1. - THE ZOO OF SINGULARITIES

We shall start by showing a small set of singularities which occur in very different problems. These singularities are as fundamental as, say, the ellipse, the hyperbolas and the parabola. Their occurence in very different theories depends on the universality of the same kind as the universal occurence of quadratic forms in all branches of mathematics and physics.

1.1 - Morse theory of functions

Let us consider a function $y = f(x)$. (Here and always all the functions and mappings are supposed to be sufficiently smooth, say - infinitely differentiable or even analytical, that is developable into a Taylor series convergent to those functions or mappings at a neighbourhood of every point. The main part of the theory remains valid and nontrivial even in the case where all the functions are supposed to be polynomials.)

The singularities we shall discuss are *critical points*. For a typical function the critical points are non-degenerate. For functions depending on n variables (such that x belongs to the n-space $\mathbb{R}^n$ or to an n-dimensional manifold), a critical point of a function is called *nondegenerate*, if its second differential is a nondegenerate quadratic form.

MORSE LEMMA. *In a neighbourhood of a nondegenerate critical point, a function may be reduced to its quadratic part, i.e. it may be written in the normal form*

$$y = -x_1^2 - \ldots - x_k^2 + x_{k+1}^2 + \ldots + x_n^2$$

for a suitable choice of the local coordinate system $(x_1, \ldots, x_n)$, *whose origin is at the critical point (the origin at the axis of the values, y, is supposed to be choosen at the critical value).*

The Morse Lemma explains the occurrence of quadratic forms (and hence of ellipses, of hyperbolas and so on) in most of the problems of calculus, geometry and physics: they are the normal forms of *arbitrary* generic functions in the vicinity of their critical points. (Similar reasoning forms the *raison d'être* of the whole of algebraic geometry: polynomials are either local approximations or the local normal forms of arbitrary functions or mappings.)

In the functional space of all functions, degenerate functions form a hypersurface (a submanifold of codimension one, that is defined by one equation). This hypersurface is called the *bifurcation set*. The bifurcation set is not smooth: it consists generally of smooth manifolds of different dimensions, adjacent to each other in a very special way. The study of the structure of this bifurcation set (and of similar bifurcation sets in other problems) is the main content of singularity theory.

A typical function has only nondegenerate critical points, and its local structure is completely described by the Morse lemma, if we consider two functions equivalent when they are transformed one into another by a smooth change of the independent and of the dependent variables.

Globally, the set of critical points does not define a function up to equivalence (even in the case when we know the index k, which allows us to distinguish the maxima, the minima and different kinds of saddle-points). This is true even for functions of one variable.

PROBLEM. Find the number $K(n)$ of pairwise nonequivalent functions of one variable having n nondegenerate critical points with pairwise different critical values, supposing that at infinity the function behaves like x for even n and like x^2 for odd n.

ANSWER. $K = 1, 1, 1, 2, 5, 16, \ldots$ for $n = 0, 1, \ldots$;

$$\sum K(n)\frac{t^n}{n!} = \sec t + \tan t.$$

[REMARK (for the experts). This result shows, for instance, that Euler numbers and Bernoulli numbers together form a single sequence. It also opens the way for the definition of Euler and Bernoulli numbers associated to any simple Lie algebra (the usual case corresponding to unitary groups), or even to general singularities.]

The bifurcation set divides the function space into components. Two functions in the same component will be equivalent to each other if we include in the bifurcation set two parts: the hypersurface of functions having degenerate critical points and the hypersurface of functions with coinciding critical values. The first hypersurface is called the *caustic* (in the corresponding optical problem it is the place of the concentration of light). The second hypersurface is called the *Maxwell set* (because of the Maxwell rule in phase transition theory; this rule is the condition for the coincidence of two critical values of some function).

PROBLEM. Find the number of parts into which the full bifurcation set (consisting of functions with degenerate critical points or multiple critical values) divides the space of polynomials of the form

$$x^{n+1} + a_1x^n + \ldots + a_{n+1}.$$

ANSWER. $K(n) + K(n-2) + K(n-4) + \ldots = 1, 1, 2, 3, 7, 19, \ldots$ for $n = 0, 1, \ldots$.

EXAMPLE. The plane of polynomials $x^5 - x^3 + ax^2 + bx$ is divided by the bifurcation set into 7 parts, shown in fig. 1.

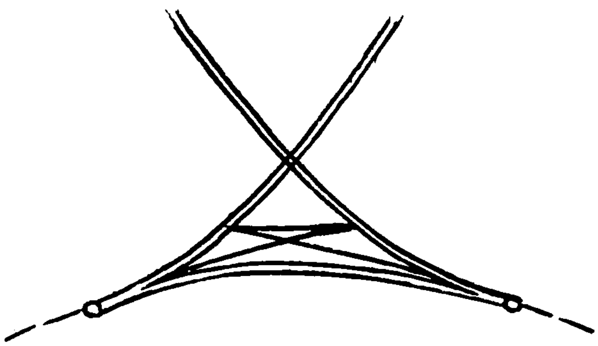

Fig. 1 - The caustic and the Maxwell stratum of a family of functions.

In this case the caustic (shown in fig. 1 as a double line) and the Maxwell set (completed by its analytical continuation, which is represented in fig. 1 by a hatched line) are diffeomorphic: these two plane curves can be transformed one into the other by a diffeomorphism of the plane (each curve has two cusps and one self-intersection point).

The decomposition of the bifurcation set into submanifolds of different dimensions is represented in fig. 1 by the decomposition of the curve into 6 points and 10 intervals.

A generic function has neither degenerate critical points nor multiple critical values. However, degenerate points and multiple critical values occur unavoidably in the families of functions depending on parameters.

Let us consider, for instance, one-parameter families. A one-parameter family is represented in function space by a curve. This curve may intersect the bifurcation hypersurface. If the intersection is transversal ("the angle between the curve and the hypersurface" being nonzero; and the intersection point being a generic point of the hypersurface), then the intersection is stable (it cannot be destroyed by a small variation of the family). Of course, for a neighbouring family, the intersection will happen for a slightly different value of the parameter, when compared with the non perturbed family, and the point of intersection itself is slightly different. But it is impossible to remove the intersection with the bifurcation set for all the values of the parameter simultaneously by a small variation of the family.

What can be achieved by a small variation of a one-parameter family is transversality: we may force the curve representing the family to intersect the bifurcation hypersurface in its generic points only and "at nonzero angles".

In fig. 1 a curve representing a typical one-parameter family, should not contain any of the 6 singular points of the bifurcation curve and should not be tangent (neither to the caustic curve nor to the Maxwell curve). Of course, those events (tangency or passage through a singular point) are possible for some special families. But we may destroy them by a small variation of the family.

The same way a k-parameter family is represented by a k-dimensional submanifold of the function space. A typical k-parameter family intersects only those parts (strata) of the natural decomposition of the bifurcation set whose codimension is k or smaller. And the intersections are transversal (no tangency). Hence the trace of the bifurcation set on the k-manifold, representing the family, provides a correct picture of the singularity of the bifurcation set at the stratum corresponding to the point of intersection. This trace is called the *bifurcation diagram of the family.* It may be considered as living in the parameter space of the family.

EXAMPLE. The two-parameter family of functions in fig. 1 is transversal to the bifurcation set in the function space. Fig. 1 provides a correct picture of the singularities of the bifurcation set at its strata of codimension 1 and 2.

Thus, in a typical one-parameter family of functions, degenerate critical points occur for some special values of the parameter. Those degenerate critical points correspond to a transversal intersection of the curve, representing the family in function space, with the caustic.

In a neighbourhood of such a point, the function of one variable may be reduced (by a smooth change of variables) to the normal form $y = x^3$.

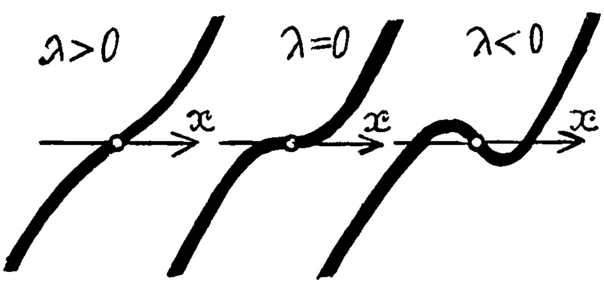

Fig. 2 - A perestroika of the birth of two critical points.

Moreover, the whole family may be reduced to the normal form of fig. 2,

$$y = x^2 + \lambda x$$

where λ is the parameter (the reduction is achieved by a smooth change of the parameter and a smooth change of coordinates, which depend smoothly on the parameter).

[In the case of more than 1 independent variable, to the normal form x^3 we have to add a nondegenerate quadratic form of the other variables, so the final normal form of the family is

$$y = -x_1^2 - \ldots - x_k^2 + x_{k+1}^2 + \ldots + x_{n-1}^2 + x_n^3 + \lambda x_n.]$$

Thus, the transversal intersection of a one-parameter family of functions with the caustic implies the birth or death of a pair of neighbouring critical points, whose indexes differ by one (whether the birth or the death takes place depends on the direction of the variation of the parameter).

In a typical two-parameter family, a further degeneration may occur. The corresponding surface may pass through the (simplest) singular points of the caustic. The corresponding normal form (for the families of functions of one variable) is

$$y = x^4 + \lambda_1 x^2 + \lambda_2 x$$

(if the number of independent variables is larger, one has to add a nondegenerate quadratic form of the other variables).

The bifurcation diagram of the family (the trace of the caustic on the plane of the parameters (λ_1, λ_2)) is a semicubical parabola $\Delta = 0$, where $\Delta = 8\lambda_1^3 + 27\lambda_2^2$. It means that the bifurcation set in function space has a codimension two stratum (the "cusped edge") of a semicubical type. (The adjacent bifurcation hypersurface intersects the two-dimensional surface transversal to the codimension two stratum along a curve having a semicubical cusp point - the bifurcation diagram, also called in this case the *caustic* of the family, which is represented by the surface.)

For typical 3-parameter families a further degeneration occurs stably (corresponding to the codimension 3 stratum of the caustic). The corresponding normal form of the family (of functions of one variable) is

$$y = x^5 + \lambda_1 x^3 + \lambda_2 x^2 + \lambda_3 x.$$

The bifurcation diagram (the caustic), in the space of parameters $\{(\lambda_1, \lambda_2, \lambda_3)\}$, is a surface with a singularity shown in fig. 3, studied by Kroneker and called the *swallowtail*.

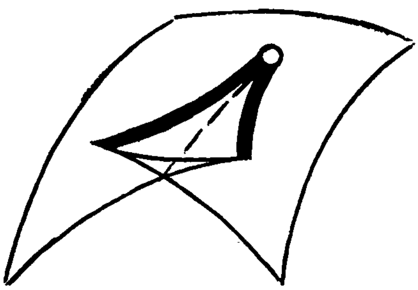

Fig. 3 - The swallowtail.

Its section by the plane $\lambda_1 = -1$ is represented in fig. 1. It is interesting to note that the trace of the Maxwell set (continued analytically) is diffeomorphic to the trace of the caustic: the trace of the Maxwell set in the space of parameters $(\lambda_1, \lambda_2, \lambda_3)$ is also a swallowtail (compare fig. 1).

The topological properties of different strata of the natural stratification of the space of functions on a manifold are, in principle, the invariants of the smooth structure of the manifold. But very little is known about these topological properties.

EXAMPLE. The space of functions on a line, coinciding with x at infinity and having no critical points more complicated than x^3, is weakly homotopy equivalent to the loop space of the two-sphere, ΩS^2.

Let us associate to any point x the vector $(f'(x),\ f''(x),\ f'''(x))$ of 3-space. Since this vector is nowhere zero, we obtain a mapping from the axis of x to the two-sphere. The conditions at infinity imply that this mapping defines a loop. We have thus constructed a mapping from our function space to the space of loops. This mapping happens to be a weak homotopy equivalence.

REMARK (for the experts). Thus, in this case, the "principle of nonrelevance of integrability conditions" of Smale-Gromov holds, while this does not follow from any known general theorem (of course, replacing x^3 by x^n we have to substitute S^{n-1} to S^2, see [1]).

1.2 - Whitney theory of mappings

H. Whitney found in 1955 the local normal form of the singularities of typical mappings from two-dimensional manifolds to the plane (or to another two-dimensional manifold):

$$y_1 = x_1^3 + x_2 x_1, \;\; y_2 = x_2.$$

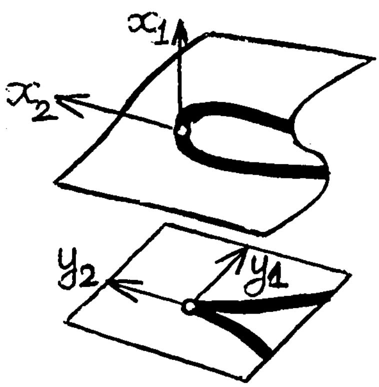

Fig. 4 - The Whitney tuck mapping.

EXAMPLE. Let us consider the surface, defined by the first equation (fig. 4) in the 3-space with coordinates (x_1, x_2, y_1). Its projection to the (y_1, x_2) plane along the x_1 axis has the required singularity. This singularity is stable (any neighbouring mapping has, at some neighbouring point, an equivalent singularity, i.e. a singularity which may be reduced to the same normal form by some smooth change of the independent and of the dependent variables).

The Whitney formula may be considered as defining a one-parameter family of functions $y_1(x_1)$, depending on $x_2(= y_2)$ as a parameter. This family experiences a perestroika of fig. 2 at the value 0 of the parameter (two critical points are born when the parameters move from the positive to the negative values).

The set of critical points of the Whitney mapping (where its jacobian vanishes) is a smooth curve (a parabola on the plane of independent variables x). The set of critical values is a curve with a semicubical cusp on the plane of the dependent variables y. For the projection mapping of fig. 4, the critical points are the points of the surface where the tangent plane is vertical, and the critical values form the "apparent contour" of the surface seen from above (we suppose the surface to be transparent).

The apparent contours of the generic smooth surfaces (projected along the typical directions) have no other singularities besides the Whitney cusps (and the obvious self-intersections). In particular, one can find the cusps at the apparent contours of the faces of people (fig. 5), but we usually do not perceive them since the faces are not transparent.

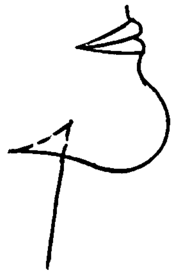

Fig. 5 - A singularity of an apparent contour.

A semicubical singularity is usually a sign of a Whitney mapping hidden somewhere in the neighbourhood. The generic singularities of typical mappings of three-dimensional manifolds to the plane may be considered as being the combinations of Morse and Whitney singularities. The normal form is

$$y_1 = x_1^3 + x_1 x_2 + x_3^2, \quad y_2 = x_2.$$

The set of critical values is still the semicubical parabola in the plane of dependent variables y. The preimages of the values of y, which do not lie on this parabola, are smooth elliptic curves on the planes $\{x_1, x_2, x_3 : x_2 = \text{const}\}$. The variables y may be considered as presenting the parameters of this family of curves. The semicubical parabola of the critical values is the bifurcation diagram of this family of curves.

In the same way, the semicubical parabola of the critical values of the Whitney mapping from a surface to the plane may be viewed as being the bifurcation diagram of the corresponding family of the 0-dimensional subsets of the line (the subsets being the intersections of the vertical lines, along which we project the surface, with the surface).

EXAMPLE. Let us consider the family of normals to an Euclidean plane curve, say, to an ellipse (fig. 6).

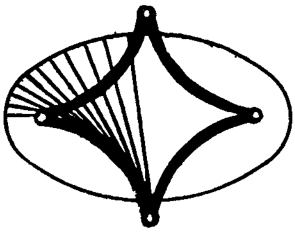

Fig. 6 - The caustic of an ellipse.

This family has an envelope (which is an astroide in the case of an ellipse). The envelope (called also the *caustic* of the curve, its *evolute*, its *focal set* and its *curvature centre set*) has, generically, singular points - semicubical cusps. These singularities of the caustics are stable: they do not disappear under a small variation of the initial generic smooth curve, they move smoothly with it.

As semicubical cusps are the signs of Whitney singularities, we seek for the corresponding surface mapping. In this case, it is the *normal mapping* associated to the curve. The source surface is the manifold of all the vectors normal to the curve. This 2-manifold is called the *normal bundle space* of the curve. The normal mapping associates to a normal vector (considered as being an arrow in the plane starting from a point of the curve) its end point.

A critical point of the normal mapping is a vector joining a point of the curve with the curvature centre of the curve at that point. The critical values form consequently the curvature centre set of the curve. But the critical values of the normal mapping are also the intersection points of infinitesimally close normals of the curve. Hence, the critical value set of a generic curve is the envelope of its normals.

The normal mapping of a generic curve has only the singularities typical for the generic mappings between two-dimensional manifolds, that is only the Whitney singularities: the "folds" have the normal form

$$y_1 = x_1^2, \;\; y_2 = x_2$$

displayed as the caustic line, and the "cusps" (or "tucks") of fig. 4, displayed as the cusps of the caustic line.

WARNING. The genericity of the singularities of the normal mapping associated to a generic curve is not evident. Indeed, the normal mappings associated to curves form a narrow subclass (of infinite codimension) in the functional space of all the surface mappings. This might lead, in principle, to the following two phenomena. 1°. Some singularities, typical for the general surface mappings, might become nontypical (or even impossible) for normal mappings. 2°. Some singularities, typical for normal mappings, might become nontypical for general surface mappings (they might be destroyed by arbitrary small perturbations in the class of general surface mappings, but not in the class of the normal mappings).

While neither of these two phenomena does occur for the normal mappings associated to plane curves, the situation in higher dimensions is different. Indeed, the typical singularities of the caustics (or of the curvature centre varieties) of dimension 2 or more, differ from the typical singularities of the critical sets of the generic mappings of the manifolds of the corresponding dimension (n for the curvature centre varieties of the hypersurfaces in Euclidean or Riemannian n-spaces). The singularities of the normal mappings are better understood in the theory of the Lagrange singularities of symplectic geometry [2], [3].

1.3 - The Whitney-Cayley umbrella

The normal form of a typical mapping of a surface $\{(u, v)\}$ into the three-space $\{(x, y, z)\}$ was also found by Whitney:

$$x = u, \ \ y = uv, \ \ z = v^2.$$

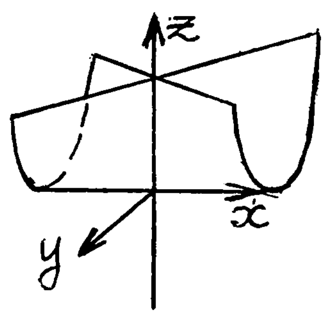

Fig. 7 - The Whitney-Cayley umbrella.

The image of this mapping is the surface $y^2 = zx^2$ (fig. 7), called the Whitney umbrella (or the Cayley umbrella, since this and most of the other important singularities first appeared in the works of Cayley).

Strictly speaking, this eccentric umbrella (which is not really any use in rain) also contains a handle $\{x = y = 0, \ \ z < 0\}$, which is not included in the image of the real plane $\{(u, v)\}$ under the above Whitney mapping.

This mapping has just one critical point (the origin). The critical value is also the origin ($x = y = z = 0$). The other points of the (positive) z axis are the points where a transversal intersection of two "smooth branches" of the umbrella surface occurs. Each self-intersection point has two preimages at the plane of the independent variables (u, v). In a neighbourhood of each of these preimage points, the mapping is nondegenerate (it defines a local embedding of a plane in 3-space).

It is interesting to note that the surfaces of the swallowtail (fig. 3) and of the umbrella (fig. 7) are homeomorphic (topologically equivalent). Moreover, there exists a homeomorphism (a one-to-one continuous mapping whose inverse mapping is also continuous) of the ambient 3-space, transforming the swallowtail surface into the umbrella surface.

This homeomorphism is a manifestation of general multidimensional phenomena (B.A. Hessin, B.Z. Shapiro, 1989).

The umbrella surface is related to the perestroikas of the projections of