CATASTROPHE THEORY

ROBERT GILMORE, *Department of Physics and Atmospheric Science, Drexel University, Philadelphia, Pennsylvania, U.S.A.*

INTRODUCTION

Push on something. It will move. Push just a little bit harder and it will move just a little bit more. This incremental response to incremental stresses is very typical. But under rare conditions a small increase in the level of stress will produce a large and dramatic response. Such a response is called a "catastrophe." This kind of behavior has been summarized succinctly in the phrase "the straw that broke the camel's back."

Although this phenomenon occurs under rare conditions, it is also "typical." That is, although it is unlikely that any *particular* straw will break the camel's back, it is certain that after enough straw has been loaded, *some* straw *will* break the camel's back.

Situations in which gradually increasing stress leads to gradually increasing response, followed by a sudden catastrophic jump to a qualitatively different response state, are all too common. Many examples can be given:

1. Under gradually increased loading a bridge sags to a greater degree, followed by a sudden collapse under the last bit of stress (Zeeman, 1977; Poston and Stewart, 1978; Gilmore, 1981).
2. As temperature is gradually decreased at

1-56081-062-9/92/$5.00+.50

constant pressure, a gas gradually contracts until a certain temperature is reached, at which it condenses to its liquid (or solid) state, with a sudden and very large change in its volume (Zeeman, 1977; Poston and Stewart, 1978; Gilmore, 1981).

3. Slow variation of the aileron trim settings (wing flaps) of an airplane leads to slow change in the attitude of the plane, until a certain threshold value is passed, at which point there occurs a large change in aircraft attitude (Gilmore, 1981).
4. When a glass tube filled with a helium–neon or carbon dioxide–neon mixture is set inside a cavity formed by two highly reflecting mirrors and a gradually increasing current is passed through it, the amount of incoherent light emitted gradually increases. After a current threshold is passed, the light which is emitted increases rapidly in intensity and coherence: the laser has turned "on" (Poston and Stewart, 1978; Gilmore, 1981).
5. The amount of sunlight falling on the Earth varies gradually over about 1000 years due to variation of the Earth's orbital parameters. Sudden dramatic changes leading to the occurrence and disappearance of the Ice Ages occur during this period and seem to be precipitated only by this gradual variation of the Earth's orbital parameters (Gilmore, 1981).
6. It is even possible to conceive of small causes producing big effects in biological, economic, social, and political systems (Thom, 1975; Zeeman, 1977; Poston and Stewart, 1978).

Catastrophes are widespread, occurring throughout all fields in the scientific and engineering disciplines, and even beyond. Moreover, the mathematical description of a catastrophe follows the same procedure and draws from the same restricted set of functions, independent of whether the catastrophe occurs in the area of physics, chemistry, structural engineering, aircraft dynamics, climate dynamics, etc. The mathematical description of catastrophes, involving mathematical functions called elementary catastrophes, provides a language in terms of which discontinuous phenomena can adequately be described.

It is easy to imagine how to describe systems in which a small push gives rise to a small response. One expects the mathematics of continuity, the calculus of Newton, to be applicable. One also expects that linearization about the local state (linear response function) will give a quantitative estimate of how much the system will respond to a small push.

But what of the mathematics of discontinuity? How does one describe catastrophes? Is it necessary to give up the ideas of continuity?

Roughly speaking, the state of a system is an equilibrium—in fact, a stable equilibrium. By changing some external parameters called control parameters (stress, loading), the equilibrium is displaced. A small change in these parameters usually results in a small displacement of the equilibrium. Sometimes small parameter changes result in the appearance of new equilibria or the disappearance of old equilibria. It is the latter instance in particular that can lead to a catastrophic sudden jump. A systematic study of catastrophes is closely related to a systematic study of equilibria, and especially the appearance and disappearance of equilibria.

Families of functions depending on (control) parameters are called catastrophes when the number of equilibria they possess changes as the parameters are varied. There are only a small number of catastrophes. A very small number have been used to model sudden jumps in physical systems. We will study the mathematics of catastrophes in two steps:

1. We first study functions representing situations in which the number of equilibria is about to change. This occurs when two or more equilibria occur at the same point (become degenerate).
2. We then study the effects of perturbations on the degenerate equilibria. In fact, we identify the simplest perturbation which can reproduce the effects of the most general perturbation.

That both these programs can be carried out successfully is remarkable.

A major obstacle in applying the mathematics of catastrophe theory to physical systems is in identifying the underlying catastrophe. A multiplicity of phenomena occur in the presence of a catastrophe (catastrophe flags). The occurrence of any one of these is an indication that others are present and can be found, and that a catastrophe is ultimately responsible for all. The catastrophe flags are easy to recognize and provide an inordinate amount of information about the underlying catastrophe. This information includes the type of catastrophe, a rough indication of where the sudden jump may occur, and—most important—how to avoid it if that is a desirable objective.

This article is organized into three parts. Section 1 describes what catastrophe theory is, Sec. 2 describes how the catastrophe functions are constructed, and Sec. 3 describes how catastrophe theory is applied to the description of phenomena that occur in the science and engineering disciplines.

In Sec. 1 we describe a progression of three theorems of elementary calculus. These theorems describe local standard forms for functions in the neighborhood of a point. The first, the implicit function theorem, tells us that a function can be replaced by its linear approximation when its slope is nonzero at a point. The second theorem, the Morse lemma, tells us that under suitable conditions a function can be well approximated by a quadratic form in the neighborhood of an equilibrium. The third result, the Thom splitting lemma, describes what happens when the "suitable conditions" required above are not satisfied. In this case two or more equilibria occur very close together (are degenerate), making it possible for a small perturbation either to split or annihilate the equilibria. This change in the number of equilibria is a "catastrophe." Tables containing Thom's list of elementary catastrophes and a complete list of all elementary catastrophes are provided. This is followed by a discussion of the geometric properties of the very simplest catastrophe functions. This section provides a clear answer to one of the questions posed above: it is possible to describe discontinuous phenomena without giving up the ideas of continuity. In fact, this mathematics of discontinuity is an essential part of the calculus of Newton.

In Sec. 2 we compute some explicit catastrophe functions. The first example starts with a family of functions depending on one state variable and two control parameters. Following an algorithmic procedure, we reduce this to a standard form in the neighborhood of its most degenerate equilibrium. This degenerate equilibrium is then perturbed in order to determine how these critical points can be created and annihilated as a function of the control parameters. The algorithm developed to effect this reduction to normal form is then summarized and applied to a more complicated family of functions.

In Sec. 3 we address the question of how to apply the mathematics of catastrophe theory to real world processes that exhibit discontinuous phenomena. Two widely adopted conventions are first described. These are assumptions about the mathematical characterization of the equilibrium state of a physical system: whether it is determined by a local minimum or the global minimum of a potential, and the conditions under which a jump from one minimum to another occurs. The mathematics tells us only where, how many, and what type of equilibria a catastrophe function possesses; the convention isolates the physically important equilibria. When a physical system exhibits a catastrophe, a multiplicity of phenomena occur. It is useful to be able to recognize them—particularly if one wants to avoid the physical catastrophe (e.g., bridge collapse). These phenomena are called catastrophe flags. The presence of any one is an indication that the others are present. Their recognition provides a great deal of information about the underlying catastrophe. The use of catastrophe conventions and catastrophe flags is illustrated in the context of an important example in Sec. 3.3. This illustrates unexpected dangers which may arise in the design of structures following standard optimization criteria employed to reduce costs. In Sec. 3.4 we indicate how the elementary catastrophes may make their appearance within the broader program of catastrophe theory, in the field of dynamical systems theory. That is, the fold catastrophe appears in the guise of saddle-node bifurcations while the cusp catastrophe appears in the guise of pitchfork and Hopf bifurcations.

In a short Appendix we outline the early turbulent and confusing history of catastrophe theory, which was at one point heralded as the greatest advance in mathematics since the development of the calculus by Newton.

1. WHAT IT IS

In Sec. 1 of this article we describe the enormous mathematical program called catastrophe theory and the much smaller and more manageable mathematical program called elementary catastrophe theory (Sec.

1.1). This latter program is not only the starting point for the study of the larger program, but also a continuation of a very important and rather simple program in elementary calculus. This is the program of determining and classifying the standard, or canonical, forms that functions can assume in any of their neighborhoods.

The first two stages in this program are well known at an intuitive level (Sec. 1.2). These are the implicit function theorem and the Morse lemma, which are the mathematical justifications underlying the approximation of a function by a linear function at a point where its slope is nonzero, or by a quadratic form at an equilibrium (where the slope is zero). The third stage, the Thom splitting lemma, describes what happens when two or more equilibria become degenerate. Under this condition a perturbation can either split the equilibria or annihilate them. Change in the number of equilibria is closely related to the occurrence of sudden jumps in physical systems.

Functions that describe sudden jumps, or changes in the number of equilibria of a system, are called catastrophes. There is a small number of elementary catastrophes. They are classified in Sec. 1.3 and presented in Tables 1 and 2 (Sec. 1.4). Since the number of catastrophes is small, the properties of each can be studied in detail (Sec. 1.5); the results are then directly applicable to any physical system described by that mathematical function. The geometric properties of the two simplest of the elementary catastrophes are studied in detail in Sec. 1.6. In Sec. 1.7 we make quantitative the statement that a small push will displace an equilibrium by just a little bit by computing the linear response function for a large class of physical systems, and showing that this function diverges precisely when a sudden jump is imminent.

1.1 The Program of Catastrophe Theory

Catastrophe theory is a program. The purpose of the program is to determine how the qualitative properties of solutions of equations change as the parameters that appear in the equations change (Gilmore, 1981).

It often happens that small changes in the values of parameters that appear in equations produce only small quantitative changes in the solutions of the equations. However, there may be parameter values for which a small change, either in parameter values, initial conditions, or boundary conditions, produces a large quantitative change in the solutions to these equations. Large quantitative changes in solutions describe qualitative changes in the behavior of the system. Catastrophe theory is concerned with determining the parameter values at which qualitative changes occur in solutions of equations described by parameters (Thom, 1975; Zeeman, 1977; Poston and Stewart, 1978; Gilmore, 1981).

This is an ambitious and difficult program. For example, for systems of equations of the form

$$F_\alpha(x,x',t;c)=0, \tag{1}$$

where F_α is a set of functions; x is an n-vector, $x=(x_1, x_2,\ldots,x_n)\in R^n$, called a state vector; c is a k-vector, $c=(c_1, c_2,\ldots,c_k)\in R^k$, called control parameters; and $'=d/dt$, there are no general results. When the set of equations is restricted to the simpler form of coupled nonlinear first-order ordinary differential equations (also called dynamical systems) of the form

$$x_i'=f_i(x,t;c), \tag{2}$$

very little can be said in general. Many results are known when $n=2$ and f is independent of t. A few results are known when $n=2$ and the forcing term is periodic, $f(x,t;c)=f(x,t+T;c)$. Much less is known when $n=2$ and f is not periodic. The case $n>2$ invites a lifetime of work.

When the forcing function in the dynamical system equations (2) is independent of time and can be written as the gradient of some potential,

$$f_i=-\partial V(x;c)/\partial x_i, \tag{3}$$

then the system

$$x_i'=-\partial V(x;c)/\partial x_i \tag{4}$$

is called a gradient dynamical system. For such systems many results are available.

The qualitative properties of a gradient dynamical system can be constructed by investigating the phase-space portrait of its flow. This can be done by plotting the value of the potential as a function of the phase-space coordinates x_i. The phase-space flow is "downhill" on the potential function. It is easily determined in the neighborhood of each equilibrium, or critical point, independent of the stability of the equilibrium. The local flow portraits around each critical point can then be pasted together to determine a global phase portrait. The potential

$$V(x,y;a,b)=\tfrac{1}{4}x^4+\tfrac{1}{2}ax^2+bx+\tfrac{1}{2}y^2 \qquad (5)$$

is illustrated in Fig. 1, together with the phase-space portrait in the state-variable space.

Small changes in the control-parameter values (a,b) typically produce small changes in the location of the critical points. In turn, this produces only small quantitative changes, and therefore no qualitative change, in the phase-space portrait. Qualitative changes will only occur when changes in the control-parameter values result in changes in the number of critical points. This number can change only when two or more critical points coalesce and annihilate or, viewed from the other direction, two or more critical points are created in phase space and then move apart from each other as the control parameters are varied.

Elementary catastrophe theory is the study of how the critical points of a potential, $V(x;c)$, move about, coalesce and annihilate each other, or are created and disperse from each other, in state space $x \in R^n$ as the control parameters $c \in R^k$ are varied.

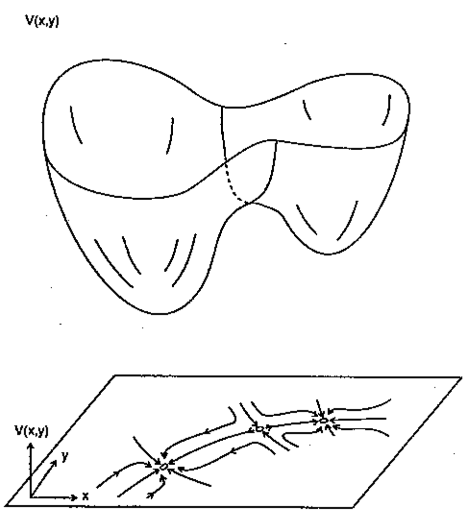

FIG. 1. The dynamics of a gradient system are governed by a potential. The potential of Eq. (5) is shown, together with the flows to and from the stable and unstable equilibria. These are projected down into the *x-y* plane.

1.2 Three Theorems from Elementary Calculus

Elementary catastrophe theory is found at the intersection of two lines of mathematical development, one old, one new. On the one hand, it is the latest development in the quest, in elementary calculus, for standard local forms for functions. On the other hand, it is the first result in the quest, in catastrophe theory, for canonical representations of functions that show qualitative changes when control parameters are varied.

Elementary catastrophe theory is the third in a series of reduction-to-standard-form theorems in elementary calculus. The three developments are the implicit function theorem, which depends on the first derivatives of a function; the Morse lemma, which depends on the second derivatives of a function; and the Thom splitting lemma, which depends on the third (and higher) derivatives of a function. Each of these results provides a standard, or canonical, form for a function in the neighborhood of a point.

We summarize these results now.

1.2.1 Implicit Function Theorem. The implicit function theorem tells us that if the slope of a function is nonzero at a point, the function can be represented locally by a linear approximation to that function. In a rough sense, it tells us that it is justified to linearize a function about a point at which its derivative is nonvanishing.

Implicit Function Theorem: Let $f(x)=f(x_1, x_2,\ldots,x_n)$ be a function with nonzero gradient at x_0:

$$\nabla f|_{x_0} \neq 0. \qquad (6)$$

Then it is possible to find a new coordinate system, $y=(y_1,y_2,\ldots,y_n)$, $y=y(x)$, so that

$$f=y_1. \qquad (7)$$

That is, f is equal, after a smooth change of coordinates, to y_1.

1.2.2 Morse Lemma. The Morse lemma takes over where the implicit function theorem leaves off. Suppose the gradient of a function does vanish at a point—what then? Such a point is called an equilibrium, or critical point. Provided that the function has enough "curvature" at the critical point, it can be represented locally by a quadratic form. In a rough sense, the Morse lemma tells us that it is justified to represent a function at an equilibrium by a quadratic form, provided none of the eigenvalues vanish.

Morse Lemma: Let $f(x)=f(x_1,x_2,\ldots,x_n)$ be a function with equilibrium at x_0 and nonsingular stability matrix at x_0:

$$\begin{aligned} &\nabla f|_{x_0}=0, && \text{equilibrium,} \\ &\det[\partial^2 f/\partial x_i \partial x_j]|_{x_0}\neq 0, && \text{nonsingular.} \end{aligned} \tag{8}$$

Then there is a smooth change of coordinates, $x'=x'(x)$, so that

$$f=\sum \lambda_i (x_i')^2, \tag{9}$$

where λ_i are the eigenvalues (all nonzero) of the stability matrix.

A critical point satisfying the conditions (8) is called variously a Morse critical point or an isolated critical point.

The quadratic form (9) can be put into a canonical form by rescaling the coordinates:

$$y_i=|\lambda_i|^{1/2}x_i'. \tag{10}$$

Under this scale transformation the function at equilibrium assumes Morse canonical form:

$$\begin{aligned} &f=M_i^n, \\ &M_i^n=-y_1^2-\cdots-y_i^2+y_{i+1}^2+\cdots+y_n^2. \end{aligned} \tag{11}$$

The quadratic forms (11) are called Morse saddles. The Morse saddle M_0^n has a minimum at $y=0$, while M_n^n has a maximum at $y=0$. The remaining Morse saddles M_i^n, $i\neq 0,n$, have equilibria at $y=0$ that are neither maxima nor minima.

1.2.3 Thom Splitting Lemma. The Thom splitting lemma takes over where the Morse lemma leaves off. Suppose the stability matrix of a function is singular at an equilibrium. Then one or more eigenvalues (λ_i) vanish. What then? The Thom splitting lemma tells us that there is a smooth change of coordinates, $x'=x'(x)$, where $x_1',\ldots,x_l'$ are tangent to the eigenvectors with vanishing eigenvalues at the critical point, and $x_{l+1}',\ldots,x_n'$ are tangent to the eigenvectors with nonvanishing eigenvalues at the critical point, so that the function can be broken down into two parts. One part, associated with the nonzero eigenvalues, is simple and can be put into Morse canonical form. The other part, associated with the vanishing eigenvalues, is interesting and has all its second derivatives equal to zero at the critical point. This non-Morse function is the principal object of study in elementary catastrophe theory.

Thom Splitting Lemma: Let $f(x)=f(x_1,x_2,\ldots,x_n)$ be a function with equilibrium and singular stability matrix at x_0:

$$\begin{aligned} &\nabla f|_{x_0}=0, && \text{equilibrium,} \\ &\det[\partial^2 f/\partial x_i \partial x_j]|_{x_0}=0, && \text{singular.} \end{aligned} \tag{12}$$

If the stability matrix has exactly l vanishing eigenvalues, then there is a smooth change of coordinates, $x'=x'(x)$, so that

$$\begin{aligned} &f(x)=f_{\text{NM}}(x_1',\ldots,x_l')+M_i^{n-l}(x_{l+1}',\ldots,x_n'), \\ &\partial^2 f_{\text{NM}}/\partial x_i \partial x_j|_{x_0}=0, \quad 1\leqslant i,j\leqslant l. \end{aligned} \tag{13}$$

The Thom splitting lemma can be proven by the methods of elementary calculus. It tells us that when the Morse lemma is not applicable, the function can be split into two functions, a "good" function in $n-l$ coordinates which can be put into Morse canonical form and a "bad," or non-Morse, function of l variables which bears further scrutiny. It tells us nothing about the non-Morse function except that its Taylor series expansion about the critical point begins with at least third-degree terms.

We emphasize here that the three results of elementary calculus, the implicit function theorem, the Morse lemma, and the Thom splitting lemma, depending on first, second, and third derivatives, are local in nature. The theorems do not provide an estimate for the size of the neighborhood for which the statement of the result is true.

1.3 Thom Classification Theorem

For a typical function, $f(x)$, the gradient at a random ("typical") point will be nonvanishing, so that the implicit function theorem is applicable. There are, however, typically isolated points at which the gradient vanishes. At such points the stability matrix is typically nonsingular, so that the Morse lemma is applicable. How, then, does it come about that the machinery of elementary catastrophe theory becomes useful?

When the function depends on control parameters c as well as state variables x, $f=f(x;c)$, then the eigenvalues of the stability matrix at a critical point, $x_0=x_0(c)$, depend on the control-parameter values: $\lambda_i=\lambda_i(c)$. As a result, there may be choices of the control-parameter values that annihilate one or more of the eigenvalues.

As a result, the structure of the non-Morse function in Eqs. (13) will depend on control parameters. The Thom classification theorem, which is outside the scope of calculus (elementary or otherwise), provides a further resolution of the non-Morse function into two functions. One of these, the catastrophe germ, depends only on the l state variables $y_1, y_2, \ldots, y_l$ and summarizes the nature of the singularity at the non-Morse critical point. The other function, the universal perturbation, is a function of both the l state variables and k control parameters. This function summarizes what can happen to the singularity, or degenerate critical point, under the most general possible ("universal") perturbation.

Thom Classification Theorem: Let $f_{NM}(y;c)=f(y_1,\ldots,y_l;c_1,\ldots,c_k)$ be a non-Morse function of l state variables and k control parameters. Then there is a smooth change of coordinates so that

$$f_{NM}(y;c)=\mathrm{Cat}(l,k). \qquad (14)$$

The elementary catastrophe function, Cat(l,k), is the sum of two terms:

$$\mathrm{Cat}(l,k)=\mathrm{CG}(l)+\mathrm{Pert}(l,k). \qquad (15)$$

The catastrophe germ, CG(l), depends only on the l state variables. All its second partial derivatives vanish at the critical point. The universal perturbation depends on the k control parameters as well as the l state variables. The dependence of Pert(l,k) on the control-parameter values is linear. For "most" choices of control-parameter values (all but a set of measure zero) the function Cat(l,k) has isolated critical points.

The Thom classification theorem, like the three results of elementary calculus described in the previous section, is local in nature. The theorem does not provide an estimate of the size of the neighborhood for which the statement of the theorem is true.

1.4 Thom's List of Elementary Catastrophes

Thom's original classification theorem provided a list of the elementary catastrophes (Thom, 1975). A slightly expanded version of this list is provided in Table 1 (Arnol'd, 1981, 1986). This list contains the canonical catastrophe functions for $k<6$ and therefore (cf. Sec. 2.3) $l<3$. This list consists of the classification of the function following the beautiful convention introduced by Arnol'd (1981), the original descriptive name, when it exists (Thom, 1975; Zeeman, 1977; Poston and Stewart, 1978), values for k and l, the catastrophe germ, CG(l), and the universal perturbation, Pert(l,k). Thom's original list contained only the seven members with $k\leqslant 4$ (dimension of spacetime) for unsupportable historical reasons.

The catastrophe functions listed in Table 1 are elementary in the sense that all coefficients in the catastrophe germ can be assigned canonical values. There are no free parameters; every coefficient in the catastrophe germ can be given canonical numerical values such as $\pm 1,0$ by a coordinate change. For example, a term of the form $-3x^4$ in the A_{-3} catastrophe could be transformed to the canonical

Table 1. All catastrophes up to control-parameter dimension five are elementary and are listed below by dimension of control-parameter space.

k	l	Classification	Name	CG(l)	Pert(l,k)
1	1	A_2	Fold	x^3	a_1x^3
2	1	$A_{\pm 3}$	Cusp	$\pm x^4$	$a_1x^2+a_2x^2$
3	1	A_4	Swallowtail	x^5	$a_1x^2+a_2x^2+a_3x^3$
3	2	D_{-4}	Elliptic umbilic	x^2y-y^3	$a_1x+a_2y+a_3x^2$
3	2	D_{+4}	Hyperbolic umbilic	x^2y+y^3	$a_1x+a_2y+a_3x^2$
4	1	$A_{\pm 5}$	Butterfly	$\pm x^6$	$a_1x+a_2x^2+a_3x^3+a_4x^4$
4	2	D_5	Parabolic umbilic	$\pm(x^2y+y^4)$	$a_1x+a_2y+a_3x^2+a_4y^2$
5	1	A_6	Wigwam	x^7	$a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5$
5	2	D_{-6}	Second elliptic umbilic	x^2y-y^5	$a_1x+a_2y+a_3x^2+a_4y^2+a_5y^3$
5	2	D_{+6}	Second hyperbolic umbilic	x^2y+y^5	$a_1x+a_2y+a_3x^2+a_4y^2+a_5y^3$
5	2	$E_{\pm 6}$	Symbolic umbilic	$\pm(x^3+y^4)$	$a_1x+a_2y+a_3xy+a_4y^2+a_5xy^2$

form $-x'^4$ (or $-\frac{1}{4}x'^4$) by an appropriate scale change $x'=\lambda x$.

Thom's list of elementary catastrophes is not a complete list. The complete list is provided in Table 2. There are two infinite series of elementary catastrophes and one finite series. One infinite series, the cuspoids $A_{\pm k}$, depends on only one state variable. The other infinite series, the umbilics $D_{\pm k}$, depends on two state variables. The exceptional elementary catastrophes, $E_{\pm 6}$, E_7, E_8, depend on two state variables (Arnol'd, 1981, 1986; Poston and Stewart, 1978; Gilmore, 1981).

There is a remarkable correspondence between the classification theory of elementary catastrophes and the classification theory for Lie algebras all of whose roots (in the root space diagram) have the same length (Arnol'd, 1981, 1986; Gilmore, 1981). The correspondence is as follows. The phase-space portraits of the elementary catastrophes with maximum number of isolated critical points can be summarized by drawing the flow from each equilibrium to any of the others to which a flow is possible. The phase-space portraits so obtained are exactly the Dynkin diagrams which classify all the simple Lie algebras whose roots have equal length; these are A_{n-1}, D_n, and the exceptional simple Lie algebras E_6, E_7, and E_8. This nomenclature for simple Lie algebras has accordingly been adapted to the classification of elementary catastrophes. For Lie algebras the subscript (e.g., 8 for E_8) denotes its rank; for elementary catastrophes the subscript denotes the number of isolated (complex) critical points generated by an arbitrary perturbation of the function. This is the maximum number of real critical points into which the non-Morse critical point splits under a general perturbation.

Tables 1 and 2 differ in a subtle way, indicating that they are responses to somewhat different questions. The question to which Table 1 responds is: "Up to what control-parameter dimension are all catastrophes elementary, and what are they?" The question to which Table 2 responds is: "For each control dimension $k=1,2,\ldots$, what are the elementary catastrophes?"

The difference between the two tables indicates that for control dimension $k\geqslant 6$ there are catastrophes that are elementary and those that are not, while for $k<6$ all catastrophes are elementary. We will explore what happens at $k=6$ to generate nonelementary (modal) catastrophes as well as elementary catastrophes in Sec. 2.3. Briefly, the result is as follows. To annihilate l eigenvalues of the stability matrix requires $k\geqslant l(l+1)/2$ control parameters. A linear transformation can be used in an attempt to provide canonical values for cubic terms in the Taylor series expansion

Table 2. There are three series of elementary catastrophes. The cuspoids A_k depend on one state variable while the umbilics D_k and the exceptional catastrophes E_k, $k=6,7,8$, depend on two state variables. The subscript k indicates the maximum number of real Morse critical points the catastrophe splits into under an arbitrary perturbation. The number of control parameters required in the universal perturbation is $k-1$.

Symbol	Catastrophe germ	Universal perturbation
$A_{\pm k}$[a]	$\pm x^{k+1}$	$\sum_{j=1}^{k-1} a_j x^j$
$D_{\pm k}$	$x^2y \pm y^{k-1}$, k even $\pm(x^2y+y^{k-1})$, k odd	$\sum_{j=1}^{k-3} a_j y^j + \sum_{j=k-2}^{k-1} a_j x^{j-(k-3)}$
$E_{\pm 6}$	$\pm(x^3+y^4)$	$\sum_{j=1}^{2} a_j y^j + \sum_{j=3}^{5} a_j xy^{j-3}$
E_7	x^3+xy^4	$\sum_{j=1}^{4} a_j y^j + \sum_{j=5}^{6} a_j xy^{j-5}$
E_8	x^3+y^5	$\sum_{j=1}^{3} a_j y^j + \sum_{j=4}^{7} a_j xy^{j-4}$

[a] $A_{+k}=A_{-k}$ if k is even.

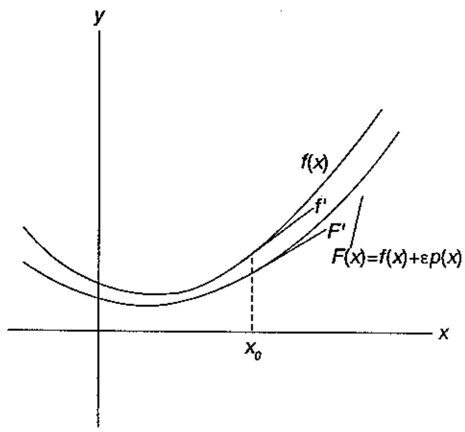

FIG. 2. At a point where the implicit function theorem is applicable, perturbation of a function does not produce a qualitative change in the function.

of the non-Morse function. When the number of cubic coefficients $[(l+3-1)!/(l-1)!3!]$ exceeds the number of degrees of freedom in the $l\times l$ linear transformation (l^2), then it is not possible to assign canonical values to all cubic terms in the Taylor series expansion of the non-Morse function, and the resulting catastrophe germ cannot be elementary. This first occurs for $l=3$ ($\Rightarrow k=6$). For $k\geqslant 6$ there are catastrophes that are elementary and those that are not. For $k<6$ all catastrophes are elementary.

1.5 Why a List of Perturbations Is Required

The first two canonical form theorems of elementary calculus are clean and simple. If the function has certain properties at a point, then the canonical form in the neighborhood of the point is provided by the statement of the theorem (implicit function theorem, Morse lemma). By contrast, the third result is not nearly so clean cut. The Thom splitting lemma tells us that we can decompose a function at a non-Morse critical point into the sum of two functions, one Morse, the other interesting. The classification theorem provides a list of the interesting functions by number of state variables (l) and control parameters (k). Why is it that the classification theorem, in addition to providing a list of canonical forms for catastrophe germs, in the spirit of the implicit function theorem and the Morse lemma, also provides a list of canonical perturbations? The reason is that perturbation of the canonical linear form which is provided by the implicit function theorem does not change its qualitative properties. The same is true for the canonical quadratic form which is provided by the Morse lemma. However, perturbation of the canonical singularity CG(l) provided in Tables 1 and 2 produces dramatic changes in its qualitative properties. Different perturbations produce different qualitative changes. The canonical perturbation, Pert(l,k), of each catastrophe germ is the "smallest" function, in the sense of number of control parameters required, which incorporates all distinct qualitative changes produced by all possible perturbations of the catastrophe germ.

We illustrate these statements in Figs. 2–4. In Fig. 2 we show a function, $f(x)$, which satisfies the conditions of the implicit function theorem at x_0. Under a perturbation, $\varepsilon p(x)$, where $p(x)$ is a well-behaved function and ε is a small parameter, the new function, $F(x)=f(x)+\varepsilon p(x)$, also satisfies the conditions of the implicit function theorem at x_0 for ε sufficiently small. Therefore, perturbation of $f(x)$ at x_0 does not change its qualitative properties (perturbation "commutes" with the implicit function theorem).

In Fig. 3 we show a function, $f(x)$, which satisfies the conditions of the Morse lemma at x_0. Under a perturbation, $\varepsilon p(x)$, the new function, $F(x)=f(x)+\varepsilon p(x)$, no longer appears to satisfy the conditions of the Morse lemma at x_0, since typically $p'(x_0)\neq 0$. However, $F'(x_0)=\varepsilon p'(x_0)$ is small for small ε so

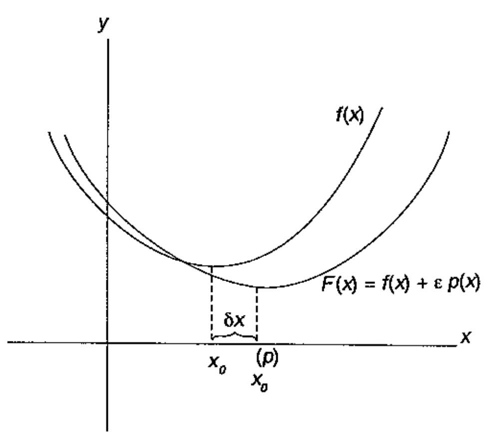

FIG. 3. At a point where the Morse lemma is applicable, perturbation of a function does not produce a qualitative change in the function.

that the implicit function theorem is on the verge of not being applicable. It is more useful to regard the perturbation as moving the location of the critical point to a nearby point, $x_0^{(p)} = x_0 + \delta x_0$. At this perturbed critical point the conditions for the Morse lemma are fulfilled. Thus, a small perturbation at an equilibrium produces only a small displacement of that equilibrium. Not only that, but the curvature or, more generally, the Morse saddle type of the canonical form remains unchanged. As a result, perturbation of a function that satisfies the conditions for the Morse lemma does not produce a qualitative change of the function in the neighborhood of the critical point.

The situation is quite different for a catastrophe germ. In Fig. 4 we plot the catastrophe function A_2: $f(x;a_1)=\frac{1}{3}x^3+a_1x$ for three values of the control parameter a_1. For $a_1=0$ the catastrophe germ $f(x;a_1=0)=x^3/3$ has a doubly degenerate critical point at $x=0$. The perturbation with $a_1<0$ splits this doubly degenerate critical point into two isolated critical points at $x_-=-(-a_1)^{1/2}$ and $x_+=+(-a_1)^{1/2}$. The perturbation with $a_1>0$ removes the critical point altogether. These are the only two qualitatively distinct things that can occur to a doubly degenerate critical point under an arbitrary perturbation. These correspond to scattering of the solutions of $\nabla f(x;a_1)=0$ from the real axis to the imaginary axis as a_1 increases through zero, with the point of double degeneracy occurring at $a_1=0$. It is a remarkable result that the simple perturbation $P(1,1)=a_1x$ encapsulates all distinct possibilities under generic perturbations.

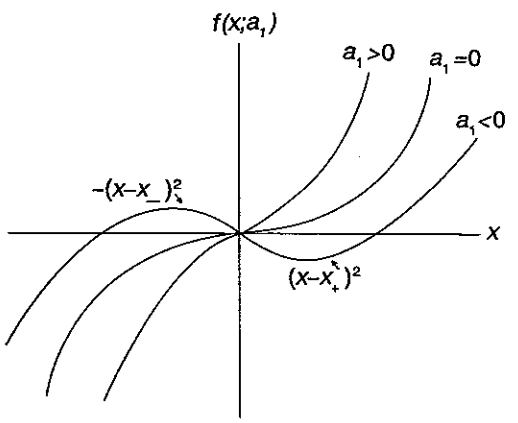

FIG. 4. At a point where a function has a degenerate critical point, so that neither the implicit function theorem nor the Morse lemma is applicable, perturbation of a function produces a qualitative change in the function. In the case shown, perturbation of the function x^3 either annihilates the critical points or splits them into two nondegenerate critical points.

It should now be clear why the catastrophe germs listed in the classification theorem must be accompanied by a list of universal perturbations while the implicit function theorem and the Morse lemma are not encumbered by such baggage. The canonical linear and quadratic forms are invariant under perturbation: perturbation produces no qualitative change. However, the catastrophe germs undergo a wide spectrum of distinct qualitative changes under perturbation. The perturbation functions listed are those of minimal control-parameter dimension which are capable of reproducing the entire spectrum of distinct qualitative changes induced by the most general perturbation.

It is a remarkable result that the control parameters appear linearly in these perturbations.

1.6 Geometry of the Fold and the Cusp

In this section we review the properties of the two simplest elementary catastrophes, the fold catastrophe A_2 and the cusp catastrophe $A_{\pm 3}$. Since the cusp catastrophe A_{+3} occurs more frequently in physical applications than A_{-3} (which is not bounded below), we study specifically the properties of A_{+3}. The properties of A_{-3} are related by appropriate sign changes. We also review a restricted set of properties of the co(ntrol)-dimension three catastrophes A_4, $D_{\pm 4}$.

For the fold and the cusp we study the following properties:

1. typical functions in the family of functions as well as the bifurcation set;
2. location of the critical points;
3. values of the function at the critical points;
4. curvature of the function at the critical points.

We present only the bifurcation set for the three catastrophes A_4, $D_{\pm 4}$.

1.6.1 Geometry of the Fold Catastrophe. The fold catastrophe is

$$A_2:\ f(x;a)=\tfrac{1}{3}x^3+ax. \qquad (16)$$

The canonical properties of this function are shown in Fig. 5. In Fig. 5(a) we show members of this family with $a>0$, $a=0$, $a<0$. The bifurcation set is the set of points in the control parameter space at which there is a qualitative change in the nature of the function. This occurs when two or more critical points become degenerate. For the fold catastrophe this consists of the single point $a=0$, at which there is a doubly degenerate critical point at $x=0$.

The location of the critical points, the solution of $\nabla f(x;a)=0$, is shown in Fig. 5(b). The critical points, $x_\pm(a)=\pm(-a)^{1/2}$, have a standard 1/2 power-law dependence on the control parameter a. Note that the critical points $x_\pm(a)$ exist only for $a\leqslant 0$ and that the graph of $x_\pm(a)$ as a function of a is a smooth manifold embedded in the space R^1 (state space) $\times R^1$ (control parameter space). This is a general result. The fold catastrophe derives its name from the shape of its critical set $\nabla f(x;a)=0$, which looks like a curve folded over itself.

The value of the function at the critical points,

$$f_c(x_\pm(a);a)=\pm\tfrac{2}{3}(-a)^{3/2}, \qquad (17)$$

is plotted in Fig. 5(c). This curve has a canonical 3/2 power-law dependence. This graph is not generally a manifold.

The curvature of the function,

$$f''(x_\pm(a);a)=\pm 2(-a)^{1/2}, \qquad (18)$$

is shown in Fig. 5(d). Although the critical curvature happens to be a manifold in the present case, this is not generally true for the remaining catastrophes.

1.6.2 Geometry of the Cusp Catastrophe. The cusp catastrophe is

$$A_{+3}\text{: } f(x;a,b)=\tfrac{1}{4}x^4+\tfrac{1}{2}ax^2+bx. \qquad (19)$$

The canonical properties are shown in Fig. 6. In Fig. 6(a) we show the control-parameter plane $R^2=(a,b)$, various points in this plane, and the function $f(x;a,b)$ evaluated at these points. Within the cusp-shaped region the

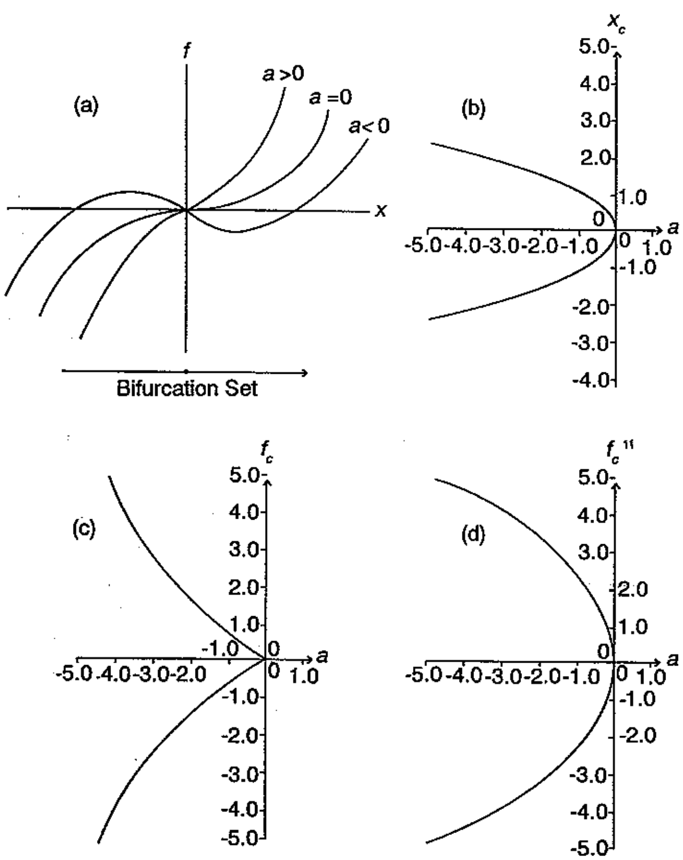

FIG. 5. (a) Members of the fold family of functions $f(x;a)=\frac{1}{3}x^3+ax$ for various values of the control parameter a. (b) Location of the critical points as a function of a. (c) Value of the function at its critical points. (d) Curvature of the function at its critical points.

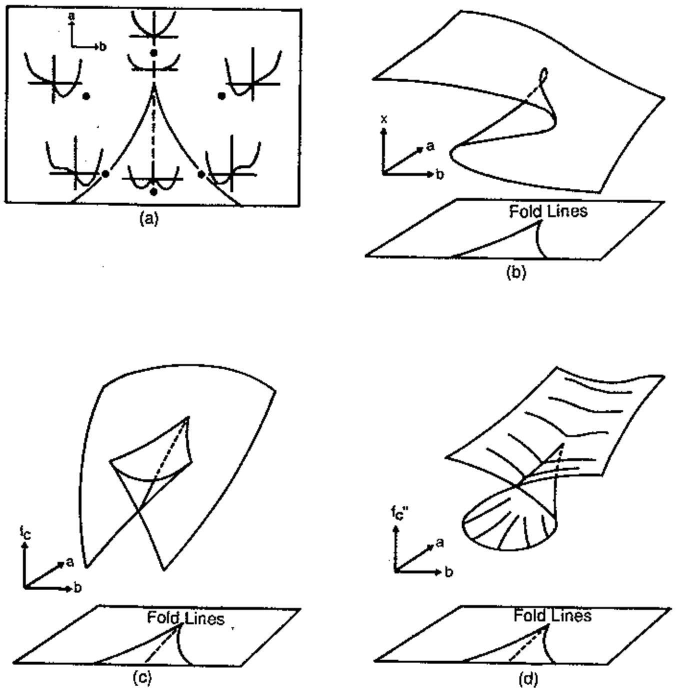

FIG. 6. (a) Members of the cusp family of functions $f(x;a,b)=\frac{1}{4}x^4+\frac{1}{2}ax^2+bx$ for various values of the control parameters a,b. (b) Location of the critical points as a function of position over the control parameter plane. (c) Value of the function at its critical points. (d) Curvature of the function at its critical points.

function has three isolated critical points, two minima separated by a local maximum. Outside the cusp-shaped region the function has only a single minimum. These two regions in the control plane parametrize two qualitatively distinct types of functions. Any path from one region to the other must pass through the cusp-shaped curve, along which there is a doubly degenerate critical point (triply degenerate at the tip of the cusp). This degeneracy occurs when the local maximum collides with one of the two minima. Entering the larger region complementary to the cusp-shaped region, the two degenerate critical points annihilate each other in a catastrophe which is of type A_2.

The cusp-shaped bifurcation set is determined by the condition that a critical point $[\nabla f(x;a,b)=0]$ is degenerate $[\nabla^2 f(x;a,b)=0]$:

$$\begin{aligned}&\text{critical point:} \quad f'(x;a,b)=x^3+ax+b=0,\\&\text{degenerate:} \quad f''(x;a,b)=3x^2+a=0.\end{aligned} \tag{20}$$

From these two equations we compute the semicubical parabola

$$\begin{aligned}a&=-3x_c^2,\\b&=2x_c^3,\end{aligned} \tag{21a}$$

along which a critical point is degenerate. The projection of this space curve into the control parameter plane is

$$(a/3)^3+(b/2)^2=0. \tag{21b}$$

This is the bifurcation set shown in Fig. 6(a).

In Fig. 6(b) we show the critical point(s), $x_c(a,b)$, as a function of the control parameters a,b. These points lie on the critical manifold or catastrophe manifold $\nabla f(x;a,b)=0$. Outside the cusp-shaped region there is a single critical point. Over the cusp-shaped region there are three. The middle critical point is the local maximum which separates the two minima. Moving toward the edge of the cusp-shaped region, two of the critical points move together. They collide on the bifurcation set and annihilate each other beyond the bifurcation set. The graph $x_c(a,b)$ in R^1 (state space) $\times R^2$ (control parameter space) is a smooth two-dimensional manifold. The locus of points on this manifold where the tangent is "vertical" is the semicubical parabola (21). From another point of view, the cusp-shaped bifurcation set in R^2 is the projection into R^2 of the fold in the manifold $x_c(a,b)$ in $R^1\times R^2$. The singularity in this catastrophe lies not in the catastrophe manifold itself, which is smooth, but in the projection of this two-dimensional manifold down into the two-dimensional control parameter space. In general, the graph of $\nabla f(x;c)=0$, with $x \in R^n$ and $c \in R^k$, is a smooth k-dimensional manifold embedded in $R^n \times R^k$. The only singularity occurs in the projection of this k-dimensional manifold into the k-dimensional space of control parameters.

In Fig. 6(c) we present the critical function, the value(s) of the function at the critical point(s). This graph is not a manifold because of the sharp corners and self-intersections. The two lower pieces of the graph are the values of the function at the two minima. Where these pieces intersect, the minima are equally deep (Maxwell set, $a<0$, $b=0$, Sec. 3.1). The remaining piece of this graph, which looks like the seat of an Art Moderne chair, is the value of the function at the local maximum. The creases at which the values at the local maximum and minimum join have canonical power-law dependence familiar from the behavior of the Gibbs free energy of a function exhibiting a second-order phase transition (Gilmore, 1981). This canonical power-law dependence is that of the fold catastrophe, namely, 3/2.

The critical curvature, $f''(x_c(a,b);a,b)$, or curvature of the function at its critical point(s), is shown in Fig. 6(d). The curvature is positive at the local minima and negative on the intermediate local maximum. Although there are no creases as in Fig. 6(d), this graph is not a manifold because of the self-intersection. Over the bifurcation set the critical curvature vanishes because the second derivatives vanish and the tangent is "vertical."

1.6.3 Bifurcation Sets for the Three-Dimensional Catastrophes. The geometry of the fold and the cusp was relatively easy to visualize because their graphs could be embedded in low-dimensional spaces: $R^1 \times R^1$ for the fold and $R^1 \times R^2$ for the cusp. Higher-dimensional catastrophes are more difficult to visualize. The catastrophe A_4 should be viewed in $R^1 \times R^3$ while the catastrophes $D_{\pm 4}$ should be viewed in $R^2 \times R^3$. However, the bifurcation sets for these three catastrophes are relatively simple to visualize, since they are embedded in the control-parameter space R^3. These three catastrophes are

$$A_4:\quad f(x;a,b;c)=\tfrac{1}{5}x^5+\tfrac{1}{3}ax^3+\tfrac{1}{2}bx^2+cx,$$

$$D_{+4}:\quad f(x,y;a,b,c)=x^2y+\tfrac{1}{3}y^3+a(y^2-x^2)+bx+cy,$$

$$D_{-4}:\quad f(x,y;a,b,c)=x^2y-\tfrac{1}{3}y^3+a(y^2+x^2)+bx+cy. \tag{22}$$

In each instance the three-dimensional control-parameter space is partitioned into open regions by two-, one-, and zero-dimensional manifolds, components of the bifurcation set on which two, three, and all four of the critical points are degenerate. Within each open region the critical points are isolated; their number and type are unchanged by a sufficiently small perturbation. The number of critical points can change only when passing from one open region to another through the bifurcation set. The bifurcation sets for these three catastrophes are shown in Fig. 7. Shown in each figure is the number of critical points possessed by the catastrophe function in each of the open regions in its control parameter space. The catastrophes A_4, D_{+4} can each have zero, two, or four nondegenerate critical points while D_{-4} can have only two or four nondegenerate critical points.

1.7 Perturbations of Gradient Dynamical Systems

The qualitative properties of a gradient dynamical system are determined by the number, saddle type, and distribution of its critical points. If the critical points are isolated, then the dynamical system is structurally stable against perturbations. If one or more critical points are degenerate, the system is structurally unstable—a perturbation will produce a qualitative change in the properties of the system by splitting or annihilating the degenerate critical points.

As a result it is sufficient to use perturbation theory to describe the effect of a perturbation on a structurally stable system. In the case of a structurally unstable system it is useful to reduce the degenerate critical point to canonical form (a catastrophe) and then discuss the effect of a perturbation by using the catastrophe germ's universal perturbation.

To illustrate the effect of a perturbation in the structurally stable case, we consider a family of potentials, $V(x;c)$, depending on n state variables and k control parameters. As-

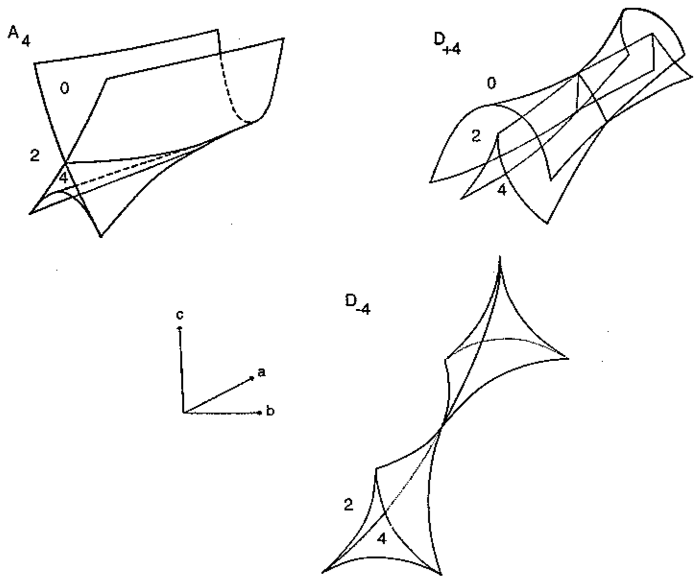

FIG. 7. Bifurcation sets for the three elementary catastrophes [Eq. (22)] of control dimension three as a function of the three control parameters *a,b,c*. The number of isolated critical points is shown in each of the open regions into which the control-parameter space is divided by the bifurcation set.

sume x_0 is a critical point for control-parameter value c_0. What happens under a perturbation $c_0 \to c' = c_0 + \delta c$? We expect that the critical point x_0 will move to a nearby point: $x_0 \to x' = x_0 + \delta x$. The displacement of the critical point as a function of the change in control-parameter value is determined by expanding $V(x_0+\delta x;c_0+\delta c)$ in a Taylor series about x_0, c_0:

$$V(x_0+\delta x;c_0+\delta c) = V(x_0;c_0) + V_i\delta x_i + V_\alpha \delta c_\alpha + \tfrac{1}{2}V_{ij}\delta x_i \delta x_j + V_{i\alpha}\delta x_i \delta c_\alpha + \tfrac{1}{2}V_{\alpha\beta}\delta c_\alpha \delta c_\beta + \text{higher-order terms.} \tag{23}$$

The coefficient $V_i(x_0;c_0)=0$, since x_0 is assumed to be a critical point for $c=c_0$. The value of δx is computed by solving $\nabla V(x_0+\delta x;c_0+\delta c)=0$. To lowest order (linear), we find

$$V_{ij}\delta x_j + V_{i\alpha}\delta c_\alpha = 0. \tag{24}$$

If the stability matrix V_{ij} is nonsingular,

$$\partial x_i/\partial c_\alpha = -(V^{-1})_{ij}V_{j\alpha}, \tag{25}$$

where (V^{-1}) is the matrix inverse of the nonsingular stability matrix: $(V^{-1})_{ij}V_{jk} = \delta_{ik}$. That is, a small change in control-parameter value produces a small change in the location of the critical point, as long as the stability matrix is nonsingular. The matrix (25) is the linear response function for the potential at the equilibrium x_0: it describes how much the equilibrium is displaced by a small change in the control parameters.

To second order the value of the potential at the displaced critical point is

$$V(x_0+\delta x;c_0+\delta c) = V(x_0;c_0) + V_\alpha \delta c_\alpha + \tfrac{1}{2}[V_{\alpha\beta} - V_{\alpha i}(V^{-1})_{ij}V_{j\beta}]\delta c_\alpha \delta c_\beta. \tag{26}$$

The stability matrix at the displaced critical point, $V_{ij}(x_0+\delta x;c_0+\delta c)$, is related to the stability matrix at the original critical point, $V_{ij}(x_0;c_0)$, by

$$V_{ij}(x_0+\delta x;c_0+\delta c) = V_{ij}(x_0;c_0) + P_{ij\alpha}\delta c_\alpha,$$

$$P_{ij\alpha} = V_{ij\alpha}(x_0;c_0) - V_{ijk}(V^{-1})_{kl}V_{l\alpha}. \tag{27}$$

As a result, for sufficiently small perturbations the Morse saddle type cannot change if the stability matrix is nonsingular.

This application of perturbation theory, and the analytic results constructed in Eqs. (25)–(27), evaporate when the stability matrix V_{ij} becomes singular. Under these conditions the evolution of the dynamical system under change in the control-parameter values ("perestroika") is computed by expressing the potential, $V(x;c)$, in the neighborhood of a non-Morse critical point by an appropriate catastrophe, following the perturbation through the well-defined bifurcation sets, and then resuming the perturbation treatment sufficiently far from the critical point degeneracy.

2. WHY IT EXISTS

The reduction of a family of functions to the sum of a non-Morse function and a Morse i-saddle at a non-Morse critical point, which is guaranteed by the Thom splitting lemma, can be carried out by the methods of elementary calculus. It will therefore not be reviewed here.

The reduction of a non-Morse function in the neighborhood of a degenerate (non-Morse) critical point cannot be carried out by the methods of elementary calculus. It will therefore be discussed in Sec. 2 of this article. This reduction to canonical form can be accomplished following a simple algorithm. The algorithm is illustrated in Sec. 2.1 in the context of a simple example: the reduction of a two-parameter family of functions of a single state variable to canonical form in the neighborhood of its most degenerate critical point. The steps involved in this algorithm are summarized in Sec. 2.2, and applied in Sec. 2.3 to an important example in which the canonical form is not at all evident.

The algorithm involves two main procedures. The first is the determination of the most degenerate critical point, and the catastrophe germ at this point. This is accomplished by finding critical points, using the control parameter degrees of freedom to eliminate the leading terms in the Taylor series expansion of the function in the neighborhood of that critical point, and then using a nonlinear change of coordinates to eliminate the higher-degree terms in the Taylor series expansion. What is left over, between the eliminated terms of low and high degree, is the catastrophe germ, CG(l), describing the degenerate critical point.

The second procedure in this algorithm is the determination of a universal perturbation. This follows the steps described above in more or less reverse order. First, an arbitrary function (perturbation) is added to the catastrophe germ. Then the high-degree terms are eliminated by a nonlinear change of coordinates. Finally, as many as possible of the low-degree terms are eliminated by a rigid translation of the coordinate system. The small number of terms remaining from the original general perturbation form the universal perturbation, Pert(l,k), of the catastrophe germ CG(l).

The mathematics used in these examples appears naively simple. However, this simplicity should not hide the fact that the mathematical approach used here is correct in spirit and can be (has been) made rigorous with sufficient effort (Poston and Stewart, 1978; Arnol'd, 1981; Gilmore, 1981).

2.1 A Simple Example

We carry out the reduction of a family of functions with a triply degenerate critical point to the canonical form of the cusp catastrophe $A_{\pm 3}$. We begin by assuming $f(x;c_1,c_2)$ is a family of functions depending smoothly on one state variable, x, and two control parameters, c_1 and c_2. We wish to determine the qualitative properties of any member of this family and, in particular, to determine how these qualitative properties change as the control parameter values change.

To facilitate this study, we expand $f(x;c_1,c_2)$ about some point x_0:

$$f(x;c_1,c_2)=\sum_{n=0}^{\infty} f_n(x_0;c_1,c_2)(x-x_0)^n. \qquad (28)$$

The Taylor series expansion of this function is shown in line 1 of Fig. 8. The remainder of the discussion of this example will refer to subsequent lines of this figure. At points x_0 where $f_1 \neq 0$ the implicit function theorem is applicable and nothing qualitatively interesting will happen. We therefore search over the range of the state variables to find a critical point. At that critical point $f_1=0$. Since we are interested in the qualitative properties (shape) of the potential, we readjust our y axis so the function vanishes at the critical point. As a result of using these two degrees of freedom—choice of origin in both ordinate and coordinate—the first two terms in the Taylor series expansion of f vanish. The expansion is shown in line 2 of Fig. 8.

There remain the two degrees of freedom represented by the two control-parameter values c_1, c_2. It might be expected that we can use these two degrees of freedom to annihilate one, or at most two, of the remaining coefficients in the Taylor series expansion, but that

Object	Line / Procedure	Coefficients of x_n x^0	x^1	x^2	x^3	x^4	x^5	x^6
Find Canonical Germ	1. Taylor Expansion	f_0	f_1	f_2	f_3	f_4	f_5	f_6
	2. Adjust Ordinate; Locate Critical Point	0	0	f_2	f_3	f_4	f_5	f_6
	3. Exploit Control Parameter Degrees of Freedom			0	0	f_4	f_5	f_6
	4. Smooth Change of Coordinates					±1	0	0
Find Canonical Perturbation	5. Add Arbitrary Perturbation	ε_0	ε_1	ε_2	ε_3	$\pm1+\varepsilon_4$	ε_5	ε_6
	6. Smooth Change of Coordinates	ε_0	ε_1	ε_2	ε_3	±1	0	0
	7. Shift Origin	0	ε_1	ε_2	0	±1		

$$f(x{:}C_1,c_2) = \underbrace{\pm x^4}_{CG(1)} + \underbrace{\varepsilon_1 x^1 + \varepsilon_2 x^2}_{Pert(1,2)}$$

FIG. 8. The steps used to reduce a two-parameter family of functions to its canonical form at the most degenerate possible critical point are illustrated in lines 1–4. The steps used to determine the universal perturbation are shown in lines 5–7.

we cannot generically (some deep mathematics is required at this point: cf. Poston and Stewart, 1978; Arnol'd, 1981, 1986; Gilmore, 1981) annihilate more than two coefficients. If coefficients of some high powers (f_5 of x^5, f_7 of x^7) are annihilated, then the coefficient f_2 of x^2 remains nonzero, the Morse lemma is applicable, and nothing qualitatively interesting will happen. The most interesting things happen when the control parameter degrees of freedom are used to annihilate the leading nonzero terms in the Taylor series expansion. In the present case the following two interesting possibilities arise:

(i) $f_2=0$, $f_3\neq0$,

(ii) $f_2=0$, $f_3=0$.

The second case is more interesting than the first, which can already be encountered in functions depending on one state variable and a single control parameter.

We assume therefore that the two control-parameter degrees of freedom can be used to annihilate the two leading nonzero terms in the Taylor series expansion given on line 2 of Fig. 8. The resulting Taylor series expansion shown on line 3 of Fig. 8 begins with a term of degree 4.

This is as far as we can go using the simple "linear" degrees of freedom represented by choice of origin in the ordinate and coordinate and the two control-parameter degrees of freedom. However, we have not yet exploited any "nonlinear" degrees of freedom. That is, there is the possibility to perform a smooth nonlinear change of coordinates to remove some of the higher-degree terms in the Taylor series expansion. To do this we seek a nonlinear change of variables,

$$x'=A_1x+A_2x^2+\cdots, \qquad (29)$$

which will eliminate as many of the higher-degree terms as possible from the remaining Taylor series expansion (line 3, Fig. 8). It is possible to find a nonlinear change of variables that eliminates the Taylor tail, the tail of the Taylor series expansion (the term x^4 is determinate; more powerful mathematics is needed at this point also, cf. Poston and Stewart, 1978; Arnol'd, 1981, 1986; Gilmore, 1981). The first two coefficients in the nonlinear change of variables above are $A_1=|f_4|^{1/4}$, $A_2=\pm f_5/4|f_4|^{3/4}$. The expansion is analytic and the sum converges locally. In the new coordinate system the transformed function is $\pm x^4$. The Taylor series after this nonlinear change of variables is shown on line 4 of Fig. 8.

This procedure produces the catastrophe germ of the two-parameter family of functions $f(x;c_1,c_2)$:

$$f(x;c_1,c_2)=\pm x^4 \quad \text{(catastrophe germ)}. \tag{30}$$

The second step in the process of reducing a non-Morse function to canonical form is computation of the universal perturbation of minimum dimension. To do this, we repeat the procedure described above in almost the reverse order. We begin by adding an arbitrary perturbation

$$\epsilon(x)=\sum_{i=0}^{\infty} \epsilon_i x^i \tag{31}$$

to the catastrophe germ. The Taylor series expansion of the perturbed function is presented on line 5 of Fig. 8.

A nonlinear change of variables can once again be called upon to eliminate the Taylor tail. The nonlinear transformation (29) fails to converge unless the base term from which the elimination is made has a large coefficient. Thus, trying to eliminate the Taylor tail of terms above x, x^2, or x^3 will fail because the coefficients of these terms, ϵ_1, ϵ_2, ϵ_3, are small and not bounded away from zero. However, since the coefficient $(\pm 1+\epsilon_4)$ of x^4 is large and bounded away from zero when ϵ_4 is small, all terms above x^4 may be eliminated by a nonlinear change of coordinates. To state this in another way: Any terms in the Taylor tail that may be eliminated in the initial construction of the catastrophe germ (line 3 to line 4) can be eliminated from the perturbation (line 5 to line 6). The result of this nonlinear change of variables is shown in line 6 of Fig. 8.

As a final step, we can choose a new origin of ordinates (rigid vertical shift) to eliminate ϵ_0 and of coordinates (rigid horizontal shift) to eliminate one of the three remaining terms. Since the linear and quadratic terms cannot always be eliminated but the cubic term can be, we choose to eliminate this term. The final result,

$$f(x;c_1,c_2)=\pm x^4+\epsilon_1 x^1+\epsilon_2 x^2, \tag{32}$$

is shown in line 7 of Fig. 8.

We summarize briefly the steps taken to answer the two equations implicit in the statement of the classification theorem presented at the beginning of this section:

1. How to construct the catastrophe germs:
   (a) Locate a critical point and move the origin of coordinates to that critical point.
   (b) Eliminate the leading Taylor series coefficients by using the control parameters.
   (c) Remove the Taylor tail by a smooth nonlinear change of variables.
2. How to construct the universal perturbation:
   (d) Add an arbitrary perturbation.
   (e) Remove the Taylor tail by the same smooth nonlinear change of variables.
   (f) Move the origin of coordinates to eliminate unnecessary terms.

2.2 General Procedure

The general procedure for reducing a non-Morse function to the sum of a catastrophe germ and a universal perturbation has been illustrated by example in the previous section. In some respect, the procedure followed in the reduction follows in spirit the procedure used in the study of electrodynamics. We often find that the near-field regime (low-degree Taylor series terms) is tractable, as is the far-field regime (Taylor tail). The really interesting part is between these two regimes (intermediate-field regime ~catastrophe germ).

We summarize the general procedure below and refer to Fig. 9. The following steps are taken to compute the catastrophe germ of a function $f(x;c)$ with $x\in R^n$ and $c\in R^k$:

1. Computation of the catastrophe germ CG(l):
   (a) The function is expanded in a Taylor series. The origin of ordinates and coordinates is rigidly (linearly) shifted to a critical point. This eliminates the zeroth- and first-degree terms in the Taylor series expansion.
   (b) The control parameter degrees of freedom are used to eliminate as many of the remaining low-degree Taylor series coefficients as possible. No more than k coefficients can typically be annihi-

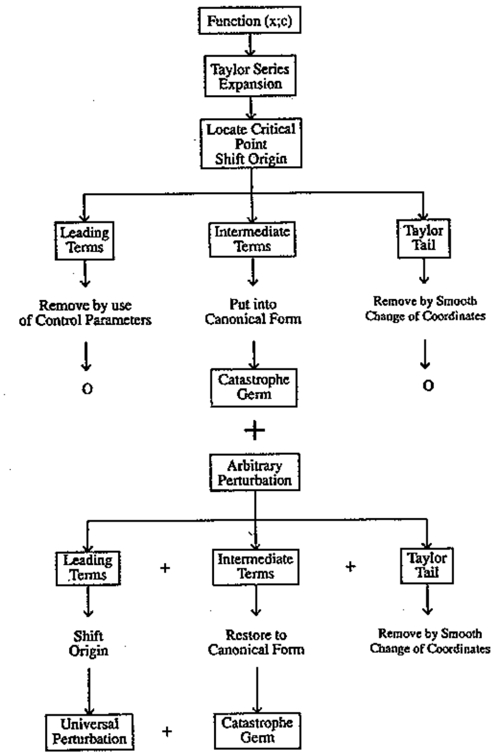

FIG. 9. General procedure for computing the catastrophe germ and the universal perturbation of a family of functions.

lated (by transversality: cf. Poston and Stewart, 1978; Arnol'd, 1981, 1986; Gilmore, 1981).

(c) A smooth nonlinear change of variables is introduced to eliminate as many terms in the Taylor tail as possible. Whether all terms above some finite degree, or only most terms, can be eliminated is determined algorithmically (determinacy algorithm: cf. Poston and Stewart, 1978; Arnol'd, 1981, 1986; Gilmore, 1981).

After the catastrophe germ has been computed, the steps followed above are taken in reverse order to compute the universal perturbation of the catastrophe germ:

2. Computation of the universal perturbation Pert(l,k):
   (d) An arbitrary perturbation is added to the germ. The smooth transformation used to eliminate the Taylor tail in step (c) above is used again to eliminate the same terms of the perturbed germ and reduce the coefficients in the germ to the same canonical values.
   (e) A shift of ordinate and coordinate is introduced to eliminate the constant term in the perturbation and l (state space dimension of the germ) of the remaining terms of the perturbation.

This procedure reduces a family of functions to canonical form in the neighborhood of a degenerate critical point.

2.3 A More Complicated Example

In this closing section of Sec. 2 we tackle a more complicated example. This is done for two reasons:

1. to illustrate the methodology described in the previous section for a multidimensional state-variable space.
2. to illustrate exactly why the classification of elementary catastrophes terminates at control dimension $k=5$.

The particular example we consider is the reduction to standard form of the six-control-parameter family of functions defined on a three-dimensional state-variable space,

$$f(x;c)=f(x,y,z;c_1,c_2,c_3,c_4,c_5,c_6). \tag{33}$$

To effect this reduction we follow the steps described in the previous two sections. The computation of the catastrophe germ is illustrated in Fig. 10(a). The computation of the universal perturbation is shown in Fig. 10(b). The steps 1–5 used to compute the germ below are indicated by the corresponding numbers over the arrows, which show the effect of these steps, in Fig. 10(a). Steps 6–9, used to compute the universal perturbation, are similarly keyed to the arrows in Fig. 10(b).

1. The Taylor series expansion for this function is

$$f(x;c)=\sum_{\substack{i\geq 0\\ j\geq 0\\ k\geq 0}} \frac{i!j!k!}{(i+j+k)!} f_{ijk}(x-x_0)^i \times(y-y_0)^j(z-z_0)^k. \tag{34}$$

   This Taylor series is illustrated schematically in Fig. 10(a), where we have grouped together all terms of the same degree and suppressed factorials.
2. We search over state-variable space for a

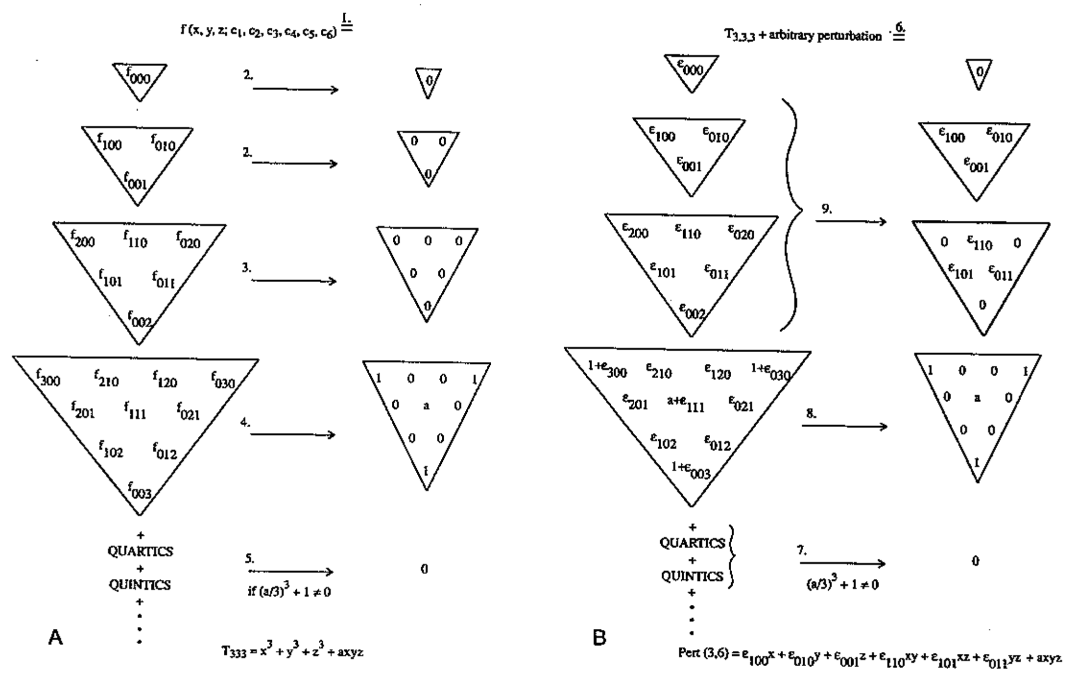

FIG. 10. The procedure used to compute the catastrophe germ and universal perturbation for a six-parameter family of functions depending on three state variables. The numbers over the arrows are keyed to the steps taken in the text.

critical point, then move the origin of coordinates to that point. This removes the constant and linear terms from the Taylor series expansion.

3. The six control-parameter degrees of freedom can be used to eliminate up to six of the remaining terms in the Taylor series expansion. If any of the quadratic terms remain, the catastrophe ultimately constructed appears in the list given in Table 2. However, if we use these six control-parameter degrees of freedom to eliminate all six second-degree terms from the Taylor series expansion, something new appears. Accordingly, we make this choice. The resulting expansion begins with the third-degree terms.
4. Using a homogeneous linear transformation we can attempt to put the cubic terms into some canonical form. Since a linear transformation on the state-variable space is a 3×3 matrix with 3^2 degrees of freedom and there are ten cubic coefficients, it is possible to give canonical values to only nine of these ten coefficients. Symmetry dictates the choice

$$\sum_{i+j+k\geqslant 3} f_{ijk}x^iy^jz^k = x^3+y^3+z^3+axyz \quad (+\text{ higher-degree terms}). \tag{35}$$

This linear transformation maps nth-degree terms ($n\geqslant 3$) into terms of the same degree. The parameter, a, which appears in this function cannot be given a canonical value, and is the reason why this catastrophe is not elementary.

5. We then search for a smooth nonlinear change of variables which will eliminate as much of the tail of the Taylor series expansion as possible. If

$$(a/3)^3+1\neq 0, \tag{36}$$

the cubic terms are determinate. This means that a smooth transformation exists which transforms away all terms in the Taylor series expansion of degree in excess of some finite value, in this case three. The resulting catastrophe germ is

$$f(x;c)=T_{3,3,3}=x^3+y^3+z^3+axyz. \tag{37}$$

This catastrophe germ is shown at the bottom of Fig. 10(a).

6. To compute the universal perturbation we add a general function of three state variables,

$$\epsilon(x,y,z) = \sum_{\substack{i\geq 0\\ j\geq 0\\ k\geq 0}} \epsilon_{ijk} x^i y^j z^k, \qquad (38)$$

to the catastrophe germ $T_{3,3,3}$. All coefficients ϵ_{ijk} are assumed to be small. The Taylor series of the resulting perturbed function is shown in Fig. 10(b).

7. Working backwards, we find a linear transformation which transforms the cubic terms to the form (37).

8. If the original function $T_{3,3,3}$ is determinate $[(a/3)^3+1\neq 0]$, the perturbed function must also be determinate $[(a'/3)^3+1\neq 0,\ a'=a+\epsilon_{111}]$ for a sufficiently small perturbation. This means that a smooth transformation exists which transforms away all terms in the Taylor series expansion of the perturbed germ of degree greater than 3.

9. Finally, we make a rigid displacement of coordinate system to eliminate the constant term ϵ_{000} in the perturbation and three of the remaining nine linear and quadratic terms. Since it is always possible to eliminate the three terms x^2, y^2, z^2, the universal perturbation is

$$\begin{aligned}\mathrm{Pert}(l=3,k=6) &= \epsilon_{100}x+\epsilon_{010}y+\epsilon_{001}z\\ &\quad +\epsilon_{110}xy+\epsilon_{101}xz\\ &\quad +\epsilon_{011}yz+axyz. \end{aligned} \qquad (39)$$

The seventh term, $axyz$, has been added because it is not possible to reduce the coefficient of xyz in the catastrophe germ $T_{3,3,3}$ to a standard value. The universal perturbation is shown at the bottom of Fig. 10(b).

The net result of this calculation is a reduction of the function $f(x;c)$ given in Eq. (33) to canonical form:

$$\begin{aligned} f(x;c) &= T_{3,3,3} + \mathrm{Pert}(3,6)\\ (33) &= (37) + (39) \end{aligned} \qquad (40)$$

We emphasize once more that for $k<6$ control parameters it is possible to assure that any catastrophes that are encountered are elementary, but for $k=6$ (and $k>6$) this is no longer possible, as $T_{3,3,3}$ shows.

3. HOW IT WORKS

Elementary catastrophe theory consists of a collection of theorems about the canonical forms of functions in the neighborhood of degenerate critical points. Catastrophes in the real world consist of processes that exhibit discontinuities and sudden jumps. It is clear that the mathematics provides the right language for describing the physical processes. But how? The mathematics describes potentials that depend on control parameters; processes depend on time. It is tempting to allow control parameters to depend on time, but then the catastrophe functions are no longer strictly potentials. And what of the dynamics describing a physical system? If many equilibria are available, which does the system choose? Finally, how can we recognize when a physical process can be modeled by a mathematical catastrophe function?

These practical questions are the subject of Sec. 3 of this article. The difficulty of relating the dynamics of a system to a potential that describes it is discussed in Sec. 3.1. In most instances the system state is governed either by a local or the global minimum of a potential, and sudden jumps from one state to another occur when the control parameters pass through an appropriate component of the bifurcation set. Although the two conventions presented (Maxwell, Delay) are not complete, they serve well in a preponderance of situations.

When a catastrophe occurs in some physical process, a multiplicity of related phenomena occur. These phenomena are called catastrophe flags. They are treated in Sec. 3.2. The identification of any is a clear indication that the others are present and will be recognized if sought. These flags may be used to identify the mathematical catastrophe function that most accurately describes the discontinuous physical process.

In Sec. 3.3 we illustrate the use of catastrophe theory in determining the sensitivity of failure modes of a complex structure to hidden or unforeseen imperfections. This is done by investigating the catastrophe germ for an optimized system and then determining the universal perturbation for that germ. In the particular example considered, that of a propped cantilever, strong coupling of two "soft" collapse modes yields a "soft" collapse direction and a "hard" direction, one in which a severe and unexpected collapse is possible. Further, analysis of the universal perturbation for that collapse mode reveals that the mode is extremely sensitive to imperfections in either the design or fabrication stages of construction. Finally, the use of catastrophe theory to suggest tests to locate the actual (as opposed to designed) failure load, as well as ways to increase the load at which failure occurs, are suggested.

Since elementary catastrophe theory is the first step in the program of catastrophe theory, all its results and phenomenology will necessarily be present in subsequent developments in this program. The next stage in this program following elementary catastrophe theory is the study of dynamical systems, sets of coupled first-order ordinary differential equations. We show in Sec. 3.4 how the fold and the cusp catastrophe appear in dynamical systems theory under the guise of three elementary bifurcations: the saddle-node bifurcation (A_2), the pitchfork bifurcation (A_3), and the Hopf bifurcation (A_3).

3.1 Catastrophe Conventions

Elementary catastrophe theory is the first tentative step in the program of catastrophe theory. It was motivated above by our desire to describe the qualitative properties of gradient dynamical systems of the form (4). The qualitative dynamics of such systems are determined by a potential function. The local flow in the neighborhood of each nondegenerate critical point (Morse i-saddle) is canonical and the separate flows can be pasted together to construct a global flow.

As the study of elementary catastrophe theory evolved, it became clear that the key to our understanding of non-Morse critical points was the determination of how the critical points moved about, coalesced and disappeared, or were born degenerate and moved apart as the control-parameters changed. The dependence of critical points on control-parameter values, as described by the canonical decomposition (15), provides a complete resolution to the mathematical question posed by elementary catastrophe theory. However, it does not address a corresponding physical question. That is, how does the dynamics depend on the control parameters? Once the control parameters are allowed to vary in time, the system (4) is no longer a gradient system—the potential is time dependent. As a result, knowledge of the shape of the potential is not sufficient to determine the state of the system unless the equations of motion are known.

In the absence of a detailed set of equations of motion, the state of the physical system described by an elementary catastrophe must be determined by some infusion of intuition. This is done by adopting one of two standard assumptions. These are the Delay convention and the Maxwell convention (Zeeman, 1977; Poston and Stewart, 1978; Gilmore, 1981).

Delay Convention: The state of a system is determined by a (local) minimum of a potential. As the control parameters change, the state remains at the local minimum as long as that minimum exists. When the local minimum ceases to exist (at a bifurcation set) the system state jumps to a new local minimum.

Maxwell Convention: The state of a system is determined by the global minimum of a potential. As the control parameters change, the state remains at the minimum as long as that minimum remains the global minimum of the potential. When the minimum ceases to be the global minimum, the system state jumps to the new global minimum.

We stress again that these assumptions about the behavior of a physical system are required only by our lack of knowledge of the system's equations of motion.

The difference between these two assumptions is illustrated in Fig. 11 for the cusp catastrophe. When $a<0$ the potential may have one minimum or two. We assume the system begins in the unique minimum of the potential. As the control parameter b is changed, a new minimum is created in a fold catastrophe. The original minimum eventually becomes metastable with respect to the new minimum, and finally the original minimum disappears in a fold catastrophe. According to the Delay convention, the system jumps to the new minimum only when the original minimum is annihilated. According to the Maxwell convention, the jump occurs when the two minima become equally deep (at $b=0$). These two conventions are illustrated on the canonical critical-value surface for the cusp in Fig. 11(b). The jumps occur at the parts of the critical-value graph

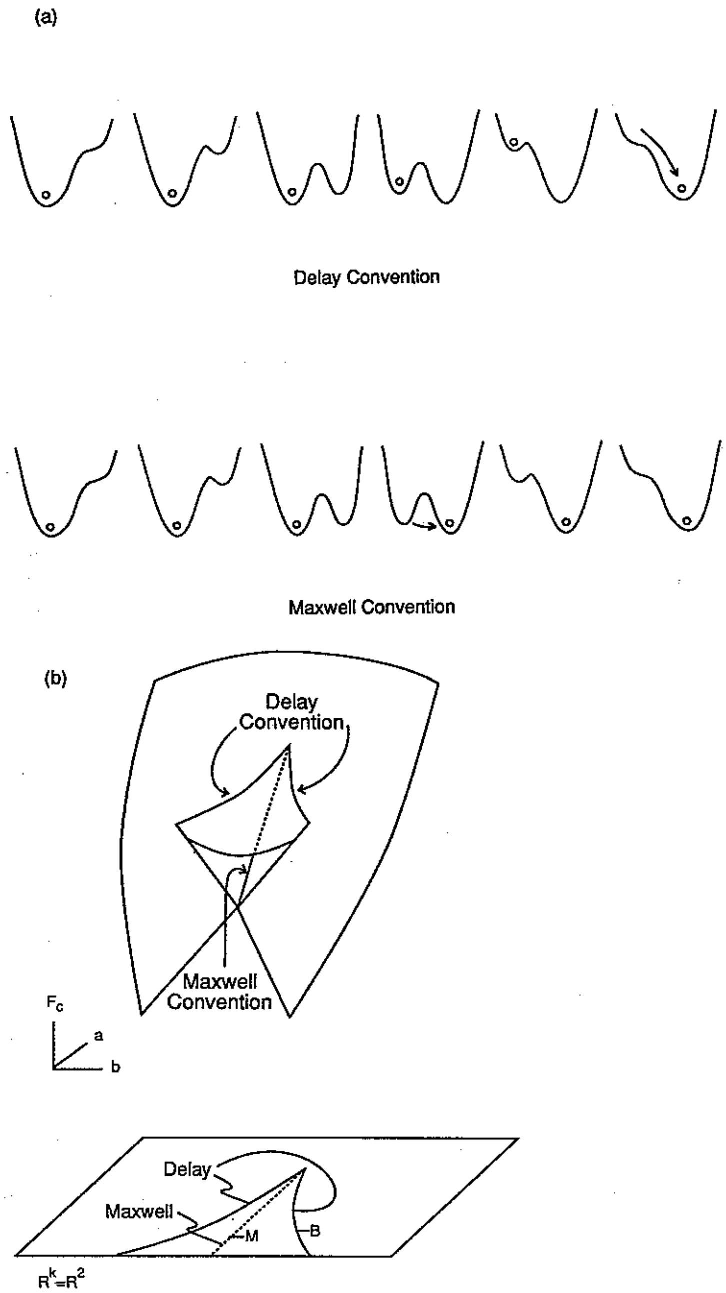

FIG. 11. (a) The Delay convention and the Maxwell convention determine when the jump from one local minimum to another takes place as the control parameters are varied. (b) The two bifurcation sets correspond to pieces of the critical surface which prevent the surface from being a manifold. Under the Delay convention, jumps occur at the creases in the surface, where a local minimum and maximum become degenerate. Under the Maxwell convention, jumps occur at the intersection of two of the leaves of this surface, where the two minima become equally deep.

$$f=f(x_c(a,b);a,b),$$

which prevent the surface from being a manifold. The Delay convention requires jumps to occur at the creases which represent the collision of a local minimum and maximum and which lie over the cusp-shaped curve (21). The Maxwell convention requires jumps to occur at the self-intersection which represents the surface on which the two minima are equally deep and which lies over the half-line $b=0$, $a<0$.

The bifurcation sets for the Delay and Maxwell conventions are sets in the control-parameter space at which jumps to new states take place. These sets are determined by local and global considerations for the two conventions. The bifurcation set for the Delay convention consists of that subset of the bifurcation set for the catastrophe which involves a Morse 0-saddle. For catastrophes more complicated than $A_{\pm 3}$, some components of the bifurcation set involve Morse saddles M_i^n and M_{i+1}^n, $i > 0$. On these components of the bifurcation set the two saddles M_i^n and M_{i+1}^n are degenerate. These components of the catastrophe's bifurcation set are not components of the bifurcation set determined by the Delay convention, since local minima (M_0^n) must be present. The bifurcation set for the Maxwell convention is determined by nonlocal arguments. Once a point on the Maxwell set has been located, the remainder of the set may be constructed by integrating a set of equations of Clausius–Clapeyron type (Gilmore, 1981).

The appropriate choice of convention must balance the rate at which the control-parameters are varied against the noise level of the system—specifically, the probability that a fluctuation in energy will occur in excess of the barrier height separating a local minimum from other minima, including the global minimum. When the noise level is low, it is usually safe to adopt the Delay convention. When large fluctuations occur more rapidly than the control parameters are swept, the Maxwell convention is more suitable.

From Fig. 11 it can be observed that the bifurcation set for the Maxwell convention "lies inside" the bifurcation set for the Delay convention. This is a general feature of catastrophes and has the following physical interpretation. If a system whose phase transition properties are usually described by the Maxwell convention is handled very gently, the jump from a metastable state to a global minimum may not occur at the bifurcation set for the Maxwell convention. The transition may be delayed—but it cannot be delayed beyond the point at which the local minimum ceases to exist—the bifurcation set for the Delay convention. The outer limits beyond which metastability cannot occur for a physical system are called spinodal curves (or surfaces). The spinodal curves surrounding the bifurcation set for the Maxwell convention consist of the bifurcation set for the Delay convention. The spinodal curves for the symmetry-restricted catastrophe

$$A_{+5}\text{: } f(x;a,c)=\tfrac{1}{6}x^6+\tfrac{1}{4}ax^4+\tfrac{1}{2}cx^2 \qquad (41)$$

are shown in Fig. 12.

When the equations of motion for the system are known explicitly, there is no need to make an assumption (i.e., adopt a convention) about the system's behavior. For example, the system state may be governed by a probability distribution, $P(x,t)$, which satisfies a Fokker–Planck equation

$$\partial P/\partial t=\nabla(P\nabla V)+\nabla^2(DP), \qquad (42)$$

where V has a time dependence due to the time dependence of some control parameters:

$$V(x;t)=V(x;c(t)) \qquad (43)$$

and D is a diffusion constant. Then the probability distribution for the system state can be determined explicitly as a function of time. In this case there is no well-defined bifurcation set—there is an occupation probability for the two minima (and surrounding regions) and that probability changes smoothly in time.

It has been shown that the Fokker–Planck equation can be used to interpolate between the two extreme limits represented by the Delay convention and the Maxwell convention (Gilmore, 1981).

We remark, finally, that the problem of system state is not necessarily resolved by adopting the Delay convention, as appears to be the case in Fig. 11. In that figure there is one remaining minimum when the metastable minimum disappears in a fold catastrophe, so the system has a unique minimum to jump to. But what happens when there are two or more remaining minima that the system can jump to? Does the system jump to the deepest minimum, the nearest minimum, or yet another minimum? On this the Delay

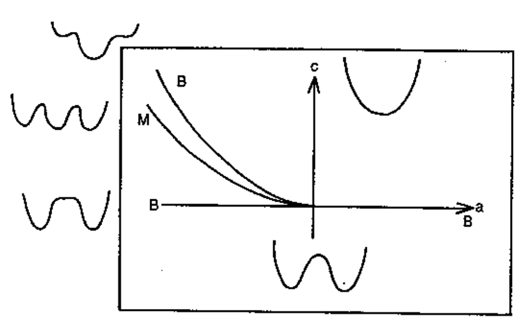

FIG. 12. Shapes of the symmetry-restricted butterfly catastrophe $f(x;a,c)=\frac{1}{6}x^6+\frac{1}{4}ax^4+\frac{1}{2}cx^2$, together with bifurcation sets under the Delay and Maxwell conventions. Spinodal curve for the Maxwell set is the bifurcation set for the Delay convention.

convention has nothing to say. In short, the conventions are incomplete. They can be very useful but they are not substitutes for detailed knowledge of the equations of motion.

3.2 Catastrophe Flags

An important reason for studying the elementary catastrophes is that their properties are canonical. Therefore, predictions based on the properties of an elementary catastrophe in a physical system are canonical. This means that their properties need be studied only once. They can then be applied directly, without change, to any physical system described by that catastrophe. As a result, it is important to be able to determine the presence and type of a catastrophe underlying the behavior of a physical system.

This may be done either deductively or inductively. In the first procedure there is full knowledge of the equations of motion describing the system. When this is the case it is generally possible to make whatever predictions are desired by computing solutions of these equations. It is then possible to predict the behavior of the system without any knowledge of the elementary catastrophes. This, in fact, has been a long-standing operational mode in the physical sciences.

There are, however, instances in which the equations of motion are not known. Then the presence of a catastrophe cannot be deduced (from the equations). It may, however, be induced from the behavior of the system. When a catastrophe is present, the system undergoes (or may undergo) a qualitative change in its behavior. This qualitative change is often accompanied by a number of other phenomena. For any particular system these may be more or less difficult to recognize. The point is that the observation of any will imply the presence of all and that if we look hard enough we will find all these correlated phenomena.

Eight phenomena typically occur when a catastrophe is present. These are called catastrophe flags: in a sense the catastrophe "waves flags" to attract our attention. Of the eight, five are classical: they occur when there is a qualitative change in the system. The remaining three also occur when there is a qualitative change, but they may also be observed before a phase change occurs. This is very important when the phase change represents the transition of a structure from one state to another (i.e., collapse of a bridge). The five classical catastrophe flags are modality, sudden jumps, inaccessibility, sensitivity, and hysteresis (Zeeman, 1977; Poston and Stewart, 1978). The three diagnostic catastrophe flags are divergence of linear response, time dilation, and anomalous variance (Gilmore, 1981). These are now discussed and illustrated in Fig. 13 for the cusp catastrophe.

3.2.1 Modality. In the neighborhood of a catastrophe the system can exhibit two or

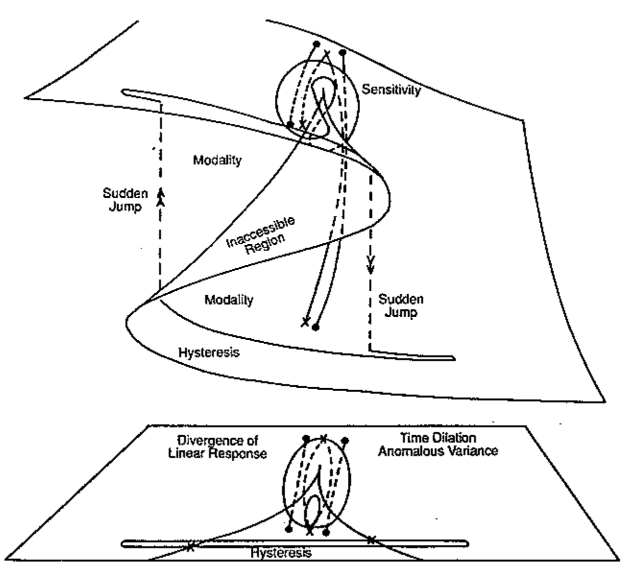

FIG. 13. Eight phenomena (flags) associated with the presence of a catastrophe.

more distinct types of behavior (near a critical point: liquid, gas; near a triple point: solid, liquid, gas). The upper and lower surfaces of the cusp catastrophe manifold represent two distinct types of locally stable behavior in Fig. 13. The intermediate part of this manifold represents an unstable mode of behavior which nevertheless may be observed with sufficient care.

3.2.2 Sudden Jumps. The system may suddenly jump from one mode of behavior to another mode of behavior (liquid, high density, to gas, low density) as the control parameters (intensive thermodynamic variables: temperature and pressure) are varied. These jumps represent the transition from one local minimum to a global or another local minimum. The location of the sudden jumps may or may not be adequately described by one of the two conventions presented in Sec. 3.1.

3.2.3 Inaccessibility. Separating the two or more local minima (M_0^n) responsible for "modalities" is at least one Morse saddle (M_1^n). The Morse saddles M_i^n ($i > 0$) are all dynamically unstable. In the neighborhood of such a saddle the system, if perturbed, will move quickly to a locally stable state. As a result, modality will always be accompanied by inaccessible regimes as represented by the middle sheet of the cusp catastrophe manifold.

3.2.4 Sensitivity. If the system is brought from the single-mode regime to the multimode regime the question arises: what is the final state of the system? For most initial conditions and processes, the final state is robust against perturbations. That is, if the final state lies on the upper sheet of the cusp catastrophe manifold, then a perturbation of either initial conditions and/or processes will leave the final state on the upper sheet. However, there is a small set of initial conditions and/or processes for which this is not true. This sensitivity to (some) initial conditions is shown in Fig. 13. For the process $db/dt=0$, $da/dt<0$ the initial condition with $b=\epsilon>0$ will result in a final state on the lower sheet while the nearby initial condition $b=-\epsilon<0$ lies on the upper sheet. A similar phenomenon occurs for processes that are almost identical, beginning from appropriate initial conditions. We remark that for the elementary catastrophes there is sensitive dependence on *some* initial conditions while in the study of chaos there is sensitive dependence on *all* initial conditions.

3.2.5 Hysteresis. A sudden jump from one modality to another may occur at the bifurcation set for the Delay or Maxwell convention, or anywhere in between. If the process is reversed, the jump back to the original modality may occur at the same control-parameter value as the initial jump (Maxwell convention) or at different control parameter values (any other convention). Unless the Maxwell convention is observed, jumps between the different modalities will exhibit hysteresis. Hysteresis, while not necessarily a part of the catastrophe scenario, can usually be observed with sufficient care even when the Maxwell convention suggests itself (e.g., superheating and supercooling in a liquid–gas phase transition).

The five classical catastrophe flags just described are of little use to the structural engineer looking for the limits of stability of a large structure (bridge, building) or a naval engineer testing a ship's stability. In many instances, transition from one modality to another is not only a catastrophe, but a disaster as well (collapsed bridge, building; capsized ship). For this practical reason it is essential to have a set of catastrophe flags that identify the presence of a catastrophe, and can be used to map out approximately the bounds of the catastrophe (i.e., bifurcation sets) without actually having to enter the multimode regime, or at least make the transition between modes. The three diagnostic catastrophe flags accomplish just this.

3.2.6 Divergence of Linear Response. The response, δx_i, of an equilibrium in state-variable space to a small change in the control parameters, δc_α, is given by

$$\delta x_i=\chi_{i\alpha}\delta c_\alpha, \tag{44}$$

where the linear susceptibility tensor $\chi_{i\alpha}$ is given explicitly by Eq. (25) for a system governed by a potential. As a local bifurcation set is approached, at least one of the eigendirections has a very large response. This is the direction in which two or more critical points are approaching each other. This direction is an eigenvector of the stability matrix, V_{ij}, in Eq. (25). The eigendirection, and the dependence of the diverging eigenvalue (λ_i^{-1}) on the control parameters c_α, can be used to map out the bifurcation set without actually penetrating it (Gilmore, 1981). In the liquid–gas phase transition divergence of linear response manifests itself as the divergence of the compressibility as the critical point is approached from either the single-phase or the multiphase region. Divergence of the compressibility may also be seen, with greater difficulty, as the spinodal lines are approached from within the multiphase region.

3.2.7 Time Dilation (Critical Slowing Down, Mode Softening). In the neighborhood of a nondegenerate critical point the equations of motion for a dissipative gradient system are

$$dx_i/dt = -V_{ij}x_j. \tag{45}$$

For a dissipative system the relaxation to equilibrium following a perturbation occurs on a time scale that is the reciprocal of the smallest eigenvalue of the stability matrix V_{ij}. For a conservative system the spectrum of oscillation frequencies, ω_i, consists of the square roots of the eigenvalues of V_{ij}. Upon approach to a bifurcation set one or more of the eigenvalues of V_{ij} approaches zero. Relaxation to equilibrium slows down (critical slowing down) as the bifurcation set is approached. The dependence of the vanishing eigenvalue on the control parameters can be used to determine where in control-parameter space the bifurcation set occurs. The associated eigenvector can be used to locate where in state space the critical points become degenerate. A similar analysis is possible when one (or more) of the frequencies of a conservative system approaches zero (mode softening). Time dilation is well known to experimentalists. The time required for equilibration near the critical point in the liquid–gas phase transition grows very rapidly as the critical point is approached, making experiments on the equilibrium properties of fluids near their critical points very difficult and time consuming.

3.2.8 Anomalous Variance. When the state of a physical system is described by a Morse critical point, motion about the equilibrium is confined by a quadratic potential well. On approach of another critical point the quadratic potential becomes flattened and motion about the equilibrium becomes less confined. This reduction in localization, or deconfinement, is called anomalous variance. In a gas–liquid system anomalous variance is dramatically observed at the critical point as critical opalescence. This occurs when fluctuations in the size of liquid droplets in the gas phase and gas bubbles in the liquid phase can occur at all length scales, including length scales comparable to optical radiation, producing an anomalously strong scattering of light.

Although catastrophe theory provides a qualitatively correct description of physical properties of a fluid near its critical point, it does not provide a quantitatively correct description of these properties and must be replaced by a better descriptive mechanism for the purposes of making accurate predictions of physical behavior.

3.3 The Dangers of Design Optimization

Large structures are built from many smaller parts. Cost is almost always a major factor in the design and construction of a complex structure, such as a bridge. As a result, a design philosophy has been widely adopted. This is the philosophy of design optimization. The basic idea is to design each component to meet but not exceed its specifications. For example, if an elevated roadway designed to support a load L is to be supported by ten cantilevers, each cantilever should be designed to support some fraction of that load. It makes no sense to design some cantilevers to be stronger (and more expensive) than others: what good will they do when the others have already collapsed, taking the roadway with them?

This is the philosophy of design optimization. It has hidden dangers because complex structures may have hidden or unexpected collapse modes. Further, these modes may have exceptionally sensitive dependence to imperfections in the fabrication/construction stages. We illustrate these ideas and dangers by treating a simple example. This is a propped cantilever designed to support a load F up to some critical load, F_c.

The cantilever is shown in Fig. 14. We assume it has unit length and a force F is applied vertically at the top of the cantilever. The mass of the beam is assumed small compared to the loading force F.

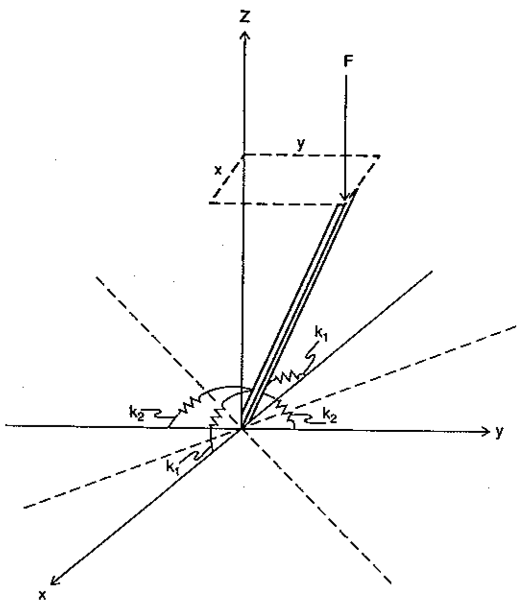

FIG. 14. The cantilever consists of a vertically propped beam supported by springs of spring constant k_1 providing restoring forces in the x–z plane and springs of spring constant k_2 providing restoring forces in the y–z plane. The length of this beam is scaled to 1. Displacements from the vertical are described by the (x,y) coordinates of the top of the beam. Solid axes (x,y) are axes along which a "soft" A_{+3} catastrophe occurs. Dashed axes ($x=\pm y$) are axes along which a "hard" A_{-3} catastrophe occurs.

We assume the collapse modes of this cantilever involve deflection in the x-z plane through an angle θ_1 from the vertical, deflection in the y-z plane through an angle θ_2 from the vertical, or any combination of these motions. We also assume that deflection is counteracted by springs with spring constants k_1, k_2 providing restoring forces in the x-z and y-z planes, respectively. The potential describing this cantilever is

$$V(\theta_1,\theta_2;F)=(2/2)k_1\theta_1^2+(2/2)k_2\theta_2^2+Fz, \quad (46)$$

where z is the height of the top of the cantilever.

It is useful to express this potential as a function of the displacement coordinates $x=\sin\theta_1$ and $y=\sin\theta_2$ of the top of the cantilever. The potential, up to quartic terms, is

$$V(x,y;F)=k_1x^2(1+x^2/3)+k_2y^2(1+y^2/3) +F(1-x^2-y^2)^{1/2}. \quad (47)$$

For displacements in the x-z plane with $y=0$,

$$V(x,0;F)=F+(k_1-F/2)x^2+(k_1/3-F/8)x^4. \quad (48)$$

This can immediately be recognized as a symmetry-restricted cusp catastrophe, A_{+3}. A similar potential describes displacements in the y-z plane $x=0$. Since the potential $V(x,y;F)$ exhibits a bifurcation at $k_1-F/2=0$ and $V(0;y;F)$ exhibits one at $k_2-F/2=0$, it makes sense to design the cantilever with equal spring constants $k_1=k_2=F_c/2$, where F_c is the critical load to be supported by the cantilever.

Or does it?

The potential for the optimized cantilever under an arbitrary displacement is

$$V(x,y;F)=F+\tfrac{1}{2}(F_c-F)(x^2+y^2) +(F_c/6-F/8)(x^4+y^4)-\tfrac{1}{4}Fx^2y^2. \quad (49)$$

The last term represents a strong coupling between the two displacement directions. This coupling term has severe consequences for the collapse behavior of the cantilever.

As the force F is increased, the potential $V(x,y;F)$ remains locally stable until the critical load F_c is reached. At this load, the fourth-degree term $F_c(x^4-6x^2y^2+y^4)/24$ allows sudden collapse of the cantilever in particular directions. This is not what is expected from investigating the planar potentials $V(x,0;F)$ and $V(0,y;F)$: for these potentials the coefficient of the quartic term is positive at the critical load, so that only a small in-plane displacement occurs after the critical load is exceeded. In other words, the postbuckling behavior of the cantilever in either the x-z plane or the y-z plane is "soft" in the sense that after the bifurcation the stable equilibrium is located near the original stable equilibrium $(x,y)=(0,0)$ in the state-variable space.

In cylindrical coordinates $x=r\cos\phi$ and $y=r\sin\phi$ the potential is

$$V(r,\phi;F)=F+(k-F/2)r^2 +(k/3)r^4(\cos^4\phi+\sin^4\phi)-(F/8)r^4. \quad (50)$$

The coefficient of r^4 is largest $(=+1)$ for displacements along the $\pm x$ or $\pm y$ axes ($\phi=n\pi/2$, n integer) and smallest $(=1/2)$ for displacements along the axes $x=\pm y$ ($\phi=\pi/4$

$+n\pi/2$, n integer). Along these two different sets of displacement directions the potentials are

$$V(r,\phi=0;F)=F+(k-F/2)r^2+(k/3-F/8)r^4, \quad (51a)$$

$$V(r,\phi=\pi/4;F)=F+(k-F/2)r^2 + (k/6-F/8)r^4. \quad (51b)$$

The first potential, representing displacements in the x-z or y-z planes, is a catastrophe of type A_{+3}, with "soft" postbuckling behavior. That is, for loads in excess of the critical load the displacement from vertical increases rapidly but continuously. The second potential, representing coupled displacements ($x=\pm y$), is a catastrophe of type A_{-3}, with "hard" postbuckling behavior. For loads in excess of the critical load the only stable position is horizontal—completely collapsed.

These observations may appear academic. After all, the cantilever remained vertical and supported loads F below the critical load. However, we have not yet considered imperfection sensitivities. These are encountered in the fabrication stages (of the springs, for example) and the construction stages in every large project. The most general perturbation, representing arbitrary imperfections for the cantilever, are obtained by investigating the germ of the potential, which is $(x^4-6x^2y^2+y^4)/24$. The universal perturbation of this catastrophe germ is the eight-parameter function

$$\text{Pert}(x,y;a_{ij})=\sum a_{ij}x^iy^j, \quad i\leqslant 2,\ j\leqslant 2,\ (ij)\neq(00). \quad (52)$$

It is truly remarkable that every imaginable imperfection in fabrication of the individual components (nonuniformity in density of material from which the beam is constructed, unwanted bends in the beam, air bubbles, etc.) as well as every imaginable imperfection in construction (springs along the x axis with different spring constants, springs along x and y axes slightly out of perpendicular, rest height not quite vertical, loading force not quite vertical, etc.) can be represented by a universal perturbation with only eight parameters.

However, even eight parameters will present a complicated problem of analysis. In the present case we can do even better. The failure eigendirections for the soft A_{+3} catastrophe are the x and y axes $(\pm x,0)$, $(0,\pm y)$. Under this catastrophe [Eq. (51a)] the original unperturbed stable solution $(x,y)=(0,0)$ becomes unstable for $F>F_c$, giving rise to new stable equilibria $(\pm[6(F/F_c-1)]^{1/2},0)$, $(0, \pm[6(F/F_c-1)]^{1/2})$. Since the new equilibria are relatively near the old one and are stable, this structural failure is a "soft" bifurcation. This "soft" bifurcation is shown in Fig. 15(a).

The other failure eigendirection, $x=\pm y$, is a "hard" bifurcation and a completely different story. In this case [Eq. (51b)] the new equilibria exist for $F<F_c$ at $(\pm[6(1-F/F_c)]^{1/2}, \pm[6(1-F/F_c)]^{1/2})$. These equilibria are saddles of type M_1^2. If for $F<F_c$ a perturbation moves the system state (x,y) to the neighborhood of any of these saddles, there is the possibility of sudden collapse of the cantilever. This "hard" bifurcation is shown in Fig. 15(b).

As a result of these considerations, in order to study the severe failure modes of the propped cantilever, and in particular the imperfection sensitivity of these modes, it is sufficient to study the imperfection sensitivity of the A_{-3} catastrophe [Eq. (51b)]. Under the most general perturbation, the potential describing failure modes in the weak eigendirection is

$$V(r;b,F)=-\tfrac{1}{24}F_c r^4+\tfrac{1}{2}(F_c'-F)r^2+br, \quad (53)$$

where $F_c' = F_c + a$, and a is one of the two control parameters for the cusp catastrophe. The other parameter, b, describes symmetry-breaking perturbations. The load $F_f = F(F_c',b)$ at which the propped cantilever collapses is determined by the values of F_c', b at which the locally stable equilibrium becomes degenerate with either of the two saddles it separates. This locus is the semicubical parabola (21b), which gives for the failure load

$$F_f=F_c'-\tfrac{1}{2}(3bF_c^{1/2})^{2/3}. \quad (54)$$

This simple expression reveals that the failure load is severely decreased by any symmetry-breaking imperfection ($b\neq 0$). The severity is indicated by the power-law dependence: 2/3. An imperfection leading to a small asymmetry in either plane $x=\pm y$ will lead to an unexpectedly large reduction in the cantilever's failure load. The imperfection sensitivity

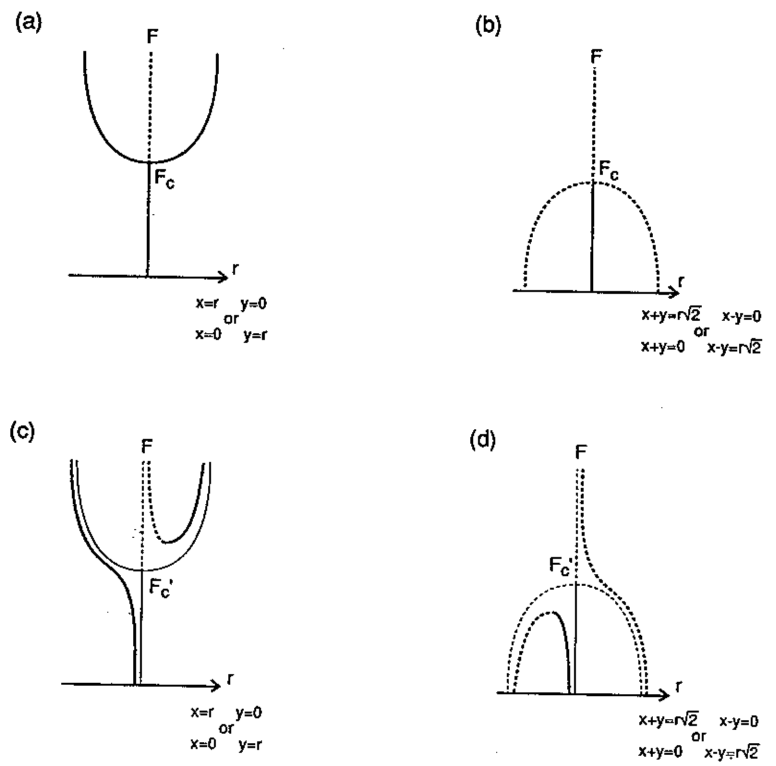

FIG. 15. In all figures stable equilibria (Morse saddles M_0^2) are shown by solid curves and unstable equilibria (M_1^2) are shown by dashed curves. (a) Equilibrium displacement for the "soft" A_{+3} catastrophe along the axes $(x,y=0)$ or $(x=0,y)$. Prebuckling path ($F<F_c$) is $(x,y)=(0,0)$. Beyond the critical load, $F>F_c$, the postbuckling path is parabolic, with $r\sim(F-F_c)^{1/2}$. (b) Equilibrium displacement for the "hard" A_{-3} catastrophe along the directions $x=\pm y$. Prebuckling path $(x,y)=(0,0)$ is stable for $F<F_c$. There are no stable equilibria for $F>F_c$. (c) The most general perturbation of a cusp catastrophe consists of a linear plus a quadratic term: $(a/2)r^2+br$. The quadratic term $(a/2)r^2$ shifts the critical load: $\frac{1}{2}(F-F_c)r^2+(a/2)r^2+br\rightarrow\frac{1}{2}(F-F_c')r^2+br$, where $F_c'=F_c-a$. The linear term breaks the symmetry. For the "soft" catastrophe A_{+3} the perturbation displaces the equilibrium as shown. (d) For the "hard" catastrophe A_{-3} the perturbation can dramatically reduce the load at which failure occurs, as shown.

of the "soft" and "hard" bifurcations are shown in Figs. 15(c) and 15(d).

Having analyzed the problem, can we do anything about it? The first step is to determine how severe the problem is. To determine the failure load, we make use of the catastrophe flags. The first five (modality, sudden jumps, inaccessibility, sensitivity, hysteresis) are observable when the cantilever collapses. This is precisely the problem we would like to avoid. The remaining three catastrophe flags (divergence of linear response, critical slowing down, anomalous variance) are diagnostics designed precisely for the purposes at hand. That is, the system is subjected to gradually increasing loading. The following data are then recorded:

1. the displacement, r, from equilibrium;
2. the normal-mode frequency, ω, for small oscillations;
3. the amplitude, A, for small oscillations about equilibrium.

These three observables depend in a canonical way on the control parameters $F-F_c'$ and b. Testing is carried out gently and terminated before the collapse of the cantilever. By extrapolating the equilibrium displacement $r(F)$, the normal-mode frequency $\omega(F)$, and the amplitude of small oscillations $A(F)$, the failure load F_f and the control-parameter values F_c' and b at the failure load can be determined (Gilmore, 1981). Such extrapolations are shown in Fig. 16.

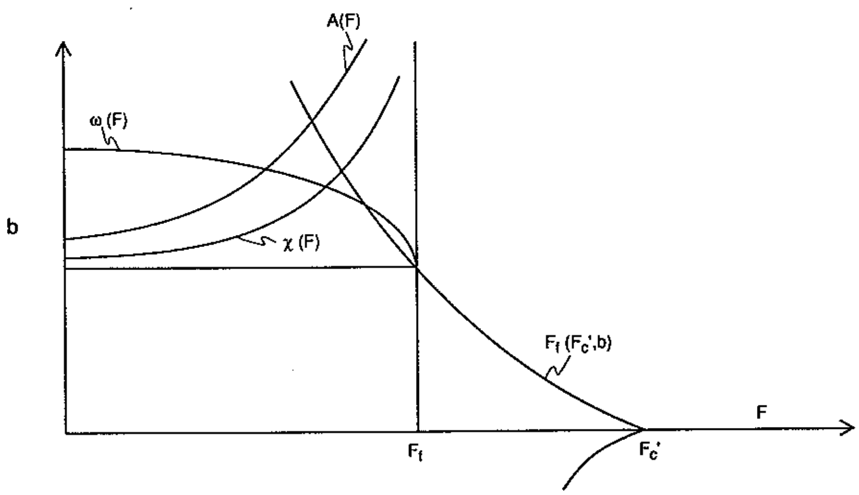

FIG. 16. The failure load $F_f(F_c,b)$ has a 2/3 power-law dependence on the symmetry-breaking parameter *b*. The value of the load at which collapse occurs (vertical line at F_f) can be located by exploiting the catastrophe flags. This is done by plotting the frequency of small oscillations, $\omega(F)$ (critical slowing down), by determining the linear susceptibility, $\chi(F)$ (divergence of linear response), and by measuring the amplitude of small equal-energy oscillations, $A(F)$ (anomalous variance), all as functions of increasing load *F*. The amplitude diverges and causes collapse before the failure load F_f is reached.

Once the symmetry-breaking parameter *b* has been determined, small adjustments to the cantilever can be made to reduce the magnitude of *b*. Such adjustments may possibly include the nonintuitive steps of decreasing the strength of some component (e.g., reducing one of the spring constants) in order to increase the strength of the compound system.

Much larger compound structures have many more degrees of freedom, and consequently many more failure modes. While some failure modes are clear-cut, others may be very subtle—depending on unexpected coupling between "soft" failure modes, as the example above illustrates. Characterization of the imperfection sensitivity is algorithmic: for any catastrophe germ (the potential of the perfect optimized structure at the failure load) there is a universal perturbation. This is finite dimensional provided that the germ is determinate. It is important to understand the universal perturbation in order to catalog all possible failure modes of the system, both obvious and not, in order to prevent catastrophes from occurring in such complex structures.

3.4 Elementary Catastrophes in Nonlinear Dynamics

Elementary catastrophe theory is the first result in the program of catastrophe theory. As a result, it stands at the base of the pyramid of bifurcations to be expected in nonlinear dynamical systems (see CHAOTIC PHENOMENA). Any bifurcations encountered in elementary catastrophe theory will be encountered in nonlinear systems with greater complexity.

To illustrate how the elementary catastrophes are encountered in more complex nonlinear systems, we survey here how the two simplest elementary catastrophes appear in nonlinear systems at the next level of complexity in the hierarchy of nonlinearity. Such systems are low-dimensional periodically driven dynamical systems or, equivalently, autonomous (time independent) dynamical systems with one more dimension.

In treating two-dimensional periodically driven dynamical systems and three-dimensional autonomous systems it is typical to explore the bifurcation diagram by sweeping one control parameter and locating periodic and aperiodic orbits. When a single control parameter is varied, it is typical to encounter among the catastrophes only the fold. If, however, there is a symmetry present, then the symmetry-restricted cusp catastrophe $[V(x;a)=\frac{1}{4}x^4+\frac{1}{2}ax^2=V(-x;a)]$ may also be encountered.

The fold catastrophe appears in bifurcation diagrams for autonomous dynamical systems as saddle-node bifurcations. These are shown in Fig. 17. In this figure there are several control parameter values at which new stable (solid curve) and unstable (dashed curve) solutions appear. If the system is periodically driven the saddle-node pair consists of a stable periodic orbit and a saddle orbit with one unstable direction. Both orbits have the same period.

Under a broad range of conditions the saddle remains unchanged as the control-parameter values are further increased, but the node undergoes additional bifurcations. Such bifurcations are often of the period-doubling variety. The period-doubling bifurcation is essentially a symmetry-restricted cusp catastrophe. Its relation to the cusp is shown in Fig. 18. As a control-parameter value is increased past some threshold, a stable periodic orbit loses its stability. In a transverse section (Poincaré plane) the unstable point is surrounded by two stable nodes. As the Poincaré plane is swept around the unstable periodic orbit, the two nodes will rotate around the direction of propagation of the Poincaré section either an integer number of times or a half-integer number of times. If the two nodes rotate an integer number of times around the saddle that is between them [Fig. 18(a)], the two nodes belong to two separate periodic orbits, both of which have the same period as the unstable periodic orbit from which they bifurcated. On the other hand, if the two nodes rotate a half-integer number of times around the saddle that they straddle [Fig. 18(b)], they belong to a single orbit whose period is twice the period of the unstable periodic orbit from which they bifurcated.

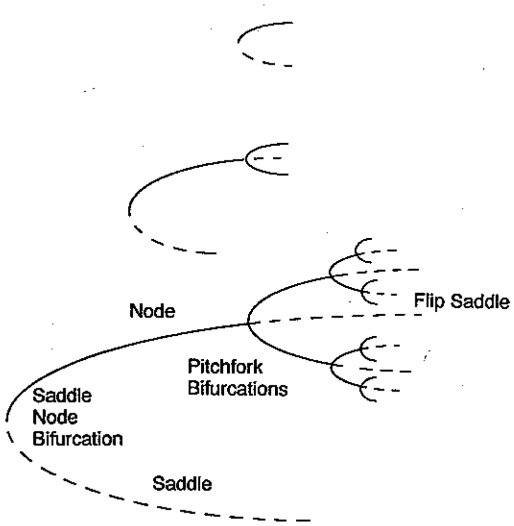

FIG. 17. Fold and symmetry-restricted cusp (also called pitchfork) bifurcations are commonly encountered in nonlinear dynamical systems. At a fold or saddle-node bifurcation a stable orbit (solid lines) and a saddle orbit (dashed lines) are created. At a pitchfork bifurcation, period doubling may occur, leading to creation of orbits of twice the period of the original orbit, which loses its stability.

Saddle-node bifurcations and period-doubling cascades are general features of autonomous dynamical systems. The pitchfork bifurcation with an integer number of twists is atypical in autonomous dynamical systems. However, if there is a symmetry, as occurs, for example, in the Lorenz equations, a period-doubling cascade cannot occur beginning at a symmetric orbit (Gilmore, 1981). As a result, the first pitchfork bifurcation occurs at a symmetric orbit of period T, creating an asymmetric pair of orbits of the same period. Subsequent pitchfork bifurcations from the asymmetric periodic orbits then generate period-doubling cascades.

A simple Hopf bifurcation can also be viewed as a symmetry-restricted cusp catastrophe (Gilmore, 1981). In the neighborhood of a Hopf bifurcation the dynamical equations can be written in polar coordinates as

$$dr/dt=\lambda r\pm r^3, \quad d\theta/dt=\omega. \tag{55}$$

During the bifurcation the equation of motion for θ remains essentially unchanged. The bifurcation is therefore described by the radial equation, which can be written in gradient form:

$$dr/dt=-\partial V/\partial r, \quad V(r;\lambda)=-\tfrac{1}{2}\lambda r^2-(\pm\tfrac{1}{4}r^4). \tag{56}$$

A typical Hopf bifurcation is illustrated in Fig. 19. A section of the bifurcation surface in a plane containing the control parameter axis has the standard pitchfork shape.

Since the saddle-node, period-doubling, symmetry-breaking, and Hopf bifurcations are all related to elementary catastrophes, all

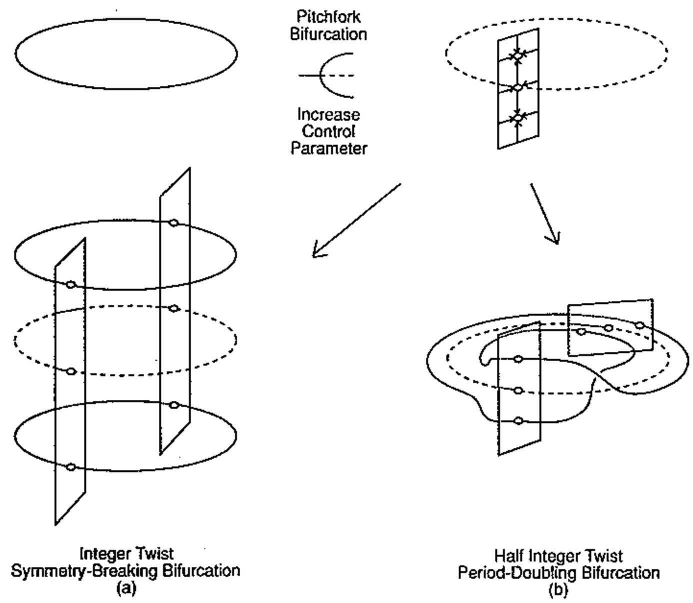

FIG. 18. When a periodic orbit loses its stability in a pitchfork bifurcation, either two new orbits of the same period may be created or one new orbit of twice the period may be created, depending on the local torsion about the original orbit (integer or half-integer).

the phenomenology associated with the elementary catastrophes is present and observable for these bifurcations as well. For example, one can observe modality, sudden jumps, inaccessible regions, sensitivity, hysteresis, divergence of linear response, time dilation, and anomalous variance in all these bifurcations. Furthermore, the behavior of each bifurcation has the same canonical form as that of the corresponding catastrophe (e.g., power-law dependence).

APPENDIX: A BRIEF HISTORY OF CATASTROPHE THEORY

Catastrophe theory burst upon public consciousness in the mid-1970s with a series of articles in widely available sources: *Scientific American, Science, London Times, Time,* and *The New York Times.* The articles themselves were bimodal. At one extreme the *London Times* heralded it as the "main intellectual movement of the century" while the *Science* article announced "The emperor has no clothes."

The theory of singularities of mappings was intensively developed by Whitney, Thom, and Mather in the decade 1955–1965. Thom presented his formulation of the theory of singularities in his eagerly awaited work *Stabilité Structurelle et Morphogénese* (1972). This book was an enigma in both form and substance. It was largely inaccessible to the mathematics community because it was written in the language of biologists, and inaccessible to the biological community due to its presentation of mathematical concepts which seemed to be deep and mysterious.

Yet it held the promise of describing discontinuous phenomena in a systematic way. Zeeman crowned the subject with the florid name "catastrophe theory" and took up the challenge of presenting these mathematical ideas to the larger community of scientists. When challenged to show concrete applications in the real world, he responded with a series of articles of stunning originality which explored the range of possible applications of this new subject.

These applications eventually worked their way into public consciousness through the popular journals. A reaction to the value of catastrophe theory set in. This reaction was

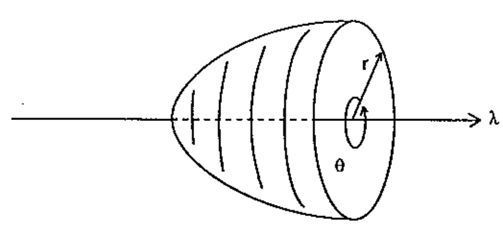

FIG. 19. A Hopf bifurcation can be viewed as a symmetry-restricted cusp catastrophe in the radial direction.

partly as a result of the overblown claims made in its name, partly as a result of the neglect offered to the workers who created the field of dynamical systems theory: Poincaré, Andronov, Pontryagin, Smale, etc., and those currently working in this field. The result was a multiyear public dialogue on the merits of this subject of discontinuities using arguments now long forgotten and best left unearthed, a dialogue of which the public eventually tired.

The appearance of the monographs by Poston and Stewart (1978) and Gilmore (1981) made it clear that this was a subject of substance, which had to be taken seriously, one capable of providing a useful language for the description of discontinuities at both a qualitative and a quantitative level.

GLOSSARY

Anomalous Variance: A catastrophe flag. Amplitude of motion about an equilibrium becomes increasingly large as a degenerate critical point is approached.

Bifurcation: A qualitative change in the properties of a system.

Bifurcation Set: Set of values in control-parameter space at which qualitative changes occur.

Catastrophe: Mathematical—A family of functions, depending on control parameters, in which the number of equilibria changes as the control parameters are varied. Physical—A sudden, discontinuous change in the state of a system.

Catastrophe Flags: Phenomena that occur when a catastrophe (sudden jump) is present.

Catastrophe Germ: A function which describes a degenerate critical point.

Catastrophe Theory: The program for determining how the qualitative properties of the solutions of equations change as parameters appearing in these equations change.

Clausius–Clapeyon Equations: A set of equations which determine the Maxwell set of a catastrophe.

Control Parameters: A set of parameters $c=(c_1,c_2,\ldots,c_k)$ which appear in the equations that describe a physical system.

Critical Curvature: The curvature d^2f/dx^2 of a function evaluated at its critical points $\nabla f=0$.

Critical Point: An equilibrium: a point at which motion does not occur.

Critical Set: The k-dimensional manifold $\nabla_x f(x;c)=0$ for $x\in R^n$ and $c\in R^k$.

Critical Slowing Down: A catastrophe flag. Relaxation to equilibrium takes increasingly long times as a degenerate critical point is approached.

Critical Values: Values of a function at its critical points.

Cusp Catastrophe: The function $A_3(x)=x^4/4+ax^2/2+bx$. So named because the critical set $\nabla_x A_3(x;a,b)=0$ exhibits a cusp on projection to the a-b control parameter plane.

Cuspoids A_{n-1}: A class of functions depending on one state variable x and $n-2$ control parameters. The family of functions $A_{n-1}(x)=x^n/n+\Sigma_{i=1}^{n-2}a_ix^i$ has up to $n-1$ isolated critical points.

Degenerate Critical Point: An equilibrium ($\nabla V=0$) at which the stability matrix $(\partial^2V/\partial x_i\partial x_j)$ is singular. This requires two or more critical points to be arbitrarily close to each other (degenerate).

Delay Convention: The system state remains in a local minimum until that minimum ceases to exist.

Divergence of Linear Response: A catastrophe flag. Linear response coefficients increase very rapidly as a degenerate critical point is approached.

Dynamical System: A set of coupled first-order ordinary differential equations which may be nonlinear: $x_i' = f(x,c;t)$, where x is a state vector and c are control parameters.

Elementary Catastrophe: A function of one or two state variables (x,y) and one or more control parameters. These functions are used to describe systems in which small causes can produce large effects.

Equilibrium: A point at which all forces vanish; a critical point.

Exceptional Elementary Catastrophes: E_6, E_7, E_8 depend on two state variables and have up to 6, 7, 8 critical points, respectively.

Flow: The motion of a point in phase space under the equations of motion for the system.

Fokker–Planck Equation: An equation that describes the evolution of a probability distribution in the presence of diffusion and driving forces.

Fold Catastrophe: The function $A_2(x) = x^3/3+ax$. So named because the critical set exhibits a fold.

Generic: Typical. Occurs with "probability one."

Gradient Dynamical System: A dynamical system in which the time-independent forcing term is the gradient of a potential function: $f_i(x,c) = -\partial V(x,c)/\partial x_i$.

Hopf Bifurcation: Bifurcation in which a stable fixed point becomes unstable and emits a stable periodic orbit (limit cycle).

Hysteresis: Transitions from one state to another do not occur at the same values of the control parameters when the control parameters are changed in the opposite direction.

Implicit Function Theorem: A theorem of elementary calculus. If a function has a nonzero slope at a point, $\nabla f \neq 0$ at x_0, then it can always be approximated by a linear function at x_0.

Inaccessibility: An unstable physical state separating two stable physical states.

Isolated Critical Point: An equilibrium ($\nabla V=0$) at which the stability matrix ($\partial^2 V/\partial x_i \partial x_j$) is nonsingular.

Maxwell Convention: The system state is always in the deepest minimum.

Maxwell Set: The set of control parameter values at which a function has two or more equally deep minima.

Metastable: Stable to small perturbations but not to large perturbations.

Modal Catastrophe: A catastrophe germ which depends on one or more parameters.

Modality: Distinct types of behavior that a system can exhibit under identical or nearly identical conditions (e.g., liquid–gas states for a fluid).

Mode Softening: A catastrophe flag. Oscillation frequencies approach zero as a degenerate critical point is approached.

Morse *i*-Saddle: The quadratic function $M_i^n(x) = -x_1^2 - \cdots - x_i^2 + x_{i+1}^2 + \cdots + x_n^2$.

Morse Lemma: A normal-form theorem like the implicit function theorem. If a function of n variables has a nonsingular stability matrix (det $\partial^2 f/\partial x_i \partial x_j$) at an equilibrium x_0, it can always be approximated by a quadratic function at x_0. Such a function is called a Morse saddle.

Nongeneric: Atypical. Occurs with "probability zero."

Perestroika: Change in the properties of a dynamical system due to changes in its control-parameter values.

Perturbation: A function which is everywhere small.

Phase Space: A space in which the coordinates of a point define the state of the system.

Phase-Space Portrait: A trajectory (or trajectories) in phase space which describes the evolution of a physical system from one (or more) initial conditions.

Pitchfork Bifurcation: Bifurcation in which a stable fixed point becomes unstable and two new stable fixed points are emitted. Corresponds to a single-well potential developing into a double-well potential.

Quadratic Form: A function of n variables $x=(x_1,x_2,\ldots,x_n)$ which can be written $x^t Mx$, where M is a symmetric $n \times n$ matrix.

R^n: Real n-dimensional space.

Saddle-Node Bifurcation: Bifurcation in which two fixed points, one saddle and one node, are created.

Sensitivity: The final state of a system may change under small perturbations of either the initial conditions or processes applied to the system.

Stability Matrix: The matrix of mixed second partial derivatives of a function: $\partial^2 f/\partial x_i \partial x_j$.

State Vector: A set of parameters $x=(x_1, x_2,\ldots,x_n)$ which describes the state of a system.

Structural Stability: Invariance of qualitative properties under a perturbation.

Taylor Tail: Tail of a Taylor series expansion: all terms beyond a certain degree.

Thom Classification Theorem: A list of elementary catastrophes depending on one or two state variables and up to five control parameters.

Thom Splitting Lemma: A normal form theorem for functions. If a function has an equilibrium ($\nabla f=0$) at which the stability matrix $\partial^2 f/\partial x_i \partial x_j$ has l vanishing eigenvalues, the function can always be written as the sum of two functions, $f(x)=f_{\mathrm{NM}}(y_1,\ldots,y_l)+M_i^{n-l}(y_{l+1},\ldots,y_n)$, where $f_{\mathrm{NM}}(y_1,\ldots,y_l)$ is a function with vanishing first and second derivatives in the l variables corresponding to the vanishing eigenvalues and M_i^{n-l} is a Morse i-saddle in the remaining $n-l$ variables.

Time Dilation: A catastrophe flag. System response takes increasingly long times as a degenerate critical point is approached.

Umbilics D_{n-1}: A class of functions depending on two state variables and $n-2$ control parameters. The family of functions $D_{n-1}(x,y)$ has up to $n-1$ isolated critical points.

Universal Perturbation: Simplest function (i.e., lowest dimension, or number of control parameters) that describes all possibilities that can result when a degenerate critical point is perturbed.

List of Works Cited

Arnol'd, V. I. (1981), *Singularity Theory,* Cambridge, England: Cambridge University.

Arnol'd, V. I. (1986), *Catastrophe Theory,* 2nd ed., Berlin: Springer-Verlag.

Gilmore, R. (1981), *Catastrophe Theory for Scientists and Engineers,* New York: Wiley.

Poston, T., Stewart, I. N. (1978), *Catastrophe Theory and Its Applications,* London: Pitman.

Thom, R. (1975), *Structural Stability and Morphogenesis,* New York: Benjamin-Addison-Wesley.

Zeeman, E. C. (1977), *Catastrophe Theory: Selected Papers (1972–1977),* Reading, MA: Addison-Wesley.

Further Reading

Anon. (1973), *Manifold* **14**, 6–13.

Anon. (1975), *Times Higher Educational Supplement,* Dec. 5, p. 12.

Arnol'd, V. I. (1974), *Proc. Int. Cong. Math. Vancouver 1974,* pp. 19–39.

Croll, J. (1976), *New Sci.* **17**, 630–632.

Fussbudget, H. J., Znarler, R. S. (1979), *Math. Intell.* **2**, 56–59.

Guckenheimer, J. (1978), *Math. Intell.* **1**, 15–20.

Saunders, P. T. (1980), *An Introduction to Catastrophe Theory,* Cambridge, England: Cambridge University.

Stewart, I. N. (1975), *New Sci.* **68**, 447–464.

Stewart, I. N. (1977), in: *Encyclopedia Britannica,* Special Supplement to the Yearbook.

Sussman, H. J., Zahler, R. S. (1978a), *Synthêse* **37**, 117–126.

Sussman, H. J., Zahler, R. S. (1978b), *Behav. Sci.* **23**, 383–389.

Thom, R. (1975), in: *Dynamical Systems—Warwick 1974,* Lecture Notes in Mathematics 468, Berlin: Springer-Verlag, pp. 366–372.

Thompson, J. M. T., Hunt, G. W. (1974), *Eng. Optim.* **1**, 99–110.

Thompson, J. M. T. (1982), *Instabilities and Catastrophes in Science and Engineering,* New York: Wiley.

Whitney, H. (1955), *Ann. Math.* **62**, 374–410.

Woodcock, A., Davis, M. (1980), *Catastrophe Theory,* New York: Avon.

Zahler, R. S., Sussman, H. J. (1977), *Nature* **269**, 759–763 (Correspondence Dec. 1, Dec. 29, 1977).

Zeeman, E. C. (1975), in: *Dynamical Systems–Warwick 1974,* Lecture Notes in Mathematics 468, Berlin: Springer-Verlag, pp. 373–383.

Zeeman, E. C. (1976), *Sci. Am.* **234**, 65–83.