Controversy in Science: on the Ideas of Daniel Bernoulli and René Thom

E. C. Zeeman*
Gresham Professor of Geometry, Hertford College
Oxford, OXI 3BW England

*Text of the 1992-1993 Johann Bernoulli Lecture given by Prof Sir E.C. Zeeman FRS at the University of Groningen on February 16, 1993. The Johann Bernoulli Foundation for Mathematics founded in Groningen in 1988 organizes each year a Johann Bernoulli lecture for which it invites prominent scientists, in particular mathematics. Johann Bernoulli was professor of mathematics at the University of Groningen from 1695-1705.

<table>
<tr><td colspan="2"></td><td colspan="2">THINGS</td></tr>
<tr><td colspan="2"></td><td>discrete</td><td>continuous</td></tr>
<tr><td rowspan="2">BEHAVIOUR</td><td>discrete</td><td>Dice
Symmetry

DISCRETE BOX

Finite probability
Finite groups</td><td>Music, harmony
Light
Discontinuities

PANDORAS BOX

Fourier series
Quantum theory
Catastrophe theory</td></tr>
<tr><td>continuous</td><td>Planets
Populations

TIME BOX

Ordinary differential equations</td><td>Waves
Elasticity

CONTINUOUS BOX

Partial differential equations</td></tr>
</table>

FIGURE 1. The four types of applied mathematics.

Finally we come to the top right box of continuous things behaving discretely: this is an unexpected box because continuous things normally behave continuously, and when they do not it tends to be anti-intuitive. Consequently there are frequently arguments about things in this box, especially if new mathematics has to be invented to explain phenomena that previously had been inexplicable. Such explanations may not be accepted at first, and may become controversial. This is why I have called it Pandora's Box: as soon as you open it and let something out it is liable to set off a controversy.

The first example is music, for musical instruments tend to be continuous things vibrating in harmonics. It is really an eigenvalue problem, and the discreteness of the harmonics underlies the whole harmonic structure of music. For example the string of a harpsichord or a violin is a continuous thing, but it vibrates in a combination of harmonics, and the mathematical modelling of vibrating strings led to a major controversy between d'Alembert, Euler and Daniel Bernoulli; it began around 1750 and lasted some 80 years before it was finally resolved by the work of Fourier and Dirichlet, as I shall explain below.

The second example is quantum theory, which was invented at the beginning of this century to explain the wave/particle nature of light, but which was staunchly opposed by Einstein all his life. The foundations of quantum theory are still controversial even today.

The third example is catastrophe theory, which is a method of modelling invented by René Thom in the 1960's, based on theorems in differential topology. It is particularly applicable to phenomena in which continuous causes give rise to discontinuous effects, so it sits firmly inside Pandora's Box and when it was first let out it gave rise to a controversy in the 1970's, which I shall describe below.

THE VIBRATING STRINGS CONTROVERSY

The mathematical study of musical sounds has a long history dating back to the Pythagoreans in the fifth century BC, but significant mathematical discoveries were not made until the eighteenth century. Thanks largely to the experimental work of Sauveur it was already known by 1700 that a stretched string can vibrate in many different modes corresponding to the harmonics. Today we can describe the harmonics as follows (although Sauveur did not know this). Suppose we choose the unit of length so that the string has length $\cancel{\vartheta}$. Then π

the first harmonic is the fundamental

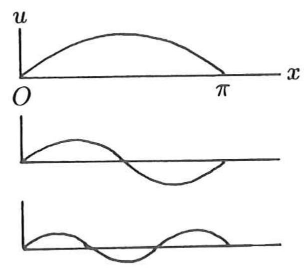
$$u = \sin x \cos ct,$$

the second harmonic is the octave above

$$u = \sin 2x \cos 2ct,$$

the third harmonic is the fifth above that

$$u = \sin 3x \cos 3ct,$$



and so on. Here $x, 0 \leq x \leq \pi$, denotes the distance along the string, u denotes the displacement (assumed small) of the point x at time t, and c denotes a constant, which is the speed at which waves can travel along the string, and which is equal to the square root of the tension in the string divided by the mass per unit length.

In 1713 BROOK TAYLOR [17] (after whom the Taylor series is named) found the shape and frequency of the fundamental. In 1727 JOHANN BERNOULLI [5] wrote to his son Daniel in St Petersburg showing how to obtain this shape as the limit, as $n \to \infty$, of n beads vibrating on a weightless elastic string. In 1733 DANIEL BERNOULLI [2, 3] began looking at the mathematics of the harmonics and their superposition.

It is important to remember that this was a period of enormously fruitful musical development. In 1722 Bach had completed the first half of his 48 preludes and fugues under the title of *The well-tempered clavier*. By this he meant that if a keyboard instrument is tuned so that the octave is divided into 12 equal semitones then, although this will have the disadvantage that every single key is slightly out of tune, nevertheless the ear can learn to live with this small amount of mistuning, and what is gained is the great advantage of being able to play in any key. This gave composers the opportunity of making all possible key changes and harmonic modulations, leading to the richness of western music as we know it today. Bach completed his 48 preludes and fugues by 1744. So harmonics were very much in the air, and Daniel Bernoulli was right at the cutting edge of modern research at the newly founded Academy in St Petersburg, a city which itself had only been founded in 1703. The stage was now set for the three major papers by d'Alembert, Euler and Bernoulli that triggered off the controversy. The history is summarised in Figure 2.

THE PRECURSORS

1700	SAUVEUR: 1653-1716	harmonics (experiment)
1713	TAYLOR [17]: 1685-1731	shape and frequency of the fundamental
1722	BACH: 1665-1750	the well-tempered clavier
1727	JOHANN BERNOULLI [5, and letter to Daniel]: 1667-1748	beads on a string
1733	DANIEL BERNOULLI [2, 3]: 1700-1782	harmonics (mathematics)

THE CONTROVERSY

1747	D'ALEMBERT [1]: 1717-1783	the wave equation&its solution in terms of two travelling waves
1748	EULER [8]: 1707-1783	non-smooth solutions
1753	DANIEL BERNOULLI [4]: 1700-1782	Fourier series solutions

THE RESOLUTION

1807	FOURIER [9]: 1768-1830	calculated the Fourier coefficients
1829	DIRICHLET [7]: 1805-1869	proved the Fourier series convergent

FIGURE 2. The vibrating strings controversy.

In 1747 D'ALEMBERT [1] discovered that the string satisfied the wave equation

$$u_{tt} = c^2 u_{xx}.$$

Then he solved it most elegantly as the average of two travelling waves, one going forwards and the other going backwards, both starting at the initial position of the string. More precisely, given initial conditions $u = f(x)$ and $u_t = 0$ then the complete solution is

$$u(x,t) = \frac{1}{2}[f(x+ct) + f(x-ct)],$$

where f is the odd 2π-periodic extension to the real line $\mathbb{R}$ of the given initial condition in $[o, \pi]$. This was the first ever really successful use of partial differential equations, and so in terms of Figure 1 d'Alembert had achieved a spectacular opening of the Continuous Box.

Meanwhile Euler was a more practical man and he knew that in the harpsichord, which was the main keyboard instrument of the day, the strings were plucked with a quill. So he suggested in a paper [8] in 1748, published a few months after d'Alembert's, that the initial position of a plucked string was unlikely to be smooth: it was more likely to be as shown in Figure 3(a), consisting of two straight segments OA and $A\pi$, with a kink at A.

He then applied, not d'Alembert's *equation*, but d'Alembert's *solution*: he deduced from the two travelling waves that the position at time t would be as in Figure 3(b), consisting of three straight segments with kinks at B and C. The outer segments OB and $C\pi$ are stationary, while the interior of the middle segment BC is moving downwards with constant speed, as shown in Figure 3(c). One complete cycle is given by the two kinks B and C running steadily down the upper edges of the parallelogram $OA\pi D$ in Figure 3(d), bouncing successively off the ends O and π, then running down the lower edges of the parallelogram until they meet at D, bouncing off each other at D, and finally reversing the whole motion back to A again.

One can just imagine d'Alembert's indignation at Euler stealing his solution of the wave equation without the slightest justification, because the function could not possibly satisfy the equation itself since it was not differentiable at A. Nor could the non-smoothness of f be localised at A because the kinks B and C travelled the entire length of the string and back again. Not only had Euler moved the problem from the beautiful Continuous Box into the awkward Pandora's Box, but he had also undercut and devalued d'Alembert's pioneering achievement of applying calculus rigorously to partial differential equations.

Of course today we can happily compute u_x as the discontinuous function in Figure 3(e), and u_{xx} as the Dirac function in Figure 3(f), zero everywhere except below B and C where it is minus infinity. Moreover this Dirac function is precisely the impulse that is needed in u_{tt} to give the points B and C an infinite acceleration in zero time, in order to suddenly change them from being stationary on OB and $C\pi$ to moving at constant speed on BC. Thanks to Schwarz's modern theory of distributions, these Dirac functions are rigorous solutions of the wave equation. Alternatively it is possible to smooth out the kinks with an arbitrarily close C^∞ approximation, which incidentally is probably a more realistic model of the string.

But in d'Alembert's day the very notion of function was in the process of evolution [see 10, 12]; d'Alembert thought of a function as given by an analytic formula, while Euler was moving towards a more general concept. It was not possible to base the notion of function on set theory as we do today because the set of real numbers had not been properly defined - indeed it was not until

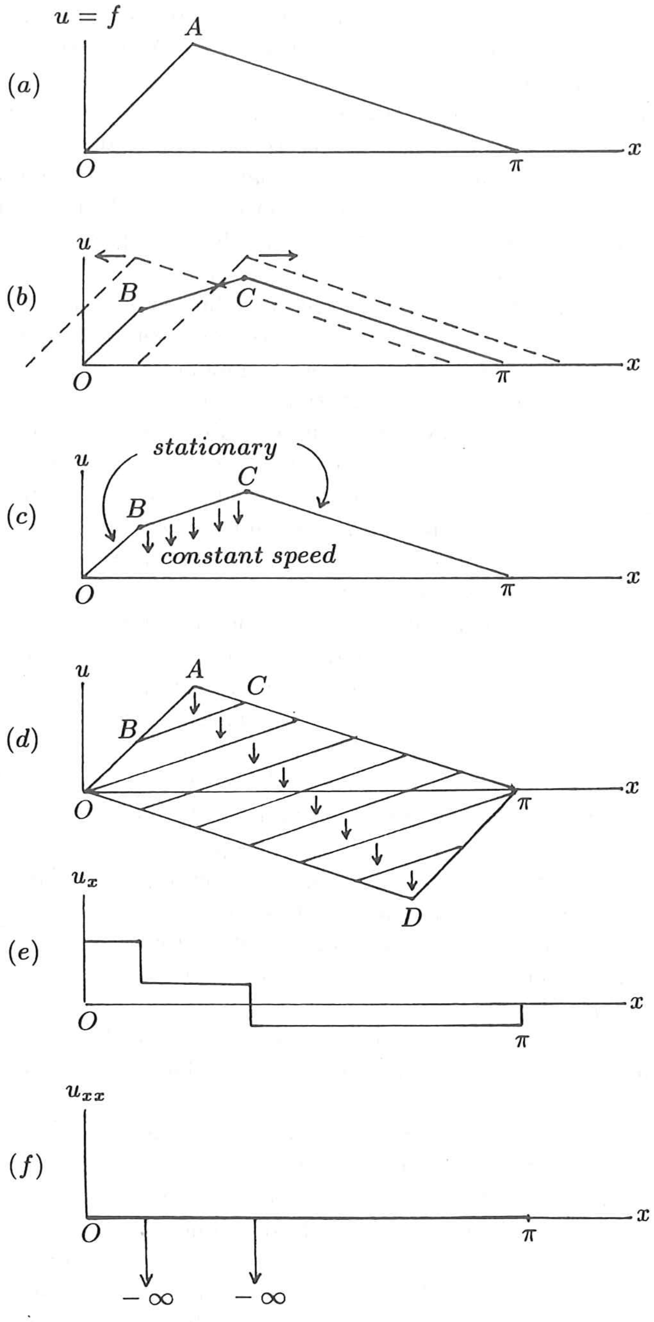

$u = f$
A
(a)
O
π
x
u
(b)
B
C
O
π
x
u
stationary
C
B
(c)
constant speed
O
π
x
u
A
C
B
(d)
O
π
x
D
u_x
(e)
O
π
x
u_{xx}
(f)
O
π
x
$-\infty$
$-\infty$

FIGURE 3.

a century later in 1854 that Dedekind invented the first rigorous definition of the reals.

Today we are in a position to appreciate both d'Alembert's intellectual honesty as a pure mathematician, and Euler's imaginative flexibility as an applied mathematician. But at the time the controversy between them was long lasting and hurtful, as is shown by the following extract in a letter from Euler to Lagrange ten years later in 1759 [12, p512]:

> *'I am delighted to learn that you approve my solution . . . which d'Alembert has tried to undermine by various cavils, and that for the sole reason that he did not get it himself. He has threatened to publish a weighty refutation; whether he really will I do not know. He thinks he can deceive the semi-learned by his eloquence. I doubt whether he is serious, unless perhaps he is thoroughly blinded by self- love.'*

Meanwhile DANIEL BERNOULLI was justifiably upset because both D'ALEMBERT and EULER had ignored his earlier work on the harmonics. In 1753 he replied [4] claiming that the general solution was merely an infinite sum of harmonics:

$$u(x,t) = \sum_{n=1}^{\infty} a_n \sin nx \cos nct,$$

where the coefficients a_n are determined by the initial condition:

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} a_n \sin nx.$$

Today we call it a Fourier series, but at that time Fourier had not yet been born.

Consider the ages of the three men involved; Bernoulli was the senior mathematician at 50, Euler was well established at 40, and d'Alembert was the young upstart at 30. What must have particularly annoyed Bernoulli was that d'Alembert had successfully hijacked the problem from Pandora's Box into the Continuous Box, therefore missing the whole point. The most amazing thing about music is the harmonic relationship between the notes, which d'Alembert had totally ignored. Of course Euler with his kinks had moved the problem back into Pandora's Box to d'Alembert's disgust, but somehow this was even worse in Bernoulli's eyes because Euler had moved it there for the wrong reason: he too had missed the main point. One can sympathise with Bernoulli's frustration because his own approach was the deepest, giving the greatest insight into the structure of music.

Both d'Alembert and Euler, meanwhile, were very pleased with themselves because they claimed they had discovered a whole wealth of new shapes and new types of vibration for the string – in fact all possible types of vibration. Bernoulli responded that, on the contrary, there was nothing new because, he claimed, his sums of harmonics already included all their new shapes and vibrations. He could not prove this claim, but here is a translation of an extract from the introduction to his 1753 paper [4] in which he emphasises the strength of his convictions by using the French word '*absolument*':

> '*We remark therefore that the string can make not only [Taylor] vibrations of the first, second and third kinds, and so on to infinity, but also any sum of these vibrations in all possible combinations. Meanwhile all these new curves and new types of vibration given by Messrs. d'Alembert and Euler are absolutely nothing more than sums of several kinds of Taylor vibrations.*'

D'Alembert and Euler replied separately (they could not reply together since they were not on speaking terms because of their own controversy) that one could not possibly represent an arbitrary continuous function by a mere sum of harmonics, because the latter sums were too simple compared with the wealth of possibilities amongst all functions. Bernoulli protested that he had an infinite number of coefficients to play with, but since he did not know how to calculate them people did not believe him.

There followed a heated debate that lasted for over 50 years, with many other people taking sides, until eventually in 1807 FOURIER [9] showed how to calculate the coefficients. At last Bernoulli's point of view was vindicated and people realised that he must have been right after all. In 1829 DIRICHLET [7] finally clinched the matter by proving that the resulting Fourier series converged to the original function.

Fourier was in fact looking at a different problem, which happened to be in the Continuous Box. He was investigating the way heat diffuses away from a bar whose ends are kept at zero temperature. What he was really thinking about was himself; Napoleon had posted him to Egypt for the Egyptian campaign, where it was much too hot, and when that collapsed had posted him to the French alps, where it was much too cold. The ends of the bar were in fact his frozen hands and feet, and as he drank hot drinks to try and keep warm he gloomily watched the heat diffusing away to his frozen extremities. He discovered the heat equation

$$u_t = c^2 u_{xx},$$

where $u(x,t)$ is the temperature at the point x at time t. Suppose that the bar has length π, and that the given initial condition is $u = f(x)$. Then he found the solution

$$u(x,t) = \sum_{n=1}^{\infty} a_n e^{-c^2 n^2 t} \sin nx.$$

Initially

$$f(x) = \sum_{n=1}^{\infty} a_n \sin nx,$$

and he showed that the Fourier coefficients must be given by

$$a_n = \frac{2}{\pi}\int_0^\pi f(x)\sin nx dx.$$

However he could not prove that the resulting Fourier series converged to f, so he suggested the problem to Dirichlet who had come to study with him as a young student of 17. In 1829 DIRICHLET [7] solved the problem for all functions having a finite number of maxima and minima, and at most a finite number of discontinuities. Dirichlet's paper is a jewel. It is rigorous and beautifully written, and could be said to be the first paper in modern analysis. In one fell swoop he had disposed of much misleading intuition and fallacious folklore amongst the mathematical community. For instance here is a list of some of the plausible but mistaken beliefs held by d'Alembert and Euler and their followers, which Dirichlet's result proved to be false.

(i) An arbitrary continuous function in $[0, \pi]$ cannot be expressed as a Fourier series. (False)

(ii) Anyway there is no way to calculate the coefficients. (False)

(iii) If f is not periodic it is impossible to express it as a sum of periodic functions in $[0, \pi]$. (False)

(iv) The limit of a convergent series of analytic functions must be analytic. (False)

(v) By the uniqueness of analytic continuation, if $\sum_n f_n$ converges to f in an interval, and if both f and all the f_n are analytic in a larger interval, then $\sum_n f_n$ must also converge to f in the larger interval. (False)

(vi) The limit of a convergent series of continuous functions must be continuous ('Theorem' of Cauchy 1823: False – it needs to be uniformly convergent).

Let us now summarise:

CONCLUSIONS ABOUT THE VIBRATING STRINGS CONTROVERSY

1. Each of d'Alembert, Euler and Daniel Bernoulli made major positive contributions to the understanding of vibrating strings.
2. Most of their criticisms of each other's work turned out to be wrong in retrospect.
3. Bernoulli's contribution was the deepest, because it gave the most insight into music and harmony, and his conjectures eventually proved to be correct.
4. The controversy lasted for 80 years because the mathematics was not proved until Dirichlet's paper.

5. The controversy was mathematically very fruitful, because it stimulated research that gave a deeper understanding of functions, differential equations, Fourier series, uniform convergence and harmonic analysis.

THE CATASTROPHE THEORY CONTROVERSY

We now turn to the more recent controversy about catastrophe theory. I propose to use the same historical technique as before, highlighting the main players and indicating the publications that were the more important from the point of view of the controversy itself, as shown in Figure 4. I will then illustrate the nature of the controversy by giving an example. Since I myself was involved I should warn the reader that I am biased in favour of catastrophe theory.

The two most eminent mathematicians involved were René Thom and Stephen Smale, both Fields medalists. Thom was the creator of the theory and Smale the main critic.

The first thing to observe is that catastrophe theory is not a theory in the usual scientific sense of the word, because it does not postulate any scientific hypotheses. It could be called a mathematical theory in the sense that it does embrace a collection of definitions, theorems, proofs and conjectures, but this would miss the point because the main emphasis is on applications. Catastrophe theory is better described as a method of modelling. To put it boldly, Thom wanted to understand the emergence of form in nature. His philosophy was the opposite of that of ergodic theory, in which whatever spacial structures you start with you finish up with a homogeneous soup. Thom started with the soup and wanted to explain the emergence of structure. This is exactly the subject matter of Pandora's Box, and so a controversy was not so unexpected.

The mathematical precursor to catastrophe theory was Whitney's theorem [22] classifying the stable singularities of smooth maps from the plane to itself (the fold and the cusp), and Thom's own generalisation of this to higher dimensions [18]. Having secured his reputation in pure mathematics with his Fields medal in 1958, Thom felt emboldened to explore the philosophy of applications. He had always been interested in the natural world and had a healthy scepticism of conventional science. He was influenced not only by his own work on singularities, but also by some experiments that he did with light caustics (which I shall explain below), by some graphic models that he once saw of the unfolding embryo, by D'Arcy Thompson's book *On growth and form* [20], and by an earlier paper of mine on a topological model of the brain [24]. In moving from differential topology towards dynamical systems and applications Thom was to a certain extent moving into Smale's territory, although the types of applications that they were interested in lay in different boxes: Smale's were primarily in the Time Box, whereas Thom's were firmly in Pandora's Box.

Thom discovered the main theorems classifying the elementary catastrophes in the early 1960's; he invented the mathematical concepts necessary to prove them and sketched out the broad lines of proof, but there were still some gaps. In particular he got stuck at the preparation theorem, so he persuaded Malgrange (against his will according to MALGRANGE [13]) first that it was true, and then to prove it. He also inspired Mather to develop the analysis of germs [14]. The proofs were complete by 1968, and Thom summarised the results in his book *Structural stability and morphogenesis* [19], which was mainly about applications and an outline of his general theory of models. After some delay the book appeared in 1972.

THE PRECURSORS

1945	Whitney's theorem [22] classifying stable singularities of maps of the plane (the fold and the cusp)
1956	Thom's extension to higher dimensions [18]
1958	Thom is awarded a Fields Medal
1966	Smale is awarded a Fields Medal

CATASTROPHE THEORY

1968	Thom's theorem classifying elementary catastrophes (with help from Malgrange and Mather)
1972	Thom's book *Structural stability and morphogenesis* [19]

THE CONTROVERSY

1974	Zeeman's address to the Vancouver Congress [25]
1976	Zeeman's article in Scientific American [26]
1977	Zahler & Sussman's article and the replies in Nature [23]
1977	Zeeman's book *Catastrophe theory* [27]
1978	Smale's review in the *Bulletin of the American Mathematical Society* [15]
1980	Smale reprints the review in his book *The mathematics of time* [16]

FIGURE 4. The catastrophe theory controversy.

I myself became involved when I took a sabbatical year during 1969/70 at the Institut des Hautes Etudes Scientifiques in Paris, where Thom was one of the permanent professors. I had just spent five years setting up the Mathematics Institute at Warwick, and was ready to explore a new line of research. I was fascinated by Thom's ideas and began to see how they could be applied to all branches of science. I was impressed by the depth and beauty of the mathematics. The geometric approach particularly appealed to me because it gave an overall grasp and insight into phenomena that had previously appeared to be opaque or unmodellable. I also liked the elegance of the models, in that the theorems guaranteed the existence of intrinsic coordinates in an application with respect to which the model took on canonical form. My own approach was complementary to that of Thom; I was the pragmatic anglo-saxon compared with his gallic philosopher. While he painted the broad canvas I tried to work with scientists suggesting simple concrete models, designing experiments, and making predictions that they could test. In particular some of my models in the biological and behavioural sciences began to attract attention, and led to some exaggerated claims in the press. After my address [25] to the Vancouver International Congress of Mathematicians in 1974, and an article [26] in Scientific American in 1976, I received over ten thousand requests for reprints. Soon hundreds of scientists began using catastrophe models in the physical, engineering, biological, medical and social sciences [21].

Of course the word 'catastrophe' itself was partly to blame for the publicity, and at the time I did have some misgivings about Thom's choice of this word. In retrospect, however, I think it was a wise and courageous choice on his part, because it focused attention upon the sudden jumps, enabling researchers to perceive them, discuss them, look for them, design experiments to test for them, and process the data so as to reveal them.

For example I was talking to a psychiatrist who was showing me some oscillatory graphs of the mood-swings of his manic-depressive patients, and so I naturally asked him whether or not the swings were catastrophic, meaning whether or not they were sudden, and whether or not the moods in between were stable. To find out I suggested that he record the patients' moods hourly instead of weekly. To his surprise he discovered that they were indeed catastrophic, much more so than he had previously suspected.

Smale, meanwhile, became increasingly sceptical about catastrophe theory. He felt it had limited substance, great pretension and that catastrophe theorists had created a false picture in the mathematical community and the public as to its power to solve problems in the social and natural sciences. He encouraged other critics like Zahler and Sussman to challenge this false picture in print. It was the sharpness of their attack that attracted attention and gave rise to the perception of a controversy. Their criticism, however, came to an abrupt end after they had published an article [23] in *Nature* in 1977, because all the replies in *Nature* [23] from mathematicians, physicists and biologists were unanimously in favour of catastrophe theory; the replies pointed out the misquotations and misrepresentations in Zahler and Sussman's article, and drew attention to a variety of successful applications that they had ignored.

When my book *Catastrophe theory* [27] appeared later that year Smale decided it was time to show his hand, so in 1978 he offered to review the book for the *Bulletin of the American Mathematical Society.* This review [15] was evidently an important paper for Smale, because in 1980 when he published his own book *The mathematics of time* [16] containing his eleven best papers in dynamical systems he rather surprisingly included amongst them this somewhat negative review.

And there the controversy more or less ended. The main reason why it was so short-lived compared with the vibrating strings controversy was that the underlying mathematics of elementary catastrophe theory had been established before the controversy began. Today this mathematics continues to develop and merge seamlessly with that of the neighbouring fields of differential topology, singularity theory, global analysis, dynamical systems, bifurcation theory and chaos. And scientists in many disciplines continue to quietly use catastrophe models whenever they are appropriate.

Smale's chief complaint was that the publicity had given a misleading impression of the importance of catastrophe theory compared with the rest of dynamical systems. At the time this criticism had some validity, but today it looks as ephemeral as the publicity itself. Yesterday's fashion was catastrophe, today's fashions are chaos and string theory, and tomorrow's fashions will be something else. The exaggerated claims of journalism today on behalf of any new theory are in fact the reverse side of a good coin, namely the increasing public interest in science; and Smale admits that it is important for scientists to be aware of mathematical possibilities for models.

Smale's main criticism of Thom was to say that when he (Thom) proceeds beyond elementary catastrophe theory he loses pretty much any direct contact with mathematics. The elementary theory concerns point attractors, and the interest comes from the variation due to changing parameters. Non-elementary theory concerns more complicated attractors, such as the catastrophic jump into a chaotic attractor at the onset of turbulence. It also includes generalised catastrophes such as the boundary of a set of elementary catastrophes, like the surface of a tree being the limit of a set of branching twigs, or the surface of a lung being the limit of a set of branching bronchial tubes. It is true that the mathematics of such examples is still in its infancy, and nowhere near any kind of classification as there is in the elementary theory. Developmental biology is also still in its infancy, because there are very few models yet to describe how the DNA code determines the unfolding of the embryo. But Thom's vision was to try and imagine in his general theory of models what will be needed in the mathematics and biology of the future in order to solve these problems. His writing is often provocative, because he wants to provoke the reader's own imagination as to what will be needed. And he finishes his book with the words:

> *A mathematician cannot enter on subjects that seem so far removed from his usual preoccupations without some bad conscience. Many of my assertions depend on pure speculation and may be treated as day-dreams, and I accept this qualification - is not day-dream the virtual catastrophe in which knowledge is initiated? At a time when so many scholars in the world are calculating, is it not desirable that some, who can, dream?*

Smale disappointingly chooses not to address Thom's main idea, apart from sarcastically quoting a few of Thom's choicer remarks. This may have been because Thom had deliberately chosen to write at the levels of mathematics and the philosophy of science, rather than at the scientific level in between. He knew that the mathematics was unassailable, and he wanted to stimulate debate about the science of the future (as opposed to science of the past, which is what most philosophers of science usually talk about). Consequently he was quite amused at the controversy, while at the same time standing somewhat aloof from it. He had hoped that people would take up his philosophical challenge, and was disappointed at the level of debate.

Realising that in D'Arcy Thompson's 1917 book *On growth and form* [20] most of the general principles had been right, but many of the specific details had subsequently turned out to be wrong, Thom sought to avoid making the same mistake by steering away from experimental prediction, and leaving that to others. This was a quite a sensible policy from the point of view of keeping his main ideas intact for the future, because although the failure of a few specific models might be irrelevant to the success of the general theory, nevertheless there was a danger that the former might be seen as a blemish on the latter. On the other hand he was apt to complain that more biologists had not taken up his ideas – but of course to achieve that it may be necessary to get one's hands dirty.

Meanwhile I was quite prepared to get mine dirty; consequently my models were much more open to attack and Smale turned on them with relish. Let me therefore give an illustration of the type of example that annoyed Smale and the other critics, and for your amusement I will choose one in economics that is new and topical. Then I will attack it as Smale would have done, using his own words from his review, and then I will respond with a defence.

MODEL OF THE DEVALUATION OF THE POUND

The model concerns the currency markets and the recent catastrophic devaluation of the pound in September 1992. It could also apply to future devaluations of other currencies. If the exchange rate is flexible then the value of the pound floats freely up and down with the rise and fall of confidence in the pound, as shown in Figure 5(i).

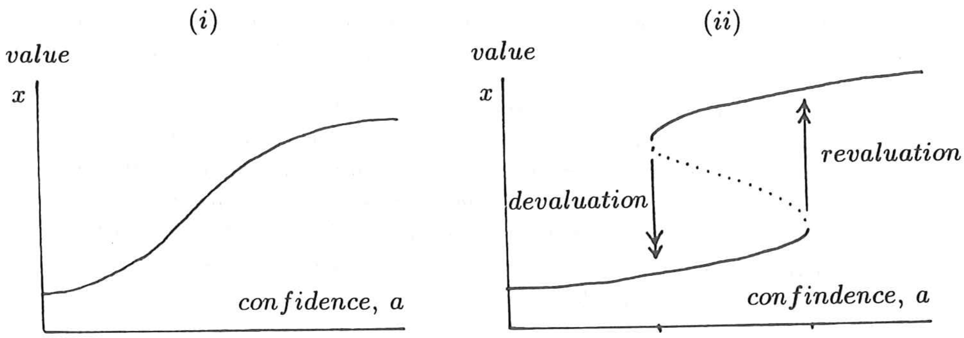


FIGURE 5. Graphs showing how the value x depends on confidence a when the exchange rate is (i) flexible or (ii) constrained.

If the exchange rate is constrained, however, as it was in the ERM (the European Exchange Rate Mechanism) then a fall in confidence does not induce a corresponding fall in value because of the defensive intervention by the central banks, until a sufficiently low threshold is reached where it triggers a sudden devaluation. Similarly a rise in confidence again does not induce a rise in value until a much higher threshold is reached, where it triggers a sudden revaluation, as shown in Figure 5(ii). The middle dotted part of the graph represents unstable equilibria, and so is not observed in practice.

Suppose now that the level of flexibility can be made into a continuous parameter, b. Then we can combine Figures 1 and 2 into the cusp catastrophe shown in Figure 6. Here let $\mathbb{R}^3$ denote 3-dimensional space; the two horizontal axes represent the two parameters, confidence a and flexibility b, while the vertical axis represents the value x. The surface M in $\mathbb{R}^3$ is the graph of x over a, b and shows how the value depends upon both market confidence and exchange rate flexibility.

The recent history of the pound can now be visualised as a path in the parameter space, inducing a corresponding path on the surface of the graph above. The path starts in 1985 when the pound began tracking the deutschmark, thus increasing market confidence and reducing exchange rate flexibility (in other words increasing both a and b). When the UK joined the ERM in 1990 flexibility was further reduced. Confidence in the pound then began to fall but the value was protected, because it was being held on the upper surface of the graph near the front. In September 1992 the parameters had reached the point B. There was still some way for the confidence to fall before it reached the threshold C that would trigger devaluation. The Bundesbank made a suggestion, however, of a possible basket of currency realignments, which introduced a new perception of flexibility within the ERM. This had the effect of moving the parameter point from B to D, thus triggering the devaluation at D. The interesting mistake that the UK made was to concentrate on defending itself against the move from B to C while overlooking the more vulnerable move from B to D. As a result the UK was cross with Germany for triggering this move, while Germany was cross at being criticised for trying to be more flexible. The pound meanwhile is now floating back to its original starting point. We can justify the cusp catastrophe in Figure 6 by making two hypotheses as follows.

HYPOTHESIS 1. There is a value b_o such that the sections of the graph in Figure 6 are equivalent to Figure 5(i) if $b < b_o$, and Figure 5(ii) if $b > b_o$.

HYPOTHESIS 2. The market can be implicitly modelled by a parametrised dynamical system D on a space X. X may be either 1-dimensional with coordinate x, or multi-dimensional with x as one of its coordinates (the other coordinates including possibly the values of other currencies). The parameter space is $\mathbb{R}^2$ with coordinates a, b. The dynamic D has a generic parametrised Lyapunov function f. Let M denote the subset of $\mathbb{R}^2 \times X$ consisting of the equilibria of D, which are the same as the critical points of f. If X is 1-dimensional then $\mathbb{R}^2 \times X = \mathbb{R}^3$ and M is the surface shown in Figure 6. If X is multi-dimensional then the projection of X onto the x-axis induces a map of

$\mathbb{R}^2 \times X$ onto $\mathbb{R}^3$ throwing M onto the surface shown in Figure 6. The unshaded part of M corresponds to stable equilibria (minima of f), and the shaded part to unstable equilibria (maxima or saddles of f).

DEDUCTION. By using Thom's deep classification theorem we can then deduce that Figure 6 is a cusp catastrophe.

Let us now criticise the model as Smale would have done, using his own words from his review. Then I will respond with an answer to each criticism, and give an explanation of each answer below.

CRITICISM 1. No justification is given for the model in terms of existing data or economic theory. It fits the caricature of a mathematician throwing a model to the economists to pick up and develop.
ANSWER 1. True, but intentional.

CRITICISM 2. Hypothesis 1 alone gives the structure of the surface in 3-dimensions without using any mathematics at all. There is no need to involve 'Thom's deep classification theorem', and to do so is not only misleading but also mystifying and intimidating to non- mathematicians.
ANSWER 2. Criticism 2 is an elementary mathematical mistake on Smale's part, since there is a simple counterexample, given below, showing that Hypothesis 1 alone does not imply a cusp catastrophe.

CRITICISM 3. Anyway the cusp is not Thom's theorem but Whitney's theorem.
ANSWER 3. Criticism 3 is a serious mathematical mistake on Smale's part, revealing a misunderstanding of the difference between the two theories.

Let me now explain the answers.

EXPLANATION 1. The statements in Criticism 1 are valid, but I do not necessarily regard them as criticism. I gave several examples in my book of such models that have led to useful scientific prediction and successful experimental confirmation, which Smale chose to ignore. As far as I know Smale in his own writing does not make any scientific predictions, nor propose any scientific experiments, and so he has no experience in the matter. He makes creative *mathematical conjectures*, which are sometimes right and sometimes wrong, but he never ventures to make *scientific predictions.* To give an example of such a prediction let me digress by describing a simple model from Chapter 1 of my book which later led to successful confirmation by experiment.

DIGRESSION ON TERRITORIAL FISH
This is a model of the mood of a territorial fish defending his nest against invasion by another male of the same species. The parameter in this case is the distance of the invader from the nest. When the invader is far away the mood is protective while the defender sits on his nest. When the invader crosses the threshold A the stability of that mood breaks down and there is switch of mood into an aggressive mood, causing the defender to attack the invader.

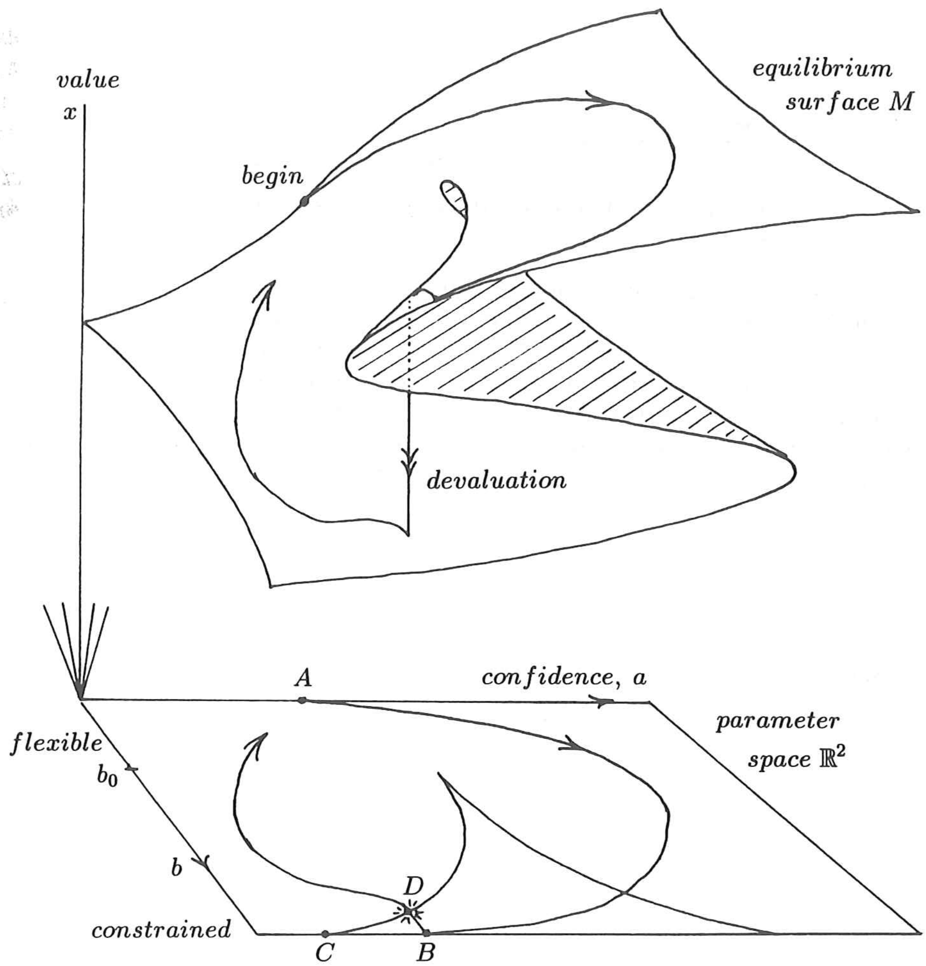


FIGURE 6. The cusp catastrophe, showing the graph of the value of the pound as a function of market confidence and exchange rate flexibility. The recent history of the pound is shown by the path on the graph.

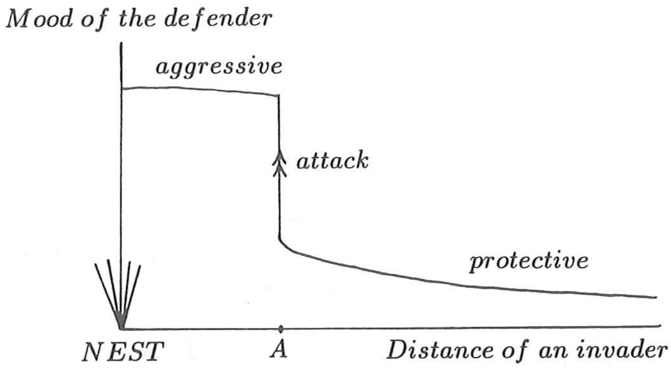


FIGURE 7. Switch of mood as the invader approaches.

We now make a hypothesis that the mood can be implicitly modelled by a parametrised multi-dimensional dynamical system describing the neurological activity in the brain of the fish. Moods are represented by stable attractors of the system. But the attractors of a parametrised dynamical system exhibit hysteresis. In other words the threshold B where the stability of the aggressive mood breaks down, causing the reverse switch back to the protective mood, is different from A. Therefore the graph is as in Figure 8.

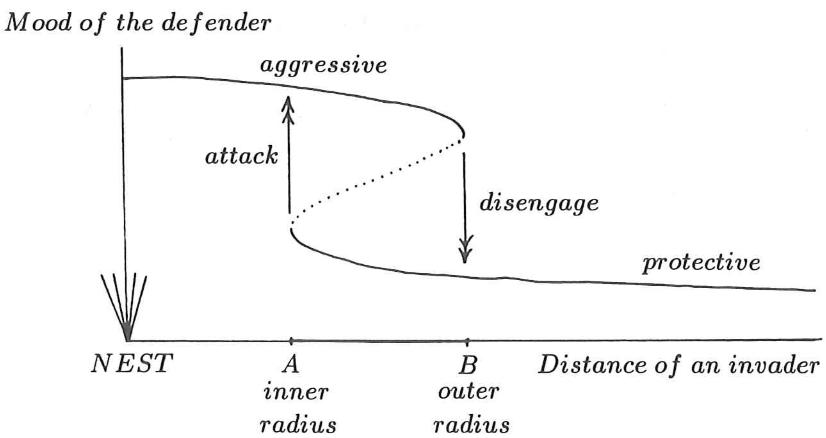


FIGURE 8. Graph of mood over distance.

Here we are implicitly mapping the multi-dimensional space of the dynamical system onto $\mathbb{R}$ in such a way as to map the attractors underlying the protective and aggressive moods to different points. In an experiment, however, there is no need to measure the mood neurologically because the switches of mood can be observed psychologically by the change of behaviour, and the distance from the nest of the points A, B where these switches are observed to take place can be measured. Therefore, although the hypothesis is implicit and the model qualitative, nevertheless the prediction is quantitative; the fish's territory will have two radii that can be measured, an inner radius where he attacks and an outer radius where he disengages. When I discussed this idea with biologists I found that the concept of a territory having two radii was new to them, so in 1976 I published the model in *Scientific American* [26] and reprinted it in my book [27]. In effect I was following Smale's caricature of a mathematician throwing a model to the biologists to pick up and develop, as described in Criticism 1.

And sure enough, without my knowing, the biologist P.W. COLGAN [6] picked it up and decided to test the prediction on pumpkinseed sunfish nesting in their natural habitat on the bottom of Lake Opinicon near Kingston, Ontario. The water was clear and about a metre deep. He and his students made dummy wooden models of invaders in an aggressive posture, and marked the points where the defenders attacked and disengaged. They confirmed the prediction and found that in several hundred trials during the nest-building phase the average values of the two radii were 13cm and 18cm ($\pm$ 0.2cm). Moreover they found that an interesting social structure resulted from this behaviour. The nests were close-packed so that the inner perimeter of one abutted onto the outer perimeter of the next; indeed one can see that if they were to overlap then the two fishes could get into a fight, both locked in the aggressive mood, until one was injured causing his nest to disappear. Between the inner territories was a buffer zone where other fishes could swim freely without being attacked, and through which the females could thread their way safely to visit their partners, as shown in Figure 9. The buffer zone is thus an evolutionary advantage arising from the naturally occurring hysteresis in the dynamics of the brain.

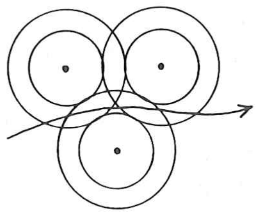


FIGURE 9. Close-packing of the territories.

One is tempted to say 'how very human', thereby tacitly admitting that the territorial imperative in the human species is instinctive rather than rational. Indeed this may be true because the territorial instinct appears to be located in the R-complex, which is the philogenetically oldest part of the human forebrain that evolved some three hundred million years ago in our ancestors the mammal-like reptiles, whereas rational thought appears to be located in the cortex, which is the newest part of the forebrain that evolved some thirty million years ago in primates.

We end the digression and return to Criticism 1 of the economics model. I do not have the expertise to justify it in terms of existing data or economic theory, but in spite of this I believe the model is potentially useful if it gives insight into the relationship between confidence and flexibility. In particular if economists were able to measure flexibility this might lead to the design of exchange rate mechanisms that were more flexible than the ERM but less flexible than a free market, in which it was possible to make realignments of currency in times of falling confidence that were more controlled and less catastrophic.

EXPLANATION 2. Smale devoted a third of his review to Criticism 2, and made an elementary mathematical mistake. At first sight it may look as if Hypothesis 1 implies a cusp catastrophe, but that is false, and I give below a simple counterexample to prove it is false. The reason why it looks true is that I have deliberately arranged the hypotheses so that they are transparently as plausible as possible to social scientists, exactly as I did for the two particular examples in my book that Smale chose to criticize. In effect I am saying to the reader: if you are willing to accept the geometrically explicit Figure 5 and Hypothesis 1 then I can prove rigorously that Figure 6 is a cusp catastrophe, provided that you also allow me to assume the implicit Hypothesis 2. The crucial word in Hypothesis 2 is *generic*[1] and it is plausible because by Thom's theorem generic functions are open dense in the space of all functions. This is the real power behind the theorem; openness implies that generic models are stable under perturbation, and density implies that any model can be approximated by a generic one. In other words generic models are universally good for applied mathematics in the same way that Bernoulli's solutions were universally good for the vibrating string.

The counterexample shows that Hypothesis 2 is necessary, and Thom's theorem shows that it is sufficient. Here sufficient means that Hypotheses 1 and 2 together imply that Figure 6 is cusp catastrophe (in other words is equivalent to the canonical model - see the definition below). The simplest way to prove this is to appeal to Thom's classification. One can also prove it directly from the hypotheses without appealing to Thom's theorem, but the direct proof (even in the case when X is only 1-dimensional) is almost as long as the proof of Thom's theorem itself, since it involves having to show the equivalence of both germs and unfoldings. When Smale claimed in his review that it followed 'without using any mathematics at all' he evidently was not aware of the definition of a cusp catastrophe and did not realise it was necessary to prove that it was equivalent to the canonical model.

Before writing my book [27] I tried for some years to shorten the proof of Thom's theorem, particularly in the case when X was 1-dimensional, in order to make it available to the more mathematically inclined biologists and social scientists, but without success. In Chapter 18 of the book I managed to reduce the complete proof to 66 pages, but this is not easy reading for non-

[1] A parametrised function is defined to be *generic* if the induced map into the germ space is transversal to the natural stratification of the germ space [see 27],

mathematicians. That is why I use the word 'deep', as a warning to scientists that if they need to be sure of handling the mathematics correctly then it may be advisable to collaborate with a mathematician who does understand the proof.

Before constructing the counterexample we need to define the canonical model.

DEFINITION OF THE CANONICAL CUSP CATASTROPHE

The canonical cusp catastrophe is defined as follows.

- The *parametrised Lyapunov function* $f : \mathbb{R}^3 \to \mathbb{R}$ is given by

$$f(a,b,x) = \frac{1}{4}x^4 - ax - \frac{1}{2}bx^2.$$

- The *parametrised dynamic* on $\mathbb{R}$ is given by

$$\dot{x} = -f_x.$$

- The *equilibrium surface* $M, \subset \mathbb{R}^3$, is given by

$$f_x = x^3 - a - bx = 0.$$

The map $\chi : M \to \mathbb{R}^2$ is induced by the projection π from onto $\mathbb{R}^3$ the parameter space $\mathbb{R}^2$ given by $\pi(a,b,x) = (a,b)$.

The *bifurcation set* B, is defined as the image in $\mathbb{R}^2$ of the singularities of χ, and is obtained by eliminating x from $f_x = f_{xx} = 0$, giving the cusp

$$27a^2 = 4b^3.$$

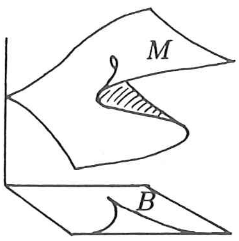
There is analogous definition for when the space X of the dynamic is multi-dimensional rather than 1-dimensional [see 27].

DEFINITION OF A CUSP CATASTROPHE

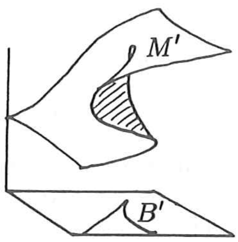
Let $f' : \mathbb{R}^3 \to \mathbb{R}$ be another parametrised Lyapunov function for some dynamic (not necessarily the gradient of f'). Suppose that f' determines M', χ', B' as above. We say that f' is a *cusp catastrophe* if it is *parametrised-function-equivalent* to f, in other words there are diffeomorphisms α, β, γ such that the diagram commutes:

$$\begin{tikzcd}
\mathbb{R}^3 \arrow[r, "\alpha", "\text{diffeo}"'] \arrow[d, "\pi\times f"'] & \mathbb{R}^3 \arrow[d, "\pi\times f'"] \\
\mathbb{R}^2\times\mathbb{R} \arrow[r, "\beta", "\text{diffeo}"'] \arrow[d, "\pi"'] & \mathbb{R}^2\times\mathbb{R} \arrow[d, "\pi"] \\
\mathbb{R}^2 \arrow[r, "\gamma", "\text{diffeo}"'] & \mathbb{R}^2
\end{tikzcd}$$

where π always denotes projection onto the parameter space $\mathbb{R}^2$. It follows that χ, χ' are *map-equivalent*, in other words that the diagram commutes:



$$\begin{tikzcd}
\mathbb{M} \arrow[r, "\alpha|M", "\text{diffeo}"'] \arrow[d, "\chi"'] & \mathbb{M}' \arrow[d, "\chi'"] \\
\mathbb{R}^2 \arrow[r, "\gamma", "\text{diffeo}"'] & \mathbb{R}^2
\end{tikzcd}$$



Therefore γ maps B to B'. Therefore the cusp on B is mapped to a cusp on B', since a cusp on a curve is a diffeomorphism invariant. Therefore B' contains a cusp.

THE COUNTEREXAMPLE

To construct a counterexample to Criticism 2 it suffices to find an f' such that M' satisfies Hypothesis 1 and B' does <u>not</u> contain a cusp. The trick is to replace a by a^3 in the formula for f. This does not qualitatively alter the sections of M' so they remain equivalent to those in Figure 5 and continue to satisfy Hypothesis 1. But it converts the bifurcation set B' into the parabola

$$3a^2 = 4^{1/3}b.$$

Therefore the cusp is lost. Therefore it is not a cusp catastrophe. And in the economics application the subtlety of the discussion on what triggered the currency devaluation is lost.

The explanation is that f' is not generic and so does not satisfy Hypothesis 2. Furthermore one can make worse examples B'' by replacing a^3 by a^{3k}, for any positive integer k as shown in Figure 10.

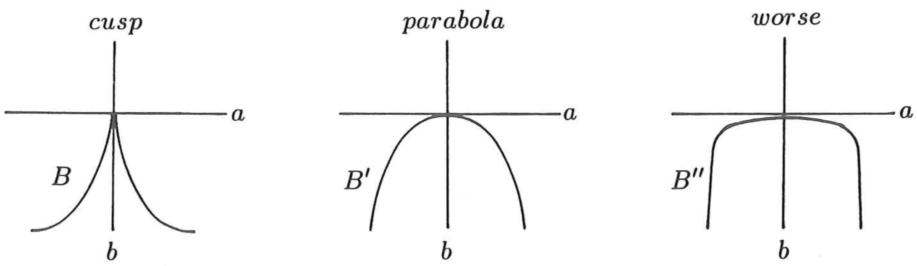


FIGURE 10. Bifurcations sets.

I actually explained this little counterexample to Smale in the presence of Moe Hirsch, Mike Shub and John Guckenheimer while I was Hitchcock Professor at Berkeley in 1977. He was extraordinarily resistant to receiving it, and it took a long time to get across much to the surprise of the others because it normally takes only a few minutes to explain to a dynamical systemist. By the time Smale came to write his review he had evidently forgotten it again; perhaps Criticism 2 was one of the things that he acknowledges he learnt from Zahler and Sussman, since they also made the same mistake in an earlier paper.

EXPLANATION 3. Criticism 3 is a serious mistake which suggests that Smale did not fully understand the difference between Whitney's theorem and Thom's theorem. Whitney classifies stable singularities of maps, whereas Thom classifies elementary catastrophes. Now the elementary catastrophes also give rise to certain singularities of maps (from the equilibrium manifold to the parameter space), which are stable within the context of catastrophe theory, but which are not the same as the Whitney singularities in dimensions greater than 2. The two classifications happen to coincide in dimension 1 (the fold) and dimension 2 (the cusp), but to establish that fact it is necessary to separately prove both Whitney's theorem and Thom's theorem. You cannot deduce Thom's theorem in dimension 2 from Whitney's theorem as Smale seems to think; it is a question of extending the diffeomorphism $\alpha|M$ of the map-equivalence to the diffeomorphism α of the parametrised-function-equivalence (see above), which requires the full panoply of the preparation theorem. That was why it took most of the 1960's to prove Thom's theorem.

The fact that the two classifications differ in dimension 3 is historically interesting because that was how Thom actually discovered catastrophe theory. He was experimenting with 3-dimensional light caustics expecting to find only the swallowtail (because that is the only Whitney singularity in $\mathbb{R}^3$), and was astonished to observe also the umbilics. He had been thinking it was a map problem, and suddenly realised it was a function problem; the function in this case is the geodesic distance from points on the wave front to points near the caustic, and the rays of light follow the critical geodesics perpendicular to the wave front [see 19].

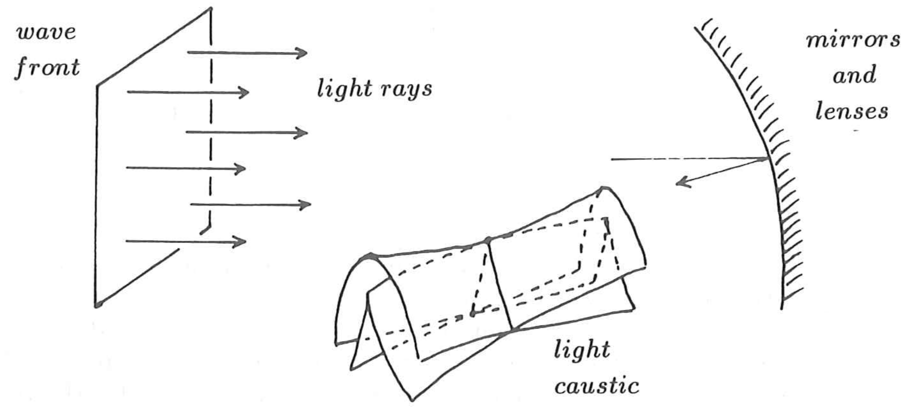


FIGURE 11. Light caustics.

If Smale thought that it was possible to deduce Thom's theorem from Whitney's theorem then I can understand why he used to complain about my use of the word 'deep', because the proof of Whitney's theorem is much more accessible to the scientist than that of Thom's. It is ironic that I explained in detail the differ ce between the two classifications in the very book that Smale was purporting to review.

CONCLUSIONS ABOUT THE CATASTROPHE THEORY CONTROVERSY

1. The controversy was relatively short-lived because the underlying mathematics had already been established and was uncontroversial.

2. The controversy was mainly about applications, and between mathematicians rather than between experts in the fields of those applications.

3. The controversy was similar to that between d'Alembert, Euler and Bernoulli in that both Thom and Smale made major positive contributions to topology and dynamical systems, but Smale's negative contribution was not so great. As I said in my address [11] at the banquet for his 60th birthday:

The real evidence for the excellence of Steve's mathematical judgement is the number of mathematicians world-wide who now follow his taste. In fact I have only known him to make one serious error of judgement, and that was his opinion of catastrophe theory.

REFERENCES

1. J. LER. D'ALEMBERT, 1747, *Récherches sur la courbe que forme une corde tendue mise en vibration,* Hist. de l'Acad. Roy. de Berlin 3, pp. 214-249.
2. D. BERNOULLI, 1732/3, *Theoremata de oscillationibus corporum filo flexili connexorum et catenae verticaliter suspensae,* Comm. Acad. Sci. Imp. Petrop. 6, pp. 108-122.
3. D. BERNOULLI, 1740, *De oscillationibus compositis praesertim iis quae fiunt in corporibus ex filo flexili suspensis,* Comm. Acad. Sci. Imp. Petrop. 12, pp. 97-108.
4. D.BERNOULLI, 1753, *Réflexions et élaircissemens sur les nouvelles vibrations des cordes exposées dans les mémoires de l'Académie de 1747 et 1748,* Hist. de l'Acad. Roy. de Berlin 9, pp. 147-195.
5. J. BERNOULLI, 1728, *Meditationes de chordis vibrantibus, cum pondusculis aequali intervallo a se invicem dissitis, ubi nimirum ex principio virium vivarum quaeritur numerus vibrationum chordae pro una oscillatione penduli datae longitudinis D,* Comm. Acad. Sci. Imp. Petrop. 3, pp. 13-28.
6. P.W. COLGAN, W.NOWELL & N.W.STOKES, 1981, *Spacial aspects of nest defence by pumpkinseed sunfish (Lepomis gibbosus): stimulus features and an application of catastrophe theory,* Animal behaviour 29, pp. 433-442.
7. G.L.DIRICHLET, 1829, *Sur la convergence des séries trigonométriques qui servent a représenter une fonction arbitraire entre des limites données,* Jour. für Math. (Crelle's Journal) 4, pp. 157-169.
8. L.EULER, 1748, *Sur la vibration des cordes,* Hist. de l'Acad. Roy. de Berlin 4, pp. 69-85.
9. J.FOURIER, 1819, *Théorie du mouvement de la chaleur dans les corps solides,* Mém. de l'Acad. Roy. des Sci. de l'Inst. de France, Paris 4, pp. 185-555 (submitted 1807).
10. E.A.GONZÁLEZ-VELASCO, 1992, *Connections in mathematical analysis: the case of Fourier series,* Am. Math. Monthly 99, pp. 427-441.
11. M.W.HIRSCH, J.E.MARSDEN & M.SHUB (Editors), 1993, *From topology to computation; proceedings of the Smalefest,* Springer-Verlag New-York. pp. 79-80.
12. M.KLINE, 1972, *Mathematical thought from ancient to modern times,* Vol 2, Oxford U.P.
13. B.MALGRANGE, 1966, *Ideals of differentiable functions,* Oxford U.P.
14. J.N.MATHER, 1968, *Stability of C^∞ mapping I,* Annals of Math. 87, pp. 69-104; and, 1968, III, Publ. Math. IHES 35, pp. 127-156.
15. S.SMALE, 1978, *Review of E.C. Zeeman: Catastrophe theory, selected papers* 1972-1977, Bull. Amer. Math. Soc. 84, pp. 1360-1368.
16. S.SMALE, 1980, *The mathematics of time,* Springer-Verlag, New York.
17. B.TAYLOR, 1713, *De motu Nervi tensi,* Phil. Trans 28, pp. 26-32.
18. R. THOM, 1956, *Les singularités des applications différentiables,* Ann. Inst. Fourier 6, pp. 43-87.
19. R.THOM, 1972, *Stabilité structurelle et morphogenèse,* Benjamin, New York, English translation: *Structural stability and morphogenesis,* Benjamin, 1975.

20. D'ARCY W.THOMPSON, 1917, *On growth and form*, republished Cambridge U.P. 1961.
21. B.W.WETHERALL and E.C.ZEEMAN, 1981, *Bibliography on catastrophe theory,* Warwick University preprint.
22. H.WHITNEY, 1955, *Mappings of the plane into the plane,* Ann. Math. 62, pp. 374-470.
23. R.S.ZAHLER & H.SUSSMANN, 1977, *Claims and accomplishments of applied catastrophe theory,* Nature 269, pp. 759-763; and replies in Nature 270 pp. 381-384 and 658.
24. E.C.ZEEMAN, 1962, *The topology of the brain and visual perception,* Topology of 3-manifolds (ed. M.K.Fort) Prentice Hall, NJ, pp. 240-256.
25. E.C.ZEEMAN, 1974, *Levels of structure in catastrophe theory,* Proc. Int. Cong. Math. Vancouver 2, pp. 533-546.
26. E.C.ZEEMAN, 1976, *Catastrophe theory,* Scientific American 234 (April 1976) pp. 65-83.
27. E.C.ZEEMAN, 1977, *Catastrophe theory, selected papers* 1972-1977, Addison-Wesley, Reading, Mass.