SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NR 170-938 FINAL
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): CATASTROPHE THEORY: STATE OF THE ART AND POTENTIAL APPLICATIONS
5. TYPE OF REPORT & PERIOD COVERED: Final 6/82-8/83
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s):
F. Craig Johnson
R. C. Lacher
8. CONTRACT OR GRANT NUMBER(s): N00014-82-G-0065
9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Florida State University
Tallahassee, Florida 32306
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 170-951
11. CONTROLLING OFFICE NAME AND ADDRESS: Organizational Effectiveness Research Program, Office of Naval Research (Code 452), 800 North Quincy, Arlington, Virginia 22217
12. REPORT DATE: August 31, 1983
13. NUMBER OF PAGES: 138
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approval for public release of this report is unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
Catastrophe theory, dynamical systems, stochastic differential equations, bifurcation theory
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):
Position papers and commentaries are presented covering some of the controversy surrounding applications of elementary catastrophe theory to social science in general and to organizational effectiveness in particular. Recommendations are made for research in the near and long term.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

CATASTROPHE THEORY: STATE OF THE ART

AND

POTENTIAL APPLICATIONS

F. Craig Johnson
R. C. Lacher

Florida State University
314 Stone Building
Tallahassee, Florida 32306

Preparation of this report was sponsored by Contract NR 170-938 through the Office of Naval Research (ONR). This contract was funded jointly by ONR's psychological, mathematical, and statistical programs. The contracted project was designed to review what is known about dynamical systems in general and bifurcation and catastrophe theory in particular. Several symposia, involving ranking experts in the fields of mathematics and statistics as well as in the social and behavioral sciences, were conducted at Florida State University.

TABLE OF CONTENTS

INTRODUCTION, PROCEDURES, AND RECOMMENDATIONS

SCIENTIFIC BACKGROUND

In February of 1979, the Navy Personnel Research and Development Center in San Diego, California, issued technical report 79-8, "Catastrophe Theory in the Behavioral Sciences" written by W. A. Hillix et al. The characteristics of catastrophe theory were described, applications reviewed, and some experimentation on perception conducted. The conclusions and recommendations included the following:

> Catastrophe theory is conceptually rich and provides a new way of looking at and conducting psychological experiments. While its promise has not yet been realized, it is far too early to dismiss the theory.
>
> Work should proceed toward further developing statistical procedures that will help in evaluating the fit of catastrophe theory models to data.
>
> Renewed emphasis should be placed on the sequential and systematic manipulation of independent variables as a tool in behavioral methodology. (Pages 45 and 47)

In April of 1980, E. C. Zeeman presented a paper on "Catastrophe Models in Administration" in which he observed that in administration there are many "phenomena where continuously changing forces produce discontinuous effects." He detailed examples in endangered institutions, obsolescence versus need, territorial defense, job satisfaction, committee behavior, and decision making, among others.

At recent meetings of the European Association for Institutional Research, papers (Cossu, 1980; Johnson, 1981; Lacher, 1981) were presented on cost accounting, budget planning, and decision making which applied the principles of catastrophe theory to administration in institutions of higher education in several different countries.

At this time there are many people who are beginning to be interested in the potential of catastrophe theory in administrative practice. One such expression of interest was made by President Donald E. Walker (1980) of Southeastern Massachusetts University when he said, "I think we (administrators) repair to the oracle of the computer too quickly with rather fixed stereotypes in our minds, and because of the questions we ask, get answers that are both unrealistic and not useful. The probelm is more complicated by the fact that even practiced administrators have the wrong models in their heads of what they do; they end up explaining catastrophe only in terms of pathology, rather than the end result of the predictable 'natural process'."

The administrators who share this view are not typically mathematicians or statisticians and require technical assistance to make applications of the theory. The mathematicians and statisticians who have the technical knowledge do not have the experience nor the insights into the problems needed to understand the underlying dynamics of organizational systems.

OBJECTIVE

The purpose of this report is to bridge the communications gap between mathematicians and statisticians who have worked on catastrophe theory in the social sciences, and administrators who understand the dynamics of organizational behavior.

PROCEDURES

Originally we envisioned the development of an encyclopedia from a lexicon prepared for this project (see Appendix), and Professor Tim Poston from University of Warwick contributed to this early planning. It soon became apparent, however, that this was an unrealistic expectation because of the differences in mathematical backgrounds among the participants. It was decided, rather, to bring together on the Florida State University campus five small groups of two to three participants each in order to carry on intensive discussions of catastrophe theory applications and, following this, to ask each participant to write his reflections upon the state-of-the-art and the potential developments of catastrophe theory. These work products were then reviewed by all other participants and comments were collected from this exchange. From these materials a draft report was prepared, circulated and taken to the University of Warwick in England for review by Professors E. C. Zeeman, Ian Stewart and J. Q. Smith. The participants who contributed to this process were:

Mathematics

Dr. J. J. Callahan, Professor, Smith College
Dr. H. M. Hastings, Professor, Hofstra College
Dr. L. Markus, Professor, University of Minnesota
Dr. C. P. Simon, Professor, University of Michigan
Dr. I. N. Stewart, Professor, University of Warwick
Dr. E. C. Zeeman, Professor, University of Warwick

Statistics

Dr. L. Cobb, Professor, South Carolina Medical School
Dr. I. R. Savage, Professor, Yale University

Dr. J. Q. Smith, Professor, University of London
Dr. J. E. Stecklein, Professor, University of Minnesota

Administration

Dr. R. T. Holt, Graduate Dean, University of Minnesota
Dr. B. Lawrence, Director, NCHEMS
Dr. D. E. Walker, President, Southeastern Massachusetts University

In addition, several other people agreed to review the work products and to comment. These included:

Dr. K. S. Cameron, Director of Organizational Studies, NCHEMS
Dr. P. Peregoy, Professor, University of Pittsburgh at Johnstown
Dr. W. A. Simpson, Professor, Michigan State University

During the several symposia on the Florida State University campus a seminar was held with the following graduate students: Paul Carney, Amilcar Castellano, Julia Duckwall, and Gail Fletcher. A special contribution to the project was made by Gail Fletcher and Julia Duckwall who took primary responsibility for the typing, proofreading and production of this report. Also, we are indebted to the special contribution of Dr. James Callahan for his preparation of all the drawings.

END PRODUCT

What follows are the principal investigators' recommendations; views of selected participants who contributed papers on the key issues of the state-of-the-art of catastrophe theory; and the reactions, in the form of letters, from other participants on these issues. Finally, the lexicon, which formed the mathematical base, is appended to this report.

The end products found in this report approach the qualitative methods in the social sciences from the theory of parametrized systems of (not necessarily linear) differential equations, or from another point of view, in the theory of slowly evolving (nonlinear) dynamical systems. A qualitative language is being developed in this active field of mathematical research and offers one of the best settings for the study, classification, and eventual understandingof many of the difficult problems in organizational effectiveness. Our recommendations suggest some first steps toward that understanding.

RECOMMENDATIONS FOR FUTURE RESEARCH SUPPORT

Many of the details of these recommendations were obtained during written and verbal interaction with the participants whose contributions are in this report.

Catastrophe theory, in its generalized form, is a mathematically rich and active field. Competent mathematical researchers are currently being supported by the National Science Foundation (NSF) and since this support is expected to continue for the foreseeable future, there seems little reason to recommend that the Office of Naval Research (ONR) direct its attention to catastrophe theory in the general form.

Elementary catastrophe theory, where initial applications in the social sciences can be expected to be found, is essentially complete, known, organized, and exposited. There seems little reason to recommend support of more research in this area, by either NSF or ONR, with two possible exceptions: the **elementary catastrophe theory of stochastic differential equations** seems very important and is almost completely unstudied, and practical **geometric understanding of otherwise known complicated singularities** is interesting and potentially useful.

Statistical research on elementary catastrophe theory is another matter. Practically nothing has been done and we encourage support of two types of non-goal-oriented pure research proposals: statistical problems associated with **detection and calibration of elementary catastrophes** over the next five to ten years, and **parametrized stochastic differential equations** for an indefinite future.

Organizational effectiveness research, we feel, should wait upon a determination as to whether catastrophe theory can be useful in advancing knowledge in the social sciences in general. This determination should be made by supporting some projects through their fruition. These projects must be more than conjectural qualitative model building; they must involve experiment and data collection. For the near term, we recommend support of **applied projects** conforming to a fairly rigid set of guidelines.

Thus we have three types of projects to recommend: for the near term, applied projects of specified standard (described in 3.3); for the middle term, research into the statistics of elementary catastrophe theory; and for the long term, support of parametrized stochastic differential equations. We conclude by listing these project types and commenting on evaluation of proposals of the applied type.

1 Parametrized stochastic differential equations. (Long term support recommended)

1.1 **Brief description.** Research in stochastic differential equations (SDE), particularly parametrized families of same (PSDE); nonlinear bifurcation problems in SDE; singularities and catastrophes in SDE; genericity in SDE.

1.2 **Potential importance.** SDE is a relatively new area of interest in the mathematical sciences; PSDE is virtually embryonic. SDE will likely be one of the great intellectual bridges between determinists (users of differential equations or dynamical systems models, as in mathematics and classical physics) and stochastists (users of statistical and probabilistic models, as in statistics and certain branches of modern physics). PSDE has the potential of playing the same role, expanded to include such areas as bifurcation theory, catastrophe theory and differential dynamics. SDE and PSDE would seem to be the most natural mathematical settings for studying phenomena which have both a deterministic and a stochastic component driving their dynamics.

2 Statistics of elementary catastrophe theory. (Middle term support recommended--5 to 10 years)

2.1 **Brief description.** Problems associated with detection and/or quantification of elementary catastrophes in data; theoretical statistics of multi-modal decision problems, distributions, parametrized distributions, genericity, and other non-single-valued phenomena.

2.2 **Potential importance.** The statistics of elementary catastrophe theory would play an essential role in applied projects of a quantitative nature, just as present-day statistics is indispensable in current social science research.

3 Quantification of models in the social sciences. (Short term support recommended--3-5 years)

3.1 **Brief description.** Projects to collect data or otherwise empirically verify, calibrate, or quantify catastrophe-theoretic models of specified phenomena in the social sciences.

3.2 **Potential importance.** Studies of this type are necessary, and possibly sufficient, to judge on a factual basis the potential of catastrophe theory in the social sciences.

3.3 **Recommended structure of proposals.** Proposals for Type 3 projects should be supported only if they meet each of the following criteria:

3.3.1 contain a catastrophe-theoretic model with variables already named (but not necessarily scaled)

3.3.2 argue creditably for the plausibility of the model as far as possible on qualitative grounds

3.3.3 discuss which variables in the model can and should be scaled and measured

3.3.4 indicate what methods would be used to accomplish 3.3.3.

3.4 **Evaluation.** In evaluating proposals of Type 3 meeting the criteria set forth in 3.3.1 through 3.3.4, some possible **negative features** are as follows:

3.4.1 confusion of or lack of proper distinction between proposed model and reality

3.4.2 imprecise or inadequate enunciation of hypotheses of proposed model

3.4.3 naive or mathematically unsound discussions of catastrophe theory and Thom's Classification Theorem

3.4.4 overambitious output/input ratio, i.e. proposals to solve major problems or understand complex phenomena with relatively little effort.

In addition, proposals of Type 3 should be evaluated on the basis of significance of the problem under study, soundness of methods proposed to gather data or otherwise scale variables, plausibility of the proposed model, and competence of the research team.

To elaborate on the last point, the research team should display competence in mathematics, statistics, and social science. Mathematical competence should include an understanding of elementary catastrophe theory, including the concepts of topological equivalence, genericity, and local classification of singularities. Statistical competence should include working knowledge of procedures likely to arise in proposed data gathering and analysis. Social science competence should include thorough knowledge of relevant and background literature as well as experience in research in or near the proposed topic of study. These competencies could conceivably be found in one investigator, but it would seem more likely that teams of two or more would be necessary to obtain appropriate competency levels in all three of these directions.

OVERVIEW

OVERVIEW OF CATASTROPHE THEORY

Ian Stewart

Activity in the field divides into (a) the mathematics and (b) its applications. As usual, there is a constant trade of ideas and problems between the two.

A comprehensive bibliography is:

Wetherhilt & Zeeman. *1981 bibliography on catastrophe theory*

Introductory texts in increasing order of difficulty include:

Woodcock & Davis. *Catastrophe theory*

Saunders. *Introduction to catastrophe theory*

Thompson. *Instabilities and catastrophes in science and engineering*

Gilmore. *Catastrophe theory for scientists and engineers*

Poston & Stewart. *Catastrophe theory and its applications*

Zeeman. *Catastrophe theory: Selected papers 1972-1977*

THE MATHEMATICS

For mathematicians, the main interest in Thom's "catastrophe theory" is that it lays down a broad program for understanding bifurcating systems from the qualitative viewpoint. What most expository articles have (for simplicity) called "catastrophe theory" is really only a part of the subject, more properly called "elementary catastrophe theory." While certain technical questions do remain open, elementary catastrophe theory may be considered as a completed piece of mathematics: the main challenge is to extend it to other settings. Figure 1 shows some twenty variations on the theme of Thom that have been pushed through during the past decade.

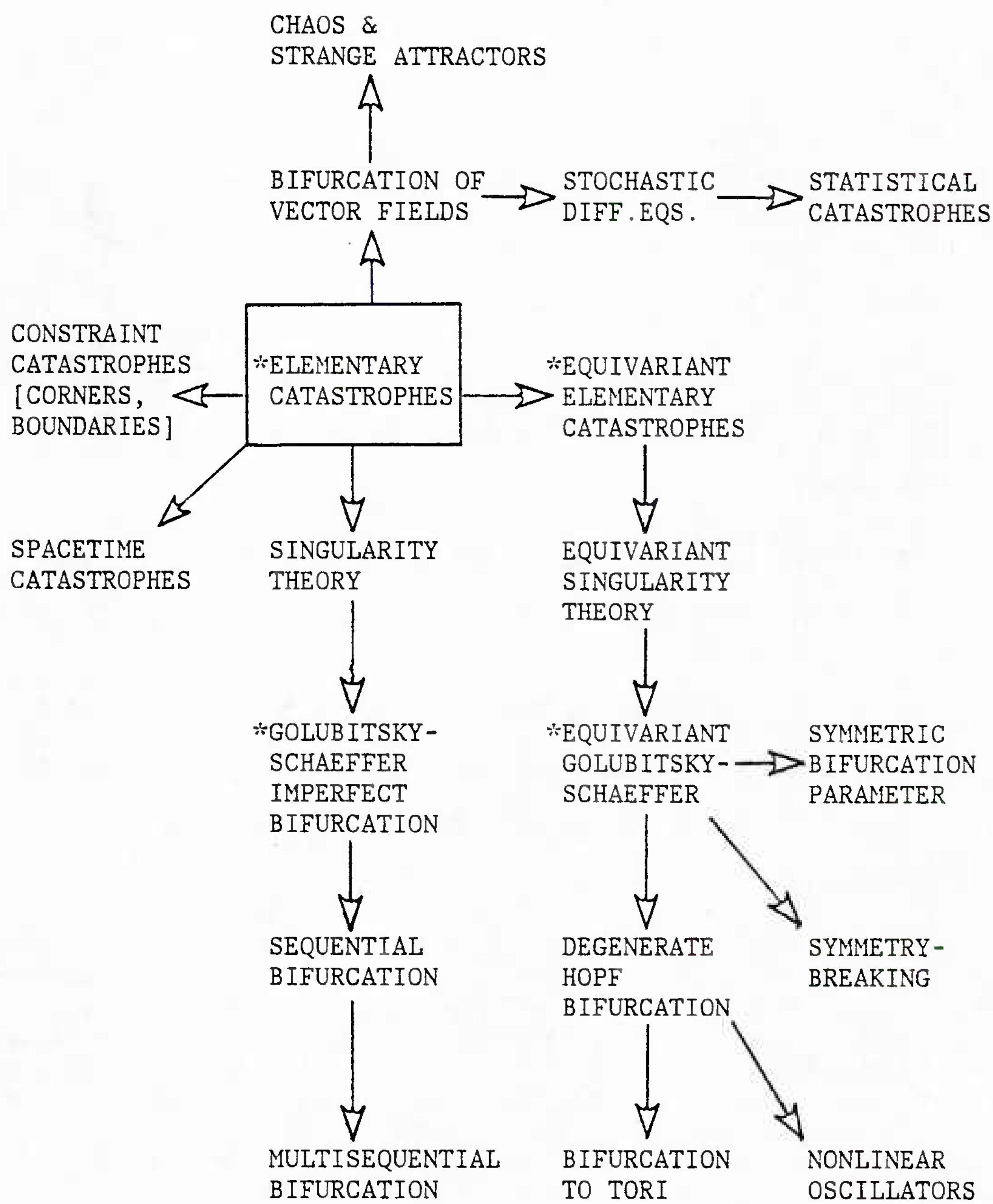


FIGURE 1. The main extensions of elementary catastrophe theory and their interrelations over the past decade. (* indicates a variation which has been developed both in finite and infinite dimensions.)

The major works in the field are:

Arnold (various papers)

Golubitsky & Langford. Classification and unfolding of degenerate Hopf bifurcations

Golubitsky & Schaeffer. A theory for imperfect bifurcation via singularity theory

Golubitsky & Schaeffer. Imperfect bifurcation in the presence of symmetry

Malgrange. Ideals of differentiable functions

Mather (various papers)

Thom. Structural stability and morphogenesis

Good expositions of the technical mathematics may be found in the books of Gibson (1979), Lu (1976), Martinet (1982); and Brocker & Lander (1975) in the Wetherhilt & Zeeman (1981) bibliography.

The mathematical problems which have been solved in each of the areas shown in Figure 1 may be grouped as

(a) classification: What are the possible phenomena?
(b) recognition: How do we tell which one we've got?
(c) perturbation: If small disturbances are acting, what effect do they have?

Important unsolved mathematical problems include:

(a) globalize the existing local theories,
(b) extend the range of non-elementary catastrophes that are well understood,
(c) apply the methods to bifurcations to tori,
(d) deal with Hamiltonian systems,
(e) deal with chaotic dynamical systems, and
(f) deal with forced oscillations and subharmonic bifurcation.

The area is one of extensive mathematical research activity.

RELATED THEORIES

Several alternative approaches to a similar range of problems exist, in particular:

(a) topological dynamics
(b) bifurcation theory
(c) synergetics
(d) non-equilibrium thermodynamics.

Catastrophe theory is not in competition with these: it is an attempt to add a new viewpoint and to supplement (not supplant) them. Connections between all of these theories are extensive and deserve to be examined systematically and thoroughly.

APPLICATIONS

Broadly, the applications fall into two distinct classes, each with its own methodology:

(a) physical science
(b) social and biological science.

[For example, see Thom. <u>The two-fold way of catastrophe theory</u>.]

In physics, the starting-point in an application will be an existing mathematical theory (such as a partial differential equation) which is analyzed in a chosen setting (quantum mechanics, chemical kinetics, wave optics) using catastrophe theory as a mathematical tool. The catastrophes are deduced directly from existing models and theories.

In social and biological science, catastrophe models are hypothesized on the basis of empirical observations. Sometimes the process is more deductive (game theory, population dynamics), especially in parts of biology (ecology, biochemistry) that more closely follow the physical paradigm; but the area in general is not already set up along deductive lines, and catastrophe theory can hardly be expected to make deductions from a non-existent basis.

Physical applications are surveyed in:

Berry & Upstill. <u>Catastrophe optics: Morphologies of caustics and their diffraction patterns</u>

Guttinger & Eikemeier. <u>Structural stability in physics</u>

Poston & Stewart. <u>Catastrophe theory and its applications.</u>

Stewart. <u>Applications of catastrophe theory to the physical sciences</u>

Stewart. <u>Catastrophe theory in physics</u>

The methods are successful and sound.

SOCIAL SCIENCE

As is widely recognized, in this area methodological problems differ from those in physics. Catastrophe theory does not alter this general fact, although it suggests some potential for improvement.

The majority of models in the existing literature are qualitative and speculative: "What if the world were this way?" This is the result of historical factors rather than an unavoidable feature of catastrophe modeling, and an increasing number of workers are taking the next, and crucial, step of performing well-designed experiments and testing their theories against data. The total number of such papers, however, is not large.

The modeling process even for elementary catastrophes raises many unsolved problems, and there has been little attempt to progress towards even more complicated non-elementary catastrophe models. This appears sensible: it is best to learn to walk before attempting to run. The mathematical background can thus be considered as "understood" as regards setting up the model and deriving its consequences.

The main problems are the standard ones in social science:

(a) definition of variables
(b) measurement of variables
(c) presentation of data
(d) analysis of data
(e) design of controlled experiments.

Of these, (a),(b) and (e) yield to existing techniques; (c) poses new problems because the structure sought for is nonlinear, and it is difficult to visualize multidimensional nonlinear forms; and (d) raises many new problems in nonlinear data-fitting. Some progress has been made by Cobb (1980a, 1980b, 1980c), enough to allow suitable models to be pushed through to satisfactory conclusions.

Relevant literature include:

Cobb (various papers)
Colgan, Nowell, & Stokes (1980)
Seif (1979)
Smith, J. Q. (various papers)
Stewart & Peregoy (in press)
Zeeman (1977)

The modeling of administrative processes may be considered a typical part of the general area of social science, and the above remarks apply specifically to it. Some of the more developed catastrophe theoretic models involve problems bearing a close resemblance to topics of concern in administration, and these may be used to assess the prospects.

CONCLUSIONS

The reasons for wishing to use catastrophe theory are not that it provides a universal classification of sudden changes: the belief that this is the case stems from a misunderstanding of expository material aimed at the layman. But it does offer a well-defined and natural range of models of discontinuous processes, with a good mathematical pedigree, of proven value in physical science, and appearing to have some potential in social science by virtue of the following features: structural stability (robustness) and genericity (avoidance of the unnaturally special). Its qualitative generality is also an advantage since in social science it is unwise to place too much reliance on a specific system of equations or on an over-detailed construct.

The objectives of current research, and the needs of future research, include:

(a) identification of appropriate areas for modeling by catastrophes;

(b) design of models in appropriate settings;

(c) design of experiments, collection of data, and testing of data (more studies will be needed to gain enough experience);

(d) improvement of statistical methods for analyzing nonlinear models and data; and,

(e) provision of theoretical underpinning for successful empirical models, aiming at a more deductive overall structure. But this must be a long-term aim except in special cases.

These objectives involve expertise in a wide range of topics: abstract mathematics, statistics, and numerous branches of social science. It will be necessary to assemble teams of qualified researchers to achieve an adequate breadth and depth of expertise.

CONTRIBUTED PAPERS

CATASTROPHE-THEORETIC AND DYNAMICAL SYSTEMS MODELING IN SOCIAL SCIENCE: STATE OF THE ART AND PROSPECTS

Ian Stewart

****

Mathematical modeling in social science goes through a number of stages, including the construction of the model, experimental testing and data-collection, statistical analysis, and evaluation and modification.

The article examines these processes first in a general context, and then more specifically with reference to catastrophe theory and (to a lesser extent) dynamical systems. Emphasis is placed on the new problems and possibilities raised by nonlinear models.

There follow three case histories of work that (unusually in this area at this time) have passed through the main stages of experiment and evaluation. These are:

A. the work of Seif (1979) on hypothyroidism
B. the work of Colgan, Nowell and Stokes (1980) on territorial fish
C. the work of Poston and Stewart (1978b), Peregoy and Zeeman (in press) on perceptual multistability.

In each case it is shown that, as a result of the data-fitting exercise, new ideas and further topics for study have emerged which could not have been anticipated without the original model.

Finally, some advantages and disadvantages of catastrophe models are discussed, and suggestions for future work are listed.

****

The aim of this paper is to examine the general processes of mathematical model construction and testing in connection with the social sciences, as these relate to models based on dynamical systems theory and catastrophe theory. A general discussion of the problems and prospects is followed by a representative but not exhaustive series of case histories; a final summary is aimed more specifically at the modeling of administrative processes.

I. THE MODELING PROCESS

In most branches of science, the development and testing of a mathematical model will go through (at least) the following stages:

A. selection or construction of a model
B. acquisition of data
C. presentation of data

D. comparison of theory with experiment
E. adjustment of model
F. extensions and further implications

The process may be repeated several times until a satisfactory theory results or until the effort is abandoned.

Let me consider each stage in detail.

(A) SELECTION OR CONSTRUCTION OF A MODEL

Until fairly recently, as far as the physical sciences were concerned, the selection and construction took this form: "write down the equations." "The equations" would almost always be some system of ordinary or partial differential equations which supposedly described the behavior of the physical system being studied; and the object was to solve these equations for a given set of initial or boundary conditions.

As mathematical models began to be introduced into other areas of science, such as biology and sociology, and as the range of mathematical models required in physical science grew, many other mathematical structures (such as groups, singularities, probability distributions, and networks) began to be used. Thus the process now takes two steps:

1. Decide on an appropriate mathematical setting.
2. Select or build a plausible model within that setting.

The role played by Step 1 should not be underestimated: a different choice of setting, even for the same phenomenon, may suggest a radically different model. For example, a continuum-mechanical model of fluid flow will take a completely different form from a finite-element model, and require different mathematical techniques.

Holt emphasizes the importance of getting the setting right before embarking on the construction of a model in his contribution to this study: "Getting the mathematical structure right is, at the early stages, as important as getting the numbers right." (page 40 of this report). I agree.

(B) ACQUISITION OF DATA

That is, experiment. Substages include:

1. definition of variables
2. measurement of variables
3. design of experiment
4. elimination of unwanted outside effects

It is worth remarking that the choice of a model conditions the selection and measurement of data. The "wrong" model may suggest data that do not reveal the essential phenomena. In particular, any smoothing or averaging of data may conceal multi-valued responses or other nonlinear relationships.

(C) PRESENTATION OF DATA

I mention this because it tends to be ignored, but can be very important, especially when more than two or three variables are involved or when the relation between them is nonlinear. A mass of numerical data is seldom very informative. It is necessary to find effective ways of presenting the data to make it possible to see the general structure involved.

(D) COMPARISON OF THEORY WITH EXPERIMENT

In the social and biological sciences this generally involves the application of statistical techniques to elucidate the important relationships between variables, or even to decide which variables or groups of variables are the most important. It may require suitable transformations of the variables before these methods can be used.

(E) ADJUSTMENT OF MODEL

Usually the model will involve adjustable parameters which must be chosen to give the best fit with experiment. Some measure of the goodness-of-fit is required.

(F) EXTENSIONS AND FURTHER IMPLICATIONS

A useful theory will suggest further conjectures, models, problems, experiments. A theory that explains only the data around which it was built is not especially useful!

II. CATASTROPHE-THEORETIC MODELS

Next, I want to take a look at each of these stages when a catastrophe-theoretic or dynamical systems model is involved. For simplicity I will talk of catastrophe models, but the problems and differences remain very similar if a dynamic is present, except where the detailed design of an experiment or statistical methods of testing the model is concerned. In these latter cases, it is hard to see how progress can be made except by first treating the simplest "elementary catastrophe" cases. In fact, I would suggest that a major reason for concentrating on the relatively simple catastrophe models is that they provide a good testing-ground for this style of modeling. They are admittedly far more simple in structure than might be desirable in some contexts, but they raise questions which ought to be tackled first in a relatively simple setting. If they cannot be dealt with successfully in the simple setting, the prospects for more complex models cannot be very good.

A. SELECTION OR CONSTRUCTION OF A MODEL

The traditional models used in social science are overwhelmingly linear as regards the underlying mathematical structure. Statistical techniques, such as linear regression, correlation, cluster analysis, and factor analysis, rest on the often unstated assumption that a significant relationship between variables must be a linear one. Similarly, game theory is largely linear in character, and linear differential or difference equations are frequently used in preference to nonlinear ones.

A major reason for this emphasis on linearity is not (as one would hope) that Nature tends to be linear. On the contrary, Nature is increasingly recognized as being highly nonlinear. Instead, linear methods are used because nonlinear mathematics is usually much harder than linear mathematics.

In the absence of any effective techniques for dealing with nonlinearity, this approach can be defended on pragmatic grounds. But, as nonlinear mathematics increases in power, it would seem advisable to consider putting it to good use. Of course the greater simplicity of linear methods has much to recommend it, but if a choice must be made between a linear model that is computable but wrong, and a nonlinear one that is harder to compute but might conceivably be right, then presumably the latter suggests itself more strongly. Of course there are many areas of statistics that deal with "nonlinear" relationships. But most of these are really just linear methods in disguise, coming in three stages:

1. transform the variables somehow
2. seek a linear relation between the transformed variables
3. deduce a nonlinear relationship between the original variables.

For example, suppose that two variables x and y are related as follows:

x	-5	-4	-3	0	3	4	5
y	0	3	4	5	4	3	0

Then the correlation coefficient is zero. That is, there is no (linear) relationship between them. But in fact there is an exact functional relationship

$$y = (25 - x^2)^{1/2}$$

If we transform the data from x and y to $X = x^2$ and $Y = y^2$ then the correlation coefficient between X and Y becomes -1: perfect negative correlation. If instead we transform from x to

$$Z = (25 - x^2)^{1/2}$$

then we get a correlation coefficient of 1: perfect positive correlation. So the same data, depending on the choice of transformation, are unrelated, positively related, or negatively related!

What this shows is that linearity is crucial to the interpretation of a correlation coefficient. But it also shows that undue reliance on linear models might cause significant relationships to be overlooked, especially when there is a tendency to put data on a computer and analyze it using a standard linear regression package, rather than taking a serious look at the shape the data takes.

In fact, consider any functional (i.e. single-valued) relationship

$$y = f(x)$$

Transform the independent variable x to

$$X = f(x)$$

Then

$$y = X$$

and there is a linear relationship between the transformed variables.

Let me call any relationship, such as this one, which is a transform of a linear relationship, "pseudolinear." Then the majority of "nonlinear" models and methods used in statistics and social science turn out to be pseudolinear.

The crucial feature of pseudolinearity is that the dependent variable y is single-valued in terms of x. For each value of x there must correspond a unique y.

The types of relationship considered in catastrophe theory are in general multi-valued: there may be several possible values of y for a given x. These models cannot be pseudolinear: they are genuinely nonlinear.

The experience of physics and biology is that Nature is quite often genuinely nonlinear. (I mean by this that it appears to require genuinely nonlinear models: nonlinearity is a property of mathematical systems, not of Nature.) If, as seems reasonable, this also holds good in the social sciences, then there is a need for methods that can handle genuine nonlinearity.

Curved things are much more complicated, and diverse, than straight things. The possible range of nonlinear models is so vast that some guiding principle is required to select significant types of models. The principle of parsimony--choose the simplest--presents itself. But what do we mean by "simplest"? The catastrophe theory answer is that we should avoid unnecessary "accidents" and use generic nonlinearities. In the context of equilibrium models, this means "assemble the model from the standard catastrophe surfaces." In many cases, the principle of parsimony can then be invoked to reduce the model to a single catastrophe form. In other contexts similar reasoning should apply, although the appropriate class of "natural" models will vary from one setting to another.

None of this proves anything nor is it meant to. The process of model-selection is necessarily heuristic. But the catatrophe theory approach is a natural one, based upon a proper understanding of the mathematical possibilities and their relative simplicity or complexity.

The importance of the Thom Classification Theorem is not that it "describes all possible types of discontinuous relationships." It does not. What it does do is demonstrate that certain types are of more fundamental importance than others. In the search for appropriate models, it is not unreasonable to try out these fundamental forms. [See Brocker & Lander, 1975.]

To summarize: Catastrophe theory has definite potential in the area of model construction because it provides a versatile and flexible range of nonlinear models which are simple and natural from the mathematical viewpoint. This is one of the points that Markus makes in his paper.

B. ACQUISITION OF DATA

The problems of defining and measuring variables in the social sciences are well known. They derive from the difficulty of performing controlled experiments as much as from anything else: in the physical sciences, historically, definition and measurement have also proved difficult (even as simple a variable as time has caused tremendous problems, and heat and temperature were only treated successfully within the last century). While it may be true that catastrophe models do not make these problems easier, it is certainly true that they do not make them any harder. Although certain models in the catastrophe theory literature make no serious attempt to face these problems, this does not imply that they cannot be faced; merely that the authors were not at that point concerned to face them. (In expository articles, particularly, this is not especially surprising.) Existing methods will be adequate, or not, depending on the topic under discussion and the selection of variables treated, not the particular mathematical theory being used.

It is at least conceivable that catastrophe theory might be able to ease some problems of definition and measurement. For example, the models may suggest natural choices of variables. The multi-valued aspects may reveal relationships, hence significant variables, that would otherwise be missed. I know of no strong examples where this has actually occurred, and I wouldn't want to emphasize it too much, but it is definitely a possibility.

There may also be variables that depend on a catastrophe model for their very definition, for example, the coordinates of a cusp point or the size of a catastrophic jump. It is worth bearing in mind that correlation coefficients and regression coefficients depend similarly on a previous choice of model--in their case a linear one. They are not items which an experimenter would naturally think of measuring. One of the functions of a model may be to suggest useful quantities to be calculated; and it is again conceivable that the range of nonlinear models suggested by catastrophe theory might suggest fruitful new quantities to describe or detect important aspects of nonlinearity--just as the correlation coefficient describes an important aspect of linearity.

There is one additional feature of catastrophe theory that is often misunderstood: its ability to reduce the number of variables in a problem. Thom's Theorem says that, subject to certain technical hypotheses, with any number of state variables, under the action of up to four control variables, the typical behavior may be reduced to one of the standard models, having at most two state variables. This is a genuine reduction of complexity; the limitation to one or two state variables is a deduction from the mathematics, not a hypothesis of the model.

It is, for instance, futile to criticize the model by Zeeman (1977, p. 3-8) of aggression in dogs, which uses a cusp catastrophe, on the grounds that this has a 1-dimensional state variable, whereas aggression is too complex to be measured by a single variable. This is doubtless true; but what catastrophe theory says is that even if one starts with, say, a thousand-dimensional set of variables that define the aggressive behavior, then, provided only two control variables are effectively operating, the model necessarily reduces to one in which some single combination of those thousand variables bifurcates according to the standard cusp geometry. Indeed any "random" combination of the thousand variables should suffice. So when Zeeman draws a single arrow and labels it "aggression," he is not claiming that all the complex details of aggressive behavior can be measured by a single variable. He is merely using the standard topologist's convention of drawing a multi-dimensional space as if it were 1-dimensional, or else abusing language slightly by labelling the reduced 1-dimensional variable "aggression" rather than something like "reduced measure of aggression."

Again, the practical importance of this kind of reduction is less than clear. Conceptually, any reduction of dimensionality is extremely useful, but for practical experiments it would be desirable to exhibit the reduction explicitly. The real problem is to reduce the number of control variables by holding as many as possible constant. The inability to do this effectively is a major source of "noise" in experiments. In physics, one can purify copper, produce a uniform wire, and hold its temperature constant to discover Ohm's Law. But no one has yet been able to purify a dog, let alone draw it out to uniform thickness or hold its temperature constant.

C. PRESENTATION OF DATA

It is surprisingly difficult to visualize multi-dimensional data. If points are plotted in the plane then the eye can often "see" whether they appear to lie on a curve, but it is already difficult to see whether points in 3-space lie on a surface, and it takes an unusual mind to be able to see whether points in 4-space lie on a 3-dimensional submanifold. In fact it is not at all easy to find an effective way to present data in 4-space.

One possible way to make progress is to guess at the most likely shape for the surface (if there is one) and then to try to transform the data in a way that will permit such a surface to be located. For example, if one suspects that the data lie on the surface of a sphere, then it is possible to select a candidate for the center point, compute distances to the data points, and see whether these are roughly equal.

This method relies on making the right initial guesses, but it appears to be orders-of-magnitude more difficult to seek a surface of unknown form. Since catastrophe theory suggests a specific range of types of surface, it may sometimes assist in the presentation of multi-dimensional data.

D. COMPARISON OF THEORY WITH EXPERIMENT

Despite the cautionary remarks made by Savage, he acknowledges that catastrophe models do pose interesting questions for statistics, especially as regards multi-valued relationships or complex nonlinearities. In fact I would suggest that catastrophe theory has an important contribution to make to the development of statistics.

Firstly, as I have remarked above, traditional statistics is largely linear (or pseudolinear, which is no more adequate in many cases). The usual goodness-of-fit or confidence tests, based on the normal probability distribution, also possess this (pseudo) linearity. From the qualitative point of view, the important feature of the normal curve is that it is unimodal: it has only one maximum. It represents the deviations of a variable about a single most-likely value.

Variables that undergo statistical fluctuations about several likely values are much more rarely treated in statistics, but the experience in physical science is that such variables are likely to be relatively common. Genuine nonlinearity requires multi-modal distributions. Catastrophe theory requires that such questions be tackled systematically, but it also assists in this by ensuring that a limited range of basic types of nonlinearity naturally present themselves for initial treatment. Considering how much work and how many useful statistical ideas derive from the normal distribution, it is likely that generalization to suitable multi-modal distributions will be a very fruitful direction for research. The work of Cobb (1978) takes some useful steps in this direction, but much more remains to be done.

Cobb in particular can deal only with the "cuspoid" catastrophes, of co-rank 1. A student of mine, Alexander Galis, has derived some extensions to umbilic catastrophes, and indeed, to arbitrary co-rank, but his work is in a very preliminary form. However, I see little harm in concentrating on, say, just the cusp catastrophe in initial research. Before decrying this as a much too limited aim, it is wise to recall the vast effort expended on the normal distribution, which from the qualitative point of view is even more limited than the cusp. I believe we should learn to walk before we try to run, and I suspect that many attempts to assess the likely role of catastrophe theory models slip up by asking for too much, too early. Savage would like to see a theory of the statistical estimation of a dynamical system. So would I. (Indeed I think we need more: a coherent method for estimating the simultaneous properties of an entire family of dynamical systems, undergoing bifurcation.) However, there isn't yet even a good statistical theory of equilibrium states of dynamical systems or of their bifurcations. By all means let us bear in mind that what we are trying to do is but one step towards something that is obviously far more versatile, but let us also avoid asking for the moon.

E. ADJUSTMENT OF MODEL

It is not enough merely to find a model that fits data fairly well. One would like to find one that does this as well as possible. Catastrophe models come with several adjustable parameters--indeed, one can adjust parameters indefinitely by making changes of variables. In practice, however, these changes have to be of fairly limited form, and the initial data-fitting can be done by assuming that changes in control variables are linear. (This may sound curious, given my criticisms of all things linear, but the point is that the theory says that the relevant coordinate changes are smooth, and hence, approximable by linear mappings.)

Adequate measures of goodness-of-fit of nonlinear models pose, I feel, a difficult but crucial problem. It is not a problem that can be solved by routine application of techniques developed for linear models: the whole phenomenology is completely different. For example, imagine data consisting of four points at the vertices of a square. These fit exactly any of the following nonlinear models:

a circle,
a square,
two diagonals crossing at the center,
two parallel sides of the square,
the other two parallel sides of the square,

and many other alternatives can be invented. So which fits best? I doubt that this is a meaningful question, as it stands. It cannot be decided just by looking at the existing data. It seems to require the collection of more data to distinguish between the plausible alternatives.

It should be much less hard to devise measures of goodness-of-fit within a particular special class of models. In practice, that is what happens now with linear models: one measures how well the model fits in comparison with other linear models. Similarly, one might be able to say of a cusp model that it fits better than other cusp models or even that it also fits better than any linear model. But it is unlikely that it will be possible to assert with confidence that it fits better than any other conceivable nonlinear model might. This is simply asking too much: no scientific theory yet has been able to prove that no other possible theories can be better.

It is, therefore, very important indeed to select an appropriate class of models in advance of detailed data-fitting. I suspect that the only way to do this, in practice, will be by acquiring experience from such an exercise. This is how most practicing statisticians work: They have experience of a standard range of models which suggests to them which one might be appropriate. Catastrophe theory models, or any other additions to the existing range, will require similar experience. At the present time, experience is very limited indeed, and the only remedy for that is to make repeated attempts to fit catastrophe models to actual data and to observe what happens. In order to get a good feel for this kind of data-fitting, it would not be a bad idea to concoct "fake" data, known to lie on (say) a cusp catastrophe surface or be drawn from some family of multi-modal distributions, and subject it to independent analysis using statistical procedures intended to test for such things. I would also like to suggest a parallel and somewhat iconoclastic numerical experiment: subject similarly concocted nonlinear data to the standard procedures of factor analysis, cluster analysis, and regression. The results might well be surprising . . .

F. EXTENSIONS AND FURTHER IMPLICATIONS

Critics of catastrophe theoretic modeling have often asserted that it is of no advantage to know that a catastrophe theory model fits the data, because if it does, it is simply restating the data itself. "Even if it looks like a catastrophe surface, this tells you nothing you didn't know already."

I don't find this argument very compelling for a number of reasons. One is that it seems to apply to any exercise in empirical data-fitting. Kepler's ellipses told him nothing he didn't know already. Mendel's observation of genetic proportions told him nothing; indeed worse than that, it merely gave him a bad approximation (1:2:1) to data that were doubtless far more accurate (1001:2001:999 or whatever). Ohm's Law (voltage is proportional to current) says no more than the set of data points (lying on a straight line) that inspired it.

But the point about Ohm's Law is that while it says nothing apparently new about a single piece of wire, it applies to all pieces of wire; and it suggests a new concept: electrical resistance. One may then ask: "How does the resistance vary?", either from one substance to the next or with temperature or with the state of the weather in Honolulu. Similarly, while the fact that aggression in a particular species of fish fits a hysteresis loop may tell us nothing we didn't know when we first wrote down the data, we cannot ask "How does the size of the hysteresis loop vary from one species to another or one individual to another?" until we have recognized that there is a loop to be talked about.

In other words, what data-fitting can do is to suggest that a particular mathematical structure is appropriate. That structure will itself raise questions. What makes orbits elliptical? Why do genetic proportions look like the top end of the Pascal triangle? Is something combinatorial going on? It is not necessary to agree with Dirac's remark to the effect that "it's more important whether your equations are beautiful than whether they agree with experiment" to appreciate that finding a good mathematical setting is often a necessary step towards properly understanding what is going on.

More specifically, however, even the current rather limited experience of fitting catastrophes to data shows at once that this exercise does lead to new ideas that, while no doubt implicit in the data, are not sufficiently clear in the data to attract any attention. (I have little sympathy for those who argue that if A logically implies B, then B tells us nothing that A didn't. In particular, we only can say this after the logical dependence of B on A has been established. To put it another way, implicit results aren't much use until someone makes them explicit.) The case histories that follow this section provide concrete instances to supplement the above generalities.

III. CASE HISTORIES

This section will give brief descriptions of three attempts to apply catastrophe models and to illustrate how these led to further ideas, theories, or experiments.

A. HYPOTHYROIDISM

This work is due to Fritz Seif (1974) at the University of Tubingen. In a normal individual the brain controls metabolism through a hormone chain

$$\begin{tikzcd} \text{HYPOTHALAMUS} \arrow[r, "a"'] & \text{PITUITARY} \arrow[r, "b"'] & \text{THYROID} \arrow[r, "c"'] \arrow[l, bend left=60, "\text{negative feedback}"] & \text{METABOLISM} \end{tikzcd}$$

where the hormones are:

a = protirelin
b = thyrotropin
c = thyroxine + triiodothyronine

Hypothyroidism occurs when too little c is produced; **hyperthyroidism** occurs when too much c is produced.

To treat hyperthyroid patients, some of the thyroid gland is removed surgically to reduce the production of c.

However, after such treatment, about one patient in three begins to display symptoms of hypothyroidism (call these "treated" patients) and the other two-thirds return to normal. This effect is due to the pituitary losing its ability to respond to a. Seif devised a measure, X, of this response-ability and noticed that on data from 422 measurements, X is bimodally distributed for some values of (b,c) but not for others. (It is interesting to conjecture what would have happened had he merely shoved the data through a standard linear regression package.) This suggested to him that a cusp surface model might be appropriate, with (b,c) as control variables and X as state variable. This in turn suggested a cure: inject b until the state X undergoes a catastrophic jump, then stop the injections, and let the patient return to normal. This cure worked. (Note that at this point the model was entirely qualitative and no data-fitting had yet been done except very roughly, by eye.)

However, Seif went further. He fitted data quantitatively by an iterative least-squares method, obtaining excellent results. He then devised a micro-model of the behaviour of the pituitary to explain, more dynamically, why such a cusp surface might arise.

B. TERRITORIAL FISH

Zeeman (1975, p. 13-14) suggested, starting from observations by Konrad Lorenz (1967), that the aggressive behavior of territorial fish should involve, among other things, a hysteresis loop in the level of aggression plotted against the distance of the invader from the nest. That is, the fish has two different radii of territory: the defense radius RD (to which distance another fish must approach before being attacked), and the attack radius RA (beyond which it must retreat before the attack ceases).

Colgan, Nowell and Stokes (1980) tested this theory using a wooden dummy fish in Lake Opinicon, Ontario, to mount fake attacks on Pumpkinseed Sunfish. They observed a distinct hysteresis effect with RD being about 13 cm. and RA about 18cm.

They also realized that the existence of two radii had important consequences for the geometry of the packing together of nest territories in the lake. With only a single perimeter, territories would evolve until they were packed rim-to-rim. Any fish attempting to move around the lake in search of food would spend all of its time fighting one battle after another with the enraged property-owner.

With the double perimeter, however, territories would be arranged so that the attack perimeter of one would coincide with the defence perimeter of the other, leaving a "demilitarized zone" between the nests, through which any fish might pass unmolested. This was a novel concept in the subject area. It had not been obvious to anyone until the theory had been confirmed and its implications began to be studied seriously.

C. MULTI-STABLE PERCEPTION

I will now describe some work done jointly with Tim Poston and (at a later stage) the psychologist Peter Peregoy, on the perception of ambiguous figures. The work with Poston was published in Behavioral Science (1978b); the work with Peregoy has recently been accepted by Psychological Bulletin (in press); and a further study by Peregoy and Zeeman is in preparation.

There are certain standard drawings that people can perceive as representing more than one object. The particular drawing in our work resembled both the face of a man and the figure of a kneeling woman. Hysteresis effects in the perceptual process had been observed by researchers who devised a whole sequence of drawings in which features were selectively biased to make the result more manlike or more womanlike. Poston and I suggested that a cusp catastrophe model might be suitable, and if so that it would be possible to devise a 2-dimensional array of pictures that would give the whole cusp surface.

Poston and I are both mathematicians and we did not have the expertise to test our ideas experimentally. But recently, Peregoy and I (in press) carried out a series of experiments using both British and American students. We drew the array of drawings on separate cards and asked students to rate these on a scale from -5 to +5. We used a version of Cobb's computer program to find the best fitting cusp catastrophe surface, obtaining a good fit with the data and one that was better than any linear model. When the group of subjects was broken down by categories, small variations in the cusp position emerged, but its position was remarkably consistent.

In these experiments, the drawings were shown in a random pattern. We decided to test the "delay rule" of catastrophe theory by presenting drawings in sequences and observing the responses (which, according to the convention, can now be predicted and are single-valued; the value, of course, depends on the choice of path through the control plane). We found that on the first pass along a sequence, the delay convention appeared to be operating; but even on a second (reversed) path, the subjects typically had learned to anticipate the change in perception. They jumped not on the catastrophe bifurcation set, as predicted by delay convention, but close to the Maxwell set. In fact they appeared to jump sooner than the Maxwell set, anticipating the change slightly before it really happened. [For a further discussion of the delay rule and Maxwell's rule, see Zeeman, 1977, p. 306-310.]

Thus there appears to be a learning process involved whereby the subject's choice from the conflicting stimuli is conditioned by his prior experience in a similar situation. It is as if the original random selection from the two possible percepts is replaced by a more considered judgement, based on previous experience and recent history. A theory of decision-making devised by Zeeman some years ago involves precisely this type of effect, and Peregoy and Zeeman are currently adapting it to the perceptual context.

Without the original catastrophe surface model, it would not have been likely that anyone would have tested the convention whereby subjects jumped from one surface to another; because it would not have been clear that there were any surfaces there to jump between.

IV. MODELS IN ADMINISTRATION

Finally, let me turn to the main purpose of this report. While a fairly large number of catastrophe theoretic models have been proposed in connection with social science, the whole subject is in an early state of development. Few models have been adequately tested against data, and many are "off-the-cuff" suggestions by researchers new to catastrophe theory or by mathematicians new to the social sciences. In such circumstances enthusiasm can outweigh due caution and novelty alone may provide a reason for publication. However, the subject cannot be permitted to remain in such an immature state indefinitely.

It would not be possible at this time to assess the likely role of catastrophe theoretic or dynamical systems models in administration by restricting attention entirely to that field. The relevant body of work is not sufficiently large. In any case, the nature of this rather interdisciplinary exercise means that experience in one area of application tends to carry implications for other areas.

It seems wisest to judge prospects by looking at those studies that have carried the modeling process through to a proper conclusion, at least up to the point of obtaining data and comparing theory with experiment. Several such models relate fairly closely to administration in that they study some kind of decision-making process. Of the case histories above, those on territorial fish and on perception fall into this category: The fish must make a decision when to start or cease his attack (it is the male Pumpkinseed Sunfish who defends the nest), and the experimental subject must decide between two conflicting and ambiguous perceptions of the stimulus materials.

The other case history concerns an attempt to influence the "natural" course of events in order to achieve a desired goal (curing the patient), paralleling an important administrative function. Therefore, two distinct types of model appear feasible and have been pursued with reasonable success:

A. understanding how decision-making processes act

B. understanding how to control a system which, in principle, may take up several states, so as to force it into the required state.

These two functions, comprehension and control, have obvious value.

V. ADVANTAGES

What advantages do catastrophe theoretic or dynamical systems models appear to offer?

Several of the participants in this study have emphasized the same basic points. The theory suggests a natural range of models for processes exhibiting typical features of nonlinearity, such as multiplicity of states, sudden discontinuities, hysteresis and other history-dependent phenomena, divergent behavior, and so forth. If I understand Savage's remarks correctly, it appears that a common reaction to such observed phenomena is to seek ways to eliminate them, such as, introducing additional variables to separate out a multiplicity of states. This has definite attractions but I doubt its wisdom. First, a major problem in the area is the influence of large numbers of variables, and it seems less than sensible to introduce new ones if that can be avoided. Second, if this approach had been tried in other areas of science (notably physics), most of today's theories would never have been discovered. The nonlinear phenomena listed above are to be expected; they are natural and commonplace. To treat them as pathologies to be avoided even at the cost of complicating the model is to misunderstand the way that a complex system is likely to function. To some extent the social sciences are trying to work within a paradigm of the physical sciences that does not, in fact, correspond to what happens in that area.

VI. DISADVANTAGES

Nonlinear mathematics is less familiar to most scientists than its linear counterpart. It tends to be more subtle and more difficult. Fundamental notions such as estimates of goodness-of-fit are hard to define.

To estimate the quantitative behavior of a dynamical system from empirical observations is a virtual impossibility unless the system is unusually simple or unless it can be observed repeatedly with different initial conditions. This is seldom possible in the social sciences. However desirable such a project may be, it is often going to be necessary to settle for much more partial information.

To some extent, this is why catastrophe theory (or more accurately, elementary catastrophe theory) concentrates on the attractors, that is, the long-term behaviour of the system or its equilibrium behavior. There is a greater chance of saying something definite.

It should not be forgotten that many traditional models also involve taking a far more limited view. When using linear regression to analyze a trend, it is seldom asked whether this trend has settled into some kind of dynamic equilibrium so that it will persist in the long run. Nor is it asked, "What dyanamic is holding the system to the trend line?" But few people fail to ask similar questions of a catastrophe theoretic model. I suspect that one of the roles of nonlinear modeling will be to reveal hidden assumptions in the traditional linear ones.

An additional disadvantage is the current lack of experience in using and testing catastrophe theoretic models and an almost total lack of experience in more general dynamical ones. This creates, among other things, misunderstandings. It can only be remedied in the longer term as experience accumulates. Currently, it is best kept under control by ensuring that any research project has access to workers with expertise in all the fields involved. There is a shortage of, for example, mathematicians with an interest in developing models in social science.

VII. FUTURE WORK

The advantages of nonlinear modeling via catastrophe theory or dynamical systems are manifest, provided the approach can actually be made to work. Some critics have claimed to be able to prove that this kind of modeling cannot succeed, but such extensive claims must be viewed with some skepticism (especially as many rule out useful applications to the physical sciences where the theory has been highly successful). On the other hand, the disadvantages should urge a degree of caution. The case histories above show that the problems are not insuperable. I do not expect to see any rapid or explosive development of nonlinear modeling in social science, but I do think that it is important in the long run to develop experience in such modeling because "Nature is nonlinear" and will not be tamed by a linear approach, however elaborate. (Witness the general failure of linear macroeconomic models.) Future work should, among other things, deal with the following:

A. obtaining a sufficient body of experience of nonlinear modeling in an adequate variety of sociological systems

B. paying proper attention to the novel problems raised in the design of experiments required to test nonlinear models

C. ditto for the statistical analysis of data. (Much theoretical work is needed on nonlinear statistics. In the absence of suitable field data it may be possible to conduct numerical experiments with "fake" data. The traditional methods might be tested in a similar way with data known to be derived from, say, a catastrophe surface; but field data would be preferable.)

D. encouraging contacts across the necessary disciplines and developing suitably expert teams of researchers

E. identifying topics suitable for catastrophe theoretic modeling in which the problems of defining and measuring variables can be dealt with, allowing the modeling process to be pushed through to a proper conclusion (positive or negative).

If the approach appears to be successful in enough cases, it will be necessary to extend the techniques to cope with the great variety of nonlinear and dynamic phenomena not reducible to the elementary catastrophes. Mathematically, catastrophe theory has matured well beyond the parts that are most commonly presented in nonspecialist expository articles. There are for example at least twenty different contexts in which the general point of view of elementary catastrophe theory can be pushed through to useful conclusions (catastrophes with symmetry, boundaries, in infinite-dimensional spaces, breaking symmetry, referring to periodic rather than steady-state behaviour, etc.) Work on applications will eventually have to reflect this breadth.

VIII. CONCLUSIONS

While catastrophe theory or dynamical systems models have not always been as successful in the past as the more exuberant expositions of the theory have tended to suggest, they have equally not been as unsuccessful as the more exuberant critics have indicated.

In view of this, it would be unwise either to support or to not support a particular research project purely because it makes use of (some version of) catastrophe theory or dynamical systems theory. As always, what should count is the competence of the people involved, their suitability for the project, and the potential importance of the topic.

Catastrophe theory is sometimes considered to be too simple to handle sociological modeling and sometimes too complex to be useful. While I can see that in principle both views might be correct, it might also be argued that it is treading, with some chance of success, the very fine line between what appears complex today and what will be simple tomorrow. If that is so, then its pursuit is amply justified.

It is pointless to predict the likely knowledge to be gained if the catastrophe theory approach works. Each individual application which has succeeded (even in limited terms) has shown its own unexpected twists and problems. But there are definite grounds for optimism that these techniques are worth developing, at least in some areas. Administrative and decision-making processes bear a close resemblance to some existing work that can be counted a reasonable success, and this suggests that it may be worth taking a closer look at those areas in the future, in particular by identifying suitable topics, not just for suggesting speculative models, but for experimental analysis and the collection of real data.

COMMENTARY

This is an excellent contribution interweaving three major themes:

1. a general discussion of mathematical modeling in the social sciences
2. a critique of the appropriateness of linear models, the reasons for their pervacity, and the suppression of modeling assumptions inherent in their use
3. an argument for an increase in the use of nonlinear methods in general and catastrophe theory in particular.

The discussion of mathematical modeling is clearly the work of someone who has thought through and worked on social science research problems. It is a timely discussion and would be a good candidate for replacing the standard sort of "hypothesis testing" blurb usually used as a basis for "research methods in the social sciences"-type courses (which usually only address nuts and bolts aspects of Stewart's second stage with all kinds of hidden assumptions about linearity).

Some thoughts on linearity are stimulated by Stewart's article. "Linearity," the existence of some kind of naturally occuring linear relationship among variables, is something that <u>cannot even make sense</u> in the social sciences. Scales in the social sciences are not universal, not based on naturally occuring constants, and essentially arbitrary inventions of researchers. The concept of linearity is defined only at the level of scale. Therefore, linearity is not in the class of properties that social science variables are allowed by nature to possess. In fact, one of the ways social science research can turn out to "look good" is for the reasearcher, through intuition, experience, and insight, to <u>choose</u> scales for variables in such a way as to end up with a linear relationship among scaled variables. This "choice of scale" should not be seen as unworthy. Indeed, this is often a contribution of great value. We should keep in mind, however, what is proved: the variables themselves are pseudolinearly related.

Stewart's term "pseudolinear" helps to make the point that scales are man-made objects in the social sciences. In a more mathematical discussion, pseudolinear is really a topological concept. A collection of variables (unscaled) would be <u>pseudolinearly related</u> if it is theoretically possible to scale each of them in such a way that the scaled variables are linearly related. (Scaled variables x, y, z are <u>linearly related</u> if there exist non-zero constants a, b, c and a constant A such that the equation $ax + by + cz = A$ holds while x, y, z vary. Linear is an algebraic concept.) Then "genuinely nonlinear" could be defined as "not pseudolinear," again a topological concept. The philosophical point seems to be that (social) science is more topological than algebraic in character.

Stewart's argument for the need to break away from pseudolinear models is compelling. His argument for settling on catastrophe theory as a replacement for linear models, on the other hand, is one that could be used to argue for keeping the linear ones: There is, in fact, no point in restricting social science to any class of models just because they are simpler and easier to deal with than the ones we are choosing to ignore. The point is, really, that elementary catastrophe theory represents the first step beyond pseudolinear theories in a topological hierarchy of complexity. Stewart is advocating walking before running, but he isn't saying we should never run. On the other hand, Stewart illustrates as well as possible (at this early stage of its use) that catastrophe theory can play an important role in modeling genuinely nonlinear phenomena in the social sciences. He also points out that catastrophe theory "contains" all pseudolinear relationships among variables, so that pseudolinear models are not being tossed aside. A point to make explicit is that, while nature is almost never linear, it is quite often pseudolinear, so in fact the research scheme

$$\begin{array}{c}\text{define variables}\\ \downarrow\\ \text{scale variables}\\ \downarrow\\ \text{look for linear relationships}\end{array}$$

is not destined to failure all the time. (The emphasis, however, should be shifted from the second arrow to the first.)

The point about scaling, and the issue of "linear vs. nonlinear," is largely irrelevant to catastrophe theory applications. If variables x, y, z,... are psuedolinearly related, then any one of them can be shown to be a (possibly nonlinear) single-valued function of the others. Catastrophe theory would offer a setting for modeling phenomena in which some measureable quantity is "determined by" certain parameters but is not a (single-valued) function of those parameters.

A way of expanding (pseudo)linear models goes as follows. Suppose we have "Theory X" that states the observed phenomonon P is determined by parameters a,b. (A natural presumption to make at this point is that Theory X states P is a function of a,b. This is the assumption of pseudolinearity!) Experiment or scaling/measurement, however, produces evidence that P may be in different states for the same values of a,b. That is, P is not a function of a,b; P is multivalued. There are two ways to fix the situation:

1. Find more parameters upon which P may depend (thus contradicting Theory X and/or casting doubt on the experimenter's ability to control unwanted parameter variation).

2. Re-examine the phrase "is determined by" in Theory X. Perhaps the phenomonon P can be observed only when certain utility (or probability, or energy) functions are maximized (or minimized). For example, it could be that each pair of values for a,b determine a utility function F(P), where P ranges over a continuum of posible manifestations, and the observed phenomenon is one of maximal utility. Thus, the observed phenomenon is determined by a,b to satisfy the

equation

$$\delta F/\delta P = 0.$$

Allowing a,b,P to vary, this equation defines a surface in a,b,P-space, and Theory X says an observable with parameter values a,b must lie on the surface. If the surface is "single-sheeted" over every point in the a,b-plane then we have verified pseudolinearity--observed $P = fcn(a,b)$. If the surface is "multisheeted," as experiments predicted, we have a genuinely nonlinear model to which catastrophe theory applies directly. Thus, neither Theory X nor experimental findings are contradicted.

APPLICATIONS OF CATASTROPHE THEORY IN THE SOCIAL SCIENCES: ISSUES IN A PHILOSOPHY OF INQUIRY

Robert T. Holt

****

This article begins with a "methodological paradox" concerning the difference between testing theories in social science and physical science and the difference between point predictions and correlations. The social sciences must find a different way to improve the process of theory evaluation since improved measuring techniques will not in practice solve the problem. The role of classification schemes is raised as one possible line of progress. What distinguishes a "good" (i.e., fruitful) classification scheme from a "bad" one? The suggested answer is that a good scheme "gets the right structure" for the problem, that is, finds an appropriate setting.

Catastrophe theory has the potential for developing "good" classifications because

(a) it provides structurally stable, i.e., robust models

(b) it is qualitative

(c) it makes predictions that, unlike correlations, are closer to the physicist's "point-prediction" paradigm, and in particular can easily be falsified if wrong.

****

Every scholar during his career encounters certain works he finds particularly cogent to his research and that affect its direction and development. Over 10 years ago I was struck by a trenchant article by Paul Meehl (1967) entitled, "Theory testing in psychology and physics: A methodological paradox." While concerned specifically with weaknesses in research design and measurement in psychology, I think its disturbing argument applies to much quantitative empirical work in the social sciences more generally. I shall argue in this paper that a philosophy of inquiry employing a mathematical system with the properties of catastrophe theory as an aid to description and classification can help overcome the weaknesses that Meehl identifies.

MEEHL'S PARADOX

Let me highlight the paradox by quoting Meehl's statement of it and then summarize his argument.

> In the physical sciences, the usual result of an improvement in experimental design, instrumentation, or numerical mass of data is to

> increase the difficulty of the "observational hurdle" which the physical theory of interest must successfully surmount; whereas, in psychology and some of the allied behavior sciences, the usual effect of such improvement in experimental position is to provide an easier hurdle for the theory to surmount. Hence, what we would normally think of as improvements in our experimental methods tend (when predictions materialize) to yield stronger corroboration of the theory in physics, since to remain unrefuted, the theory must have survived a more difficult test; by contrast, such experimental improvement in psychology typically results in a weaker corroboration of the theory, since it has now been required to survive a more lenient test. (p.103-4)

Meehl's explication of a paradox is detailed and complex. For the purposes of this paper let me present a simplified version which is, in effect, a special case of a more general and powerful argument.

Hypotheses in the behavioral sciences are typically relational. In the simplest case, a linear relationship is posited between two variables and is tested by stating the relationship as a correlation coefficient. (Sometimes the research design calls for a statement in the form of a null hypothesis.) A correlation coefficient is a number in the range from -1 to +1. If there is no linear relationship between the variables, the correlation coefficient would be zero. But there are an infinite number of points between -1 and +1. A correlation coefficient of zero is infinitely improbable. Thus, there is an *a priori* probability of 1 that any two variables will be linearly correlated and that any relational hypotheses will be confirmed (or the null hypothesis rejected).

Anyone familiar with the type of work that follows in the tradition outlined above would immediately raise two objections to what is but a caricature. First, "no relationship" between variables is not defined as a zero correlation; there is some range $\pm e$ that is treated as effectively zero, and there is, of course, a finite *a priori* probability that the value of the coefficient will fall in that range. The second point is that most hypotheses in the tradition that Meehl is dealing with are directional. They posit not only linear relationship, but a relationship in a specific direction.

The first objection gets to the heart of the Meehl paradox. Of course there is a range of values around zero that is treated as effectively zero. The greater the sampling and/or measurement error and the less demanding the research design, the larger is that range. Thus, improvements in research methodology allow the researcher to be more and more confident that correlations near zero are meaningful, and thus make it easier for hypotheses and the theories from which they are derived to be confirmed.

But if the hypothesis being tested is directional, then only coefficients of greater than zero or less than zero (but not both) serve as confirming evidence. This means that the *a priori* probability of any directional hypotheses being confirmed is only 50%. Those, however, are not bad odds but still present a dismal picture to one who demands tough standards for theory testing.

The actual situation is even worse. Suppose a scholar posits a certain relationship as being consistent with, if not actually derived from, a given theory. His empirical tests, however, provide evidence that there is a relationship opposite to the one he initially hypothesized. There are so many "theories" floating around social sciences that it is not difficult to develop a seemingly respectable *ad hoc* explanation of this contrary finding. The clever investigator can write an article expounding on the results of his research and the "theory" it supports. Meehl ridicules this specious scientific procedure in psychology. Think what his reaction would be to the scholar who collects masses of data, develops an intercorrelation matrix, and devises clever explanations of "significant" relationships.

In effect, then, even for the directional hypotheses there is an *a priori* probability of 1 that some theory will be supported by the finding, even though it may be different from that which gave rise to the initial investigation.

How do the natural sciences differ? Briefly, their hypotheses are typically point predictions: they predict a specific value. Empirical findings on either side of the point disprove the hypotheses. Elimination of measurement error and sampling error and improving research design reduce the size of the space within which empirical findings support the hypotheses, and thus, make verification more difficult.

I do not wish to linger over the paradox Meehl raises. It is certainly something to discuss and debate. At this time I hope you will agree that at least some of the theory testing in the fields in which we are interested suffers from the weaknesses he identifies. What can we do to improve matters?

Let us recognize at the outset that we cannot simply mimic the procedures followed in physics that he holds up as an ideal. Our theoretical formulations simply do not have the deductive power to produce hypotheses in the form of point predictions. It is to strengthen that deductive power that we must turn if we wish to eliminate the fatal weakness.

In physics that deductive power is largely the product of the statement of theories in mathematical form. The hypotheses to be tested are formally derived mathematical theorems. Rules of interpretation link the symbols in these mathematical statements to empirical reality and measuring instruments are calibrated in terms which are commensurate. While economics has moved a long way in this direction, theories and research in political science and sociology--particularly those which deal with phenomena that have long time constance like wars, revolutions, and social change--have not.

If the history of science provides any guide, there are a number of ways to proceed to increase the deductive power of our theories. One is to work on measuring instruments and quantification letting the theory follow the data generated by improved instruments. The path of progress would be from new and better measuring instruments to empirical generalizations from the data collected to stronger and more deductively powerful theories. Coleman (1975) has made a compelling argument that this is a promising way to proceed. I am, however, very skeptical about achieving any significant breakthroughs in measuring instruments for the study of social processes with long time constraints.

If improved instrumentation and measurement (i.e., quantification) is an unlikely first step towards hypotheses that yield point predictions, one might explore the opposite tack. One could attempt first to state principles as formal axioms and develop a rigorous theory through strict deduction. This task is far too mammoth to attack directly. If the axioms concerned fundamental human behavior (perhaps principles of social learning), then an enormous intellectual edifice would be erected by the time we had any theories related to processes for which the time constance is measured in decades or centuries. But furthermore, the axiomatization of scientific theories does not typically occur until a great body of reliable theory has been built up in other ways.

DESCRIPTION, CLASSIFICATION, AND GENERALIZATION

Let us turn to something more prosaic. Any handbook on methodology and theory building will have chapters on description, on classification, and on empirical generalization which emphasize the importance of each, but which rarely link them together. We are told that "meaningful" or "proper" classification schemes (like the taxonomies of Linneaus or Mendeleev) are important to generalization and theory building. But we are not given any criteria on the basis of which one can distinguish the "good" from the "bad" typology before a valid theory has been developed. If this is the case how do we know what is a good classification in the sense that it will facilitate tneory building and particularly theories with deductive power?

Consider the problem as one beginning with description. How would one describe a killer whale and a great white shark in such a way that one would classify the former as a mammal along with a bat, a human, and a giraffe, and the latter as a primitive fish? That is the kind of classification that facilitated theory building. But it is surely not the one which emphasizes obvious similarities.

When my colleagues and I attempted to describe the time paths of variables involved in the outbreak and ending of World Wars I and II, it was a classificatory problem which, among others, concerned us. Did the two wars belong to the same class of phenomena in the sense that both killer whales and Bengal tigers are mammals, or in the sense that both killer whales and great white sharks are large carnivores of the sea? If the former were the proper classification, then we would like to have a single theory which could explain both wars. If the latter is the case, we would accept the similarities as being relatively unimportant from a scientific point of view.

World War I started suddenly and ended suddenly, while World War II came on more smoothly and ended more smoothly. Perhaps even more interesting, historians who have looked for "causes" have emphasized the forces of nationalism (incompatible basic objectives), the arms race, and the tight system of alliances as leading to the outbreak of World War I. While they find nationalism a "cause" of World War II, they also see the absence of a tight alliance system and the failure of Germany's opponents to rearm aggressively in the 1930's as other causes. If the opposite of what caused World War I caused World War II, is it a fruitful classification scheme that would lump them together and, thus, call for a single theory to explain them both?

If one can describe two different phenomena in terms of the same mathematical structure, then we believe there is a high potential that a single theory can be developed to explain them both, even though obvious characteristics would suggest that they are very different. We attempted to describe the time paths of the relevant variables in the cases of the first and second World Wars in terms of the butterfly catastrophe. To the extent that we were successful, we believe we took a step towards building a deductively more powerful theory of war.

SOME EXAMPLES FROM THE HISTORY OF SCIENCE

The very fact that the description is done in terms of a mathematical system means that some significant generalizing has been undertaken. Indeed, the description becomes an empirical generalization. There are a number of simple examples of this from the history of science. Mendel described the offspring of two different purebred parents as occuring in the ratio of 1:1:2 (one like one parent, one like the other parent, and two hybrids). The actual number of smooth, wrinkled, and mixed peas he counted approached these ratios but were different. He generalized his findings to produce his law which is really a description of his findings in generalized form. It is that generalized description of the pattern that became the meaningful puzzle--it is the phenomena to be explained. The early theory of genes provides the explanation of the Mendel ratios. It is the theory of a mechanism at a more micro level that can generate the phenomena as mathematically described.

The descriptions of observed phenomena in terms of some mathematical structure is one way of generalizing. In the example used it also leads immediately to point predictions. But this latter is not the crucial point I wish to make here. Getting the mathematical structures right is, at the early stages, as important as getting the numbers right. Let me discuss another classic example to illustrate this point.

One of the most fruitful generalizations in the history of science was Kepler's portrayal of the orbit of the planet Mars as an ellipse. His establishment of the orbit as an ellipse, rather than as a circle, not only provided a generalization which accounted for all of the observations of the discrete points at which Mars had been observed from Earth, but it also provided the conceptual basis for explaining the structure of the solar system by locating the sun at one focal point of the orbit of each planet, and demanded a physical theory to account for the variable orbital radii of the planets. Kepler's achievement (which essentially laid the foundations for modern astronomy) has been termed the "finest retroduction ever made . . . . perhaps the most significant systematic hypothesis yet conceived" (Hanson, 1958, p.73). Beginning with accurate data on the positions of the planets, Kepler struggled for a period of over five years to discover a mathematical structure which would account for these observations.

Before suggesting how he arrived at his final generalization, let us speculate how many modern social scientists would have proceeded in the same situation. The typical social scientist would have connected the points with a smooth line and if mathematically inclined, written a polynominal which would have approximated such a curve. He probably would have hypothesized that the orbit of the planet was a circle, since the path of the orbit deviated only by 858 parts in 100,000 from being a perfect circle (Hanson, 1958). Proudly he would have proclaimed that he could account for over 99% of the variation in the orbit of the planet Mars and thus had established that it was a circle. But, he would have been wrong.

Kepler himself struggled with the circular conceptualization for several years before arriving at the ellipse as the proper curve. It is important to recognize that he was not simply motivated by a desire for greater precision but, instead, by a desire to specify the mathematical structure involved. Of course, we do not know Kepler's psychology or personal philosophical motivations for investigating, and finally accepting, the ellipse as the geometric description of planetary motion. However, it seems reasonable to suppose that the traditional mathematical studies of conic sections, as well as the special physical and geometric properties of the ellipse, must have guided Kepler's intellectual interest and his aesthetic judgment. In other words, pure mathematical research gave an impetus to this astronomical application in a manner that could not be predicted or defended by deductive logic alone.

Kepler was not engaged in induction. The shift from circle to ellipse entailed a momentous conceptual leap--from an orbit centered on a single focus (e.g., circles or ovids) to one centered on two foci; from a Greek perfect static form to a dynamical interchange of forces. Once the ellipse was recognized, it was a relatively simple task to go on to demonstrate that the orbit of every planet is elliptical with the sun as one of the two foci--thus specifying the geometric structure of the entire planetary system.

This example illustrates two critical points relevant to this discussion. First of all, as Hanson argues, it exposes the logic and the importance of inference by retroduction. Scientists do not search for deductive systems <u>per se</u> nor do they achieve fruitful results through induction by enumeration and approximation. The scientist is in search of explanations of data. "His goal is a conceptual pattern in terms of which his data will fit alongside better-known data" (Hanson, 1958, p.72). This is a process of inference by retroduction from observation to explanation. Typically, contemporary researchers in international relations do not proceed in this fashion.

The second point illustrated by Kepler's activity is the central import of achieving description and of establishing a classificatory schema as prior steps in the uncovering of mathematical structure to account for systematic behavior. Once the ellipse was presented as an astronomical reality, scientists could note the important fact that the ellipse is one of a class of curves defined by planar sections of a cone. The trajectory of a cannonball on earth is a parabola, another curve in the same class. The orbit of a comet is either an ellipse, a parabola, or a hyperbola, all curves in the same class. There is an important sense, then, in which comets, cannonballs, and planets, which appear to be very different phenomena, exhibit the same behavior. Newton was able to expand the law of inverse squares to account for all curves that are planar sections of a cone. His basic laws of gravitational dynamics which produce the trajectories of all three would have been more difficult to discover if earlier scientists had not developed generalizations that facilitated classification schemes for phenomena that were later shown to obey the same physical laws. Planets, comets, and cannonballs are governed by a single law; from the point of view of Newtonian mechanics they are in the same class. The differences in their individual behaviors can be accounted for by different initial conditions in the individual instances.

The typical social scientist tends to neglect these principles. By concentrating upon approximation and accounting for variation without having first adequately determined if the phenomena under examination are, or can be, classified together, the development of theoretical explanations is impeded rather than facilitated. Part of this difficulty stems from the paucity of well-articulated mathematical structures that are appropriate for the classification and description of social phenomena, including international violence. We are suggesting that catastrophe theory models, while they do not themselves provide a theoretical dynamic, do provide a very useful classificatory structure in which to construct a theory, just as the geometry of conic sections provides models in terms of which orbits and trajectories can be classified. While astronomy involves extremely accurate observations of a dynamical system following a deterministic evolution, behavioral sciences deal with much less precise data having a different sort of causal basis. These epistomological diversities are reflected in our choice of classificatory geometric models as we turn from the quantitative elliptic curve of Kepler to the qualitative butterfly surface of catastrophe theory. But the principle of describing empirical phenomena in terms of a mathematical structure remains the same.

THE POTENTIAL OF CATASTROPHE THEORY

Let us recapitulate the argument thus far. In the social sciences, better measurement, sampling, and research design make it easier (not more dificult) to confirm theories. Thus, the better our theory-testing research technology, the less confidence we can have in our theories. The problem lies not with the research designs and techniques themselves but with the absence of deductive power in our theories. Descriptions as genralizations done in terms of some mathematical system or structure will aid in classification decisions. These generalized descriptions, in turn, become statements of significant phenomena to be explained. But why catastrophe theory?

The first reason is because of the structural stability of the basic mathematical system.

The force of Thom's theory of catastrophes lies in a mathematical uniqueness theorem that holds that the elementary catastrophes are the only possible geometric configurations sufficiently robust ("structurally stable" in mathematical terminology) to be maintained in a real-world situation and not to be destroyed by random noise disturbances. (The context of gradient-like systems is part of Thom's hypothesis. See the comments of Stewart regarding popular misconceptions of elementary catastrophe theory.) In other words, if any predictable regularity is observed in a process of the physical, biological, or social sciences, and this observable is subject to multi-model behavior, radical divergences in development upon slight modifications of causal history, and even discontinuous jumps of observed behavior, then such a scientific process is most likely describable in terms of one of the elementary catastrophe geometric models. Of course, any such mathematical limitation on physical reality must be itself based on philosophical and mathematical assumptions, such as dynamical continuity, structural stability, and further requirements of simplicity.

To understand the significance of the structural stability of a mathematical system, it must be distinguished from the stability of an attractor. The mathematical system itself is stable if, when the system is perturbed, its qualitative properties remain invariant. The qualitative appearance of the parameterized phase portrait does not change as a result of an alteration in the parametrized functions.

What is the significance of this structural stability of a mathematical system for empirical research? It means the small errors in setting up one's model and errors in measurement and sampling will not distort the major features of the portrayal of empirical reality. Because it may be very difficult to set up the model of change when the time constants are long and when there are likely to be severe measurement problems, sturctural stability is a valuable property.

A second characteristic of catastrophe theory that we find attractive is its qualitative nature. While skeptical of many of the attempts at precise measurement in the study of war and long-term social change, we have much more confidence in qualitative comparisons. While hard pressed, for example, to come up with a precise measurement of the "tightness of the coalition structure" in the European state system in 1910 and in 1936, we are confident that it was tighter in 1910 than in 1936. We also have more confidence in answering the direction of change in some variable over long time periods than in specific measures at some point in time.

The ability to work with qualitative data in a structurally stable mathematical system allows us to be reasonably confident of our descriptions of reality. One should not, however, take this to mean that the descriptions are easy--that any findings can be fit into the model. In our study of World Wars I and II, for example, we described the time paths of five variables. Very few time paths would have conformed to the mathematical requirements of catastrophe theory.

Thus it would have been very easy to demonstrate that the butterfly catastrophe was an inappropriate mathematical structure for the description of the relevant phenomena.

While we have emphasized the use of catastrophe theory as an aid to description and have never claimed that it "explains" the phenomenon being observed (any more than the geometry of conic sections "explains" the orbital paths of heavenly bodies), the mathematics constrains the description so tightly that one can make necessary "if-then" predictive statements. For example, if our description of World Wars I and II is valid, then our theory of the international system must hold that the most peaceful states occur when each nation has a comprehensive treaty with every other nation (world federalism?) and that the highest states of international violence occur when no nation has an alliance with any other nation (the war of all against all?). These may sound like trivial statements that are just common sense. It is encouraging, however, when common sense emerges as a necessary conclusion from complex mathematics. There are other less obvious necessary conclusions. For example, when the level of arms is low and the coalition structure loose, minor differences will lead to localized hostilities. If the coalition structure is tight, major wars break out over smaller differences when the arms level is low rather than high.

These predictions are qualitative. They emerge from real world interpretations of regions within a topological space. They can best be interpreted comparatively. But the predictions do follow necessarily from the mathematics. Thus they have some of the desirable characteristics of the point predictions Meehl finds so valuable in the physical sciences. They would be easy to disprove.

We are making a limited but important claim for the application of catastrophe theory. We should be explicit in what we are not claiming. There is no substantive theory that is explicit. We have posited no dynamic. While we can use the theory to help characterize states of the international system, we have no dynamic by which states change.

The next step in our work involves developing that dynamic. This will be first a theory (also qualitative) of how the control variables interact with one another. If that can be worked out we will move to the more micro level of nation behavior to demonstrate how nations' actions give rise to the dynamic in the control variable. This lies in the future but is part of the continuing quest to increase the deductive power of the social science theories.

COMMENTARY

There is substantial agreement between Holt and Stewart. In particular, both emphasize the need to choose an appropriate setting before collecting or attempting to interpret data. A "bad" setting can introduce artifactual quantities, obscure actual phenomena, and suggest fruitless lines of research.

The selection of the "right" setting appears to be an art more than a science. It involves aesthetic and philosophical considerations.

Holt considers direct approaches to the problem of improving social models--improve the measurements and make the theories more deductive. He concludes that while both are desirable, neither looks feasible at this time and suggests trying something less ambitious but having some chance of success.

The analogy of the killer whale and the shark is a vivid one and the author's experience of a genuine piece of modeling--World Wars I and II--adds "practical" experience to the general discussion.

Particularly interesting is the discussion of Kepler's famous exercise in datafitting, the discovery of elliptical orbits. Holt does not emphasize that Kepler's work did not stem from any lack of accuracy of existing models. The Ptolemaic system of epicycles was extremely accurate; it could easily be used today with a few corrections to adjust it to the prevailing state of the solar system. But the Ptolemaic system was cumbersome and complicated-- "chewing-gum and string". It began with a circular orbit, calculated the errors thus produced, corrected the errors by superimposing a second circular orbit centred on the first, and repeated this process indefinitely to obtain more and more accurate empirical predictions at the expense of introducing more and more epicycles (a total somewhere into the fifties).

In passing, it might be remarked that this approach is not unlike the statistical technique of accounting for as much variance as possible, then correcting for "residuals", and continuing this process until a good fit occurs. To what extent are today's statisticians playing the role of Ptolemy, and might there be a Kepler lurking in the wings?

Kepler did not just want a good set of predictions or a good method for computing "the answers;" he wanted a satisfying aesthetic structure that would help him to understand the phenomena.

Another interesting argument in Holt's paper is that the predictions made by a catastrophe-theoretic model are relatively "tight" and readily falsified if wrong. This is in direct conflict with a common criticism of catastrophe theory--because it is topological and hence highly flexible, it can be made to fit almost any set of data and hence (?) is "unscientific." The same argument would appear to apply to (e.g.) Fourier analysis, cubic splines, and the theory of differential equations; hence the query. A possible resolution of the argument might be this: Any specific catastrophe theory model is testable and "scientific," but catastrophe theory itself is not--it is a branch of mathematics. We may ask "does it have applications?" and we may ask of any single application "is this right?" but it is pointless to ask simply "is catastrophe theory right?" just as nobody would think of asking "is calculus right?" (expecting thereby to be able to judge every application of it simultaneously: Either all of them are correct or all of them are false).

THEOREMS ABOUT ADMINISTRATION AND/OR CATASTROPHES

Donald E. Walker

****

President Walker draws upon a lifetime of experience as an administrator to present, in the form of aphorisms, axioms, and case histories some of the regular features which appear to be present in the administration of complex institutions. According to Thom (1977), the stating of analogies or metaphors is the proper aim of catastrophe theory and Walker moves toward that goal in this paper. Walker begins by observing that "getting even" is likely to be an expensive process and often invites rather than avoids catastrophe. He emphasizes the need to maintain homeostatic equilibrium in order to avoid catastrophe even when the solutions seem direct and obvious. Finally, he illustrates the cyclical nature of innovative plans and begins to introduce a model to describe the intuition gained from his experiences.

****

A. "You can never get even with the world. It takes too long and too many lawyers." (Woody Allen)

I suppose the model in the back of my head that makes this an important maxim of administration is the notion that campuses and similar organizations tend to be highly resistant to cusp catastrophes unless interfered with by outside forces or events. I suppose as a corollary to this I believe that, even when such organizations are traumatized by outside forces, they still respond best (and sometimes avoid catastrophes) if they are managed in such a way that everybody is involved and people have the right and responsibility to make some choices, however painful those decisions may be. There are self-correcting and self-regulating tendencies at work on most campuses. Universities are homeostatic, at least to a degree. To throw some additional light on the homeostatic hypothesis, I include a brief essay by Lewis Thomas from his book <u>The Lives of a Cell</u>. How do these pages fit? The chapter fits in kind of a negative way. I have found that cusp catastrophes or even more complicated catastrophes can indeed occur in the face of even mild outside threats, if the institution is badly managed at the moment when the crisis threatens. Also, a crisis can be created by administrators alone in the absence of significant outside threats. One of the "dark seeds" of mismanagement is the attempt to "get even." This impulse stems from the fact that the administrator sees him- or herself as morally superior to the people with whom he works, better informed, or less self-serving. The administrator proposes a plan of action which is resisted or subverted. Then the old ego comes into play and the game plan becomes to "get even" with those who resist the administrator on the ground that this will prevent further mischief. Wrong! If catastrophes are to be avoided and not to be created by administrators, administrators need to respect the organizations they serve enough to change course and to simply absorb the slings and arrows. Resistance isn't personal and it's bad news when administrators decide to churn the unrighteous into rectitude.

B. Axiom: "All of us are smarter than any of us."

I believe that complicated organizations, particularly organizations such as universities with multiple constituencies (including government agencies) emphatically do not operate on a pyramidal model of leadership. The problem is that most campus leaders participate in a prevailing stereotype of the culture about how "normal" campuses should be organized. The model may derive from the examples of the family or the church. The stereotype is that the organization is hierarchical with a strong leader at the top who should make all of the critical decisions. This is not a realistic view. If organizations ever operated like this, or if there is a class of organizations that still does, they are not relevant to the way in which most of society's large public service institutions presently operate. The leader may play a crucial role in the decisions of the organization and may even serve symbolically as the embodiment of the organization's ideals. The process is really much more broadly interpersonal, however, than traditional models allow us to perceive. Example: A university is faced with a significant budget cut. The information comes to the university late in the spring. The rationale for the cut is a shortfall in revenue for the state but also an anticipated decline in enrollment for the institution. What to do? There are a number of decisions which must be made--proposals come forward from campus constituencies, some of them realistic and some of them not. For example:

1. The faculty proposes, initially, in hurried consultations with relevant peer groups, that the budget deficit be met by cutting back on equipment budgets, telephones, supplies, travel and, if personnel dismissals are necessary, that they come from part-time employees and from the administrative cadre. The administration sees this as partly reasonable and partly not. Certainly supplies, equipment, telephone, and travel must be cut. Some part-timers must be let go, but administrative cuts, primarily through attrition, have been taking place for some time and people in administrative offices are already howling about working conditions and the need for additional help. The cut into the administrative ranks will be as explosive in the administration building as cuts in full-time faculty would be in "faculty land."

2. The suggestion comes forward from some of the less initiated in the academic community and from the members of the institution's trustees that the president and the development office should launch a huge campaign for support from private industry, alums, and other groups in order to raise a "war chest" of several million dollars to prevent any cutbacks in the institution. The faculty are very enthusiastic about this idea and release statements to the paper heralding the "leadership" of the board of trustees in a time of crisis. The administration is wary and unconvinced. After all, it is

now late in the spring. Critical decisions must be made about personnel soon. Lawsuits may already be inevitable because of AAUP guidelines for notification of non-reappointment and, in unionized sectors of the campus, union deadlines have already been passed and can be transgressed only after a declaration of financial emergency. To wait until the middle of summer to see if a fund raising campaign will be successful doesn't seem to administrators to be the jazziest idea in town. After all, neighboring institutions have already launched fund raising appeals. The university's alumni association has conducted a fairly successful fund drive earlier in the year. The industries in the area seem to be only moderately supportive of the university and then primarily in technical areas.

3. The university administration feels that because of the size of the deficit the only possibility for realistically meeting the crisis is to send letters of non-reappointment to three-quarters of the part-time faculty. (Some part-timers are absolutely essential to teach laboratories and freshman sections.) They also believe that a freeze must be put on all hiring for the fall. Such action would mean aborting a fairly sizable number of "search and screen" processes already underway. Finally, they feel trustees should declare the institution to be in financial crisis and to issue letters of non-reappointment to a number of nontenured faculty members.

4. There are two departments in the university that, in the view of the president, seem "expendable". They could be closed without damaging the main thrust of the mission of the institution. One is the department of geology and the other is the department of education. The department of geology has been steadily losing enrollment. It has at best only a dozen majors remaining. The faculty has been reduced to a half dozen people. The average age of the department faculty is sixty-two and would be higher had not the department been given a position four years before to hire a petroleum geologist--a man now tenured and in his early forties.

The reasoning of the president is that the geology of the area has never been exciting and that it has already been well mapped and explored. Geology students receive insufficient field experience to fully qualify them as consulting or research professionals in the field. Students have been encouraged to do summer work in Colorado and other "interesting" geological states but the program has met with indifferent success. The human costs of closing the department would be relatively low. Practically all of the professors involved would have livable incomes were they to retire. Majors could be recounseled into other areas. The younger man could be kept on for two or three years to permit students enrolled in the major who chose to stay to at least complete a minor. The presence of one remaining member of the staff might also mitigate the possibility of losing costly lawsuits to students enrolled under catalogues with a geology major listed.

The education department is more of a problem. It is divided. It has a good tournament-level research cadre consisting of five or six people. They might be transferred into other departments or even placed in an office of institutional research which the institution needs. Most of the work of the department is, after all, the training of elementary and secondary school teachers for whom there are few employment opportunities. The problem is further complicated by the fact that, in spite of the slender job market, a great many students still wish to earn a teaching credential. There would be contract and legal obstacles to closing the department. The university would win most cases under the financial exigency clause but some of the teachers could successfully argue that they could and should have been absorbed in other departments. Even assuming that half of these people were to win their cases, the savings to a university would be considerable. There would be fussing from departments that were required to absorb transfers since, regardless of the qualifications of the individual, transfer is never a popular solution to budget problems on campuses. Besides, some of the people who could argue competence in other fields, and therfore be retained, would not be considered the strongest members of the department.

5. I should have mentioned that the first scenario would be that the president is instructed immediately by the trustees and by faculty groups to go to the legislature and the governor and howl bloody murder. I am assuming this strategy has been ineffective and that the legislative leadership has assured the president and the trustees that no changes in budget will be made at this late point in the session unless, indeed, further cuts are required. Now what? The answer, in my view, is that a process of "workup" has to take place. Whatever solution or combination of solutions is finally settled on will be the result of involving the multiple constituencies of the campus in dialogues and even confrontations with each other in order that best-fit solutions may be developed. It's a political process. It's a negotiating process. Forgive the pun, it's a process process. For the president to arrive at a solution before a considerable workup in which the shamans from every tribe put on the buffalo horns, pick up the gourds and dance around the patient will be simply a waste of time and will place the institution in danger of a catastrophe in which the president will be seen as the major problem.

The process will need to be initiated and fostered by the president who must call meetings and offer his proposals for meeting the crisis. Often the only way that people decide in a stress situation what they really want is by a process of elimination by deciding first what they do not want--what is intolerable. This is a legitimate and a proper way to arrive at decisions. If pyramidal models of administration are cherished, however, this appears weak, vacillating, and as an attempt to avoid responsibility by getting others involved in the confusion. Pressures will come from every direction for the president or other administrators to be "strong," to exercise "leadership." What is meant by these demands is that partisan groups seek someone in a leadership position who will pursue their pet objectives ruthlessly. It simply isn't possible without creating a catastrophe. The president and other administrators must be perceptive enough to sense when consent (not approval) exists in sufficient degree to declare a portion of the puzzle to be in place and to announce the decisions that have been reached. Even when the solution is obvious from the start because there are no others that are reasonable, the process must still be transacted. We are accomplices to one another in complex institutions. There are appropriate timings in the metabolism of complex institutions like universities which are highly relevant to decision making and to the induction or prevention of "catastrophes" in stressed times. After a certain amount of maneuvering, there's a point at which permission is given for decisions to take place, even painful decisions. The cues are subtle and when administrators discuss them they use the word "intuition" or "instinct" more often than is usual in either academic or administrative discussions. It's almost as though there were a biological metabolism at work.

The Life Cycle of Innovative Plans

It occurs to me that the natural cycle of innovative plans on a campus may have some relevance to catastrophe theory. I've noticed that innovations are very difficult to arrange in bureaucratic structures when the innovative idea comes from the top. There is a prevalent tendency for those in the affected department to say, "Fine, you have a good idea. Give us the money and we will work it out through our normal departmental procedures." The problem is that somehow things never quite get worked out. The extra money, if it comes, is absorbed to make those already in place and functioning in traditional ways more comfortable. If the suggestion for innovation comes from the department to the dean and president, the answer is, "Thank you for the splendid suggestion. We will consider it in the light of other university priorities," and again nothing happens. Usually, innovation actually occurs in universities because some one person, or occasionally two or three people, catch on fire about an idea and simply will not put it down. They push and test the existing structure to its limits, often arranging end runs that are annoying to those above them in the hierarchy. Basically, they subsidize the new scheme out of their own perspiration and overtime. Having succeeded in attracting acceptance from "the establishment" within the organization, but still sailing under the colors of "innovative verve," they then apply successfully for subsidization. Once they are included in the formal budget structure of the organization, the innovations of the program begin to cost more and more or less and less. Again, the normal tendency in organizations is to spend new monies to make existing programs more comfortable rather than to expand and to continue to innovate.

Let me offer an example: A campus with which I am familiar had as a member of its faculty a very bright artist (painter). This artist was also interested in folk music. The idea occurred, as a result of acquaintances made and experiences he had had on a sabbatical in England, that a gathering of folk musicians might be a successful event at his university in the states. He talked to key people in the university about an allotment of money for such an event but was discouraged. After all, state monies really could not be expended to subsidize a folk music festival. He decided to go ahead on his own. With the help of like-minded students and some friends on the faculty, he advertised a gathering for the following fall. He wrote to folk musicians he had met on his sabbatical year abroad and asked if there was any possibility of their stopping by, free of charge, of course, to perform during what he hoped would be a stellar week where good friends and musicians simply shared fellowship and music with one another and treated each other to the mutual gift of talent. The organizing professor indicated that although there would be no recompense either for travel or performance, he and his friends would be able to offer housing and meals in their own homes for their guests during the period of their visit. A few responded favorably. Attracted by "name" performers from other countries, folk musicians from nearby communities also attended at their own expense. The event provided an emotional "high" for those who were present. It was stipulated by all to have been a highly successful gathering which should be repeated. It had the fresh ego-enhancing smell of a new automobile about it. The festival was repeated for four or five years with the gatherings being somewhat larger each year. In the meantime, the administration had bent a little and had offered first, postage for mailings, next, a little help with typing, in another year, some assistance with phone bills which were becoming larger, and much more active help in arranging space for the performers and in publicizing the event. By this time, the event was drawing favorable attention from all over the nation, at least among folk music illuminati. The university began to take pride in the yearly festival as though it had been officially and centrally sponsored and arranged. The president began to mention it in his community and national talks. The founder of the event meanwhile became increasingly restless. He complained, and legitimately, about the tax the event placed on his time and his energy. He wasn't certain that he could continue unless he received some help. He asked nothing for himself but thought that a paid half-time assistant to handle more of the detail would be very useful. The half-time assistant was arranged. Within a year or so, the half-time assistant had become full time. A part-time secretary was given to the program. The director himself finally declared that although it would be unfair to ask to be recompensed for all the time he spent throughout the year on the program, nevertheless, some additional stipend or reward for the period of most active work would be fair and even essential if the program were to continue. In the meantime, the practice had developed of arranging and paying for commercial housing for some of the visiting folk musicians. There was more talk about the need for travel money and other support for the program. All of these legitimate requests were honored. Finally, the ultimatum arrived that not only was the program in the fall necessary, but several events of a similar character in other fields of music were essential to complete the weave of the colorful tapestry that had been such an ornament to the university. A full-time office must be established to run the festival or else the director had better things to do with his time. In this particular instance the program could have gone either way. Had it been included as a part of the university activities under university sponsorship and full funding, it would have then been in place and the cycle would have been completed. I should add that for the last two or three years of the program neither the number of participants nor the audience had expanded greatly. The program had become increasingly better known and costs had increased but service had declined.

This scenario is, it seems to me, characteristic of the life cycle of such programs. To complete the description of the scenario at the university with which I am familiar, the demand for expansion and full subsidization occurred in a bad budget year. The director was told that regretfully the university could not expand its commitment and, indeed, might have to cut back some. The director with equal regret indicated that the event should be dropped. Nothing happened for two or three months. Then another individual came to the administration, indicated that she thought the program should and could be rescued by contracting commitments and events, by a return to the principle of calling on volunteers, by limiting the celebration to a weekend in the fall as in the beginning, and by relying on faculty families to supply housing to visiting musical groups. Modest postage and telephone subsidization and a small financial stipend for the new organizer were all that was required. The cycle began again. The next year the new organizer felt that more time was required and a partial salary was agreed to. The last I heard, the new organizer had applied for a part-time assistant.

The relevance to the catastrophe theory may be that in innovative enterprises, on a campus at least, the enterprise is healthiest and most vigorous and most useful when it is subsidized out of overtime and enthusiasm of faculty and staff. As more and more of the operation is subsidized, services are reduced and costs go up. It may be that catastrophes occur in organizations at stages roughly comparable to the life cycle of biological "organisms." The biological metaphor has been used and overused before but it does seem suggestive.

COMMENTARY

Walker's remarks bring out some points from the practical side of administration. First, he shows that the actual administrative process is much more complicated, in fine detail, than any model could be. (In fact, if such a model could be constructed, it would be far too complicated to be useful.) So the aim of a model should be to provide insight into the coarse structure and to set up some signposts and useful metaphors that can be borne in mind when making decisions. (The <u>metaphor</u> is one of our most useful and traditional ways of transferring knowledge gained through experience; the metaphor provides a mental setting in which a phenomenon can be "experienced" cerebrally rather than physically.) Thus a catastrophe (or other) model would give a broad framework. After setting that up, a finer structure may be introduced.

Another point brought out by Walker is that catastrophe really does mean <u>catastrophe</u> after all! He feels that abrupt changes are an undesirable property of leadership, that gradual change through maintenance of homeostasis is a desirable property. Even though the outcomes might be the same, it is preferable to achieve the outcome through continuous change rather than abrupt change ("catastrophic jump"). Given the "maintenance of homeostatis" as a possible minimization principle and the potential for abrupt change, a catastrophe model would seem a natural beginning point. Understanding such a model might help to generalize some of these properties of "good" and "bad" leadership.

Walker's last section describes a cyclical dynamic process which might be used as a starting point for a genuinely dynamic (non-equilibrium) model.

CATASTROPHE THEORY AND STATISTICS IN THE SOCIAL SCIENCES

I. R. Savage

****

This article concerns the application of current statistical methods to catastrophe models and the new methods that might be required--in particular, methods for estimating the structure of a dynamical system. The question of equilibrium/disequilibrium is raised.

At present, it is not easy to detect multivaluedness in a set of data and new techniques might be required here. Some specific suggestions are made.

Finally there is a brief survey and commentary on existing work relating statistics to catastrophe theory.

****

The usual way to develop new methods and theory in statistics is to fill a need from empirical research. Thus one presumes that as empirical applications of catastrophe theory appear when existing statistical knowledge is not adequately supportive of that research then new statistical ideas will be developed. It is a bit speculative to guess at what will be required before much work has appeared. With these thoughts in mind I will make brief comments on where statistics is, some directions of development which might be required to implement catastrophe theory, and the impact catastrophe theory has had on statistics.

Where Statistics Is

Statistical methods are highly developed for discrete and continuous, and for univariate and multivariate models when the errors are independent or have a simple pattern of dependence (as in survey sampling). New techniques and theory continue to be developed--ridge regression, optimum designs, Bayes procedures, etc. There is a gigantic body of material in this area developed in the last 100 years.

There is a growing and pressing need for statistical procedures to be applied when the observations have relatively complicated patterns of dependencies. Discrete time, continuous, univariate, and multivariate time-series theory is highly developed and is an active area of research. This is a central part of engineering and economics. Other areas, such as continuous time, discrete Markov processes are being actively developed as a response to needs of sociology. A great variety of techniques have been developed to explore response surfaces, such as optimal design or bio-assay.

New Statistics for Catastrophe Theory

Thinking of catastrophe theory as the analysis of the steady states of a dynamic system, it is necessary to be able to give a statistical analysis of the entire system. Although the steady states are the focus of catastrophe theory it should be realized that in the social sciences attention will linger on the dynamics since steady states are seldom approached when the obvious examples of social institutions (such as business organizations, professions, families, etc.) are considered.

If the dynamic system is deterministic, with superimposed measurement errors, then the resulting statistical analysis would stand a good chance of being routine or easily obtained from currently available techniques. If the dynamic system is stochastic, possibly with measurement errors, then there is less chance that existing techniques will be helpful with the statistical analysis. It is not possible to guess which stochastic models will play an important role in the emerging need to use stochastic processes in social science research. At the moment the research effort is to develop the statistical properties of relatively simple and often used processes, such as time series and special classes of Markov processes. This effort is not likely to change course without compelling reasons.

For many dynamic systems the steady state is quickly approximated in terms of the times required to make observations. For social systems the time to reach stability might be considerable although good approximations might occur quickly, as in the theory of stable populations. So if we wish to study social systems with either a stochastic drive or measurement error then we need analytic and statistical techniques to see if the system is approximately in a steady state or to measure the response time for the system to reach a steady state after it has been shocked out of a steady state. These topics need to be covered in applications but it is doubtful that basic new statistic topics would arise. The possible things that might occur are so numerous as to make it impossible to do the statistical research before people have thought of the kinds of empirical problems they wish to explore.

With systems having slow response times, the most direct approach to catastrophe theory is to observe and model the dynamic system. When the structure of the system has been well estimated we can also estimate the steady states and catastrophes. One can then observe steady state behavior, consequences of shocks, moves along paths in the control states, etc. That is, we can check the predictions of the model. In rapidly moving systems one can not estimate the model by directly observing the dynamic system. Then one might be forced to focus on the steady states.

So now let us consider what, if anything, can be learned from the steady states. In the social sciences this is likely to be a strategy of second choice since the dynamic process can often be directly observed and steady may not be obtained.

The analysis of the steady state surface without reference to the dynamic system imposes several requirements. First of all one must be sure they are on the steady state surface. One must develop observational plans that allow one to check that the control variables have been fixed long enough to make sure the system has come to rest. Different techniques would be required depending on which variables, control and response, can be observed repeatedly. (There may be situations where it is not possible to fix control variables long enough to have the system settle down.) At this time it is doubtful that any new statistical techniques will be required.

Just thinking of cusp catastrophes, very little can be learned from single observations of each realization of the system. If the scatter plot of the response surface appears in part to be multivalued, that is slightly suggestive of a cusp catastrophe. The "slight" stems from several considerations.

Although current graphic techniques are very good in examining data in three-space, it is not clear that visual clues would be adequate in most situations with the social data to make a strong argument for or against a multivalued response surface. This class of problems might require new statistical techiques. Also, if multivalued response surfaces were suspected there are alternatives to catastrophes. In particular, many scientists and statisticians might search for additional control variables in order to obtain a nicer surface. This traditional approach would be appealing in those situations where the dynamics have not been explored and a detailed modeling effort has not been made.

A more fruitful situation occurs when several observations can be made on the same system in steady states under different values of the control variables. (We have in mind here the possibility of moving an individual through a sequence of steady states.) These multiple observations begin to allow us to search out the existence of cusp catastrophes. It must be remembered at all times that we need assurance of making the observations in steady states. Further, there is likely to be substantial noise in the data so catastrophe surfaces are not going to reveal themselves in a crisp manner. If data can be obtained in this form there are a host of statistical techniques that need development.

(A) How to design sequential and nonsequential experiments to locate and measure a catastrophe.
(B) Repeat (A) with observational data.
(C) Define operationally what is a discontinuity. Then give statistical techniques to locate and measure the sizes of discontinuities.
(D) Repeat (C) for the other "qualitative" features introduced by catastrophe theory.

Treatment of these topics is not out of range of statistical theory. The lack of technique appears more as a lack of demand than lack of ability. The bottleneck in applying catastrophe theory to social science material is to find appropriate situations where adequate data could be obtained.

Impact of Catastrophe Theory on Statistics

The work of J. Q. Smith (1979) employs catastrophe theory to solve problems in statistical theory and the work of Cobb (1978, 1980b, 1980c) creates new topics in statistical theory. These are the major if not the sole authors. I will cite a few examples.

Smith considers the problem of finding the Bayes strategy when the posterior risk has several minima. He motivates the problem and further motivation is given by B. Spencer. Catastrophe theory as used here helps to organize a moderately messy problem in differential calculus. In the same spirit Smith has studied qualitative properties of families of density functions when the underlying variable is subject to a monotone transformation.

Cobb has opened a new area in catastrophe theory by replacing the deterministic differential equations of the dynamic system by stochastic differential equations; he perhaps has only scratched the surface. Cobb's theory replaces the steady states by the density function of the steady states. The study of the qualitative properties--multimodality as a function of the control parameters--becomes the catastrophe theory.

Smith's work is generated by existing problems so that it might find immediate use. On the other hand, Cobb is more speculative and more time will be required to assess its usefulness.

COMMENTARY

At the present time, a gulf exists between current statistics and applied catastrophe theory. No doubt this gulf will be filled in time after enough data sets are collected which exhibit catastrophe-like properties when viewed naively and after enough statisticians are attracted to help understand such data sets. There will be little interest among statisticians to study hypothetical situations and there may be little interest among mathematicians to collect large data sets without prior guidance from statistics, but there is no reason to think social scientists won't collect data without predetermined analytical guidelines--that's never held them back before. It will take time for these natural forces of curiosity to work, but the gulf will surely be filled, or judged irrelevant, in time.

It would certainly be very nice to be able to estimate an entire dynamical system. But, as Stewart remarks, this is probably too much to ask at this stage. It is reasonable to concentrate on the first few steps.

Equilibrium is a sufficient condition for a catastrophe model, but not a necessary one. In particular there is no requirement that the control variables be in equilibrium; only that the state is responding fast enough for a quasi-static approximation to be valid. One attractive area for catastrophe-theoretic attack is the possibility of simplifying certain types of dynamical system into a fast flow towards quasi-equilibrium and a slow dynamic on control space. There is a paper by Vasilis Angelis (1980) dealing with population changes in towns which uses this modeling technique to simplify the problem. The analysis of the observed data is made much simpler after the decomposition into fast and slow variables has been made.

Smith's work has direct relevance to decision-making. Cobb has found the beginnings of a very nice mathematical generalization of important results in stochastic differential equations. While Cobb's own results are aimed at fairly practical aspects of the problem (least-squares estimation, etc.) and thus require practical testing before their value can be judged, the study of singularities, bifurcations, and catastrophes in parametrized stochastic differential equations should be an interesting and rich area of mathematics. The relations between singularity/catastrophe theory and stochastic differential equations appear to be a fruitful source of research.

It has often been asked, and it is certainly a natural question, what is the distinction between, or advantage one over the other of, a "cusp" model and a "merging-normals" model in dealing with situations where unimodal distributions may become bimodal? We remark on this question, based on work of Ian Stewart (1983), at this time.

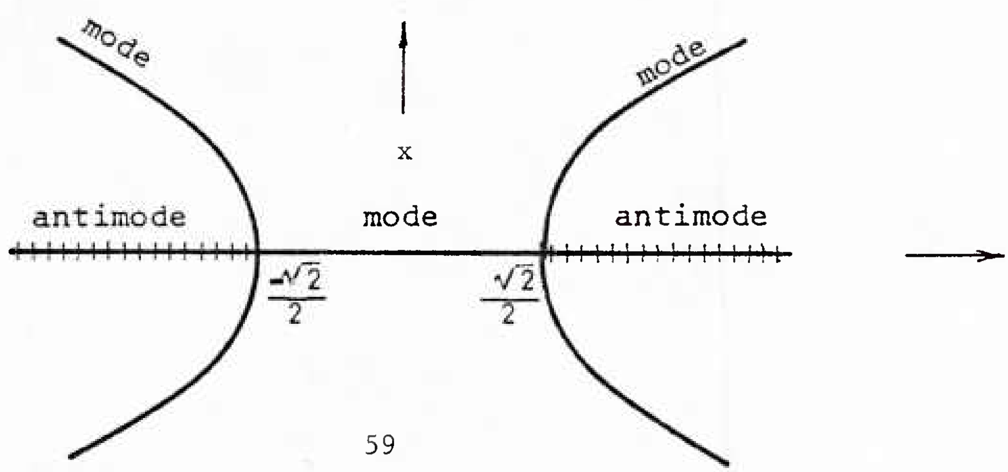
Consider the two parametrized families of distributions

$$F_{a,b}(x) = F_0 \exp(-x^4/4 + a(x^2/2) + bx) \text{ and}$$

$$G_\lambda(x) = G_0(\exp[-(x-\lambda)^2] + \exp[-(x+\lambda)^2]),$$

where $F_0 G_0$ are normalizing constants (parameter dependent). We refer to F as the "cusp" model and G as the "merging-normals" model for the rest of this discussion. Note that F is Cobb's canonical cusp distribution; G is certainly close to what people have in mind when they speak of "merging normal distributions". The parameters a,b are the "bifurcation (splitting) factor" and the "weight (normal) factor," respectively. The parameter λ is a "merge" factor.

The model G represents two superimposed normal distributions whose modes merge and separate along the lines $x = \pm\lambda$ (as λ passes through $\lambda = 0$). One might guess that G is thus bimodal for $\lambda \neq 0$, with modes merging and separating at $\lambda = 0$. Something more interesting happens. Stewart shows that the modes and antimodes of G appear as

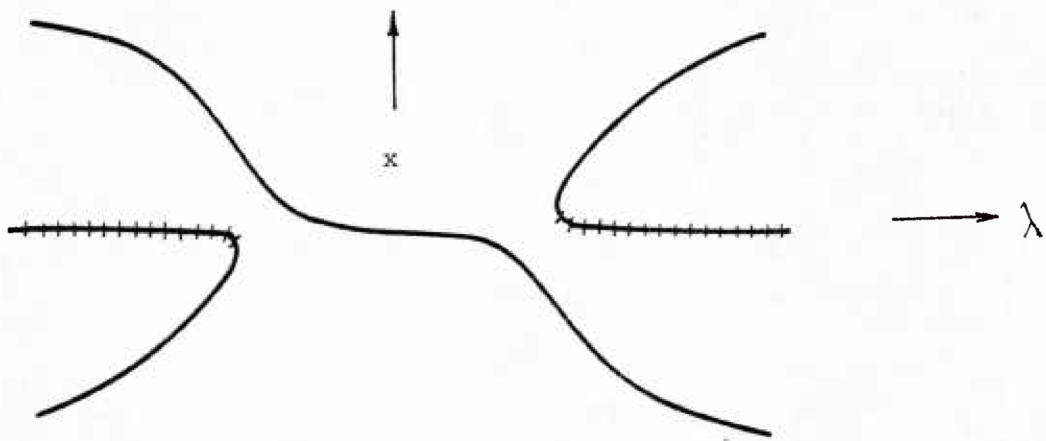


That is, G_λ is bimodal for $\lambda < -\sqrt{2}/2$, becomes unimodal for $-\sqrt{2}/2 \leq \lambda \leq \sqrt{2}/2$, and goes bimodal again for $\sqrt{2}/2 < \lambda$. Note the similarity between this picture and a "two-ended pitchfork!"

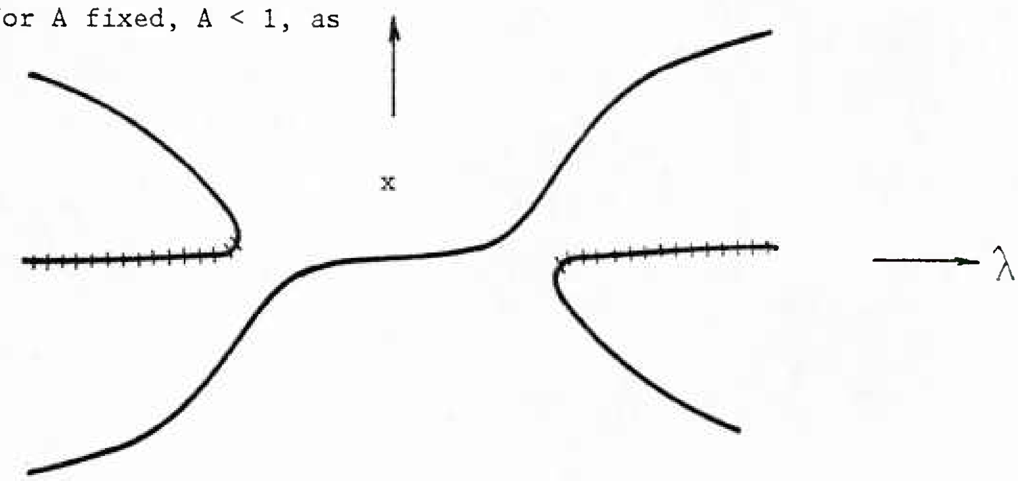
In the absence of outside symmetry constraints, a bimodal distribution cannot maintain symmetry between its two modes. Therefore, G is a structurally unstable model. Consider the introduction of another parameter A, "weighting," that allows the two modes to differ in size:

$$G_{\lambda,A}(x) = G_0(\exp[-(x-\lambda)^2] + A \exp[-(x+\lambda)^2])$$

Stewart shows, for fixed $A \neq 1$, the larger mode assimilates the smaller one at some value $\lambda = -\lambda_A$, G remains unimodal for $-\lambda_A \leq \lambda \leq \lambda_A$, and then reforms a smaller mode on the other side for $\lambda_A < \lambda$. The modes of $G_{\lambda,A}$ appear for A fixed, A > 1, as



and for A fixed, A < 1, as



All other perturbations of the double pitchfork are realized when λ, A are moved through $\lambda = 0$, $A = 1$ along various paths. Graphing all modes and antimodes of $G_{\lambda,A}$ in λ,A,x-space gives the picture shown in Figure 1. Thus $G_{\lambda,A}(x)$ is a "double cusp."

Considering the case $\lambda > 0$ only, the two models $F_{a,b}(x)$ and $G_{\lambda,A}(x)$ are equivalent. It seems that the major distinction might be in point of view: G comes encumbered with the idea that there are "two superimposed populations," while F brings with it the idea of "one population with two behavior patterns." Cobb's cusp seems to have the following advantages:

(a) it is more easily derived from a stochastic differential equation

(b) it is simpler (one singularity instead of two, if one does not restrict $\lambda > 0$)

(c) it is more easily comparable with the cusp (one example of structurally stable singularities).

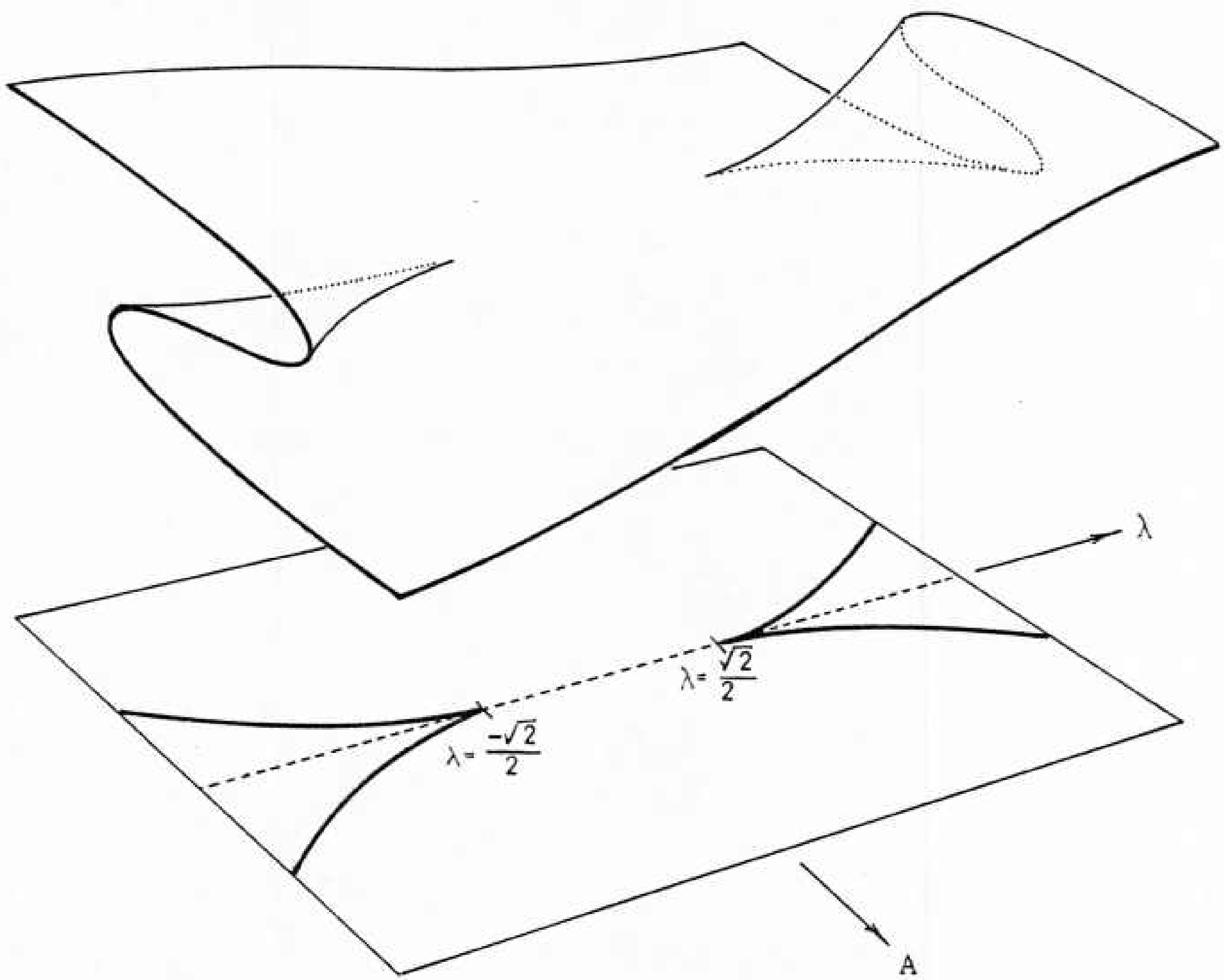


Figure 1.

TWO THOUSAND YEARS OF CATASTROPHES: PURE AND APPLIED MATHEMATICS

L. Markus

****

This paper demonstrates that catastrophe theory is part of a long and successful mathematical tradition of geometric methods in science by providing an often amusing description of the historical development.

Emphasis is placed on the search for natural qualitiative forms as the basis of models, and the history is used to support this line of attack.

Advantages of catastrophe theory include structural stability and genericity and it appears a good place to start in seeking to advance beyond the current linear or pseudolinear modeling techniques.

Some mathematical and applications-oriented problems are described.

****

The Tradition.

First of all I wish to pay tribute to the Great Geometer who launched the mathematical program that forms the basis of modern catastrophe theory. I refer, of course, to Apollonius of Perga (262-200 B.C.), who wrote his famous treatise On Conics at the Museum at Alexandria (Apollonius, 1952).

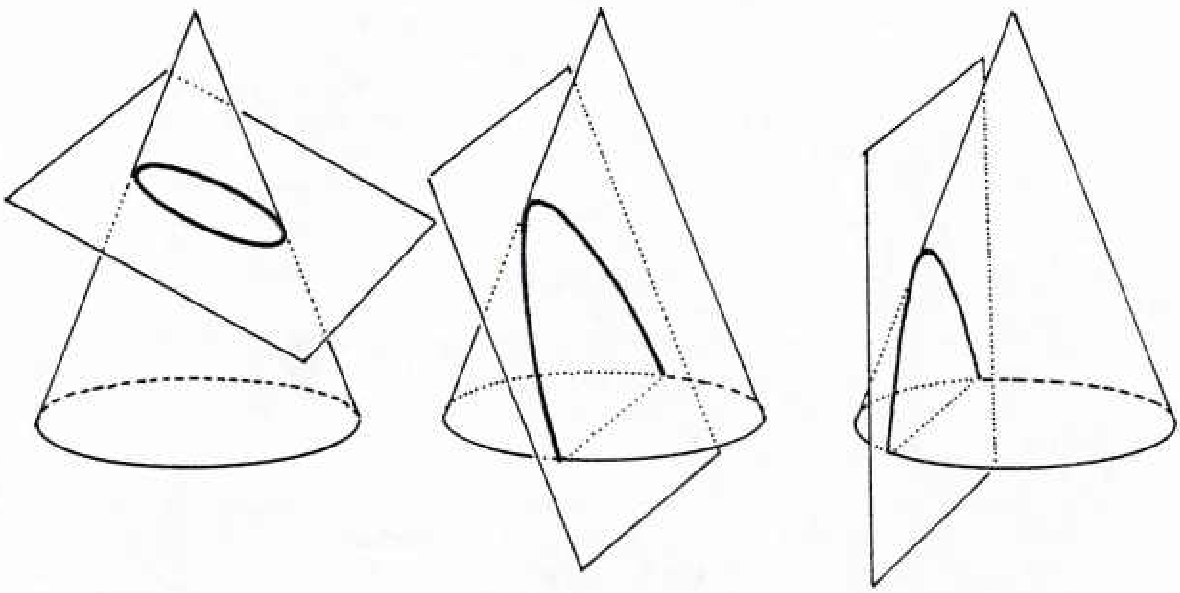
In the current generation there has appeared another great geometer, Rene Thom (1972), who has reformulated and continued this mathematical tradition in his important research volume Stabilite structurelle et morphogenese, written at the Institut des Hautes Etudes Scientifiques at Paris.

In this paper I intend to discuss the geometric researches of Apollonius, of Thom, and of some of the great geometers from the intervening two thousand years who developed the mathematical area that we now refer to as "catastrophe theory." In particular, I plan to indicate the development of the geometrical and scientific concepts with the accompanying theories of pure and applied mathematics and to comment on how these ideas have interacted with the methodology and philosophy of science over the past two millenia.

In order to set the record straight, the central mathematical theory traced in this lecture is that of conic curves and their generalizations to quadric surfaces, and our theme is that catastrophe theory is the latest fillip (but a worthy fillip) in this millenia old tradition.

Helenistic Catastrophe Theory

Euclid left Athens to found the School of Mathematics in the Museum of Alexandria and, a generation or so later, the most distinguished scientist on the staff was Apollonius of Perga. Apollonius (1952) established his mathematical fame (during the Middle Ages he was referred to by the title "The Great Geometer") by a research treatise On Conics in which he defined the ellipse, parabola, and hyperbola curves as conic sections, and demonstrated their properties that now are familiar topics in a standard college course.



Ellipse Parabola Hyperbola

The method of Apollonius was to cut or section a right circular cone by a plane and to obtain the ellipse, parabola, or hyperbola as the intersection of the plane with the conic surface. When the plane was horizontal--that is perpendicular to the axis of the cone--the section was a circle. When the plane was inclined, the section was an ellipse whose eccentricity was determined by the angle of inclination. As the inclination was increased, the family of ellipses could be parametrized by the corresponding eccentricities. For the very special inclination that brought the plane into parallel with a generator line of the cone, the family of ellipses terminated abruptly (a catastrophe) and the conic section was a parabola. For still steeper angles of inclination the resulting curves formed a family of hyperbolas.

We are interested here in the tacit assumptions of philosophy and psychology in the approach of Apollonius rather than his technical mathematical results about conic sections. In particular, we shall comment on how these assumptions and the resulting mathematical discoveries fit into our modern concepts of catastrophes.

First recall that the Greek philosophy of mathematics, pure and applied, held that straight lines and circles were the only natural curves that had a legitimate role in science. This view was emphasized in Euclid's Elements (1952) where triangles, squares, polygons (formed from line segments), and circles are the only plane figures of significance. [Note that one of the famous Delian problems was to "square the circle."]

Aristotlian mechanics (Aristotle, 1952) asserted that terrestrial bodies fell in straight lines towards the earth and celestial bodies moved in circular paths in the heavens. Later astronomical theories of Ptolemy (1952) required some modifications of this view to allow compound circular motions (epicycles) to explain the motions of the planets, but the epicycle configuration and the pre-occupation with the circular form was assumed inviolate through the work of Copernicus (1952) over a millenium later. [It is of interest to note that this pre-occupation with lines and circles still exists in Newton's first law or Einstein's law on geodesics and for the spherical expanding universe. See Einstein, 1923.]

From this viewpoint the mathematical significance of the research of Apollonius was that he brought the new curves, ellipse, parabola, and hyperbola into legitimacy by relating them directly to constructions involving lines and circles. [The inclined plane is generated by a moving line and the right circular cone is generated by a line joining the vertex to the circular section.]

Moreover, besides becoming legitimate in a certain psychological sense, these conic curves become practical, workable objects that could be handled successfully by skillful mathematicians. In other words, the conic sections entered the mathematical repertory of natural and useful constructs that should always be kept in mind by future scientists.

A similar claim could be made for Archimedes (1952), the contemporary of Apollonius, with regard to his treatise On Spirals. Thus when Kepler reached (mentally) for the ellipse to describe the orbit of Mars or when Cornu reached for the spiral to describe the diffraction of light, these constructs presented themselves as natural, practical possibilities. In fact, there is some evidence that spirals and their generalizations tend to become more important than conics whenever terrestrial dissipative dynamics displaces celestial conservative dynamics in mathematical research fashions. However, a catastrophe theory for families of spirals (that is, bifurcation theory for dissipative dynamical systems) is still in its infancy.

Another feature of Apollonius' approach to conic sections is more technically related to the concepts of catastrophe theory. Namely, a certain qualitative form (the ellipse) is maintained through a parametrized family of similar geometric objects, but at a special "catastrophe point" (the inclined plane parallel to the generator of the cone) the family of ellipses changes abruptly into a family of hyperbolas. The special "catastrophe section" defines a degenerate conic, a parabola in this example. Of course, the concept of "qualitative form" here refers to the "global figure" of the ellipse (or hyperbola), that is, the shape of the full ellipse, as a closed curve with certain special symmetries. This is not the same as the "local figure" that enters into the modern catastrophe theory of Thom.

But the central idea in both geometries concerns abrupt radical transformations of form, with emphasis on the structure of the transformation itself and, even more, abrupt or catastrophic transformations of conics.

It is my view, expounded later in more detail, that the main consequence of Thom's catastrophe theory is to legitimize certain geometric forms, and to bring these new and interesting geometric configurations into the mathematical repertory of natural and useful constructs, and that also should be kept in mind by future scientists. What are Thom's elementary catastrophes? In what sense are they philosophically and psychologically natural? And how can they be interpreted as straight-forward generalizations of Apollonius' conic sections, in theory as well as in practice? These questions are the subject of this paper.

Renaissance Catastrophe Theory

During the 16th and early 17th century there were two brilliant applications of the methods of conic sections to the newly arising sciences of kinematics (by Kepler, 1952) and of dynamics (by Galileo, 1952). The introduction of the ellipse into astronomy and the parabola into terrestrial ballistics were scientific discoveries of the highest order of creativity and profundity. Yet neither innovation would have been plausible, despite the detailed observational evidence, without the philosophical and psychological framework created by Apollonius and other classical geometers in their studies of conic sections.

Kepler made the gigantic leap of the imagination to replace the circular complex of epicycles of Copernicus by elliptical orbits for the planets around the Sun. From the mathematical viewpoint, Kepler was much more of a revolutionary than was Copernicus. But Kepler still could not break totally from the circular tradition, and was able to accept the ellipse as a physical reality only because of its global geometry, its symmetry, and its close relation to the classical circular form (and, of course, because the choice of an ellipse for the orbit of Mars gave a very good fit to the observational data). As an instance of the strength of the Greek tradition, Kepler regularly computed the elliptical orbit using the inscribed and circumscribed circles as guides.

Galileo's discovery of the parabola as the trajectory of a cannon ball was revolutionary in a different and more modern sense. Galileo was led to the parabola, not because of the global geometry, but because of its local or even infinitesimal geometry that describes the instantaneous acceleration of the cannon ball at each moment. In modern terminology, the parabolic trajectory is the solution of the Newtonian differential equations of motion, where the only significant force is the constant gravitation of the Earth. Thus Galileo led the philosophical revolution away from the worship of global symmetry to the slavery to infinitesimal analysis--the revolution from Aristotelian perfection to Newtonian determinism (although Aristotelian teleology seems to be making a comeback in the group theoretic methods of atomic physics).

Enlightenment Catastrophe Theory.

Newton, through his discoveries of differential calculus and his techniques of power series, cast the laws of nature into the language of infinitesimal analysis. An arbitrary variable quantity as portrayed geometrically by a curve should be considered to be (approximately) a straight line or linear function in the infinitesimal neighborhood of each point. Further, relative to higher order infinitesimals, a curve should be regarded (locally) as a conic and a surface should approximate a quadric. In fact, Newton (1952) in his System of the World devised a theory of nature that emphasized second-order approximations, for instance, concepts such as acceleration and curvature. [Note: curvature is still central in Einstein's general relativity and the recent Yang-Mills field theories.]

The new philosophy of the mathematization of natural law, through infinitesimal calculus, gave a new psychological acceptance to the importance of the role of conic curves and quadric surfaces in physical reality. According, an important task for mathematicians, both pure and applied, was to clarify and classify the geometrical properties of quadrics in two or more coordinates. This problem was resolved by L. Euler (1797) when he classified the real quadric surfaces into three types: ellipsoid, hyperboloid of 1-sheet, and hyperbolic of 2-sheets, with the transition or "catastrophe" quadrics designated as elliptic paraboloid and hyperbolic paraboloid (saddle surface). The method of Euler also gave immediately the geometric generalizations for higher dimensional quadrics, a topic now covered in undergraduate mathematics classes.

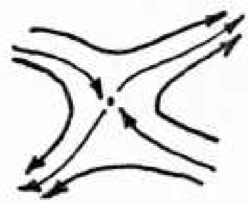
But it was for J. Lagrange (1762) to demonstrate the power of this geometry of quadrics in applications by his method of "small vibrations," probably the most useful mathematical tool in the history of dynamics, and still the most widely-used method in current engineering design practice. Lagrange studied dynamical systems at and near equilibrium--that is, with stable or unstable small vibrations about the equilibrium state. By means of the mathematical approximations of linearization ("dropping higher order terms"), Lagrange found that a good model for all such small deviations from equilibrium could be postulated to be the movements of a particle under a gravitational force but constrained to slide on a quadric surface (or higher dimension quadric for dynamical systems with many degrees of freedom). For example, the small vibrations near a stable equilibrium are modeled by a particle sliding near the bottom of an ellipsoid, completely unstable motions correspond to deviations away from the top of the ellipsoid, and partially unstable motions to a saddle on a hyperbolic paraboloid. In technical terminology, the nature of the stability of the small vibration is determined by the eigenvalues of the linear approximation to the dynamical forces (or of the quadratic approximation for the potential function).

Modern-Era Catastrophe Theory

The dynamics of Newton, Euler, and Lagrange, through the methods of infinitesimal analysis, brought the geometry of quadrics to the fore because quadrics naturally arise through the second-order approximations of general functions. In this sense the "naturalness" of quadrics was based on the choice of seemingly non-natural mathematical calculations of power-series alegebra. While this basis was certainly practical and workable, it lacked any strong philosophical and aesthetic appeal. During the century 1850-1950 A.D., a new philosophical and psychological foundation for this usage of quadrics was provided via the concepts of "qualitative form," "structural stability," and "generic behavior."

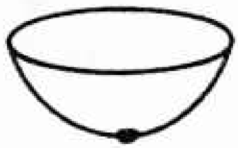
It was in his doctoral thesis that H. Poincare (1879) made precise the method of small vibrations that had been developed by Lagrange as an approximation method for dynamical systems near a stable equilibrium. Namely, Poincare showed that if the dynamical system is viewed or measured in suitably flexible coordinates then its behavior is exactly that of the linear approximation of Lagrange. Thus, up to "qualitative form" the nonlinear dynamical system is precisely linearizable near its stable equilibrium state. In more recent times the qualitative linearization theory of dynamical systems has been extended to include unstable systems with quite general behavior but always locally near an equilibrium state.

Linear

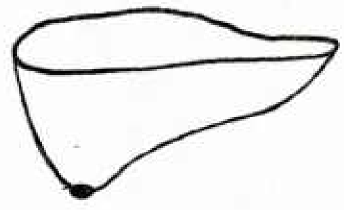


Nonlinear

Poincare showed that, up to qualitative form, nonlinear vibrations are precisely the same as linear vibrations which Lagrange had treated as particles sliding on quadric surfaces. The achievement of Poincare in perfecting the small vibration dynamics of Lagrange was duplicated later by M. Morse (1934) in his theory of critical points of functions which he related to the geometry of conics and quadrics of Apollonius and Euler. That is, Morse considered general surfaces and graphs of differentiable functions in higher dimensions, and showed that near each critical point (where the function has zero gradient) the qualitative form of the function is exactly that of a quadric. While Newton and Euler had recognized that an arbitrary function near a critical point could be closely approximated by a quadric, Morse showed that the graph of the function was exactly a quadric, up to considerations of qualitative form.



(Nondegenerate) Quadric



Nondegenerate critical point

[Technical Remark: By power series expansion methods near the origin

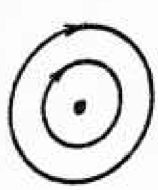
$$x_i = 0,\ f(x_1,x_2,\ldots,x_n) = (\delta f/\delta x_i)_0 x_i + (\delta^2 f/\delta x_i \delta x_j)_0 x_i x_j + \ldots,$$

where we assume $f(0,0,\ldots,0) = 0$. In case the origin is a critical point, $(\delta f/\delta x_i)_0 = 0$ and the quadratic approximation to $f(x)$ is given by $(\delta^2 f/\delta x_i \delta x_j)_0\, x_i x_j$; and this is non-degenerate in case $\det \left|\delta f/\delta x_i \delta x_j\right|_o \neq 0$.]

It is important to note how this modern concept of qualitative form differs from the corresponding concept as it appears implicitly in the geometry of conics propounded by Apollonius. For Poincare and Morse the geometry is local, that is, significant only in the neighborhood of some point, whereas for Apollonius, the global geometry of the conic-as-a-whole is the important feature. In other words, the local analysis is concerned only that a small part of a curve (a surface) resembles a small part of a conic (a quadric). Another distinction is that Poincare and Morse both allow general (diffeomorphic) nonlinear transformation or distortions in recognizing the qualitative form, whereas Apollonius allowed only linear transformation or scale changes in recognizing the common qualitative form for the family of ellipses. Thus the modern concept of qualitative form is broader, yet it is deficient in that it fails to capture even the most obvious or gross quantitative features.

Even well into the twentieth century the motivation for this classification and local qualitative analysis still rested on the tradition of power series and a more general philosophical approach seemed desirable. This goal was achieved when L. Pontryagin (1937) provided a definition and criterion of structural stability.

We indicate first the significance of the fundamental concept of "structural stability" for dynamical systems. Begin by classifying dynamical systems up to qualitative form (allowing nonlinear coordinate transformations) and think of formalizing the condition that such a qualitative form be maintained as a permanent feature under arbitrary exogeneous perturbations. That is, we seek to study qualitative forms that are stable in structure. Hence, a dynamical system is structurally stable in case it maintains its given qualitative form under all suitably small perturbations. For example, the linear oscillation $\ddot{x} + x = 0$ (all solutions periodic, with the constant period 2π) is not structurally stable since an arbitrarily small frictional force perturbs this dynamical system to a damped oscillator $\ddot{x} + \varepsilon\dot{x} + x = 0$ (small ε) having no periodic orbits. On the other, hand the damped linear oscillator $\ddot{x} + \dot{x} + x = 0$ can be shown to be structurally stable and hence represents a more practical model of engineering reality.



While the concept of structural stability was first proposed for dynamical systems, the analogous property is easily understood for real functions, whose graphs may specify curves, surfaces, or higher dimensional manifolds in many variables. The two uses of the concept are closely related (although technically not identical) if we take the dynamical system along the gradient of the real function--that is, we consider the potential function for the dynamics.

A real function $f(x_1,x_2,...,x_n)$ is called structurally stable in case all perturbations, to functions $f(x_1,x_2,...,x_n) + \varepsilon(x_1,...,x_n)$ (for all suitably small $\varepsilon(x_1,...x_n)$), still maintain the same qualitative form as $f(x_1,...,x_n)$. Sometimes it is desirable to localize this concept and to consider (in an obvious way) a function to be structurally stable in the locality of some point on its graph. It was immediately recognized that a differentiable function $f(x_1,...,x_n)$ was automatically structurally stable near each non-critical point and also near each critical point that was non-degenerate (that is, with the qualitative form of a non-degenerate quadric). In fact, Morse's theory showed that a critical point of $f(x_1,...,x_n)$ is structurally stable if and only if it is a non-degenerate critical point [technically: $\delta f/\delta x^i = 0$ and $\det\left|\delta^2 f/\delta x_i \delta x_j\right| \neq 0$].

The concept of structural stability thus provides a new approach to the classification of critical points for a real function. Moreover, the non-degenerate critical points--and hence, non-degenerate quadrics--now have a sound philosophical legitimization because they are precisely the geometric forms that maintain their qualitative form under all perturbations or exogeneous variations.

Contemporary-Era Catastrophe Theory.

Beyond the concepts of "qualitative form" and "structural stability," a further important idea "genericity" was introduced into geometry in a far-reaching and practical manner by Rene Thom (1956), in particular, Thom's transversality lemma. These intertwining concepts also formed the core of the contemporary theory of dynamical systems as advanced by S. Smale (1967).

A class of mathematical objects, for instance dynamical systems, real functions, or geometric surfaces, is called generic in case almost all of the possible objects of this type belong to the specified generic class. In other words, a generic class contains all the mathematical objects under consideration, excepting relatively few, very special pathological or artificial examples. So generic behavior means typical behavior, with the unnatural and atypical cases excluded. Of course, the precise description of the allowable class of mathematical objects and the definition of the atypical excluded objects is necessary for a rigorous mathematical theory, and this has been satisfactorily accomplished in most problems of interest, in particular for functions, or even function-families that depend on parameters.

For instance, among all real functions of one variable (with graphs that are smooth curves in the (x,y)-plane), those having only non-degenerate critical points (locally of the qualitative type of non-degenerate quadrics) form a generic class and the same conclusion holds for real functions of many variables. That is, functions with degenerate critical points are hereby declared to be atypical and thus excluded from the generic class. It is remarkable that this generic class of functions can also be characterized as those functions with the property of structural stability.

The spectacular innovative idea of R. Thom was that any generic transition between distinct quadrics (or between functions whose critical points are qualitative, like distinct quadrics) must necessarily involve higher order geometric forms that are not themselves quadrics. That is, if we seek an interpolation between two generic functions and if we demand that the interpolation family itself be generic, as a parametrized family of functions, then this family must involve some individual functions that are nongeneric (that is, functions that possess certain degenerate critical points).

In Apollonius' theory of the catastrophes of conic sections, a family of ellipses changes abruptly to a family of hyperbolas via a catastrophe at the parabolic section. But the catastrophe theory of Thom shows that Apollonius' transition between these distinct conics is not generic and that any generic transition must involve certain higher curves that are not themselves conics, and, in particular, not parabolas. In this way, Thom brought certain new types of curves and surfaces into the realm of "natural" geometry and physical reality. These new geometrical forms are not quadrics but they arise as the natural interpolations between quadrics. Thom named these new geometrical forms the elementary catastrophe models. Thus, from the modern viewpoint of "qualitative form, structural stability, and generic behavior," Thom showed that Apollonius had found the wrong interpolation between conics (and quadrics) and that the correct interpolation must pass outside the class of quadrics and must involve the new geometric forms classified and described as the elementary catastrophes.

In accord with the modern viewpoint, Thom's theory of catastrophes was local in nature, that is, it treated a family of functions in the local neighborhood of the catastrophe with the region of space restricted near the changing critical point and the parameter for the functions of the family near the catastrophe value. But Thom's theory was recast in global form by E. C. Zeeman (Trotman & Zeeman, 1974) to deal with real functions in the whole n-dimensional coordinate space and generic families of such global functions. Later L. Markus (1979) rephrased the work of Zeeman to give a global version of Thom's theory of catastrophes that was hinged on the concept of "structurally stable function families," rather than on generic function families. Markus verified that both approaches are logically equivalent and produce the same elementary catastrophe models in the transitions between the non-degenerate critical points of the functions of the family.

In summary, the Thom-Zeeman theory deals with real-valued functions $f(x_1,x_2,\ldots,x_n,\alpha_1,\alpha_2,\ldots,\alpha_r)$ depending on n state variables $(x_1,\ldots,x_n)$ and r parameters $(\alpha_1,\ldots,\alpha_r)$. These r-parametrized function families are assumed smooth (differentiable) for all real values of $x = (x_1,\ldots,x_n)$ and $\alpha = (\alpha_1,\ldots,\alpha_r)$. For most values of $\alpha = \bar{\alpha}$, the functions $f(x,\bar{\alpha})$ are generic, that is, have only non-degenerate critical points. But for certain catastrophe values $\alpha = \alpha^*$, the function $f(x,\alpha^*)$ will have degenerate critical points (x^* where $\delta f/\delta x^i = 0$, but $\det|\delta f/\delta x^i \delta x^j| = 0$). Moreover, for generic function families only certain kinds of degenerate critical points arise, and the transition through these degenerate critical points is given geometrically by the elementary catastrophe models.

This construction can also be noted from the viewpoint of dynamics and applied mathematics. For instance, assume we observe some phenomenon with state or measured effect x (possibly with n-components, so we write the state vector $x = (x_1,x_2,\ldots,x_n)$) for each fixed setting of the causal or control parameters $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_r)$ that describe the environment or general exogenous levels for the background parameters. The state x will usually be found to be located at some value x^* where some "potential function" $f(x_1,x_2,\ldots,x_n, \alpha_1,\ldots,\alpha_r)$ (or merely $f(x,\alpha)$ for abbreviation) has a local minimum, or at least a stationary or critical point where the gradient $\delta f/\delta x^i(x^*,\alpha)(i=1,\ldots,n)$ is zero. For instance, we could postulate a "fast or microdynamics system" satisfying the gradient law

$$dx^i/dt = -\delta f/\delta x^i , \quad i=1,\ldots,n ,$$

with equilibria states x^* where $\delta f/\delta x^i = 0$.

It is therefore of great interest to plot the loci $x^*(\alpha)$ describing the observed equilibrium states as they vary slowly when the control or environmental parameters α are modified in some possible "slow or macrodynamical system." It is precisely such a graph M_f called the critical manifold (see Figures 1 and 2) of x^* versus α that indicates the locus of all critical points of $f(x,\alpha)$ including the local minima (stable equilibria for the microdynamics), local maxima (unstable equilibria for the microdynamics), as well as other critical points including possible degenerate critical points that will be described in terms of the elementary catastrophes.

The behavior of the observed state $x^*(\alpha)$ would show a tendency to adjust rapidly to the critical manifold M_f and to stick on M_f as α varies. At various catastrophe points of $\alpha = \hat{\alpha}$ (where M_f no longer projects onto x-space in a smooth 1-to-1 manner), the observed state $x^*(\alpha)$ is forced to jump discontinuously to another sheet or branch of

M_f. In physical terminology the state x jumps abruptly from one observational mode to another mode.

These jump phenomena for multi-modal phenomena are observed in physical, biological, and behavioral systems and it seems paradoxical that such a sudden discontinuity can arise from a gradual continuous variation in the parameter α. However, in terms of the function family $f(x,\alpha)$ the jump phenomenon is not mysterious, but merely rests on the obvious geometrical fact that the number of bends in a curve (or critical points on the graph of a surface) can increase or decrease by an integral number, even though the curve itself is being slowly modified in a continuous manner.

These geometric analyses of function families and the applications of these ideas will be pursued in the next sections of the paper.

Today's Catastrophe Theory.

Mathematics is eternal, according to Plato, and so today's catastrophe theory is merely the way we regard the geometrical elementary catastrophes today. Let us first describe a few of the seven elementary catastrophe models and then conclude this section with a formal statement of Thom's discovery in the language of Zeeman and Markus.

In Figure 1, we show a transition between a curve displaying a minimum to a curve with no critical points (see Holt, 1977). Then a replay of the sequence with suitable reversals produces a curve displaying a maximum. For this purpose we study the generic function family of curves $y = f(x,\alpha)$ that depend on a single real state variable x and a real parameter α.

The generic transition occurs when a local minimum meets a local maximum and these two critical points annihilate each other at the catastrophe to produce a curve with no critical points. [Note in Figure 1(a) the parameter α decreases leftwards, to show this sequence].

Often we consider the parameter α as quasi-static or slowly varying in some macrodynamics as the environment of the phenomena is gradually shifted. But the state variable x follows a fast microdynamics for each value of the parameter α. We study the case where x follows the gradient of $f(x,\alpha)$ towards the equilibrium states x^* where $df/dx = 0$; that is, the critical points x^* are the only observable values of x as a steady state. The fast dynamics $dx/dt = -f_x$ are shown in Figure 1(b).

A further plot of these steady state observable values of x for each parameter value α shows a folded curve, the "critical manifold" in Figure 1(c). Hence we attach the name "fold catastrophe" to the function family $f(x,\alpha) = x^3 - \alpha x$. This function-family $(x^3 - \alpha x)$, containing the non-generic function x^3, is the first of the elementary catastrophe models.

If we consider a function family $f(x_1,x_2,\alpha)$ with two state variables (x_1,x_2) and one parameter α, then the catastrophe model is still called a fold, since only one of the variables at a time, say x_1, can enter into the precise moment of transition under generic conditions. That is, near $x_1 = 0$, $x_2 = 0$, $\alpha = 0$ the family has the qualitative form

$$f(x_1,x_2,\alpha) = (x_1)^3 - \alpha(x_1) \pm (x_2)^2.$$

This still displays the fold catastrophe in (x_1) and the gradient of f in the direction of (x_2) remains unchanged and uninteresting. It is in this sense that Thom asserts the uniqueness of the catastrophe that depends on a single real parameter α, and this explains the importance of the fold catastrophe $f(x,\alpha) = x^3 - \alpha x$.

If we now allow two real parameters (α_1,α_2), then there is exactly one catastrophe (in addition to the fold) and this is the model of the cusp catastrophe illustrated in Figure 2. The appropriate function family is $f(x,\alpha_1,\alpha_2) = x^4 - \alpha_2 x^2 - \alpha_1 x$ and this describes a generic transition from $(x^4 - x^2)$ with a single critical point. (The cubic term is eliminated by simple preliminary coordinate transformations without affecting the qualitative form.) The cusp catastrophe is of primary importance in discussing multi-modal phenomena that depend on two exogenous parameters. Since this geometric situation is so important in both pure and applied mathematics, we offer next a somewhat technical description and analysis of this elementary catastrophe that is pictured in Figure 2 (see Holt & Markus, 1978; Holt, 1977).

Consider first the cusp as described in terms of the 4th degree function family

$$f(x,\alpha_1,\alpha_2) = x^4 - \alpha_2 x^2 - \alpha_1 x.$$

Here the real-valued function of the single real state variable x can possess two local minima (where $f' = 4x^3 - 2\alpha_2 x - \alpha_1$ has two or more real zeroes, for instance at $\alpha_1 = 0$ and $\alpha_2 > 0$ where we compute $x^* = \pm\sqrt{\alpha_2/2}$ as local minima and $x^* = 0$ as local maximum). On the other hand f possesses only one local minimum when f' has only one real zero (for instance $\alpha_1 = 0$ and $\alpha_2 < 0$ where $x^* = 0$). The geometric locus in the (α_1,α_2)-plane separating the bimodal from the single-modal behavior of the equilibrium state x^* (as observed at the local minima of f) is called the bifurcation locus and is here described by the cusp curve

$$27(\alpha_1)^2 - 8(\alpha_2)^3 = 0.$$

In more detail the phenomenon described by observations of the local minimizing states for

$$f(x,\alpha_1,\alpha_2) = x^4 - \alpha_2 x^2 - \alpha_1 x$$

bifurcates from a single-modal to a bimodal behavior whenever

$$f' = 4x^3 - 2\alpha_2 x - \alpha_1$$

has a single real root x^* that splits into two real roots as (α_1,α_2) changes. This bifurcation occurs at a double root of f', that is, at a common zero of

$$f' = 4x^3 - 2\alpha_2 x - \alpha_1 = 0 \quad \text{and}$$

$$f'' = 12x^2 - 2\alpha_2 = 0.$$

Upon elimination of x from the equations f' = 0 and f" = 0, we find the required cusp bifurcation locus in the (α_1,α_2)-plane. In Figure 2, we illustrate the corresponding configurations for f, f', the bifurcation cusp, and the observed state on the critical manifold M_f.

The region bounded by the cusp bifurcation curve is covered by three sheets of the critical manifold M_f, but only the top (upper) and bottom (lower) sheets correspond to local minima of the potential function f, and so these sheets yield observable values of the state x^*. The middle sheet corresponds to a local maximum of f, and hence the state x would diverge rapidly (under the assumed microdynamics) away from the middle sheet of M_f, because of measurement noise and external disturbances, so x would not register an observable value there.

In terms of the microdynamics for the state x at each fixed setting of the parameter $\alpha = (\alpha_1,\alpha_2)$, only the upper and lower sheets of M_f constitute "attractors" or stable equilibria for the state variable x; but the nature and orientation of these attractor sheets depends on the setting of the control or environmental parameter α. Of couse, outside the region of the (α_1,α_2)-plane bounded by the cusp curve, the single sheet of M_f provides the unique attractor for the observable mode of the state.

If we next follow a macrodynamical time trajectory for $(\alpha_1(t),$

$\alpha_2(t))$ at slow speed in the control plane, then $x^*(t)$ will follow the attractor surface M_f as closely as the geometry permits. For instance, if we take the macrodynamics along $\alpha_1 = -t$, $\alpha_2 = +1$ to define a slow parameter adjustment along the line $\alpha_2 = +1$ moving from right to left in Figure 2, then for $t < 0$ the state $x^*(t)$ rides on the upper sheet of M_f, and evolves continuously until the moment $t = \sqrt{8/27}$ at the control point $\alpha_1 = -\sqrt{8/27}$, $\alpha_2 = +1$; whereupon the upper sheet of M_f disappears and the observed state x^* jumps discontinuously to the lower sheet of M_f for its subsequent evolution. Upon reversing the macrodynamical adjustment, the jump to the upper sheet of M_f occurs at the point $\alpha_1 = \sqrt{8/27}$, $\alpha_2 = 1$. Thus an asymmetry occurs in the behavior of the state x^* and this shows that current values of the causal parameters (α_1, α_2) are not sufficient to determine the state of the system, but the whole past history of the process is required to specify the observed state. Such jump behavior and history-controlled violent processes are familiar in many contexts, particularly in the behavioral sciences.

For the moment we only wish to point out that this classification of discontinuities is local in format (usually normalized to a neighborhood of $x = 0$, $\alpha_1 = 0$, $\alpha_2 = 0$) and is qualitative in nature. For example, jumps along the "fold line" of M_f (near which the surface M_f is merely a fold curve displaced trivially to form the surface) are essentially like the fold curve illustrated in Figure 1. The only new type of catastrophe in Figure 2 occurs in the locality of the cusp point $\alpha_1 = 0$, $\alpha_2 = 0$, where the two fold lines run together. At this cusp point, the cusp catastrophe arises from the "unfolding" of the function family generated by x^4.

Similar analyses hold for higher order catastrophes (the swallowtail, butterfly, and wigwam catastrophes) but these have 3-, 4-, and 5-dimensional control spaces. In these cases we cannot give any helpful pictorial representation of the critical manifold M_f. There are even more complicated catastrophes that involve two or more state variables, but we shall not go beyond the familiar fold and cusp catastrophes in this paper.

With these preliminary examples and motivations in mind, we can now assert the mathematical theorem of R. Thom, as modified by Zeeman and rephrased by Markus. The mathematical terminology is technical but standard and has a precise meaning for which we refer to more advanced mathematical texts.

Theorem Among the totality of smooth function families of the form $f(x_1, x_2, \ldots, x_n, \alpha_1, \ldots, \alpha_r)$ (with $r \leq 5$), those families that are locally structurally stable constitute a generic set. Moreover, near each degenerate critical point the qualitative form of such a locally structurally-stable function family is that of an elementary catastrophe. The number of types of elementary catastrophes are as follows:

r	1	2	3	4	5
# Elem. Cat.	1	2	5	7	11.

For $n = 1$ and $r = 1$ there is only the fold catastrophe described by $f(x,\alpha) = x^3 - \alpha x$ near $x = 0$, $\alpha = 0$. For $n = 1$ and $r = 2$, there is in addition the cusp catastrophe described by the "Thom polynomial"

$$f(x,\alpha_1,\alpha_2) = x^4 - (\alpha_2)x^2 - (\alpha_1)x \quad \text{near } x = \alpha = 0.$$

For $n > 1$ these Thom polynomials in the state variable $x = x_1$ must be supplemented by non-degenerate quadratic forms in the remaining $(n-1)$ state coordinates.

The fundamental existence and uniqueness theorem of Thom asserts, in the case of greatest interest for applications to behavioral sciences:

For a single state observable ($n = 1$) there exists a unique 2-modal qualitative geometric form having at most two significant causal or control parameters--namely, the cusp catastrophe. Further, there exists a unique 3-modal qualitative form having at most four significant control parameters--namely, the butterfly catastrophe. Thus the cusp catastrophe yields the simplest description of a 2-modal phenomenon, just as the butterfly catastrophe gives the simplest description of a 3-modal phenomenon.

Tomorrow's Catastrophe Theory.

We conclude this paper with some tentative examples and incomplete investigations on the theory and the practicality of the method of catastrophes. First we discuss some research problems in the pure mathematics of catastrophe theory within the tradition of geometry. Later we conclude with some remarks on topics in applied mathematics, linking catastrophe analysis to behavioral science and to general questions of scientific methodology.

A. Theory of Catastrophes

Let us consider a curve, like $y = x^4 - x^2$ in Figure 2, that approaches $+\infty$ as $|x| \to \infty$. We count the number of critical points, all either local maxima (just one in this example) or local minima (two in this example), since these are the only possible types under assumptions of genericity. It is easy to determine that, for such generic curves $y = f(x)$, a universal algebraic formula holds concerning the number of maxima and minima:

$$\#\min - \#\max = 1.$$

This is a very simple example of the general algebraic theory of critical points developed by Morse. But Morse found analogous complicated results for the numbers of critical points of generic functions defined on higher dimensional spaces, and he related these critical points to the geometry and topology of the underlying space. For instance, if we study $f(\theta)$, where θ is an angular coordinate on a circle (topologically different from the real x-axis) then a different algebraic result holds:

$$\#\min - \#\max = 0.$$

Next, suppose two different generic curves are given f_0 and f_1 (say, defined for θ on the circle--to avoid analysis at ∞). Now let us try to interpolate between these two functions using a real parameter α on $0 \leq \alpha \leq 1$. That is, we seek a function family

$$f(\theta,\alpha), \text{ for } \theta \text{ on circle, } 0 \leq \alpha \leq 1,$$

so $f(\theta,0) = f_0(\theta)$ and $f(\theta,1) = f_1(\theta)$. Moreover let us demand that $f(\theta,\alpha)$ is a generic function family; then we know that only fold catastrophes can enter among the critical points.

We propose a research problem to classify and count the types and numbers of catastrophes that can (or must) enter into such generic function families on the circle or on higher dimensional state and parameter manifolds. In other words, construct some kind of a generalization of Morse-theory for catastrophes. A start has been made in this project by Markus and others, but no satisfactory theory has been developed (see Markus, 1979).

Now let us turn to a different type of problem involving dynamical systems for the state x of a particle moving along a line, subject to a force $-df/dx$ arising from a potential $f(x)$. Then Newton's dynamical law yields $\ddot{x} = -df/dx$.

If $f(x) = x^2$, then the dynamical system is a linear oscillator about the critical point $\ddot{x} + 2x = 0$. In terms of the $(x,y = \dot{x})$ state-plane, the system becomes $\dot{x} = y$, $\dot{y} = -2x$ with solution portrait, shown in Figure 3. On the other hand, if $f(x) = -x^2$, then the dynamical system is that for a linear repulsive force, namely an unstable saddle point, also shown in Figure 3.

The research problem concerns the possible interpolation between these two linear systems, with a parametrized function family $f(x,\alpha)$. If α is a single real parameter then the transition is at a fold catastrophe $f = x^3$ with $\ddot{x} + 3x^2 = 0$, and the corresponding solution portrait shows no periodic orbits.

On the other hand if $\alpha = (\alpha_1,\alpha_2)$ then the nature of the cusp catastrophe gives the certainty of an intermediate dynamical system like $\ddot{x} + x^3 = 0$ (translations of the coordinates to bring the critical point to the origin, and qualitative transformations of coordinates are used freely in this casual analysis). But this last system does have an infinite family of periodic orbits with the remarkable novelty that there are arbitrarily long periods (see Markus, 1971).

Can we construct a coherent and sensible mathematical theory of parametrized dynamical system that utilizes the full knowledge of the elementary catastrophe critical points?

B. Applications of Catastrophes

As far as practical programs and actions go, can we take this subject seriously--all this talk about conics, quadrics, bifurcations along bending wires, and pleating paper sheets? Is this a worthwhile intellectual pursuit for healthy adults? The query is blasphemous.

No subject studied in universities is more important than the contrast of an elliptic-paraboloid (hilltop) and a hyperbolic-paraboloid (saddle col). The distinction between an ellipse and a hyperbola is precisely that between an attractive and a repulsive force; that between an electron and a proton; that between Newtonian mechanics and Einsteinian mechanics; that between welfare socialism and market capitalism. So enough of that!

But can we become highly emotional over this subject? How can catastrophe theory become controversial, make life-long personal enemies, and invoke rage and fear? Why is there an enthusiastic pro-catastrophe article in the Scientific American (Zeeman, 1976) while there is a sarcastic anti-catastrophe article in The Sciences (Sussmann & Zahler, 1977)?

The irritation, rage, fear, and controversy have been induced by the manner of the exposition of some of the more popular versions of the applications of catastrophe theory. In these examples, the dynamics are not presented in deterministic terms, say rocks rolling under prescribed forces, as much as in teleologic terms describing new insights into the overriding purposes and preferences of natural phenomena. In these popular approaches, catastrophe theory is claimed to stake out limitations on the human imagination and on the concepts of subjective and objective reality. In other words, catastrophe theory is presented as a new philosophical theory of epistemology.

In order to bring these rather fragmentary philosophical comments into focus, let us consider, in abbreviated form, the famous example explaining how a dog bites a man (see Figure 4).

AGGRESSION IN DOGS can be described by a model based on the cusp catastrophe. The model assumes that the aggressive behavior of the dog is controlled by two conflicting factors, rage and fear, which are plotted on a horizontal plane, the control plane. The behavior of the dog directed against the man ranges from attacking to retreating and is represented on the vertical axis. For any combination of rage and fear there are as many as three possible states of aggression for the dog, as read from the graph of the behavior cusp-surface. We assume that the cusp catastrophe must be the central feature of this problem in animal behavior since there are two control parameters, and Thom has proved that the cusp-surface is the only surface displaying the significant geometric phenomena of multi-modal behavior with jump discontinuities between the sheets of the surface. In other words, we select the cusp-catastrophe because we have made a mental commitment to utilize one of the elementary catastrophes of Thom and this is the only elementary catastrophe involving just two control parameters.

If an angry dog is made more fearful, its mood follows the trajectory A on the control plane, as interpreted in Figure 4. Similarly, a frightened dog that is angered follows the trajectory B. The dog's behavior remains on the bottom sheet of the cusp-surface until that sheet disappears, then it jumps to the top sheet and the dog attacks and jumps at the man. This explains how a dog bites a man.

How are we to interpret this story? Is it a serious study in animal behaviorism or is it a useful illustration of the geometric features of the cusp-catastrophe with some casual indications on how this approach might conceivably be useful in some later psychological studies? I am more inclined to the latter choice. It is not difficult to criticize this example; the difficulty is to defend it.

The defense of this method of dealing with natural phenomena directly by the geometry of the elementary catastrophes rests on two principles in the philosophy of mathematical modeling:

1. Structural Stability (repeatability or robustness)
   Small perturbations of the data defining a natural phenomenon will not change the qualitative structure of our model of the phenomenon.

2. Genericity
   Our model should ignore or eliminate any unnecessary assumptions of symmetry, causation, or special numerical coincidences. (Newton said, "I make no hypotheses.")

In summary, we recapitulate our basic philosophy, namely, the elementary catastrophes are a good place to start in the modeling of natural processes that display the phenomena of multi-modal behavior, jump discontinuities, hysteresis, and divergence of development under slight variations in the environmental modifications. Moreover, catastrophe theory provides a conceptually coherent, philosophically attractive, and mathematically tractable and workable mechanism for the analysis of complex natural phenomena. But one must always be aware that catastrophe theory is merely an analytical tool, and a tool is practical and useful only in the hands of an experienced and skillful master of his discipline or craft.

COMMENTARY

Markus demonstrates that catastrophe theory, as mathematics, is part of a long tradition of geometric results. It has both depth and substance.

A strong theme is the use of quadratic polynomials or conic sections as the basis of scientific models. But Markus is most concerned about the transition between one such form and another, such as that between ellipse and hyperbola. The traditional Apollonian transition, the parabola, appears from the modern viewpoint to be the "wrong" intermediary. It is too degenerate and the correct transition should bring in cubic or higher degree terms to the equations.

To elucidate Markus's meaning here, consider the mathematically simpler but analogous problem of finding a transition between x^2 and $-x^2$. The "obvious" way to achieve this is to take the 1-parameter family of quadratics ax^2 and let a (the parameter) run from -1 to +1. This gives the transitional form as 0.

But to a topologist, zero is hopelessly and enormously degenerate, far too special to be appropriate. So what should it be replaced by? Qualitatively the significance of the original two quadratics is that they define certain standard structurally stable forms, and if higher terms such as cubics are added to them, the result remains equivalent to the original. These higher order terms may thus be ignored precisely when structural stability holds, and the place they may not be ignored is at the transition point. So the "hidden" cubic terms must be added back in. Now we are looking for the transition between, say, $x^2 + x^3$ and $-x^2 + x^3$ so we introduce the parametrized family $ax^2 + x^3$ and run a from -1 to 1. The transition then becomes not zero, but x^3.

Even this, topologically, turns out not to be the "right" answer. The most natural transition requires a constant fourth-order term and a linear term with an additional variable parameter--precisely, the cusp catastrophe.

It is natural because it renders the entire transition process structurally stable. Although the catastrophe theorist's emphasis on that property should be treated with caution and in a suitable context, critics take it far too lightly. Structural stability and genericity are highly important and desirable properties, once correctly formulated within a given context, and this point of view has been urged by many eminent mathematicians during the past fifty years: Pontryagin, Lefschetz, Smale, Arnold, Thom.

While professors Hastings and Markus were on campus, a discussion was held on "competitive vs. cooperative systems." (Participants in this discussion included professors Hastings, Markus, Johnson, Lacher, graduate students Carney, Castellano, Duckwall, and Fletcher.) A model emerged which we now describe.

<u>A model describing transition between cooperative and competitive systems.</u>

A society consists of two (groups of) citizens A and B who communicate all information freely, but who can act individually (with only self-interest) to improve his (or her) own position. A "state" of the society is a point (x,y) in the x,y plane.

Citizen A has access only to the x-coordinate; he can increase or decrease x in small (or infinitesimal) steps. Similarly B has access and can change only the y-coordinate. Each citizen moves (relatively slowly) his coordinate until he is satisfied that no further improvement is available to him personally. Personal satisfaction is determined by net pay which in turn is affected by the state of society (i.e., the x,y coordinates) and the rules governing distribution of wealth.

<u>Cooperative system.</u> Citizens share wealth equally, so each is paid an amount of one-half of society's total wealth, (i.e., $(T-x^2-y^2)/2$). The government gets x^2+y^2 for communal services.

<u>Competitive system.</u> Citizen A is paid $(T-x^2+y^2)/2$ and B is paid $(T+x^2-y^2)/2$. The government gets nothing (except, in an expanded model, what it <u>earns</u> as a citizen).

In either system, the constant T is a theoretical total wealth of society (sort of a gross national product (GNP)).

Note that in either model, each citizen maximizes his personal income by moving his coordinate toward zero, so that the origin $x=0$, $y=0$ is a (the only) stable equilibrium in either case. In the cooperative system, tax revenue x^2+y^2 is minimized to zero. In the competitive model, if A is the government, say, then A's income (as well as B's) is optimized.

Examining the two income functions, for B, say

$$(T-x^2-y^2)/2 \quad \text{vs.} \quad (T+x^2-y^2)/2,$$

we recognize a change of form of the type discussed above.

In fact, a linear change of coordinates (affine transformation) reduces these two income functions to

$$x^2/2 + y^2/2 \quad \text{vs.} \quad -x^2/2 + y^2/2$$

Catastrophe theory says that the simplest structurally stable form containing these two is

$$ax + bx^2/2 + x^4/4 + y^2/2$$

a cusp catastrophe. (Here "simplest" means up to pseudolinear coordinate changes, assuming no symmetry constraints).

We will not speculate on the possible meaning of the two control parameters a,b in this context; they are given to us from the "unfolding" of the x^2 singularity, so they arise from legitimate mathematical considerations. The "cooperative" and "competitive" models are embedded into a structurally stable parametrized family of models. Investigation of these models and possible transition paths between competitive and cooperative could prove interesting.

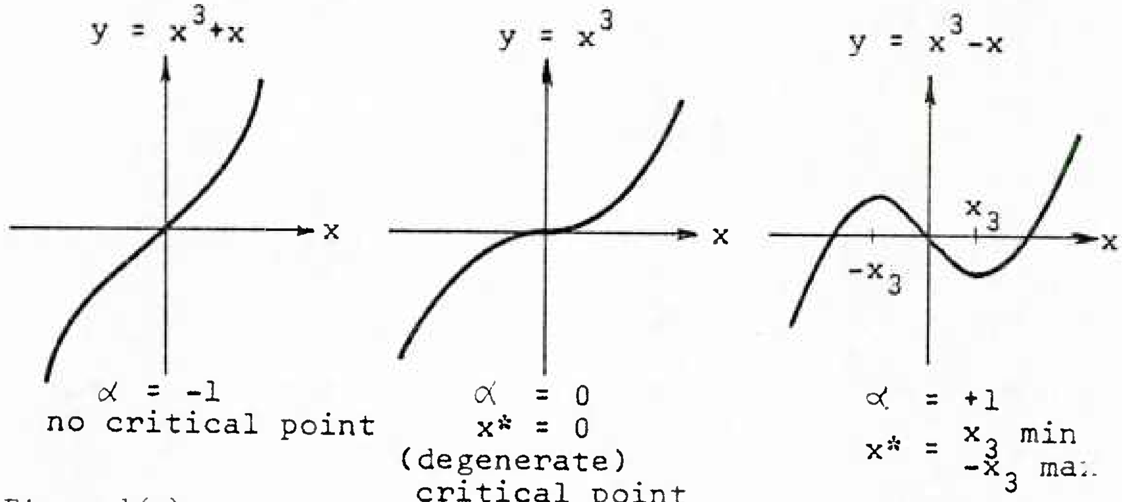


Figure 1(a).

Microdynamics $\dot{x} = -f'$

←— ←←←← ←— $\dot{x} = -3x^2-1$

←— ←←•←← ←— $\dot{x} = -3x^2$

←—←←•→ —→ →•←← ←— $\dot{x} = -3x^2+1$

Figure 1(b).

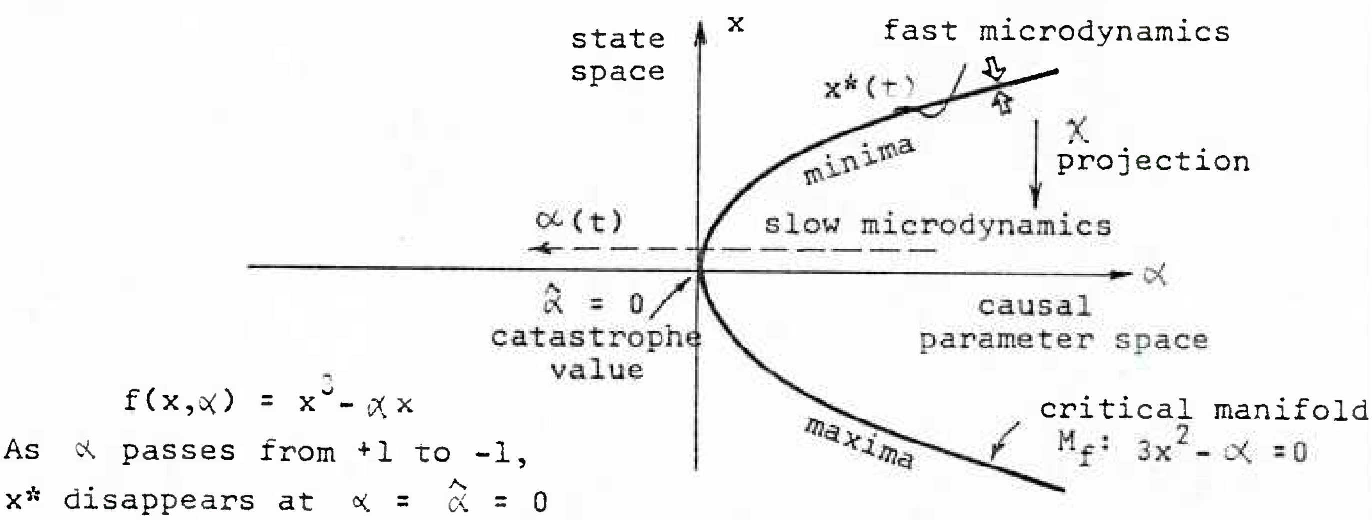


Figure 1(c). FOLD CATASTROPHE $f = x^3 - \alpha x$

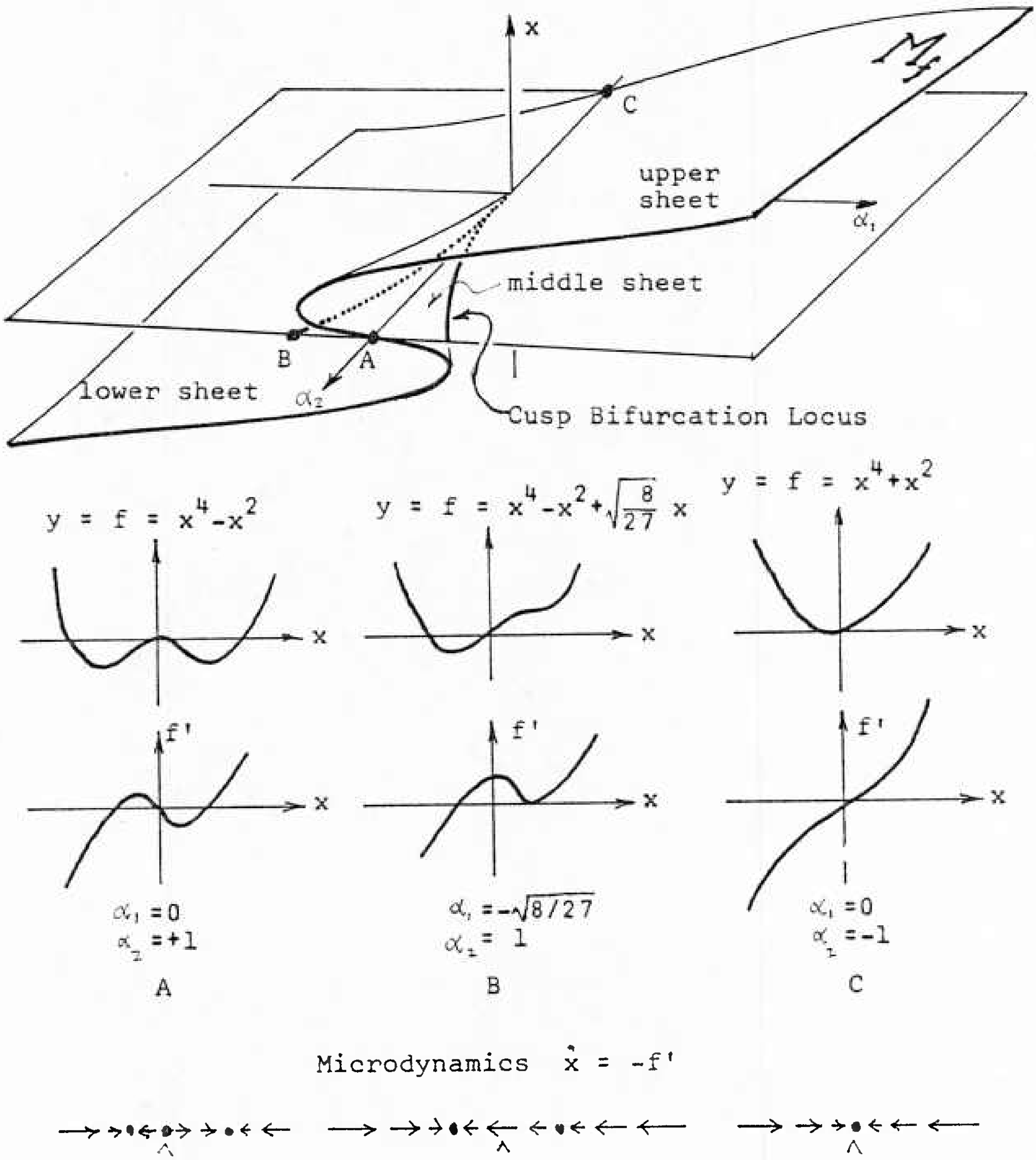


Figure 2. CUSP CATASTROPHE $f = x^4 - \alpha_2 x^2 - \alpha_1 x$

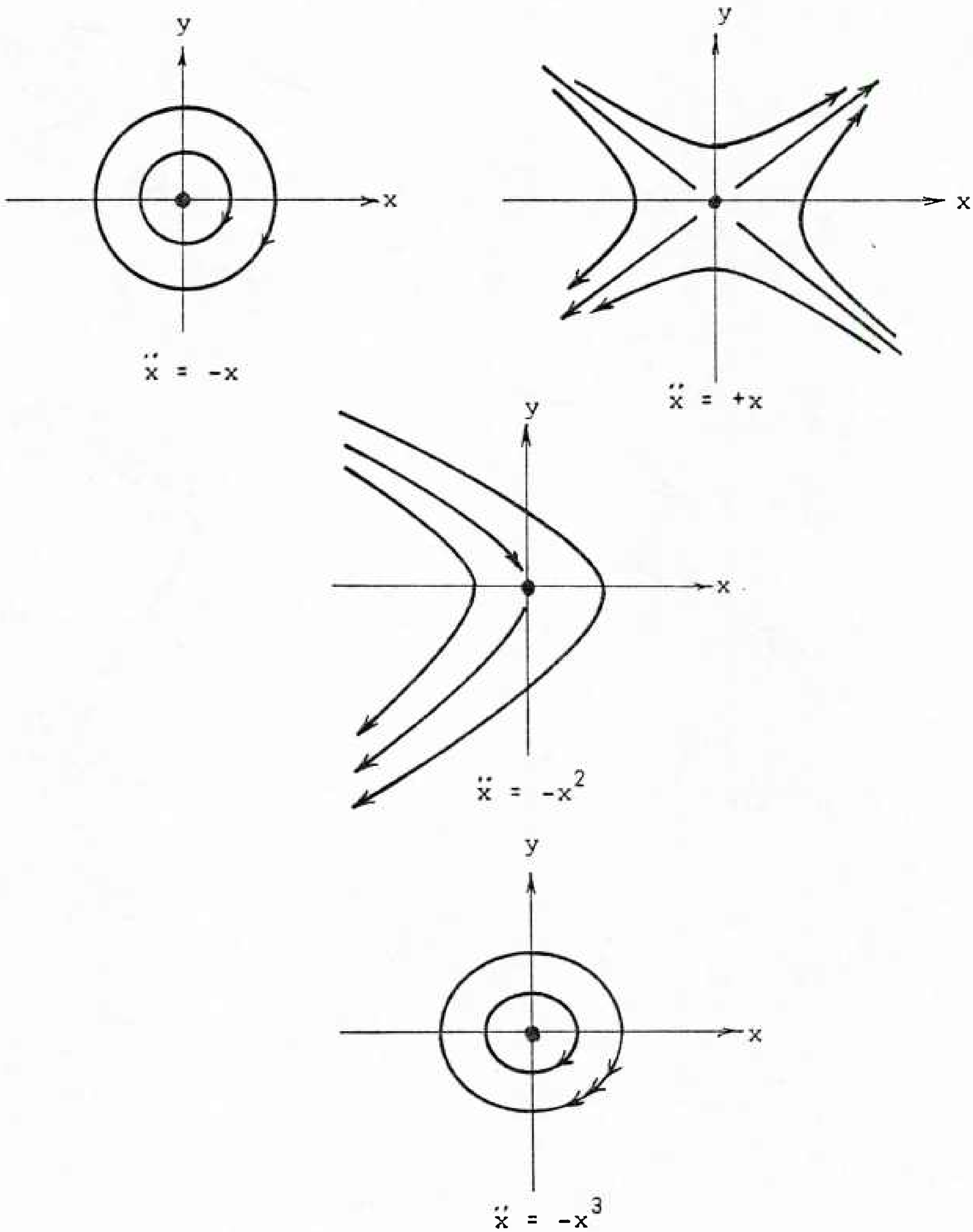


Figure 3. BIFURCATIONS THROUGH CATASTROPHES

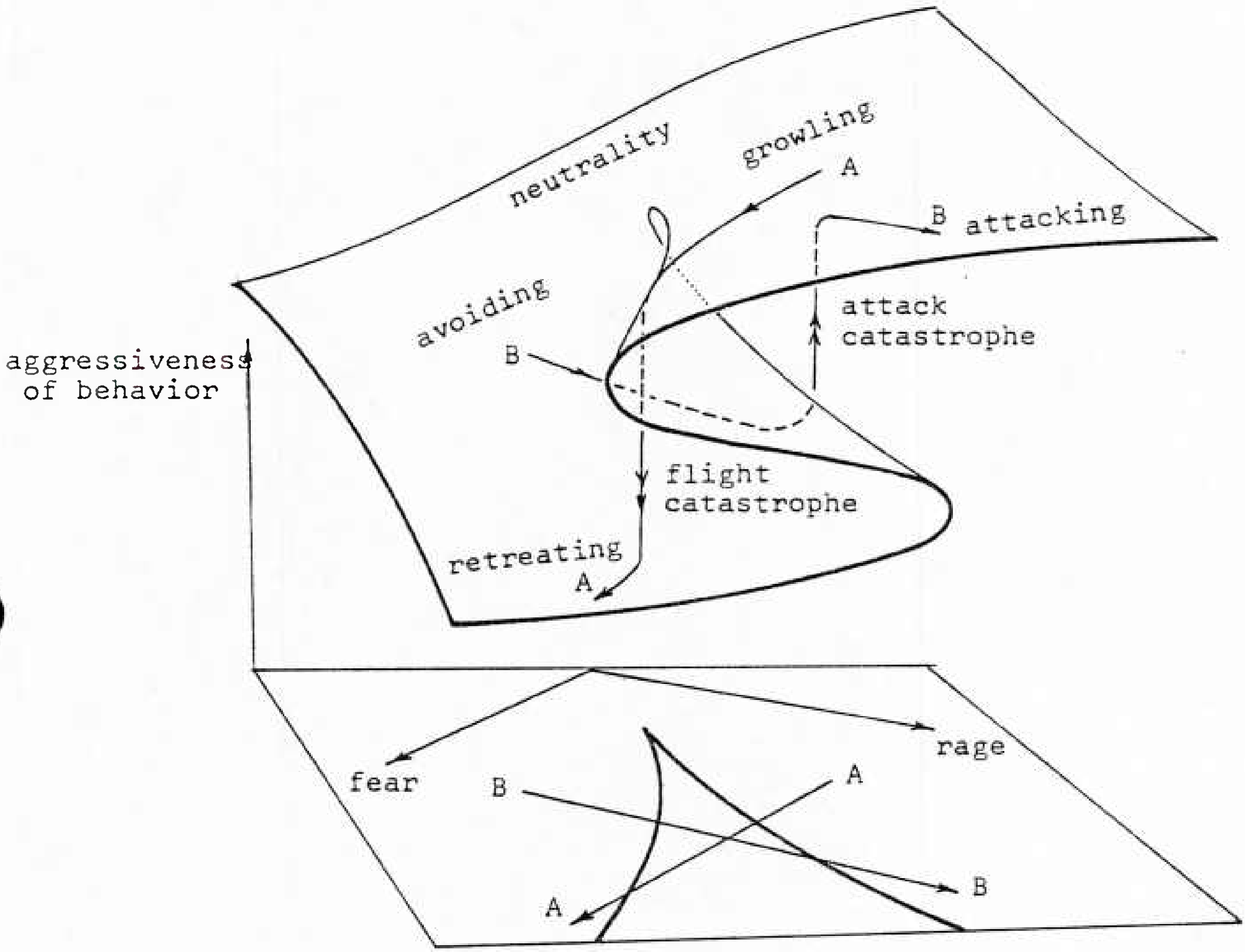


Figure 4. AGGRESSION IN DOGS. This figure is sketched from E. C. Zeeman's well-known Scientific American article, 1976.

LETTERS

COMMENTS ON THE ROLE OF CATASTROPHE THEORY IN THE SOCIAL SCIENCES

Carl P. Simon

In their papers in this collection, Professors Holt and Stewart are enthusiastic about catastrophe theory and its applications in the social sciences. Both authors can back up their enthusiasm by pointing to the incisive research articles they have written applying catastrophe theory in a variety of efficient areas. Furthermore, Stewart's books with Poston (1976, 1978a) on catastrophe theory are the best written, detailed introductions to the subject.

As Stewart (and Markus) points out in these proceedings, catastrophe theory "provides a versatile and flexible range of nonlinear models which are simple and natural from the mathematical viewpoint" (page 21 of this work). It is an aesthetic and potent technique for handling modeling situations which involve sudden discontinuities, hysteresis, multimodal distributions or divergent behavior from similar input paths. When carefully applied, catastrophe theory can also provide a rich modeling vocabulary which draws the modeler's attention to such key concepts as genericity, structural stability, and behavioral surfaces. Finally, although it is based on some incredibly sophisticated mathematical theorems, it can usually be explained rather quickly and completely to anyone who is comfortable with elementary mathematical concepts so that the student can not only read papers which use catastrophe theory but even consider writing one himself.

However, this ease in quickly explaining the concepts of catastrophe theory to a neophyte leads to one of the major problems that has arisen in the controversy over catastrophe theory, namely the belief by some that one can make catastrophe theory the cornerstone of mathematical modeling in the social sciences. I do not think that Holt and Stewart share this belief, but some of the passages in their papers appear to encourage it. For example, after lamenting "the paucity of well-articulated mathematical structures that are appropriate for the classification and description of social phenomena" (page 42 of this work), Holt writes in his article:

> We are suggesting that catastrophe theory models . . . do provide a very useful classificatory structure in which to construct a theory, just as the geometry of conic sections provides models in terms of which orbits and trajectories can be classified. While astronomy involves extremely accurate observations of a dynamical system following a deterministic evolution, behavioral sciences deal with much less precise data having a different sort of causal basis. These epistomological diversities are reflected in our choice of classificatory geometric models, as we turn from the quantitative elliptic

> curve of Kepler to the qualitative butterfly surface of catastrophe theory. But the principle of describing empirical phenomena in terms of a mathematical structure remains the same.

Stewart is not quite so direct. In his section on model selection, he points out that the "traditional models used in social science are overwhelmingly linear as regards the underlying mathematical structure . . . . A major reason for this emphasis on linearity . . . is that nonlinear mathematics is usually much harder than linear mathematics" (page 18 of this work). After commenting that "the process of model-selection is necessarily heuristic," Stewart concludes that "the catastrophe theory approach is a natural one, based upon a proper understanding of the mathematical possibilities and their relative simplicity or complexity" (page 21 of this work).

As enthused as I am about the successes--both actual and potential--in using catastrophe theory in social science models, I do not think it is at all wise to encourage hopes that catastrophe theory be <u>the</u> (or even <u>a</u>) cornerstone of social science modeling. After all, catastrophe theory is just one tool in the standard math modeler's toolbox--a toolbox which should certainly include other techniques such as matrix theory, calculus, optimization (constrained and unconstrained, linear and nonlinear), the implicit function theorem and "comparative statistics," ordinary and partial differential equations, difference equations, optimal control, topology, functional analysis, graph theory, combinatorics, game theory, statistics, probability, and stochastic processes. A good physicist or engineer is familiar with most of these areas. I don't think that it is a coincidence that some of the most exciting uses of catastrophe theory, such as buckling of beams, stability of ships, and geometric optics (see Poston & Stewart, 1978a; Guttinger & Eikemeier, 1979) have come in physics and engineering and have involved research in which catastrophe theory did not play the major role but mostly helped to choose the most effective functional forms at certain stages or to sort out the analysis at certain bifurcation points.

There are certainly other reasons why catastrophe models have had more success in the physical sciences than in the social sciences. One can quickly point in the former to the exactness of measurement, the longer history of quantitative modeling, and the absence of unpredictable interactions with living organisms, such as human beings and their communities. [See Chapter 17 of Poston & Stewart, 1978, for a more complete discussion of the difficulties of mathematical modeling in the social sciences.] However, I am still struck by a comparison of the inclination of physicists and engineers to use catastrophe theory as simply one important technique among many with the tendency of many social scientists to base their modeling almost completely on the tenets of catastrophe theory. Even among social scientists, I have found that the richer the array of mathematics used in models which involve catastrophe theory, the more convincing are the conclusions of the model, with the shallowest models simple attempts to show that the global graph of one of the catastrophe surfaces is the behavioral surface of some social phenomenon.

Both Holt and Stewart place much more emphasis on fitting data to curves and surfaces as a role of mathematical modeling than I feel comfortable with. [See, for example, Holt's discussion of Kepler's work at the foundation of modern astronomy and mechanics, page 41 of this work.]

In addition, I find it strange that social scientists should embrace catastrophe theory as a possible cornerstone of modeling in the field when most social science journals (outside economics) will reject as too mathematical any article which contains significant amounts of simpler, more basic mathematical concepts, such as the implicit function theorem, constrained non-linear optimization, and optimal control theory.

So, while Stewart decries "a shortage of . . . mathematicians with an interest in developing models in social sciences" (page 31 of this work), I feel that one should also lament the lack of mathematical interest that seems common in programs in political science and sociology. It is certainly very easy to earn doctorates in either of these fields without ever taking a course in calculus as an undergraduate or graduate. Researchers in these fields with some mathematical sophistication are still somewhat rare and are still usually compelled to publish their calculus-based results in journals which are out of the mainstream of their fields.

In discussing recent work on the use of the cusp catastrophe in bifurcations to multimodal distributions, Stewart asserts that "we should learn to walk before we try to run" (page 24 of this work). In the same spirit, I would summarize my comments by encouraging social scientists to bring calculus-based modeling techniques into the mainstream of social science research before we turn them on to catastrophe theory as a possible cornerstone of their basic nonlinear modeling. Besides making these models richer and more versatile, this will also help these scientists understand better the actual ingredients of catastrophe theory, such as its reliance on gradient dynamics (at least in dimensions greater than one) and the fact that the theory gives a quantitative description of behavioral surfaces only in small regions around bifurcation points.

COMMENTS ON SOME CURRENT DYNAMIC PROBLEMS IN ECONOMIC MODELING

Carl P. Simon

Economics appears to sit right in the middle of the scientific spectrum--much less exact than physics, engineering, or even biology, but more quantifiable than political science and sociology. Unlike most political science students, undergraduate economics majors usually must take a solid calculus course while graduate economics students at top universities are usually required to learn most of the mathematical topics I listed in the previous comments section before they can begin to write their dissertation. (Nevertheless, unlike physical science graduate students, they usually squeeze their math and statistics courses into one intense graduate year instead of spreading them out throughout the program and basing other courses on their earlier mathematical ones.) Just about any economics journal will accept articles describing sophisticated mathematical models. No one would dream of encouraging economists to make catastrophe theory a modeling cornerstone, even though abrupt change, hysteresis, and stability are important economic phenomena. Economists, such as Balasko (1978) and Varian (1979), have successfully used catastrophe theory as one technique among many which can shed light on important economic phenomena.

I guess my suggestions for the use of catastrophe theory in administration should be obvious. I would encourage researchers building and studying models of administration to learn about and use techniques of catastrophe theory. But more important, they should develop a working knowledge of most of the topics listed earlier and should not single out catastrophe theory as a centerpiece. Many economics departments have a one or two semester course which tries to cover in some detail many of the mathematical topics I listed in the previous comments section. [I help teach such a course at the University of Michigan and Ph.D. students in the University's public administration program (IPPS) usually take this course along with the basic graduate courses in microeconomic and macroeconomic theory. I would sugest that this is the minimal background that anyone interested in models in administration should have. After this, a working knowledge of catastrophe theory would be icing on the cake.]

I would like to close this commentary by discussing a phrase which Stewart used often but discussed very little--"dynamical systems." I include among dynamical systems the study of ordinary and partial differential equations (linear and non-linear), difference equations, dynamic optimization, and control theory. In fact, I would probably include catastrophe theory as a subfield of dynamical systems because it really does focus on the behavior of the attractors of gradient differential equations. (I am using dynamical systems a bit more broadly than it is sometimes used, since it often focuses on stability and periodicity questions of autonomous (usually non-linear) differential equations on non-linear surfaces, based on the pioneering work of Poincare and the more recent work of Smale. See for example Smale, 1967.) There are two trends in the use of dynamical systems in economics which deserve special attention and which could bring about some significant changes in the subject. The first is the growing extensive use of partial differential equations in economic analysis. This use has occurred in a number of different areas without much connection between them. One of the first areas involved the consistency of consumer choice with rational preference relations. One says that a commodity bundle x is "revealed preferred" to commodity bundle y (abbreviated xSy) if the consumer chooses x when he can afford y too. If the binary relation S is asymmetric, i.e., xSy implies not ySx, then we say that S satisfies the weak axiom of revealed preference (WARP). If S is acyclic, i.e., $x_1Sx_2S\ldots Sx_n$ implies not x_nSx_1, then we say that S satisfies the strong axiom of revealed preference (SARP). The earliest proofs that a preference ordering that satisfies SARP can be represented by a continuous utility function U involved showing that a certain system of partial differential equations for U satisfied the necessary Frobenius conditions and were therefore solvable for U. In fact, Hurwicz and Richter's (1979) proof of this result illustrated close connections with the two fundamental principles of the thermodynamics and the existence of entropy. [See Chipman, Hurwicz, Richter, & Sonnenschein, 1971, and Sondermann, 1982, for more complete discussions of this integrability question.]

A second, rather different application of partial differential equations in economics has arisen in portfolio theory and option valuation. (Actually, both of these applications do have one major factor in common--both studies were pioneered by the work of Paul Samuelson.) A call option on a stock with exercise price c and expiration date T is basically a ticket which entitles its bearer to buy a share of the stock at the terminal date T, if he wishes, for a specified price of c dollars. If the stock price S_T at time T is greater than c, then the option bearer can buy a share at T for c dollars, resell it for S_T dollars, and make a profit of S_T-c (less brokerage and option costs). Using principles of Brownian motion, Black and Scholes (1973) developed and solved a second-order partial differential equation which showed exactly how a rational investor should value such a call option. This is still an active area of research in both economics and finance departments, now involving martingales and stochastic integration. [See for example, Harrison, & Pliska, 1981.]

More recently, partial differential equations have arisen naturally in dynamic models involving differentiated commodities and consumers. In neoclassical Arrow-Debreu economic models, a finite number of consumers choose a commodity bundle, i.e., an amount x_i for each of n possible commodities; each commodity has a price per unit p_i. So the price of bundle $(x_1,\ldots,x_n)$ is simply $p_1x_1+\ldots+p_nx_n$. One can ask how these prices change over time and are then led to some interesting ordinary differential equations for the p_i's.

However, in the newer models with differentiated commodities and consumers, commodities are classified by their characteristics or qualities which lie in some subset of Euclidean space. Nearby parameters are used to characterize similar qualities. On the one hand, there are firms manufacturing goods with different characteristics; on the other hand there is a large set of consumers parametrized by their tastes who are choosing which characteristics to consume. There are natural ways of assigning prices (hedonic prices) to characteristics. [See Rosen, 1974.] So let $p(x,t)$ denote the price at time t of one unit of a good with characteristics x. Now, there is no reason to assume that prices are linear in x or even that x lies on a linear space. If one studies the process in which a short run equilibrium (supply = demand) is reached or if one goes further and studies the movement of firms as they progress from a short-run to a long-run equilibrium, one is naturally led to a partial differential equation for $p(x,t)$. These equations are usually second order and non-linear in x. However, in recent papers Sonnenschein (1981, 1983) developed an example in which the classical heat equation arose as the natural dynamic for myopic firms using short-run profits as signals for movement toward a long-run equilibrium. This author and others are working on generalizing Sonnenschein's work. Catastrophe theory has played a non-trivial role in one paper (Simon, 1979) in this direction.

The other type of dynamical system which is drawing considerable attention among economists is the non-linear difference equation. Until recently, economists automatically used differential equations to model non-linear dynamics, mostly because the mathematical theory of non-linear differential equations (e.g., the Poincare-Smale school of dynamical systems) is much more developed than the corresponding theory of difference equations (unless the difference equations are characterized by smooth, invertible functions, i.e., diffeomorphisms). On the one hand, I believe economists realize that most dynamic models involve changes over discrete units of time so that difference equations are probably more appropriate modeling vehicles. On the other hand, there is a general belief--based on experiences with the linear theory--that the solutions of the differential equations are good approximations of the solutions of the corresponding difference equations. To see how far from the truth this approximation expectation can be, consider the first-order difference equation $x_{n+1} = f(x_n)$ and its corresponding first-order autonomous differential equation $x = F(x)$. Solutions of the latter can have exactly one of the following three simple motions: (1) stationary point, (2) a path which moves asymptotically and monotonically toward a stationary point, and (3) a path which moves monotonically toward (plus or minus) infinity. Nothing more complicated can occur on the line in forward or backward time--not even simple periodic motion. On the other hand, almost anything can occur for the simplest nonlinear first-order difference equation--the quadratic $x_{n+1} = rx_n(1-x_n)$, where r is a parameter usually between 0 and 4 and $0 \le x_n \le 1$. For small values of r, all orbits tend asymptotically toward zero. However, for certain larger values of r, the orbit of a generic point fills up the whole unit interval in the sense that given any small subinterval J of

[0,1], some iterate of the point under the difference equation will eventually hit J. Furthermore, the motion is essentially random in that if one divides the interval into two subintervals $J_0=[0,.5]$ and $J_1=[.5,1]$, and if one knows which J_i contained each of the first N iterates of a generic point, one cannot predict which J_i the (N+1)st iterate will land in. To make matters even worse, these "ergodic properties" of this quadratic difference equation vary wildly with r. So, to say the least, the orbit structure of non-linear, first-order, autonomous difference equations can be an order of magnitude more complicated than the corresponding structure for the corresponding differential equation. This dichotomy increases in dimension two where difference equations can be wilder still, yet the Poincare-Bendixson Theorem tells us that bounded orbits of planar, autonomous ODE's are either stationary, periodic, or asymptotic to a stationary or periodic orbit. It is only in three dimensions that truly complicated behavior can arise for bounded solutions of autonomous ODE's.

So, economists (and applied mathematicians) have two challenges here. First, understand better the trade-offs one makes when one approximates a discrete dynamic phenomenon by an ordinary differential equation. Second, understand more clearly the complicated dynamics which I described previously for simple non-linear difference equations. Naturally enough, such dynamical systems have been called "chaotic systems." One reason they are particularly interesting to economists is the new realization that behavior that appears chaotic or random can sometimes be modeled by a simple, deterministic, possibly even first-order difference equation. Recently, Richard Day has been especially active in demonstrating that chaotic difference equations arise naturally in many economic settings and in illustrating in some cases just what this entails for the corresponding economic phenomenon.

I have described two aspects of dynamical systems which are playing a major role in current economic modeling--partial differential equations and chaotic difference equations. There is, of course, a rich diversity of other developments in dynamic economic modeling, e.g., Smale's dynamic models of repeated Prisoner's Dilemma games (1980) and, of course, the growing interest in catastrophes, singularities, and bifurcations. These developments, both separately and in their interconnections, are making dynamic economics one of the most active and exciting areas of current mathematical modeling.

CATASTROPHE THEORY AND ADMINISTRATIVE MODELS-
REACTIONS ON THE STATE OF THE ART

Harold M. Hastings

****

As evidence of the potential for applications, this article points to the ability of catastrophe-theoretic models to deal in a simple way with jump phenomena and history-dependent behavior. It places emphasis on the need for team efforts in any research projects on this topic. The subject is multidisciplinary and it is unlikely that a single individual will be able to combine the necessary expertise.

It suggests two classes of problems for further work: fairly specific and limited problems that appear feasible now, and more ambitious problems that might prove fruitful in the future.

****

1 Accomplishments.

1.1 Catastrophe theory has provided a class of building blocks to model situations where small, continuous changes in inputs produce abrupt jumps in behavior. This provides a prescription for the modeling of many phenomena in the social and administrative sciences: first describe the qualitative features of the phenomenon to be studied; then find a suitable catastrophe; and finally, build the model. However, caution is still needed in applying these models.

1.2 Catastrophe theory has provided a simple way to include the role of past history in the present dynamics of a model.

1.3 Catastrophe theory has provided some impressive models, of which the model for anorexia nervosa is the most dramatic. IT IS MOST LIKELY THAT THE BEST RESULTS WILL CONTINUE TO COME FROM TEAM EFFORTS INVOLVING MATHEMATICIANS AND OTHER SCIENTISTS.

2 Remarks on the lexicon. The lexicon will perform its most useful service in facilitating the communication mentioned above. To serve this purpose the lexicon must be as well written as a good text on engineering science, including explicit statements of the hypotheses needed for a specific application. It was apparent to me that we still need to bridge the gap between our own discussions of administrative science and of catastrophe theory.

3 Feasible goals for the near future. I will briefly list some of the projects which I feel can be accomplished in the near future and which will be needed in most applications.

3.1 Continued development of a language for modeling using catastrophe theory

3.2 Developing useful descriptions of the dynamics near the catastrophe surface (The terms FAST and SLOW must be incorporated.)

3.3 Understanding the role of noise in catastrophe theoretic modeling

3.4 Developing a theory of stability with respect to a time series of perturbations of not-necessarily very small size

3.5 Clarifying the relationship between descriptive and mechanistic models (I feel that descriptive models are helpful, both in and of themselves, and as precursors for mechanistic models. However, the prospective user must be wary of placing too much reliance on purely descriptive models.)

3.6 Clarifying the role of intuition in catastrophe theoretic and other approaches to modeling.

4 **More difficult projects.** The following projects are more difficult but are still quite likely to be beneficial.

4.1 Incorporate catastrophe theoretic dynamics into automata or general systems (The biologist's understanding of the Hodgkin-Huxley oscillator may serve as an initial prototype.)

4.2 Develop a theory of systems identification for catastrophe theoretic models

4.3 Combine at least elementary aspects of control theory with catastrophe theory to provide one possible approach to non-linear control theory.

COMMENTARY

There already exists some intertesting work relating the Hodgkin-Huxley equations to catastrophe theory, due to Isabel Labouriau (at present unpublished). This uses the imperfect bifurcation theory of Golubitsky and Schaeffer (1979a, 1979b) (a variation of Thom's elementary catastrophe theory) and the subsequent work of Golubitsky and Langford (1981), to study periodic solutions to the space-clamped Hodgkin-Huxley equations. There is an organizing center of topological codimension 2 which predicts the qualitative form of the bifurcation diagrams obtained numerically by Rinzel and Miller (1980).

THE ROLE OF CATASTROPHE THEORY IN SOCIAL ORGANIZATION MANAGEMENT

John E. Stecklein

****

This presents the researcher's view of the problems. Many observed phenomena appear to display features remeniscent of the elementary catastrophes so there is no shortage of areas to which the theory might be applied. There is, however, a shortage of fully developed work.

Teamwork is going to be crucial to any such project, and it is not going to be easy to assemble suitably qualified teams.

The main immediate prospect is a general increase in understanding rather than a specific breakthrough on a practical problem.

Analogies between one field of application and another could be useful.

Specific applications might include sudden shifts in behavior and decision-making processes.

****

Torgerson (1969) has defined an organization as a system of continuously and consciously coordinated activities of two or more persons. An organization is a dynamic system, thus appropriate for analysis and attempts for greater understanding by the possible application of the mathematics of nonlinear dynamical systems. The central concern of this paper is ways in which a mathematical model of a special kind of dynamical system can be utilized by managers, researchers, or others involved in and concerned about the effective operation of organizations, by citing possible appplication to one kind of social organization--a college or university.

I agree with Poston and Stewart (1978a) when they say that one must understand in some detail the precise mathematical hypotheses involved in these mathematical models, and the way they lead to conclusions, in order to even begin to have a thorough grasp of when catastrophe theory may and may not be an appropriate choice for modeling. Although I thought I had, I do not believe that I have yet achieved that level of understanding. It is difficult, therefore, to speculate with any degree of certainty, how useful nonlinear modeling might be in dealing with organizations. Certainly some similarities and parallels to situations in which catastrophe theory has already been applied occur to me. As to how relevant they are and how easily they might be refuted, I have no idea. Nevertheless, it might be of some use to list such speculative situations or conditions of organizational functioning that might lend themselves to at least the more elementary models of catastrophe theory.

I do believe that there is potential here, but achieving this potential will take much time and effort and the combined imaginations, expertise, perspectives, and insights of teams of mathematicians and practitioners in the operation and study of organizations. Because I have been convinced of the desirability of such team-attacks to the problems, I thought that the idea to attempt to bring such teams together was excellent and I was eager to participate. My participation has served to clarify the difficulties involved in finding individuals who are confident enough in their fields of specialization to go afield, so to speak, in considering this type of cross-over, and perhaps, more importantly, to find individuals who, although highly specialized in their fields, are able to communicate in other than their own specialized jargon. Your approach to bring in several teams with different mixes of mathematicians, statisticians, and field specialists will give you more insight into the kinds of combinations that can communicate best and work most effectively.

Poston and Stewart (1978a) also make the point that catastrophe theory is not purely qualitative as its critics sometimes claim. They cite the progress made in applying catastrophe theory to the physical sciences and evidence of such progress in the biological sciences. In the former the opportunity to select "simple" systems, and "more recently, [of] those with <u>disorganized</u> complexity, which can be 'statistically simple'" (page *x*), provide opportunities to apply catastrophe-theoretic methods useful in furnishing quantitative information which may be confirmed by experiment. It may be that the social sciences will again follow the physical and biological sciences in the evolution of new tools useful in understanding the fields.

Until we reach the stage where the theoretic modeling can be combined with quantification, however, it seems that catastrophe theory will only be useful in trying to better understand our fields. I think this is important, and the trial and error process of fitting models to various aspects of organizational functioning may increase our understanding of the process. However, I do not see any short-range prospect of using catastrophe theory to establish a common qualitative language to enable better comparisons of situations, conditions, or problems among institutions. Perhaps such an event may come to pass if enough people can be induced to participate in the trial-and-error approaches that might be attacked by team combinations of experts. But as of now, the most I can anticipate is the opening of new perspectives about social organizations and an increased awareness of the similarities of organizational phenomena with those in other fields.

Having stated my misgivings, I will now speculate as to some possible applications of the simple cusp model of catastrophe theory in a social organization--specifically, the university or college. In the cusp model, we are concerned with a three-dimensional situation in which the joint occurrence of two variables may interact over some large period of time without provoking a major shift in a third variable. The classic case cited in the literature as an appropriate setting for catastrophe theory modeling is the situation of the combination of tension and unrest in a prison, the ebb and flow of interaction between tension and unrest until some critical point in either one or the other or both variables causes a sudden shift resulting in a prison riot. Such sudden shifts can take place in many aspects of university operation as well as in other social organizations. One possibility is the usually innate resistance to unionization that is encountered in a university. Since variables [factors] that might lead to a desire to unionize typically involve general working conditions on the one hand and remuneration [salary, rewards, or benefits] on the other, one might visualize a situation in which the combination of these two variables--one reaching an acceptable level and the other not, or both being less than acceptable, or both being reasonably unacceptable--may still provide insufficient impetus to cause the majority of the faculty to decide they wanted or needed a union. Some critical point is usually reached, however, involving either or both of the variables which causes a sudden change of mind and a faculty vote to unionize.

Another way of looking at a very similar and not unrelated situation would be to visualize a cusp model in which one variable is job satisfaction and the other variable is perceived value of the psychological trait of loyalty to the institution. It is my perception that individuals--and probably organizational units as a whole--often remain steadfastly loyal to the institution in which they work, despite negative forces such as deteriorating work conditions, noncompetitive salaries, lack of recognition or reward for excellent performance, and signs that others do not reciprocate or share one's valuation of loyalty, until suddenly, often with no noticeable overt action, loyalty to the institution suddenly disappears. Perhaps the descriptions of the variables suggested in these two examples should be changed. I can think of alternative suggestions, but I think the process of refining the definitions of the variables will be part of the beneficial process of using catastrophe theory models to understand better what transpires in sudden shifts of the kind represented by the two suggestions.

Johnson and Lacher (1981) have demonstrated how some of the features unified by the cusp catastrophe, such as local minimum, stationary bifurcation, and hystereses loop, might be applied to a decision involving student admission policies. Such models could be applied to many kinds of decisions in an organization, but particularly those that might be considered political decisions. At some point every decision-maker is answerable to someone--a group of faculty members, an administrator, a board, or in private enterprise, the stockholders. Many decisions, therefore, are made according to what they perceive the sentiment to be of those to whom they are answerable. Such shifts and changes in decisions can be modeled by one or more of the three models mentioned above. A fourth model--saddle point--would appear to be most useful in any decision in which there are two opposing points of view, both of which must be brought to bear simultaneously in making a decision. It is unlikely, assuming that the proponent of each point of view is equally firm and equally powerful, that either side will achieve the highest degree of satisfaction in whatever decision is made. Rather, it is more likely that some sort of compromise must be reached with both sides achieving less than their ultimate desires.

Perhaps attention should be given to the possible application of catastrophe theory models differentiated according to types of decisions or manner of decision-making that vary among several types of organizations. For example, decision-making in a military organization will be different than decision-making in a business or industrial organization, both of which would be different from the process in an educational institution, other social service, or nonprofit organization. Obviously some similarities in the decision-making process prevail, but the pressures of politics seem to vary, the obeisance to authority varies, and the extent of participatory democracy as a part of the decision-making process certainly varies. Thus common catastrophe theory models might be applicable but the variables representing the different dimensions might be quite differently defined. For example, the model might be most useful in clarifying the differential effects of variables that enter into a promotion decision for persons in the military, in business operation, or in an educational institution. Furthermore, such differentiation might provide insight into the ways in which various psychological variables (such as loyalty referred to above) are influenced differentially by conditions innate to or imposed by a particular organizational structure.

COMMENTARY

Note once more the emphasis on teamwork, expressed by Stewart and Hastings, and the need to develop experience in many more specific applications, expressed by Stewart.

PERSONAL OBSERVATIONS WITH RESPECT TO THE APPLICATION OF CATASTROPHE THEORY TO HIGHER EDUCATION MANAGEMENT

Ben Lawrence

****

Like Walker's article, these are the observations of someone who is actively involved in the "sharp end" of administration, hence deserving careful attention.

It focuses on the simplest catastrophe model, the unembellished cusp, and suggests that while this has its place in teaching inexperienced administrators and in theoretical research on administrative processes, it is less useful to the experienced administrator who will already possess an intuition for "catastrophic" effects.

If additional quantitative information could be found (e.g., the relative sizes of poles of opinion as well as the qualitative feature of bimodality) this might prove useful to the practicing administrator.

An interesting problem to consider is the effect of changes in the weighting of priorities.

****

The utility of the catastrophe model lies in its graphic conceptual representation of real and potential shifts in behavior. The model will be intuitively acceptable to most college and university administrators: It can be used to describe real and potential situations of which they are well aware. The utility to experienced administrators I believe to be limited. I will return to this assessment later.

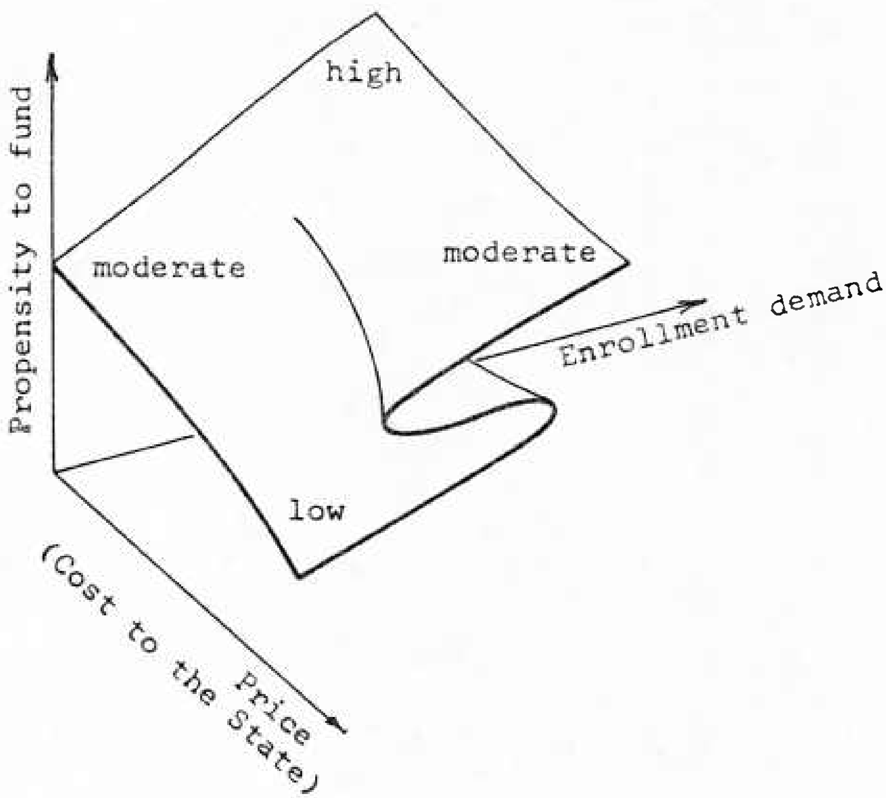
The model has high utility as a teaching device for aspiring or beginning administrators. While most experienced administrators, at least in higher education, will find the model intuitively acceptable, their intuition for the most part was gained from experience. Instructional exercises using this model could be very useful in conveying to aspiring or beginning administrators the kinds of problems they will generally face on a day-to-day basis and in providing a means of thinking about them.

Researchers may also find the model useful as a heuristic tool. I do not believe that needs explanation. Ray Zammuto of the National Center for Higher Education Management Systems (NCHEMS), for example, has used it in this way in his studies of the impact of decline on colleges and universities.

As I stated above, I believe the model has little utility for the experienced administrator. However, it would be interesting to pursue the matter.

Most college and university administrators, particularly presidents, are well aware that some changes (especially if they are sudden) in the behavior of an organization or its environment can have serious negative impacts on their institution. They do not eschew change, however, and indeed they often promote change. One of their primary responsibilities is to be on the lookout for change--be it planned or as a result of external forces that can cause harm to the effectiveness of the institution. Accordingly, they not only plan but also develop contingency strategies. While often not articulated, there is an intuitive understanding of catastrophe. Their problem is not in understanding but in predicting and devising interventions. The complexity of the interacting variables is the major deterrent and while catastrophe theory improves upon current methodologies by one dimension, it does not appear to be that significant.

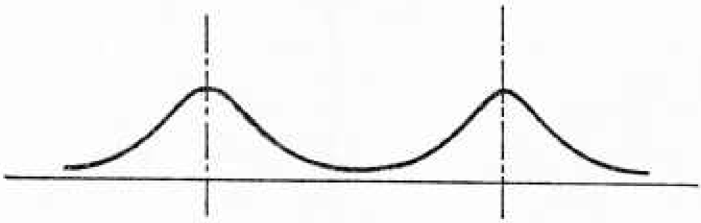
The following example will serve to illustrate. A major concern of a public college or university president or the C.E.O. of a state system is financing. To reduce this major concern to the catastrophe model, one must utilize variables that in themselves may be susceptible to analysis using the catastrophe model.



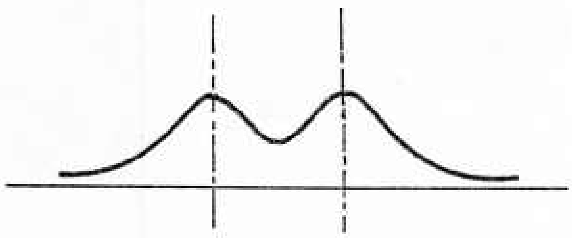
In this model, as enrollment demand increases, the propensity of state government to fund the institution will increase. As the price (cost to state) increases, the propensity to fund will decrease. As both enrollment demand and cost to the state get high, few administrators would deny that a state of uncertainty will arise because of the conflicting propensities on the one hand to satisfy enrollment demand and on the other to reduce cost to the state. Legislative bodies will polarize around those two general propensities. Administrators seek to ameliorate such polarization and generally keep the propensity to fund high.

The model, however, is of little practical use, even though I suspect with enough money and time that data could be gathered to use the model in a predictive sense as Zeeman has done in the case of the Gartree Prison (1977, p. 387-406).

The model does not tell me the distance between the two poles:

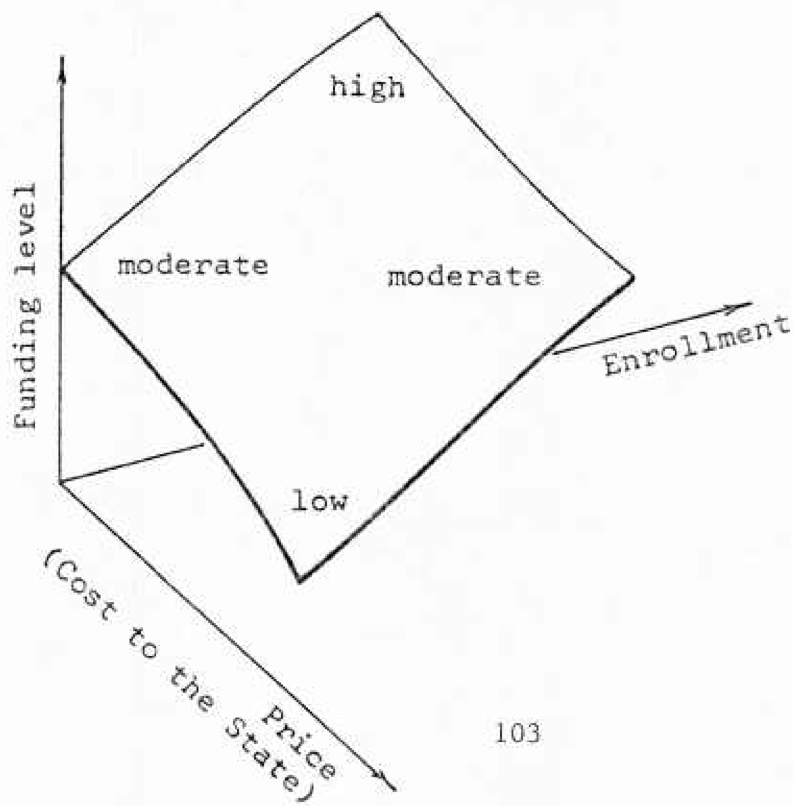


or



This is an extremely important factor for it helps determine intervention strategy. If the distance between the two poles is great, a strategy of compromise could have devastating short-term effects on the institution. We have seen this occur in the last two years at both the state and federal levels. The strategy then needed is to shift the weight of opinion to a clear majority on the higher propensity to fund. If the distance between the two poles is small, compromise is probably the best strategy for the long term. Not compromising will intensify the polarization and creates animosities towards the institution that will compound the problem in coming years.

This raises a point of interest. Imagine the following model superimposed on the catastrophe model.



Note that I have changed "Propensity to Fund" to "Funding Level" and "Enrollment Demand" to "Enrollment." I have also introduced a dotted line in place of the cusp. I have, in effect, reduced the model to three easily measured dimensions. The surface of the new model now describes actual funding levels. Intuitively, I know there is a relationship between the surfaces of the two models.

If, for example, the particular "funding event" was a bond issue (a binary issue) and the funding proposal lost because of the propensity not to fund, the dotted line would show a step down under the cusp and indeed that might be viewed as a catastrophe. While there is a relationship, the model itself does not help us predict the outcome. The weight of and the distance between the two modes is more predictive and the data collected in many instances provides insight that is helpful in developing intervention strategies to reduce the polarization.

As I understand it, the reason this happens is that the dimensions themselves are not linear. Enrollment demand itself is a complex function and one of the variables is price. So, too, price is a complex function and one of its variables is enrollment.

As I have reviewed the several journal articles describing potential applications of catastrophe theory to higher education management, I observe that all are subject to the same general limitations described above.

NCHEMS early attempts at simulation offered the theoretical possibility of prediction. And indeed, in a research setting (with considerable satisfaction) we were able to develop models that were essentially a series of interrelated equations with the output of one becoming one of the inputs to the next. The data requirements were, however, huge--foreclosing the feasibility of application. As we simplified the models to curtail data requirements the models became less satisfying.

Our current search for solutions to these kinds of problems is taking us in the opposite direction--more disaggregate analysis. The integration of the results of these disaggregate analyses for the time being will be made in the context of decision-making itself.

I found my visit helpful and informative. While I cannot point to any one thing that NCHEMS will do as a result, the discussion has enhanced my perspective, creating the potential for some future practical connection between theory and practice.

COMMENTARY

The recent paper by Peregoy and Zeeman (in preparation) sets up a model which corresponds quite closely to the idea of changes in weighting of priorities playing an important role in the production of sudden changes.

It should be emphasized here that "catastrophe theory" is not just a body of dead knowledge. If it is to be of any use to anyone, it must develop. It has not stood still by any means since the early work of Thom and Zeeman; Stewart's contribution in particular makes the point that at least twenty different variations on the basic theme exist. The aim of catastrophe theory should be the development of a coherent body of mathematical techniques for the analysis of sudden change. The idea that Thom's Classification Theorem lists all possible sudden changes is an unfortunate misinterpretation of expository articles and needs to be corrected. The importance of the theorem is that it lists several natural and basic ways that sudden changes might occur and that it lays down a paradigm for developing analogous theories in other contexts.

Another point worth re-emphasizing is that catastrophe theory consists of more than a "cusp model" and is at its best as a <u>unifying</u> structure. Thus, Lawrence's comment near the end that the current trend at NCHEMS is to "disaggregate" analyses of specialized topics, leaving the integration of knowledge to "the decision making process itself" (i.e., to a process not well understood), might be the very place where catastrophe theory could play a meaningful role.

A final comment relative to both Walker and Lawrence: It is difficult if not impossible to get administrators to spend much time speculating on mathematical modeling of simplified versions of what they themselves do for a living. They are more interested in doing what they do than in theorizing about it. (Presumably this last sentence is true of almost anyone.) Thus, if new types of models are to impact the effectiveness of administrators, the initial investigations and demonstrations of utility will have to come from some other source. Administrators themselves will be, and should be, the last to be convinced.

CATASTROPHE THEORY AND THE SOCIAL SCIENCES

Loren Cobb

****

Cobb looks at the impact of dynamical systems and statistics for stochastic processes on mathematical models in the social sciences. Early applications of catastrophe theory suggest four major statistical implications, including the importance of historical studies, the need for a statistical theory of nonlinear time series, the need for statistical detection of multiple modes, and the importance of stochastic differential equations. He warns of the need to place catastrophe theory within the larger domain of nonlinear dynamical systems.

****

IMPACT

The significance of catastrophe theory for the social sciences (here meaning sociology, economics, anthropology, history, political science, psychology, psychiatry, education, and administration) is twofold. First, by presenting several remarkable and appealing behavioral characteristics (bifurcation and hysteresis) not previously seen in social science mathematical models, catastrophe theory has encouraged many social scientists to learn about the broader field which contains catastrophe theory, namely nonlinear dynamical systems. This cannot fail to have a profoundly positive impact on the quality and subtlety of the mathematical models and metaphors used by social scientists. Second, catastrophe models and their relatives pose severe problems of empirical verification, the solution of which will encourage a greater sophistication in statistical methodology than is now found in the social sciences. In fact, new statistical concepts and methods are now being developed to cope with the unique problems posed by catastrophe models.

The above two effects of catastrophe theory are certainly not independent. Traditionally, social scientists have relied on statistical tests for the empirical verification of their research hypotheses. These tests are possible only when the random elements in the hypothesis have been clearly identified. Thus a statistical verification of a catastrophe model requires that it be expressed in a suitable stochastic form. But to do this one must first thoroughly understand stochastic nonlinear dynamical systems. The net effect is that social scientists who become seriously interested in catastrophe theory will learn a lot about two subjects--dynamical systems and statistics for stochastic processes. This effect is already clearly visible in the social science literature.

The actual substantive impact of catastrophe theory on the social sciences has been quite small. To date there have been no truly convincing models of social phenomena that contain catastrophes. However, this almost certainly reflects the difficulty of the subject and the impoverishment of social science methodology. J. Q. Smith (1983), a statistician, has demonstrated the existence of catastrophes in certain decision-theory situations and several authors have proven the existence of catastrophes in certain economic and voting models, but these applications are relatively unimportant and are primarily of theoretical interest. On the other hand, experimental psychologists may be on the verge of some very successful applications and there are several promising applications in political science. All in all, it is still much too early to tell. It is reasonable to suppose that there will be a gradual increase in the quality of social science applications of catastrophe theory as researchers become familiar with the models and their associated statistical methods and problems.

STATISTICAL CONSIDERATIONS

There are four major statistical implications of catastrophe theory:

1. The hysteresis effects observed in catastrophe models suggest that scientific experiments MUST take into explicit account the history of each experimental unit. Failure to do this will result in a grossly inflated within-cell variance if there is more than one stable equilibrium value in the given cell or experimental condition. This phenomenon is already well-known in certain branches of engineering which routinely deal with hysteresis effects but is virtually unknown in all other areas, particularly psychology and economics. An inflated within-cell variance estimate results in an inflated estimate of the total error variance in an experiment which will degrade or possibly eliminate the statistical significance of the test of the experimental hypotheses. Perhaps many social science studies have shown inconclusive results because of unsuspected essential nonlinearities in the dynamics of the processes under study. Certainly the usual methodology in the analysis of variance is completely inadequate for detecting such effects.

2. There are so many difficulties associated with reliably deducing the presence of a catastrophe from snap-shot (i.e. single point in time) studies that the use of time-series methods will surely increase. But this will necessitate the development of better statistical theory for <u>nonlinear</u> time series models with time-varying controls. This is bound to become an active research area in statistics, as it already is in engineering. Certainly the mainstay of time-series analysis, the Box-Jenkins method, is completely inadequate for this purpose.

3. When snap-shot data are all that are available, it is still possible to infer much about the underlying dynamics of a system by analyzing the shapes and higher-order moments of the observed frequency distributions. This approach has a long history in certain limited areas of the social sciences (e.g. the study of income distributions), but with the advent of statistical catastrophe theory it should become much more widespread. Detecting and explaining multimodality in frequency distributions with hypotheses that use

multi-stable systems instead of multiple subpopulations will become an active research area in many disciplines.

4. With very few exceptions, social scientists have avoided differential equations as a tool for theory construction, primarily because of the patently nondeterministic nature of social processes. Statistical catastrophe theory, however, begins from a formulation using stochastic differential equations which incorporate the relevant random effects and simultaneously allow the full power of differential equations to be brought to bear on the theory. Almost as a fringe benefit, these stochastic models carry with them the essential ingredients for statistical tests. Thus social scientists now have access to methods that are mathematically powerful, substantively interesting, and statistically verifiable.

CAVEATS

All of these implications tend to make catastrophe theory sound a good deal more important than it really is. To bring things into proper perspective, it should be recognized that many of the same points made previously could also be made about the recent use of modern algebra to describe a very wide range of social and cognitive structures. Catastrophe theory is a relatively small area within the larger domain of nonlinear dynamical systems and it is to this larger area that I attribute most of the benefits stated here. However, catastrophe theory has had more than its share of publicity and controversy which has in turn called our collective attention to the larger issues and problems.

The controversy surrounding catastrophe theory can, I believe, be distilled down to the question of the scientific validity of basing the justification for a given model on the mathematical property of genericity. (Genericity is a mathematical term that denotes a kind of typicality. Perfect symmetry in a mathematical model is usually non-generic so that when symmetry occurs in natural phenomena there is always an important and nontrivial reason for it.) For myself, I find the genericity argument satisfactory only for sorting out promising models from unpromising ones. If a nongeneric model is empirically better than a generic model, then it is important to find out why. Beyond this it is neither necessary nor productive to use genericity. This point of view essentially abandons one of the principle reasons for using catastrophe models. However, it is not at all clear how genericity extends to statistical models, and in this context an alternative and easier justification follows from the parametrized form of Laplace's theorem. Thus the major part of the controversy about catastrophe theory is irrelevant to sciences that will be using statistical forms for their models.

COMMENTARY

What is the motivation for this approach? It is a very natural development from a series of different mathematical points of view.

First, there is the dynamical systems (differential equations) idea of a gradient flow (things flow downhill on a landscape and accumulate in valleys) which is at the basis of a great deal of mathematical physics. Potential theory and variational principles are two examples.

Second, interest has been aroused recently in stochastic dynamics--a dynamical system perturbed by statistical noise. (Stuff flows downhill but random effects kick it about and disturb its equilibrium.) Mostly this topic arose in engineering control problems (e.g. the behavior of an aircraft) but they have since been widely applied elsewhere.

Third, there is pure mathematical input--the nature of critical points. Usually these are "nice" (i.e. Morse functions, with hilltops, saddles, valleys) but they can be degenerate and more complicated (see Markus, this work). Catastrophe theory arose in exactly this setting and its mathematical importance has never been challenged.

Most work in stochastic dynamics concerns the behavior of a gradient dynamical system near a nondegenerate critical point under white noise. Cobb takes the next, and entirely natural, step of studying the degenerate points. The mathematics pretty much demands that such points be unfolded (universally perturbed) into parametrized families of systems.

Cobb's class of partial differential equations arises as the stationary equations of such an unfolded system. Poston has shown how the elementary catastrophe potentials arise as the low-noise asymptotic limit of stochastic differential equations. They are thus one very natural way (to a mathematician) to study stochastic differential equation models whose stationary equations may bifurcate.

Incidentally, dynamical systems theory has been forced to pay extensive attention to bifurcations--not just what one system will do, but how that behavior can change. There is a need for a parallel stochastic theory.

The actual results that Cobb has obtained are only one tiny step in an enormous program. His main interest is in practical estimation of empirical distributions and the use of the simpler stochastic differential equations as models for real processes. But the point of view that his work opens up is of clear importance and has a long mathematical pedigree. It is not a concocted side-issue: It is part of the mainstream.

****

In response to a request for comments on the first draft of this report, we received several positive and even congratulatory replies. Richard Savage wrote the following letter which raises some substantive issues which could form the basis for an additional exchange of ideas. We reproduce that letter here, with sentences deleted which refer to items that were modified for this final version of the report, in order to stimulate that exchange.

****

Yale University
Department of Statistics
Box 2179 Yale Station
New Haven, Connecticut 06520-2179
June 1, 1983

Professor F. Craig Johnson
Department of Educational Research
College of Education
Florida State University
Tallahassee, Florida 32306

Dear Craig,

Thank you for your note of 13 May and the draft manuscript "Applications of Catastrophe Models to Organizational Effectiveness." I have read through the material and have a few general comments as well as specific comments.

As the report now stands, it is moderately interesting. It suffers from having many authors. It really would be well to have it edited and unified in some ways. In addition to having many authors, it suffers from the disease that everyone has with catastrophe theory. Everyone seems to have some self-interest that they need to protect. Each person likes to introduce their own vocabulary. (I will come back to a separate paragraph on the problems of "linear".) Many of the statements really appear quite wild. They are not tempered by facts or by reality. There is much backpatting. There are very few results relevant to the discussion at hand. I suspect that the main difficulty is that the authors are typically not experts in the areas that they are interested in writing about, that is, the social sciences.

One thing that seems essential to make this document readable is an essay on meaning of linear and nonlinear . . . . As far as the analysis of the stationary states goes, the basic question is not linear or nonlinear. Or at least that is my opinion. The critical thing is that in the catastrophe theory models the regression surface is not single valued. Putting the emphasis on multivariate functions instead of linear or nonlinear would catch the eye as to what is exactly going on.* Questions of scaling and nonlinearity are for the birds. Those were road blocks set up by the psychologists 40 years ago. They have in fact never been relevant to the development of good science. If you are going to get beyond the talk stage, you must become quantitative. Scales are developed so that linear operations become useful. That is the way science operates. Statisticians are quite competent to look at situations which are nonlinear in various ways. In particular, all of statistical regression theory is concerned with the possibility that the regression equations are not linear in the independent variables but linear in the parameters. (One must be careful to make sure that these terms do not get switched around. I noticed that at some point in the document the control variables which I would call the independent variables become the parameters.) In addition to that large group of situations, statisticians have been developing extensive techniques for handling situations where the parameters enter into the regression function in a nonlinear way. I really think that it is the multivalued aspect of the regression function that brings a new dimension to the problem from the viewpoint of current statistical activity.

Stewart's "Overview" will be practically useless for this audience. None of the terms seem to make any sense to me . . . . In general, his writing is too fragmentary here to make it of any use.+

. . . .

In Stewart's second article, second paragraph from the last on page 21, is that idle speculation or does he really have something concrete in mind. It is not sufficiently spelled out to even be suggestive. Is it worth including? Does he have examples from other fields? Page 26, the first of several irrelevant examples. (Can we have he data to look at it?) Near the [top of the next] page, if the data had been processed by good statistical package, it would have looked at the residuals and, likely, would have found what we are now discussing. The parenthetical remark . . . is just that kind of expression which seems entirely out of place. With bimodal residuals we may not have gone to the idea of a multi-valued regression surface but we certainly would have been looking for trouble. In fact, we might find a better explanation than the bimodal regression surface. In the last paragraph on page 34 in the second line, the word "parameter" is used as I would use "independent variable." I strongly urge that a formal discussion be made of the words "linear" and "nonlinear" in the context that are being used here, that is, statistics - social science and mathematical physics.

In Holt's article on page 37 the second complete paragraph reads (to put it mildly) bizarre. He has technical terms confused, he does

---

* We agree completely. Some of the commentary was rewritten in order to make this point less obscure.

+ Stewart was asked to give a brief overview, in outline form, to bring the reader up to date on issues developed in previous publications. Perhaps it is too brief.

not present a realistic or interesting form of science, he brings in fallacious probabilistic arguments, and I am doubtful that he understands the difference between Bayesian and non-Bayesian statistics. It is a terrible paragraph. The 50% [at the top] of page 38 is ludicrous. Ask yourself where it could come from in a serious argument. Page 38, third complete paragraph, is ridiculous. (Paul Meehl is a great man and I do not want to contradict him. I hope the current author just has the story screwed up.) How many examples do you know of natural science predicting something exactly. Notice the inverse square law is not even trusted by natural scientists; they have for years been checking to see if that could be right. Even so, in the inverse square law there's an undetermined parameter that needed to be estimated, and that is the gravitation constant. I have no idea of what the author is talking about. When I see the handbook of chemistry and physics, I wonder what world he is in. Page 41, first complete paragraph, hypothetical history has no fun for me . . . . [At bottom of] page 41, how in the world did international relations get in here? . . . . Last paragraph I find madness. Without a good warehouse of relevant examples, this seems outrageous. Is this what is really going on in modern social science? *

. . . .

In Cobb's article, page 106, the last complete paragraph reads a little bit strangely compared to some of the other things that we will see. In particular, there are many people working in catastrophe theory who do not bother to look at the differential equation but only look at the steady states. Certainly much of social science has got along for years and done some interesting work with neither looking at the dynamics nor looking at the possibility of multi-valued response functions. Thus his claim is much too strong. I do agree with him that the natural approach to catastrophe theory for the social scientist is through the dynamics. In fact, I say that several times in my little piece. But he has to explain that that is just one route. Near the top of page 107 is Cobb trying to say that some of these results are in normative (in contrast to behavioral) situations. And what we are primarily interested in is finding examples of catastrophes in behavioral models and not in normative models. Page 107, point 1, is complaining, incorrectly, that the analysis of variance would not be capable of detecting a phenomenon which has not yet been shown to exist. The analysis of residuals would be indicative of just this type of phenomenon. Routine analyses of residuals are made in most good practices of statistical packages. It might be interesting to show point 2 to Professor Box. It would have to be done in context so he could understand what was being claimed. Page 108, last paragraph, I never did understand the meaning of "genericity." And the "parameterized form of the Laplace theorem" does not ring a bell with me either. To whom is the author writing?

---

* On the other hand, E. C. Zeeman feels that those papers which reflect research using catastrophe theory, such as the works of Cobb and Stewart, are the strongest in the report. Zeeman was less impressed with papers by people who had not done any research.

The Markus article appeared rather grand. I am not too sure it was relevant . . . .

In summary, I think there are two main difficulties with the report. There are often wild and unsubstantiated or meaningless statements. Further, the amount of relevant material is quite small. The first point could be taken care of by a rather careful internal refereeing. You might send some of my comments to the authors and see if they would like to modify or support statements which I found outrageous. That sort of thing could be done on a systematic basis. Actually, let some of the involved authors act as referees for the other papers. On the question of relevance, I think the problem is more difficult. Perhaps the person who threw in thyroid gland would be willing to take it out. I would rather replace it with nothing than to leave something that is so distant from what we are interested in. But maybe if the pressure were applied to him he could come up with a better example. Maybe a person like Cobb could at least speculate on what a nonlinear time series model would look like for some special issue in the social sciences. He has not done this in the past. It would be good to apply pressure to him to get out what he thinks would be interesting models. The one data set of his that I am familiar with is his analysis of the birthrates as a bimodal distribution. To me it is an extremely bad example. One would guess that in this situation the data came from more than one population. Put a little presure on some of these people and see if they can say something that is relevant.

Best of luck. Please do not squeeze me.

Sincerely,

Richard I. Savage

REFERENCES

REFERENCES

Angelis, V. A new approach in modelling the growth and decline of industrial cities. Preprint, 1980. (Available from 11 Paidriadon Street, Athens 808, Greece)

Apollonius. On conics. _Great Books of the Western World_ (Vol. 11). Chicago: Encyclopedia Britannica, 1952.

Archimedes. On spirals. _Great Books of the Western World_ (Vol. 11). Chicago: Encyclopedia Britannica, 1952.

Aristotle. Physics. _Great Books of the Western World_ (Vol. 8). Chicago: Encyclopedia Britannica, 1952.

Arnold, V. I. Classification of bimodal critical points of functions. _Functional Analysis & Applications_, 1975, _9_, 43-44.

Arnold, V. I. Classification of unimodal critical points of functions. _Functional Analysis & Applications_, 1973, _7_, 230-231.

Arnold, V. I. Critical points of functions on manifolds with boundary, simple Lie groups B_k, C_k, F_k and evolute singularities. _Uspehi Math. Nauk_, 1978, _33_ (5), 91-105.

Arnold, V. I. Critical points of smooth functions. _Proceedings of the International Congress of Mathematicians._ Vancouver, 1974, 19-39.

Arnold, V. I. Critical points of smooth functions and their normal forms. _Russian Math. Surveys_, 1975, _30_, 1-75.

Arnold, V. I. Integrals of rapidly oscillating functions and singularities of projections of Lagrangian manifolds. _Functional Analysis & Applications_, 1972, _6_, 222-224.

Arnold, V. I. Lectures on bifurcations and versal families. _Russian Math. Surveys_, 1972, _27_, 54-123.

Arnold, V. I. Local normal forms of functions. _Invent. Math._, 1976, _35_, 87-109.

Arnold, V. I. Normal forms for functions in the neighbourhood of degenerate critical points. _Russian Math. Surveys_, 1973, _29_, 10-50.

Arnold, V. I. Normal forms for functions near degenerate critical points, the Weyl groups of A_k, D_k, and E_k, and Lagrangian singularities. _Functional Analysis & Applications_, 1972, _6_, 254-272.

Arnold, V. I. On matrices depending on parameters. _Uspehi Mat. Nauk_, 1971, 26, 101-114; Russian Math. Surveys, 1971, 26, 29-43.

Arnold, V. I. Singularities of smooth maps. Russian Math. Surveys, 1968, 23, 1-43.

Arnold, V. I., Varchenko, A. N., & Gusein-Zade, S. M. Singularities of differentiable mappings, I and II. Moscow: Nauk, 1981. (Book of 700 pages, Birkhauser translation in press.)

Balasko, Y. Economic equilibrium and catastrophe theory: An introduction. Econometrica, 1978, 46, 557-569.

Berry, M. V., & Upstill, C. Catastrophe optics: Morphologies of caustics and their diffraction patterns. In E. Wolf (Ed.), Progress in Optics. North Holland, 1980, 258-345.

Black, F. & Scholes, M. The pricing of options and corporate liabilities. Journal of Political Economy, 1973, 81,637-659.

Brocker, T., & Lander, L. Differentiable germs and catastrophes. London Mathematical Society Lecture Notes (Vol. 17). Cambridge: Cambridge University Press, 1975.

Chipman, J., Hurwicz, L., Richter, M., & Sonnenschien, H. (Eds.). Preferences, utility, and demand. New York: Harcourt Brace Jovanovich, 1971.

Cobb, L. Cusp surface estimation. Computer program. Charleston: Medical University of South Carolina, 1980. (a)

Cobb, L. Parameter estimation for ensembles of non-linear stochastic systems. Preprint. Charleston: Medical University of South Carolina, 1980. (b)

Cobb, L. Parameter estimation for the cusp catastrophe model. Behavioural Science, 1980. (c)

Cobb, L. Stochastic catastrophic models and multi-modal distributions. Behavioural Sciences, 1978, 23, 360-374.

Cobb, L., & Watson, W. B. Statistical catastrophe theory: An overview. International Journal of Mathematical Modelling, 1980.

Coleman, S. Measurement and analysis of political systems. New York, 1975.

Colgan, P. W., Nowell, W., & Stokes, N. W. Spacial aspects of nest defense by pumpkinseed sunfish (Lepomis Gibbosus): Stimulus features and an application of applied catastrophe theory. Animal Behavior, 1980.

Copernicus. On the revolution of the heavenly spheres. Great Books of the Western World (Vol 16). Encyclopedia Britannica, 1952.

Cossu, C. Financial attrition, fixity of variable costs, ... and catastrophe theory. Paper presented at the meeting of the European Association for Institutional Research, London, England, November, 1980.

Einstein, A. Cosmological considerations on the general theory of relativity. The Principles of Relativity. Dover, 1923. (Originally published, 1917.)

Euclid. Elements. Great Books of the Western World. (Vol 11). Encyclopedia Britannica, 1952.

Euler, L. Introductic (Rev. ed.). Lyons, 1797.

Galileo, G. The two new sciences. Great Books of the Western World (Vol. 28). Encyclopedia Britannica, 1952.

Gibson, C. G. Singular points of smooth mappings. Research Notes in Mathematics. 25, London, England: Pitman Publishing, 1979.

Gilmore, R. Catastrophe theory for scientists and engineers. New York: Wiley-Interscience, 1981.

Golubitsky, M., & Langford, W. Classification and unfolding of degenerate Hopf bifurcations. Journal of Differential Equations, 1981, 41, 375-415.

Golubitsky, M., & Schaeffer, D. A theory for imperfect bifurcation via singularity theory. Communication in Pure and Applied Mathematics, 1979, 32, 21-98. (a)

Golubitsky, M., & Schaeffer, D. Imperfect bifurcation in the presence of symmetry. Communications in Mathematical Physics, 1979, 67, 205-232. (b)

Guttinger, W., & Eikemeier, H. (Eds.). Structural stability in physics. Synergetics, 1979, 4, Berlin: Springer-Verlag.

Hanson, N. R. Patterns of discovery. Cambridge: Cambridge University Press, 1958, 72-73.

Harrison, J.M., & Pliska, S. Martingales and stochastic integrals in the theory of continuous trading. Northwestern University Discussion Paper, 1981.

Hillix, W. A., Hershman, R. L., & Wickes, F. D. Catastrophe theory in the behavioral sciences (Report No. TR 79-8). San Diego, Calif.: U. S. Navy Personnel Research and Development Center, February 1979.

Holt, R. Catastrophe theory and the origins of World Wars I and II. In R. Aris, (Ed.), Catastrophes and other important matters. Minneapolis, MN.: University of Minnesota Department of Chemical Engineering, 1977, 13-39.

Holt, R., Job, B., & Markus, L. Catastrophe theory and the study of war. Journal of Conflict Resolution, 1978, 22, 171-207.

Hurwicz, L., & Richter, M. K. An integrability condition with applications to utility theory and thermodynamics. Journal of Mathematical Economics, 1979, 6, 7-14.

Johnson, F. C. Some implications of qualitative mathematical models for institutional research: Basic assumptions and implementation. Paper presented at the meeting of the European Association for Institutional Research, Louvain-la-Neuve, Belgium, November, 1981.

Johnson, F. C., & Lacher, R. C. Qualitative analysis of decisions. Paper presented at the meeting of the Florida Statewide Conference on Institutional Research, Fort Lauderdale, Florida, June 1981.

Kepler, J. Epitome of Copernican astronomy. Great Books of the Western World (Vol 16). Chicago: Encyclopedia Britannica, 1952.

Labouriau, I. Degenerate Hopf bifurcation and nerve impulse. Coventry, England: Warwick University, in press.

Lacher, R. C. Qualitative mathematical phenomena in a model of decision-making. Paper presented at the meeting of the European Association for Institutional Research, Louvain-la-Neuve, Belgium, November 1981.

Lagrange, J. [Oeuvres (Vol. 1).] Paris (Originally published, 1762).

Lorenz, K. [On aggression.] London: Methuen, 1967. (Originally published, 1963.)

Lu, Y-C. Singularity theory and an introduction to catastrophe theory. Berlin: Springer, 1976.

Malgrange, B. Ideals of differentiable functions. Oxford and New York: Oxford University Press, 1966.

Markus, L. Catastrophe theory: Mathematical theory or philosophical catastrophe? In R. Aris (Ed.), Catastrophes and other important matters. Minneapolis, MN.: University of Minnesota Department of Chemical Engineering, 1977, 1-10.

Markus, L. Extension and interpolation of catastrophes. Annals of the New York Academy of Sciences, 1979, 316, 134-149.

Markus, L. Lectures in differentiable dynamics (CBMS Regional Conference Series in Mathematics 3). Providence, R. I.: American Mathematical Society, 1971.

Martinet, J. Singularities of smooth functions and maps. LMS Lecture Notes, 1982, 58 C.U.P.

Mather, J. How to stratify mappings and jet spaces. Singularities d'Applications Differentiables, Plans-sur-Bex 1975 (Eds.: O. Burlet and F. Ronga) Lecture Notes in Mathematics 535, Springer, Berlin 1976, 128-176.

Mather, J. Stability of C-mappings

I The division theorem, Annals of Mathematics, 1968, 87, 89-104.

II Infinitesimal stability, Annals of Mathematics, 1969, 89, 254-291.

III Finitely determined map germs, Pub. Math. IHES, 1968, 35, 127-156.

IV Classification of stable germs by R-algebras, Pub. Math. IHES, 1969, 37, 223-248.

V Transversality, Adv. Math., 1970, 4, 301-336.

VI The nice dimensions, Proceedings of the Liverpool Singularities Symposium I. Springer Lecture Notes in Mathematics (Ed.: C. T. C. Wall), Berlin, 1971, 192, 207-253.

Meehl, P. Theory testing in psychology and physics: A methodological paradox. Philosophy of Science, June 1967, 103-115.

Morse, M. Calculus of variations in the large. American Mathematical Society Colloquium Publications, 1934, 18.

Newton, I. The system of the world. Great Books of the Western World (Vol. 34). Chicago: Encyclopedia Britannica, 1952.

Peregoy, P., & Zeeman, E. C. (paper in preparation)

Poincare, H. [Sur les proprietes des fonctions definies par les equations aux differentielles partielles.] Oeuvres (Vol. 1). Paris (Originally published, 1879).

Pontryagin, L., & Andronov, A. [Systemes grossiers.] Dok. Acad. Nauk, 1937, 14, 247-251.

Poston, T., & Stewart, I. N. Catastrophe theory and its applications. London, England: Pitman Publishing Limited, 1978. (a)

Poston, T., & Stewart, I. N. Non-linear modeling of multistable perception. Behavioural Science, 1978, 23, 319-334. (b)

Poston, T., & Stewart, I. N. Taylor expansions and catastrophes. London, England: Pitman Publishing Limited, 1976.

Ptolemy. The almagest. Great Books of the Western World (Vol. 16). Chicago: Encyclopedia Britannica, 1952.

Rinzel, J., & Miller, R. N. Numerical calculation of stable and unstable periodic solutions to the Hodgkin-Huxley equations. Math. Biosci. 1980, 49, 27-59.

Rosen, S. Hedonic prices and implicit markets. Journal of Political Economy, 1974, 34-55.

Saunders, P. T. An introduction to catastrophy theory. Cambridge: Cambridge University Press, 1980.

Seif, F. J. Cusp bifurcation in pituitary thyrotropin secretion. In W. Guttinger & H. Eikemeier (Eds.), Structural Stability in Physics. Synergetics, 1979, 4, Berlin: Springer-Verlag, 275-289.

Simon, C. Ellet's Transportation Model of an economy with differential products and consumers. University of Michigan C.R.E.S.T. Discussion Paper, 1979.

Smale, S. Differentiable dynamical systems. Bulletin of the American Mathematical Society, 1967, 73, 747-817.

Smale, S. The prisoner's dilemma game and dynamical systems associated to non-cooperative games. Econometrica, 1980, 48 (7), 1617-1634.

Smith, J. Q. Catastrophes and geometrical sufficiency. Preprint. London, England: Department of Stattistical Science, University College, 1981.

Smith, J. Q. Catastrophes in statistical models and decision theory: A way of seeing (Research Report No. 26). London, England: Department of Statistical Science, University College, June 1983.

Smith, J. Q. Mixture catastrophes and Bayes decision theory. Mathematical Proceedings of the Cambridge Philosophical Society, 1979, 86, 91-101.

Smith, J. Q., Harrison, P. J., & Zeeman, E. C. The analysis of some discontinuous decision processes. European Journal of Operations Research, 1981. 7, 30-43.

Sondermann, D. Revealed preference: An elementary treatment. Econometrica, 1982, 50, 777-779.

Sonnenschien, H. Price dynamics and the disappearance of short-run profits. Journal of Mathematical Economics, 1981, 201-204.

Sonnenschein, H. Price dynamics based on the adjustment of firms. American Economic Review, 1983.

Stewart, I. N. Applications of catastrophe theory to the physical sciences. Physica 2D, 1981, 245-305.

Stewart, I. N. Catastrophe theory in physics. Rep. Prog. Phys., 1982, 45, 185-221.

Stewart, I. N., & Peregoy, P. Catastrophe modelling in psychology. Psychological Bulletin, in press.

Sussmann, H. J., & Zahler, R. S. Catastrophe theory: Mathematics misused. The Sciences, 1977, 17, 20-23.

Thom, R. Un lemme sur les application differentiables. Bol. Soc. Mat. Mexicana, 1956, 2, 59-71.

Thom, R. [Structural stability and morphogenesis] (D. H. Fowler, trans.). New York: Benjamin-Addison Wesley Publishing Companies, 1975. (Originally published, 1972.)

Thom, R. The two-fold way of catastrophe theory. In P. J. Hilton (Ed.), Structural Stability, the Theory of Catastrophes and Applications in the Sciences, Lecture Notes in Mathematics, 1976, 525, Berlin: Springer, 235-252.

Thom, R. Answer to Christopher Zeeman's reply. In E. C. Zeeman, Catastrophe theory: Selected papers 1972-1977. Reading, MA.: Addison-Wesley Publishing Company, 1977, 633-638.

Thomas, L. The life cycle of innovative plans. The lives of a cell.

Thompson, J. M. T. Instabilities and catastrophes in science and engineering. London: John Wiley, 1981.

Torgerson, P. E. A concept of organization. New York: Van Nostrand Company, 1969.

Trotman, D. J. A., & Zeeman, E. C. The classification of elementary catastrophes of codimension $\leq$ 5. In P. J. Hilton (Ed.), Structural Stability, the Theory of Catastrophes, and Applications in the Sciences, Lecture Notes in Mathematics, 1976, 525, Berlin: Springer-Verlag, 263-327.

Varian, H. R. Catastrophic theory and the business cycle. Journal of Economic Inquiry, January 1979.

Walker, D. E. Personal communication, August 13, 1983.

Wetherhilt, B. W., & Zeeman, E. C. 1981 Bibliography on catastrophe theory. Coventry, England: Mathematics Institute, University of Warwick, 1981.

Woodcock, A. E. R., & Davies, M. Catastrophe theory. New York: E. P. Dutton, 1978.

Zeeman, E. C. Catastrophe models in administration. Address presented at the Forum of the Association for Institutional Research, Atlanta, Georgia, April 1980.

Zeeman, E. C. Catastrophe theory. Scientific American, 1976, 234, 65-83.

Zeeman, E. C. Catastrophe theory: Selected papers 1972-1977. Reading, MA.: Addison-Wesley Publishing Company, 1977.

APPENDIX

APPENDIX: MATHEMATICAL LEXICON

It is beyond the scope of this lexicon to review the breadth or depth of interaction between dynamical systems theory and social sciences. However, it should be noted that a bibliography (Wetherhilt and Zeeman, 1981) prepared recently at the University of Warwick in Coventry, England, contains 58 pages of references and is still incomplete. Fields of science (other than dynamical systems) well represented in this bibliography include: statistics, physics, chemistry, geology, biology, medicine, fluid dynamics, meteorology, economics, management, psychology, sociology, education, political science, archaeology, and several others (depending on how finely one subdivides disciplines). Scientists and mathematicians alike are realizing that many perplexing observations are better understood and described through the qualitative universal language of dynamical systems. Some basic terms of the language are presented here. The end of the decade will see a richer language with wider usage in all sciences.

Dynamical systems. A dynamical system consists of a number of states (1-dimensional variables) $x^{(1)}, \ldots, x^{(n)}$ and rules governing changes of state. The rules may be thought of as differential equations of the form

$$dx^{(j)}/dt = f_j(x^{(1)}, \ldots, x^{(n)})$$

where f_j $(j = 1, \ldots, n)$ are smooth functions of n variables and t represents time. The manifold parametrized (locally) by the various states of the system is called the **phase space**. Thus, each point of the phase space represents a state of the system. Alternative to a system of differential equations, the rules for a dynamical system may be taken to be a specification of speed and direction for each point x in phase space, with these specifications being smooth functions of x. For a given x, the rules determine a path which x must follow through phase space. Extending these paths indefinitely (forward and backward in time) produces a collection of paths in phase space satisfying three principle constraints: (1) wherever speed is nonzero the path is smooth, (2) every point in phase space lies on a path, and (3) no two paths intersect. A specification of paths in phase space satisfying (1), (2), and (3) locally and globally is equivalent to a system of differential equations like the one mentioned at the beginning of this paragraph.

A path of the system together with the direction of travel along the path is called a **trajectory** of the system. A **phase portrait** of a dynamical system consists of its phase space together with an indication of its trajectories. In practice, one cannot sketch all trajectories without coloring every point in phase space (resulting in nothing more than a phase space of a different color), but one must sketch enough trajectories to indicate the dynamics of the system. Figures 1 through 8 are phase portraits of various dynamical systems.

A **parametrized dynamical system** is really an entire collection of dynamical systems parametrized by one or more real parameters $c^{(1)}, \ldots, c^{(m)}$ often referred to as "external" or "control" parameters. The parametrization is smooth (the functions f_j are smooth functions of the parameters) and often slow (thinking of the system as evolving, the rate of evolution is slow compared to the rates of travel on the paths of the system). A parametrized dynamical system retains the same phase space as it evolves, but the phase portrait may evolve with the system.

A point p in the phase space of a dynamical system is called **non-wandering** if neighbors of p do not move away forever: Given any neighborhood N of p, there is some point q of N so that q returns to N at some later time. Alternatively, p is **wandering** if p and its neighbors move away permanently: There is some neighborhood N of p, each point of which moves out of N and never returns. The **non-wandering set** of a dynamical system consists of its non-wandering points. This set is usually denoted by Ω. The non-wandering set is **invariant** (that is, points of Ω go to points of Ω under the action of the dynamical system) and contains all "eventual" behavior of the system. Assuming that the speed of travel of trajectories is fast relative to our powers of observation, the non-wandering set contains all "observable" states of the system. Understanding and describing the non-wandering set is the most important (and often the most tractable) aspect of forming a phase portrait.

In a parametrized dynamical system, the non-wandering set evolves with the system, and understanding the evolution of Ω along with associated changes in stability is crucial.

The various phase portraits or non-wandering sets of a parametrized dynamical system may be graphed simultaneously in coordinates $c^{(1)}, \ldots, c^{(m)}, x^{(1)}, \ldots, x^{(n)}$, obtaining a phase portrait or non-wandering set for the entire parametrized dynamical system. The examples that follow serve to illustrate the ideas just introduced as well as a number of other concepts.

Discrete time systems. A discrete time system is a dynamical system in which time is measured in discrete units. The rules governing change of state may be thought of as difference equations

$$\Delta x^{(j)} = f_j\,(x^{(1)}, \ldots, x^{(n)})$$

where f_j are smooth functions or simply as a recursion: Points move about phase space in discrete jumps according to certain dynamical rules. Most of the concepts of (continuous-time) dynamical systems have analogs in the theory of discrete time systems.

It is difficult for the principal investigators to imagine a deterministic system in any setting (physical, biological, or social) that is not accurately modeled by some parametrized dynamical system (or a discrete-time system). Therefore, unless ruled out on scientific grounds, dynamical systems phenomena must be considered at least potentially useful in understanding real phenomena. In fact, evidence seems to indicate at present that facsimiles of all systems phenomena do occur in the physical and biological sciences and there is little reason to believe the social sciences are any simpler. It would follow that the language of dynamical systems is essential to communication and understanding in the social sciences.

Qualitative behavior type. In cases where a definitive description requires mathematical technicalities inappropriate for this writing, the discussion is limited to the major attributes of the phenomenon and gives a **canonical model**, meaning that (1) the canonical model exhibits the behavior described, and (2) by definition, the **behavior type** is anything qualitatively equivalent to the canonical model. Deeper understanding of these qualitative types requires, among other things, analyses of the canonical models.

Two dynamical systems phenomena are **qualitatively equivalent** if a smooth coordinate change ("psuedolinear" coordinate change, in the language of Stewart) transforms the phase portraits of one to the other. Similarly, two parametrized dynamical systems phenomena are **qualitatively equivalent** if there is a smoothly parametrized family of smooth coordinate changes transforming the parametrized phase portraits of one to the other.

An **equilibrium** is a state in which assigned speed is zero--a steady state, a stationary state, a state which constitutes an entire trajectory. When speed assigned to the state $x_o = (x_o^{(1)}, \ldots, x_o^{(n)})$ is zero, the dynamical system cannot change into or out of the state x_o. These are the theoretical steady states of the system, though they may not be observable in a practical sense. A **stable equilibrium** is an equilibrium x_o with the property: There is a neighborhood N of x_o in phase space such that every trajectory passing through a point of N heads toward x_o as t becomes infinite. Stable equilibria are the observable equilibria of the system. If the system is in state x_o, where x_o is a stable equilibrium, and if some perturbation jolts the system into a slightly different state in N, then the dynamical rules will drive the system back toward x_o. If the system has fast "response time," the net effect from an observer's viewpoint is that the system remains in state x_o even in the presence of small perturbations. "Reponse time" is the time required for a point of N to get so close to x_o that our powers of observation cannot distinguish it from x_o.

Bifurcation is used in two senses. As an external parameter c passes through a certain value c_o, a qualitative change in the non-wandering set occurs, usually accompanied by a shift in stability. The term **bifurcation** is used when such change is continuous.

Bifurcation also is used to mean any of a number of particular qualitative types of bifurcation, such as stationary bifurcation and Hopf bifurcation. A qualitative change in Ω usually means a qualitative change in the observable states of the system as an external parameter slowly changes. A qualitative change may occur continuously or discontinuously, distinguishing "bifurcation" from "catastrophe" (discussed later).

A **stationary bifurcation** is a type of bifurcation in which one equilibrium divides into three equilibria as some external parameter c passes through a certain value c_o. In the most useful setting, the original equilibrium becomes unstable while two stable equilibria are created.

Canonical model: $dx/dt = cx-x^3$

A **cycle** is a trajectory that returns to itself; i.e., a closed loop of states through which the system flows indefinitely, in periodic fashion, always returning to a given state in some finite period of time. If a trajectory ever returns to the same state, this trajectory must be a cycle (otherwise constraint (3) would be violated for the dynamical system localized near the point of self-intersection). The return time is the same for any point of the cycle; the time of first return is called the **period** of the cycle. A **stable cycle** is a cycle Γ with the property: There is a neighborhood N of Γ in phase space such that every trajectory passing through a point of N heads toward Γ as t approaches infinity.

Comments similar to the ones for stable equilibria apply here. Stable cycles are the "observable" cycles of the system. Stable cycles are sometimes called **limit cycles**.

A **Hopf bifurcation** is a type of bifurcation in which a stable equilibrium becomes unstable while a stable cycle is created as an external parameter passes through a certain value.

Canonical model: $d\theta/dt = 1$

$$dr/dt = cr - r^3$$

[(θ,r) are polar coordinates in the x_1x_2-plane]. The first equation insists on rotation at a constant rate; the second equation forces r (the radius of rotation) to behave as a stationary bifurcation. A line of observable steady states disappears during the creation of a line of stable cycles of increasing radius.

Catastrophe is used in two senses. As an external parameter c passes through a certain value c_o, a discontinuity in the non-wandering set occurs. **Catastrophe** is also used to mean any of a number of particular qualitative types of catastrophe, such as fold catastrophe and cusp catastrophe. A discontinuous change in the non-wandering set usually means a discontinuous change in the observable states of the system.

A **fold catastrophe** is a type of catastrophe in which a parametrized stable equilibrium and a parametrized unstable equilibrium coalesce, annihilating both equilibria, as an external parameter passes through c_o. In particular, a parametrized stable equilibrium simply disappears at $c = c_o$. If the system is in the state corresponding to this stable equilibrium, then beyond $c = c_o$ the system must seek another state. Assuming fast response time, the system might appear to jump from one state to another.

Canonical model: $dx/dt = c - x^2$.

A **double fold hysteresis loop** occurs as an external parameter varies and catastrophe occurs in two places, $c = c_1$ and $c = c_2$, with c_1 less than c_2. As c changes from $c < c_1$ to $c > c_2$, one catastrophe occurs at $c = c_2$. As c changes from $c > c_2$ to $c < c_1$, one catastrophe occurs at $c = c_1$. For $c_1 < c < c_2$, there are two stable equilibria; for $c < c_1$ or $c > c_2$ there is one stable equilibrium.

Canonical model: $dx/dt = c + x - x^3$.

The fold catastrophe can usually constitute only part of the dynamics of a system that models real phenomena, because there is no other line of stable equilibria (or other attractors) to "catch" the system after the coalescing equilibria disappear. The double fold hysteresis loop is the simplest system containing catastrophic change from one "observable" state to another. Notice that for $c_1 < c < c_2$, there are two observable states for each value of c. Knowledge of the external parameter is not sufficient to predict the state of the system. However, if one knows the history of c, it may be possible to predict the state.

A **cusp catastrophe** occurs when a system has two external parameters, one (say, $c^{(1)}$) exhibiting a double fold hysteresis loop, and the other (say, $c^{(2)}$) exhibiting stationary bifurcation. The hysteresis loop grows in size (from zero) as $c^{(2)}$ moves into its post-bifurcation zone.

Canonical model: $dx/dt = c^{(1)} + c^{(2)}x - x^3$.

Notice that when $c^{(1)} = 0$, the canonical model for the cusp reduces to the canonical model for stationary bifurcation and, when $c^{(2)} = 1$, it reduces to the canonical model for double fold hysteresis loop. The cusp is a localized phenomenon (like stationary bifurcation) and a structurally stable phenomenon (like the double fold hysteresis loop). These concepts are discussed in more detail later.

An **attractor** is a non-void set A in the phase space of a dynamical system satisfying the following: **invariance**, i.e., A consists solely of entire trajectories, so that points do not move into or out of A under the dynamical rules; **closure**, i.e., A contains all of its limit points; **stability**, i.e., there is some neighborhood N of A in phase space such that every trajectory through a point of N heads toward A as t approaches infinity; and **indecomposability**, i.e., some trajectory heads toward "all" of A as t approaches infinity. Notice that an attractor is a subset of the non-wandering set. Stable equilibria and stable cycles are attractors as are some more bizarre objects discovered in recent years. (See "strange" attractors.) A fact about attractors that is enlightening is A must contain a dense trajectory, i.e., some point of A passes "as close as you care to choose" to every other point of A as t approaches infinity. The **basin of attraction** is the largest neighborhood N that satisfies the property in the definition of attractors. A "strange" attractor is an attractor that is neither a stable equilibrium nor a stable cycle. One of the earliest examples was discovered by E. N. Lorenz during an investigation of a model (obtained by truncating actual equations of motion) in fluid dynamics. The behavior observed by Lorenz was that the system oscillates in one mode for several cycles, then switches to another mode for several cycles, then switches back to the first mode, etc. The number of oscillations before a switch of modes appears to be a random sequence of integers. (See Lorenz, 1963.) The Lorenz attractor seems to be a "fractal" object with non-integral Hausdorff dimension. Other "strange" attractors have been discovered, sometimes in a purely mathematical setting, other times in connection with attempts to model real phenomena. Most of these fail to have even the most basic topological amenities such as local connectedness. "Strange" attractors are not well understood at the present time. Their study and classification is one of the major problems facing dynamical systems theorists in coming years.

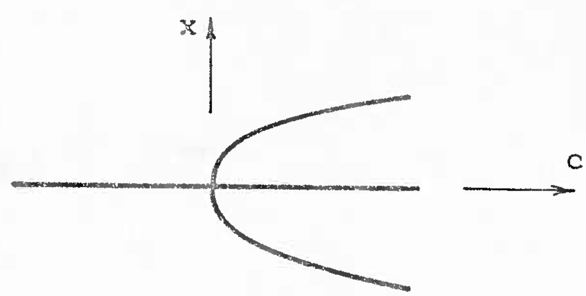
A phenomenon associated with "strange" attractors is **sensitive dependence on initial conditions**. Throughout A there are points whose trajectories are cycles of arbitrarily long periods as well as points whose trjectories wander through A forever, coming arbitrarily near every other point of A. In the presence of even undetectably small perturbations, the behavior displayed by such a system may appear "random" even though it is theoretically deterministic. This phenomenon is sometimes called **chaos**.

An Ω explosion occurs when one or more external parameters pass through certain values and a stable equilibrium or cycle transfers stability to an entire non-wandering set, creating a "strange" attractor. Such phenomena are not completely understood or classified, but their existence is established. From an observer's viewpoint, the system suddenly changes from well behaved to chaotic. Some of the latest research into fluid turbulence centers on this concept.

A **local phenomenon** occurs when a qualitative type of parametrized dynamical system possesses an **organizing center**, i.e., a point about which the qualitative type of the system remains constant no matter how near to the point we restrict our observations. In terms of canonical models, this means that there exist arbitrarily small neighborhoods N of p in (parametrized) phase space such that (1) N is invariant, and (2) the restriction of the system to N is qualitatively equivalent to the entire system. A stationary or Hopf bifurcation is local as is a cusp. A double fold hysteresis loop is not local. Local phenomona are independent of scale, so the concept is especially important in investigations where scale is artificially introduced.

Structural stability is not to be confused with stability in the attractor sense. This is a property of the (parametrized) system itself. Roughly, a system is structurally stable if the qualitative properties of the system do not change when the system is perturbed to a slightly different system. A slight change in the parametrized functions f_j does not alter the qualitative appearance of the parametrized phase portrait. A stationary bifurcation is not structurally stable; both the cusp and the Hopf bifurcation represent stable types to which a stationary bifurcation can be perturbed. Structural stability is important in two ways. First, if a model is not structurally stable then even undetectably small errors in setting up the model may change its qualitative behavior. Therefore, in the "non-rigid" sciences (such as social and biological) structural stability is a natural and desirable property of models. Second, structurally stable systems are easier to describe and classify mathematically. The cusp and Hopf bifurcation, being both local and structurally stable, are extremely important local phenomena in any science.

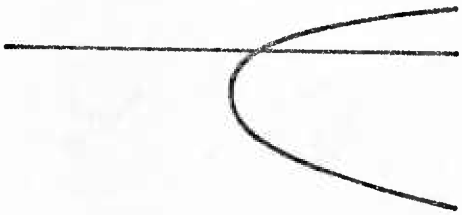
There are more restricted notions of structural stability which may also be useful in social science. Consider a stationary bifurcation. The canonical model $dx/dt = cx-x^3$ has symmetry about the c-axis in phase space. This symmetry can be seen in its bifurcation diagram (often called a "pitchfork").



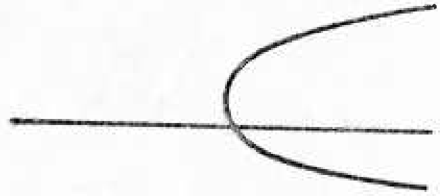
Reflection about the c-axis leaves the pitchfork, along with all the dynamics of the model, invariant. (This is Z/2-symmetry--the 2-element group acts on the cx-plane by sending (c,x) to (c,-x).) If one asks for the simplest structurally stable model containing the stationary bifurcation, one first finds all structurally stable diagrams to which the pitchfork can be perturbed and then seeks a model that combines all these forms. (This is a heuristic description of the Golubitski-Schaefer version of universal unfolding.) The pitchfork perturbs to structurally stable diagrams

and

but also to two asymmetric pitchforks



and

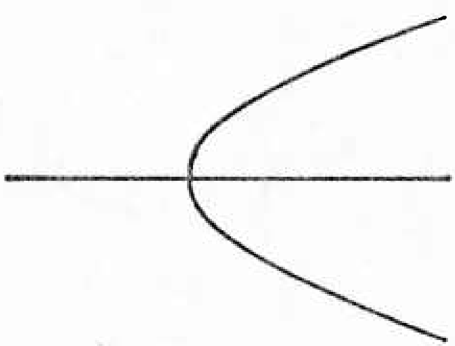


which perturb to the structurally stable diagrams

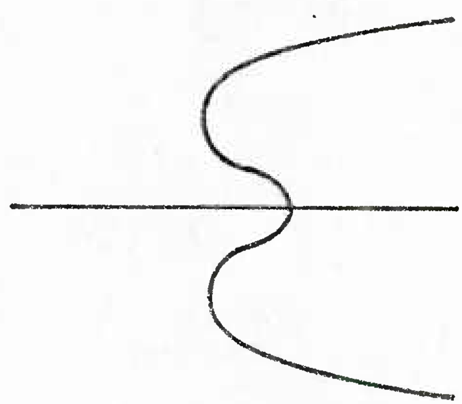
and

respectively. The canonical cusp contains all four of these structurally stable forms; it is the "universal unfolding" of the canonical stationary bifurcation.

However, if the Z/2-symmetry inherent in the canonical stationary bifurcation is an important feature we wish to maintain in our model, then we will consider only "Z/2-symmetric perturbations." The pitchfork then perturbs to two structurally stable (with Z/2-symmetry) diagrams:



and



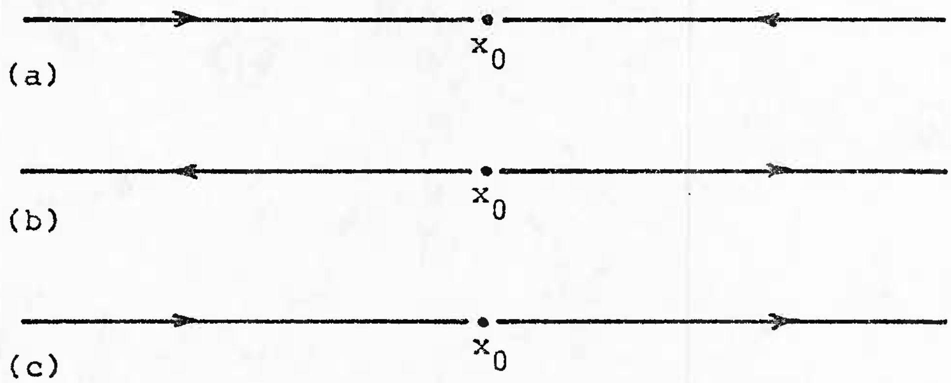


Figure 1. **Equilibria.** These illustrate typical stable (a) and unstable (b) and (c) equilibria in systems with one internal parameter x. Each consists of three trajectories, one of which is the equilibrium point x_o. The non-wandering set consists solely of x_o.

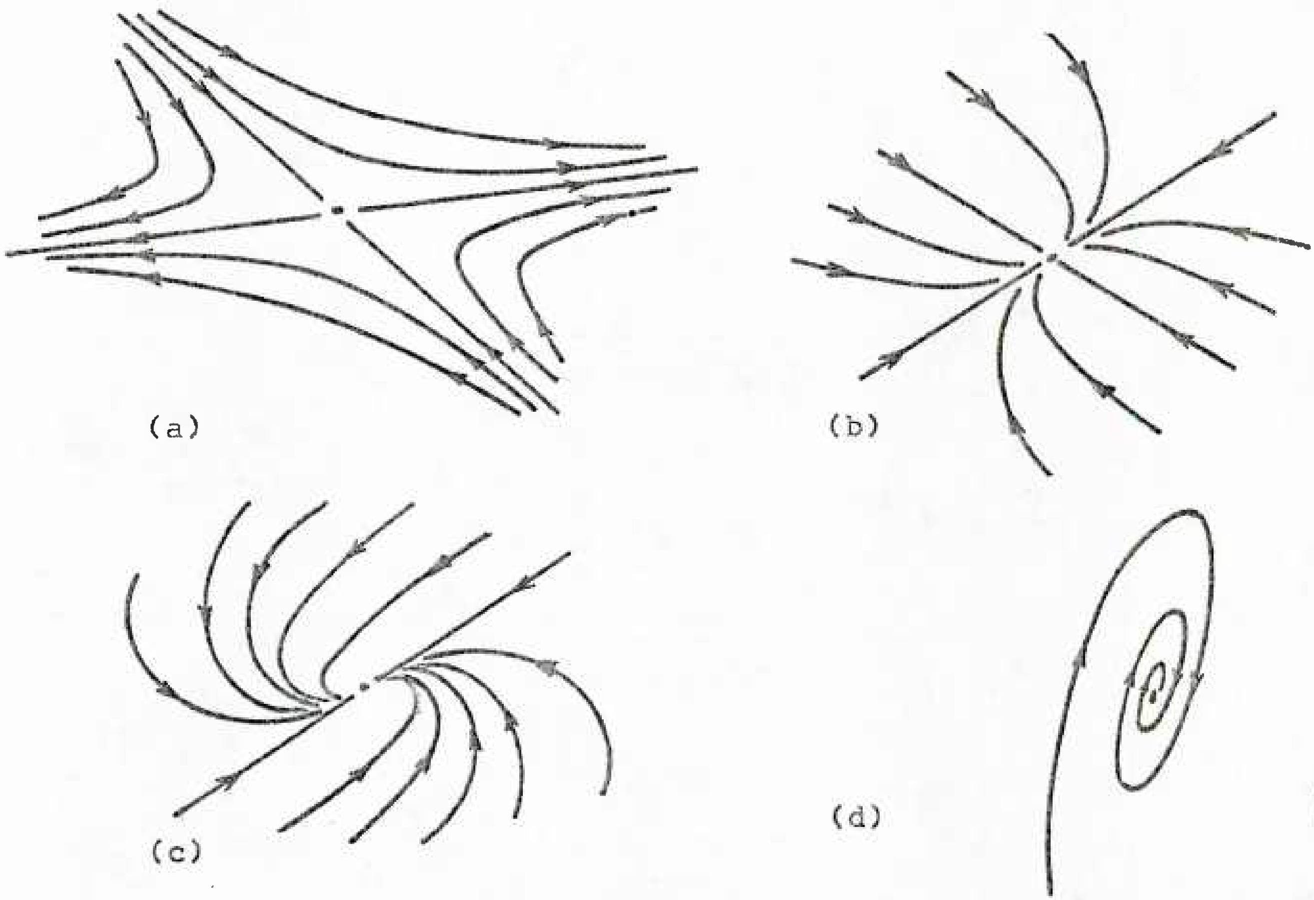


Figure 2. Equilibria. When there are two or more internal parameters, the flow near an equilibrium can exhibit diverse forms of behavior. Each illustration assumes a 2-dimensional phase space (the plane in which the trajectories are sketched)--(a) is called a saddle point and is unstable; (b), (c), and (d) are stable; reversing the arrows in (b), (c), and (d) results in typical unstable equilibria. The non-wandering set in each illustration consists of the equilibrium point alone.

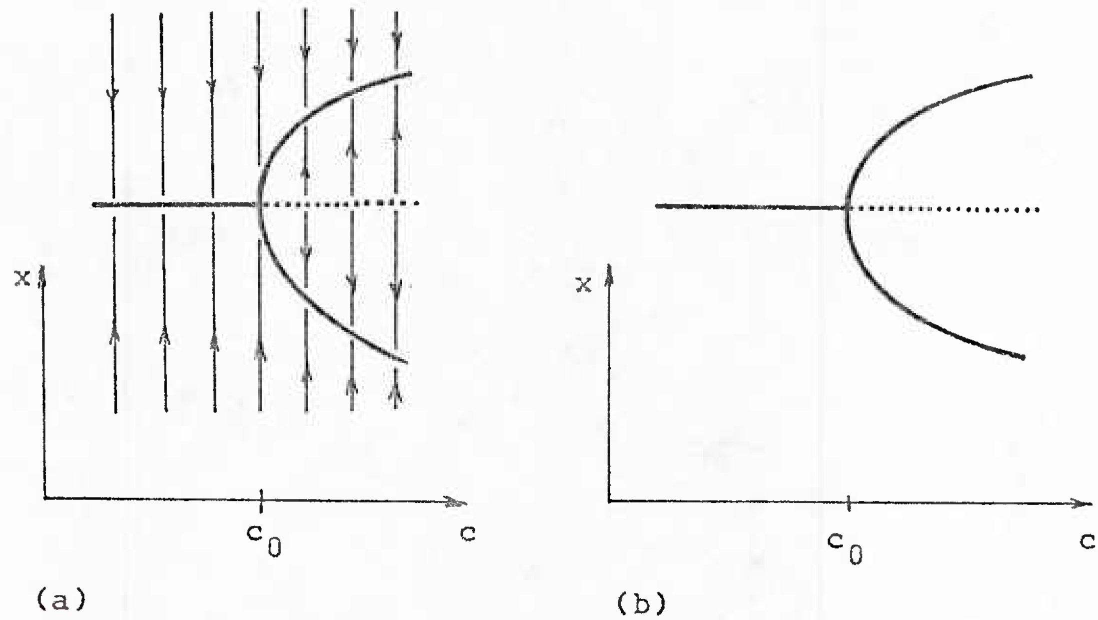


Figure 3. **Stationary bifurcation.** These illustrate the parametrized phase portrait (a) and parametrized non-wandering set (b) of a stationary bifurcation with one internal parameter x. For $c < c_o$ the system moves (in the x-direction only) toward the stable equilibrium above c. (These pictures are each vertical versions of Figure 1 (a).) For $c > c_o$, the system moves toward either of the outer equilibria but away from the inner one. Notice that the same information is conveyed in the less cluttered (b) where only the parametrized non-wandering set is displayed. The dotted line in (b) indicates unstable equilibria. A picture indicating the parametrized non-wandering set in parametrized phase space is often called a bifurcation diagram.

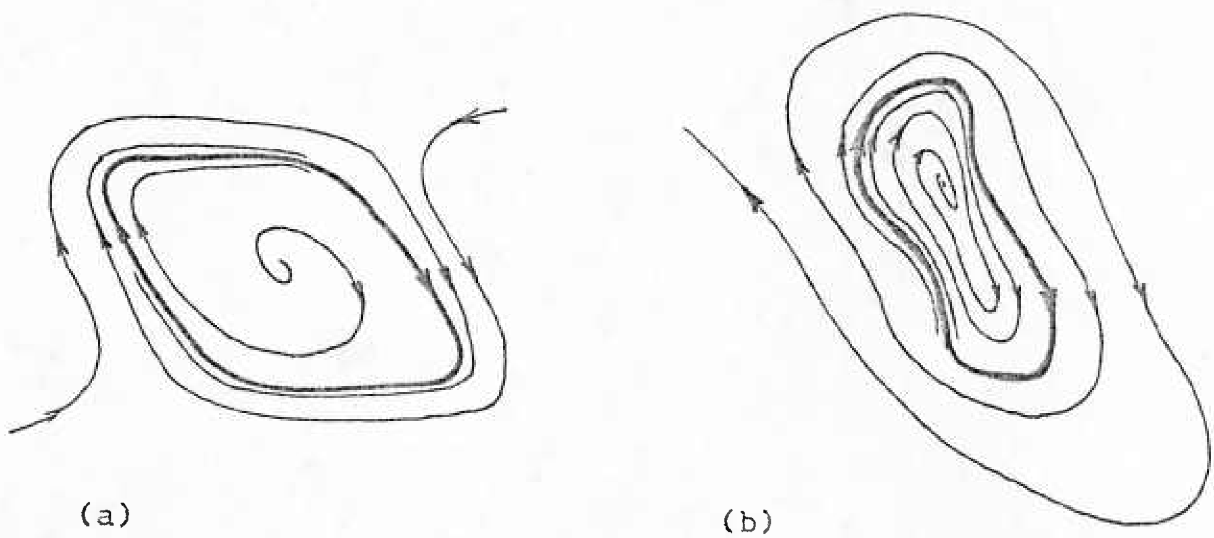


Figure 4. **Cycles.** At least two internal parameters are required for a (continuous time) dynamical system to exhibit periodic or oscillatory behavior. (One of these parameters may be time or, as often happens in mechanical systems, momentum.) Phase portraits containing cycles are illustrated with (a) stable and (b) unstable. In each case the non-wandering set consists of the cycle together with the equilibrium (which is unstable in (a) and stable in (b).) In terms of "observable" behavior, we would "see" oscillatory behavior in (a) and static behavior in (b) since omnipresent perturbation would not allow the system to remain in an unstable state.

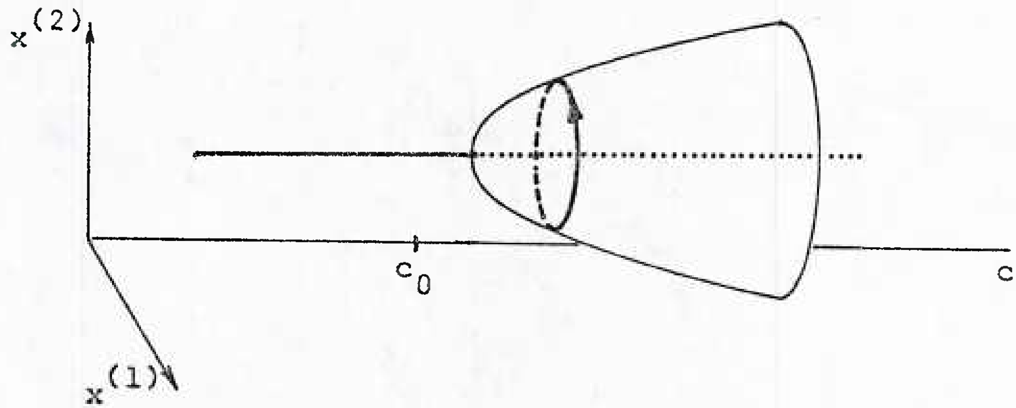


Figure 5. **Hopf bifurcation.** The parametrized non-wandering set of a Hopf bifurcation is depicted. The solid line above $c < c_o$ indicates a stable equilibrium for each $c < c_o$. As c passes through c_o, the stability of this equilibrium is "transferred" to a cycle whose size grows from zero at c_o. Note that at least three parameters are required for this model--two internal and one external.

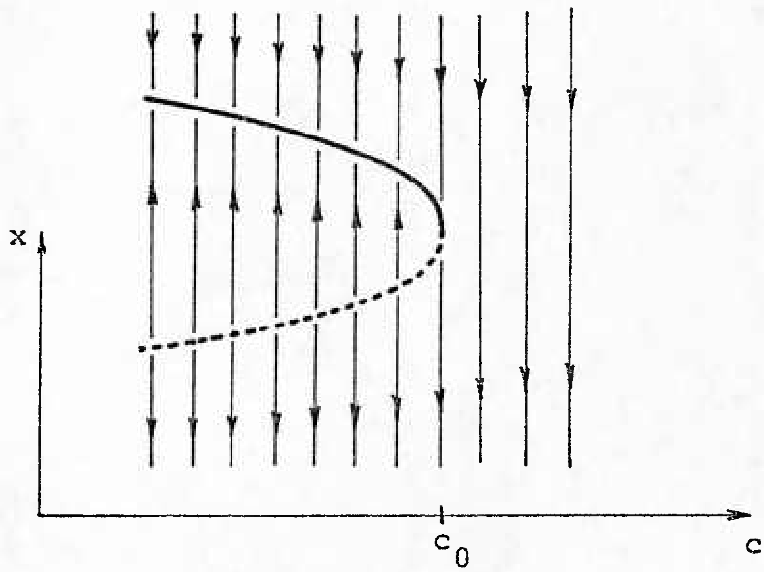


Figure 6. Parametrized phase portrait of a fold catastrophe. As c passes through c_o (from left to right in this picture) the parametrized stable equilibrium (represented by the solid line) disappears, leaving the system with no "observable" states.

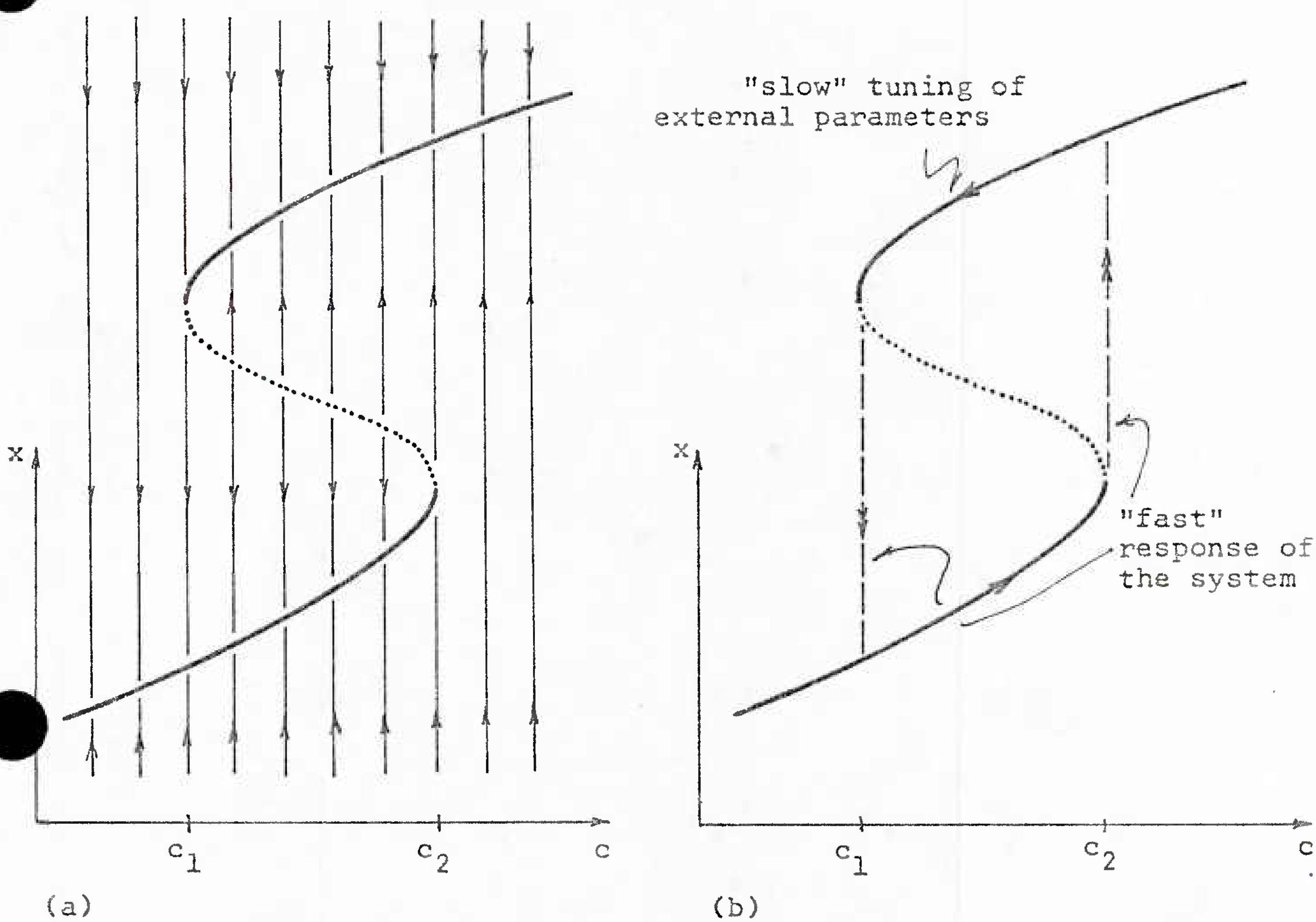


Figure 7. Parametrized phase portrait (a) and parametrized non-wandering set (b) of a double fold hysteresis loop. With $c < c_1$, there is only one steady state. As c moves slowly through $c = c_2$, the line of stable equilibria (representing observed behavior) disappears in a fold catastrophe, and the system responds by (rapidly) seeking another state. The only one available is on the upper line of stable equilibria. As c slowly returns to original setting, a similar "jump" occurs at $c = c_1$. The hysteresis loop itself is indicated in (b).

U211035

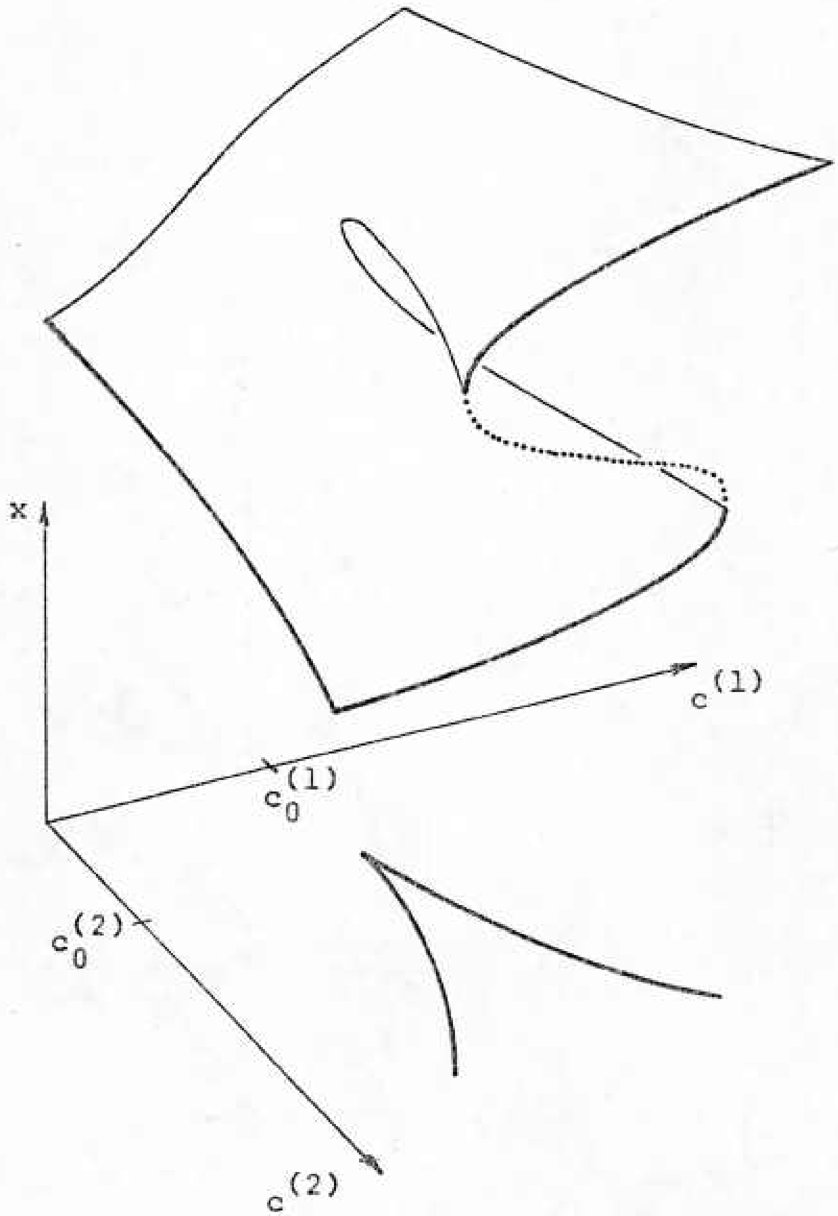


Figure 8. Cusp catastrophe. The parametrized non-wandering set of a cusp catastrophe is shown. Sections $c^{(2)}$ = constant $c_o^{(2)}$ are double fold hysteresis loops. The section $c^{(1)} = c_o^{(1)}$ is a stationary bifurcation. Note that three parameters are required for this model--one internal and two external.