Working Paper No.274.

AN INTRODUCTION TO ELEMENTARY CATASTROPHE THEORY, WITH APPLICATIONS IN URBAN PLANNING.

John R. Beaumont and Martin Clarke.

School of Geography,
University of Leeds,
LEEDS LS2 9JT.

March 1980.

CONTENTS:

1. INTRODUCTION.

The use of mathematical models is now well established as a tool for the analysis of socio-economic systems. Over the past twenty years attention has mainly focussed on static or at best cross-sectional representations. However, although some dynamical models have been formulated, recent developments in dynamical systems analysis suggest alternative and new avenues for exploration. During the 1970's elementary catastrophe theory, widely promoted by Thom (1975) and Zeeman (1977), has generated much interest in both the physical and social sciences. In this paper we outline in a relatively straight-forward fashion the main concepts of this theory and investigate the possible applications and limitations in urban planning.

The traditional mathematical technique for comprehending dynamic behaviour is *calculus*, as developed by Newton and Leibniz in the seventeenth century. Calculus is, however, limited in the types of behaviour it can handle. Specifically, it is generally restricted to those systems displaying continuously smooth behaviour. Given a function, for example,

$$y = f(x), \tag{1.1}$$

this means that, calculus is restricted to the set of functions, $\underline{f}$, that, for a smooth change in x, result in a smooth change in y. These types of functions are illustrated in Figure 1.

Functions that describe discontinuous behaviour, of the type illustrated in Figure 2, are, therefore, excluded from this type of analysis.

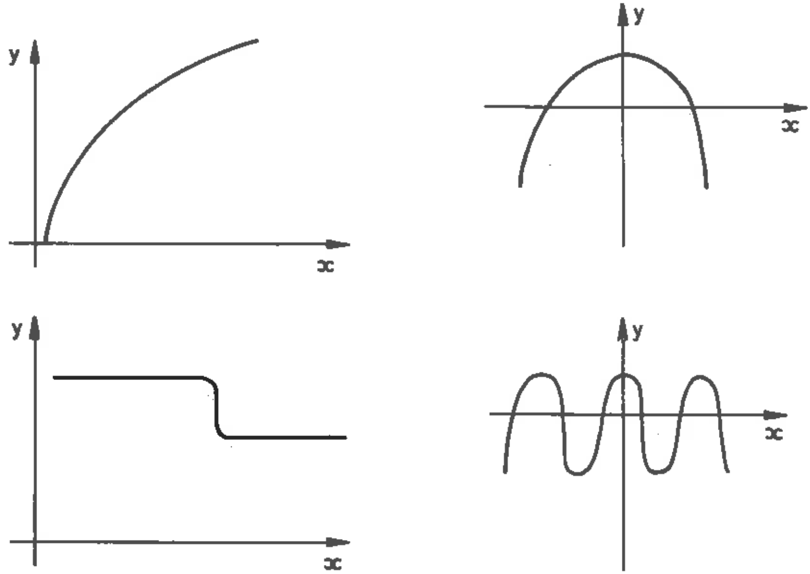


Fig. 1

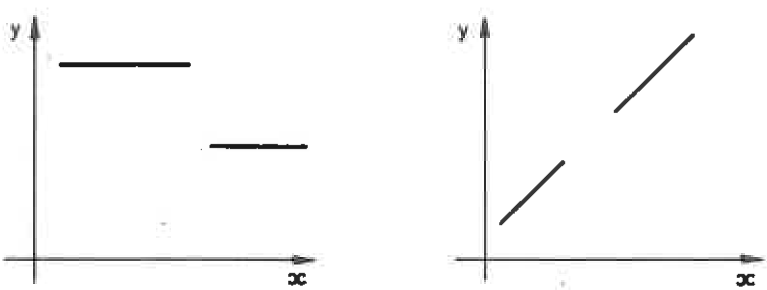


Fig. 2

In the real world, it is clear that both continuous and discontinuous behaviour are common. An everyday example of discontinuous change is a kettle boiling - the phase transition between water and steam arising from the steady application of heat.

Rene Thom's book "Structural Stability and Morphogenesis", published in French in 1972† and translated into English in 1975, offers a profound model of discontinuous changes in natural phenomena. Thom's main interests are in biological processes, but subsequently a wide range of applications of catastrophe theory have appeared, with the subject matter being as diverse as anorexa nervosa and literary symbolism, the stock market, prison riots, laser action, and so on. These applications have often generated novel insights and have led to interesting, if not uncontentious, predictions of potential future behaviours. (For a discussion of these, and other examples, the reader is referred to the books by Poston and Stewart, 1978; Woodcock and Davis, 1978; and Zeeman, 1977 . In addition, a number of selected papers are given in the bibliography). A useful pedagogic device to visually illustrate catastrophe theory is Zeeman's so-called "Catastrophe Machine", the construction of which is described in Zeeman (1972).

Recently, catastrophe theory has attracted the attention of a number of geographers (for example, Wagstaff, 1976; Baker, 1977; Area, 1979; Wilson, 1980), economists (for example, Balasko, 1978;

---

† It is somewhat unfortunate that many people associate catastrophe theory with natural disasters or dramatic change. The French word '*catastrophe*' translates as 'sudden change', for example, 'a sudden change in the weather'.

Varian, 1979) and regional scientists (Isard and Liossatos, 1979). In Section 5 we discuss in some detail three different examples, highlighting a number of features of catastrophe theory. Further examples have employed the framework to examine urban growth (Amison, 1972) and Wagstaff (1978) has presented a novel description of the change of settlement patterns in the Mani peninsula, Greece, between the second and seventeenth century; Mees (1975) has studies the revival of cities in Medieval Europe.

The purpose here is to present an elementary introduction to the main ideas that catastrophe theory offers. This should enable geographers and planners to form their own opinion as to the usefulness or otherwise of the approach, and to be aware of the applications that may exist in the field. The mathematics and terminology of catastrophe theory, especially those found in Thom's book, are, without doubt, difficult. We have restricted the mathematics in our account to the absolute minimum, and in doing so have sacrificed the rigour of mathematical proofs for the ease of introducing the basic concepts. In addition, the pictorial representation that accompanies the mathematics reinforces the viewpoint that the results of the complex proofs can be stated simply and are relatively easy to understand and to apply.

In Section 2, we present, in a straightforward manner, one of the substantial underpinnings of catastrophe theory, that of *structural stability*. (Anyone with a working knowledge of differentiation should easily be able to follow this exposition). In Section 3, Thom's main theorem is presented, by describing the seven elementary catastrophes. The first two of these, the well-known *fold* and *cusp* catastrophes, are then given a more detailed, but still elementary, treatment to enable the reader to grasp the underlying features of catastrophe theory.

In Section 4, there is a general discussion of planning and possible applications of catastrophe theory; this forms a foundation for a detailed examination of specific examples undertaken in Section 5. Some of the criticisms levelled at the application of catastrophe theory in the social sciences are briefly considered in Section 6, and further extension of the methodology are suggested in Section 7.

2. SOME BASIC CONCEPTS.

At the outset, it is useful to fix ideas on the basic representation which will be employed here; a brief resumé of elementary mathematics is also required. Concern, for instance, focuses on the behaviour of a particular system, represented by a set of state variables $\underline{x}$, in relation to its governing controls represented by a set of control parameters $\underline{\alpha}$. A system can be described as consisting of n state variables, $x_i (i = 1, \ldots, n)$, and m control parameters, $\alpha_j (j = 1, \ldots, m)$. It is commonplace in the exposition of catastrophe theory to visualize the system as determined by the maximisation (or with appropriate changes, the minimization) of a potential function (f).

$$\max_{\underline{x}} f(\underline{x}, \underline{\alpha}). \tag{2.1}$$

Catastrophe theory is concerned with the analysis of the equilibrium (steady states, critical points, or stationary points) of a system. A stationary point exists (and hence the solution to equation (2.1)) when

$$\frac{\partial f(\underline{x}, \underline{\alpha})}{\partial \underline{x}} = 0 \tag{2.2}$$

that is, when the derivative equals zero. As an example, we shall calculate the stationary point of a simple function, f(x). Let

$$f(x) = x^2 \tag{2.3}$$

which can be differentiated with respect to x to produce

$$\frac{\partial f(x)}{\partial x} = 2x. \tag{2.4}$$

A function's optimum (maximum or minimum) occurs at the point where its gradient or first order derivative is zero. Thus, from equation (2.4) $f(x) = x^2$ has an optimum point at $x = 0$ (see Figure 3) because this is the only value of x which will satisfy the condition $2x = 0$. To ascertain whether the point is a maximum or minimum point (or something else), it is necessary to look at the sign of the second order derivative $(\partial^2 f(x)/\partial x^2)$; the conditions are described in Table 1.

TABLE 1: Minimum and Maximum Points, and Points of Inflexion.

Type of critical point	$\frac{\partial f(x)}{\partial x}$	$\frac{\partial^2 f(x)}{\partial x^2}$
minimum	0	>0
maximum	0	<0
point of inflexion	0	0

It should be noted that the minimum point of a function, $f(x)$, is also the maximum point of the function, $-f(x)$; for example, $f(x) = x^2$ is illustrated in Figure 3 (and it is left as an exercise for the reader to demonstrate that the maximum point is at x equals zero). It is, therefore, possible to restrict our discussion to minimum points without loss of generality.

To understand catastrophe theory, it is necessary to appreciate that its basic concept is that of *structural stability*. A system is deemed to be structurally stable if a minor alteration of a control parameter does not produce a change in the overall form taken by the function of the state variables. The notion of a system's stability of form can be

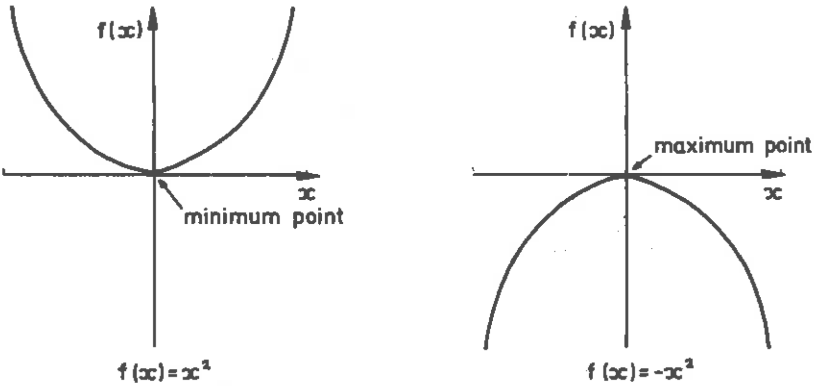


Fig. 3

illustrated by considering the shape of the simple functions shown in Figure 4.

From Figure 4, we know the form of the curve $f(x) = x^2$. Figure 4 portrays a whole family of curves for the function,

$$f(x) = x^2 + a_i \tag{2.5}$$

where a_i is any real constant. We note that the minimum point is always at $x = 0$, because the constant term is dropped in differentiation.

In contrast, consider the function,

$$f(x) = x^3. \tag{2.6}$$

By employing the general rule for differentiation, it is easily demonstrated that

$$\frac{\partial f(x)}{\partial x} = 3x^2 \tag{2.7}$$

which equals zero when $x = 0$. In addition it is known that

$$\frac{\partial^2 f(x)}{\partial x^2} = 6x \tag{2.8}$$

which also equals zero when $x = 0$ (this is a singular point). From Table 1, we know that $x = 0$ is, therefore, a point of inflexion, and this is plotted in Figure 5.

If we now consider the function,

$$g(x) = x^3 - ax \tag{2.9}$$

where a is any real constant, we know that

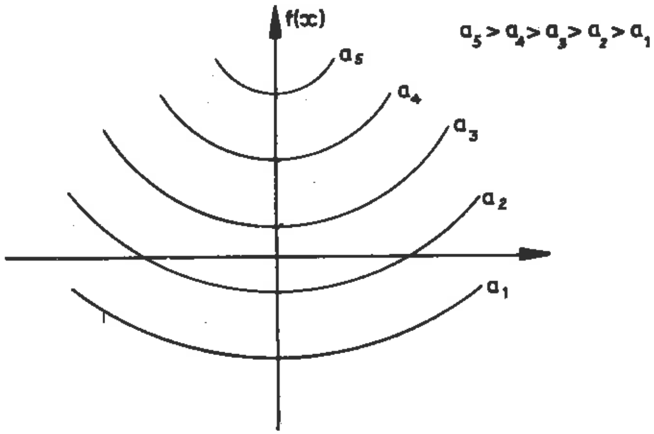


Fig. 4

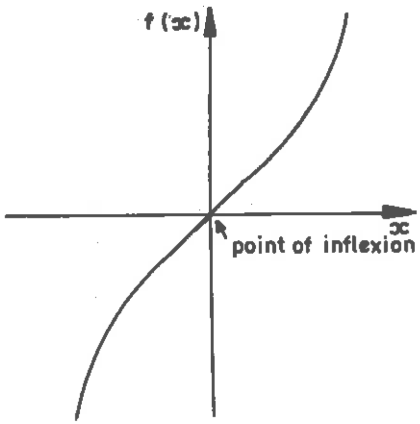


Fig.5

$$\frac{\partial g(x)}{\partial x} = 3x^2 - a. \qquad (2.10)$$

It is possible to examine three different cases, depending on the value of the constant:

(i) if <u>a = 0</u>, the function, $g(x)$, is the same as the function $f(x)$, given by equation (2.6).

(ii) if <u>a > 0</u>, equating equation (2.10) to zero gives

$$3x^2 - a = 0 \qquad (2.11)$$

which can be simplified to show that the function $g(x)$, has two stationary points, $x = +\sqrt{\frac{a}{3}}$ and $x = -\sqrt{\frac{a}{3}}$.

(iii) if <u>a < 0</u>, the function $g(x)$ has no stationary points, because there is no real solution to the problem of finding the square root of a negative number,

$$x^2 = \frac{-a}{3}. \qquad (2.12)$$

These three cases are described in Figure 6, and it is clearly demonstrated that the forms of the curves are different. (It can be noted that

$$\frac{\partial^2 g(x)}{\partial x^2} = 6x \qquad (2.13)$$

and, therefore, as depicted in Figure 6(ii), $x = +\sqrt{\frac{a}{3}}$ is a minimum point, and $x = -\sqrt{\frac{a}{3}}$ is a maximum point). Thus, the fact that a change in a function's stationary points can occur for a small alteration in a parameter has been graphically represented; a more rigorous mathematical means of analysing the (local) stability of a singularity is by ascertaining whether the *determinant* of the function's *Hessian matrix* is zero or not (that is, whether the matrix is singular or not). A Hessian

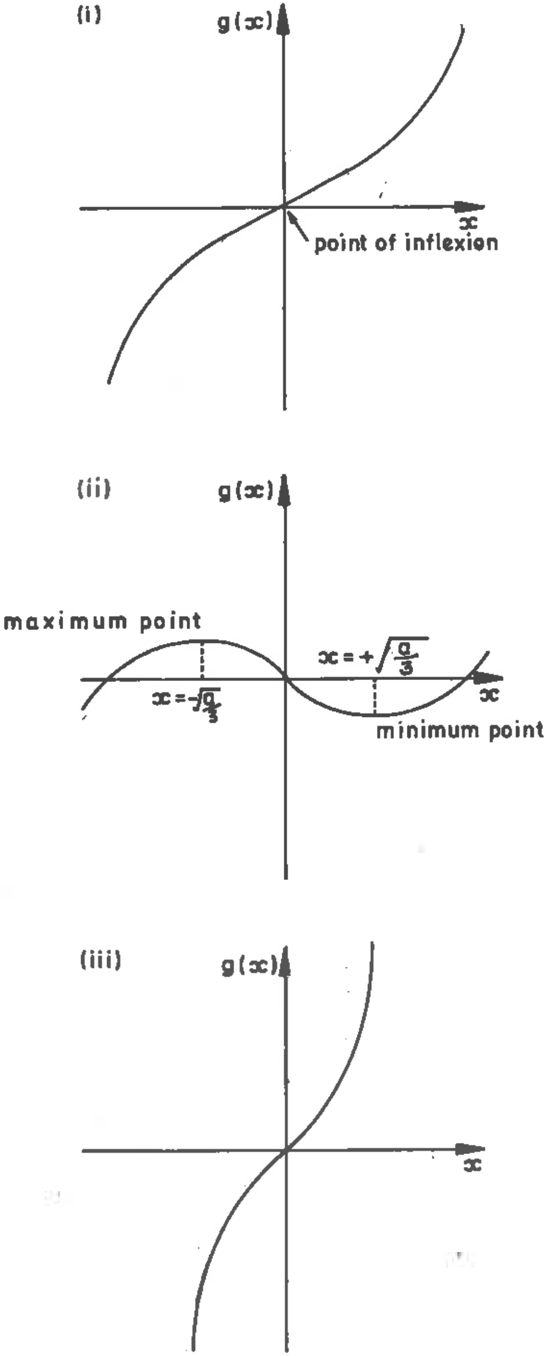

(i)
g(x)
x
point of inflexion
(ii)
g(x)
maximum point
x = +√(a/3)
x = −√(a/3)
x
minimum point
(iii)
g(x)
x

Fig.6

matrix, $H(\underline{x}^*)$, is the matrix of second order partial derivatives of the function $f(\underline{x})$ at $\underline{x}^*$ (which is a stationary point),

$$H(\underline{x}^*) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\underline{x}^*) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x^*) \ldots\ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x^*) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x^*) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x^*) \ldots\ldots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x^*) \end{bmatrix} \quad (2.14)$$

Expressed in formal terms we can say if the Hessian matrix at $\underline{x}^*$, $H(\underline{x}^*)$, is a singular matrix, the function, $f(\underline{x})$ has a *degenerate singularity* at $\underline{x}^*$. (Beware of the confusing terminology - a singular point is different from a singular matrix). That is, by definition a critical point $\underline{x}^*$ of a function, $f(\underline{x})$, is called degenerate or non-degenerate depending on whether the determinant of a Hessian matrix of $f(\underline{x})$ at $\underline{x}^*$ vanishes or not. It can be shown that non-degenerate singularities are structurally stable under small perturbations, and it is the degenerate singularities which catastrophe theory addresses.

Using matrix algebra it is possible to determine critical points and in so doing it enhances the interpretation of system behaviour. The product of the eigenvalues of a matrix is the matrix's determinant; this has the interesting corollary that a degenerate singularity occurs when at least one of the eigenvalues is zero. It is therefore possible to examine changes in the values of individual eigenvalues for alterations in the parameter values (see Beaumont and Clarke, 1980).

It should be noted that computer programmes are available to calculate the determinant of a given matrix, and we describe how to calculate the determinants of a 2 x 2 matrix and a 3 x 3 matrix for illustrative purposes, remembering from elementary matrix algebra, that a singular matrix has no inverse.

$$\text{Det } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc. \tag{2.15}$$

$$\text{Det } B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

$$= b_{11} \text{ Det } \begin{bmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{bmatrix} - b_{12} \text{ Det } \begin{bmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{bmatrix} + b_{13} \begin{bmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \tag{2.16}$$

The locus of degenerate singularities defines the critical area, the so-called 'bifurcation set', in which "catastrophic" behaviour can be recognised. The above representation is a multivariate description which for the simple function

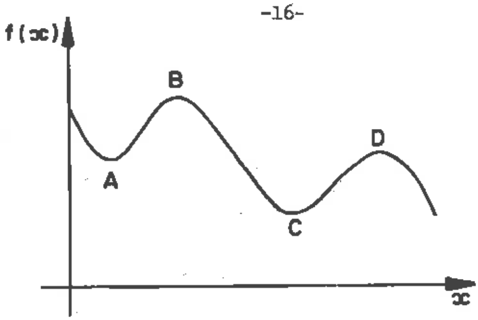
$$f(x_i, \alpha), \quad i = 1 \tag{2.17}$$

is equivalent to the disappearance of the second derivative at a point of inflexion. This feature is mirrored in the later discussion of Poston and Wilson's (1977) work in which the alteration of local minimum is associated with a point of inflexion.

Before describing the power of Thom's theorem, it is necessary to point out that catastrophe theory inexorably involves systems that have multiple steady states and that it involves a *local* analysis of these critical points. For example, a function may have the form pictured in Figure 7, in which it is clearly recognised that A, B, C and D are all stationary points: A and C are minimum points, and B and D are maximum points. Now point A, for instance, is a *local* minimum with regard to the function's values in its immediate neighbourhood; with regard to the whole function, it is not the function's lowest value.

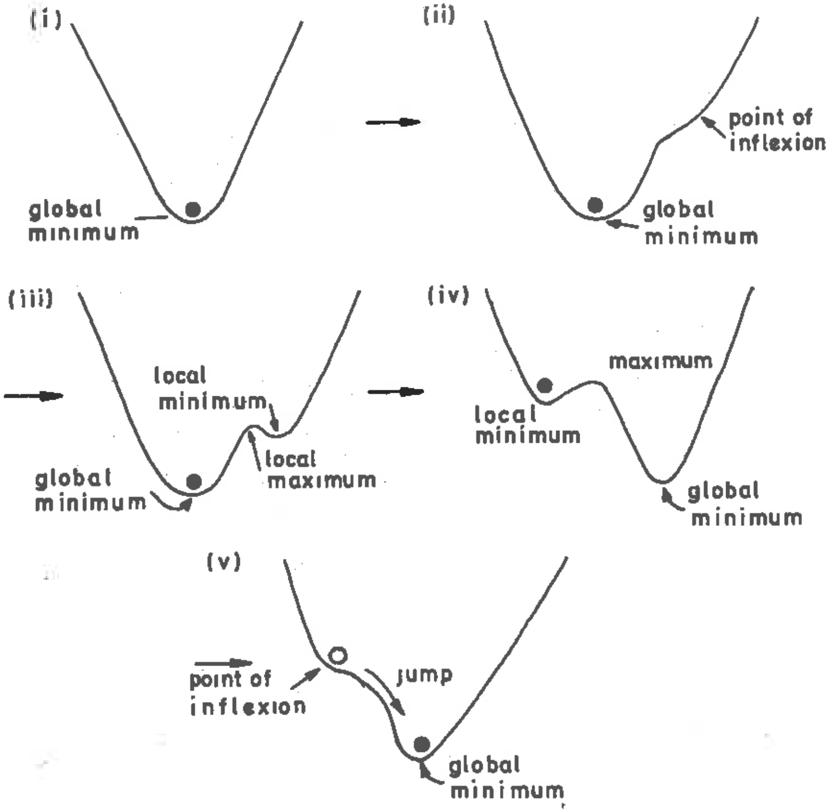
Given these ideas, one can easily understand how a system can suddenly jump from one equilibrium state to another when smooth changes in a control parameter occur. For example, as a control parameter changes a function's pattern may alter in the sequence pictured in Figure 8, where we can imagine the stable position of a rolling ball to be as illustrated (Stewart, 1975).

This diagrammatic representation of a rolling ball offers the possibility of gaining an intuitive feel for "jump" behaviour. It should be noted that in this specific situation, the evolution is described by the '*delay convention*'. By this a system remains in equilibrium relating to a particular minimum for as long as possible, and only alters when the initial one vanishes (as in Zeeman's catastrophe machine). An alternative, widely applied by Thom in his book, is the '*Maxwell convention*'. In this a system is visualised as selecting an equilibrium which relates to the global optimum. Applied to Figure 8, this convention would result in a jump between diagrams (iii) and (iv). Although it is possible to demonstrate that, "Either convention could approximate reality in different circumstances" (Poston and Wilson, 1977, p.684), Isnard and Zeeman (1976, pp.51-54) justify the use of the delay rule in the social sciences because of the lack of information, intuition,



f(x)
B
A
D
C
x

Fig.7.



(i)
global
minimum
(ii)
point of
inflexion
global
minimum
(iii)
local
minimum
local
maximum
global
minimum
(iv)
maximum
local
minimum
global
minimum
(v)
point of
inflexion
jump
global
minimum

Fig.8.

sociological pressures, inertia and past history. Basically any application of catastrophe theory necessitates deciding between the conventions; it is important to appreciate that they result in slightly dissimilar models. In relation to planning, optimal systems are related to global optima and therefore are associated with the so-called Maxwell convention.

3. THOM'S THEOREM: THE SEVEN ELEMENTARY CATASTROPHES.

At this stage in the discussion it might appear as if we are faced with the daunting task of considering an infinite number of possible forms which the curve through the points of degenerate singularity might take. The power of Thom's theorem is that it allows the classification of their (topological) forms, because it shows that the forms are dependent only on the number of control parameters ($\underline{\alpha}$) in any system of interest and not on the number of state variables ($\underline{x}$). Thom demonstrated that given up to four control variables, α_1, α_2, α_3 and α_4, only seven (topologically) unique forms can occur: the seven so-called 'elementary catastrophes' (see Table 2 and Figure 9).

More recently, analysis has shown that for five ocntrol parameters there are eleven fundamental catastrophe forms, and that for more than five control parameters there is an infinite number. (Zeeman and Trotman, 1976).

Two additional remarks, relating to gradient systems and *canonical transformations*, are necessary. Firstly, elementary catastrophe theory only pertains to a *gradient* (dynamical) *system*.† As the name suggests, such a system involves a definition of the dynamics of its state variables by the gradient of its function. Formally, this is described as

---

† Although see Section 7 for extension to non-gradient systems.

$$\frac{dx_1}{dt} = - \frac{\partial f}{\partial x_1} (x_1, x_2, \ldots, x_n, \alpha_1, \alpha_2, \ldots, \alpha_m)$$

$$\frac{dx_2}{dt} = - \frac{\partial f}{\partial x_2} (x_1, x_2, \ldots, x_n, \alpha_1, \alpha_2, \ldots, \alpha_m)$$

$$\vdots \qquad (3.1)$$

$$\frac{dx_n}{dt} = - \frac{\partial f}{\partial x_n} (x_1, x_2, \ldots, x_n, \alpha_1, \alpha_2, \ldots, \alpha_m).$$

Whilst this restriction may seem to undermine the potential applications of elementary catastrophe theory, it should be remembered that theoretical science has widely employed such evolutionary trajectories, and, that it can be recognised at least as an initial approximation to the real dynamics (see Beaumont (1979) for a more detailed discussion of this point).

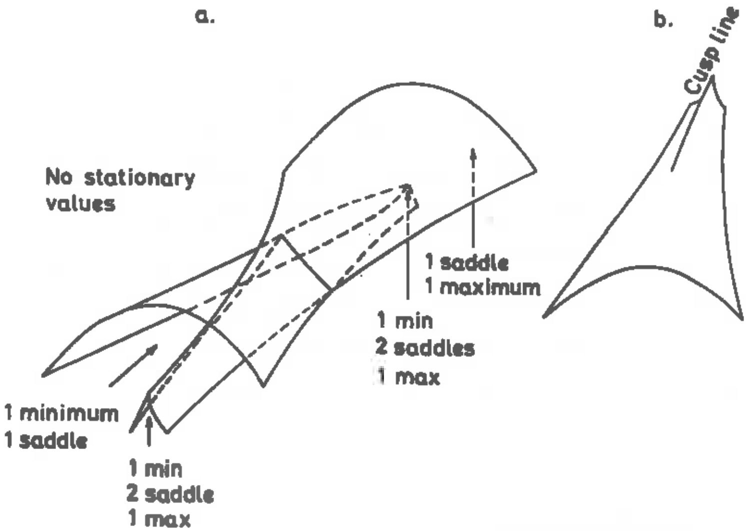
Secondly, whilst Table 2 presents the equations of the seven elementary catastrophes, it would be surprising if a particular equation fitted a specific application exactly. (This point should be borne in mind with regard to Isard and Liossatos's (1977) model of transition processes which is described below). One of the significant facets of Thom's theorem, in fact, is that for a given number of (up to four) control parameters, the form of the sinfularities can be transformed locally to one of the standard (of '*canonical*') forms of the elementary catastrophes. It should be stressed, however, that such transformations present formidable tasks.

CATASTROPHE		CONTROL DIMENSIONS	BEHAVIOUR DIMENSIONS	(STABLE UNFOLDING) FUNCTION	FIRST DERIVATIVE
CUSPOIDS	FOLD	1	1	$1/3x^3+ax$	x^2+a
	CUSP	2	1	$1/4x^4+ax+\frac{1}{2}bx^2$	x^3+a+bx
	SWALLOW TAIL	3	1	$1/5x^5+ax+\frac{1}{2}bx^2$ $+ 1/3cx^3$	$x^4+a+bx+cx^2$
	BUTTER FLY	4	1	$1/6x^6+ax+\frac{1}{2}bx^2$ $+1/5x^3+1/4dx^4$	$x^5+a+bx+cx^2+dx^3$
UMBILICS	HYPER-BOLIC	3	2	$x^3+y^3+ax+by+cxy$	$3x^2+a+cy$ $3y^2+b+cx$
	ELLIPTIC	3	2	$x^3-xy^2+ax+by$ $+c(x^2+y^2)$	$3x^2-3y^2+a+2cx$ $-6xy+b+2cy$
	PARA-BOLIC	4	2	$x^2y+y^4+ax+by$ $+cx^2+dy^2$	$2xy+a+2cx$ $x^2+4y^3+b+2dy$

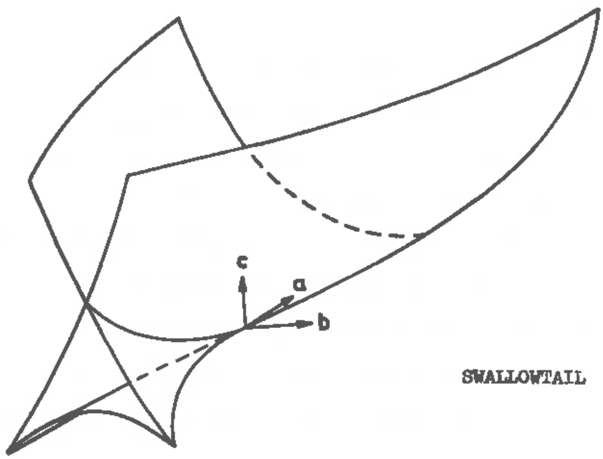
State variables: x and y

Control parameters: a, b, c and d.

<u>TABLE 2</u>. Equations of the seven elementary catastrophes.



HYPERBOLIC UMBILIC



SWALLOWTAIL

Fig.9.

Before presenting a few examples of the application of catastrophe theory, a detailed examination of the fold and cusp catastrophes will indicate how some of the preceeding concepts link together. This will also allow an introduction of some additional terminology. The most simple catastrophe, the fold, has one state variable (x), and one control variable (α). Its function, given in Table 2, is similar to that already discussed in equation (2.8). We know that the first derivative of this function is equal to

$$\frac{\partial f(x, \alpha)}{\partial x} = x^2 + \alpha \qquad (= 0) \tag{3.2}$$

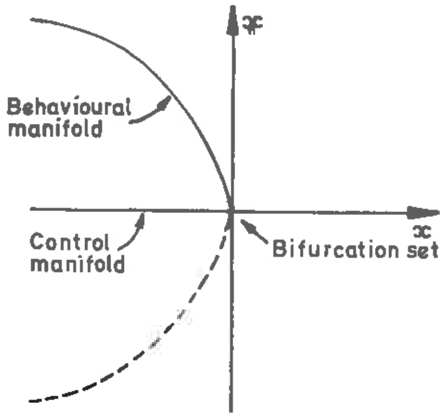
which means that the equilibrium states are given by

$$x = \pm \sqrt{-\alpha}. \tag{3.3}$$

Once again, we can recognise that only real roots exist when $\underline{\alpha}$ is less than zero. If we differentiate to ascertain which are the minimum and maximum points, we produce

$$\frac{\partial^2 f(x, \alpha)}{\partial x^2} = 2x. \tag{3.4}$$

Bearing in mind Table 1, minimum and maximum points occur when x is greater than zero and less than zero, respectively; interest focuses on the stable minimum points. The features are portrayed in Figure 10 below. The parabola described by the equation $x^2 + \alpha = 0$ is drawn to display its stable (complete line) and unstable (dashed line) portions. In the terms of catastrophe theory - the *behaviour*, *control manifolds*, and the *bifurcation set* are indicated. The behaviour manifold $(\underline{x}, \underline{\alpha})$ relates the system's state variables to specific values of the control parameters; the control manifold $(\underline{\alpha})$ is the projection of the behaviour



x
Behavioural manifold
Control manifold
x
Bifurcation set

Fig.10.

manifold onto the parameter space (which in this case is the horizontal line related to the control parameter, α). We have already mentioned that the bifurcation set is the locus of degenerate sinfularities on the control manifold. In the fold catastrophe the bifurcation set is only the origin, because that is where

$$\frac{\partial f(x,\ \alpha)}{\partial x} = 0 \qquad \text{and} \qquad \frac{\partial^2 f(x,\ \alpha)}{\partial x^2} = 0. \tag{3.5}$$

Unlike the other forms of elementary catastrophe, multiple minimum points for particular values of the control parameters do not exist, and, therefore, the so-called '*catastrophe set*', which is the projection of these states onto the control manifold, is not found in the fold catastrophe. (Further exemplification of these aspects is given in the following discussion of the cusp catastrophe and in the specific examples used to illustrate the application of catastrophe theory in urban planning).

Continuing the discussion of specific catastrophe forms, we will now examine the *cusp catastrophe*, which is the most commonly applied archetype. It has one state variable (x) and two control parameters (α and β). Formally, the function is represented as

$$f(x,\alpha,\beta) = \frac{1}{4}x^4 + \alpha x + \tfrac{1}{2}\beta x^2 \tag{3.6}$$

which has the first order derivative

$$\frac{\partial f(x,\alpha,\beta)}{\partial x} = x^3 + \alpha + \beta x. \tag{3.7}$$

If this is equated to zero, we have the *behavioural surface*, which is a function of the control parameters, α and β. Now we know that interest centres on the set of degenerate singularities (when the Hessian matrix's determinant is zero), and this area of the behavioural surface is known as the *singularity set*. Projection of this set of critical points on to the *control surface* gives an image termed the *bifurcation set*. This locus of points is associated with structural instability; the number and form of the critical points alter here. These features can be analysed algebraically by a consideration of the real roots of the above cubic equation.

$$x^3 + \alpha x + \beta = 0. \tag{3.8}$$

It is necessary to make use of a standard formula, which is described here for completeness (see Slazer et.al. (1958) for more details). Given a cubic equation of the general form

$$ax^3 + bx + c = 0 \tag{3.9}$$

its roots may be represented in the form, $(-c/b)f_1(\gamma)$; $(-c/b)f_2(\gamma)$; and $(-c/b)f_3(\gamma)$, (where $\gamma = ac^2/b^3$):

$$f_1(\gamma) = \frac{[1 + \sqrt{1 + (4/27\gamma)}]^{1/3} + [1 - \sqrt{1 + (4/27\gamma)}]^{1/3}}{(2\gamma)^{1/3}} \tag{3.10}$$

$$f_2(\gamma) = \frac{-f_1(\gamma)}{2} + \sqrt{\frac{f_1(\gamma)^2}{4} - \frac{1}{\gamma f_1(\gamma)}} \tag{3.11}$$

$$f_3(\gamma) = \frac{-f_1(\gamma)}{2} - \sqrt{\frac{f_1(\gamma)^2}{4} - \frac{1}{\gamma f_1(\gamma)}} \, . \tag{3.12}$$

Interest is in the situation where there are three real roots (that is, multiple solutions); for real roots to exist,

$$1 + (4/27\gamma) > 0 \tag{3.13}$$

(recall we need to take the square root of a *positive* real number) which implies that

$$27\gamma + 4 > 0. \tag{3.14}$$

Substituting $\gamma = ac^2/b^3$ into this equation gives, after simplification,

$$27\ ac^2 + 4\ b^3 > 0. \tag{3.15}$$

This can now be rewritten in terms of the original cusp equation (where it is noted that the term x^3 has no associated control parameter), that is,

$$27\ \beta^2 + 4\ \alpha^3 > 0. \tag{3.16}$$

For the roots, $f_2(\gamma)$ and $f_3(\gamma)$, to be real, the second term must also be positive. It can be shown that this arises when

$$\gamma < 0 \tag{3.17}$$

or, in terms of the original cusp equation, when

$$\beta^2 < \alpha^3 \tag{3.18}$$

which implies that

$$\alpha < 0. \tag{3.19}$$

Given these conditions, it is possible to demonstrate that the function's three stationary points actually consist of two local minima and one local maximum (see the graphical illustration - Figure 11). By taking the equality form of equation (3.15), cusp shaped curves, representing the *bifurcation set*, are formed on the *control manifold* (the plane of the control parameters, α and β). The inequality depicts the whole *catastrophe set*, where multiple states exist. It should be noted that outside the catastrophe set, where

$$\frac{\beta^2}{\alpha^3} < \frac{-4}{27} \quad \text{and} \quad \frac{\beta^2}{\alpha^3} > 0 \tag{3.20}$$

there is only one real solution exists (and a pair of complex conjugates). When

$$\frac{\beta^2}{\alpha^3} = \frac{-4}{27} \quad \text{and} \quad \frac{\beta^2}{\alpha^3} = 0 \tag{3.21}$$

there are only two roots; a point of inflexion and a local minimum (this represents the *bifurcation set*). Figure 11 illustrates many of these features of the cusp catastrophe. The pliated behavioural manifold of the singularities is shown along with the projection of its degenerate singularities onto the control manifold to give the cusp curve.

To illuminate, let us examine the family of functions associated with a straight trajectory on the control manifold which goes through the catastrophe set (as described in Figure 12). In this example, for simplicity, the control parameter α has a constant value, and, therefore, the analysis considers the effect of varying the other control parameter, β, smoothly, on the system. (The seven points were selected to illustrate the function's various forms). Wilson's (1976) study of modal choice includes a similar investigation of the form of the family of functions, although he incorporates variations in both control parameters.

In situation A, one local minimum exists, and this is to be expected since the position is outside the catastrophe set. In situation B, an additional stationary point is formed, a point of inflexion, which we know to be a degenerate singularity. The point

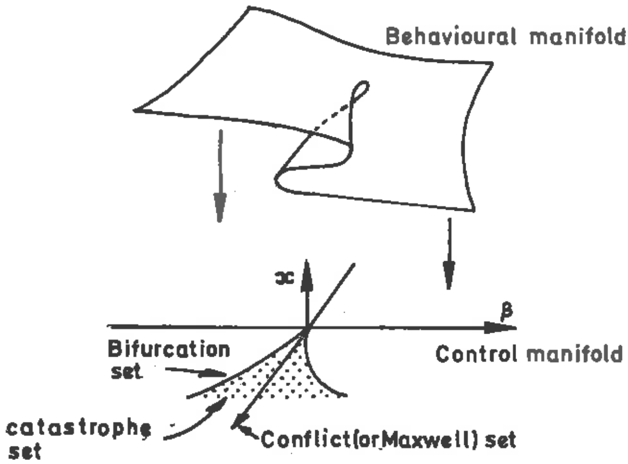

Behavioural manifold
x
β
Bifurcation set
Control manifold
catastrophe set
Conflict(or Maxwell) set

Fig.11.

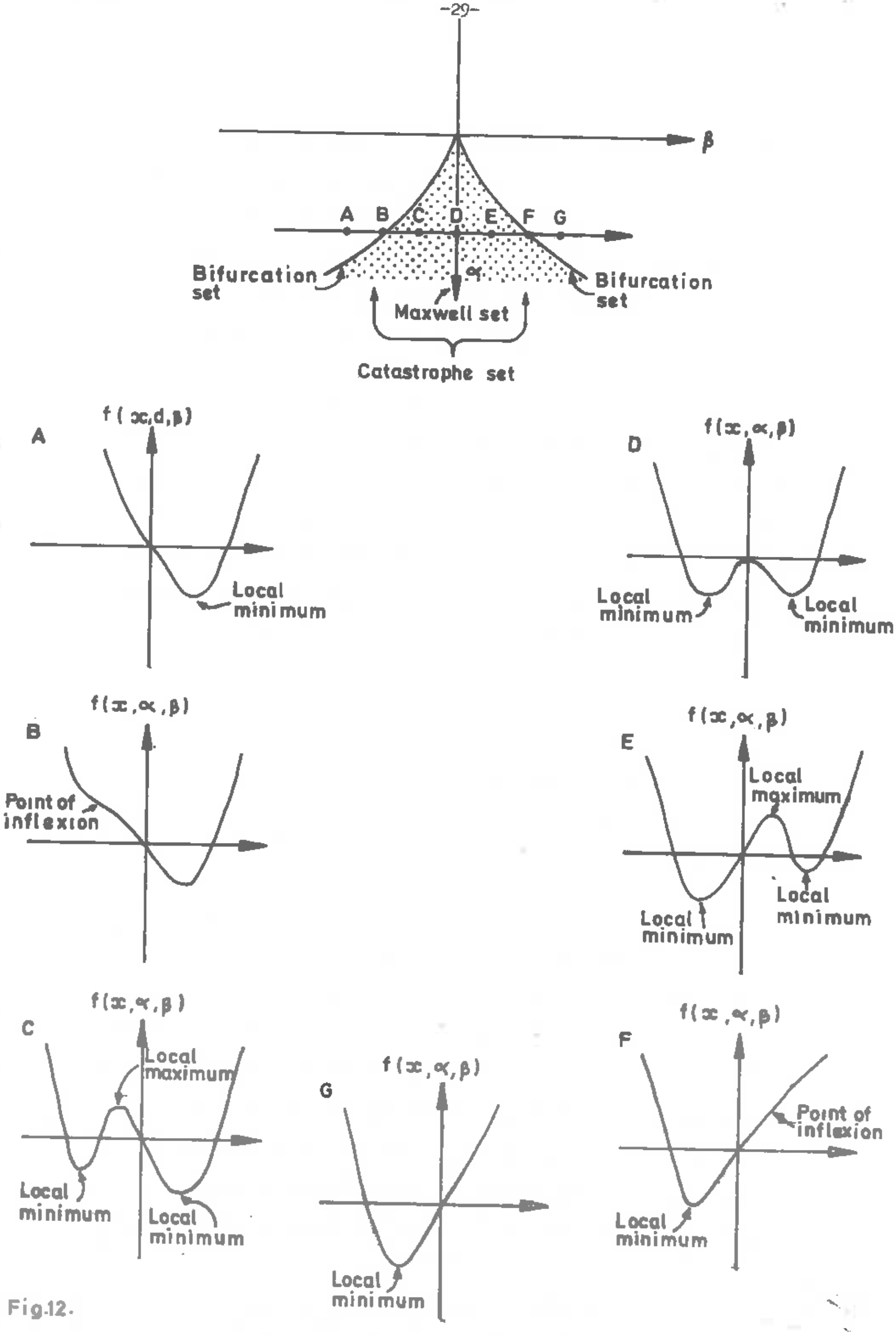

β
A B C D E F G
α
Bifurcation set
Bifurcation set
Maxwell set
Catastrophe set
A
f(x,d,β)
Local minimum
D
f(x,α,β)
Local minimum
Local minimum
B
f(x,α,β)
Point of inflexion
E
f(x,α,β)
Local maximum
Local minimum
Local minimum
C
f(x,α,β)
Local maximum
Local minimum
Local minimum
G
f(x,α,β)
Local minimum
F
f(x,α,β)
Point of inflexion
Local minimum

Fig.12.

of inflexion vanishes by situation C, where there are two local minima and one local maximum. The local minima have equal values in situation D. By situation E, the new local minimum is the one with the lowest value. The first local minimum disappears, at situation F where a point of inflexion exists; this represents the bifurcation set. At situation F, outside the catastrophe, one local minimum is found.

It is appropriate here to reconsider the *delay* and *Maxwell conventions*. If the delay convention had operated, a jump would have taken place on leaving the catastrophe set, at situation F. In contrast, if the Maxwell convention, seeking the absolute minimum, had been applied, a jump would have occurred on crossing the so-called *Maxwell set* or *conflict set* (which is the set of control parameters where the function has equal minimum values). In the cusp catastrophe, the Maxwell set of points, (β, α), where β is equal to zero and α is less than zero.

The jump behaviour and other types of behaviour associated with the cusp catastrophe can also be described on its behavioural manifold, as below (Figure 13).

The diagram of the cusp catastrophe gives the reader an awareness of the types of behaviour that can be described. The idea of *jumps* from one stable equilibrium state to another has already been examined. Associated with this discontinuous behavoiur is the so-called *hysteresis effect*. This occurs when the return to a previous state does not follow the original path away from it. As Figure 13 illustrates, an hysteresis effect can only occur when the *delay convention* is operative; that is, the jump behaviour takes place when leaving the *catastrophe set*. Interestingly, an empirical investigation in the field of transport has

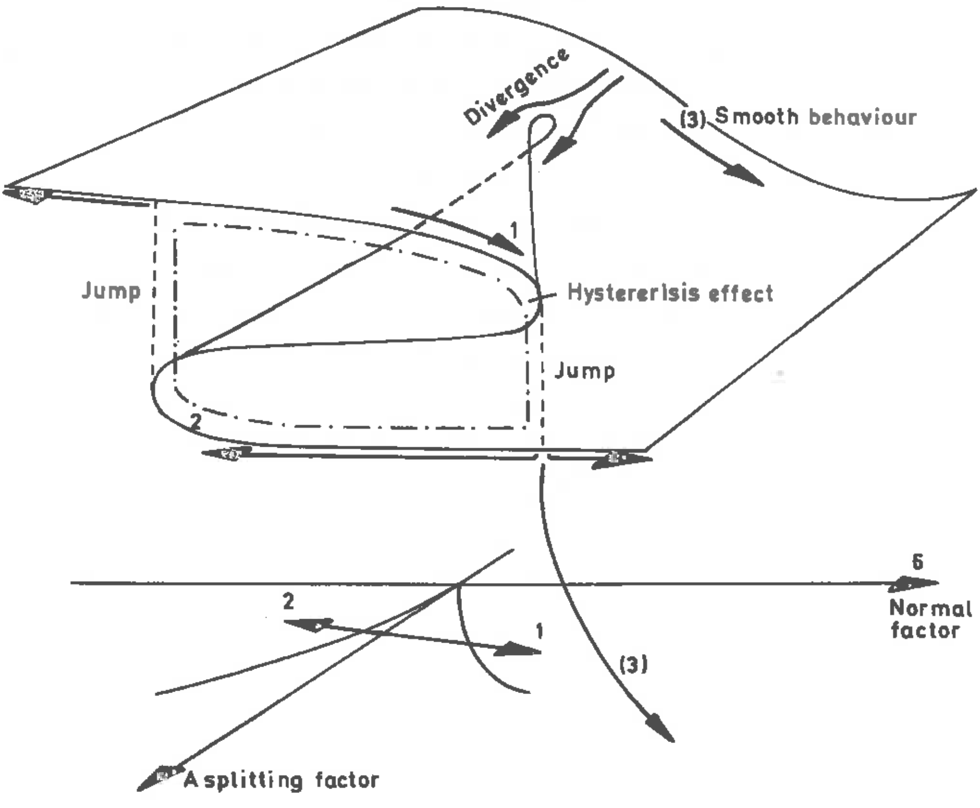

Divergence
(3) Smooth behaviour
1
Jump
Hystererisis effect
Jump
2
δ
Normal factor
2
1
(3)
A splitting factor

Fig.13.

demonstrated an hysteresis effect (Blase, 1978). *Divergent behaviour* may also be of interest to geographers. It describes the situation in which the end states of a system may be very different, though the initial states showed little variation. This type of threshold effect, separating different types of behaviour, is widely found in urban systems (see Bennett and Chorley, 1978). Finally, as displayed by the diagram of the cusp catastrophe, it is quite incorrect to think that catastrophe theory can only accommodate discontinuous behaviour - *smooth behaviour* can also be found. In the case of the cusp catastrophe, it is already known that the control parameter, α, must be less than zero for a degenerate singularity to be present; indeed, the value of α can be seen as determining whether portions of the behaviour manifold are smooth or folded. For this reason, α is commonly termed a *splitting factor*, and β is called a *normal factor*.

4. PLANNING SYSTEMS AND THE USE OF CATASTROPHE THEORY.

4.1 Continuity in Change.

From the previous discussion of elementary catastrophe theory, it can be seen that system evolution is viewed as a succession of phases of re-organisation (morphogenesis) and stabilisation (morphostasis); Lampard (1968), for instance, applied these concepts in his analysis of 'The evolving system of cities in the United States'. The basic postulate is that it is plausible to see history as a series of fairly separate eras with individual characteristics rather than as a continuous progression of incremental alterations. Whilst it is, obviously, possible to present reasons to corroborate either perspective (and, to a certain degree, there is authenticity in both views), in a variety of situations, there is a tendency for system responses to be non-incremental in nature - system organisation is restructured.

For example, in arguing that a long-term view of economic development and cultural change is necessary to adequately analyse the growth of the city system in Wales, Carter (1980) states that

> "... there are distinct phases of urban genesis. They are implicit in all texts on the history of town plan but are never translated into urban system terms. There is not a continuous process of urban creation but a number of distinct periods related to political or economic convulsion, or both".

A number of significant consequences for planning result from this conceptualisation of system evolution. There is an implicit suggestion that planning cannot indefinitely constrain development processes within a given structure; the role of planning can be thought of as one of assisting evolutionary processes. Transitions to new regimes will eventually occur, although planners could act as catalysts in the evolution of different behavioural patterns.

4.2 The Role of Planners.

There are a variety of different planning philosophies that have been articulated in recent years. These range from the rather mechanistic concept of planning as control (for example McLoughlin, 1973) to planning as a function of the state involved with the amelioration of class conflict (for example, Karnavou, 1979). Here, to paraphrase Scott and Roweis (1978), we are more concerned with what planning is and does. To this extent we will briefly discuss some of the actions that planners take and look at the ramifications of these actions. These will then be cast in such a way that allows us to contemplate the possible use of catastrophe theory as an aid to planners.

The scope of urban planning and the powers afforded to planners through successive legislative acts has progressively increased since the turn of the century. Some would argue that not enough progress has been made, others that the power of planners is too great. We would argue that in some cases while planners do not have a great deal of direct influence the consequences of planning decisions are often considerable. Let us illustrate this with reference to the land market.

One of the traditional roles of planners is in the zoning of land uses according to some strategic plan. The onus is then on developers to apply for planning permission to construct property on a certain tract of land. If planning permission is given by the planning department then it is clear that the market value of this tract of land will almost certainly rise. The effect may go beyond this and adjoining parcels of land, whether developed or not may experience changes in market value. Similarly if the planners were to release a large amount of land onto the "market", in this fashion, then, assuming that the simple law of supply and demand operates, the market may be saturated and thus land prices fall. Planners therefore do not determine land values but to a certain extent are in the position to regulate or help determine them.

In the housing system particularly in the United Kingdom, planners play an important role. Directly the public sector provides over 5 million dwelling units in England and Wales alone, some 29% of the total stock in 1975.(Housing Policy Review, 1977). The allocation rules that Local Authorities operate and the typical excess demand for council housing results in long waiting lists in most cities. Consequently the private sector benefits (landlords in particular). Thus again the local authority has both direct and indirect roles. This argument could be extended for many different sectors. Table 3 summarises several other effects.

PLANNING ACTION	POTENTIAL EFFECTS
Construction of new trunk roads.	Accessibility pattern changes land values and house prices may both increase and decrease in different areas. Transport costs decrease.
Planning permission for out of town shopping centre.	Land values change (see text). Revenue patterns of existing retail facilities will change*. Trip generation pattern changes. Shops may close.
Provision of nursery school facilities.	Allows parent of child to seek employment.
Subsidy of public transport system.	Again, change in land values, accessibility, and relation between residence and workplace.

TABLE 3: Some more examples of the consequences of planning action.

It may be clear already that there is an obvious link between these ideas and one of the basic notions of catastrophe theory, namely the definition of control variables and state variables, as outlined in Section 2. (Note that the term "control" is as used in catastrophe theory and should not be immediately associated with control in the cybernetic or conventional meaning).

There exists, as we have demonstrated, a number of variables, or parameters that come under the control of planners. Changes in these variables may trigger changes in other variables. What often proves a difficult task is to model the relationship between these two types of variables. For example, with feedback loops in a system, there is an element of ambivalence in the interpretation of dependent and independent variables, and, in so doing, creating complications in parameter estimation. Much progress has been made on this front in recent years, and a specific example, concerning urban retailing facilities is presented in the next section. Provided the system of equations can be written in the form of equation (2.1) and the number of control variables limited to four then the potential for catastrophe theory type analysis exists.

If this is the case what could the planner benefit from the approach. Arguably planners are interested in avoiding certain situations and acting as catalysts for change. If by using tools from catastrophe theory planners could expect their behaviour to produce sudden, say, undesirable, change then there exists the possibility of avoiding it, and vice versa.

Of direct relevance to planners may be the sensitivity of the solution of optimisation models (particularly non-linear ones) to values of exogenous parameters. If the solution is close to criticality (as defined above) then a rather different solution may be obtained if there are small variations in parameters, such as the β parameter in transport models (see Wilson and Clarke, 1979).

A number of present-day features of the planning process will remain of paramount importance. For example, planning must be continuous and not exhibit discontinuities associated with the system's behaviour. In addition, results attained from monitoring and forecasting procedures will continue to provide essential information to planners (Bennett, 1979). The recognition of sudden transitions between states raises a fundamental problem with respect to current forecasting methods (Bennett, 1978; Haggett, 1973). Applications of (spatial-temporal) forecasts usually involve trend extrapolations (as exemplified by Chisholm et.al. (1971) and by Cliff et.al. (1975)), and, in so doing, they do not necessarily take account of the possibility of discontinuous behaviour (Hay, 1978). Research embodying concepts of catastrophe theory within a forecasting methodology is, therefore, required. The potentiality of alternative futures presents a philosophical and methodological challenge for planners, which must be addressed in order to consider alternative solutions to contemporary societal problems (Beaumont, 1980).

Realistically, planning applications typically lag somewhat behind theoretical innovation, and this is certain to be the case with the ideas discussed here.† The development of appropriate sub-models of

---

† Especially bearing in mind the criticisms levelled at the application of catastrophe theory in the social sciences, see Section 6.

the urban system has a chequered history. Many models are particularly deficient at handling policy issues, ie. in our terminology the incorporation of the relationship between control and state variables. While we can expect this to improve in the future, catastrophe theory may play a useful role in helping to develop an awareness of different types of system behaviour, such as hysteresis, jumps, divergence, etc., and how this behaviour may be influenced by the actions of the planner.

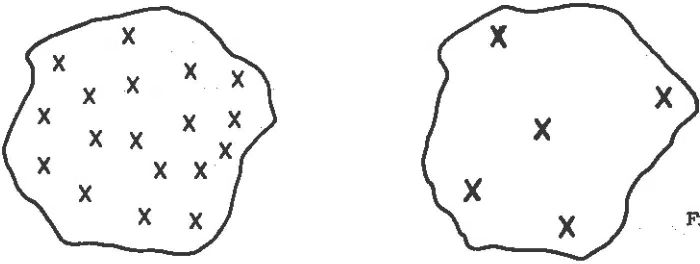
5. APPLICATIONS IN URBAN PLANNING.

The three examples considered in this section illustrate different aspects of the previous discussion. Poston and Wilson's (1977) example describes the fold catastrophe diagrammatically. Isard and Liossatos attempt to directly interpret the cusp catastrophe function in terms of social welfare. The final example is the numberical illustration of the fold catastrophe applied to urban retailing structures employing the well-known shopping model.

The first example to be described is Poston and Wilson's (1977) study of the evolution of the optimum size of shopping centres. The analysis demonstrates how the *fold catastrophe* can be derived using a relatively simple graphical approach.

Benefits are assumed to accrue to the individual that are dependent on the size of shopping centres. It is argued that a larger shopping centre offers the consumer a wider range of choice than a smaller centre, and that because of scale economies prices may be lower. However the larger the size of shopping centres, assuming constant demand, the less number of centres there can be. If we assume a relatively even distribution of population throughout the city it follows that, if there are fewer centres then the average distance that consumers travel is greater. (c.f. Figure 14). Moreover, the cost of travel is considered as a disutility, which has to be traded off against the benefits of shopping centre size.

The form that these disutility and benefit functions are assumed to take is given in Figure 15.



Many small centres - average distance to centre is low.

Few large centres - average distance to centre is high.

Fig.14.

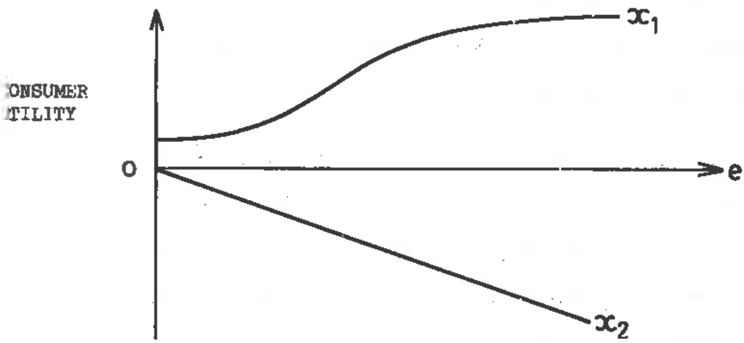


Fig.15

The benefits of increased service centre size (x_1) are assumed to increase logistically with size (in fact average distance travelled, l, as this will be directly proportional to centre size). After a certain size the rate of increase in benefits accruing to the consumer will begin to fall off until a certain size is reached beyond which benefit remains constant. Travel costs (x_2) (or disutilities) are assumed to vary linearly with average distance travelled. Total benefit (x), which is of specific interest here, is found by summing the two components of utility (ie. $x = x_1 + x_2$).

Analysis of the system focuses on a travel impediment, 'a', the marginal rate of disbenefit with distance travelled by a consumer to a service facility. 'a' thus determines the slope of the line x_2 and is used as a control parameter to study how smooth changes in its value will affect optimum facility size, the state variable. When 'a' is a large value, it implies that travel is relatively difficult, and the slope of x_2 is thus steep. When 'a' is a small value, and the slope of x_2 shallow, it implies that travel is relatively easy. To illustrate how the fold catastrophe is significant, the evolution of facility size is considered as the ease of travel progressively increases.

Optimum facility size occurs when consumers' utility is maximised (when x is a maximum). Refering back to the earlier discussion on differentiation, this takes place when the first order derivative equals zero ($\frac{\partial x}{\partial l} = 0$). Figure 16 pictures how a service system may evolve as travel impedance increases. Following Poston and Wilson's (1977, p.682) description,

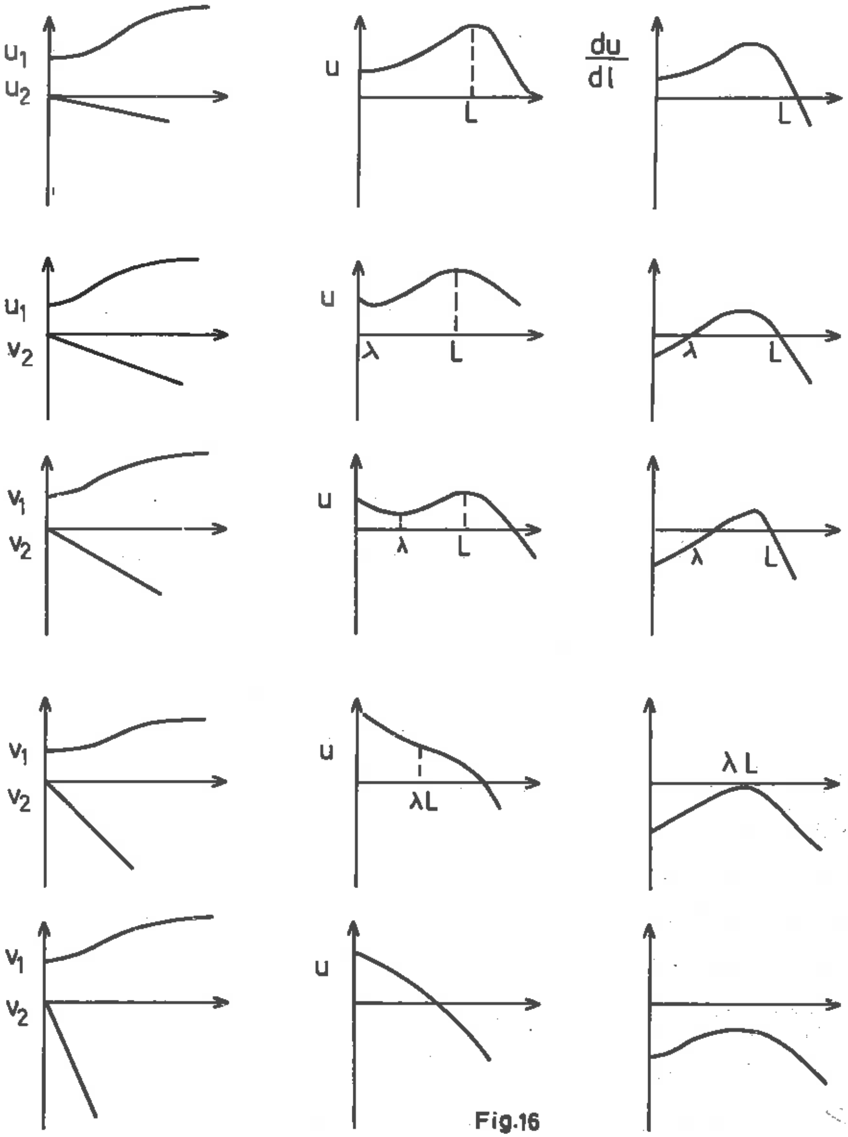


Fig.16

"Cases (1) to (5) show progressively increasing values of 'a'. Case (1) (Illustrates that) $l = L$ is a unique maximum for z. Case (2) shows the gradient of x_2 ('a') becoming steeper than of x, at $l = 0$. $l = L$ is still the global maximum but now $l = 0$ appears as a local maximum with $l = \lambda$ a local minimum. In case (3), $l = 0$ has become the global maximum and $l = L$ the local one. In case (4) the points λ and L coincide at a point of inflexion. In case (5) the local maximum at $l = L$ has disappeared and $l = 0$ remains the global maximum".

We already know that a point of inflexion is a degenerate singularity, and by plotting the optimum values of the distance to travel, l, (used as a surrogate for facility size) as a function of the control parameter 'a', it is possible to depict the system as a fold catastrophe (see Figure 17). The continuous lines are potential stable maxima; for $a_0 < a < a_1$, there are two possible optima. The specific interest here is the discontinuous changes in the system dynamics (represented by the dotted and the dotted/dashed curves in Figure 17).

"A move along the dotted curve implies increasing difficulty of travel (that is, increasing 'a') so that when $a = a_1$ is reached, the system must *jump* from the 'supermarket' maximum to a 'corner shop' maximum. The dotted/dashed curve shows the opposite phenomenon: by decreasing a, at $a = a_0$ a jump from a 'corner shop' state to a 'supermarket' state occurs". (Poston and Wilson, 1977, p.684).

In addition to the recognition of jump behaviour, an *hysteresis* effect can also be seen. That is, the paths jumping from 'supermarkets' to 'corner shops' and from 'corner shops' to 'supermarkets' are not the same. (Note the empirical investigation of this feature by Blase (1978) in connection with the choice of mode of travel and energy prices). It should be remembered that to have this hysteresis effect, the so-called *delay convention* must hold. In addition to an awareness of discontinuous behaviour, catastrophe theory, therefpre, indicates that a return to a previously desireable state may entail more than simply restoring a system to the conditions prevalent prior to the alteration. This feature obviously has significance for planners.

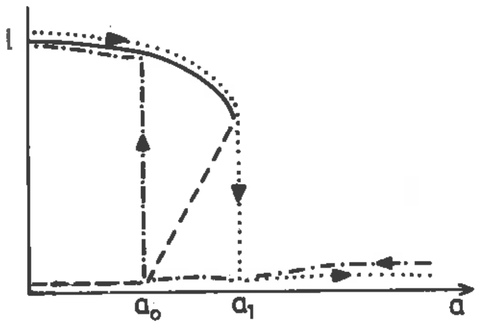


Fig. 17

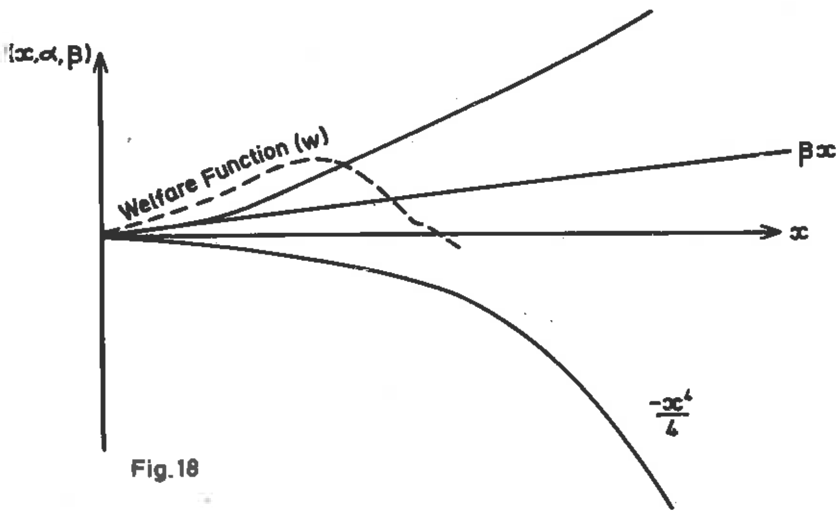


Fig. 18

1/84

Isard and Liossatos (1977) address the question of abrupt and sharply discontinuous development in a social system. The fundamental notions of equilibrium positions and structural stability are explicitly examined, and their linking of these concepts to an optimisation problem is of particular interest. An explicit attempt is made to describe and interpret a system's development by taking account of internal and external interactions. Remembering the earlier comments on the difficulties of performing canonical transformations, it is significant that Isard and Liossatos deem a 'meaningful' model to have the following welfare function (which is to be optimised),

$$W(x,\alpha,\beta) = \frac{-x^4}{4} + \frac{\alpha x^2}{2} + \beta x + C. \qquad (5.1)$$

It should be realised that the maximisation of this function is congruent with Thom's classification of the associated stable unfolding, the cusp catastrophe. Whilst this can be thought of as rather metaphoric in design, novel interpretations of social development are forthcoming; this contrasts with some poor applications of catastrophe theory which only intuitively combine a particular elementary catastrophe form with a number of recognised processes. (Such applications are considered later).

The actual rationale behind the above welfare function is explicitly described as follows:

(i) the term, $- x^4/4$, represents the negative externalities (such as, congestion, pollution, higher costs of living, ...) associated iwth the system's population growth,

(ii) the term, $\alpha x^2/2$, describes the positive externalities (agglomeration economics) associated with the system's population growth (which arises from greater specialisation),

(iii) the term, βx, can be thought of as other positive externalities (agglomeration economics) associated with population growth in the system,

and (iv) the term, C, represents the other influences, possibly from the external environment.

These components of the welfare function are illustrated separately in Figure 18. The maximum welfare condition for this given function is when

$$\frac{\partial W}{\partial x} = 0 \qquad (5.2)$$

and, that is, when

$$-x^3 + \alpha x + \beta = 0. \qquad (5.3)$$

Surprisingly, Isard and Liossatos do not explicitly refer to the cusp catastrophe. Furthermore, rather than present the usual cusp catastrophe diagram, analysis is essentially concerned with slices of the cusp for particular values of the control parameter α. This approach is of interest for a number of reasons. Multiple equilibria exist only when α is greater than zero, and only such values are examined. (The reader is reminded that, in this example, the control parameters, α and β, would be called the splitting and normal factors, respectively). This approach reinforces the idea that higher order catastrophes are made up of lower order catastrophes. For example, analysis of the equilibrium of the state variables, x, as a function of the value of the control parameter, β, (for some fixed value of α greater than zero) is shown to be a fold catastrophe in Figure 19. If all values of α were considered, such slices to be seen to together form the cusp catastrophe.

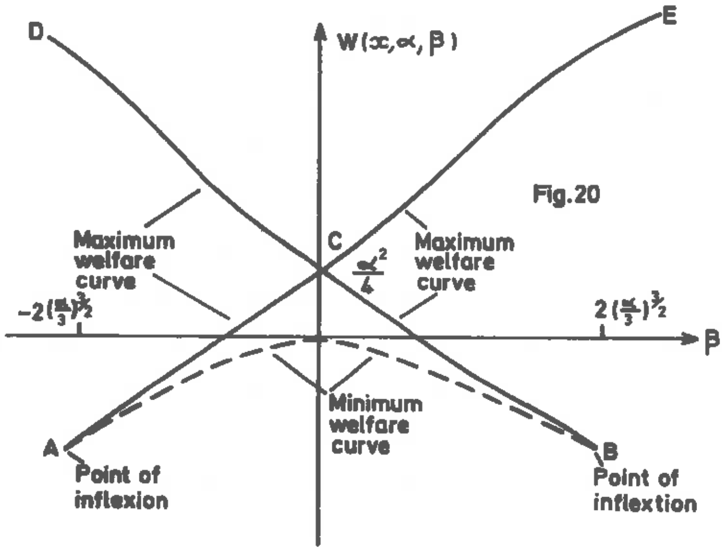

D
E
W(x, α, β)
Maximum
welfare
curve
C
α²/4
Maximum
welfare
curve
−2(α/3)^3/2
2(α/3)^3/2
β
Minimum
welfare
curve
A
Point of
inflexion
B
Point of
inflextion

Fig.20

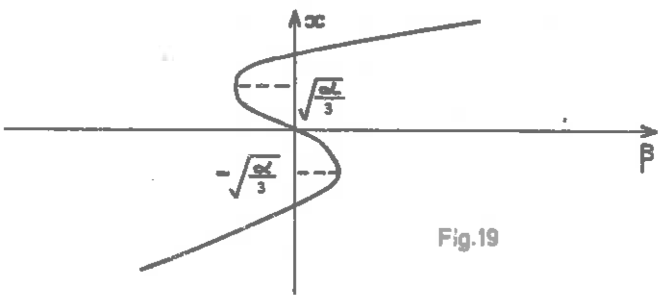

x
√(α/3)
β
−√(α/3)

Fig.19

Further insight can be generated by an analysis of welfare provision as a function of the control parameters α and β. Figure 20 illustrates how the aggregate welfare can vary with the β control parameter (for a constant value of α, which is positive). The curves AE and BD represent the points of maximum welfare for different values of β.

It should be noted that for values of β between $-2\left(\frac{\alpha}{3}\right)^{3/2}$ and $2\left(\frac{\alpha}{3}\right)^{3/2}$, multiple optima exist - two maximum points and one minimum point. This can be related to the previous discussion of the cusp catastrophe by looking at a slice of the above diagram (Figure 20); that is, examining the welfare function and the state variable with the control parameters α and β both fixed. It is also worth noting that Figure 21 demonstrates that one of the local maximum points is greater than the other (except at point C when their values are equal).

Additional analytical insights are derived by extending the study of Figure 20 to include a number of such diagrams illustrating variations in the value of the control parameter α. These are represented in Figure 22, in which it is seen that multiple equilibrium states and the corresponding values of the welfare function disappear when α equals zero.

One constant source of criticism levelled at the application of catastrophe theory in the social sciences is the lack of numerical experimentation to complement the analytical work (see Zahler and Sussmann (1977) and Section 6). In a recent paper (Wilson and Clarke, 1979) an attempt was made, building on an earlier piece of work (Harris and Wilson, 1978) to numerically illustrate how the pattern of retail

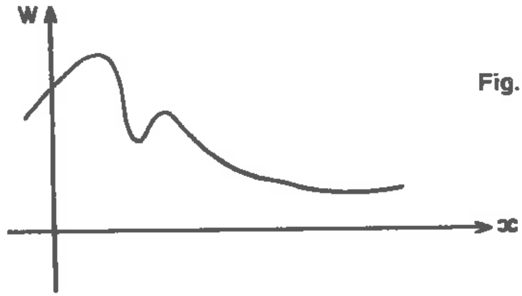


Fig. 21

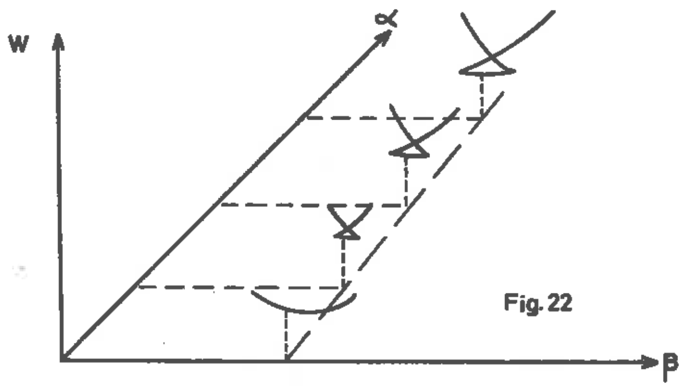


Fig. 22

facilities in a city could suddenly change, in response to a small, smooth change in a control variable.

The application uses a mathematical programming version of the well known shopping model

$$\max_{\{\underline{W},\underline{S}\}} Z = -\frac{1}{\beta} \sum_{ij} S_{ij} + \sum_{ij} S_{ij}\left(\frac{\alpha}{\beta} \log W_j - c_{ij}\right) \tag{5.4}$$

subject to

$$\sum_j S_{ij} = e_i P_i \qquad \forall_i \tag{5.5}$$

$$\sum_j W_j = W \tag{5.6}$$

where S_{ij} represents the flows of expenditure from zone i to zone j,

e_i is the percapita expenditure on retail goods in zone i,

P_i is the population in zone i,

W_j is the amount of retail floorspace in zone j,

c_{ij} is the cost of travel between zones i and j,

W is the total stock of floorspace to be allocated within the system,

α and β are parameter.

The right hand side of equation (5.5) can be interpreted as consumer surplus (c.f. Coelho and Wilson, 1976). The problem is thus the maximisation of consumer surplus subject to expenditure and stock constraints.

The mechanism that gives rise to jump type behaviour is now briefly described.

It can easily be shown that

$$\sum_{ij} S_{ij} = \sum_i e_i P_i = k \sum_j W_j = kW \tag{5.7}$$

so that

$$k = \sum_i e_i P_i / W. \tag{5.8}$$

Let us define D_j as the revenue accruing to retail facilities in zone j. That is,

$$\sum_i S_{ij} = D_j \tag{5.9}$$

thus

$$D_j = kW_j \tag{5.10}$$

and it can be shown also (Harris and Wilson, 1978) that,

$$D_j = \sum_i S_{ij} = \sum_i \frac{e_i P_i W_j^\alpha e^{-\beta c_{ij}}}{\sum_j W_j^\alpha e^{-\beta c_{ij}}} . \tag{5.11}$$

If a plot is made of the straight line (5.12) and the curve (5.11) for D_j against W_j possible zonal equilibrium points (ie. where a solution of (5.4) can occur are at the origin, and where the two lines intersect). This is shown for three different values of α in Figure 23.

The possible equilibrium points are 0, W_j^A and W_j^B. It can be shown that W_j^A is a stable point, whereas W_j^B is unstable (though it does not exist for $\alpha < 1$). The origin, $W_j = 0$, is stable for $\alpha > 1$ and for certain cases when $\alpha = 1$.

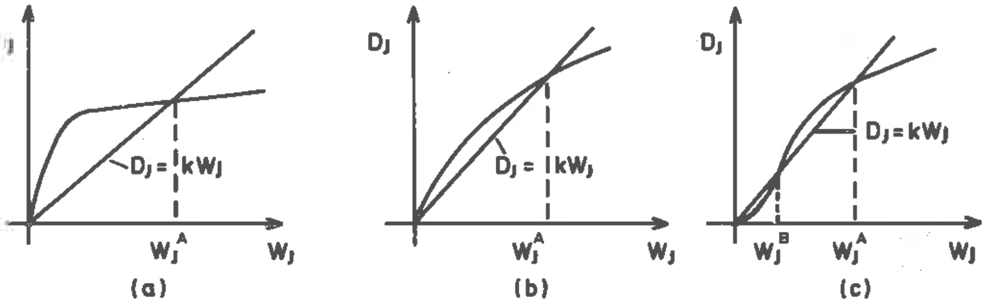
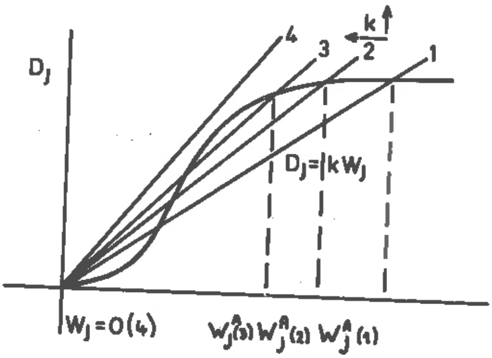


Fig.23

Now one possible mechanism for change is variation in the slope of the $D_j = kW_j$ line, that is a change in the value of k. Figure 24 illustrates how for a smooth change (increase) in k the equilibrium position jumps from the upper stable point W_j^A to the origin, $W_j = 0$.

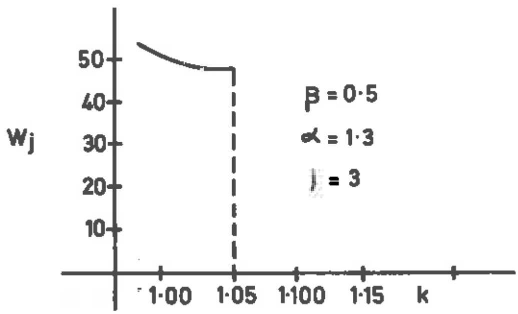
Empirical investigations of this type were undertaken and the difficulties encountered are fully described in Wilson and Clarke (1979). However, jump type behaviour was experienced for the Leeds city system. For example, the jump illustrated in Figure 25 was located in zone 3 as k increased.

The results presented in the paper are of a tentative nature and further analysis is being undertaken (see for example Beaumont and Clarke, 1980).



D_j
4
3
2
1
k
$D_j = kW_j$
$W_j = 0(4)$
$W_j^a(3)$ $W_j^a(2)$ $W_j^a(1)$

Fig. 24



W_j
50
40
30
20
10
1·00
1·05
1·100
1·15
k
$\beta = 0{\cdot}5$
$\alpha = 1{\cdot}3$
$j = 3$

Fig. 25

6. CRITICISMS OF CATASTROPHE THEORY APPLICATIONS.

The application of catastrophe theory to the biological and social sciences has attracted much debate. Most of the critics stress that they are not questioning the correctness of the mathematics or in may cases the validity of applications inthe physical sciences. Many though are sceptical as to its appropriateness as a tool of analysis in the social sciences. The most vigorous criticisms of the claims and accomplishments of applied catastrophe theory have come from Zahler and Sussman (1977, 1978) and their series of papers have provoked much response to their challanges. A fair deal of their criticism is levelled at the biological sciences and lies beyond the scope of this paper.

It is very important to be aware of the specific criticisms, which have much validity, but it must also be remembered that, unfortunately, much of the criticism has been unhelpful and, to a certain degree, over-reactive (which is perhaps an outgrowth of some of the extravagant claims). Whilst particular examples may have numerous failings, attempts to substantiate general criticisms have largely floundered. In fact, whilst the variety of examples, published at different mathematical levels, is advantageous for the social scientist without a detailed mathematical background, difficulties can also arise.

It should be noted that there is a danger of taking analogies too far. The cusp catastrophe, the most "popular" elementary catastrophe, is not appropriate to describe all phenomena exhibiting bimodal behaviour. In general qualitative descriptions disregarding processes and mechanisms of change do not aid comprehension. This does not, of couse, undermine the power of catastrophe theory to classify transitions between equilibrium states (see for example, Harris and Wilson, 1978).

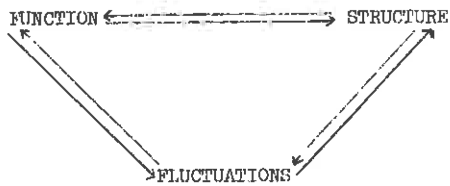
It is sufficient in this introduction to make the reader aware that criticisms of the application of catastrophe theory do exist. It is obviously necessary to critically assess the validity of any application of the technique. Specific problems are not considered, except in the discussion of the examples.

7. DISCUSSION AND FURTHER EXTENSIONS.

It has already been noted that catastrophe theory is limited to the analysis of gradient systems. This suggests that bifurcation theory, which is not restricted to gradient systems, is likely to have wider applicability. Moreover, the system of equations describes behaviour in a purely deterministic fashion, and, in so doing, excludes random elements found in reality. Both these features are of significance in system modelling, and, therefore, are briefly considered below.

Bifurcation theory can be traced back to the work of mathematicians, such as Poincare in the nineteenth century. Since then much development of the qualitative theroy of differential equations and dynamical systems has taken place (see, for example, Andronov et.al., 1973; Hirsch and Smale, 1974; and Jordan and Smith, 1977). For instance, much attention has been recently given to Hopf's bifurcation theorem (see Marsden and McCracken, 1976). Whilst research continues on the analytical front, applications of bifurcation theory have also been forthcoming (see Gurel and Roaller, 1979). Its potential wider applicability suggests it deserves and requires further attention.

Nicolis and Prigogine (1977) have presented a detailed demonstration of how a combination of deterministic and stochastic elements can facilitate a representation of dynamical processes. A general discussion of this conceptualisation has been undertaken by Beaumont (1980), and Isard and Liossatos (1979) have exemplified this approach in their study of spatial dynamics. Basically, the topic of criticality suggests that near an unstable point deviations from the mean value may be crucially significant; at criticality, it is the stochastic 'fluctuations' which actually determine the specific regime a system moves to. Nicolis and Prigogine schematically represent the complementary role of determinism and fluctuations in system evolution as



For example, a system's function is described by a set of differntial equations, it's various spatial-temporal structures arise from the instabilities inherent in the system, and the fluctuations actually trigger the transitions between states at critical points.

It should be noted that although elementary catastrophe theory classifies the types of transitions between stable regines in (gradient) systems, this framework presents a mechanism for change through the fluctuations - 'order through fluctuation'. Alternatively, it would be possible to develop Zeeman's (1977) discussion of the 'slow' equations.

8. CONCLUSIONS AND EVALUATIONS.

The basic features of elementary catastrophe theory have been described and explained in this elementary introduction, and it is hoped that geographers and planners will be encouraged to build on the initial applications in their disciplines; they can, at least, now attempt to make their own evaluation of its potential. In addition, with an understanding of the main themes presented here, the reader should find the available literature on catastrophe theory more comprehensible.

Taxanomical procedures have proved invaluable in research, and an underlying feature of Thom's theorem is the classification of the forms that degenerate singularities can take when there is up to four control parameters. This implies that catastrophe theory has the possibility of offering a coherent synthesis, and, therefore, presents geographers and planners with an opportunity to structure analysis, hypotheses, and observations in a new way.

ACKNOWLEDGEMENTS.

Several people provided useful comments on an earlier draft of this paper and to them we are grateful. Paul Keys in particular assisted us with some of the trickier mathematics. As always, Alan Wilson provided encouragement and useful advice. Of course, the responsibility for what is contained in these pages remains solely that of the authors.

9. GUIDE TO FURTHER READING AND ADDITIONAL REFERENCES.

The literature associated with catastrophe theory has grown rapidly in recent years. We present some selected avenues the reader might like to pursue after reading this introductory monograph.

1. Text-Books:

Thom, R. (Translated by Fowler, D.H.) (1975) "Structural Stability and Morphogenesis". Benjamin, London.

This is Thom's original and masterly exposition on catastrophe theory, and, for this reason alone, it is important. Without doubt, it contains very difficult mathematics and rather confusing terminology, but the indication of some of the philosophical underpinnings is of interest.

Poston, T. and Stewart, I. (1978) "Catastrophe Theory and its Applications". Pitman, London.

An excellent, clear description of catastrophe theory and of numerous examples is presented. The mathematics is rather more detailed than in the present volume, although the background given here should facilitate understanding. A comprehensive bibliography is also included.

Wilson, A.G. (1980) "Catastrophe Theory and Bifurcation with Applications in Urban Geography". Forthcoming.

This forthcoming book is the only one devoted to geographical systems, and numerous applications are included. Catastrophe theory is again introduced at an elementary, although more detailed, level (and should be seen as the natural extension of this work). An effort is made to clearly present the topic's difficult and individual terminology, which is of enormous assistance in reading many articles on the subject.

Woodcock, A. and Davis, M. (1978) "Catastrophe Theory". Dutton, New York.

This book presents an elementary introduction for the layman; technical language being consciously avoided. Examples of applications in a variety of disciplines are considered, although, by the nature of the book, they tend to be of a qualitative, rather than of an analytic, nature. The controversy that catastrophe theory has caused (which is exemplified by some of the papers below) is given a short, but balanced review.

Zeeman, E.C. (1977) "Catastrophe Theory: selected papers, 1972-1977". Addison-Wesley, Reading, Mass.

This is a collection of papers written by the leading proponent for the application of catastrophe theory. Whilst the papers are of varying complexity, the range of topics discussed and the specific forms of the numerous applications make this book worthy of careful attention.

2. ADDITIONAL REFERENCES.

Amson, J.C. (1972) Equilibrium and Catastrophic Models of Urban Growth. In London Papers in Regional Science 4: Space-Time Concepts in Urban and Regional Models. (Ed.) E.L. Cripps. Pion, London. pp.108-128.

Andronov, A.A., Leontovich, E.A., Gordon, I.I. and Maier, A.G. (1973) Qualitative Theory of Second-Order Dynamic Systems. John Wiley & Sons, New York.

Baker, A.R.H. (1977) Historical Geography. Progress in Human Geography, 1. pp.465-474.

Balasko, Y. (1978) Economic Equilibrium and Catastrophe Theory: An Introduction. Econometrica, 46(3), pp.557-569.

Beaumont, J.R. (1979) Some Issues in the Application of Mathematical Programming in Human Geography. Presented at I.B.G. Quantitative Methods Study Group's conference on 'Spatial Applications in Operational Research', University of Salford, September.

Beaumont, J.R. (1980) Towards a Conceptualisation of Evolution in Environmental Systems. Working Paper, School of Geography, University of Leeds.

Beaumont, J.R. and Clarke, M. (1980) The Determination of Critical Points in Urban Retailing Models. Working Paper, School of Geography, University of Leeds.

Bennett, R.J. (1978) Spatial Time Series: Analysis, Forecasting and Control. Pion, London.

Bennett, R.J. (1979) Space-Time Models and Urban Geographical Research. In Geography and the Urban Environment: Progress in Research and Application. Vol.2. (Eds.) Herbert, D.T. and Johnston, R.J. John Wiley, London. pp.27-58.

Bennett, R.J. and Chorley, R.J. (1978) Environmental Systems: Philosophy, Analysis and Control. Methuen, London.

Blase, J.H.(1979) Hysteresis and Catastrophe Theory: Empirical Identification in Transport Modelling. Environment and Planning A, 11, pp.675-688.

Carter, H. (1980) The Growth of the City System in Wales: A Long Term View of Economic Development and Cultural Change. Paper presented at the Annual Conference of the Institute of British Geographers, University of Lancaster, January.

Chisholm, M., Frey, A.E. and Haggett, P. (Eds.) (1971) Regional Forecasting. Butterworth, London.

Cliff, A.D., Haggett, P., Ord, J.K., Bassett, K.A. and Davies, R.B. (1975) Elements of Spatial Structure: A Quantitative Approach. Cambridge University Press, Cambridge.

Coelho, J.D. and Wilson, A.G. (1976) The Optimum Location and Size of Shopping Centres. Regional Studies, 10, pp.413-421.

Department of the Environment (1977) Housing Policy: A Consultative Document. H.M.S.O., London.

Gurel, O. and Rossler, O.E. (1979) Bifurcation Theory and Applications in Scientific Disciplines. New York Academy of Sciences, New York.

Haggett, P. (1973) Forecasting Alternative Spatial, Ecological and Regional Futures: Problems and Possibilities. In Directions in Geography. (Ed.) Chorley, R.J. Methuen, London. pp.217-236.

Harris, B. and Wilson, A.G. (1978) Equilibrium Values and Dynamics of Attractiveness Terms in Production Constrained Spatial Interaction Models. Working Paper 213, School of Geography, University of Leeds.

Hay, A.M. (1978) Some Problems in Regional Forecasting. In Regions Geographiques et Regions D'Amenagement. (Eds.) Clarke, J.I. and Pelletser, J. Collection les Rommes et les Lettres, 7, Editions Hermes, Lyon.

Hirsch, M.W. and Smale, S. (1974) Differential Equations, Dynamical Systems and Linear Algebra. Academic Press, London.

Isard, W. and Liossatos, P. (1977) Models of Transition Processes. Papres, Regional Science Association, 39, pp.27-59.

Isnard, C.A. and Zeeman, E.C. (1976) Some Models from Catastrophe Theory in the Social Sciences. In Use of Models in the Social Sciences. (Ed.) Colling, L. Tavistock Publications. pp.44-100.

Jordan, D.W. and Smith, P. (1977) Nonlinear Ordinary Differential Equations. Oxford University Press, Oxford.

Karnavou, E. (1978) Urban Planning: Theory, Practice and Ideology. Discussion Paper No.68, Graduate School of Geography, London School of Economics.

Lampard, E.E. (1968) The Evolving System of Cities in the U.S.A.: Urbanisation and Economic Development. In Issues in Urban Economics, Resources for the Future. (Eds.) Perloff, H.S. and Wingo, L. John Hopkins Press, Baltimore.

McLoughlin, J.B. (1973) Control and Urban Planning. Faber and Faber, London.

Marsden, J.E. and McCracken, M. (1976) The Hopf Bifurcation and its Applications. Springer, Heidelberg.

Mees, A.I. (1975) The Revival of Cities in Medieval Europe: An Application of Catastrophe Theory. Regional Science and Urban Economics, 5, pp.403-425.

Nicolis, G. and Prigogine, I. (1977) Self-Organisation in Non-Equilibrium Systems. Edward Arnold, London.

Poston, T. and Wilson, A.G. (1977) Facility Size Versus Distance Travelled: Urban Services and the Fold Catastrophe. Environment and Planning A, 9, pp.681-686.

Scott, A.J. and Roweis, S.T. (1977) Urban Planning in Theory and Practice: A Reappraisal. Environment and Planning, 9, pp.1077-1119.

Stewart, I. (1975) The Seven Elementary Catastrophes. New Scientist, 20th November.

Varian, H.L. (1979) Catastrophe Theory and the Business Cycle. Economic Inquiry, 17, pp.14-28.

Wagstaff, J.M. (1976) Some Thoughts about Geography and Catastrophe Theory. Area, 8, pp.316-320.

Wagstaff, J.M. (1978) A Possible Interpretation of Settlement Pattern Evolution in Terms of 'Catastrophe Theory'. Transaction, Institute of British Geographers. New Series, 3(2). pp.165-178.

Wilson, A.G. (1976) Catastrophe Theory and Urban Modelling: An Application to Modal Choice. Environment and Planning A, 8, pp.351-356.

Wilson, A.G. and Clarke, M.C. (1979) Some Illustrations of Catastrophe Theory Applied to Urban Retailing Structures. In London Papers in Regional Science. (Ed.) Breheny, M. Pion, London.

Zahler, R.S. and Sussmann, H.J. (1978) Catastrophe Thoery as Applied to the Social and Biological Sciences: A Critique. Synthese, 37, pp.117-216.

Zeeman, E.C. (1972) A Catastrophe Machine. In Towards a Theoretical Biology: Vol.4. (Ed.) Waddington, C.H. Edinburgh University Press, pp.276-286.

Zeeman, E.C. and Trotman, D.J.A. (1976) The Classification of Elementary Catastrophes of Codimension < 5. In Structural Stability, The Theory of Catastrophes and Applications in the Sciences. Springer Lecture Notes in Mathematics, 525. Heidelberg. pp.263-327.