SINGULARITIES
BANACH CENTER PUBLICATIONS, VOLUME 20
PWN–POLISH SCIENTIFIC PUBLISHERS
WARSAW 1988

# FAMILES OF SPHERES (AND CIRCLES)

J. W. BRUCE

*School of Mathematics, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, United Kingdom*

and

P. J. GIBLIN

*Department of Pure Mathematics, University of Liverpool, Liverpool, United Kingdom*

Let $\boldsymbol{R}^n$ denote Euclidean $n$-space with inner product $(x_1, \ldots, x_n)\cdot(y_1, \ldots, y_n) = x\cdot y = \sum_{i=1}^{n} x_i y_i$. The unit $(n-1)$-sphere $S^{n-1}$ given by $x\cdot x = \|x\|^2 = 1$ is the simplest (most symmetric) compact hypersurface in $\boldsymbol{R}^n$, and has been the object of innumerable investigations. Spherical geometry is a subject in itself, historically motivated, of course, by the fact that the 2-sphere is an excellent model of the earth's surface; and the circle is the most intensively studied of all geometric objects! In this survey we will consider not single circles and spheres but smooth *families* of such; they arise naturally, for example, in geometric optics. It is, perhaps, initially surprising that families of such simple geometric objects can give rise to interesting and intricate geometry; we hope to convince the reader of this below. In what follows we shall confine our attention largely to the especially relevant cases of families of circles in the plane and spheres in Euclidean 3-space; most of the results do, however, have higher dimensional analogues.

## 1. Wavefronts (spheres of fixed radius)

Let $M \subset \boldsymbol{R}^2$ be a simple closed plane curve, parametrised by arc-length by the mapping $x\colon S^1 \to \boldsymbol{R}^2$, $s \mapsto x(s)$, where $S^1$ is a circle. We can view $M$ as an initial wavefront in the homogeneous medium $\boldsymbol{R}^2$; we shall suppose that waves propogate with unit speed in this medium. By Huyghens' principle subsequent wavefronts are obtained as the envelope of families of circles of fixed radius (equal to the elapsed time) centred on $M$.

Defining

$$F\colon S^1 \times \boldsymbol{R}^2 \to \boldsymbol{R} \quad \text{by } F(s, a) = \|x(s)-a\|^2$$

we see that the circle centred at $x(s)$ of radius $r$ is given by $F(s, a) = r^2$, and the required envelope is obtained, in the traditional way, by eliminating $s$ from the equations

$$F(s, a)-r^2 = \frac{\partial}{\partial s}(F(s, a)-r^2) = 0. \tag{1}$$

Classically the envelope is thought of as being obtained as the locus of its *characteristic* points, these points being the intersection of two "consecutive circles". Since two circles intersect in 0, 1 or 2 points any envelope of circles has 0, 1 or 2 characteristic points for each parameter value (except of course in the very degenerate case when the entire circle consists of characteristic points). In the example considered here there are always two characteristic points on each circle; the wavefront can, in fact, be obtained by laying fixed distances $r$ off along each normal in both directions. One (characteristic!) property of envelopes of circles we shall need later is that at a point where the envelope is smooth, it is tangent to the corresponding circle. In the case of wavefronts this means that distinct wavefronts have parallel tangents at corresponding points; in fact they have common normals of course.

Generally speaking wavefronts have singularities at the *points of regression* of the envelope. These occur at points of the envelope where, in addition to the previous two equations (1) vanishing, we have $\partial^2/\partial s^2(F(s, a) - r^2) = 0$. To understand the nature of these singularities we naturally have to use some singularity theory, so we now introduce the family of functions which are of principal interest to us, and do so in some generality. One of the first authors to use these in differential geometry was Porteous in [22]. Let $M^m \subset \boldsymbol{R}^n$ be a smooth submanifold of Euclidean $n$-space of dimension $m$ and consider the family of *distance squared* functions

$$F\colon M \times \boldsymbol{R}^n \to \boldsymbol{R} \quad \text{defined by} \quad F(x, a) = \|x-a\|^2.$$

Below we shall require, at various stages, consequences of the following fundamental result of E. Looijenga.

THEOREM. *For generic $M \subset \boldsymbol{R}^n$ the family $F$ is a generic family of functions on $M$.*

We shall not be more specific concerning the term *generic family of functions*; the reader is referred to [20] and [23] for precise details. However for $n \leqslant 5$ the theorem implies, in particular, that this family of functions is smoothly stable, in the natural sense, and it is topologically stable for any value of $n$. The Theorem is, in fact, a transversality result, the proof of which is not too difficult.

This theorem is useful for studying wavefronts for the following reason. Staying with the general case for the moment, the wavefront propagating from an initial "disturbance" $M$ after time $r$ remains the envelope of spheres of radius $r$ centred on $M$ and is obtained by eliminating the $x$ variables from the equations

$$(2) \qquad F-r^2 = \frac{\partial F}{\partial x_1} = \ldots = \frac{\partial F}{\partial x_m} = 0$$

using $x_1, \ldots, x_m$ variables on $M$. (Of course these equations can also be interpreted as determining the centres of spheres of radius $r$ tangent to $M$ – which is one way of seeing that the wavefronts can be obtained by laying off fixed distances down normals to $M$.) Fixing our attention at some point $(x_0, a_0) \in M \times \mathbf{R}^n$ with $\|x_0 - a_0\|^2 = r^2$, we see that $F$ yields a map germ $F\colon (M^m \times \mathbf{R}^n, (x_0, a_0)) \to (\mathbf{R}, r^2)$ which we can consider as an unfolding of $f = F(-, a_0)\colon (M^m, x_0) \to (\mathbf{R}, r^2)$. Looijenga's result implies that for a fixed $r$ and generic $M$ the map $F$ is, in fact, a *versal unfolding* of the germ $f$ when $n \leqslant 6$. The equations (2) above then define the *discriminant set* of this unfolding, that is the set of parameter values $a$ with $F(-, a)$ having a critical point with critical value the base value, in this case $r^2$. It is a fundamental but straightforward fact that the discriminant of such a versal unfolding depends (up to diffeomorphism) only on $f$ and the dimension $n$ of the unfolding space.

Now in generic $n$-parameter families of functions for $n \leqslant 5$ only simple singularities of functions as defined by Arnold occur ([1]), in fact only those labelled $A_1, \ldots, A_6$, $D_4$, $D_5$, $D_6$, $E_6$. The consequence is that if the dimension of the ambient space is at most 5 we have a finite number of local models for parallels. In particular returning to the case of a plane curve we find that if $F = \partial F/\partial s = 0$, $\partial^2 F/\partial s^2 \neq 0$ at some point $(s_0, a_0) \in S^1 \times \mathbf{R}^2$ then the germ $f$ has an $A_1$ singularity at $s_0$ and the wavefront is locally smooth. On the other hand if the second derivative of $F$ *does* vanish but the third does not the germ $f$ has an $A_2$ singularity and the envelope, near the corresponding point of regression, is locally diffeomorphic to an ordinary cusp. A simple computation shows that this occurs precisely when $a$ is the centre of curvature of $M$ at $x(s)$ but this later point is not a vertex of $M$. For $M$ a curve or surface in $\mathbf{R}^3$ the generic local singularities occurring for the $f$'s are $A_1$, $A_2$ and $A_3$, and the wavefront is correspondingly locally smooth, or has a cuspidal edge or a swallowtail.

Before proceeding two remarks are in order. First, generally wavefronts will have self intersections. For example in the plane they will often have ordinary double points, but these self intersections can be dealt with very easily, they are geometrically relatively uninteresting, and we shall largely ignore them. Secondly we note that generally current understanding of function germs other than those which are simple is very poor, but thankfully only simple singularities occur when considering generic families of circles in $\boldsymbol{R}^2$ and spheres in $\boldsymbol{R}^3$.

The discussion of parallels above is, of course, all for a given *fixed* value of $r$. It is natural to enquire how, from an initial generic $M$, the wavefronts change with time. Certainly the results described above will not suffice; for example generally curves *do* have vertices, and we have, as yet, no description of the evolution of the wavefronts through the corresponding centres of curvature. To proceed we start by considering a slightly different family of functions

$$\tilde{F}\colon M \times \boldsymbol{R}^n \times \boldsymbol{R}^+ \to \boldsymbol{R} \quad \text{defined by} \quad \tilde{F}(x, a, r) = \|x-a\|^2 - r^2.$$

The result of Looijenga proves that for any $(x_0, a_0, r_0) \in M \times \boldsymbol{R}^n \times \boldsymbol{R}^+$ with $\|x_0^2 - a_0\|^2 = r_0^2$ the family $\tilde{F}$ yields a versal unfolding of the germ $\tilde{f} = \tilde{F}(-, a_0, r_0)\colon (M, x_0) \to (\boldsymbol{R}, 0)$. This means that the family of wavefronts, $\Sigma$, or "big wavefront" as we sometimes call it, parametrised by $r$ and obtained by eliminating $x$ from the equations $\tilde{F} = \partial\tilde{F}/\partial x_1 = \ldots = \partial\tilde{F}/\partial x_m = 0$, also has for $n \leqslant 5$ a finite number of local models, these being the discriminant sets of simple singularities of low codimension. We now wish to investigate how this family changes with time so we consider the projection of the set $\Sigma$ to the $r$ parameter. In [2] Arnold classified generic functions on discriminants of simple singularities, precisely so he could describe the evolution of generic wavefronts. In the case of $A_k$ singularities his results are as follows.

An $A_k$ singularity $f\colon (\boldsymbol{R}^m, 0) \to (\boldsymbol{R}, 0)$ is one which is right equivalent to a germ of the form

$$g(x) = g(x_1, \ldots, x_m) = \sum_{i=1}^{m-1} \pm x_i^2 \pm x_m^{k+1},$$

and the latter has a versal unfolding

$$G(x_1, \ldots, x_m, a_1, \ldots, a_n) = G(x, a) = \sum_{i=1}^{m-1} \pm x_i^2 \pm x_m^{k-1} + \sum_{j=0}^{k-1} a_{k-j} x^j$$

for any $n \geqslant k$. The discriminant $\mathscr{D}$ is obtained by eliminating the $x_i$ terms from the equations $G = \partial G/\partial x_1 = \ldots = \partial G/\partial x_m = 0$ (clearly the $x_i$ terms, $1 \leqslant i \leqslant m-1$, are irrelevant; in particular this discriminant does not depend on $m$). We wish to classify functions on this discriminant: we shall say that two germs $\alpha_1, \alpha_2\colon (\boldsymbol{R}^n, 0) \to (\boldsymbol{R}, 0)$ are $\mathscr{R}(\mathscr{D})$ equivalent if there is a germ of a diffeomorphism $\phi\colon (\boldsymbol{R}^n, 0) \to (\boldsymbol{R}^n, 0)$, *taking* $\mathscr{D}$ *to itself*, with $\alpha_1 \circ \phi = \alpha_2$. When $n > k$, so that the unfolding $G$ has $(n-k)$ redundant variables, the discriminant $\mathscr{D}$ or more precisely the pair $(\boldsymbol{R}^n, \mathscr{D})$ is locally diffeomorphic to the product $(\mathbf{R}^k, \mathcal{D}_1)\times \mathbf{R}^{n-k}$ where $\mathcal{D}_1$ is the discriminant of a miniversal unfolding.

For $A_k$ singularities Arnold's result can be stated as follows.

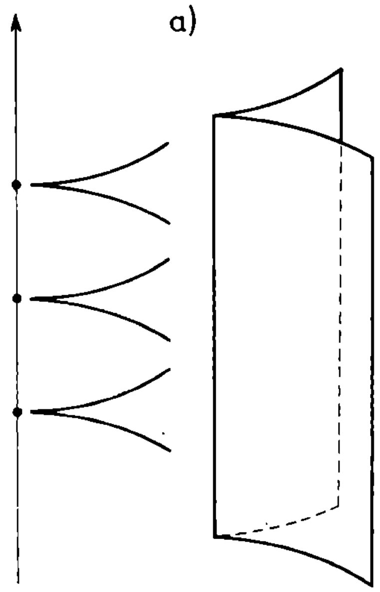
THEOREM ([2]). (i) *When $n = k$, that is $G$ is a miniversal unfolding of $g$, any germ $\alpha$: $(\mathbf{R}^k, 0) \to (\mathbf{R}, 0)$ with $\partial\alpha/\partial a_1(0) \neq 0$ is $\mathcal{R}(\mathcal{D})$ equivalent to one of the germs $\pm a_1$.*

(ii) *When $n > k$ any germ $\alpha$: $(\mathbf{R}^n, 0) \to (\mathbf{R}, 0)$ with $\partial\alpha/\partial a_j(0) \neq 0$ for some $k < j \leqslant n$ is $\mathcal{R}(\mathcal{D})$ equivalent to the projection $a \mapsto a_n$. If $\partial\alpha/\partial a_j(0) = 0$ for $k < j \leqslant n$, $\partial\alpha/\partial a_1(0) \neq 0$ and the restriction of $\alpha$ to $0 \times \mathbf{R}^{n-k}$ has a Morse singularity at 0 then $\alpha$ is $\mathcal{R}(\mathcal{D})$ equivalent to the germ $\pm a_1 + \sum_{j=k+1}^{n} \pm a_j^2$.*

Generic functions on discriminant sets of $A_k$ singularities are modelled, then, on the germs described in (i) and (ii) above. In the first case of (ii) we shall speak of a *trivial transition* (nothing essentially changes), in the second case we shall speak of a *Morse transition.*

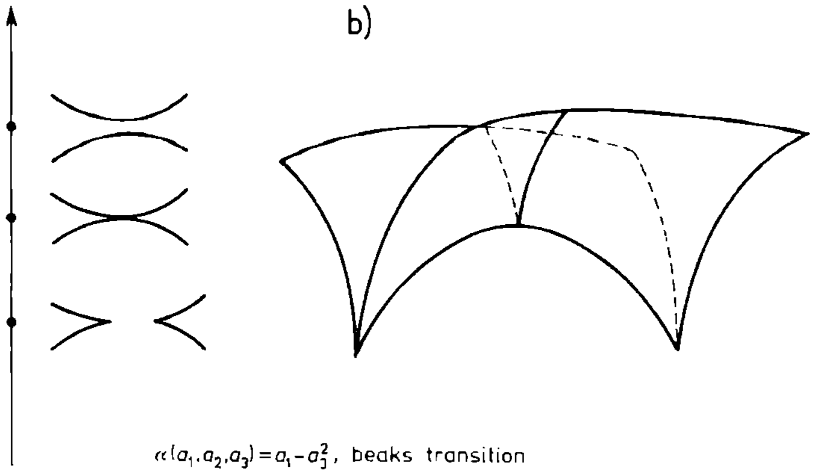
Some examples should make this clearer. For an $A_2$ singularity the discriminant of a universal unfolding is an ordinary cusp and the condition $\partial\alpha/\partial a_1(0) \neq 0$ simply means that at 0 the fibre of the map $\alpha$ is transverse to the cuspidal tangent. For a versal unfolding with three unfolding variables the discriminant is now locally a product of this cusp with an interval. The three possibilities for generic functions here are illustrated in the diagram (Fig. 1).

As another example, for the discriminant of a miniversal unfolding of an $A_3$ singularity, the swallowtail, the condition $\partial\alpha/\partial a_1(0) \neq 0$ means again that at

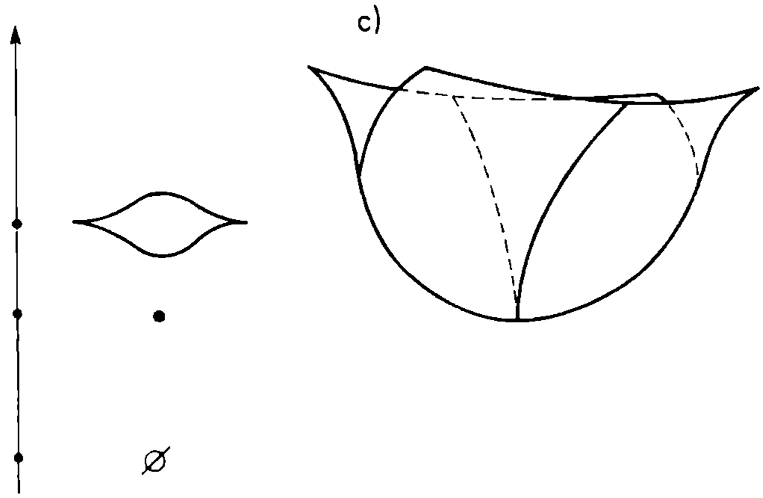



$\alpha(a_1, a_2, a_3) = a_3$, trivial projection

Fig. 1a




$\alpha(a_1, a_2, a_3) = a_1 - a_3^2$, beaks transition




$\alpha(a_1, a_2, a_3) = a_1 + a_3^2$, lips transition

Fig. 1 b, c

0 the fibre of $\alpha$ is transverse to the cuspidal tangent or equivalently the tangent to the curve of self intersection at 0. The corresponding transition is pictured in Figure 2. Arnold's results for the remaining simple singularities are of a similar nature (see [2] and also [6] for alternative proofs).

Let us now return to our wavefronts. Recall that we have a big family of wavefronts $\Sigma$ and a projection $\pi$: $\Sigma \to \boldsymbol{R}$ to the $r$ co-ordinate. We know that $\Sigma$ is locally diffeomorphic to the discriminant set of a versal unfolding of a simple singularity for $n \leqslant 5$, and for a generic choice of initial wavefront $M$ we except the projection $\pi$ to be generic, that is modelled by one of the transitions described in Arnold's theorem. Unfortunately proving this fact is not at all easy. The basic problem is that the discriminants are local models for the big wavefront only up to diffeomorphism. So for example for an $A_2$

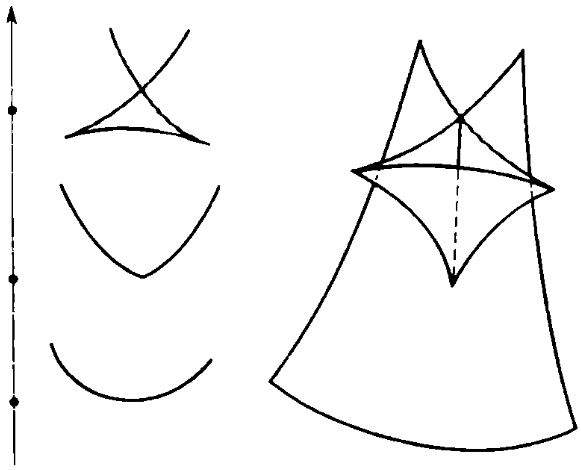


Fig. 2. Swallowtail transition

transition in $R^3$ we clearly need to locate the cuspidal edge and cuspidal tangents of the wavefronts. Techniques for essentially doing this are described in [7] and as an application it is shown that wavefront evolution in the plane and Euclidean 3-space is indeed described locally by Arnold's models. The proof of these facts requires fairly detailed and specific calculations; this is apparently a common feature of the application of Arnold's results. Moreover some of the transitions which one might expect (simply from a count of the conditions imposed by these transitions) do *not* occur. So for example for a generic initial wavefront $M$ in the plane the only transitions are of swallowtail type, which occur at the centres of curvature of vertices, while beaks and lips do not occur. (It is clear that lips transition cannot occur for the wavefront remains two connected curves throughout its evolution.) This feature whereby the geometry rules out certain generic transitions is one we shall meet again. Finally we mention that the self intersection of evolving wavefronts can be described by very much the same methods as used by Arnold; the results are in [8].

## 2. Spheres of varying radii

Returning again to the plane one may ask what one can say about envelopes of families of circles of *varying* radii centred on some curve $M \subset \mathbf{R}^2$. In one sense, of course, very little, for by varying the radius, which we may now take to be a (positive) function $r(s)$ of arc-length, we can obtain almost any curve as the corresponding envelope. In this case we consider the family of functions

$$F(s, a) = \|x(s)-a\|^2 - r^2(s).$$

The function $F(s, a)$ vanishes precisely on the circle centred at $x(s)$ of radius $r(s)$ and the envelope is obtained, as before, by eliminating $s$ from the equations $F = \partial F/\partial s = 0$. Now

$$\frac{\partial F}{\partial s} = 2(x(s)-a)\cdot T(s) - 2r(s)r'(s),$$

where $T(s)$ is the unit tangent vector to our curve at $x(s)$ and $'$ denotes derivatives with respect to $s$. Writing the characteristic points corresponding to the value $s$ of our parameter as $a = x(s)+\alpha T(s)+\beta N(s)$, where $N(s)$ is the standard normal at $x(s)$ for our oriented curve ($N = T$ turned anticlockwise through a right angle) we find that the two conditions $F = \partial F/\partial s = 0$ imply (respectively) $\alpha^2+\beta^2 = r^2(s)$ and $\alpha = -r(s)r'(s)$. Thus, as one expects, the characteristics of the envelope consist of two, one or no points and lie on a line parallel to the normal to the curve at the corresponding value of $s$. Moreover we clearly have two, one or no points according as $1-(r')^2$ is $> 0$, $= 0$ or is $< 0$. (The condition $(r')^2 > 1$ means that consecutive circles are nested.) The same techniques as for wavefronts can be applied to determine the local forms for these envelopes. Again they are generically modelled by discriminants of versal unfoldings of $A_1$ and $A_2$ singularities (envelopes of hypersurfaces are always modelled by discriminants as was observed in [9]; see also [11]). So far however this appears to be little more than an idle generalization – are these envelopes of any real interest?

As we remarked above the resulting envelopes are quite general, indeed we shall see in Section 3 that essentially any plane curve can arise as such an envelope. If we view this envelope as an outline of an object we can consider the changes in the envelope if we fix our initial curve $x(s)$ and vary $r(s)$ with time *or* fix the function $r(s)$ and allow $x(s)$ to vary with time – all the while preserving arc-length along the curve. We can view the first situation as *growth* of our object from an initial spine $x(s)$, while the second can be thought of as *motion* of the object determined by an isotopy of that spine. One can then prove the following ([12]).

THEOREM. (i) *For any plane curve $x(s)$ and generic family of radius functions $r(s, t)$ the only transitions possible in the envelope are lips, beaks and swallowtails (and of course trivial transitions too).*

(ii) *For any generic isotopy of our curve $x(s, t)$ and any family of radius functions $r(s)$ for which $(r'(s))^2-1$ and $r''(s)$ do not vanish simultaneously only swallowtail (and trivial) transitions occur.*

Some further remarks are in order here. First each circle in our envelope bounds a disk and we can consider the union of these disks as a physical object. However, only part of the envelope appears as the frontier of the union; in fact in the case of $A_2$ transitions *none* of the envelope is on the frontier. These are "internal" transitions and are not visible in the overall shape of the object. The swallowtail, on the other hand, corresponds to the birth or death of a corner. See Figure 3. Secondly the conditions $(r'(s))^2 - 1 = r''(s) = 0$ need to be avoided in (ii) since they imply a nontrivial $A_2$ transition which would persist with time $t$, and so be non-generic. Intuitively these conditions mean that "consecutive" circles are tangent to third order.

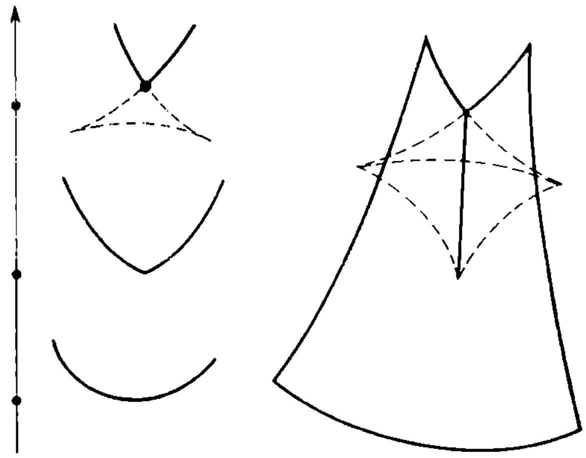


Fig. 3. Emergence of a corner

## 3. Symmetry sets

We have seen how to associate with a smooth curve in the plane, and a family of radius functions, a new curve which is the envelope of the family. Here we consider the reverse process of starting with a curve — let us say a simple closed smooth curve $M$ in $\mathbf{R}^2$ and constructing a second curve with the property that an envelope of circles on this second curve yields $M$. It follows from the tangency property of envelopes mentioned in Section 1 that this is the locus of centres of circles which touch the curve (at least) twice that is bitangent to $M$. To be precise, for a smooth embedding $x$: $S^1 \to \mathbf{R}^2$ we define the *symmetry set* of $M = x(S^1)$ to be the set of points $a$ in $\mathbf{R}^2$ for which either

(1) there exist $s_1, s_2 \in S'$ $(s_1 \neq s_2)$ with $F_a$ singular at $s_1$ and $s_2$, and $F_a(s_1) = F_a(s_2)$ or

(2) there exists $s \in S^1$ with $F_a$ having a singularity of type $A_k$ for $k \geqslant 3$ at $s$.

Thus (1) says that a circle, centre $a$, touches $M$ at $x(s_1)$ and $x(s_2)$. Also (2) allows the points of contact to coincide provided the contact is 4-point at least, and makes the symmetry set closed.

Notice that the definition allows circles such as those in Figure 4; more restrictive definitions have been considered by other authors. For example,

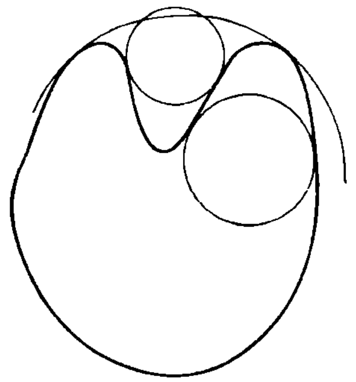


Fig. 4. Circles bitangent to a curve

the *symmetric axis transform*, or sym-ax was introduced by Blum [4] and used in theoretical biology and pattern recognition ([5]) to reduce a 2-dimensional shape to a "skeleton" representing its local symmetries. This allows only circles having interior contact at both points ($C_1$ in Fig. 4). The *central set* ([21]) allows only circles for which the radius is equal to the minimum distance from the centre to the curve ($C_1$ and $C_2$ in Fig. 4).

We use an extended version of the $A_k$ notation to describe the multi-contact of circles with curves. Thus $A_1^2 = A_1 A_1$ stands for ordinary (2-point) contact at two points, $A_1 A_2$ for ordinary contact at one point and osculating (3-point) contact at another, and so on. It is understood that collections of singularities denoted in this way are at the same *level*, that is $F_a$ has the same value at each of them so that it is *one* circle which touches the curve in various places. It is this which makes multigerms more interesting here than for discriminants.

Of course the idea can be generalized to allow $M$ to be any smooth $m$-manifold in $\boldsymbol{R}^n$: we take the centres of $(n-1)$-spheres touching $M$ twice, or having at least $A_3$ contact at some point of $M$. The locus of centres is a union of manifolds of dimension $\leqslant n-1$, and generically there will be an $(n-1)$-dimensional stratum. (An example where this fails is an ordinary circular cylinder in $\boldsymbol{R}^3$.) The cases $m = 1$, $n = 2$ or 3, and $m = 2$, $n = 3$ were studied in [14] and the theory of multi-versal unfoldings was used to obtain local normal forms for all the generically occurring singularity types. For curves in $\boldsymbol{R}^2$ there are only four, illustrated in Figure 5. For surfaces in $\boldsymbol{R}^3$ there are 11 types ([14, Fig. 6]).

One interest in the symmetry set is that its discrete smooth invariants provide a method of comparing shapes (closed curves for example) which is halfway between the (uselessly coarse) notion of diffeomorphism and the standard (uselessly fine) differential geometric invariants of the shapes.(!)

The idea is to regard the symmetry set as part of the *full bifurcation set* $\mathscr{B}(F)$ of $F$, that is the set of points $a \in \boldsymbol{R}^n$ for which either

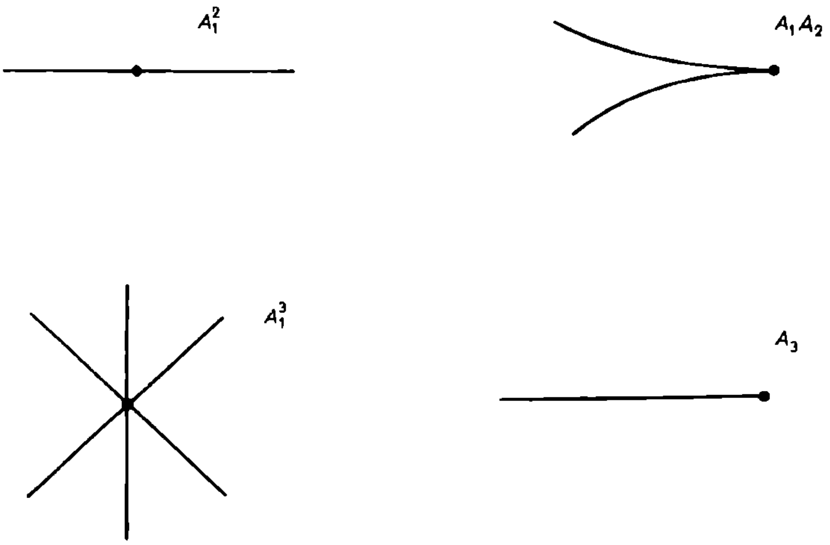



Fig. 5. Local forms for symmetry sets in the plane

(1) $F_a$ is singular at two distinct points of $M$ where $F_a$ has the same value or

(2) $F_a$ has a degenerate singularity at some point of $M$.

The points (2) form the $A_{\geq 2}$ set, or the bifurcation set of $F$, or the *focal set* or *caustic* of $M$, so that

$$\text{full bifurcation set} = \text{symmetry set} \cup \text{focal set}.$$

There is a local uniqueness theorem for full bifuraction sets of versally unfolded germ just as there is for discriminants, where versal here means as a family of potential function [11, Chap. 6]. Moreover the same is true of multi-germ unfoldings. That is, given a collection of simple singularities there is a multi-versal unfolding from which the full bifuraction set is determined; furthermore any two such unfoldings are isomorphic and the full bifuraction sets are locally diffeomorphic. (See [14, Appendix].) As usual, the significance of simplicity is that the type of a multigerm singularity (with factors $A_k$, $D_k$ or $E_k$) determines it up to right-equivalence: there are no moduli present which would result in an infinite number of local models of the full bifurcation set of a given collection of singularity types. For details see [14]; here is an example where there is only one "state" variable, which we can always assume with $A_k$ singularities (cf. Section 1 above).

Example. Singularity type $A_1 A_3$.

Let $F$: $(\boldsymbol{R}, \{0, 1\}) \to \boldsymbol{R}$ be defined by the normal form

$$f(s_1) = s_1^2, \quad s_1 \text{ close to } 0,$$

$$f(1+s_2) = s_2^4, \quad s_2 \text{ close to } 0.$$

We can also write this in the more convenient form

$$f_1(s) = s^2, \quad f_2(s) = s^4, \quad s \text{ close to } 0.$$

As an example of a multiversal unfolding we now have (using the same conventions)

$$F_1(s, a_1, a_2, a_3) = s^2 + a_3,$$

$$F_2(s, a_1, a_2, a_3) = s^4 + s^2 a_1 + s a_2 - a_3.$$

We now want to find the set

$$\mathscr{B}(F) = \{a\colon\ \partial F/\partial s = 0 \text{ at } (s, a) \text{ and } (s', a), \text{ and } F(s, a) = F(s', a) \text{ for some } s, s'; \text{ or } \partial F/\partial s = \partial^2 F/\partial s^2 = 0 \text{ at } (s, a) \text{ for some } s\}.$$

Note that $s$ and $s'$ can both be singular points for $F_2$, or one for $F_1$ and one for $F_2$, for it is possible to have two singularities at the same level near the

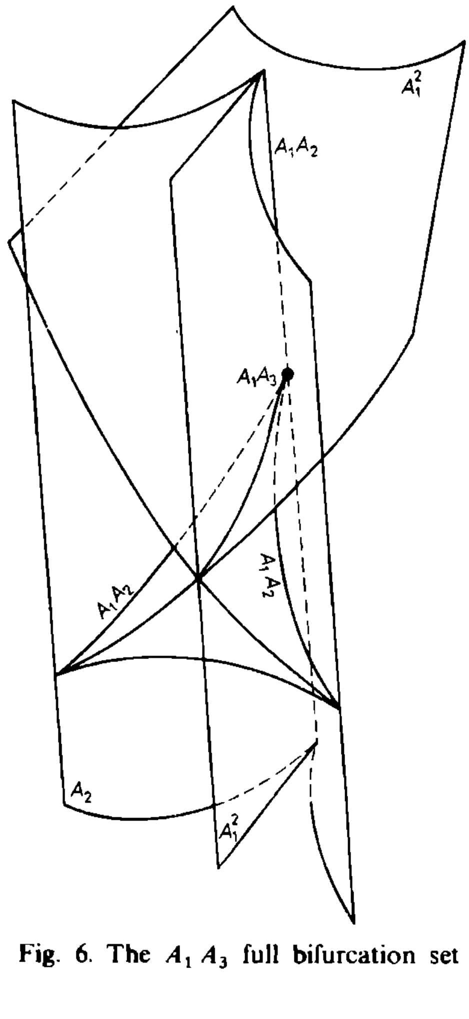



Fig. 6. The $A_1 A_3$ full bifurcation set

$A_3$ singularity. In fact this happens precisely along the line $a_2 = 0$, $a_1 < 0$ and with $s+s' = 0$. Thus the set $\mathscr{B}(F)$ is

$$\{a\colon a_2 = 0,\ a_1 < 0 \text{ or } (s^4+s^2a_1+sa_2 = 2a_3,\ 4s^3+2sa_1+a_2 = 0 \text{ for some } s) \text{ or } (4s^3+2sa_1+a_2 = 6s^2+a_1 = 0 \text{ for some } s)\}$$

This is the union of a half-plane, a swallowtail and a cusp edge (Fig. 6), the cusp edge being the $A_{\geq 2}$ set which corresponds to the focal set when $F$ is the family of distance-squared functions on a surface in $\boldsymbol{R}^3$.

It is clear from the local normal forms in the plane (Fig. 5) that the symmetry set of a generic curve $M$ is not in general a smooth curve: it has cusps and endpoints as well as triple crossings (and of course double crossings, but these correspond to *two* circles of different radius, each tangent to $M$ in two points). Taking the smooth part of the symmetry set we can in fact reconstruct most of $M$ from the envelope of circles which are bitangent to $M$; indeed $M$ is the closure of the envelope [14, § 3.5]. Thus $M$ is determined by its symmetry set and by the radius function on each smooth part of the symmetry set. The same applies to curves and surfaces in $\boldsymbol{R}^3$, and the proof depends on the following observation. Let $p_0 \in M \subset \boldsymbol{R}^n$ where $M$ is compact; then there is an $(n-1)$-sphere or hyperplane in $\boldsymbol{R}^n$ touching $M$ at $p_0$ and at another point $p \in M$, $p \neq p_0$. Presumably this is true for compact $M$ of any codimension in any $\boldsymbol{R}^n$, but we could not prove it.

There is a variant of the symmetry set for curves that is widely used in pattern and shape recognition ([5]). Instead of taking the centre of the bitangent circle, we can take the midpoint of the chord of contact, giving what we have called the *midpoint locus*, but which appears in the above cited papers as the *smoothed local symmetry*. Figure 7 shows the symmetry set and midpoint locus for a cubic oval. Note that the latter ends at the vertices of the oval and does have the appearance of a "spine" around which the oval is built. Unfortunately there does not seem to be any way of regarding the

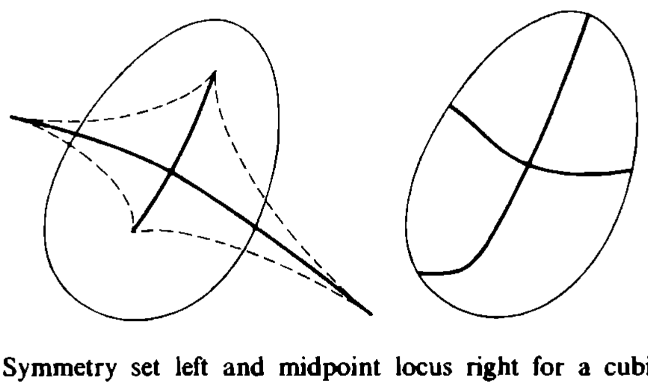


Fig. 7. Symmetry set left and midpoint locus right for a cubic oval

midpoint locus as a discriminant or bifurcation set! For plane curves it is generally smooth, as one can show by elementary arguments ([19]), apart from endpoints and of course self-intersections. It can also be shown that a

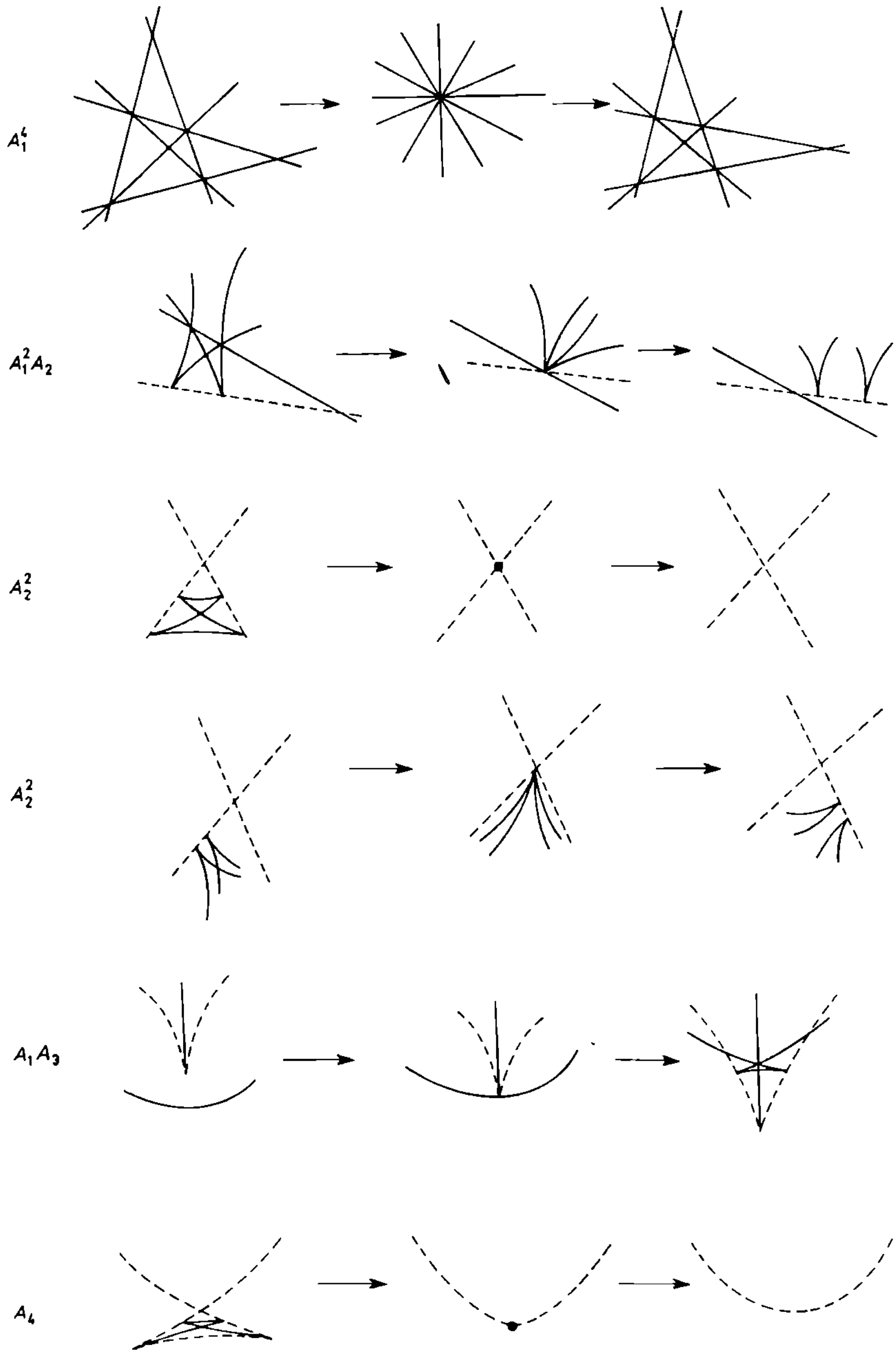



Fig. 8. All the generic transitions on 1-parameter families of symmetry sets in the plane

generic curve $M$ is determined by its midpoint locus and the radius function, for example by first reconstructing the symmetry set from these data.

Finally in this section we consider the changes which occur in the symmetry set of a curve $M$ in $\mathbf{R}^2$ when $M$ moves in a generic 1-parameter family. As for wavefronts, we find these changes by means of a "big full bifurcation set" and generic function on this set. For 1-parameter families of curves in $\mathbf{R}^2$ there are 3 "control variables" (unfolding parameters) and the big full bifuraction sets are 9 of the 11 which arise for surfaces in $\mathbf{R}^3$, corresponding to singularity types $A_1^2$, $A_1A_2$, $A_1^3$, $A_3$, $A_1^4$, $A_1^2A_2$, $A_1A_3$, $A_2^2$ and $A_4$ (the remaining two types, $D_4^+$ and $D_4^-$, cannot occur with only one state variable). *Generic functions* on these sets are found by the method of [10], and the details are in [12]. For a given singularity type the family of sections of the big full bifurcation set corresponding to one of the generic functions may not now be well-defined up to local diffeomorphism but will be up to *stratified local homeomorphism.* This is enough to ensure that, in the pictures of transitions (Fig. 8), smooth pieces, cusps and crossings are correctly represented. (Indeed, for all the sections except the "central" one, passing through the most degenerate singularity, we believe the pictures to be in fact correct to diffeomorphism.) It is particularly interesting that the geometrical constraints of the symmetry set situation, where we work with the family of distance-squared functions, do not allow the occurrence of all the possible transitions on an arbitrary full bifurcation set in $\mathbf{R}^2$. It would be interesting to know of other natural situations where the "missing" examples do occur. One such example is shown in Figure 9; of course the $D_4^+$ and $D_4^-$ do not occur for the reason given above.

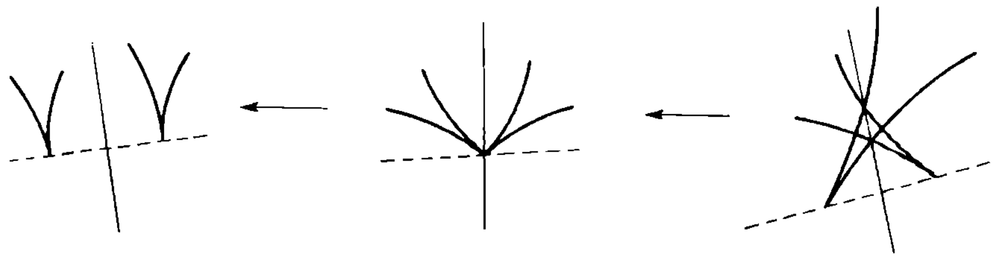


Fig. 9. A transition that does occur on planar symmetry sets

## 4. Caustics

In Section 3 we considered the full bifurcation set of $M^m \subset \mathbf{R}^n$ and noted that it contained the focal set or caustic consisting of points $a \in \mathbf{R}^n$ at which the distance-squared function $F_a$: $M \to \mathbf{R}$ has a degenerate singularity. It is of course possible to make a separate study of this set, and its evolution for a 1-parameter family of manifolds. The reason for the name "caustic" is that the normals to $M$ (think of $M$ as a source of radiation) focus on the caustic,

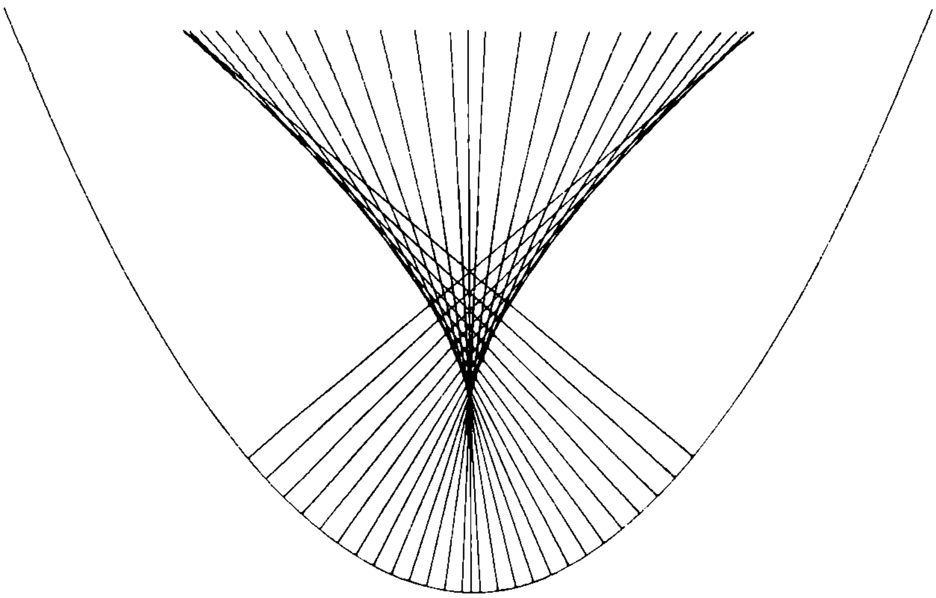


Fig. 10. Normals to a parabola focusing on the evolute

or burning set; see Figure 10 for the normals to a parabola. We shall assume $M$ is a hypersurface ($m = n-1$) in what follows. It is clear that the singular points of wavefronts will all lie on the caustic, for $F_a$ has a degenerate singularity just when the corresponding point of the discriminant of $F$ is singular. In fact as the wavefronts evolve from $M$ their singularities trace out precisely the caustic [15, § 1].

Because a caustic is a bifurcation set and not a discriminant, the evolution of caustics (as $M$ now changes in a 1-parameter family) presents more difficulties than the evolution of wavefronts (for a fixed $M$), except in the case of curves, where the single state variable makes the discriminant of an $A_k$ singularity coincide with the bifurcation set of an $A_{k+1}$ singularity. For curves the only generic transition is the swallowtail transition. Arnold and Zakalyukin [2], [3], [24] have determined the transitions for caustics of 1-parameter families of surfaces in $\mathbf{R}^3$; see also [18]. It is necessary to find diffeomorphisms which preserve a bifurcation set, which is done via the construction of vector fields tangent to the smooth strata of the bifurcation set.

Finally we mention the special case of *caustics by reflexion*. Here we regard $M$ as a mirror (we take $M$ to be convex) and consider a point "source" $L$, incident rays from $L$ being reflected from $M$ (Fig. 11). The focusing curve of the reflected rays is the caustic by reflexion of $M$ relative to $L$. It can in fact be regarded as the caustic (focal set) of an associated manifold called the *orthotomic* of $M$ relative to $L$. A study of caustics by reflexion has been made in a series of papers; see for example [16], [17].

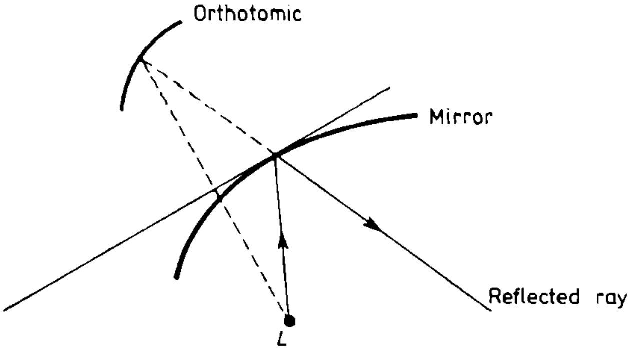



Fig. 11

A natural 2-parameter family of caustics by reflexion arises from a fixed curve $M$ in $\boldsymbol{R}^2$ and all possible positions for $L$ in $\boldsymbol{R}^2$. This has been studied in [13] and the transitions (swallowtail, butterfly) found for a generic $M$.

Returning to the orthotomic we note that this too can be obtained as an envelope of a family of circles! Namely the circles centred on $M$ passing through $L$. Moreover this orthotomic is of independent interest, since it can be viewed as an affine dual of $M$ (see [11, § 5.32], [9]). Thus when $M$ is not convex the orthotomic will have cusps and double points corresponding to inflexions and bitangents of $M$. Families of circles (or spheres if $M$ is a surface in $\boldsymbol{R}^3$) can consequently be used to study the contact of a curve or surface with its tangent lines or planes. They do indeed arise in many different situations!

## References


[1] V. I. Arnold, *Normal forms for functions near degenerate critical points, the Weyl groups $A_k$, $D_k$, $E_k$ and Lagrangian singularities*, Functional Anal. Appl. 6 (1972), 254–273.

[2] –, *Wavefront evolution and equivariant Morse lemma*, Comm. Pure Appl. Math. 29 (1976), 557–582.

[3] –, *Catastrophe Theory*, Springer-Verlag, Berlin 1984.

[4] H. Blum, *Biological shape and visual science*, Part 1, J. Theoret. Biol. 38 (1973), 205–287.

[5] M. Brady, *Criteria for representations of shape*, in Beck and Rosenfeld (eds.), *Human and Machine Vision*, Academic Press, New York 1983.

[6] J. W. Bruce, *Functions on discriminants*, J. London Math. Soc. (2) 30 (1985), 551–567.

[7] –, *Wavefronts and parallels in Euclidean space*, Math. Proc. Cambridge Philos. Soc. 93 (1983), 323–333.

[8] –, *Self intersections of wavefront evolution*, Proc. Roy. Irish Acad. Sect. A 83 (1983), 225–229.

[9] –, *On singularities, envelopes and elementary differential geometry*, Math. Proc. Cambridge Philos. Soc. 89 (1981), 43–48.

[10] –, *Generic functions on semi-algebraic sets*, Quart. J. Math. Oxford (2), 37 (1986), 137–165.

[11] J. W. Bruce and P. J. Giblin, *Curves and Singularities*, Cambridge University Press, Cambridge 1984.

[12] J. W. Bruce and P. J. Giblin, *One parameter families of symmetry sets*, Proc. Roy. Soc. Edinburgh, Sect. A 104 (1986), 179–204.

[13] –, –, *Two parameter families of caustics by reflexion*, Proc. London Math. Soc. (3) 50 (1985), 571–576.

[14] J. W. Bruce, P. J. Giblin and C. G. Gibson, *Symmetry sets*, Proc. Roy. Soc. Edinburgh, Sect. A, to appear.

[15] –, –, –, *On caustics by reflexion*, Topology 21 (1982), 179–199.

[16] –, –, –, *Caustics through the looking glass*, Math. Intelligencer 6 (1984), 47–58.

[17] –, –, –, *Genericity of caustics by reflexion*, Proc. Sympos. Pure Math. 40, American Mathematical Society, Providence 1983, Part 1, 179–193.

[18] J. W. Bruce and D. Fidal, *Generic functions on bifurcation varieties, I*, Preprint, University of Liverpool, 1985.

[19] P. J. Giblin and S. Brassett, *Local symmetry of plane curves*, Amer. Math. Monthly, 92 (1985), 689–707.

[20] E. J. N. Looijenga, *Structural stability of families of $C^\infty$-functions*, Thesis, University of Amsterdam, 1974.

[21] D. Milman, *Eine geometrische Ungleichung und ihre Anwendung* (*Die zentrale Menge des Gebiets und die Erkennung des Gebiets durchsie*), in E. F. Beckenbach (ed.), *General Inequalities*, vol. 2, Int. Series of Numerical Math. 47, Birkhauser, 1980.

[22] I. R. Porteous, *The normal singularities of a submanifold*, J. Differential Geom. 5 (1971), 543–564.

[23] C. T. C. Wall, *Geometric properties of generic differentiable manifolds*, in *Geometry and Topology, III*, Lecture Notes in Math. 597, Springer-Verlag, Berlin 1976, 707–774.

[24] V. M. Zakalyukin, *Bifurcation of front, caustics depending on a parameter*, Itogi Nauki 22 (1983), 53–93.

*Presented to the semester*
*Singularities*
*15 February–15 June, 1985*