There are butterflies everywhere. But who is to say that their flapping wings cancel each other out?

Lorenz, again:

> The average person, seeing that we can predict the tides pretty well a few months ahead would say, why can't we do the same thing with the atmosphere? It's just a different system, the laws are about as complicated. But I realized that *any* physical system that behaved nonperiodically would be unpredictable.

Weather – or not?

In this vein, Lorenz ends his 1963 paper with some speculations about the possibility of weather-forecasting. His argument is simple and original. Imagine recording a very accurate series of measurements of the state of the atmosphere, comparable to those that you wish to use for forecasting. Collect such data for a very long time.

> The crucial point is then whether analogues must have occurred since the state of the atmosphere was first observed. By analogues we mean two or more states of the atmosphere which resemble each other so closely that the differences may be ascribed to errors in observation.

If two analogues *have* occurred, then you will make identical predictions of the future weather, starting from either of them. That is, your weather-predicting scheme must predict *periodic* variation of the weather. But this is nonsense; the whole difficulty with weather-prediction is that the weather is *not* periodic.

If analogues haven't occurred, there's still hope: the entire weather system may be quasi-periodic, almost repeating the same states over again, but with tiny variations, slowly growing. In such a case, long-term weather prediction might be possible. In fact, all you have to do is look back in the records for a close analogue of today's weather, and see what happened last time.

This line of argument fails, Lorenz notes, if 'the variety of possible atmospheric states is so immense that analogues need never occur.' And he leaves one crucial question dangling: 'How long is "very long range"?' He says that he doesn't know the answer, but 'Conceivably it could be a few days or a few centuries'. Twenty-four years later, the centuries have been ruled out, and 'a few days' looks spot on.

Stretch and Fold

We've already seen an example of the butterfly effect, in Chapter 7. Smale's solenoid, or its simpler model, the mapping $x \rightarrow 10x$ on a circle. There the same sensitivity to initial conditions occurs. Two points π and π', agreeing to a billion decimal places, wander about independently of each other after a billion iterations.

That may not sound so bad. But two points agreeing to six decimal places evolve independently after only six iterations.

Where does this sensitivity come from?

It's a mixture of two conflicting tendencies in the dynamics.

The first is *stretching*. The mapping $x \rightarrow 10x$ expands distances *locally* by a factor of ten. Nearby points are torn apart.

The second is *folding*. The circle is a bounded space, there isn't room to stretch everything. It gets folded round itself many times, that's the only way to fit it in after you've expanded distances by ten. So, although points close together move apart, *some points far apart move close together*.

The expansion causes points that start off close together to evolve differently. At first, the difference grows regularly. But once the two points have moved far enough apart, they 'lose sight of each other'. No longer must one mimic the behaviour of the other.

The mixture of stretching and folding is also responsible for the irregular motion. Yes, some points must move closer together again. But which? *How can you tell?* Large differences now are due to very tiny differences many iterations back. You can't see what's coming in advance.

That's unpredictability.

You can see the stretch-and-fold process going on in Lorenz's system. Each half of the front of the surface winds round to the back and is stretched to double its width, before being 're-injected' into the front part again.

It's now pretty clear that Lorenz's strange infinitely-sheeted double-lobed surface must be a strange attractor – the *Lorenz attractor*. And his differential equations, while a somewhat hacked-down version of the physics, are down-to earth equations in three variables with some kind of physical pedigree, be it ever so littered with mongrels. They aren't artificial designer differential equations, labelled 'CAREFULLY MADE BY TOPOLOGISTS' with a green doughnut logo on the label.

And in fact you can find real physical systems which are very well

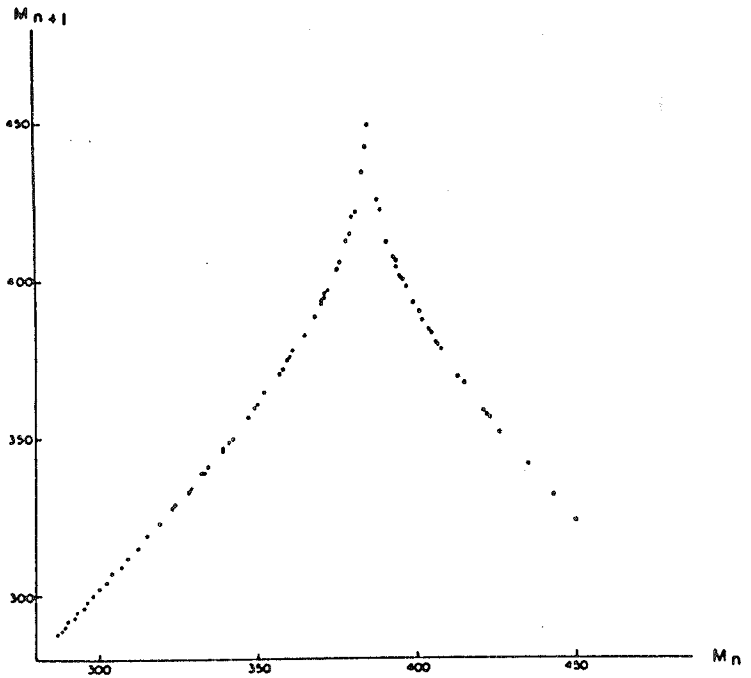

Figure 56 *Order in chaos. If the size of an oscillation is plotted against that of the previous oscillation, a precise curve results. (American Meterorological Society,* Journal of the Atmospheric Sciences, *vol. 20 (Edward N. Lorenz))*

Lorenz noticed this too. He called it the 'butterfly effect'. He discovered it by accident.

He'd had his McBee for several years, since about 1960. He used to set up model weather-systems and let them run, sometimes for days on end. The computer would type out the solution trajectory as a long series of numbers – no fancy computer graphics then. Colleagues would make bets on what Lorenz's microclimate would do next. In the winter of 1961, he was running a precursor of his now famous system. He'd calculated a solution, and he wanted to study how it behaved over a greater period of time. Rather than wait several hours, he noted down the numbers it had reached when it was in the middle of the run, fed them in as a new starting-point, and set the machine going.

What should have happened was this. First, the machine would repeat the second half of the original run, and then it would carry on from there. The repetition served as a useful check; but missing out the first half saved time.

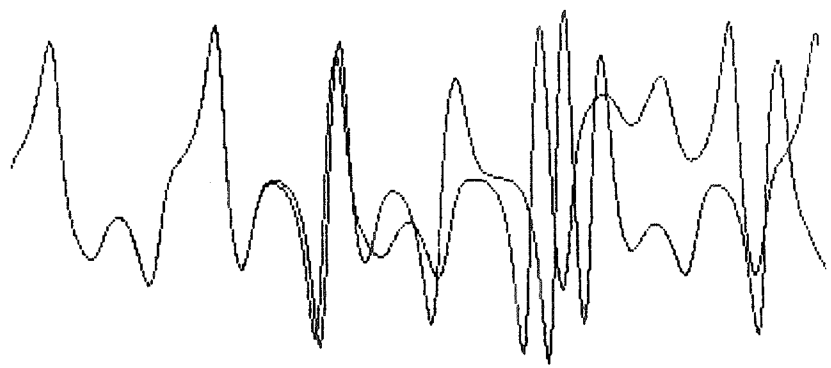

Figure 57 *The butterfly effect: a numerical simulation of one variable in the Lorenz system. The curves represent initial conditions differing by only 0.0001. At first they appear to coincide, but soon chaotic dynamics leads to independent, widely divergent trajectories.*

The meteorologist went off and had a cup of coffee. When he came back, he found that the new run had *not* repeated the second half of the old one! It started out that way, but slowly the two runs diverged, until eventually they bore no resemblance to each other.

In his book *Chaos* James Gleick, a science writer who interviewed Lorenz, tells what happened next.

> Suddenly he realized the truth. There had been no malfunction. The problem lay in the numbers he had typed. In the computer's memory, six decimal places were stored: .506127. On the print-out, to save space, just three appeared: .506. Lorenz had entered the shorter, rounded-off numbers, assuming that the difference – one part in a thousand – was inconsequential.

From the traditional way of thinking, so it should be. Lorenz realized that his equations weren't behaving the way a traditionally-minded mathematician would expect. Lorenz coined his famous phrase: 'butterfly effect' (Figure 57). The flapping of a single butterfly's wing today produces a tiny change in the state of the atmosphere. Over a period of time, what the atmosphere actually does diverges from what it would have done. So, in a month's time, a tornado that would have devastated the Indonesian coast doesn't happen. Or maybe one that wasn't going to happen, does.

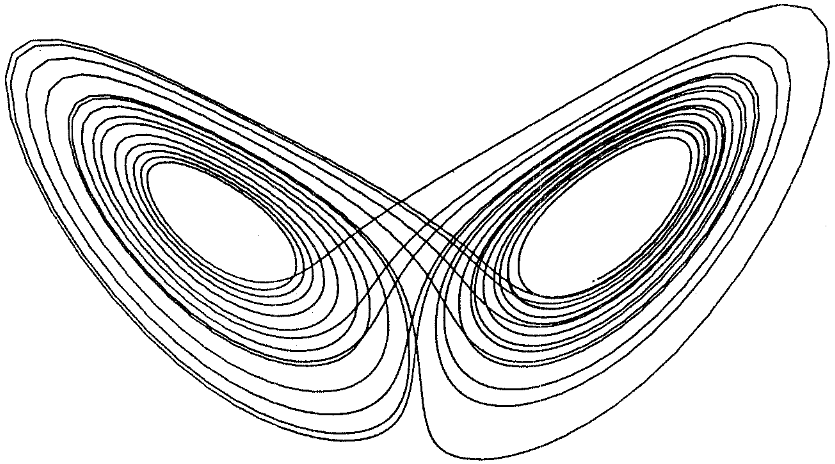

Figure 55 The Lorenz attractor: trajectories cycle, apparently at random, round the two lobes.

Lorenz had a computer. In the early 1960s this was unusual. Most scientists distrusted computers and hardly anybody had one of their own. The machine on which I'm typing this paragraph is a far better computer than Lorenz had, and I'm using it for word processing. It's like using a Rolls Royce to deliver milk. Times change. Anyway, Lorenz had a Royal McBee LGP-300 computer, a not very reliable maze of vacuum tubes and wires. So he put his equations on his Royal McBee and let it royally McBuzz away, at a speed of about one iterations per second. (My word processor is about fifty to a hundred times faster.)

Catch-22: to get out of the bind, the place, people, culture, and time must be right. Poincaré was the person, France the place – but the time and culture were wrong. Lorenz was the person, MIT the place; the culture for chaos is the computer culture, and that was well under way. When everyone has a computer, the *fact* of chaos is impossible to miss. Realizing its importance is another matter, though. For that, the time must be right too – other people have to appreciate that something really interesting is going on. The time wasn't right. More accurately, Lorenz was ahead of his time.

His paper shows the first 3,000 iterations of the value of the variable y (Figure 54). It wobbles periodically for the first 1,500 or so, but you can see the size of the wobble growing steadily. Lorenz knew from his linear stability analysis that this would happen: but what happened *next*?

Madness.

Violent oscillations, swinging first up, then down; but with hardly any pattern to them.

He drew plots of how various combinations of x, y, z varied. In the (x,y)-plane he saw a two-lobed figure like a kidney (Figure 55). Sometimes the point circled the left-hand lobe, sometimes the right.

The trajectories of his equations, he realized, lived on something rather like a squashed pretzel. A surface that had two layers at the back, but merged to a single layer at the front. The point that represented the state of the system would swing round one or other of these surfaces, pass through their junction, and then swing round again.

Lorenz knew that trajectories of a differential equation *can't* merge. So what looked like a single sheet at the front must really be two sheets very close together.

But that meant that each sheet at the back was double too; so there were four sheets at the back . . . So four at the front, so eight at the back, so . . . 'We conclude,' said Lorenz, 'that there is an infinite complex of surfaces, each extremely close to one or the other of two merging surfaces.'

It's not surprising that the meteorologists were baffled. But Lorenz was on to something big.

It's amazing what a bit of xz and xy can do for you.

The Butterfly Effect

It's not true to say that Lorenz found no pattern, that nothing was predictable. On the contrary, he found a very definite pattern. He took the peak values of the variable z, and drew a graph of how the current peak relates to the previous peak. The result was a beautifully precise curve, with a spike in the middle (Figure 56).

Lorenz's curve is a kind of poor man's Poincaré section. Instead of plotting a variable at regular periods of time, he plots z every time it hits a peak. The time intervals are then irregular, but not badly so, because there's a definite underlying rhythm to the Lorenz attractor.

Using the curve, you can *predict* the value of the next peak in z provided you know the value of the current peak. In this sense, at least some of the dynamics is predictable.

But it's only a short-term prediction. If you try to string the short-term predictions together to get a long-term prediction, tiny errors start to build up, growing faster and faster, until the predictions become total nonsense. Indeed, Lorenz's curve has the same stretch-and-fold characteristics that we've learned to associate with chaos, and the stretch makes the errors blow up.

$$\frac{dx}{dt} = -10x + 10y$$

$$\frac{dy}{dt} = 28x - y - xz$$

$$\frac{dz}{dt} = \frac{8}{3} z + xy$$

Here x, y, z are his three key variables, t is time, and d/dt is the rate of change. The constants 10 and 8/3 correspond to values chosen by Saltzman; the 28 represents the state of the system just after the onset of unsteady convection, as we'll see in a moment. These numbers can be changed, depending on the values of physical variables.

If you cross out the terms xz and xy on the right-hand sides, you get a set of equations that any mathematician worth his salt will solve with his eyes shut before breakfast. Boring, though.

But you can do something more useful along those lines. You can find the steady states of the system, where all three expressions on the right vanish, and x, y, z remain constant. There are three: one representing no convection and two others, symmetrically related, representing steady convection. You can also analyse the stability of the system near these states by a method known as *linear stability analysis*. You find that if the 28 is reduced below 24.74 then the state of steady convection is stable. At the critical value 24.74, convection starts up. Lorenz's choice of 28 occurs just after the onset of unsteady convection.

At this point linear theory abandons you. It works well *near* the steady state; but when the steady state becomes unstable, that necessarily means you have to consider what happens as the system moves away from the steady state. So linear theory can tell you where the instability occurs, but not what happens as a result. A pair of binoculars can show you where the brow of the next hill is, but not what lies beyond.

It's a start. Now you know *where the interesting behaviour occurs*. But what is it?

The Advantages of Having a Computer

There's no way out: *you have to solve the equations*. By hook, crook, cunning trickery or brute force. By far the most reliable method is brute force: compute the solution numerically.

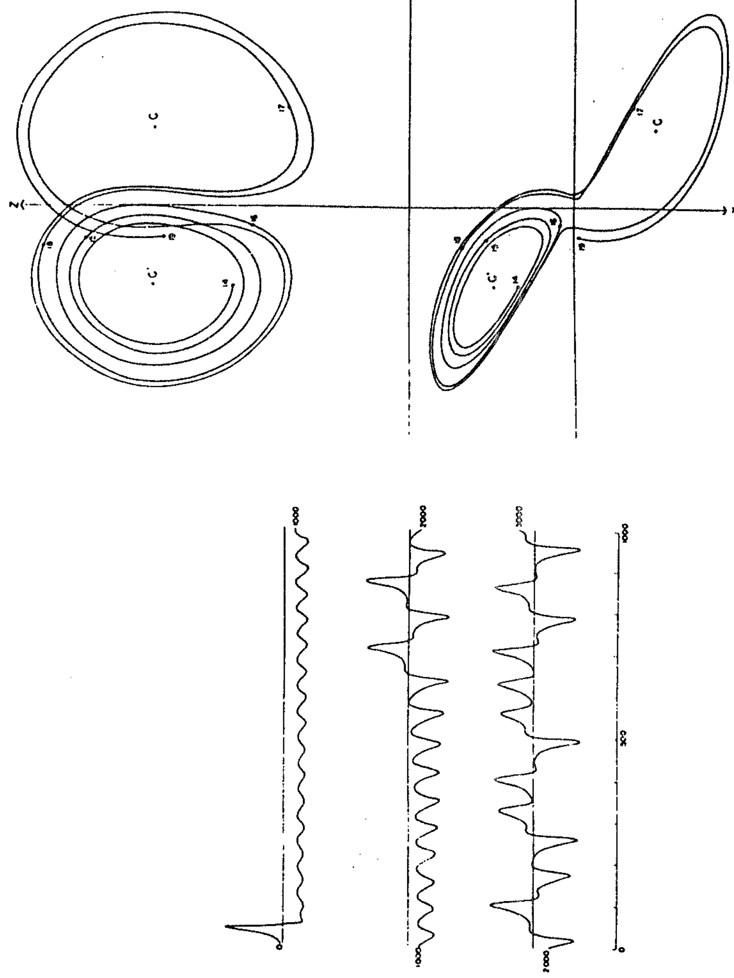

Figure 54 Lorenz's plots of 3,000 numerically computed steps in his equations for convection: (left) oscillations grow and become chaotic, (right) two views of the motion in phase space (American Meteorological Society, Journal of the Atmospheric Sciences, *vol. 20 (Edward N. Lorenz))*

and my hair stands on end. *He knew! Twenty-four years ago, he knew!* And when I look more closely, I'm even more impressed. In a mere twelve pages Lorenz anticipated several major ideas of nonlinear dynamics, before it became fasionable, before anyone else had realized that new and baffling phenomena such as chaos existed.

Lorenz, as I've said, thought he was a meteorologist, and naturally he published his paper in *Journal of the Atmospheric Sciences*. The meteorologists, who were either non-mathematical or versed only in traditional mathematics, really didn't know what to make of it. It didn't look especially important. In fact Lorenz's equations were such a mangled, lopped-off version of the real physics, that the whole thing was probably nonsense.

There are several thousand scientific journals published per year, running on average to well over a thousand pages. If you read a lot you can just about keep up with the publications in your own field. Yes, it's just barely possible that the Spring issue of the *Goatstrangler's Gazette* might contain an idea of enormous importance in dynamical systems theory, but the same goes for a thousand other obscure journals too. With the best will in the world, the best you can do is look in the places you know about. The topologists, whose necks would doubtless have prickled like mine had they come across Lorenz's seminal opus, were not in the habit of perusing the pages of the *Journal of the Atmospheric Sciences*.

And so, for a decade, his paper languished in obscurity. Lorenz knew he was on to something big, but he was ahead of his time.

Let's take a look at what he did.

Courage of his Convections

Hot air rises.

This motion is known as *convection*, and it's responsible for many important aspects of the weather (Figure 53). Thunderclouds form as a result of convection; that's why you tend to get thunderstorms on a hot humid day. Convection can be steady, with the warmer air drifting gently upwards in a constant manner; or unsteady, with the atmosphere moving about in a much more complicated way. Unsteady convection is far more interesting, and more obviously relevant to weather. Since the simplest behaviour after being steady is to change periodically, the simplest kind of unsteady convection is some sort of periodic swirling effect.

The study of convection has a distinguished history. In about 1900

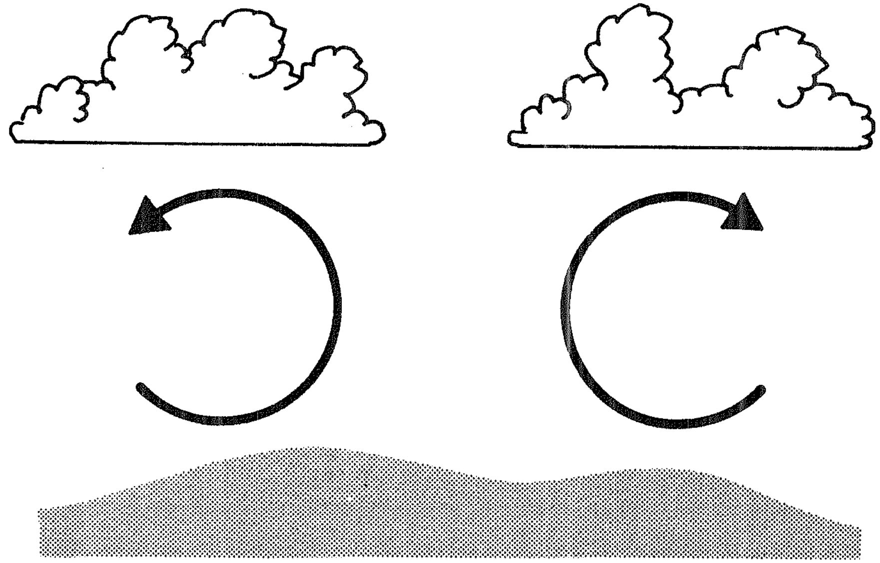

Figure 53 Convection cells, caused by hot air rising

Henri Bénard carried out a fundamental experiment, discovering that when a thin layer of fluid is heated from below it can form convection cells, looking rather like a honeycomb. Lord Rayleigh derived the basic theory of the onset of convection. But there's always more to learn. In 1962 B. Saltzman wrote down the equations for a simple type of convection. Imagine a vertical slice of atmosphere, warm the air at the bottom, keep it cool at the top, and watch it convect. What you expect to see is regularly spaced swirls, the convection cells, going round and round in a periodic fashion. In a manner typical of classical applied mathematics, Saltzman guessed an approximate form of the solution, substituted it into his equations, ignored some awkward but small terms, and took a look at the result. Even his highly truncated equations were too hard to solve by a formula, so he put them on a computer.

He noticed that the solution appeared to undergo irregular fluctuations: unsteady convection. But it didn't look at all periodic.

Lorenz was interested and decided to investigate further. Noticing that only three of Saltzman's variables played a role in this effect, Lorenz threw the rest away. This was a highly cavalier but perfectly conscious act. He obtained a system of equations that has now become a classic:

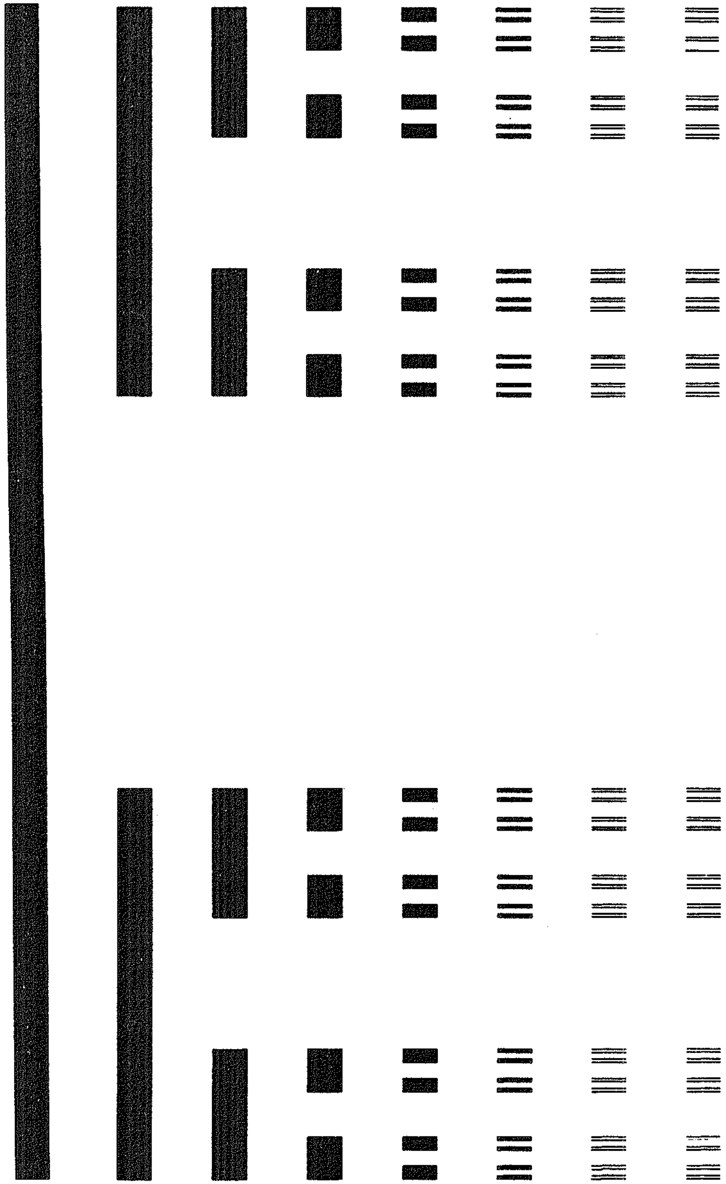

Figure 49 Construction of the Cantor set by repeated deletion of middle thirds. The vertical dimension is exaggerated for clarity: ideally the line has no width.

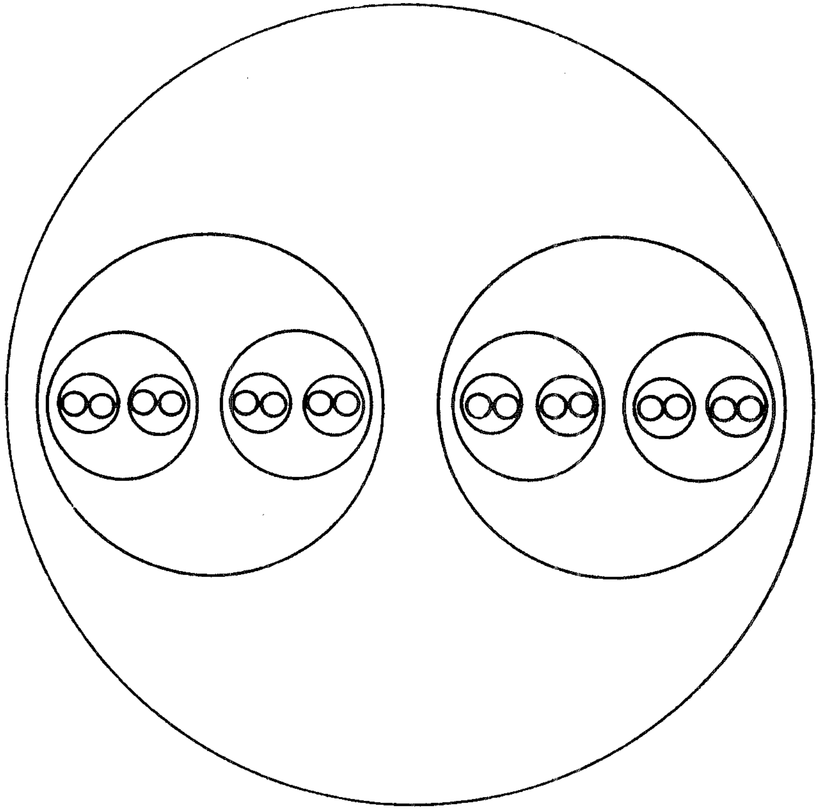

Figure 50 The Cantor cheese: alternative construction of a topological equivalent to the Cantor set, using pairs of circles

other points, as it turns out. The recipe involves expansion to base 3: if you like that sort of thing, see if you can describe exactly which points survive to make up the Cantor set.

The total length of the intervals removed is 1 – the original length of the interval you started with. So in some sense the 'length' of the Cantor set is zero! That's reasonable, the Cantor set consists mostly of holes. It's more like a dust than an interval.

There are other constructions which end up with something that is topologically equivalent to a Cantor set. One of the prettiest is to start with a circular disc, and remove everything except for two smaller discs (Figure 50). Like a button with two holes to put the thread through, except you keep the holes and throw away the button. Repeat this construction on each smaller disc, continue to just talking of the surface of the torus. Define a mapping of the torus to itself as follows. Stretch it out to ten times its circumference and roll it thin; then put it back inside itself so that it wraps ten times round, without passing through any point more than once (Figure 48). (Mathematicians normally use the number 2 rather than 10 here, but to see what goes on then you have to think in binary: I've rewritten history a little to make life easier for us.)

Imagine repeating this transformation of the doughnut. On the next application of the procedure it gets even thinner, and wraps 100 times round itself; then 1,000, 10,000, and so on.

Where does it go in the long run? You get something akin to an infinitely thin line wrapping infinitely many times round the torus. We'll examine this statement for hidden bugs in a moment; but it's not too far off the beam. There's an electrical gadget called a *solenoid*, in which miles of copper wire is wrapped around a metal core to make an electromagnet. Mathematicians borrowed this name for Smale's construct.

Two eminent dynamical systems theorists, colleagues of mine, were discussing all this in an American bar not long after its discovery, waving their hands graphically round and round, and chattering animatedly. 'Ah,' said the barman. 'You must be talking about *solenoids*!' This wasn't the kind of conversational gambit that they expected. Was the barman a mathematics graduate student working his way through college? It turned out he'd been in the navy, and what he was referring to was a real electrical solenoid.

At least the story shows that 'solenoid' is an appropriate name.

Anyway, we get this crazy mapping of a solid torus, in 3-space. Now we plunge our hands into the topological hat and extract a rabbit. Suspend Smale's solenoid mapping, and you get a flow in 4-space with his crazy mapping as a Poincaré section.

If you're not used to thinking in 4-space, you'll get the wrong picture at this point. You'll imagine a point starting in the middle of the dough, and wandering around through 3-space until it eventually ends up back inside the dough again. That's wrong. It moves out of 3-space altogether, immediately, without passing through the dough, wraps round in an entirely new dimension, and then hits the dough again somewhere else. As an analogy, using time as the fourth dimension, if you time-travel from *now* into the future, you leave the present 3-space *immediately*.

If you iterate the mapping from the torus to itself a large number of times, all initial points move closer and closer to the solenoid. So the solenoid is an attractor for the dynamics on the Poincaré section. The suspension of the solenoid – what you get when you whiz

round in the extra dimension – is therefore an attractor for the full 4-dimensional flow.

Furthermore, it's structurally stable. To see why, imagine making a very small change to the wrapping mapping. The result will still look pretty much the same. You can't change continuously from a wrap-ten-times mapping to a wrap-nine or a wrap-eleven-times. To change continuously from ten to eleven you have to pass through ten and a half, but there's no way to wrap a torus ten and a half times without breaking it. That means the dynamics after making a small change to the mapping looks topologically the same as it did to begin with; and that's what structural stability means.

Finally, the solenoid is not a single point, and it's not a circle. So it can't be one of the traditional typical attractors. Two mathematicians, Floris Takens and David Ruelle, coined a name for this new type of attractor. A structurally stable attractor that is not one of the classical types, point or circle, is said to be a *strange attractor*. The name is a declaration of ignorance: whenever mathematicians call something 'pathological', 'abnormal', 'strange', or the like, what they mean is 'I don't understand this damned thing'. But it's also a flag, signalling a message: *I may not understand it, but it sure looks important to me.*

Cantor Cheese

The solenoid is not quite as crazy as it looks. Although it isn't a nice classical point, or circle, it has a distinguished pedigree. This is highly relevant to later developments, so I'll say a little more. The appropriate object is known as the *Cantor set* (Figure 49), because it was discovered by Henry Smith in 1875. (The founder of set theory, Georg Cantor, used Smith's invention in 1883. Let's face it, 'Smith set' isn't very impressive, is it?) The Cantor set is an interval that has been got at by mice. Infinitely many vanishingly small mice, each taking tinier and tinier bites.

Less colourfully, to build a Cantor set you start with an interval of length 1, and remove its middle third (but leaving the end points of this middle third). This leaves two smaller intervals, one-third as long: remove their middle thirds too. Repeat indefinitely. You get more and more shorter and shorter intervals: pass to the limit where the construction has been repeated infinitely many times. This is the Cantor set.

You might think that nothing at all is left. But, for example, the points 1/3 and 2/3 escape removal, and so do 1/9, 2/9, 7/9, and 8/9. *All* the end-points of removed segments remain. So do quite a lot of

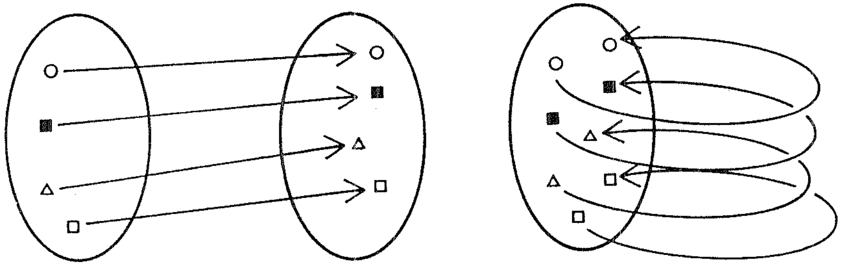

Figure 47 Suspension: a mathematical trick to turn a mapping (left) *into a flow in a space one dimension higher* (right)

the motion of a prune in a bowl of porridge being stirred by Little Baby Bear, and ask 'is there a periodic solution?', then instead of trying to solve the equations and examining the result for periodicity, you end up looking for Poincaré sections instead. 'Someone's been iterating *my* Poincaré mapping,' said Mummy Bear. You can imagine that the techniques involved are rather different.

Solenoids in Suspension

What has this to do with making the tenfold circle-wrapping mapping into respectable dynamics? Smale realized that you can work a Poincaré section backwards. Given a piece of surface – say a topological disc – and a mapping from the surface to itself, you can concoct a dynamical system for which it is a Poincaré section and the 'first return' map is the one you started with.

To do this, you introduce a new 'direction' which is like a circle that cuts the disc at right angles. An initial point on the disc flows off it, round this circle, but in such a way that when it next hits the disc it does so as prescribed by the original mapping from the disc to itself. This trick is called *suspension* (Figure 47). It's the sort of thing that's natural to a topologist asking general questions about flows in n-space, but wouldn't occur if you were a chemist trying to understand the dynamics of a nitroglycerine explosion. However, you can write down an explicit differential equation if you want one. In science, you normally start with a physical problem and extract a differential equation. But Smale moved into the Designer Differential Equation business. The subject has never been the same since.

The upshot of all this is that anything you can see in a mapping of n-dimensional space can also be seen in a flow in $(n+1)$-dimensional

space. Conversely, the way to understand flows in $(n+1)$-dimensional space is to look at mappings of n-dimensional space. In particular, flows in 3-space, not very well understood, reduce to mappings in 2-space, which we hope may be easier. Similarly flows in 4-space, which you have to work very hard even to think about, reduce to mappings in 3-space, where you can at least hope to draw pictures.

So instead of looking for a flow in 4-space, Smale looked for an unorthodox mapping in a 3-dimensional space which would have similar properties to our circle mapping when iterated. Here's what he found.

As Poincaré section, take the *interior of a solid torus*. A doughnut, American-style, with a hole. Dough included, this time we're not

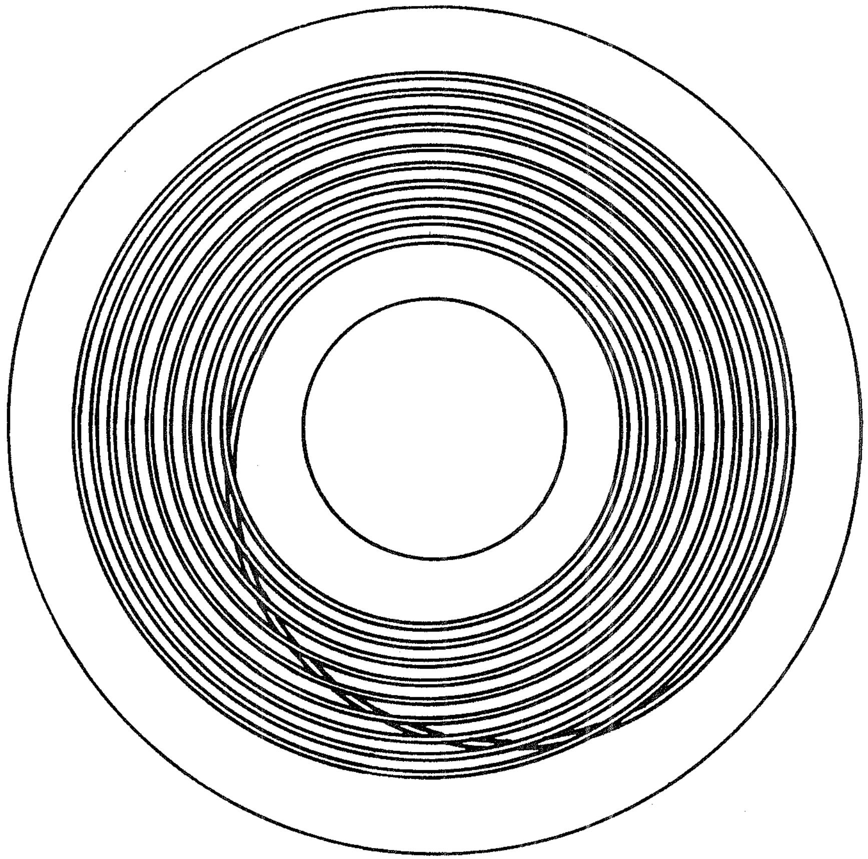

Figure 48 The tenfold wrapping applied to a solid torus to avoid self-intersections. Because the torus is three-dimensional, there is room for one winding to pass underneath the others without hitting them.