F. Zeeman

INTRODUCTION TO CATASTROPHE THEORY

E.C. Zeeman

Gresham Professor of Geometry
Principal, Hertford College, Oxford

Reprinted from:
E.C. Zeeman,
Catastrophe Theory, Chapter 1
Addison-Wesley, 1977.

Reproduced in
Shortened version in
Scientific American
1 April 1976.

1 CATASTROPHE THEORY

INTRODUCTION

Catastrophe theory is a new mathematical method for describing the evolution of forms in nature. It was created by René Thom who wrote a revolutionary book "Structural stability and morphogenesis" in 1972 expanding the philosophy behind the ideas. It is particularly applicable where gradually changing forces produce sudden effects. We often call such effects catastrophes, because our intuition about the underlying continuity of the forces makes the very discontinuity of the effects so unexpected, and this has given rise to the name. The theory depends upon some new and deep theorems in the geometry of many dimensions, which classify the way that discontinuities can occur in terms of a few archetypal forms; Thom calls these forms the elementary catastrophes. The remarkable thing about the results is that, although the proofs are sophisticated, the elementary catastrophes themselves are both surprising and relatively easy to understand, and can be profitably used by scientists who are not expert mathematicians.

In physics many classical examples can now be seen to be special cases of low dimensional catastrophes, and as a result the higher dimensional catastrophes are beginning to suggest new experiments and offer understanding of more complicated phenomena. However in the long run the more spectacular applications may well be in biology, providing models for the developing embryo, for evolution and behaviour. Much of Thom's book concerns embryology. Models in psychology and sociology suggest new insight into the complexity of human emotions and human relationships, and offer new designs for experiments.

This article is in three parts. In the first part we introduce the reader gently into the ideas by describing some simple applications to

Published (in a shortened form) in the Scientific American, April 1976, Volume 234, part 4, pages 65-83.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

E. C. Zeeman, Catastrophe Theory: Selected Papers, 1972-1977. Published by Addison-Wesley Publishing Company Inc. Advanced Book Program, Reading, Massachusetts, 01867.

elasticity, aggression, emotions, war and economics; the objective is to lead up to a precise mathematical statement of one of the classification theorems, together with Thom's list of the 7 elementary catastrophes having a control space of dimension less than or equal to 4. In the second part we go more deeply into one particular application, namely a model for the nervous disorder anorexia nervosa, in order to illustrate the profundity and qualitative power of the new language in the human sciences. In the third part we select some familiar examples from classical physics, in order to illustrate the generality and quantitative power of the new language in the physical sciences.

CONTENTS.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

PART ONE

1. THE CUSP-CATASTROPHE

The first elementary catastrophe is the fold-catastrophe but that is too simple to see what is going on, and so we shall start with the next one, the cusp-catastrophe, which contains many subtleties. The cusp-catastrophe is the 3-dimensional graph illustrated in Figure 4. But let us approach it by first considering the simpler 3-dimensional graph shown in Figure 1, which illustrates how profit x depends upon income α and costs β, by means of the simple formula, $x = \alpha - \beta$.

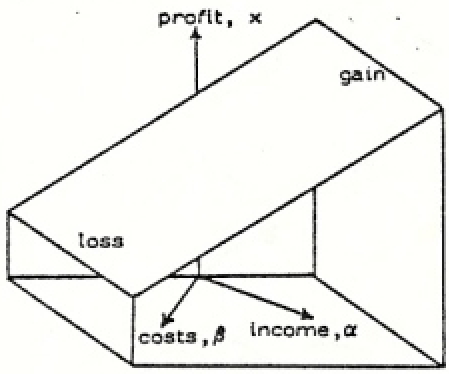


Figure 1. The graph of profit as a function of income and costs.

Here we represent α, β by axes in the horizontal plane C, and x by the vertical axis, and since the formula is linear the graph is a sloping plane in 3-dimensions. In particular :

an increase in income causes an increase in profit;

an increase in costs causes a decrease in profit;

an increase in both causes no change in profit.

We might summarise this situation by saying "income and costs are conflicting factors influencing profit".

Example 1. Aggression.

Now let us pick a similar sentence from psychology. In Konrad Lorenz's book "On Aggression" he says that rage and fear are conflicting

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

factors influencing aggression. The question is : can we represent this similar sentence by a similar graph? To make the question more specific first think of a dog, and then later we shall apply it to fish and humans. A preliminary problem arises : can we measure the rage and fear drives in a dog at any moment? Lorenz suggests that we can, and Figure 2 shows illustrations from his book; he proposes that rage can be measured by how much the mouth is open, and fear by how much the ears lay back. So let

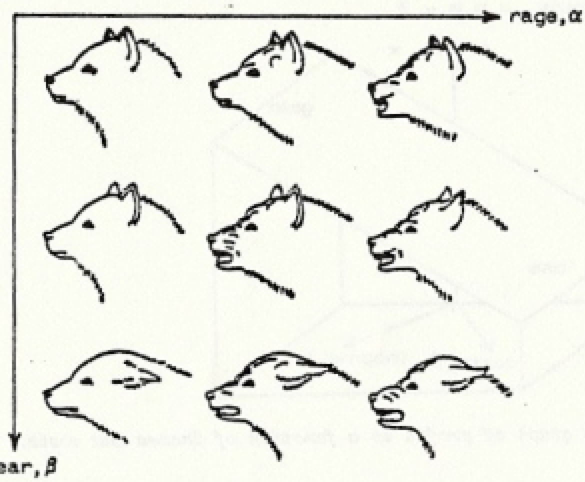


Figure 2. Rage and fear can be measured by facial expressions (after Konrad Lorenz).

us assume that rage and fear can be plotted as two horizontal axes, α and β. Meanwhile let us also assume we can devise some vertical scale x representing the resulting behaviour of the dog running from fight to flight, through intermediary behaviour such as growling, neutral and avoiding. We want to plot the graph x as a function of α and β. It is true, as before, that

an increase in rage causes an increase in aggression;

an increase in fear causes a decrease in aggression.

But what if we increase both rage and fear together? The least likely

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

behaviour is for the dog to remain neutral, and the most likely behaviour is fight or flight, although which of the two he will choose may be unpredictable. Therefore one thing is sure : there is no simple formula like $x = \alpha - \beta$, and the graph cannot look like Figure 1.

How do we analyse the situation? One answer is to look at the likelihoods. So let us imagine a likelihood distribution for the behaviour x in each of the following four cases.

	Drives	Most likely behaviour
1	Rage only	Fight
2	Fear only	Flight
3	Neither	Neutral
4	Both rage and fear	Fight or flight

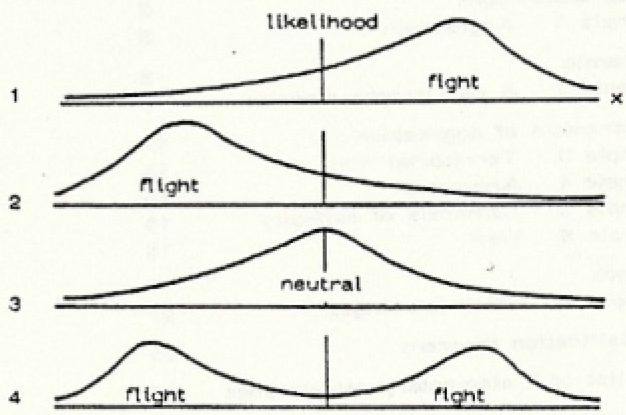


Figure 3. Likelihood of aggressive behaviour.

The interesting case is Case 4, where the distribution has gone bimodal. It should be possible to design experiments, in which the three variables are monitored by three observers for several minutes, and simultaneously fed into a computer, which could be programmed to draw the curves. From the curves the computer could then extract the 3-dimensional graph of the behaviour as a function of rage and fear. Above each point (α, β) of the horizontal plane C, representing given coordinates of rage and fear, is marked the point (or points) representing the most likely behaviour.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

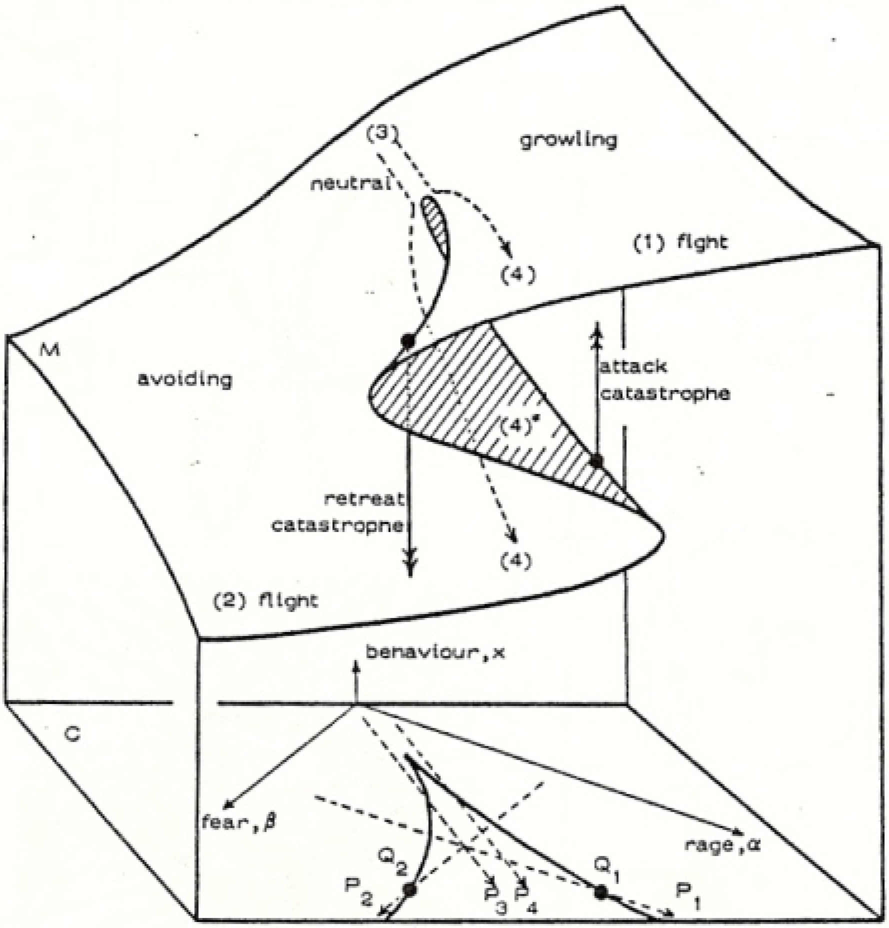


Figure 4. The cusp-catastrophe illustrating fear and rage as conflicting factors influencing aggression.

What catastrophe theory tells us is that if the likelihood distributions look like Figure 3, then the graph will look like the cusp-catastrophe surface M pictured in Figure 4. This follows from the main classification theorem, which we state later. In our experiment the graph could then be used for quantitative prediction of the dog's behaviour in the subsequent few minutes, predicting the sudden changes of mood in terms of the gradually changing facial expressions.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Already the graph has some surprising qualities, so let us analyse them. If there is rage only, then on the surface M above there is a single point marked (1), representing a fighting frame of mind, because the distribution of Case (1) of Figure 3 is unimodal. Similarly with cases (2) and (3). However in the interesting case (4) we obtain two points, marked (4) on the graph above; because the distribution has gone bimodal. Moreover there is another point, marked (4)*, in between these two, indicating the least-likely neutral behaviour. One of the reasons for including least-likely points on the graph, as well as the most-likely, is that it makes the graph M into a complete smooth surface; this is one of the consequences of the theorem. (Another reason for including these points is that it is sometimes useful to mark the threshold between the two modes of behaviour in the bimodal case). But it is important to remember when using the cusp-catastrophe that generally the middle sheet (shown shaded) represents least-likely behaviour, and only the upper and lower sheets represent most-likely behaviour.

The curve on the surface where the upper and lower sheets fold over into the middle sheet is called the fold-curve, and the projection of this down into the horizontal plane C is called the bifurcation set. Although the fold curve is a smooth curve, the bifurcation set has a sharp point, forming a cusp, and this is the reason for the name cusp-catastrophe. The cusp lines form the main thresholds for sudden behavioural change as we shall now explain.

The surface gives us a new insight into the dog's aggression mechanism. For, as his drives vary over the horizontal plane C, so his mood and behaviour will follow suit over the surface M above (except for the middle sheet). More specifically let us see what happens as his drives follow the dotted paths in C. Path P_1 begins with the dog frightened, cowering in a corner say, in a fleeing frame of mind, with his ears back. If we increase his rage, for example by approaching him too close and "invading his territory", then his mouth will begin to open, but he will remain cowering until the point Q_1 is reached. At that moment he reaches the fold curve at the edge of the lower sheet, and so the stability of his fleeing frame of mind breaks down, and he will suddenly catastrophically jump up onto

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

the upper sheet into a fighting frame of mind (indicated by the double headed arrow). Consequently he may suddenly attack. Conversely suppose he is in a fighting frame of mind, and we cause him to follow path P_2 by increasing his fear in some way, then he will nevertheless remain in a fighting frame of mind until the point Q_2 is reached, when he will suddenly and catastrophically jump down onto the lower sheet into a fleeing frame of mind. Consequently he may suddenly retreat. What actually causes these sudden changes of mind? Why should the mood jump from one surface to another? To answer these questions let us digress for a moment to a simpler mechanical example, and later return to the problem of the dog.

2. THE DYNAMIC.

Example 2. A catastrophe machine.

To understand how continuous forces can cause catastrophic jumps the reader is strongly recommended to make and play with the little toy illustrated in Figure 5a.

The materials needed are 2 elastic bands, 2 drawing pins, half a matchstick, a piece of cardboard and a piece of wood. Taking the unstretched length of an elastic band as our unit of length, cut out a cardboard disk of diameter about 1 unit. Attach the two elastic bands to a point H near the edge of the disk – the easiest way to do this is to pierce a small hole at H, push little loops of the elastic bands through the hole, and secure them by slipping the matchstick through the loops and pulling tight, as in Figure 5b. Now pin the centre of the disk to the piece of wood with drawing pin A (with the elastic bands on the top and the matchstick underneath), and make sure that it spins freely. Fix the other drawing pin B into the wood, so that AB is about 2 units, and the hook one of the elastic bands over B. The machine is now ready to go.

Hold the other end of the other elastic band : where you hold it is the control point c. Therefore the control space C is the surface of the wood. Meanwhile the state of the machine is the position of the disk, and this is measured by the angle $x = B\hat{A}H$. When the control c is moved smoothly, the state x will follow suit, except that sometimes instead of

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

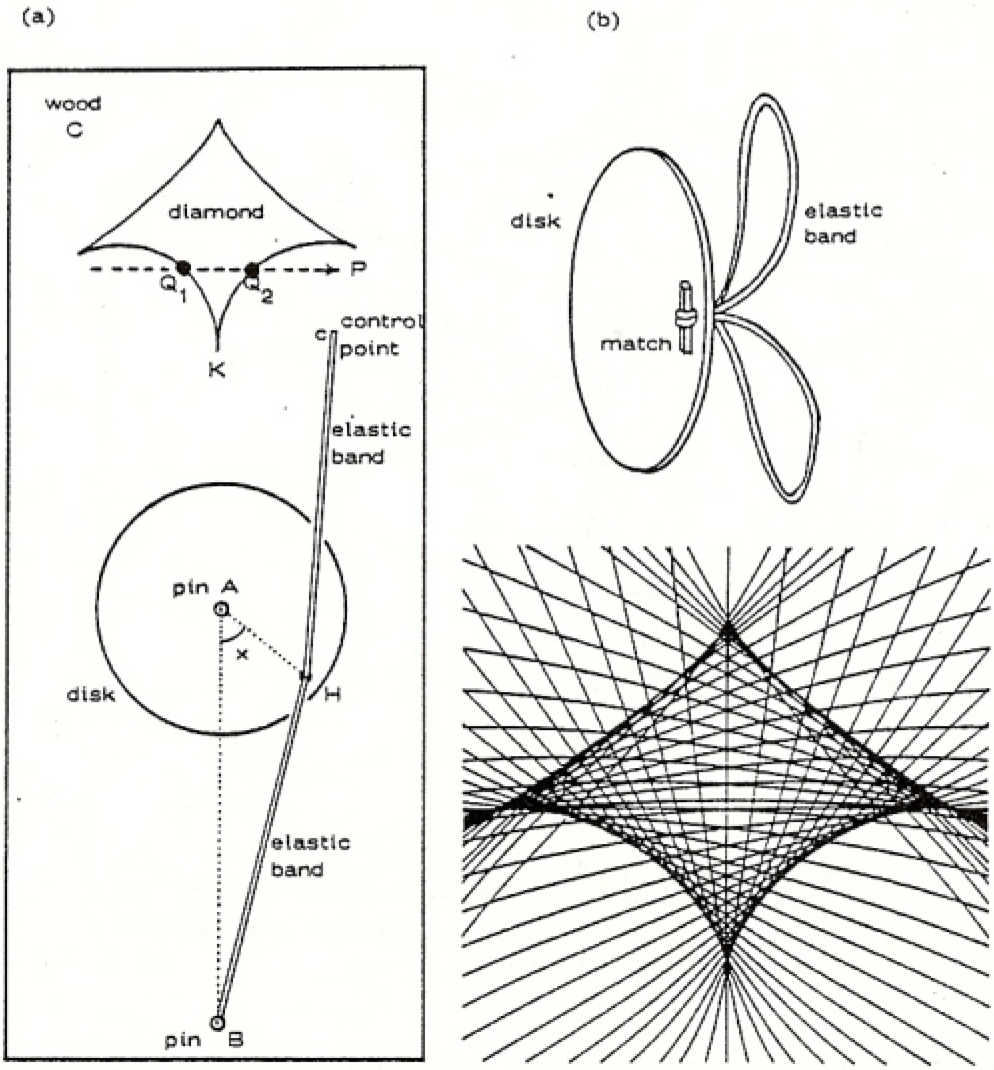


Figure 5. (a) A catastrophe machine; (b) how to attach the elastic bands; and (c) a computer-drawing by T. Poston and A.E.R. Woodcock of the diamond-shaped curve.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

moving smoothly it will suddenly jump. Every time it jumps mark the corresponding control point c with a pencil dot. Soon you will build up sufficient dots to be able to join them up in a concave diamond-shaped curve with four cusps, as shown in Figure 5a. The computer-drawing of this curve in Figure 5c shows it as the envelope of a family of lines (moving the control along any one of these lines keeps the disk stationary). The diamond-shaped curve is the bifurcation set, because if c lies outside the curve there is a unique stable equilibrium position of the disk, whereas if c lies inside there are two stable equilibria. If we restrict ourselves to control points in the neighbourhood of the lowest cusp point K, then the stable equilibria will have small angle x (this is convenient because we can then measure x along a line, whereas the full state space is in fact a circle). Let f denote the energy of the elastic bands. Then by Hooke's Law we can compute f as a function of x, for a fixed control point c. Examples of the graph of f for control points outside and inside the diamond are shown in Figure 6.

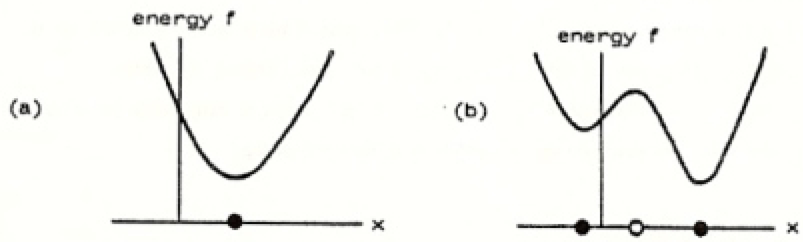


Figure 6. Graphs of the energy in the elastic bands for control points (a) outside and (b) inside the diamond.

The minima of f determine the stable equilibria (indicated by a solid circle) and the maximum of f determines the unstable equilibrium (indicated by an open circle). Figure 6 is analogous to Figure 3 with maxima and minima reversed. By the main theorem, the graph of equilibria as a function of c is equivalent to the cusp-catastrophe surface M shown in Figure 4. The upper and lower sheets represent the stable equilibria, and the middle

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

sheet the unstable equilibria separating them. The cusp in Figure 4 is exactly the same as the cusp we have drawn on the wood in Figure 5, in the neighbourhood of the point K.

In this example it is easy to understand the dynamic : by Newton's law of motion the disk rotates so as to reduce f, and any oscillations are swiftly damped out by the friction at the drawing pin. Therefore x swiftly seeks a local minimum of f. As the control is moved slowly, then f changes slowly, but the dynamic keeps x at the local minimum of f, or in other words keeps x on the surface M.

Let us now see what happens if the control is moved slowly from left to right along the dotted path P shown in Figure 5a. The sequence of energy functions is shown in Figure 7. The state x starts in the unique

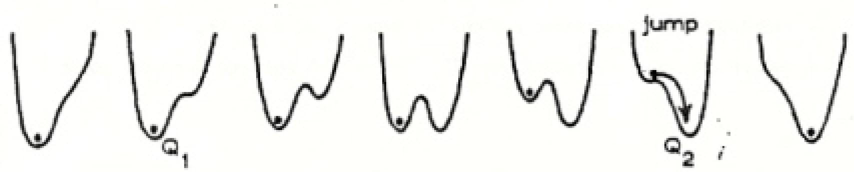


Figure 7. Changes in the energy graphs explaining the jump at Q_2.

minimum; nothing happens when a second minimum appears at Q_1. The second minimum eventually becomes deeper, but the state stays in the first minimum, held there stably by the dynamic. Eventually at Q_2 the stability of the first minimum breaks down as it coalesces with the maximum, and the state has to jump (or more precisely is swiftly carried by the dynamic) into the second minimum, which has now become the unique minimum. On the reverse journey the roles of Q_1 and Q_2 are reversed : the state stays in the right-hand minimum until Q_1, where it has to jump into the left-hand minimum again. This is called a hysteresis cycle, and is illustrated in Figure 8 (which is just the front section of Figure 4). The hysteresis of magnetism is in fact the same phenomenon happening to all the little magnets inside. Returning to the catastrophe machine, experiment

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

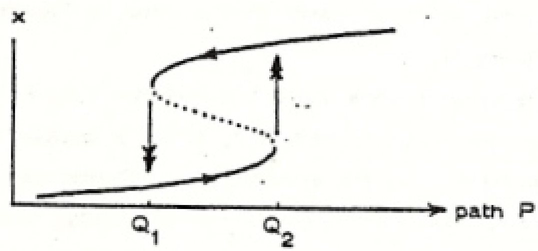


Figure 8. Going back and forth along P produces a hysteresis cycle.

will confirm that the cusp opposite K is similar, but the other two cusps are dual cusp-catastrophes, in the sense that the roles of maxima and minima are reversed : the upper and lower sheets are now unstable, and the middle sheet now represents a narrow pocket of stable equilibria. The difference can be observed by executing small circles of control around each cusp in turn.

3. THE MECHANISM OF AGGRESSION.

We now return to the rage and fear example and ask what is the dynamic? What mechanism can we add to the likelihood distributions of Figure 3 that will explain why the dog actively seeks and adopts the most-likely frame of mind? The answer lies in the underlying neurological activity of the brain. The brain may be regarded as a number of large coupled oscillators, each comprising millions of neurons. It is well known that non-linear oscillators can possess attractors (stable limit cycles), and that these attractors can typically bifurcate according to the cusp-catastrophe or higher dimensional catastrophes (See Figure 34). Therefore we may expect the elementary catastrophes to be typical models of brain activity, especially of activity in those parts of the brain such as the limbic system where the organs are more highly interconnected and consequently may tend to oscillate more as whole units (as opposed to the neocortex, whose different parts can oscillate differently at the same time, and whose activity can therefore be much more complicated).

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

According to Paul MacLean it is in the limbic system that emotions and moods are generated (while the neocortex determines the more complicated choice of behaviour within that mood). Therefore we might expect catastrophe theory to be the mathematical language with which to describe emotion and mood; and indeed it is striking that moods tend to persist, tend to delay before changing, and then tend to change suddenly, all of which qualities are typical of catastrophe models (see for example the hysteresis of Figure 8). Therefore it is not unreasonable to assume that Figure 4 is not only a model of the observed behaviour of aggression, but also a model of the underlying neural mechanism. Each point of the surface M represents an attractor of some huge dynamical system modelling the limbic activity of the brain, and the jumps occur when the stability of an attractor breaks down. The dynamical system remains implicit in the background; the only part that we need to make explicit for experimental predictions is the catastrophe model of Figure 4.

Moreover from the point of view of evolution one would expect this type of aggression mechanism to be very old, and therefore situated in a phylogenetically older part of the brain such as the limbic system. Consequently we should expect Figure 4 to be applicable not only to dogs, but to many species, under widely varying circumstances.

<u>Example 3. Territorial fish.</u>

Consider for example a territorial species such as some types of tropical fish. If we postulate that "size of invader and closeness to nest are conflicting factors influencing aggression", then the same argument leads again to the cusp-catastrophe. The resulting cusp in the control space is shown in Figure 9. Although Figure 9 is a somewhat simple picture, the argument deriving it has used a deep theorem, and we can deduce several consequences, leading to predictions which could probably be tested by fairly simple experiments. We might expect the cusp point to occur near (r_1, s_1) where s_1 is the size of the fish and r_1 the radius of his territory. If our fish meets a smaller adversary $(s < s_1)$ then he

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

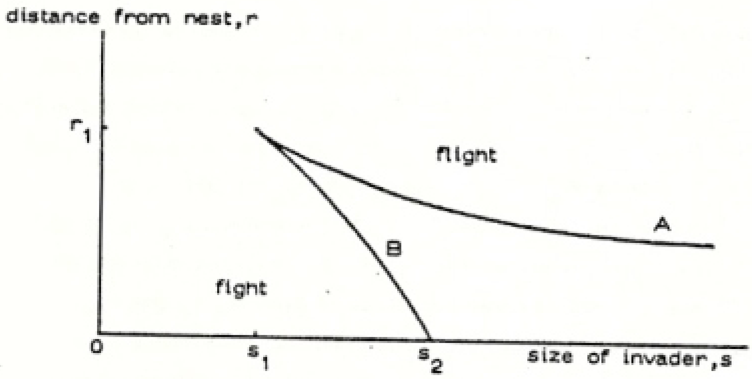


Figure 9. Size of invader and closeness to nest are conflicting factors influencing the aggession of territorial fish.

will chase him away. However if he meets a larger invader ($s > s_1$) near the nest, then he will chase the invader up to cusp line A, where his frame of mind will then jump from valour to discretion, and he will return to the nest. Therefore the radius of the territory, measured by A, should be a decreasing function of s. Now suppose our fish goes foraging and meets a large adversary far from the nest : then he will flee home, and only jump into a fighting frame of mind and turn to defend the nest when he reaches line B. Therefore the hysteresis phenomenon of Figure 8 here predicts that radius of the territory as measured by B will be noticeably smaller that that measured by A. Moreover if the invader is very large ($s > s_2$) then our fish will continue to flee; he will never turn to defend the nest, which would then be lost, were it not for his mate, who, if she has been cruising near the nest, will automatically chase the invader out to line A. Therefore the model offers a measurable test of a simple explanation of why in territorial species the partner who happens to be nearer the nest displays the more vigorous defence. Similarly one could devise many qualitative experiments of this nature on other species to test the general hypothesis that aggression mechanism is stored as a cusp-catastrophe. However now let us be a little more adventurous and make some applications of the model to man.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Example 4. Argument.

In the first application suppose that the model is describing the emotional mood of an opponent in an argument. The behaviour axis runs as follows

fight ↑ hysteria
abuse
irrational argument
rational discussion
concessions
apologies
flight | tears

If we begin to make him angry and frightened, then we will first deny him access to rational thought, and force him to jump between irrational argument and concessions. If we make him a little more so, then we will next deny him access to those behaviour modes, and force him to make bigger jumps between abuse and apology. Finally if we make him very angry and frightened, then we will limit his possible behaviour to only hysteria or tears, and the very verbal usage of the phrase "hysterical tears" is confirmation of the catastrophic jump from one end of the spectrum to the other. If our purpose is to persuade him of something that we know will make him both angry and frightened, then the best policy is to state the case and go away; for then our absence will allow his anger and fear to subside, and give him access to rational thought again, and so enable him to see our point.

Example 5. Catharsis of self-pity.

In the second application, we modify the coordinates of rage and fear to frustration and anxiety, to suit the emotions of modern civilised man. One of the behaviour patterns of a fleeing frame of mind that we learn to adopt as children is the mood of self-pity. And when a child, or an adult, falls into a persistent mood of self-pity, it often seems as if sympathy is of no avail. But a sarcastic remark may suddenly induce a loss of temper, which in turn releases the tension, and seems to act as a catharsis for the mood of self-pity. This is a fairly common pattern of events, although

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

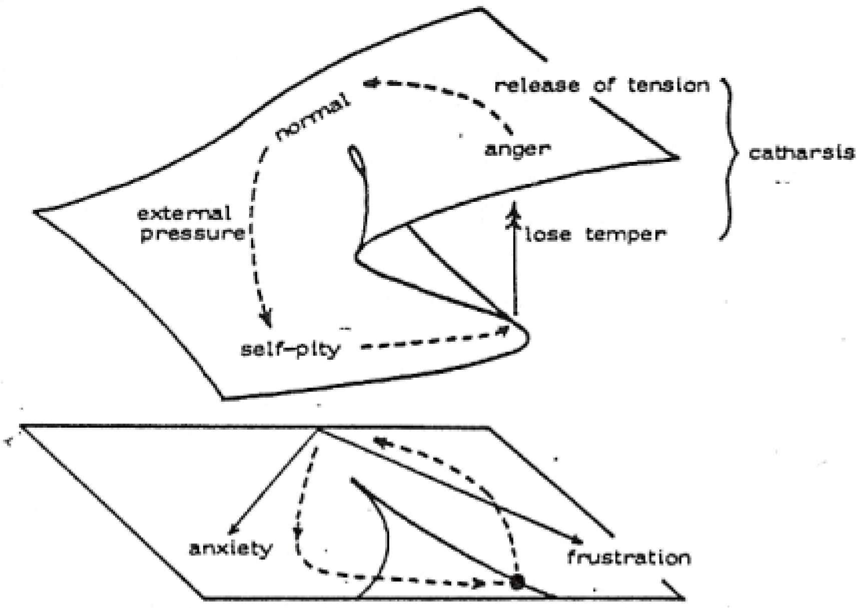


Figure 10. Anxiety can cause self-pity. Increasing frustration can cause catharsis of self-pity by loss of temper and release of tension.

It always seems a shame that sarcasm should suceed where sympathy has failed, even though the proverb tells us that it is necessary to be cruel in order to be kind. The model suggests that the pattern may in fact be just an automatic byproduct of the evolutionarily useful built-in aggression mechanism.

Example 6. War.

In the third application we go even further, passing from psychology to sociology, and apply the same model to whole nations instead of individuals. In place of rage and fear we substitute threat and cost. The behaviour axis represents the war policy of the nation, running from strong military action, through moderate or weak action to neutrality, withdrawal and surrender. The distributions of Figure 3 will represent the support amongst the population for the various policies. The two modes of case (4) are called doves and hawks. The dynamic in this example is the sensitivity of the government to its electors, continuously adapting its policy so as to increase

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

its support. The path P_1 on Figure 4 represents a nation feeling increasingly threatened, but first pursuing a policy of appeasement until it reaches point Q_1 when it suddenly declares war. The path P_2 represents a nation suffering increasing costs, but first escalating the war, until it reaches point Q_2 when it suddenly surrenders. Paths P_3 and P_4 represent two nations both experiencing similar escalating threat and cost, but one finishing up in an aggressive mood and the other in an appeasing mood, very reminiscent of the Cuban Missile Crisis in 1962. America followed path P_3 to the right of the cusp, because her leaders first felt threatened by the presence of the missiles, and then increasingly appalled by the rising likelihood of a nuclear war. Russia on the other hand followed path P_4 to the left of the cusp, because her main feeling of threat came later as the crisis escalated. The art of diplomacy is to leave your adversary a clear avenue of retreat so that he can safely follow P_4, while you yourself follow P_3, otherwise if you overthreaten him too soon, you may force him to follow P_3 as well, with disastrous results for both of you.

Admittedly these applications may be gross oversimplifications, but the very fact that our minds indulge in simplification, analogy, abstraction and synthesis may in fact reveal that the model is telling us not so much about wars, but more about the way our minds work. The oscillations of our brains can bifurcate according to the elementary catastrophes, and so our minds automatically employ the latter for the subconscious organisation of our thoughts.

4. DIVERGENCE.

Let us re-examine paths P_3 and P_4 in Figure 4 in the context of our original application to the dog. Both paths begin at the same point and end at the same point in the control space, C, but induce divergent behaviour. Following P_3 the dog first gets angry and then frightened, but persists in a fighting frame of mind; conversely following P_4 he experiences the same emotions but in the reverse order, and persists in a fleeing frame of mind. Therefore although, as we have said, the behaviour in case (4) is unpredictable, if we happen to know the recent past history then it is

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

predictable. Usually the psychologist's only defence against unpredictability is statistics, but the use of statistics in conjunction with a model of this nature is a much stronger weapon.

Notice that the change of behaviour under paths P_3 and P_4 was quite smooth without any catastrophes involved. Notice also that in the plane C the difference between the two paths may be only very marginal : all that matters is that they pass on either side of the cusp point. This phenomenon of a marginal change of path causing a major change in the behaviour we call divergence, and it is very common in biology and the social sciences. By contrast physics is generally non-divergent, because usually a small change in the initial data causes only a small change in the ensuing motion. It has long been a folk-lore that divergent phenomena in the "inexact" sciences could not be modelled by mathematics; but it is now realised that divergence is a characteristic property of stable systems, which can be both modelled and predicted, and the natural mathematical tool to use is catastrophe theory.

In fact the cusp-catastrophe shows that the five qualitative features of bimodality, inaccessibility, sudden jumps, hysteresis and divergence, are all interrelated (see Figure 11). And the deep classification theorem of catastrophe theory (stated below) permits us to enunciate the general principle that whenever we observe one of these five qualities in nature, then we should look for the other four, and if we find them then we should check whether or not the process can be modelled by the cusp-catastrophe. Indeed our verbal usage in ordinary language of pairs of opposites frequently indicates a bimodality that has grown smoothly out of some unimodality, and which may be modelled by the upper and lower sheets of a cusp-catastrophe surface. In Table 1 we give an assorted list of pairs, to indicate the scope of possible applications. In each case a fairly elaborate model can be developed, but we do not have space to develop them here, because each one needs careful analysis of what the axes should be, how they could be measured, what is the dynamic, and how to design experiments.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

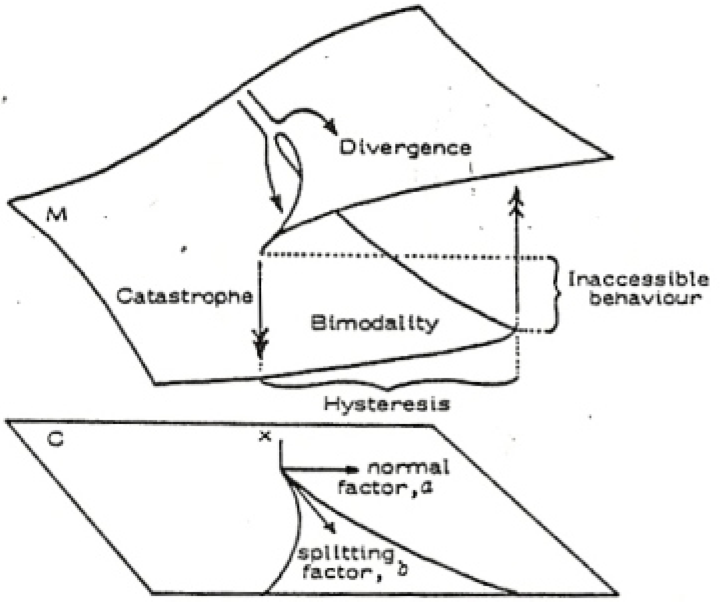


Figure 11. Five characteristic properties of the cusp-catastrophe are bimodality, inaccessibility, catastrophe, hysteresis and divergence. (The unstable middle sheet has been removed.)

In some cases the two control factors (or parameters) in C lie on either side of the cusp, such as α, β in Figure 4; in this case we call them <u>conflicting factors</u> as in all our previous examples so far. In other cases one of the control factors is perpendicular to the cusp axis, and the other lies along it, such as a, b in Figure 11 (see also Figures 12,17,30,31). In this case we call a the <u>normal factor</u>, because if $b < 0$ then x increases continuously with a, and we call b the <u>splitting factor</u>, because if $b > 0$ then M is split into two sheets. The equation of the standard cusp-catastrophe surface M illustrated in Figure 11 (with origin taken at the point on M above the cusp point) is

$$x^3 = a + bx \ .$$

By differentiating and elminating x, one obtains the equation of the cusp : $27a^2 = 4b^3$. To get the standard equation in terms of conflicting factors put $a = \alpha - \beta$, $b = \alpha + \beta$. We illustrate normal and splitting factors by Example 7, which is an application in economics.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Table 1.

	Phenomenon.	Bimodality.	Catastrophes.	Conflicting factors.
	Economic policy.	Deflation, reflation.	Stop-go.	Balance of payments, unemployment.
4	Embryology.	Ectoderm, mesoderm.	Cell determination.	Space, time.
2	More haste less speed.	Fast, slow.	Master, fumble.	Haste, skill.
				Normal/splitting factors.
13	Civil unrest.	Disorder, quiet.	Riot, restore order.	Tension/alienation.
*	Delinquency.	Harassed, lonely.	Crime, quarrel.	Approach-avoidance/anxiety.
20	Education.	Exciting, dreary.	Inspiration, disillusion.	Reward/punishment.
*	Emotional response.	Antagonism, sympathy.	Lose temper, reconcile.	Tiresomeness/demandingness of other person.
3	Heartbeat.	Systole, distole.	Contract, relax.	Pacemaker/blood pressure.

↑ Numbering of relevant papers in this book. (For * see *Proc. Roy. Inst. Gt. Br.* 49 (1976) 77-92).

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Example 7. Stock markets.

First we observe the bimodality in the common usage of the terms "bull market" and "bear market", which are situations that diverge smoothly from the more "normal market". Therefore we shall represent bull and bear by the upper and lower sheets of the cusp-catastrophe illustrated in Figure 12. Next we verify the presence of catastrophes ; a "crash" is a

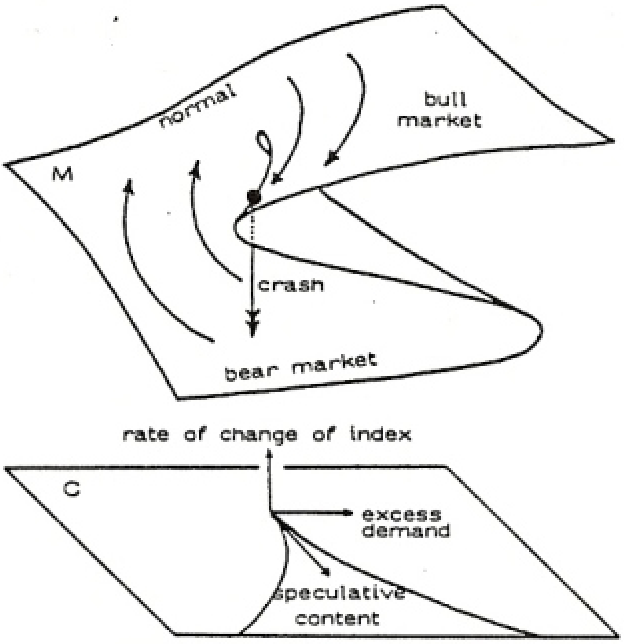


Figure 12. Excess demand is a normal factor, and speculative content a splitting factor influencing stock market behaviour.

sudden jump from bull to bear. This immediately prompts the question : why is this catastrophe more common than the opposite sudden jump from bear to bull? We shall explain the answer in a moment. First consider what is the behaviour axis? In a bull market the index is rising, while in a bear market it is falling, and so we choose rate of change of index to measure behaviour. What is the most likely reason for the index to

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

rise in a normal market? Normally the main reason is excess demand for stock by investors, and so we choose excess demand as the normal factor. What causes an abnormal market? One of the main reasons is probably the speculative content, although this may be more difficult to measure; perhaps one might be able to measure the percentage of the market held by "chartists" (investors who base their investment policy on charts of previous performance) as opposed to "fundamentalists" (who base their policy upon research into the industrial health and growth-potential of the firms involved).

If the excess demand is zero, but the speculative content high, then it is unlikely that the index will remain constant; a random event may spark off a wave of confidence that induces a stable bull market, or equally a lack of confidence that induces a stable bear market. The main feature of a stock market is its centralised sensitivity, and this is what is responsible for the rapid dynamic that causes the behaviour to go bimodal, and follow the surface M shown in Figure 12.

We can now go further and introduce a slow feedback of the index upon the two controls, representing this by the arrows shown on the surface. A bull market encourages speculation, but overvaluation causes fundamentalists to sell and invest their profits elsewhere; therefore the arrows on the upper sheet come forward and bend to the left. Conversely a bear market discourages speculation, but an undervalued recovering market presents good opportunities for fundamentalists to reinvest; therefore the arrows on the lower sheet go back and bend to the right. We can now deduce from the model the characteristic overall stock-market cycle of growth, boom, recession and recovery. The anthropologist Michael Thompson has shown that similar economic cycles happen, and are even planned, amongst primitive New Guinea tribesman.

5. CLASSIFICATION THEOREM.

We are now ready to appreciate Thom's classification of elementary catastrophes, and to go on to the higher ones. In order to keep things as understandable as possible we begin by stating a simplified version of part of the theorem, and then elaborate upon it.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Theorem. Let C be a 2-dimensional control (or parameter) space, let X be a 1-dimensional behaviour (or state) space, and let f be a smooth generic* function on X parametrised by C. Let M be the set of stationary values of f (given by $\frac{\partial f}{\partial x} = 0$, where x is a coordinate for X). Then M is a smooth surface in C x X, and the only singularities of the projection of M onto C are fold curves and cusp-catastrophes.

Remark about singularities. Here a singularity means a point where a vertical line touches M. When we say that a singularity of M is a cusp-catastrophe we mean that near that point M is equivalent** to the standard surface shown in Figure 4. Equivalence preserves all qualitative features such as the foldcurve, the cusp, the bimodality, catastrophes, hysteresis, divergence and inaccessibility. Therefore what the theorem really says is that qualitatively Figure 4 is locally the most complicated thing that can happen to a graph. That is why the cusp-catastrophe can be used with such confidence in so many different fields, whenever a process involves 2 causes and 1 effect.

Remark about the function f. In the rage and fear example the function f is the likelihood function on X parametrised by C, shown in Figure 3. The stationary values of f are the maxima and minima, representing the most-likely and least-likely behaviour respectively. In Example 6 the function f measured the support for different policies. In the catastrophe machine f was an energy function, with minima representing stable equilibria, and maxima representing unstable equilibria. In economics f might be a cost function, in evolution a fitness function, in

---

Footnotes for the mathematically minded :

* Smooth means differentiable to all orders.
Generic means that the map from C to the space of functions on X is transverse to the natural stratification. Almost all smooth functions are generic. Small perturbations of generic functions remain generic.

** Equivalence means there is a diffeomorphism from a neighbourhood N in the given C x X onto the standard C x X, throwing vertical lines to vertical lines, and throwing $M \cap N$ onto the standard surface.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

engineering a Lyapunov funcation, in light caustics a geodesic distance, etc. In most applications there is, in addition to f, a gradient-like (dissipative) dynamic that maximises or minimises f, and is ultimately responsible for the sudden jumps. However the dynamic is not involved in the statement of the theorem, and so the theorem is also applicable to phenomena such as light caustics (Example 11 below), where there are discontinuities, but no dissipative dynamic and therefore no sudden temporal jumps.

Remark about dynamics. In many applications there are dynamics on X that are not gradient-like, and consequently there is no function f that is minimised. In this situation the elementary catastrophes do not describe all the possibilities, because non-elementary catastrophes can occur. Some non-elementary catastrophes are known (such as the Hopf bifurcation) but as yet the general classification problem is unsolved.

Remark about the dimension of X. The theorem remains true, word for word, if we increase the dimension of the behaviour space X from 1 to n; that is why it is both remarkable and difficult to prove. But the beauty of this result is that we can now use the theorem implicitly, in situations that would be far too complicated to measure or put on a computer. For example we can implicitly assume X is large enough to describe the states of a cell in the embryo, with at least 10,000 dimensions for representing the concentrations of the various chemicals involved. Or we could implicitly assume X is large enough to describe the states of the brain, with at least 10,000,000,000 dimensions for representing the rates of firing of all neurons. In physics we may wish to have X infinite dimensional (as in Example 12 below). But in each case the theorem explicitly hands us back the same simple surface of Figure 4, upon which to base a model. We feed in implicit complexity, and get out explicit simplicity; all the hard work of digesting the complexity has gone on behind the scenes in the proof of the theorem.

Remark about the fold catastrophe. If we reduce the dimension of C from 2 to 1, then the analogous theorem, 1 dimension lower, says that M is

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

a smooth curve, and the only singularities are folds, such as the two that occur in Figure 8. Thus the fold-catastrophe appears in sections of the cusp-catastrophe, and the latter is made up of folds together with one new singularity at the origin. Similarly any higher dimensional catastrophe is always made up of lower dimensional ones, together with one new singularity at the origin.

6. THE SEVEN ELEMENTARY CATASTROPHES.

If we increase the dimension of C in the theorem, then this is quite a different kettle of fish. For example suppose C was 3-dimensional and X was 1-dimensional. Then C x X would be 4-dimensional, and M would be a 3-dimensional manifold (or hypersurface) in C x X. What is more, instead of being folded along curves M would now be folded along whole surfaces, and the bifuraction set instead of consisting of curves with cusp points in 2-dimensions would now consist of surfaces with cusped edges in 3-dimensions (see Figure 13). Moreover a new type of singular point would appear, called the swallowtail-catastrophe. It is impossible to draw the complete picture for M in the case of the swallowtail because this would require 4-dimensions. However just as the geometry of the cusp catastrophe can be described by drawing a cusp in 2-dimensions (and knowing that M is bimodal over the inside), so also we can derive some geometric intuition about the swallowtail by drawing its bifurcation set in 3-dimensions (see Figure 13). It is called the swallowtail or dovetail because it looks a bit like one : the name queue d'aronde was suggested by the blind French mathematician, Bernard Morin.

Now suppose we keep C 3-dimensional and allow X to be n-dimensional, $n \geq 2$. Then two more singular points become possible, the hyperbolic-umbilic-catastrophe and the elliptic-umbilic-catastrophe, which we can again describe geometrically by drawing their bifurcation sets in 3-dimensions (see Figures 13,27,32). This completes the list of elementary catastrophes in 3-dimensions. If we now allow C to be 4-dimensional then we obtain two new singularities, the butterfly-catastrophe and the parabolic-umbilic-catastrophe (see Figures 14,16). And so on, until eventually the list becomes infinite. However the infinite list is more the concern of the

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

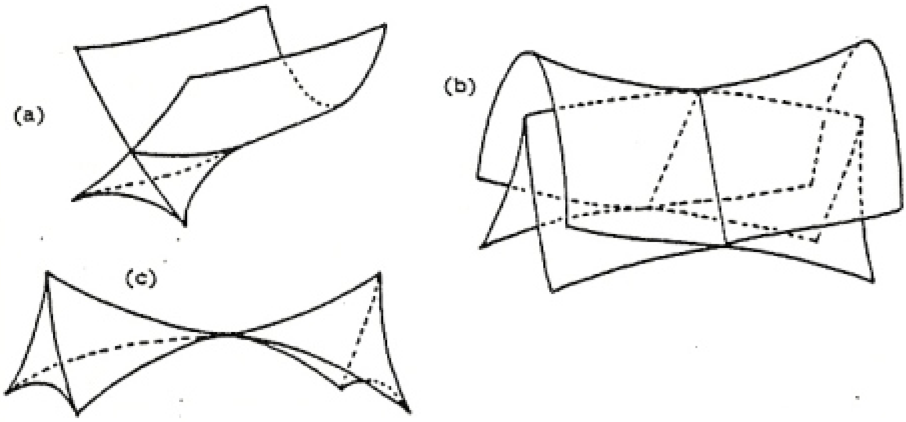


Figure 13. Bifurcation sets of (a) the swallowtail, (b) the hyperbolic umbilic and (c) the elliptic umbilic catastrophes.

mathematician, whereas here we are primarily interested in the lower dimensions because they are more relevant to applications. In particular we are interested in up to 4-dimensions, because the parameter space C often plays the role of space-time. Therefore the 7 elementary catastrophes up to 4-dimensions have all been given special names, whereas those above have not.

Table 2.

Dimension of C	1	2	3	4	5	...
Number of catastrophes	1	1	3	2	4	...
Names	fold	cusp	swallowtail hyp. umbilic ell. umbilic	butterfly par. umbilic	– – – –	...

Note that in Table 2 we have put each catastrophe in the lowest dimension where it first appears, although it also appears in all higher dimensions. For example fold points first occur when C is dimension 1 (in Figure 8) and fold curves also appear when C is dimension 2 (in Figure 4). Therefore all 7 catastrophes appear when C is dimension 4.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Each of the elementary catastrophes has a standard model, and in Table 3 we list some standard formulae for f, from which the standard models can be derived. In each case a,b,c,d are parameters for C, and x,y are variables for X. For the cuspoids M is given by $\frac{\partial f}{\partial x} = 0$, and for the umbilics M is given by $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$. Note that the "standard" formulae are by no means unique, but are chosen for convenience in applications. In particular the fractions are there only so that they disappear when we differentiate to get the equation for M. For instance the cusp-catastrophe is given by $\frac{\partial f}{\partial x} = x^3 - a - bx = 0$ (see Figure 11). The minus signs are there in order to fit in with the notions of normal and splitting factors, etc.

Table 3.

		dim X	dim C	Function f
Cuspoids	Fold	1	1	$\frac{1}{3}x^3 - ax$
	Cusp	1	2	$\frac{1}{4}x^4 - ax - \frac{1}{2}bx^2$
	Swallowtail	1	3	$\frac{1}{5}x^5 - ax - \frac{1}{2}bx^2 - \frac{1}{3}cx^3$
	Butterfly	1	4	$\frac{1}{6}x^6 - ax - \frac{1}{2}bx^2 - \frac{1}{3}cx^3 - \frac{1}{4}dx^4$
Umbilics	Hyperbolic	2	3	$x^3 + y^3 + ax + by + cxy$
	Elliptic	2	3	$x^3 - xy^2 + ax + by + c(x^2 + y^2)$
	Parabolic	2	4	$x^2y + y^4 + ax + by + cx^2 + dy^2$

Each catastrophe has its own individual and surprising geometry. We do not have space to describe them fully here, and the reader is referred to Thom's book for more details. The fold (Figure 8) and cusp (Figure 4) are easy to understand, and the bifurcation set of the swallowtail and hyperbolic and elliptic umbilics are easily visualised (Figure 13). However the two 4-dimensional ones have to be approached more obliquely by drawing sections. Figure 14 shows sections of the bifurcation set of the parabolic umbilic in the (a,b)-plane for various fixed values of c,d on the

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

unit circle in the (c,d)-plane; these sections are taken from Fowler's translation of Thom's book, and based on drawings by Chenciner and Jänich.

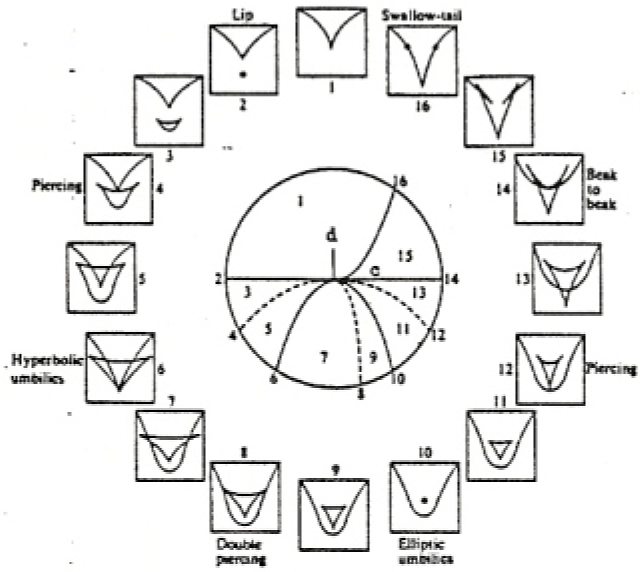


Figure 14. Sections of the 4-dimensional bifurcation set of the parabolic umbilic.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

7. THE BUTTERFLY CATASTROPHE

Finally we come to the butterfly, which deserves more attention, because after the cusp it is the most important catastrophe for the behavioural sciences. We shall illustrate it by giving a fairly elaborate application to a nervous disorder, anorexia nervosa.

What bimodality is to the cusp, so trimodality is to the butterfly. We have seen how any evolution from unimodal to bimodal behaviour determines (by the classification theorem) the unique 3-dimensional geometry of the cusp-catastrophe, with its associated jumps, hysteresis and divergence, etc. Similarly any evolution from unimodal to trimodal behaviour determines the unique and much richer 5-dimensional* geometry of the butterfly-catastrophe. Since trimodality often emerges out of bimodality, the natural way to analyse the butterfly is to regard it as an extension of the cusp, as illustrated by the following example.

Example 8. Compromise opinion.

Suppose the cusp represents the polarisation over some issue in society; then the butterfly represents the emergence of a compromise opinion. For instance in the example above of nation at war, public support might be distributed as in Figure 15 (compare with Figure 3).

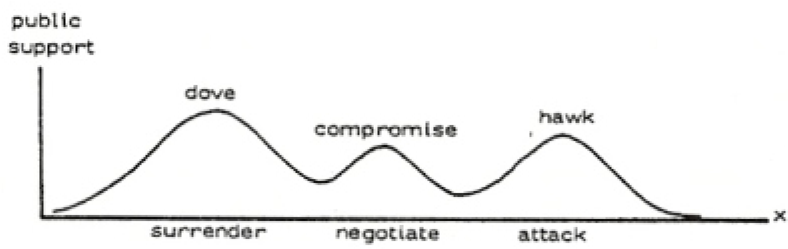


Figure 15. The emergence of compromise opinion.

---

* Each extra mode involves two more control factors, because it requires one more maximum and one more minimum (see Figure 15) and so the power of the leading term in the formula goes up by 2.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

The geometry of the butterfly. There is one state variable x, and four control factors as follows

a : normal factor

b : splitting factor

c : bias factor

d : butterfly factor.

The behaviour lies on the "surface" M in 5-dimensions given by the equation

$$x^5 = a+bx+cx^2+dx^3,$$

which is obtained by differentiating the formula in Table 3 above. Since it is impossible to draw 5-dimensional pictures we have to make do with 2- and 3-dimensional sections. The bifurcation set lies in the 4-dimensional control space, and the top six pictures of Figure 16 show 2-dimensional sections of it parallel to the (a,b)-plane for different values of c,d. The top three pictures refer to $d < 0$. When $c = 0$ the section reduces to the cusp that we know already. The effect of the bias factor c is to bias the position of the cusp : when $c < 0$ the main body of the cusp swings to the right while the tip of the cusp moves up and bends over the left; when $c > 0$ the opposite happens. Meanwhile the effect of the bias upon the behaviour surface M is to move it up and down; this can be seen in the bottom pictures (vii) and (viii) of Figure 16 which show 2-dimensional sections of M drawn with control factor, a, horizontal and x vertical, for different values of b,c,d. Meanwhile the effect of bias upon the 3-dimentional sections of M is shown in Figure 19(i) and (ii), which are drawn with (a,b) horizontal and x vertical, for different values of c,d.

Now consider the effect of the butterfly factor d. Keeping $c = 0$, as d goes positive the cusp evolves into three cusps, which form a triangular "pocket", as shown in Figure 16(v). (Turning this picture upside down looks like a butterfly, which originally suggested the name.) Meanwhile the behaviour surface develops two new folds above the sides of the pocket as illustrated in Figures 16(ix) and 20. Above the pocket itself is a new triangular sheet of stable behaviour, which has grown continuously forward from the back, in between the upper and lower sheets at the front. This new sheet represents the third behaviour mode, for instance the emergence

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

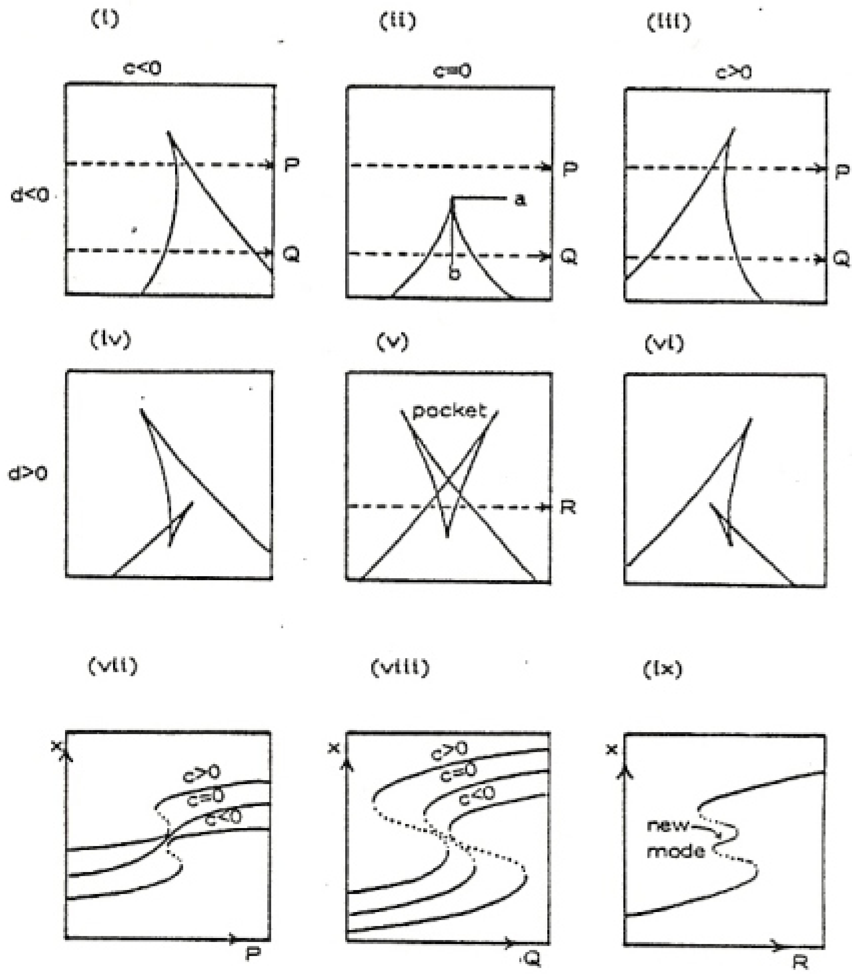


Figure 16. The butterfly catastrophe.

(i) - (vi) Sections of the bifurcation set in the (a,b)*-plane for different values of* c,d. *The effect of the bias factor* c *is to swing the cusp to and fro. The effect of the butterfly factor* d *is to create the pocket.*

(vii) - (ix) Sections of the behaviour surface over the dashed paths P,Q,R. The effect of c *is move the surface up and down. The effect of* d *is to create a new sheet above the pocket, representing the third mode.*

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

of the compromise opinion in Example 8 above. Notice that in Figure 16(iv) and (vi) the effect of the bias on the pocket is to reduce one side or the other until it disappears (by means of a swallowtail). Therefore the effect of bias is to destroy a compromise. In applications concerning the emergence of compromise, the butterfly factor will increase with time; at first the compromise is fragile, in the sense that its stability is broken by any perturbation across the nearby sides of the pocket; but as the pocket grows in size the compromise becomes stronger, in the sense of being stable under increasingly large perturbations.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

PART TWO

ANOREXIA NERVOSA

Anorexia is a nervous disorder suffered mainly by adolescent girls and young women, in whom dieting has degenerated into obsessive fasting. It generally begins between the ages of 11 and 17, although it can start as early as 9 or as late as 30. It can lead to severe malnutrition, withdrawal and even death.

The proposed model is the joint work of the author and J. Hevesi, who is a psychotherapist specialising in anorexia. Hevesi has spent some 5000 hours during the last 5 years talking to over 150 anorexics and the model is based on his close observations. Of these 150 over 60 agreed to undertake his course of treatment, and of those treated he has achieved an 80% success rate of complete cure. His innovation is the use of trance-therapy. The Anorexic Aid Society in Britain recently conducted a survey of over 1000 anorexics, and the secretary of the society, Mrs. P. Hartley, who is a psychologist, writes : "I first read of Mr. Hevesi in several letters from patients who responded to my appeal for information about anorexia nervosa, and their experience re. treatment. These patients are the only ones who claim that they have recovered completely - i.e. those whose attitude to life has changed since undergoing Hevesi's treatment. They are not just eating properly (only the awful surface problem anyway) but living a full life as a complete personality." (her underlining).

The advantage of using mathematical language for a model is that it is psychologically neutral; it permits a coherent synthesis of a large number of observations that would otherwise appear disconnected, and in particular enables us to place the trance states in relation to other behavioural modes. As yet the model is only qualitative, in the sense that the predictions that have been verified by observation have been qualitative rather than quantitative. Nevertheless it does provide a conceptual framework within which the theory could also be tested quantitatively by monitoring patients. Meanwhile we hope that it may not only give a better understanding of anorexia and its cure, but also provide a prototype for understanding other types of behavioural disorder.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

A striking feature of anorexia is that it sometimes develops a second phase after about two years, in which the victim finds herself alternately fasting and secretly gorging; the medical name for this is bulimia, and anorexics often call it stuffing or bingeing. If we regard the normal person's rhythm of eating and satiety as a continuous smooth cycle of unimodal behaviour, then we can interpret this second phase of bimodal behaviour, as a catastrophic jumping between two abnormal extremes. Therefore by the main theorem we can model the anorexic's behaviour by a cusp-catastrophe, in which she is trapped in a hysteresis cycle, as in Figure 17.

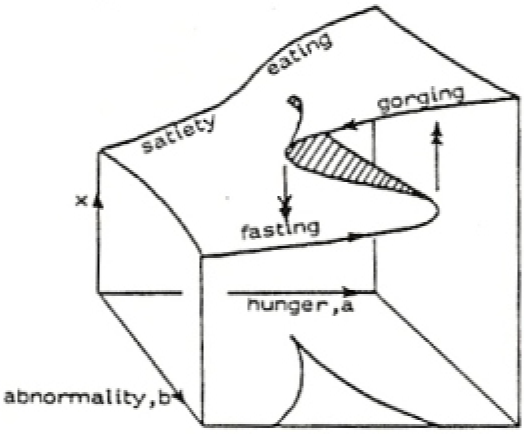


Figure 17. Initial behaviour model for anorexia.

Before we begin to analyse the model, we can immediately draw one important conclusion : <u>the victim will be denied access to the normal modes in between.</u> This denial of access to normal modes occurs already during the first phase of only fasting. Thus the main thrust of our approach will be to explain anorexia not as the complicated behaviour of a perverse neurotic, but as the logical outcome of a simple bifurcation in the underlying brain dynamics. If this is the case then catastrophe theory at once indicates a theoretical cure : if we can induce a further bifurcation according to the butterfly catastrophe, then this should open-up a new

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

pathway back to normality. The practical problem is how to devise a therapy that will induce such a bifurcation, and this is what Hevesi's treatment achieves.

In Figure 17 we have chosen hunger and abnormality as the two control factors (a,b). Hunger is the normal factor because hunger normally governs the rhythmic cycle between eating and satiety; there are various known methods for measuring hunger, but we do not yet know which will be best to use for quantitative testing. We postpone the discussion on the measurement of abnormality until later. Meanwhile to measure the behaviour, x, it would be necessary to find some psychological index that correlates with the scale of wakeful states shown in Figure 18.

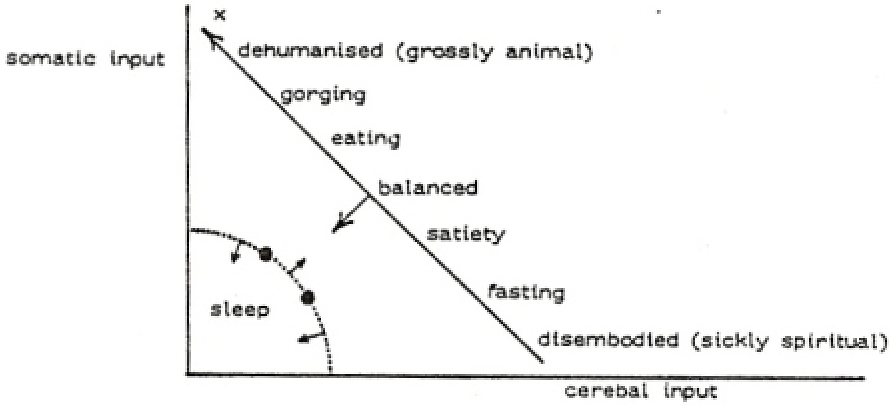


Figure 18. The x-axis measures both wakeful behaviour, and the relative weight given to the cerebal and somatic inputs to the limbic brain. The dotted line shows the boundary of the sleep basin, and the arrows show the movement of this boundary due to anorexia, leaving fixed the two nodal points.

What actually governs the behaviour is the underlying brain state, and if it is true, as MacLean suggests, that emotion and mood are generated in the limbic brain, then it is likely that x is measuring some property of limbic states. Since the limbic brain receives both cerebal inputs from the neo-cortex, and somatic inputs from the body, we might conjecture that x is some measure of the relative weight given to those

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

inputs as shown in Figure 18. Of course such a conjecture must remain speculative until it is confirmed or rejected by future brain research. Nevertheless the conjecture has already proved useful in explaining many of the symptoms of anorexia, and, what is perhaps more important, enabled us to identify what may be the key operative suggestions in the therapy, as we shall see. Meanwhile the conjecture implies that the main neurological feature of anorexia is that during wakefulness the limbic brain is dominated either by cerebal inputs or by somatic inputs, while the balanced states have become unstable, and therefore inaccessible.

Before leaving Figure 18 notice that it is 2-dimensional. From the psychological point of view the natural axes to use are x and y, which are inclined at 45° to the neurological axes, cerebal and somatic. Here x measures the different wakeful states, while y measures the difference between wakefulness and sleep, and y exhibits the familiar healthy catastrophes of falling asleep and waking up. For a more complete model we ought really to use both the behaviour variables x and y, 5 controls, and a 7-dimensional catastrophe called E_6. However this is beyond the scope of this article, and so for simplicity of presentation we shall sacrifice y and use only x.

We now introduce a third control factor, c, which will play the role of the bias factor in the butterfly catastrophe. Define c to be loss of self-control, measured by loss of weight. Geometrically the effect of bias is to swing the cusp to and fro as in Figure 16(i) and (iii). The resulting effect on the behaviour surface is shown in Figure 19.

During the first phase of the disorder the anorexic is firmly in control of herself, and so $c < 0$ as in Figure 19(i). The normal person has learnt to perform the regular smooth cycle at the back, socially structured by mealtimes. The anorexic however finds herself trapped on the lower sheet at the front by the abnormality; in other words the limbic brain oscillates continuously in states underlying a fasting frame of mind all the time she is awake, even when she goes through the motions of eating. The frame of mind is predominantly cerebal, and the victims often speak in terms of "purity"; it tends to smother instincts

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

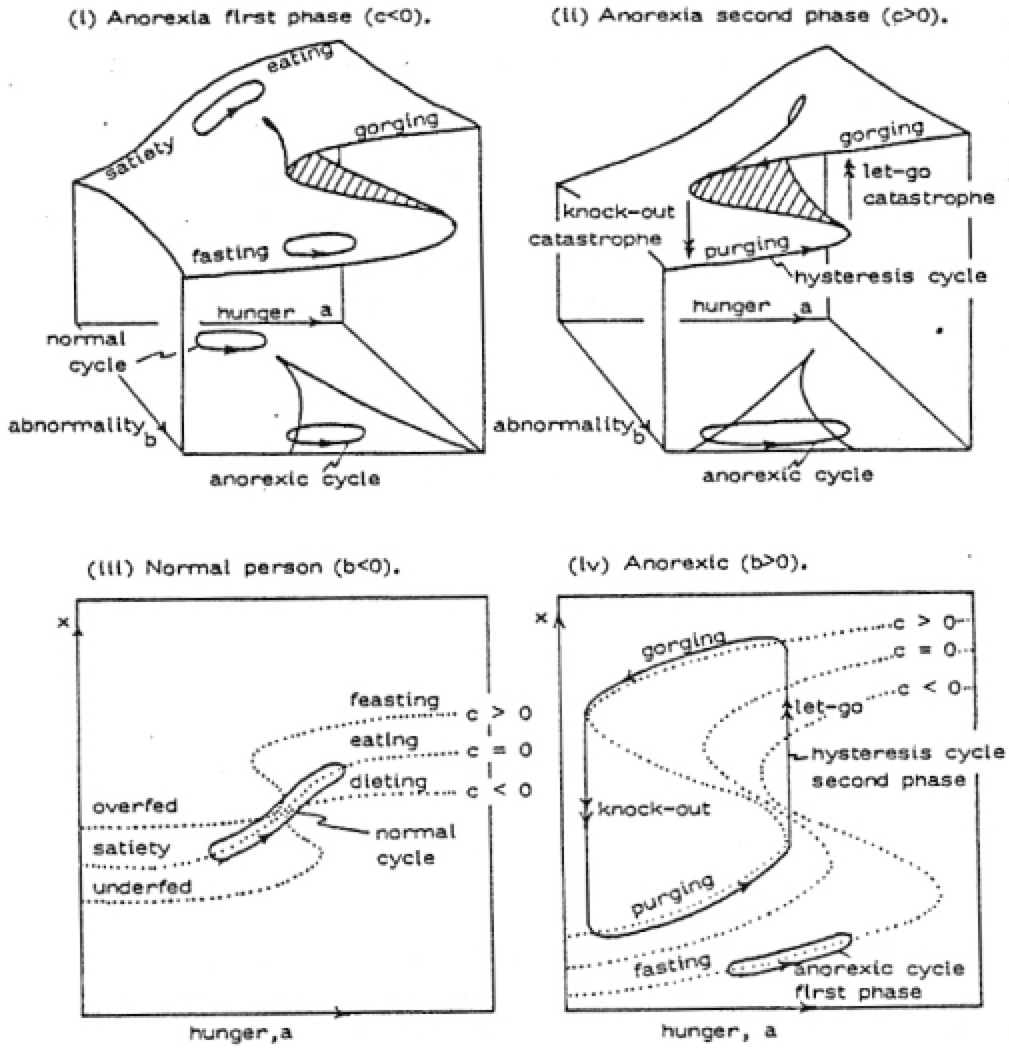


Figure 19. The effect of the bias factor, c.

(i) Anorexia first phase (c<0). *Strong self-control swings the cusp to the right, and abnormality displaces the normal cycle into fasting.*

(ii) Anorexia second phase (c>0). *Loss of self-control swings the cusp to the left, causing the anorexic to jump into the catastrophic hysteresis cycle of alternately gorging and purging.*

(iii) Normal person (b<0). *Changes in bias modify the behaviour slightly.*

(iv) Anorexic (b>0). *Changes in bias modify the behaviour dramatically.*

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

and produce excessive verbalisation. During the first phase victims often deny being ill, and refuse treatment.

Then as the anorexic gradually loses weight, she gradually loses control of herself; the bias factor c gradually increases, causing the cusp to swing gradually to the left, as in Figure 16(iii) and 19(ii). How far the cusp will eventually swing in relation to the cycle will depend upon the individual. If it swings sufficiently far for the right-hand side of the cusp to cross the right-hand end of the cycle, then this will cause the sudden onset of the second phase. For now, instead of being trapped in the smooth fasting cycle on the lower sheet, the victim finds herself trapped in the hysteresis cycle, jumping between the upper and lower sheets. The catastrophic jump from fasting to gorging occurs when she "lets go" : in the victim's own language, she watches helplessly as the apparent "monster inside herself" takes over, and devours food for several hours. Some victims vomit and gorge again, repeatedly. The catastrophic jump back occurs when exhaustion, disgust and humiliation sweep over her, and she returns to fasting for a day or several days. Some anorexics refer to this as the "knock-out". At each of the two catastrophes the limbic brain jumps from one set of states to the other, denying the victim access to the normal states between. Some anorexics even ritualise the catastrophes. The hysteresis cycle can be much longer than the previous cycle, because the after-effects of the gorge tend to prolong the fasting period.

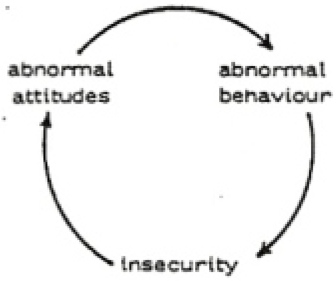
Figures 19(iii) and (iv) show how the different cycles fit onto the sections of the surface that were illustrated in Figures 16(vii) and (viii). Notice that we have labelled the fasting period of the hysteresis cycle as "purging"; this is because it occurs at a different value of x to the "pure" fasting of the first phase. Indeed the two limbic states underlie quite different frames of mind; fasting is cerebally dominated, not allowing food to enter, while purging has the somatic element of getting rid of bodily contagion.

It is not known what proportion of anorexics switch into the second phase. Sometimes the switch occurs after a hospital treatment with drugs that are used to persuade the starving first phase anorexic to eat. If the effect of such drugs is to reduce cerebal inputs to the limbic brain in

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

favour of somatic inputs, which is consistent with observed side-effects, then the drugs would be reinforcing the bias and therefore causing the switch. Thus the long-term harm caused by the use of such drugs may be greater than the short-term benefits.

Now we come to the cure. The strategic problem is how to persuade the anorexic to relinquish her abnormal attitudes, but this cannot be done directly. Therefore the practical problem is how to break the vicious circle :



The idea is to break in at the behavioural corner, by creating a third abnormal behaviour mode, during which the insecurity can be treated with reassurance. We will later show how this in turn causes a catastrophic collapse of the abnormal attitudes.

The new behaviour mode must lie between the abnormal extremes if it is going to provide a context within which reassurance can be effective. Therefore the butterfly catastrophe in Figure 20 tells us the geometric relationship that this mode must have (i.e. the dynamic relationship that the underlying brain states must have) in relationship to the existing modes.

Meanwhile Figure 18 shows that we must look for such a mode in the twilight zone between waking and sleeping for the following reasons. In the huge dynamical system modelling the limbic states of a healthy person, sleep is an attractor (i.e. a stable oscillation) with a stable boundary to its basin of attraction, separating it from wakefulness. In Figure 18 we have symbolically indicated the boundary by a dotted line. In the anorexic the boundary becomes fuzzy because the basin is being

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

shifted, as indicated by the arrows; periferally the basin is being eroded by the increasing stability of the abnormal extremes, while in between it is being enlarged by the decreasing stability of the balanced states. These changes cause the sleeping patterns to be disturbed : sleep is fragmented, shifted around and edges of the fragments become fuzzy; the anorexic goes to bed late, wakes at night, sleeps little, find herself lounging about in her night clothes. Moreover, for exactly the same mathematical reason that temporary lakes sometimes appear on the boundaries of river basins near the nodal points in between erosion and growth, so fragile attractors may appear at the boundary of the sleep basin, particularly near the two nodal points marked in Figure 18. Therefore the anorexic finds herself spontaneously falling into fragile trance-like states, in the twilight zone between waking and sleeping, between dreaming and perceiving. At the somatic node these trance-like states are filled with thoughts about food, and lists of food, while at the cerebal node they are shot through with schemes and plans how to get through the day, how to manage social occasions and avoid set mealtimes, their preparation and aftermaths, shopping, cooking and washing up.

It is these confused trance-like states that are utilised by the therapist; therapy builds upon naturally occuring processes. Hevesi's treatment consists of about 20 sessions of trance-therapy over a period of 6 to 8 weeks, each lasting 2 to 3 hours. When the sufferer asks for help, the therapist begins by pushing aside the inconclusive and confusing contents of these states, pushing them away in their respective directions so as to create a new more balanced trance. Because of the state of the sufferer quite casual remarks can carry the force of suggestions, and thus the operative suggestions are actually made quite marginally, almost incidentally. Firstly a casual but firm announcement is made at the beginning (and adhered to throughout the treatment) such as "I don't care what you eat - we are not going to talk about eating or food", because this reduces the somatic input. Secondly after the formal step of going into the trance, a suggestion is made such as "Let you mind drift - don't think - look", because this reduces the cerebal input.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Thus the patient's mind is cleared of both food and scheming, and is free to look at itself. By contrast when she is fasting she is looking all the time at the outer world with anxiety, and when she is gorging she is overwhelmed by this same world, but during trance she is cut-off and isolated. By suspending the threats, the rules, the resistance and the hunger the trance gives temporary freedom from anxiety. She is able to look at the products of her own mind, and contemplate its images and memories. In this state she is open to reassurance, and, more importantly, able to work out her own reassurance.

The more the patient practises trance, the easier it becomes; reinforcement causes an increase in stability of the new attractor, and an enlargement of its basin of attraction. The trance states begin to emerge as the new middle sheet of Figure 20(i). Therefore we introduce the last control factor, d, as reassurance, measured by time under trance.

Summarising the four control factors :

a : normal factor : hunger.

b : splitting factor : abnormality, (measurement discussed below).

c : bias factor : loss of self-control, measured by loss of weight.

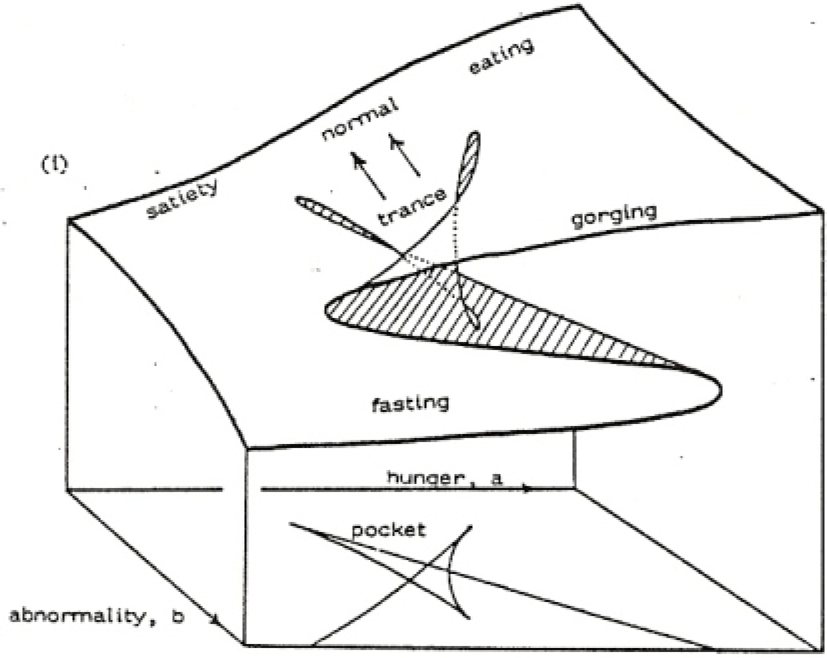
d : butterfly factor : reassurance, measured by time under trance.

Going into trance is a catastrophic jump from the lower sheet (because therapy usually takes place during the fasting part of the cycle) onto the middle sheet. Therefore the patient tends to fall into trance. What causes this jump? In fact the jump has two components, a relatively small one in the x-direction, and a larger one in the y-direction towards sleep, which is the second behavioural variable of Figure 18 that we have omitted from Figure 20 for simplicity. And it is not caused by a reduction in the abnormality, b, but by an increase in drowsiness, which is the fifth control factor, again omitted for the same reason; this is the only point where the simplicifation has caused a slight geometrical inaccuracy in our pictures.

Coming out of the trance is another catastrophe, and causes the reverse jump back onto the lower or upper sheet, depending upon whether the left or right side of the pocket is crossed, as shown in Figure 20(ii). The patients confirm that when they awake from the first few trance

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

(i)



(ii)

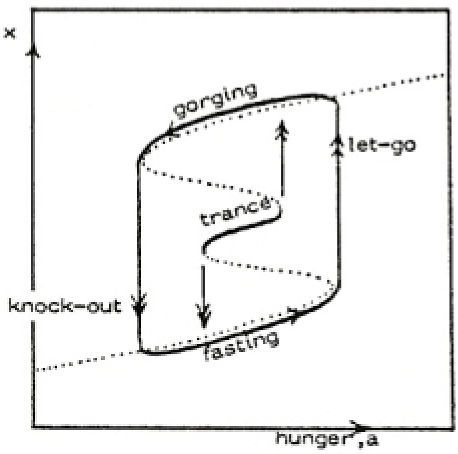


Figure 20. The effect of the butterfly factor, d>0.

(i) When therapy starts the trance states appear as a new triangular sheet of stable behaviour over the pocket (see Figure 16(v)) in between the upper and lower sheets. The new sheet opens up a pathway back to normality.

(ii) The trance states sit inside the hysteresis cycle; initially they are fragile, and coming out of trance is a catastrophic jump into either a fasting or a gorging frame of mind. (See Figure 16(ix)).

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

sessions they find themselves sometimes in a fasting and sometimes in a gorging frame of mind.

We now come to what happens during the trance. As the therapy progresses Hevesi's patients report that they experience three phenomena, of which the third is observable from the outside. Of course what the mind sees in trance it has put there, and interpreted in its own fashion, even though the images will naturally be made according to past experience, and the feelings will be such as are stored up from the past. The experience in trance may be compared to the steps in which an actor approaches a role. The first step is to envisage the part in a few simple strokes or characteristics; the second is to hear the lines the character is allotted to speak (say in a first reading through of the script); the third is to get into the part and play it, to act in front of an audience.

The first phenomenon is an experience of herself as a double personality; one personality is usually described as the "real self" and the other is called various names by different patients such as "the little one, the imp, the demon, the powers, the spirit, the voice" or merely "it". Possibly the suggestion by the therapist to look rather than to think may prepare the way for the appearance of "persons", but usually the latter appear by themselves, and we shall argue below that the patient is in fact giving a logical description of herself. It is the voice, or however she describes it, who is apparently issuing the prohibitions over food : "The little one says I musn't eat". Typically the first appearance may occur about the third session : "I've got a voice", and then perhaps a couple of sessions later "This is the first time the voice has spoken in public".

The second phenomenon is an apparent transfer of important messages between the two personalities, such as the real self promising to "pay attention" to the little one, reassuring the little one that she "will not be forgotten", while the latter in return agrees to relax the prohibitions. Sometimes the little one is symbolically given a gift, such as a teddy-bear that she once longed for and never got.

The third phenomenon is a "reconciliation" or "union" or "fusion" of the two personalities, a "welcome possession" as opposed to the earlier malignant possession. Typically "She is coming out", or "She

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

is very near the front", and then "The voice just seems to be part of myself". This third phenomenon is accompanied by a manifestation that can be witnessed by the therapist, such as speaking with a strange voice, and usually happens after about two weeks in around the seventh session. (depending of course upon the individual). When the patient awakens from this particular trance, she discovers that she has regained access to normal states, and is able to eat again without fear of gorging; she speaks of this moment as a "rebirth". Therefore during this trance the cure has taken place, a catastrophic drop in the abnormality, b, which we shall explain in a moment. Thus the trance states have opened up a path in the dynamics of the brain back to normality, indicated by the arrows in Figure 20(l). Subsequent trance sessions re-enact the reconcilitation in order to reinforce normal states and buffer them against the stresses of everyday life. At the same time the trance technique is itself reinforced, so as to provide a reliable method of self-cure, should the patient ever need to use it again at a later date.

Having dealt with the behavioural point of view, we now turn to the heart of the problem : What causes anorexia? Why can most slimmers diet without becoming anorexic? Why is there such a slow insidious apparently irreversible escalation of the disorder? How can we measure the abnormality, b? Why is the resulting neurosis/psychosis so rigid? How can there possibly be such a dramatically sudden cure?

We shall add one more cusp catastrophe to the model that will answer all these questions except the first. The reason that it cannot answer the first is that the model refers to what can be observed, whereas the original causes are probably hidden much earlier in childhood. We can offer an analogy, which may give some insight, but is not strictly part of the model. The metaphor is to describe anorexia as an "allergy to food"; of course it is <u>not</u> an allergy, but it does show some qualities similar to those of an immune system being set-up, switched-on, and inducing exposure-sensitivity. The origin of anorexia may occur in early childhood, when, perhaps for want of love or due to the inability to obtain the attention that it needs, the child retires into its shell; in other words the

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

personality sets up an immunity against disappointment by turning inwards, and leaving the shell to act out the game of life. This immunity works well enough until the shell begins to grow and get out of hand, when the encapsulated core of the personality finds that it can no longer manage. It is then that the anorexia is switched-on, instinctively identifying food as the cause of growth. This may be why anorexia so often begins at the onset of puberty, or after a period of obesity. From now on the victim is exposure-sensitive to food, and being presented with food raises deep-seated anxieties. The logical reaction is to avoid stressful situations, and so the core begins to issue prohibitions to the shell concerning food. Consequently the victim begins to feel an urge to avoid food, which she cannot explain; when she attempts to explain it she tries to capture in words some quality of the urge : e.g. "the little one" is a recognition of its origins in childhood, "the imp" describes its bad quality, "the voice" its unidentifiableness. Usually such attempts are met with disbelief, and she soon stops trying to explain.

Our metaphor breaks down when the anorexia begins to escalate. This can be observed, and so can be put into the model, as follows. Increasing insecurity is observed, associated with a gradual escalation of abnormalities over food. A typical escalation might include the following stages, but of course each anorexic will differ in the details of her own particular escalation.

escalating stages of abnormality over food

- tummy-aches at school
- give up carbohydrates
- give up cooked meat
- elaboration of diet (weight watching, calorie counting)
- excessive diet (e.g. only cheese, peanuts, black tea)
- excessive activity
- deception (e.g. secret purging after each meal)
- manipulation (of rest of family)
- nocturnal eating only
- b ↓ "I don't need to eat"

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

The important observations for our purpose are (i) there is an escalation of stages; (ii) at each stage a bimodal attitude is possible, normal or abnormal; (iii) when the anorexic reaches each stage, she will already have adopted abnormal attitudes towards all the previous stages, but as yet maintains normal attitudes towards the subsequent stages; (iv) increasing insecurity is associated with the escalation. Interpreting these facts geometrically gives the graph in Figure 21(i) showing the normality of attitude as a function of insecurity level, i, and abnormality stage, b. Abnormal attitudes begin at stage b_0, when insecurity has reached level, i_0. By the time insecurity has reached level i_1 the anorexic will have adopted abnormal attitudes towards stages up to b_1, but will so far have maintained normal attitudes towards stages beyond b_1. Thus b_1 measures the level of abnormality. Then, as the insecurity increases, so does the abnormality, following the curve in the horizontal plane, confirming that the onset of anorexia is a continual escalation by a succession of little catastrophes, little changes of attitude. Moreover as the disorder deepens, the individual catastrophes become bigger, and the attitudes towards earlier stages more abnormal (carbohydrates are at first avoided, and later feared).

We now appeal to the main theorem. The stability of memory and habit implies the existence of an implicit dynamic that holds the attitudes stably on the graph. The existence of a dynamic allows us to deduce that the graph is part of a cusp catastrophe. The right branch of the cusp marks the points where the stability of normal attitudes breaks down, and the attitude switches to abnormal, causing the gradual escalation of the disorder, while the left branch marks the points where the stability of abnormal attitudes breaks down, and the attitude switches back to normal. Thus the left branch (whose existence is a consequence of the theorem) predicts the possibility of a complete cure.

The left branch also explains why the disorder is rigid and seemingly irreversible, as follows. Suppose that after reaching i_1 the insecurity level drops again. Then the abnormality will not drop, but will stay fixed at b_1 until the insecurity has dropped to i_2 (where i_2 is given by the intersection of the line $b = b_1$ with the left branch), because all the abnormal

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

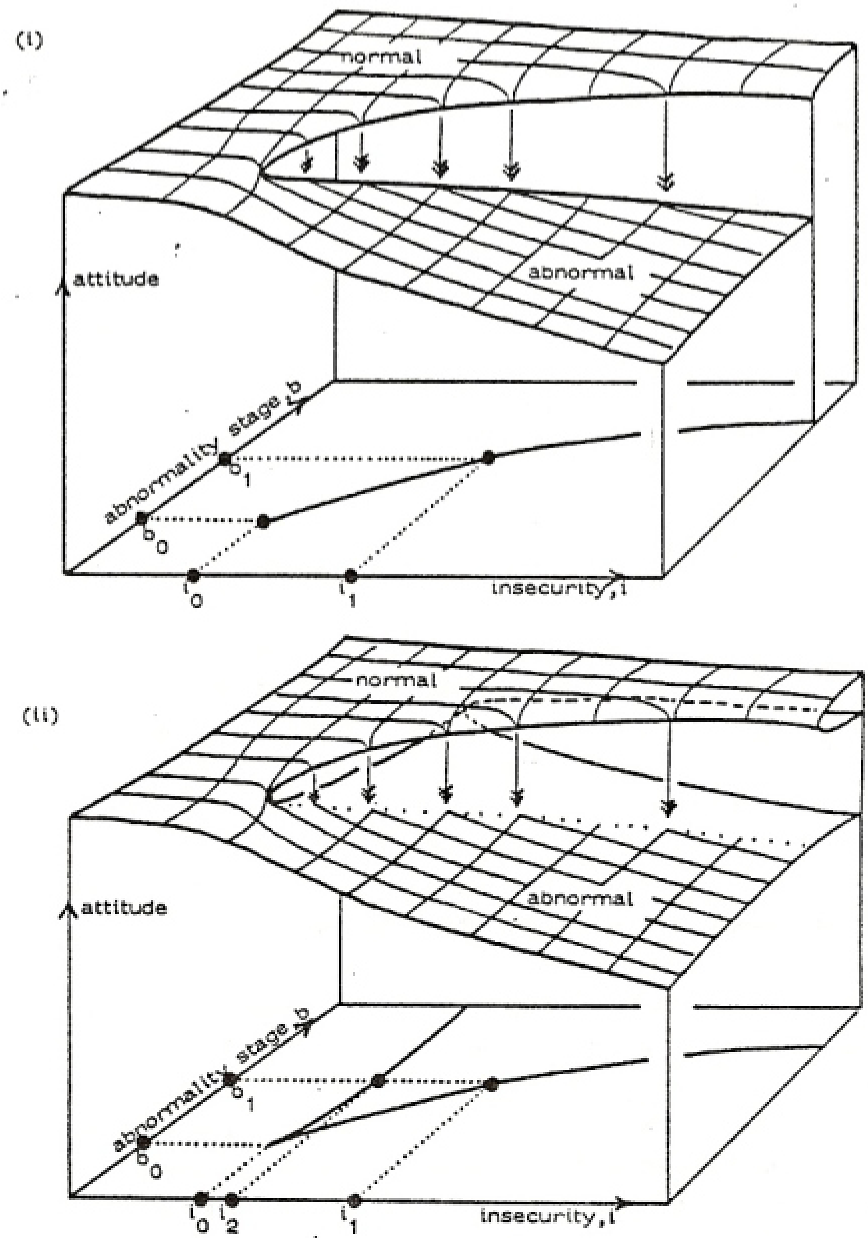


Figure 21. Abnormal anorexic attitudes.

(i) The graph showing the escalation of anorexia by a succession of little individual catastrophes, as abnormal attitudes are adopted towards each stage.

(ii) The graph embedded in a cusp-catastrophe. The left branch of the cusp indicates how far the insecurity must be reduced in order to effect a cure.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

attitudes will be held stably on the lower sheet, up to the boundary of that sheet. Then, as the insecurity drops from i_2 to i_0, there is a sudden rush of attitudes switching back to normal, along the left branch of the cusp.

Notice that the worse the anorexia is, the more rigid and irreversible it is, because as b_1 increases, so does the length of the interval $i_1 - i_2$, and hence the greater the reduction in insecurity that must be achieved before any improvement can take place. The model also explains why reasoning with the victim about her eating habits may be worse then useless, because it can only reinforce her insecurity and cannot change her attitudes; what is needed is the more fundamental reassurance about the source of insecurity. But the anorexic is not open to such reassurance while she is obsessed with, and transparently aware of, her abnormal behaviour. Hence the utilisation of the trance states, in order to give a temporary freedom from that obsession and awareness. Figure 22, which is deduced from Figure 21, shows how the therapist under these conditions can, by gently reducing the insecurity, trigger a dramatically swift catastrophic cure. Figure 22 also illustrates the difference between slimming and anorexia, showing how a quantitative difference in the initial insecurity can lead to a qualitative difference in the eventual outcome that will enable the slimmer to achieve her slimness without danger, but prevents the anorexic from escaping from her prison without help.

We have used the word "cure" in the sense of the fundamental change of attitude to life, referred to in the very revealing testimony of Mrs. Hartley quoted at the beginning; the actual physical recovery from the accompanying malnutrition and amenorrhoea will then follow naturally over the next few months. It is doubtful if this type of cure could be achieved while administering drugs that disrupt cerebal activity, because the recapturing of the whole delicate network of normal attitudes must depend not only upon reassurance, but also upon harnessing the full power of the cerebal faculties rather than suppressing them.

One of the most interesting points made by Paul MacLean is that the limbic brain is non-verbal, being phylogenetically equivalent to the brain

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

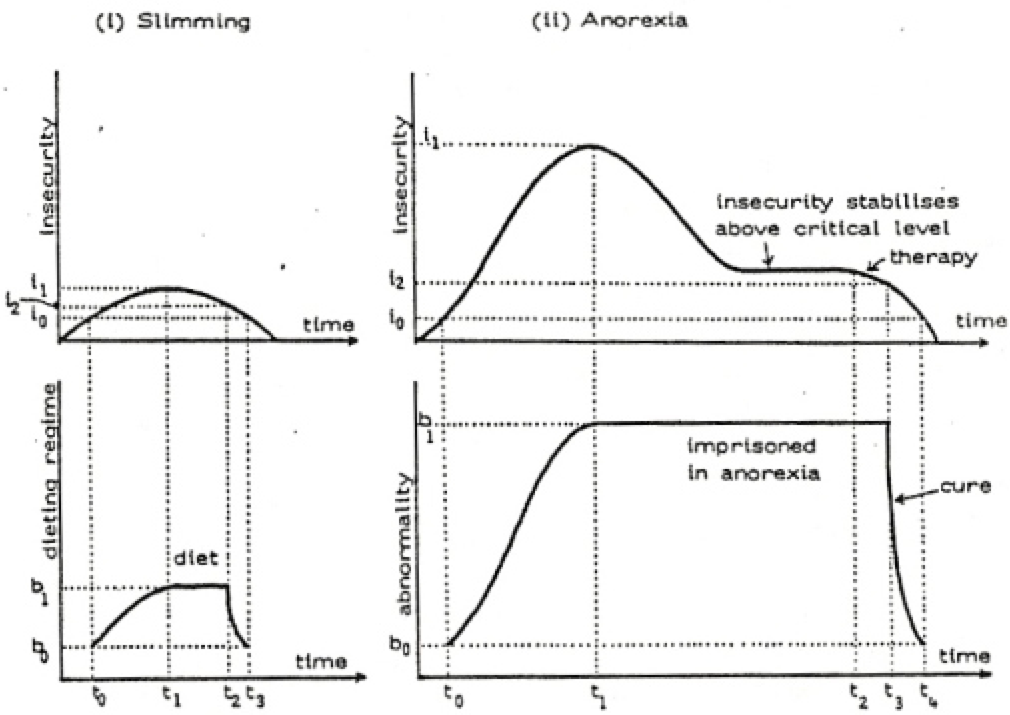


Figure 22. Comparison between slimming and anorexia.

(i) *The slimmer, anxious about her size, reaches insecurity threshold i_0 at time t_0, and therefore begins to diet with regime b_0; reaches maximum insecurity i_1 at time t_1, and therefore stabilises dieting regime at strictness level b_1; finds, as dieting succeeds, that insecurity drops to the critical level i_2 by time t_2, and therefore rapidly relaxes her dieting regime; finds that insecurity drops to threshold i_0 by t_3, and therefore gives up dieting.*

(ii) *The anorexic begins the same, except that due to deep-seated anxieties reaches a much higher maximum insecurity i_1, causing her abnormality to develop and stabilise at level b_1; is prevented from reducing her insecurity to the critical level i_2 by the feedback from the abnormal behaviour forced on her by the anorexia, and therefore remains locked in the disorder; begins therapy at time t_2, which, by reassurance during trance, reduces the insecurity to the critical level i_2 by t_3, thereby effecting the catastrophic cure by t_4.*

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

of a lower mammal. Therefore the problem of describing its activity in ordinary language is like trying to describe the conversation of a horse; no wonder anorexics have difficulty in explaining their symptoms. The patient can perceive that certain subsets of states are connected, and have boundaries, and so for her the most logical approach is to identify those

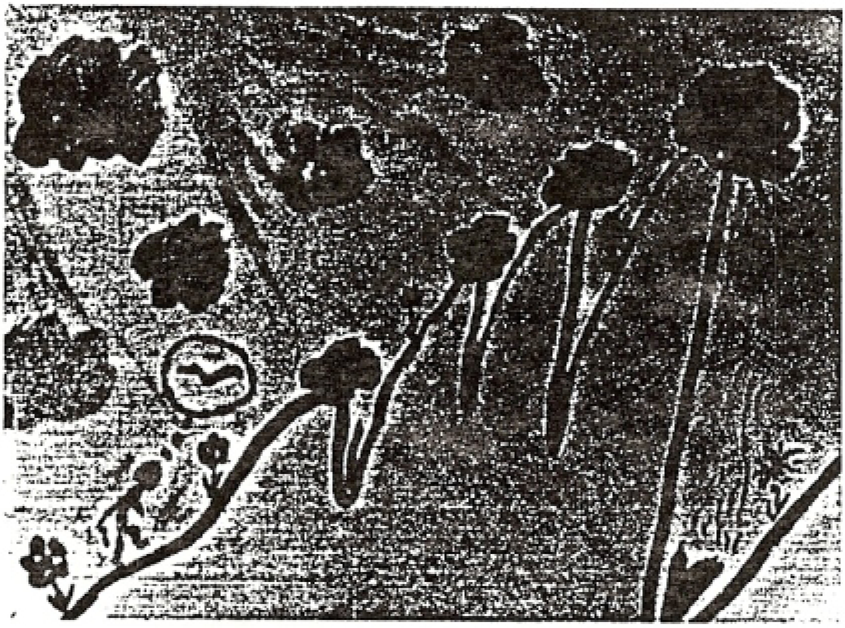


Figure 23. A painting by an anorexic of the tasks she saw ahead of her in life (reproduced with permission of A.H. Crisp).

subsets as "dissociated subpersonalities", and give them names. In the model these subsets are represented by the different sheets, and the structural relation between them is defined by the unique geometry of the catastrophe surfaces. Therefore we may identify those sheets with the patient's descriptions of her subpersonalities. For example the upper sheet of Figure 17 is often called the "monster within", and the lower sheet the "thin beautiful self". When she goes into trance the reduction of sensory input causes a shift in focus, from the close-up to the long-distance, from the immediacy of mood and behaviour to the long-term perspective of personality and insight. In terms of the model there is a shift from the perception of the states represented by the sheets of

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Figure 17 to those of Figure 21. Therefore the "monster" and "thin self" recede in importance, and are replaced by the "real self" and the "little one" corresponding to the normal and abnormal sheets of Figure 21. More precisely it is the dynamic holding the attitudes stably on the abnormal sheet that is dimly perceived and interpreted as "prohibitions" by the voice or as "malignant possession" by the little one. The "reconciliation" refers to the left branch of the cusp, which marks the boundary of the abnormal sheet, where the stability breaks down and the catastrophic cure takes place. Thus the apparent nonsense spoken by some patients makes perfectly good sense within the framework of a complete model.

Finally comes the question of how the model can be tested scientifically. It already satisfies Thom's criterion for science, because by its coherent synthesis it reduces the arbitrariness of description. Furthermore it has survived a number of qualitative experiments between myself and Hevesi of the following nature. From the mathematics I would make some prediction, or get depressed about some failing of the model, and then when we next met Hevesi could confirm the prediction, or confirm that what I had thought to be a failing was in fact another correct prediction. Let me give some examples. The mathematics predicted the location of the trance state as the middle sheet of the butterfly. However at one stage I thought the model had failed because bias destroys the middle sheet, as can be seen from Figures 16(iv) and (vi), meaning that for those patients the trance was not accessible, but to my surprise Hevesi revealed that he found very confirmed fasters or very confirmed bingers more difficult to cure. Another prediction of the mathematics was the qualitative difference between the "fasting" and "purging" frames of mind, illustrated in Figure 19(iv); the correctness of this prediction concerning the operation of the bias factor gave further evidence in favour of using the butterfly catastrophe.

Perhaps our most striking experiment concerned the finding of the operative suggestions. Hevesi says that the trance is not like hypnosis, because the therapist does not attempt to control the patient. I was curious to know what he actually did during the trance, but when I asked him he maintained that he did not do very much. Meanwhile the mathematics

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

was insisting that we ought to look at the underlying neurology as well as the psychology, even if only implicitly, in order to locate the dynamic; consequently we formulated the conjecture about inputs to the limbic brain in terms of MacLean's theories. It was only then, after watching himself with new eyes, that Hevesi was able to lay his finger on the operative suggestions that were reducing those inputs. Thus the model facilitated the communication of the therapeutic technique.

To test the model quantitatively would require monitoring a patient in different states; the prediction would be that the psychological data would give catastrophe surfaces diffeomorphic to those extracted from the accompanying neurological and physiological data, with the same bifurcation set. Different patients would have diffeomorphic bifurcation sets, or parts of them diffeomorphic.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

PART THREE

ILLUSTRATIONS IN PHYSICS.

Example 10. Liquid/gas phase transition.

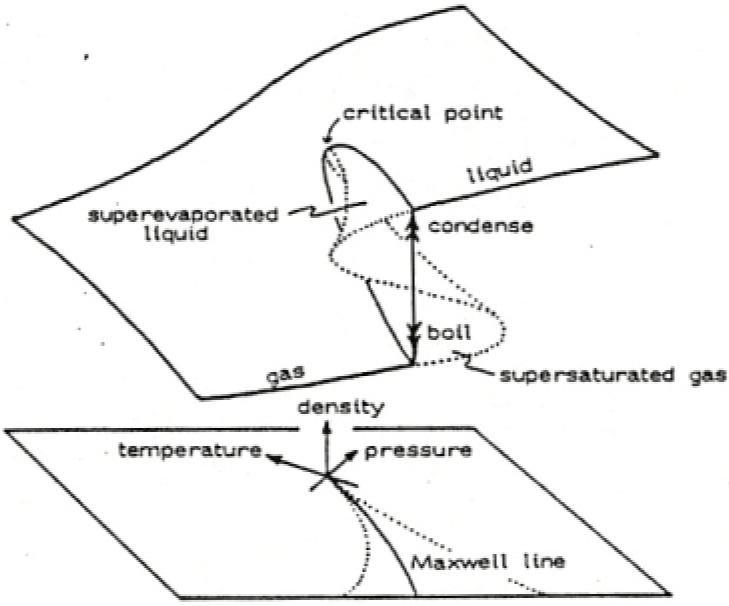


Figure 24. Temperature and pressure are conflicting factors controlling density.

Van der Waals' equation for phase transition is a cusp catastrophe surface, with temperature and pressure as conflicting factors controlling density. The upper and lower sheets represent the two phases of liquid and gas, while the two catastrophes represent the transitions of boiling and condensation. By going round the top of the cusp one can go from liquid to gas continuously. Since the controls are averaging devices, both catastrophes normally occur at the same temperature, on a line in the middle of the cusp given by Maxwell's convention. However both catastrophes can be delayed in the metastable states of superevaporated liquid or supersaturated gas. For instance clean water at atmospheric

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

pressure can be gently heated beyond 200^{0} before boiling, and when it does boil it explodes catastrophically with a sound like a pistol shot. The metastable states are exploited by high energy physicists in the bubble and cloud chambers. Since density is an averaging device, the shape of the surface very near the critical point is slightly distorted.

Example 11. Light caustics.

Light caustics are the bright geometric patterns created by reflected or refracted light. A familiar example is the cusp appearing on a cup of coffee in bright sunlight, caused by reflection of the sun's rays off the inside of the cup, as in Figure 25(i).

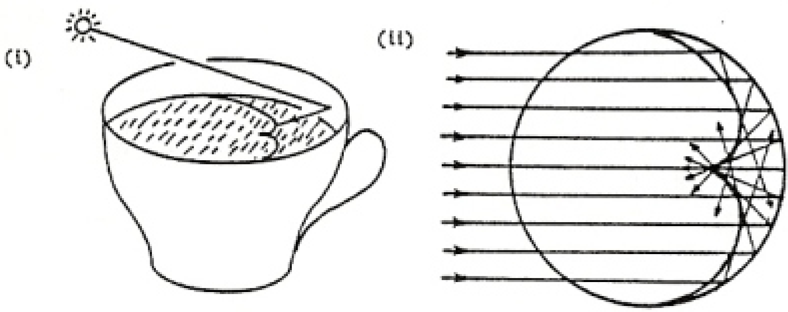


Figure 25. The caustic on a cup of coffee, caused by the sun's rays reflected off the inside of the cup, is a cusp-catastrophe.

Looking down on the cup from above, Figure 25(ii) shows that the vertical planes containing the sun's rays, after they have been reflected, envelop a vertical surface with cusp-shaped horizontal cross-section, which is called the caustic. Since the planes all touch the caustic there is a concentration of photons near the caustic (on its convex side), but of course these photons are invisible to the eye, because each is travelling along its appointed route. We can only see them if we place a screen in the way - in this case the screen is the surface of the coffee; then the concentration of photons hitting the screen will be scattered into our eye, causing the cusp-shaped section of the caustic to appear bright, with a sharp edge on the concave

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

side and soft edge on the convex side.

Another familiar caustic is the rainbow. Here each water droplet refracts and reflects the suns rays to form a caustic surface approximately cone-shaped, with radial angle between 40° and $42\frac{1}{2}^{\circ}$. The angle depends upon the wave length, so that in fact each droplet produces a co-axial family of differently coloured caustic cones. Those drops whose caustics happen to meet our eye produce the rainbow.

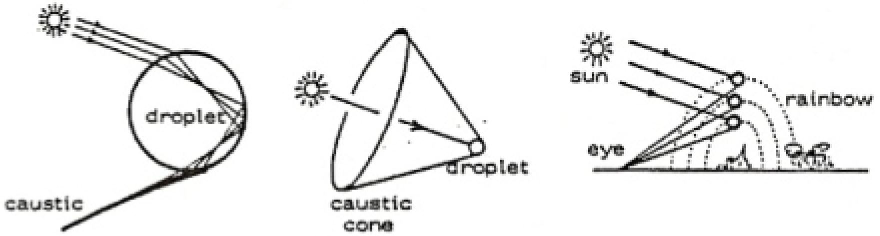


Figure 26. The rainbow is caused by a spectrum of coloured caustic cones produced by each droplet, each cone being a fold-catastrophe.

Catastrophe theory applies to light caustics because light obeys a variational principle : by Fermat's principle the light rays travel along geodesics. Let C be a 3-dimensional neighbourhood of a caustic, and let X be a 2-dimensional surface normal to the incident rays. Each point x in X determines a ray C_x in C, and the union of all these rays, parametrised by X, forms a 3-dimensional manifold.

$$M = \{(c,x);\ c \in C_x\} \subset C \times X.$$

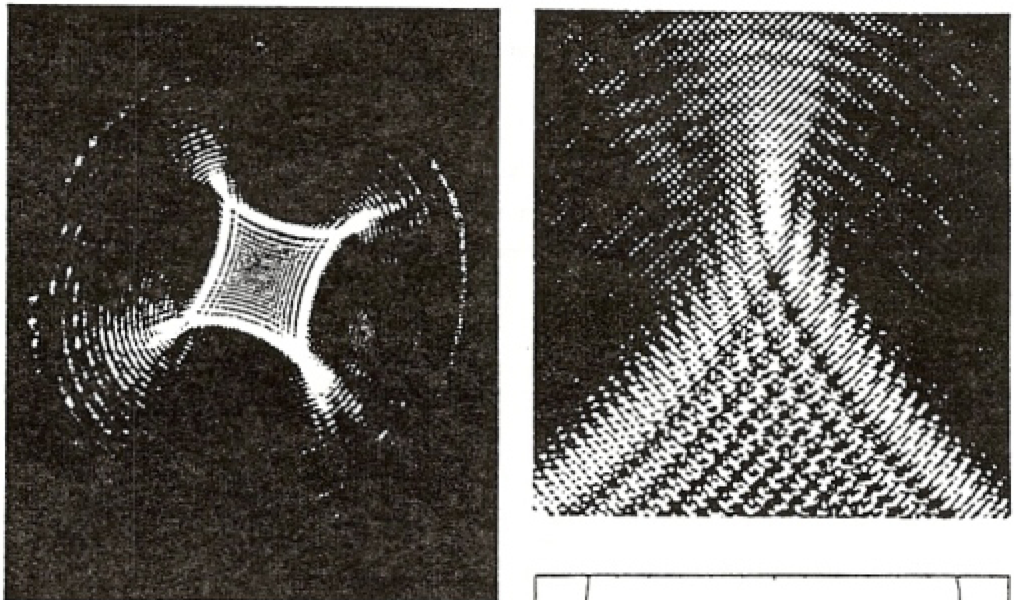
Since all the rays leave X normally, M is given by $\partial f/\partial x = 0$, where $f(c,x)$ is the geodesic distance from x to c. The intensity of light in a volume element dC of C is proportional to dM/dC, where dM is the volume of the inverse image of dC under the projection $M \to C$. The caustic is where this intensity is greatest, namely on the bifurcation set. Therefore stable caustics are stable bifurcation sets, in other words elementary catastrophes. Therefore the classification theorem gives the new result in geometric optics : the only stable singularities a caustic

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Figure 27. The caustics obtained by reflecting a point source of light in a parabolic mirror on to a screen. The three pictures correspond to three different positions of the screen, and illustrate three sections of the bifurcation set of a hyperbolic umbilic catastrophe. The catastrophe point is the top-left of the middle section. The top section shows a cusp inside a smooth fold curve; as the sections pass through the catastrophe point, the cusp pierces the curve and is itself transformed into a smooth curve on the outside, while the curve is transformed into a cusp on the inside (see Figure 13(b)). The faint repeated image is an artifact, due to a subsidiary reflection from the front surface of the mirror. Photograph by Warwick University Library Photographic Services.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

(a) (b)



(c)

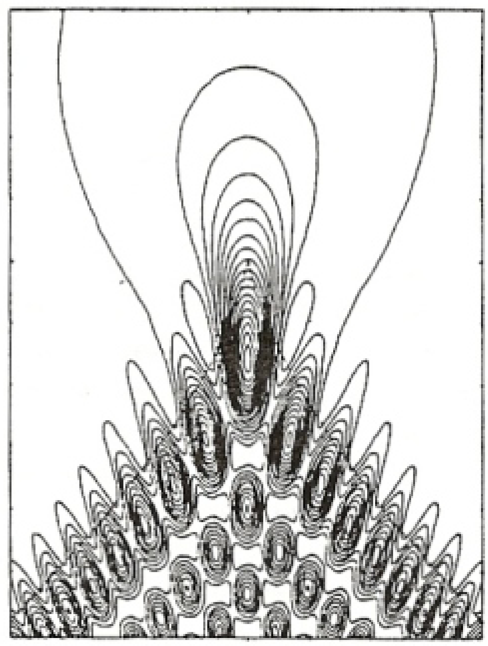


Figure 28. (a) Caustics obtained by refracting a laser beam in doubly periodic frosted glass. The interference patterns merge into Airy patterns on the fold lines of the caustic, and into Pearcey patterns at the cusp points. (b) Detail near a cusp, showing quantization of diffraction spots. (c) Computer calculation of intensity levels of Pearcey pattern. Photographs by Michael Berry.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

can have, besides cusped edges, are the three types of singular point shown in Figure 13, the swallowtail and the elliptic and hyperbolic umbilics. This discovery of Thom's about caustics was one of the reasons that stimulated him to develop catastrophe theory.

The sections of a hyperbolic umbilic catastrophe illustrated in Figure 27 were obtained by shining a torch in an old searchlight mirror. The other two types of singular point can be obtained using cylindrical and spherical lenses, such as beakers and electric light bulbs filled with water. Different sections near a singular point can be explored by moving the screen or the light.

The photographs of Michael Berry in Figure 28 were obtained by refracting a laser beam in frosted glass. Since the frosting was periodic in two directions the global picture is the projection of a torus in the plane (each little square of frosting is projected onto the whole picture). Each corner is a section of a hyperbolic umbilic. To analyse the fine structure of interference patterns merging in Airy and Pearcey patterns on the caustic, it is necessary to pass from geometric optics the more delicate wave optics. Berry uses the frosted glass as an optical analogue to study the scattering of beams of particles from a solid surface.

<u>Example 12. Euler buckling.</u>

Figure 29 shows a horizontal elastic strut subjected to a vertical load a and a horizontal compression b. (For an easy experiment hold a 1" x 4" strip of thin cardboard between the thumb and forefinger.)

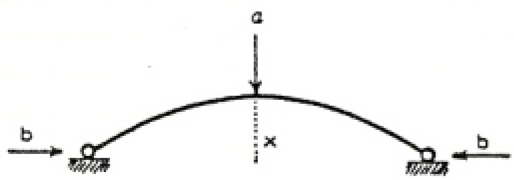


Figure 29. A free elastic strut under load a, and compression b. The resulting vertical displacement is x.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

We measure the vertical displacement of the strut by the first harmonic x (i.e. the first Fourier coefficient). Then x satisfies a cusp catastrophe with normal factor ($-a$) and splitting factor b. Buckling occurs at the

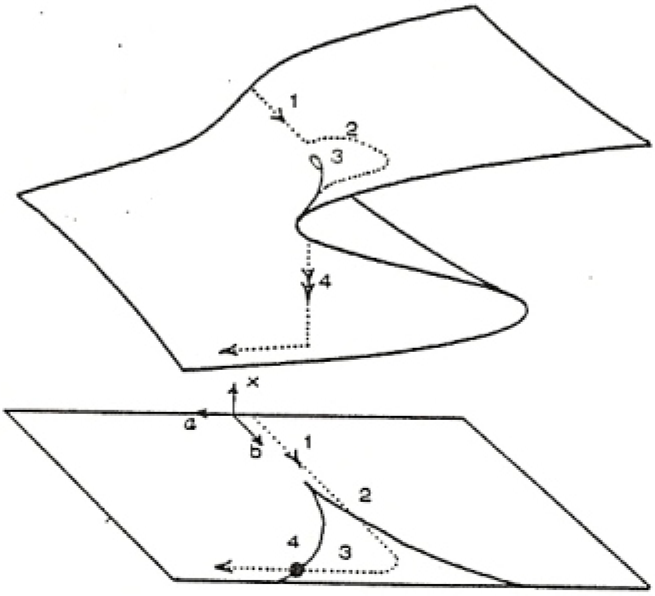


Figure 30. In the free strut the load a is the normal factor, and the compression b the splitting factor, controlling the first harmonic, x. The dotted path shows the strut (1) compressed (2) buckling (3) loaded and (4) snapping.

cusp point, which Euler showed was $a = 0$, $b = \pi^2\lambda/\ell^2$, where ℓ = length and λ = modulus of elasticity. The dotted path in Figure 30 shows the strut (1) remaining straight under increasing compression (2) buckling upwards (3) supporting an increasing load until (4) it snaps downwards.

Now suppose that the ends of the strut are fixed, so that in an unloaded state it is buckled at height x. Put on the load a, offset from the centre by a distance c (to simulate a manufacturing imperfection) and measure the displacement now by the second harmonic y (see Figure 31). The second harmonic can be seen easily with the strip of cardboard, or more dramatically in the western arch of Clare College bridge, Cambridge.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

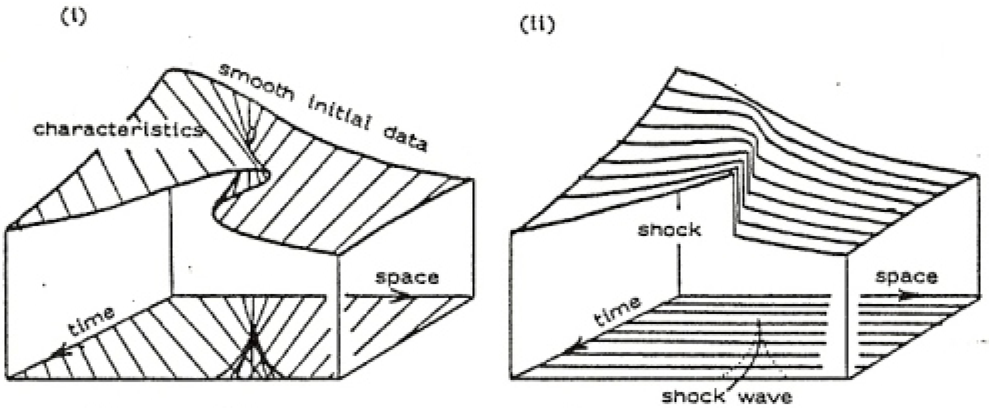


Figure 33. *(i) Propagation of the initial data along the characteristics of the partial differential equation gives a cusp catastrophe. (ii) Physically only one solution is possible, and so a shock wave has to develop, whose speed is determined by the conservation law.*

Example 14. Forced non-linear oscillations.

Forced oscillations can be modelled by Duffing's equation

$$\ddot{x} + k\dot{x} + x + \alpha x^3 = F \cos \Omega t,$$

where $k > 0$ is a small damping term, α a small non-linear term ($\alpha = -\frac{1}{6}$ for a simple pendulum), and $F \cos \Omega t$ is a small periodic forcing term with frequency Ω close to 1, the frequency of the linear oscillator. The amplitude A of the resulting oscillation depends upon the parameters, and Figure 34 shows the graph of A as a function of α and Ω (keeping k and F fixed). There are two cusp-catastrophes with α, Ω as conflicting factors. At each cusp the upper and lower sheets represent attractors (stable periodic solutions) while the middle sheet represents saddles (unstable periodic solutions). If the frequency of the forcing term is gradually changed so as to cross one of the cusp lines, going from the inside to the outside of the cusp, then the amplitude A will exhibit a catastrophic jump. There will also be a sudden phase-shift at the same time.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

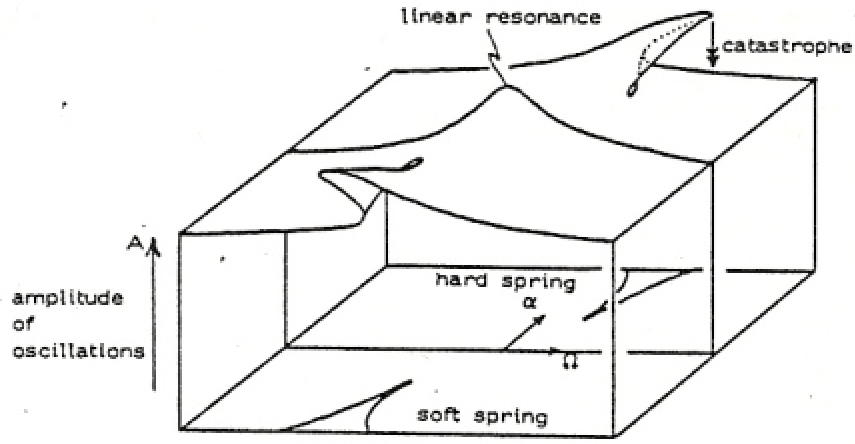


Figure 34. The oscillation of a forced non-linear oscillator bifurcates according to the cusp-catastrophe.

SUGGESTED FURTHER READING.

Books of Thom.

R.Thom, *Structural stability and morphogenesis*, Benjamin, New York, French 1972; English trans. by D.H. Fowler, 1975.

R.Thom, *Modèles mathématiques de la morphogénèse*, Union Générale d'Editions, Paris, 1974.

Introductory mathematics.

V.I.Arnol'd, Critical point of smooth functions, *Proc. Int. Cong. Math., Vancouver*, 1, (1974), 19-39.

T.Bröcker, *Differentiable germs and catastrophes*, Camb. Univ. Press, 1975.

J.Callahan, Singularities and plane maps, *Amer. Math. Monthly* 81 (1974) 211-240.

D.R.J.Chillingworth, *Differential topology with a view to applications*, Pitman, London, 1976.

P.J.Holmes & D.A.Rand, The bifurcations of Duffing's equation : an application of catastrophe theory, *J. Sound & Vib.* 44, 2 (1976)237-253.

Y.C.Lu, *Singularity theory and an introduction to catastrophe theory*, Springer-Verlag, New York, 1976.

T.Poston & I.N.Stewart, *Taylor expansions and catastrophes*, Pitman, London, 1976.

A.E.R.Woodcock & T.Poston, *A geometrical study of the elementary catastrophes*, Springer Lecture Notes in Maths, 373 (1974).

This book, papers 18, 19.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)

Anorexia, brain and behaviour.

H.Bruch, Anorexia nervosa, *American Handbook of Psychiatry;* Basic Books, New York, 4 (1975) 787-809.

A.H.Crisp, Deception, obstinacy and terror in anorexia nervosa, *Mims Magazine* 15 (March 1975) 42-53.

K.Lorenz, *On aggression*, Methuen, London, 1966.

P.D.MacLean, *A triune concept of the brain and behaviour, Hinks Memorial Lectures*, Toronto Univ. Press, 1973.

This book, papers 8,9,12.

Some applications in the physical sciences.

M.V.Berry, Cusped rainbows and incoherence effects in the rippling-mirror model for particle scattering from surfaces, *J. Phys. A* 8 (1975) 566-584.

M.V.Berry, Waves and Thom's theorem, *Advances in Physics* 25 (1976) 1-26.

M.V.Berry & M.R.Mackley, The six roll mill : unfolding an unstable persistently extensional flow, *Phil. Trans. Roy. Soc.* (1977).

M.V.Berry & J.F.Nye, Fine structure in caustic juncations, *Nature* 267 (1977) 34-36.

M.V.Berry, Focusing and twinkling : critical exponents from catastrophes in non-Gaussian random short waves, *Advances in Physics* (1977).

M.V.Berry &J.H.Hannay, Umbilic points on Gaussian random surfaces, *J. Phys. A* (1977).

T.B.Benjamin, Bifurcation phenomena in steady flows of a viscous fluid, *Proc. Roy. Soc. Lond. A* (1977).

C.DeWitt-Morette, The small disturbance equation, caustics and catastrophes, Univ. of Texas, Austin, Texas (1975).

J.Guckenheimer, Caustics and non-degenerate Hamiltonians, *Topology* 13 (1974) 127-133.

K.Jänich, Caustics and catastrophes, *Math. Ann.* 209 (1974) 161-180.

J.Komorowski, On Thom's idea concerning Guggenheim's one-third law in phase transition, Inst. Math. Methods in Phys., Warsaw University (1977).

P.D.Lax, The formation and decay of shock waves, *Amer. Math. Monthly* 79 (1972) 227-241.

J.C.Maxwell, Last essays, in L.Campbell & W.Garnett, *The life of James Clerk Maxwell*, MacMillan, London (1882) 434-444.

J.F.Nye, Disclinations and catastrophes in the vector and tensor fields of sea ice, Univ. Bristol (1976).

L.S.Schulman & M.Revzen, Phase transitions and catastrophes, *Collective phenomena*, 1 (1972) 43-47.

M.J.Sewell, Some mechanical examples of catastrophe theory, *Bull. Inst. Math. & Appl.* 12 (1976) 163-172.

R.Thom, Phase transitions as catastrophes, in *Statistical mechanics : new concepts, new problems, new applications*, Univ. Chicago Press (1977) 93-105.

J.M.T.Thompson & G.W.Hunt, *A general theory of elasticity*, Wiley, 1973.

J.M.T.Thompson & G.W.Hunt, Towards a unified bifurcation theory, *J. Appl. Maths. Phys.* 26 (1975) 581-603.

This book, papers 15 - 17.

Other applications.

At the end of each paper in this book is a bibliography related to that paper, and at the end of paper 21 there is a more general bibliography.

ISBN 0-201-09014-7, 0-201-09015-5 (pbk.)