0(R)$ of Gaussian form, and let the perturbation $\epsilon f_1(\mathbf{R})$ be a periodic undulation with square unit cells containing a minimum at the centre and maxima at each corner, and consider

$$f(\mathbf{R}) = (1 - \epsilon) f_0(R) + \epsilon f_1(\mathbf{R}) \tag{38}$$

as ϵ diminishes from 1 to 0. The progression is shown on fig. 10. When $\epsilon = 1$ only f_1 acts, and the caustic (fig. 10 (a)) is an astroid with surrounding smooth curve (Berry 1975). Slightly diminishing ϵ introduces $f_0(R)$ which

Fig. 10



Emergence of the forward peak and rainbow singularities as a periodic perturbation is switched off (eqn. (38)).

slightly perturbs the unit cells near $\mathbf{R}=0$ so that the caustic splits into an infinite concentric sequence (fig. 10 (*b*)) with ' limit lines ' from the unit cells at infinity. When ϵ is further decreased the diminishing slopes of $f_1(\mathbf{R})$ cause this caustic cluster (fig. 10 (*c*)) to shrink in $\boldsymbol{\Omega}$, while the growth of $f_0(R)$ exposes its inflexion line as a growing ' rainbow ' smooth caustic surrounding all the others. Finally, as $\epsilon \rightarrow 0$ the caustic cluster from $f_1(\mathbf{R})$ collapses (fig. 10 (*d*)) onto the forward peak (fig. 10 (*e*)) at $\boldsymbol{\Omega}=0$, leaving the rainbow circle at some finite $|\boldsymbol{\Omega}|$. Of course there are infinitely many ways apart from this in which ' superposition and collapse ' can occur, so that the infinite co-dimensionality of the forward peak is probably uncountable. (Perhaps this is one of the ' lump ' catastrophes discussed by Thom (1975, pp. 103–4).)

The third class of caustics with infinite co-dimension are those occurring at point *sources*. Of course the radiation fields are singular at source points, but we are primarily interested in the pattern of emerging trajectories. Obviously a vibrating source at rest relative to the propagation medium (or to the observer in the case of light) is an isolated point caustic (a ' focus '). No perturbation of the medium will break this even though it is structurally unstable. Motion of the source can produce fold caustics emanating from it (e.g. Cerenkov radiation, or Kelvin's ship wave pattern—see Lamb 1945), but this situation is still non-generic because the folds meet at a finite angle and not, say, at a cusp. To understand the unfolding it is necessary to imagine the time-reversed situation where there is no source and the caustic point arises from the collapse of a wavefront with a high degree of symmetry (e.g. a circle in a homogeneous medium with source at rest). Then any of an infinite-parameter class of perturbations of this wavefront will break the caustic into a structurally stable form (e.g. a cusped line in two dimensions). In the real-time problem, however, the presence of the source is a strong constraint suppressing all such perturbations, and the unfolding cannot be realized.

§ 7. Concluding remarks

Thom's theorem has enabled us to make powerful general statements about the caustics that dominate wave fields when the wavelength is short. Firstly, it lists the canonical forms that caustics in generic systems may take. Secondly, it enables canonical diffraction functions to be derived that give accurate representations of the wave fields that clothe these caustics. And, thirdly, departures from genericity, caused by special conditions corresponding to particular values of a small number of parameters describing the system, give rise to caustics that are structurally unstable ; these are easily identified because they are not elementary catastrophes with co-dimension equal to the dimension of the space in which the caustics are observed, but special sections of higher catastrophes, whose unfoldings lead to predictions of the forms of the structurally stable caustics that would result if genericity were restored by removing the special conditions.

We emphasize that structural stability of a caustic does not imply physical stability of the system generating the caustic, and vice versa. For example, a spherical raindrop is rendered physically stable by the force of surface tension, and yet it gives rise to the glory which is so structurally unstable as to require infinitely many unfolding parameters. Attempts are being made

to develop extensions of Thom's theorem to deal with such cases, by studying singularities which are structurally stable in a restricted sense, namely under perturbations of special types (e.g. those preserving spherical symmetry). Interesting as such studies undoubtedly are, they should not deflect too much effort from the more difficult task of discovering unfoldings that are structurally stable under all perturbations; in the case of the glory, for example, generic perturbations are certainly possible, in the form of oscillations of raindrops away from sphericity.

Next, we mention two aspects of short-wave theory to which Thom's theory has not so far been applied. The first is the determination of the spectrum of normal modes of bound systems (e.g. energy levels in quantum mechanics); this problem is dominated by the topologies of closed trajectories (Balian and Bloch 1974, Berry and Tabor 1976), and caustics play a relatively minor role, although they do affect the spectrum.

The second arises when the pattern of trajectories is so complicated that the caustics are separated by less than a wavelength, on the average. This case is realized, for example, when waves are reflected by a surface that is rough on all scales (Berry 1972) or when trajectories are ergodic (Percival 1973). We conjecture that in such situations no short-wave limit exists. This is an important problem that should richly repay detailed study.

It is possible that catastrophes penetrate even deeper into wave theory than this article suggests. On the finest scale, the most delicate features of waves are the singularities of their phase maps, christened 'wavefront dislocations' by Nye and Berry (1974). Now, the phase map for the cusp diffraction function, plotted by Pearcey (1946), shows a striking pattern of edge dislocation points. The three-dimensional catastrophe diffraction functions must have a pattern of dislocation lines with highly interesting topology. I conjecture that each generic caustic has an associated dislocation structure that is structurally stable. If this is correct, Thom's theorem would also classify wavefront dislocations (the classification would be incomplete because there are simple ways of creating dislocations that are not associated with caustics). This subject is now being studied in collaboration with Mr. F. J. Wright.

To conclude, let us compare the application of Thom's theorem in wave theory, as considered in this article, with the application of the theorem in statistical mechanics, to phase transitions of infinite systems. In both cases catastrophe theory gives an overall description: geometrical optics in the wave case, 'mean field theory' (e.g. Van der Waals' equation) in the statistical case. Very close to the singularities, however, more refined methods must be employed, because of diffraction in the wave case (requiring the integration techniques of § 5) and fluctuations in the statistical case (requiring special techniques for calculating 'critical exponents'). This point is well made by Schulman (1974).

Acknowledgments

I particularly wish to thank Professor John Nye for many long discussions about irregular droplet lenses. In addition, I thank Dr. Manfred Faubel for an experimental suggestion, Mr. George Keene for photography, Mr. Tony

Osman for making the undulating glass plates, and Dr. Tim Poston for some helpful comments.

References

Abramowitz, H., and Stegun, I. A., 1964, *Handbook of Mathematical Functions* (Washington : U.S. National Bureau of Standards).

Airy, G. B., 1838, *Proc. Camb. phil. Soc. math. phys. Sci.*, **6**, 379.

Arnol'd, V. I., 1973, *Russ. math. Survs*, **28**, No. 5, 19 (English translation).

Balian, R., and Bloch, C., 1974, *Ann. Phys.*, **85**, 514.

Berry, M. V., 1966, *Proc. phys. Soc.*, **89**, 479 ; 1969 a, *J. Phys.* B. **2**, 381 ; 1969 b, *Sci. Prog., Oxf.*, **57**, 43 ; 1972, *J. Phys.* A, **5**, 272 ; 1975. *Ibid.*, **8**, 566.

Berry, M. V., and Mount, K. E., 1972, *Rep. Prog. Phys.*, **35**, 315.

Berry, M. V., and Nye, J. F., 1976 (to be published).

Berry, M. V., and Tabor, M., 1976 *Proc. R. Soc.* (in the press).

Bröcker, Th., 1975, *Differential Germs and Catastrophes* (Cambridge University Press).

Buck, U., 1971, *J. chem. Phys.*, **54**, 1923 ; 1974, *Rev. mod. Phys.*, **46**, 369.

Campbell, L., and Garnett, W., 1882, *The Life of James Clerk Maxwell* (London : Macmillan), pp. 434–44.

Chester, C., Friedman, B., and Ursell, F., 1957, *Proc. Camb. phil. Soc. math. phys. Sci.*, **53**, 599.

Connor, J. N. L., 1973, *Molec. Phys.*, **26**, 1217 ; 1974, *Ibid.*, **27**, 853 ; 1976, *Ibid.*, **31**, 33.

Descartes, R., 1637, *Discours de La Méthode* (Appendix).

Dingle, R. B., 1973, *Asymptotic Expansions : Their Derivation and Interpretation* (London : Academic Press).

Duistermaat, J. J., 1974, *Communs pure appl. Maths.*, **27**, 207.

Eisenhart, L. P., 1960, *A Treastise on the Differential Geometry of Curves and Surfaces* (Dover), p. 180.

Fellini, F., 1963, $8\frac{1}{2}$ (Film distributed by Cineriz, Italy).

Ford, K. W., and Wheeler, J. A., 1959, *Ann. Phys.*, **7**, 259, 287.

Glauber, R. J., 1958, *Lectures in Theoretical Physics*, Vol. 1, edited by W. E. Brittin and L. G. Dunham (New York : Interscience), p. 315.

Goldstein, H., 1951, *Classical Mechanics* (Addison–Wesley).

Jänich, K., 1974, *Math. Annln*, **209**, 161.

Kravtsov, Yu. A., 1968, *Soviet Phys. Acoust.*, **14**, 1 (English translation).

Keller, J. B., and Levy, B. R., 1963, *Proc. Conf. Electromagn. Scattering*, 1962, edited by M. Kerker (Oxford : Pergamon), p. 3.

Lamb, H., 1945, *Hydrodynamics* (Cambridge University Press), p. 438.

Larmor, J., 1891, *Proc. Camb. phil. Soc. math. phys. Sci.*, **7**, 131.

Lawrence, J. Dennis, 1972, *A Catalog of Special Plane Curves* (New York : Dover).

Liesegang, S., 1953, *Optik*, **10**, 5.

Ludwig, D., 1966, *Communs pure appl. Math.*, **19**, 215 ; 1967, *Ibid.*, **20**, 103.

Maslov, V. P., 1972, *Theorié des Perturbations et Méthodes Asymptotiques* (Paris : Dunod). (Original Russian publication, 1965.)

Minnaert, M., 1954, *The Nature of Light and Colour in the Open Air* (Dover).

Mount, K. E., 1973, *J. Phys.* B, **6**, 1397.

Mullen, J. M., and Thomas, B. S., 1973, *J. chem. Phys.*, **58**, 5216.

Nussenzveig, H. M., 1969, *J. Math. Phys.*, **10**, 82.

Nye, J. F., and Berry, M. V., 1974, *Proc. R. Soc.* A, **336**, 365.

Pearcey, T., 1946, *Phil. Mag.*, **37**, 311.

Percival, I. C., 1973, *J. Phys.* B, **6**, L229.

Schulman, L., 1975, *Functional Integration and its Applications*, edited by A. M. Arthurs (Oxford : Clarendon Press), p. 144.

Stavroudis, O. N., 1972, *The Optics of Rays, Wavefronts and Caustics* (New York : Academic Press), p. 79.

SYNGE, J. L., 1954, *Geometrical Mechanics and de Broglie Waves* ; 1960, *Encyclopedia of Physics*, edited by S. Flügge (Berlin : Springer-Verlag), **311**, 1.
THOM, R., 1975, *Structural Stability and Morphogenesis* (Reading, Mass. : Benjamin).
TITCHMARSH, E. C., 1949, *The Theory of Functions* (Oxford University Press), p. 122.
VAN DE HULST, H. C., 1957, *Light Scattering by Small Particles* (New York : John Wiley).
WASSERMANN, G., 1974, *Stability of Unfoldings*, Vol. 393 (Springer Mathematical Notes).