

JOHN GRIBBIN

# Deep Simplicity

*Chaos, Complexity and the Emergence of Life*



PENGUIN BOOKS

PENGUIN BOOKS

Published by the Penguin Group

Penguin Books Ltd, 80 Strand, London WC2R 0RL, England

Penguin Group (USA) Inc., 375 Hudson Street, New York, New York 10014, USA

Penguin Books Australia Ltd, 250 Camberwell Road, Camberwell, Victoria 3124, Australia

Penguin Books Canada Ltd, 10 Alcorn Avenue, Toronto, Ontario, Canada M4V 3B2

Penguin Books India (P) Ltd, 11, Community Centre, Panchsheel Park, New Delhi – 110 017,  
India

Penguin Books (NZ) Ltd, cnr Airborne and Rosedale Roads, Albany, Auckland 1310, New Zealand

Penguin Books (South Africa) (Pty) Ltd, 24 Sturdee Avenue, Rosebank 2196, South Africa

Penguin Books Ltd, Registered Offices: 80 Strand, London WC2R 0RL, England

[www.penguin.com](http://www.penguin.com)

First published by Allen Lane 2004

Published in Penguin Books 2005

11

Copyright © John and Mary Gribbin, 2004

All rights reserved

The moral right of the author has been asserted

Except in the United States of America, this book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser

ISBN: 978-0-14-104221-3

For Jim Lovelock

*It always bothers me that, according to the laws as we understand them today, it takes a computing machine an infinite number of logical operations to figure out what goes on in no matter how tiny a region of space, and no matter how tiny a region of time. How can all that be going on in that tiny space? Why should it take an infinite amount of logic to figure out what one tiny piece of space/time is going to do? So I have often made the hypothesis that ultimately physics will not require a mathematical statement, that in the end the machinery will be revealed, and the laws will turn out to be simple, like the chequer board with all its apparent complexities.*

Richard Feynman  
*The Character of Physical Law*

*The simplicity of nature is not to be measured by that of our conceptions. Infinitely varied in its effects, nature is simple only in its causes, and its economy consists in producing a great number of phenomena, often very complicated, by means of a small number of general laws.*

Pierre Laplace  
1749–1827

# Contents

*List of illustrations*

*Acknowledgements*

Introduction: The Simplicity of Complexity

1 Order out of Chaos

2 The Return of Chaos

3 Chaos out of Order

4 The Edge of Chaos

5 Earthquakes, Extinctions and Emergence

6 The Facts of Life

7 Life Beyond

*Further Reading*

*Index*

## List of Illustrations

- 1.1 An unpredictable collision
- 1.2 The reversibility of time
- 1.3 The arrow of time
- 1.4 Simple equilibrium
- 1.5 'Mexican hat' equilibrium
- 2.1 An unpredictable orbit
- 2.2 Unpredictable weather
- 2.3 The 'butterfly' diagram
- 3.1 The route to chaos
- 3.2 Period-doubling and chaos
- 3.3 The line that thinks it's a plane
- 3.4 The Cheshire Cat line
- 3.5 The Cheshire Cat triangle
- 3.6 Making the Sierpinski gasket
- 3.7 A fractal fern
- 3.8 A curve with no bends
- 3.9 The Koch snowflake
- .10 The Baker transformation
- 4.1 Convective rolls

4.2 How the leopard got its spots

5.1 The earthquake power law

5.2 Fractal coastlines

5.3 The noise of a quasar

5.4 The sounds of music

5.5 The pattern of death

5.6 The law of extinction

5.7 The law of avalanches

5.8 Networking and complexity

5.9 The secret of life

7.1 Signs of life

7.2 Daisyworld

7.3 Daisyworld with rabbits

7.4 The carbon dioxide/Ice Age connection

7.5 The nuclear physics of life

7.6 A self-portrait of life on Earth

## Acknowledgements

As with all my books, Mary Gribbin played a significant role in ensuring that the ideas presented here are intelligible to someone with no background in science. Her name would appear as co-author on the front of the book if it were not for the prejudice of our editor at Penguin against joint authorship.

We are grateful to the University of Sussex for providing a base with access to an excellent library and high-speed internet access, and to the Alfred C. Munger Foundation for a financial contribution towards our research and travel expenses.

The author and publishers are grateful to the following for permission to reproduce artwork:

**Figure 2.1** from Edward Lorenz, *The Essence of Chaos*, University of Washington Press, 1993; **Figures 2.3, 3.1, 3.4, 3.5, 3.7, 3.8, 3.9** from J. C. Sprott at <http://sprott.physics.wisc.edu/fractals/chaos>; **Figures 2.2, 3.10** from James Gleick, *Chaos*, Heinemann, 1988; **Figures 4.1, 5.7** from Per Bak, *How Nature Works*, Oxford University Press, 1997; **Figure 4.2** from James Murray, *Mathematical Biology*, Springer-Verlag, 1993; **Figure 5.2** from J. Feder, *Fractals*, Plenum, 1988; **Figure 5.4** from Philip Ball, *The Self-Made Tapestry*, Oxford University Press, 1999; **Figures 5.8, 5.9** from Stuart Kauffman, *At Home in the Universe*, Oxford University Press, 1995; **Figures 7.2, 7.3** courtesy Jim Lovelock.



Every effort has been made to trace all copyright holders, and the author and publishers will gladly rectify in future editions any errors or omissions brought to their attention.

## INTRODUCTION

# The Simplicity of Complexity

The world around us seems to be a complex place. Although there are some simple truths that seem to be eternal (apples always fall to the ground, not to the sky; the Sun rises in the east, never in the west), our lives, in spite of modern technology, are still, all too often, at the mercy of complicated processes that produce dramatic changes out of the blue. Weather forecasting is still as much an art as a science; earthquakes and volcanic eruptions strike unpredictably, and seemingly at random; stock-market fluctuations continue to produce boom and bust with no obvious pattern to them. From the time of Galileo (in round numbers, the beginning of the seventeenth century) science made progress – enormous progress – largely by ignoring these complexities, and focusing on the simple questions, looking to explain why apples fall to the ground, and why the Sun rises in the east. Progress was so spectacular, indeed, that by about the middle of the twentieth century all the simple questions had been answered. Concepts such as the general theory of relativity and quantum mechanics explained the overall workings of the Universe on the very large and very small scales respectively, while the discovery of the structure of DNA and the way in which it is copied from generation to generation made life itself, and evolution, seem simple at the molecular level. And yet, the complexity of the world at the human level – at the level of life – remained. The most interesting question of all, the question of how life could have emerged from non-life, remained

unanswered.

It is no surprise that the most complex features of the Universe, which proved most reluctant to yield to the traditional methods of scientific investigation, should exist on our scale. Indeed, we may be the most complex things there are in the Universe. The reason is that on smaller scales entities such as individual atoms behave in a relatively simple way in their one-to-one interactions, and that complicated and interesting things are produced when many atoms are linked together in complicated and interesting ways, to make things like people. But this process cannot continue indefinitely, since if more and more atoms are joined together, their total mass increases to the point where gravity crushes all the interesting structure out of existence. An atom, or even a simple molecule like water, is simpler than a human being because it has little internal structure; a star, or the interior of a planet, is simpler than a human being because gravity crushes any structure out of existence. And that is why science can tell us more about the behaviour of atoms and the internal workings of the stars than it can about the way people behave.

As the simple problems yielded to investigation, it was natural that some scientists should try to tackle the more difficult puzzles associated with complex systems. As we shall see, although some valiant individual assaults on these puzzles had been made earlier, it only really became possible to begin to understand the workings of the world on the complex human scale in the 1960s, with the advent of powerful and fast (by the standards of those days) electronic computers. These new developments began to emerge into the awareness of a wider public in

the middle to late 1980s, with the publication first of the now classic book *Order out of Chaos*, by Ilya Prigogine and Isabelle Stengers, and then of James Gleick's *Chaos*. At the time, I was busily writing about the great triumphs of the old science, and although I occasionally tried to get to grips with the ideas of chaos and complexity, I found that doing so made my head hurt, so mostly I kept my distance. But after about ten years of waiting either for chaos theory to go away or for someone to write a book explaining it in language I could understand, I decided that if no one else was going to explain it in clear language, then I would have to. Which meant that I had to read everything I could find on the subject and try to understand it myself. It was while doing this that I discovered that it wasn't so difficult after all. Both relativity theory and quantum theory were regarded, when new, as ideas too difficult for anyone except the experts to understand – but both are based on simple ideas which are intelligible to the lay person willing to take the mathematics on trust. I ought not to have been surprised to find that the same is true of chaos and complexity – but I was, and I clearly recall the moment when I finally got the message about what it was all about. As I understood it, what really mattered was simply that some systems ('system' is just a jargon word for anything, like a swinging pendulum, or the Solar System, or water dripping from a tap) are very sensitive to their starting conditions, so that a tiny difference in the initial 'push' you give them causes a big difference in where they end up, and there is feedback, so that what a system does affects its own behaviour. It seemed too good to be true – too simple to be true. So I asked the cleverest person I know, Jim Lovelock, if I was on the right lines. Was it really true, I asked, that all this business of chaos and complexity is

based on two simple ideas – the sensitivity of a system to its starting conditions, and feedback? Yes, he replied, that’s all there is to it.

Now, that’s a bit like saying that ‘all there is’ to the special theory of relativity is that the speed of light is the same for all observers. It is true, it is simple, and it is easy to grasp. Nevertheless, the complexity of the structure built on that simple fact is staggering, and needs some mathematical knowledge to appreciate it fully. But I knew from experience that I could explain the essence of relativity to people with no background in science, and the realization that there were similar simple truths underpinning chaos and complexity gave me confidence to try to do the same in that area. The result is in your hands, an attempt to take the obvious step of trying to explain chaos and complexity the simple way, from the bottom up – for everybody. The great insight is that chaos and complexity obey simple laws – essentially, the same simple laws discovered by Isaac Newton more than three hundred years ago. Far from overturning four centuries of scientific endeavour, as some accounts would lead you to believe, these new developments show how the long-established scientific understanding of simple laws can explain (although not predict) the seemingly inexplicable behaviour of weather systems, stock markets, earthquakes, and even people. As we hope to convince you, the study of chaos and the emergence of complexity from simple systems is now on the brink of explaining the origin of life itself. In a phrase attributed to Murray Gell-Mann, but echoing the speculation of Richard Feynman quoted at the front of this book, the complicated behaviour of the world we see around us – even the living world – is merely ‘surface complexity arising out of deep simplicity’.<sup>1</sup> And it is the simplicity that underpins complexity that is the theme of this book.

*John Gribbin,  
January 2003*

# 1

## Order out of Chaos

Before the scientific revolution of the seventeenth century, the world seemed to be ruled by chaos in a quite different way from the way the term is used by scientists today, but in the same way that most people still apply the word. There was no suggestion that there might be simple, orderly laws underpinning the confusion of the world, and the nearest anyone came to offering a reason for the behaviour of wind and weather, the occurrence of famines or the orbits of the planets, was that they resulted from the whim of God, or the gods. Where order was perceived in the Universe, it was attributed to the response of physical objects to a need for harmony and order to be preserved wherever possible – the orbits of the planets and the Sun around the Earth (thought to be at the centre of the Universe) were supposed to be circles, because circles were perfect; things fell downwards because the centre of the Earth was at the centre of everything, the centre of symmetry in the Universe, and therefore the most desirable place to be. Even when the philosopher Aristarchos of Samos, who lived in the third century BC, dared to suggest that the Earth moved around the Sun, he still imagined that it must follow a circular orbit.

These examples highlight an absolutely crucial difference between the science of the Ancients and the science of post-Galilean times. The

Ancient Greeks were superb mathematicians, and in particular they were superb geometers, who had a very good understanding of the relationships between stationary things. This geometry had its roots in even earlier cultures, of course, and it is easy to imagine how this first science may have arisen out of the practicalities of life in the developing agricultural societies of pre-history, through the problems associated with building houses and laying out towns, and the need, as society became more complicated, to divide up land into fields. But the Ancients had no understanding at all of how things move, or the laws of motion. You have only to look at how puzzled they were by Zeno's famous paradoxes, such as the soldier who can never be killed by an arrow. If he runs away, then by the time the arrow reaches the position he was in he has moved; in the time it takes the arrow to cover that extra distance, he can move a little further; and so on.

In spite of the existence of people like Aristarchos, the Earth-centred Universe remained the established image (what scientists would now call a 'model') even after Nicolaus Copernicus published his model of a Sun-centred Universe (but one still based on circles) in 1543. His book, *De Revolutionibus Orbium Coelestrum*, had been essentially completed in 1530, and much of its contents were widely discussed before publication, leading Martin Luther to comment in 1539 that 'This fool wishes to reverse the entire science of astronomy; but sacred Scripture tells us that Joshua commanded the Sun to stand still, and not the Earth.' Responding to similar criticisms, Galileo later riposted: 'The Bible shows the way to go to Heaven, not the way the heavens go.' It was Galileo's contemporary, Johannes Kepler, using observations painstakingly compiled by Tycho Brahe, who established, for those with open eyes,



that not only did the planet Mars move round the Sun, but that it did so in an elliptical orbit, pulling the rug from under the notion that the kind of circular perfection beloved of the Ancient Greeks ruled the cosmos.

Even to people who know little about science, or the history of science, Galileo (who lived from 1564 to 1642) is famous today as the man who turned one of the first telescopes on the heavens, found evidence to support the Sun-centred Copernican model, and had a run-in with the Catholic Church, which led to his conviction for heresy and the suppression of his books in Catholic countries – which (of course) led to them selling like hot cakes everywhere else. But he did much more than this. It was Galileo, more than anyone, who laid down the principles of the scientific method of investigation, involving comparing theories (or models) with the outcome of experiment and observation, and it was Galileo who first got to grips with motion in a scientific way.

The key to Galileo's work on motion was a discovery he made while a medical student in Pisa in 1583. During a boring sermon in the cathedral there, he watched a chandelier swinging to and fro, and timed the swing with his pulse. Galileo realized that the time it took for the lamp to complete one swing was the same whether it swung through a wide arc or a shallow one, and later experiments showed that the time taken for a pendulum to swing depends on its length, not on how far it swings. This is the basis of the pendulum clock, but even without going so far as to build a clock (he did design one, later built by his son), Galileo was able to use a pendulum as an accurate timekeeper when he later carried out experiments to study the behaviour of balls rolling down a ramp. These experiments provide another insight into both Galileo's mind and the scientific method. He wanted to study falling

objects, to investigate the effect of gravity on motion. But falling balls moved too fast for him to keep track of. So he rolled the balls down an inclined ramp, realizing that this gave him a stretched-out and slowed-down version of the way balls fall under gravity. Through these experiments, Galileo developed the idea of acceleration. The velocity (or speed) of an object tells you how far it moves in a certain amount of time – say, one second. A constant velocity of 9.8 metres per second means that in every second the moving object covers a distance of 9.8 metres. But Galileo found that falling objects (or balls rolling down a ramp) move faster and faster, with the speed increasing each second. Crucially, his experiments showed that the speed increases by the same amount every second. This is uniform acceleration, and a uniform acceleration of 9.8 metres per second *per second* means that, starting from rest, in the first second the object reaches 9.8 metres per second, in the second it speeds up to 19.6 metres per second, in the third 29.4 metres per second, and so on. I have chosen this particular example because 9.8 metres per second per second is, indeed, the acceleration caused by gravity for a falling object at the surface of the Earth; because of the way time comes into the calculation twice, this is called a second order effect, with velocity a first order effect. And this acceleration due to gravity explains why pendulums behave as they do, neatly closing the loop.

Galileo did something else – something central to the story we tell in this book. He realized that the balls rolling down his inclined planes were being slowed down a little by friction. In fact, what he measured was not a perfectly uniform acceleration. But he took the dramatic and influential leap, astonishing for his time, of extrapolating from his actual

observations to work out how his balls would move *without* the effect of friction, on some idealized, perfectly slippery slope. This kind of extrapolation would be at the heart of the scientific investigation of the world for the next four centuries. When scientists – physicists in particular – tried to describe the world in terms of mathematical laws, they formulated those laws to describe the behaviour of mythical objects such as perfectly hard spheres, which bounce off one another without being deformed and roll along surfaces without feeling friction, and so on. But, unlike the Ancient Greek philosophers, they knew that their image of perfection did not represent the real world. Armed with those equations, they could then try to put in extra terms, correction factors, to take account of the imperfections of the real world, allowing, say, for the effect of air resistance on a falling object. Air resistance explains why on Earth a hammer and a feather fall at different rates, while on the airless Moon, as the Apollo astronauts demonstrated, they fall at the same rate.

All of this helped Galileo to cast out of science another aspect of the geometrical perfection which his predecessors had imagined in the real world. Before Galileo, it was thought that when a cannon fired its ball at some angle above the horizontal, the flight of the ball would consist of a straight line as it left the muzzle, then it would follow the arc of a perfect circle for a time, then it would fall vertically to the ground. Only the imagined perfection of straight lines and circles was involved in the motion. Applying his discovery that gravity produces a constant downward acceleration on the cannonball, and allowing for the initial velocity of the ball out of the muzzle, Galileo showed that the flight of the ball must actually be a single smooth curve, part of a parabola, all

the way to its target. The same calculations showed that the maximum range for the cannon (assuming the same charge of gunpowder and weight of shot) would always be achieved when it was fired at an angle of 45 degrees upwards from the horizontal. These were practical matters of great importance in the turbulent times Galileo lived in, and this kind of military work helped to establish his early reputation. Whatever philosophers and theologians might say about perfection, armies in the field had no time to quibble about the desirability of circular motion; all they wanted to know was which way to point their guns to achieve maximum effect, and Galileo told them.

It was a combination of Kepler's discovery of elliptical orbits and Galileo's insights both into acceleration and the scientific method that paved the way for the greatest scientific discovery of the seventeenth century, and perhaps of all time: Isaac Newton's universal law of gravitation. Newton was born in 1642 and died in 1727. His great work *Philosophiae Naturalis Principia Mathematica*, or just the *Principia* for short, was written over the years 1684 to 1687, and published in 1687, but was based on ideas that he had developed twenty years earlier, when he was a young Cambridge graduate about to become a Fellow of Trinity College, and had been forced to spend months at his mother's home in Lincolnshire when the University was closed because of the plague. Like Galileo, Newton stressed the importance of comparing theories and models with experiments and observations of the real world, and always carried out relevant experiments himself, whenever possible, to test his ideas. This is so deeply ingrained as part of the scientific method today that it may seem obvious, even to non-scientists, and it is hard to appreciate the extent to which, even into the seventeenth century, many

philosophers would speculate about the nature of the physical world in an abstract way, without ever getting their hands dirty in experiments. The classic example is the argument about whether two different weights dropped from the same height at the same time would hit the ground together – an argument that rumbled on for decades among those abstract thinkers even after a Flemish engineer, Simon Stevin, actually did such experiments with lead weights dropped from a height of about 10 metres, found that they fell at the same rate, and published his results in 1586.<sup>1</sup>

Newton also took on board and refined Galileo's insight into the value of deliberately simplified models (such as Galileo's frictionless planes) as descriptions of particular aspects of the real world. It is a key feature of Newton's work on gravity and orbits, for example, that in his calculations of the effects of gravity he treats an object such as Mars, or the Moon, or an apple, as if all its mass were concentrated at a single point, and that provided you are outside the object of interest, its gravitational influence is measured in terms of your distance from that point, the centre of mass of the object (which is also the geometrical centre if the object is a sphere). In the *Principia*, Newton proved that this is the case for spherical objects. He knew that the Earth is not precisely spherical (indeed, he was able to calculate how much the Earth bulges at the equator because of its rotation); but he felt it was reasonable to assume a spherical Earth (and a spherical Sun, a spherical Mars, and so on) as a first approximation, and calculate the orbits accordingly. As it happens, later calculations have established that, provided you are far enough away from them, even very irregularly shaped objects do act, gravitationally speaking, as if all their mass were concentrated at a

point, but this doesn't diminish the importance of the idea of using idealized approximations to reality where it is necessary, or helpful, to make the calculations more straightforward.

There is, though, rather more to this particular story than meets the eye. In the *Principia*, Newton proved the validity of treating the gravitational influence of a spherical object as if all its masses were concentrated at its centre, using geometrical techniques that would have been understood by the Ancient Greeks, and were certainly familiar to Newton's contemporaries. This is very difficult. But we now know that well before he wrote the *Principia* Newton had developed (or discovered) the mathematical technique now called calculus, and that this proof is very easy using calculus. Some scholars suspect that Newton actually solved the problem using calculus first, and then went through the painstaking process of translating everything into classical terminology, to make sure that his contemporaries would understand it. If so, he may, in a sense, have shot himself in the foot, because by keeping quiet about his new mathematical technique he paved the way for a bitter wrangle with the German Wilhelm Leibniz, who invented the technique independently (and gave it the name by which it is still known). Leibniz had the idea a little later than Newton did, but had the good sense to publish his work, which is partly why there were bitter rows about priority (the bitter rows were also partly because both protagonists were unwilling to make concessions, and Newton in particular was a rather unpleasant person, who had an arrogant belief in his own abilities and a spiteful reaction to anyone he perceived as an opponent). But the arguments over priority are of little concern here. What matters is that calculus is a technique that makes it possible to break up problems into

tiny components that can be manipulated mathematically and the results added up to give a solution to the overall problem. In the case of the gravitational influence of a spherical object, for example, the sphere can be treated as being divided up (differentiated) into an infinite number of infinitesimally small pieces of matter, and an equation describing the gravitational influence of a typical such piece of matter in terms of its position in the sphere written down. The overall effect of all the pieces of matter working together can then be added up, or integrated, from this equation.<sup>2</sup>

The same sort of process can be applied to time. The flight of an arrow, for example, can be differentiated into a description of how the arrow is moving at any of an infinite number of points along its path, and the same thing can be done for the path of a runner trying to escape the arrow. When the two differential equations are integrated, they will tell you exactly where the arrow meets the runner, with no paradox. When the calculus became widely understood, it seemed that Newton and Leibniz had tamed time, making it possible to describe the behaviour of moving things with the precision that the Ancient Greeks had been able to apply to their description of the relationships between static things. At least, in principle. At the level of single planets orbiting single stars, or of single runners trying to escape from single arrows, calculus is easy; like other things, it gets complicated when we are dealing with more complicated systems, although the same basic rules apply. It is usually straightforward to write down a set of differential equations to describe such a system; it is solving (integrating) the equations that can cause problems.

Whatever the mathematical techniques he used in private in the

1680s, in the *Principia* itself Newton proved, using techniques familiar to his contemporaries, that in order for the planets to follow elliptical orbits around the Sun (and therefore to match observation with theory), gravity must follow an inverse square law. Specifically, the force of attraction between any two masses a certain distance apart (the distance being measured from the centre of mass of each object) is proportional to the two masses multiplied together and divided by the square of the distance between them (hence ‘inverse square’ law). The constant of proportionality that comes into the calculation is the constant of gravity, a number which tells us how strong gravity is, and is denoted by G. Using obvious symbols,  $F = \frac{GMm}{r^2}$ .

This was not a complete surprise to Newton’s contemporaries. People such as Robert Hooke, Edmond Halley and Christopher Wren, all Fellows of the Royal Society, already knew that an inverse square law would produce elliptical orbits. Newton’s triumph was to prove that *only* an inverse square law would do this, and that therefore gravity *must* obey an inverse square law. But he went further. His work implied that this was a *universal* law. It wasn’t something that applied only to apples falling from trees, or to the Moon in orbit around the Earth, or the planets in orbit around the Sun. And it was the same all the time – it didn’t only apply in summer, or when there was an ‘r’ in the month. It applied to the force of attraction exerted by any piece of matter, anywhere in the Universe, on any other piece of matter, anywhere else in the Universe. Newton made the Universe seem an orderly place, with no room for interference from capricious gods.<sup>3</sup> And he provided three other laws, the laws of motion which describe the behaviour of moving objects in the laboratory, or in the world at large, or in the Solar System



and beyond, which, by extension, must also be thought of as universal laws, applying everywhere and at all times.

These are Newton's laws of motion, which underpin three hundred years of science, but which can be summarized very simply, and which again highlight the development of the scientific way of looking at the world. Newton's first law says that every object stays still or keeps moving in a straight line unless it experiences – or feels – a force. The 'staying still' part matches common sense (by which we mean things that we experience every day here on Earth). Things don't move unless something makes them move. But the constant motion in a straight line business flies in the face of common sense. When we set things moving here on Earth, left to their own devices they fall to the ground and stop. Newton's insight, extending the ideas of Galileo, was that they only fall down because of an external force (gravity), and they only stop moving because of other forces (friction). Without ever having seen the images familiar to us today of objects obeying Newton's laws inside the zero-gravity and almost frictionless environment of the interior of a spacecraft, Newton appreciated what the world would be like if friction and other outside forces could be ignored.

The second law tells us how much the motion of an object is affected by a force applied to it. It says that a force applied to a mass causes an acceleration given by the simple equation  $F = ma$ , or  $a = \frac{F}{m}$ . It is this law, plus the law of gravity, which explains, or describes, the orbit of a planet around the Sun. Together, the second law and the law of gravity also explain the behaviour of falling objects observed and described by Galileo. If the Earth is the mass  $M$ , then the gravitational force acting on any mass at the surface of the Earth is proportional to its own mass,  $m$

(since  $F = \frac{GmM}{r^2}$ ). But the acceleration that is caused by this force is determined by dividing. So the mass  $m$  cancels out of the equation, and all objects falling under the influence of gravity accelerate at the same rate near the surface of the Earth. They all accelerate at a different rate (but still the same as each other) at the surface of the Moon, and so on.

The third law says that every force acting on an object produces an equal and opposite force – or in Newton’s language, for every action there is an equal and opposite reaction. A simple example of this is when the force used in a rifle to send a bullet speeding on its way produces a reaction which is felt as the kick of the rifle against the sharp-shooter’s shoulder. If you thump a desk with your fist, you can feel the reaction equally obviously. Slightly less obviously, when the Sun tugs on a planet through gravity, the planet tugs on the Sun with an equal and opposite force, as if the two were joined by a piece of stretched elastic. Even an apple falling from a tree is pulling the Earth up to meet it, by a tiny amount (only a tiny amount, because although the same force applies to both the apple and the Earth, acceleration is proportional to the force but *inversely* proportional to mass, and the Earth is a whole lot more massive than an apple). As far as orbits are concerned, this means that instead of the Moon, say, being in orbit around the Earth, actually both the Moon and the Earth are in orbit around their mutual centre of mass (which, because the Earth is much more massive than the Moon, happens to lie below the surface of the Earth).

The laws apply equally well to planets and moons moving through space, and (making due allowance for friction) to events on Earth, where a good approximation to Newton’s laws in action can be seen on a pool table. But for now we’ll concentrate on the behaviour of the planets,

because it was here that problems with Newton's description of the world began to show up as soon as he developed these ideas. The problems were largely swept under the carpet for centuries, but it was in the realm of orbital mechanics, as it is known, that the first hints of chaos in the modern scientific sense eventually emerged.

Part of the problem is that although Newton's laws (including the law of gravity) can be used perfectly to calculate the orbits of two objects around each other in an otherwise empty Universe (the Moon around the Earth, or the Earth around the Sun, and so on), they cannot give exact answers for calculations involving three or more objects moving under their mutual gravitational attraction (such as the Earth, Moon and Sun moving together through space). This is called the 'three-body problem', although it applies to any number of 'bodies' bigger than two; more generally, physicists sometimes refer to the ' $N$ -body problem', where  $N$  can be any number bigger than two. The equations describing such systems can be written down, but they cannot be solved – they are not integrable and are said to have no analytical solutions (meaning that the techniques of mathematical analysis do not work with them). Equations that can be solved analytically are sometimes said to be deterministic; the equation describing the orbit of a single planet around the Sun is deterministic, and has analytical solutions in the form of ellipses. It's important to appreciate, though, that the lack of solutions to the three-body problem is not caused by our human deficiencies as mathematicians; it is built into the laws of mathematics.

There are, however, often ways around the problem. These involve approximations. For literally three bodies in orbit around one another, it is possible to carry out the calculation in repeated steps, first pretending

one of the bodies is still while you carry out the calculations to work out how the other two bodies will move a little way in their orbits. Then, starting from these new positions, another of the three bodies is held still, and the calculation carried out for the other two, and so on. The answers you get are never perfect, because in reality all three objects are moving at once. But if you make the time step involved in this repeated process (such a repeated mathematical manipulation is sometimes called iteration) small enough, in many cases the orbits you calculate will be very close to the real orbits. In the Solar System, the Sun is so much bigger than any of the planets (or, indeed, than all the planets put together) that its gravitational influence dominates, and to a first approximation the influences of the other planets can be ignored. So the orbit of Mars, for example, is calculated by pretending that the other planets do not exist, giving a perfect ellipse. The real orbit of Mars differs slightly from this analytical solution to the equations of motion, but the differences can be explained by allowing for the perturbing effects of other planets, notably the giants Jupiter and Saturn. Similarly, the orbit of the Moon around the Earth can be calculated, to a first approximation, by ignoring the effects of the Sun, which is a long way away, and then adding these in afterwards as a correction to the calculation. If you are prepared to carry out a laborious series of calculations for each successive correction involving all the planets and their influences on each other (relatively easy today using high-speed computers), you can get an approximation which is very good indeed as a prediction of the actual observed orbit of the planet or moon you are interested in. But you can *never* get a precisely accurate analytical mathematical solution which would enable you to predict how that

planet or moon will move in its orbit for ever. There is always some margin of error. And the three-body problem is completely intractable if the three objects are about the same size (that is, mass) as each other and are at roughly equal distances from each other. The absence of analytical solutions means that nature itself does not 'know' exactly how the orbits will change (evolve) as time passes. Even for the planets of the Solar System, there is always the possibility that the orbits may not stay the same for ever as they are today, but may change in literally unpredictable ways.

Newton was aware of this, but was a deeply religious man (although his religious beliefs did not entirely match those of orthodox Church teaching at the time), and solved the problem, in his own mind at least, by suggesting that if the planets began to wander from their familiar orbits (perhaps spiralling in towards the Sun or out into the depths of space) God could always intervene to put them back on course. This drew a cutting riposte from Leibniz, who picked up on the analogy between Newton's orderly, deterministic Universe and a clock to sneer that Newton's God must be a pretty poor watchmaker if he was incapable of making a clock that would not keep steady time unless He intervened whenever it broke down.

The puzzle persisted until the end of the eighteenth century, when the French mathematician Pierre Laplace (1749-1827) seemed to restore order to the Solar System. He concentrated first on calculating (using the laborious step-by-step process outlined earlier) the orbits of Jupiter and Saturn, which are the two largest planets in the Solar System, and exert the largest gravitational influence, both on each other and on the smaller planets, after the Sun itself. Laplace found that at present the orbit of

Jupiter is expanding slightly, while the orbit of Saturn is shrinking – just the sort of effect that Newton was worried about. But Laplace discovered that these changes are linked to a rhythmic change in the gravitational influence of the two planets on each other, which operates because Saturn makes two orbits of the Sun while Jupiter travels around its own orbit five times. This means that the two giant planets are at their closest to each other every fifty-nine years. Using Newton’s laws and the step-by-step iterative technique, Laplace calculated that the effect of all this was to reverse the overall changes seen in the orbits of the two planets every 929 years – after 929 years with the orbit of Jupiter expanding and the orbit of Saturn shrinking, there is an interval of 929 years with the orbit of Jupiter shrinking while the orbit of Saturn expands, and so on. Laplace thought that he had restored order to the Solar System, and famously remarked to Napoleon that he ‘had no need’ of the God ‘hypothesis’, writing:

The irregularities of these two planets appeared formerly to be inexplicable by the law of universal gravitation – they now form one of its most striking proofs.

As we shall see, he was only right up to a point. But he showed that the same kind of stability applied to other planets, and by implication to the Solar System as a whole, so that from the early nineteenth century onwards Newton’s laws really did seem to imply that the Solar System itself and the Universe at large operated with the stability and reliability of a perfect clock, requiring no outside intervention to maintain perfect time. The success of Newton’s laws enabled physicists to solve many problems, and is the rock upon which the whole of modern science is founded. Although physicists were aware, in the nineteenth and twentieth centuries, that there are many situations in which there are no

analytical solutions to the relevant equations, they could make such huge progress by solving the equations where they were deterministic, and by applying approximation techniques in other cases, that the more intractable puzzles were largely ignored. It is only natural, after all, to solve all the easy puzzles first, before worrying about the difficult ones. But just a few people did worry about problems that lay beyond the scope of Newton's equations (at least to the extent of pointing out that such problems existed), and another kind of three-body problem highlights just how limited Newton's laws really are.

This is a situation which can be visualized in terms of collisions between three balls on a pool table. If one moving ball strikes another stationary ball (or even if both balls are moving), then Newton's laws can be used to work out exactly how the two balls move after the collision, provided we know the masses of the balls, the speed the moving ball is travelling at, and make the usual allowances for friction (or ignore it). But if two stationary balls are arranged so they are touching each other, and the moving ball hits them sideways on, striking both the stationary balls simultaneously, in general Newton's laws cannot tell you how the balls will move apart after the collision. <sup>4</sup> Provided one or the other of the two touching balls is hit first, even if only by a tiny fraction of a second, the laws tell you where the three balls will go. But for a genuinely simultaneous collision, they do not. You might try to weasel your way out of this difficulty by arguing that in practice it is highly unlikely that the collision will be literally simultaneous, but it is still a worrying feature of the world that there are simple situations which, according to the laws that proved so successful when applied to everything from explaining the swing of a simple

pendulum to the flight of a spacecraft to the Moon, have unpredictable outcomes.

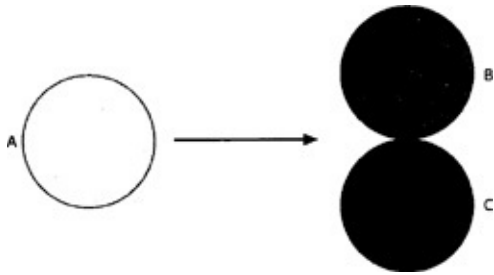


Figure 1.1 According to Newton's laws, if one perfectly elastic sphere strikes two touching perfectly elastic spheres simultaneously, it is impossible to predict where the three spheres will go.

Nobody really worried about this in the nineteenth century. They were too busy applying both Newton's laws and the scientific method developed by Galileo and Newton to produce an overall picture of an orderly, deterministic Universe. The prevailing image was the one encapsulated by Laplace in 1814 in his *Essai phibsophique sur les probabilites*:

Consider an intelligence which, at any instant, could have a knowledge of all forces controlling nature together with the momentary conditions of all the entities of which nature consists. If this intelligence were powerful enough to submit all this data to analysis it would be able to embrace in a single formula the movements of the largest bodies in the universe and those of the lighter atoms; for it nothing would be uncertain; the future and the past would be equally present to its eyes.

The modern equivalent would be a supercomputer which knew the positions and velocities of every particle in the Universe, and could use Newton's laws and the laws which describe the forces of nature (such as gravity and electromagnetism) not only to predict the future course of every particle, but to work out the entire history of where they had come from – for there is nothing in Newton's laws to reveal the direction of time, and the laws work just as well if time is running the other way, as you can easily see by imagining the reverse of a collision between two



... you can easily see by imagining the reversal of a collision between two pool balls, or reversing the motion of every planet in the Solar System in its orbit. There is no arrow of time in Newton's laws, and they seemed, to Laplace and many others, to describe a completely deterministic world in which both past and future are rigidly fixed and there is no scope for free will. What none of these people seems to have noticed is that the basic argument breaks down if, at any time and any place in the Universe, there is a simultaneous collision between three particles – although whether that would be enough to restore free will is something I shall leave for the philosophers to argue about.

The same puzzle about time emerged from one of the greatest triumphs of nineteenth-century physics: the understanding of light and other forms of electromagnetic radiation, which culminated in the work of the Scot James Clerk Maxwell (1831–79). Maxwell's description of electromagnetic radiation built on the work of Michael Faraday, who lived from 1791 to 1867, and came up with the idea of electric and magnetic 'fields' which surround magnets or electrically charged objects. Such a field can be made visible in the classic schoolroom experiment where iron filings are sprinkled on a sheet of paper held flat on top of a magnet; when the paper is gently tapped, the iron filings arrange themselves in lines linking the two magnetic poles of the magnet, marking out so-called lines of force, associated with the magnetic field. It was Faraday who first suggested that light might be produced by some sort of vibration in the lines of force associated with magnets and charged particles, twanging like plucked violin strings. But Faraday lacked the mathematical expertise to develop this idea into a fully worked-out model, and it was Maxwell who finished the job in the

1860s. He showed that all electric and magnetic phenomena known at the time, including the behaviour of light, could be described by a set of just four equations, now known as Maxwell's equations. These equations did for electromagnetic fields and radiation what Newton's laws did for solid matter – in principle, they described everything and made it possible to calculate the consequences of any electromagnetic interaction, although this might require a great deal of patience in complicated situations. Maxwell's equations describe how dynamos and electric motors work, why a compass needle points north, how big the force is between two electric charges a certain distance apart (it was already known that both electricity and magnetism, like gravity, obey inverse square laws), and much more. Together, Newton and Maxwell provided the mathematical tool kit needed to deal with everything known to physical science in the middle of the nineteenth century. And the most wonderful thing about Maxwell's equations was that they produced a description of light without being asked -the equations were set up to describe other electromagnetic phenomena, but they included within themselves a solution which described an electromagnetic wave travelling through space at a certain speed. That speed is exactly the measured speed of light (already well determined by the 1860s, and soon to be measured even more accurately), leaving no doubt that light travels in the form of an electromagnetic wave.

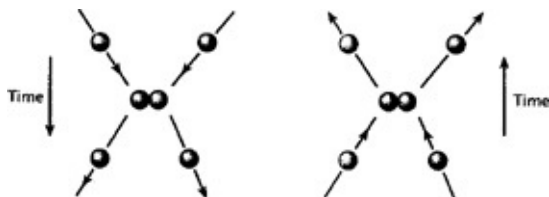


Figure 1.2 According to Newton's laws, collisions between pairs of objects are perfectly reversible. The

*picture looks equally plausible whichever way we draw the 'arrow of time'.*

But there are two curious features of Maxwell's equations, one of which was soon to have a profound impact on physics, and the other which was treated only as a minor oddity until quite recently. The first strange feature of the equations is that they give the speed of light as a constant, independent of how the source of the light is moving relative to the person (or machine) that measures its speed. According to the equations, if you shine a torch at me we both measure the speed of the light from it as this constant,  $c$ . That's fine if we are both standing still. But if I am moving past you, or you are moving past me, even at a very high speed, both you and I still measure the speed of light from the torch as  $c$ . This is true when I am approaching you (you might think, from common sense, that I ought to measure a speed equal to  $c$  plus my own speed relative to you) and when I am moving away from you (when common sense says that I should measure a speed of  $c$  minus my own speed relative to you). It is this requirement from Maxwell's equations that the speed of light is a constant for all observers, no matter how they are moving, that led Albert Einstein to develop the special theory of relativity, published in 1905, and move on to the general theory of relativity, published a decade later. Relativity theory, however, has no place in this book, except to acknowledge that it does exist and provides a more complete description of the behaviour of moving objects and gravity than Newtonian theory does. It does not overturn, or supplant, Newtonian physics, but includes Newton's ideas within itself and extends the description of the physical world to conditions where Newtonian theory is inadequate – particularly for objects moving at very high

speeds and objects in strong gravitational fields. But as a description of the motion of the Earth around the Sun, for example, Newtonian physics is accurate to within one part in a hundred million, and it is even more accurate as a description of events on the human scale.

Maxwell's equations also have their limits, particularly in the description of events on very small scales, such as the behaviour of atoms and the particles that make up atoms. There, both the classical (Maxwell) description of electromagnetic interactions and the classical (Newton) description of particle interactions have to be modified, and the rules of quantum physics apply. Quantum effects – or at least, one particular quantum effect – will come into our story later, although only in a relatively minor role. Almost everything we deal with here can be described (and perhaps understood) in terms of Newtonian mechanics. Even Maxwell's equations are only really worth mentioning because of their other odd feature.

The other strange feature of Maxwell's equations is that, like Newton's equations, they have no inbuilt arrow of time. This is not too worrying when we think about things like electrically charged particles moving in magnetic fields, and imagine reversing the direction of time. Everything would still obey the laws of physics described by Maxwell's equations and Newton's laws. But according to the equations, there is no difference between a light wave spreading out from a lamp and its time-reversed counterpart, a light wave converging on a lamp. That would be as strange, to us, as to see the scattered balls on a pool table assemble themselves into a neat triangle, using energy gained from friction, and shoot the white ball neatly on to the end of the player's cue, giving it a kick. Yet both possibilities are allowed by the combined laws of

Newton's and Maxwell's equations. Clearly, there is something odd about the arrow of time.

What was for a long time the standard explanation of why we see a preferred direction of time stems from another great triumph of nineteenth-century physics, the description of the relationship between heat and motion (thermodynamics). This was of fundamental practical importance in the industrial world when that industry was being powered by steam engines. From our point of view, the importance of the science of thermodynamics is that it provided physicists with a way to describe the behaviour of very large numbers of objects – in particular, particles of gas – in some sense working together in a complex system. This involves averaging and statistics, but it is largely based on the idea that a gas is made up of a myriad tiny particles (atoms and molecules) bouncing around and colliding with one another and the walls of their container in accordance with Newton's laws of motion. This kinetic theory of gases was a significant example of how the universal laws of physics brought order out of chaos. The very name 'gas' was coined by the Flemish physician Joannes (originally Jan) van Helmont, from the Greek word for chaos; the term first appeared in print in van Helmont's book *Ortus Medicinae*, published in 1648, four years after the death of Joannes, by his son Franciscus. The idea of gas as chaos remained appropriate for three hundred years, until Maxwell, working in Britain, and his contemporary Ludwig Boltzmann, working in Vienna, put the kinetic theory (previously a speculative idea) on a secure scientific footing based on Newton's laws. The pressure a gas applies to the walls of its container, for example, is explained in terms of action and reaction (Newton's third law) as the particles of the gas collide with

the wall and bounce off, back into the volume of the gas. This produces what feels like a steady pressure simply because there are so many particles involved, producing an enormous number of collisions every second. In round figures, the number of molecules in a box of gas the size of a matchbox is roughly one million billion billion (a 10 followed by 23 zeros), a number so big that it doesn't really make much difference if we think about a box ten times bigger or ten times smaller. A typical molecule of air in the atmosphere at sea level and at a temperature of 0 °C experiences nearly four million collisions every second.<sup>5</sup> All this is what gives the illusion that air is a continuous medium; and it also means that, whatever Laplace might have imagined, it would be completely futile to try to work out the behaviour of a gas by calculating the trajectories of every single molecule in accordance with Newton's laws – an N-body problem where  $N$  is equal to a million billion billion. This is where the statistics come into the story.

The discipline of applying the laws of mechanics in a statistical way to describe the behaviour of large numbers of atoms and molecules in a gas (or some other system) became known as statistical mechanics. The driving force behind the development of statistical mechanics was the need, in the nineteenth century, to find a scientific description of the features of thermodynamics that were observed not only in experiments, but in everyday life – for example, the observed fact that without some kind of outside intervention, heat always flows from a hotter object to a cooler one. In everyday language, hot things cool down. In the real world, if you place an ice cube on a hot surface, the ice melts as it warms up; we never see liquid water spontaneously forming itself into ice cubes and giving out heat as it does so, although, like the example of

the pool balls arranging themselves into a triangle, as far as the motion of individual atoms and molecules is concerned this would seem to be allowed by Newton's laws. There is an arrow of time built into the everyday world, and it seems to be closely tied in with the laws of thermodynamics and statistical mechanics.

Those laws were developed from observations of the way things behave in the real world. They are equivalent to starting out from an observation that apples always fall downwards from trees, never upwards, and calling this a law of nature, then going on to measure how fast they fall, how they accelerate, extending these observations to other objects (such as the Moon and Mars), and eventually to finding mathematical equations that describe this behaviour – in this case, the law of gravity. For thermodynamics, the starting point was an investigation of the way heat flows through solid objects, carried out by Joseph Fourier in France in the years leading up to 1811, when he won a prize from the French Academy of Sciences for this work. Fourier found a simple mathematical law to describe this transfer of heat – the flow of heat is proportional to the temperature difference, with (of course) heat always flowing from the hotter end of the object to the cooler end. This simple law holds an equivalent place in the development of thermodynamics as Galileo's discovery of the law of acceleration for falling objects holds in the development of gravitational theory. And, like the law of gravity, Fourier's law is universal – it applies, in fact, not just to solids but to liquids and gases as well, and it is true whatever the objects are made of, although (unlike the acceleration due to gravity) the constant of proportionality in the relationship is different for different chemical substances. As most people have learned from painful

first-hand experience, heat flows much more quickly through metal than through, say, wood.

Apart from its direct significance to the study of heat flow, Fourier's discovery of this simple law provided a profound insight into the nature of the physical world. At the level of atoms and molecules, it would be a hopeless task to try to predict the overall properties of an object big enough to be manipulated on a human scale (so-called macroscopic objects); the interactions between all the particles are so complicated that they defy direct analysis, and represent (as van Helmont appreciated) something close to the everyday meaning of chaos. But with billions and billions of molecules interacting with one another, somehow the chaos gets lost, or smoothed out, and order associated with simple law appears (or reappears). Nobody knew how this could happen. They knew that a few particles obeying Newton's laws could be described in a simple way (even if actually solving the equations involved tedious approximation techniques), and they knew that huge numbers of particles working together could be described in a simple way. But they did not know how you got from one level to the other, and it didn't matter, as long as thermodynamics could be developed in its own right as a practical science in the steam age.

That development took many decades – most of the rest of the nineteenth century – but it produced a description of the way things behave in terms of their macroscopic properties, such as temperature, pressure, density, and even their chemical behaviour, and how these properties respond to changes imposed on the system from outside, such as an increase in pressure, or a decrease in temperature. It did not try to predict such changes in terms of the behaviour of individual molecules



and atoms, although the interpretation of what was going on did acknowledge the existence of these particles and made use of their average properties in a statistical way.

An important precursor to Fourier's work had been carried out by Count Rumford in Bavaria in the 1790s. Rumford, who was born in the then British colony of Massachusetts in 1753 and lived until 1814, was originally Benjamin Thompson, but was made a Count by George III of Bavaria, whom he served as an adviser during a colourful career as a soldier, statesman, spy and public benefactor. It was through his work in Bavaria, overseeing the production of new cannon for the army, that Rumford came up with the insight that heat is a form of work. The cannon were bored out with a tool driven by horse-power, and as the horses worked harder the boring tool and the cannon blank it was cutting into got hotter. A steam engine converts heat into work; the cannon-boring process turned work into heat, as a by-product of making the cannon.

It took a long time for all these ideas to gestate. A key step was taken by James Joule in Britain in the 1840s. He went one better than Rumford by carrying out accurate experiments to measure the amount of work needed to produce a specific rise in temperature of a certain amount of water – one of his experiments, beautifully simple in concept, involved heating water in a container by stirring it with a kind of paddle wheel, which was driven through a cord attached to a falling weight. This work, and the work of Joule's contemporaries such as Hermann Helmholtz in Germany, led to the idea of the conservation of energy. This says that energy cannot be created or destroyed, but only transformed from one form into another. The energy the horse needs to

do its work comes from the hay it eats, providing chemical fuel which combines with oxygen to power its muscles; the energy stored in the hay in a chemical form comes ultimately from sunlight; and so on. The principle of conservation of energy became known also as the first law of thermodynamics, which says that in a closed system (one that does not interact in any way with the outside world, another physicists' idealization, like a frictionless plane) the total energy remains constant. But, as the example of the cannon-boring shows, no work-energy transformation is perfect, and heat is always dissipated as a by-product, so that some energy is removed from circulation, as it were. As heat always flows from a hotter place to a cooler place (the second law of thermodynamics), eventually, in any closed system, all the energy will end up as heat, and all the temperature differences will smooth out to leave a bland and featureless system where nothing interesting happens.

What is now known as the second law of thermodynamics can be expressed in many different ways, but was first spelled out by the British physicist William Thomson (later Lord Kelvin) in 1852. The essential point which Thomson drew attention to is this idea of dissipation – that although the way the natural world works can be described in terms of a great engine which converts heat into work (or into motion, which is the same thing), there must, indeed, always be some heat dissipated in the process, not really lost but spread out into the Universe at large, raising its overall temperature by a tiny bit. This goes beyond the principle, or law, of conservation of energy (the first law of thermodynamics) because now, although the total amount of energy in the world (by which the Victorians meant what we would now call the Universe) stays the same, the amount of *useful* energy is always decreasing. This implied that

physicists needed a way to quantify the amount of useful energy available in a closed system, or in the world (the Universe at large), so that they could make allowance for it and manipulate it in their equations. That led Rudolf Clausius, in Germany, to introduce the concept of entropy in the mid-1860s.

The simplest way to picture what is measured by entropy is to think in terms of the amount of order there is in a system, and the classic example of this is to imagine a box that is divided into two halves by a sliding partition. One half of the box is filled with gas, and the other half is initially empty – a vacuum. This is a system which has a certain amount of order, or structure, because there is a distinction between the two halves of the container. If you drop a microscopic robot probe into the box at random, it can tell which side of the partition it is by testing to see if it is surrounded by gas or by vacuum. Now imagine sliding the partition out. We all know, from everyday experience, what will happen. The gas spreads out to fill the box uniformly. There is less order (or, if you like, more disorder) in the system now, because from inside the bulk of the gas itself there is no way to tell which half of the box you are in. The temperature, as it happens, will also have fallen, as the gas expands into the vacuum. Now, you can certainly restore the original ordered state of the gas in the box, by pushing all the gas up one end of the box with a piston. This will also heat the gas back up to its original temperature as it is compressed. But in order to do so, the piston has to do work, and in the process heat is inevitably dissipated through friction and so on. So although the box, a closed system, is restored to its original state, in the Universe at large more heat has been dissipated as a result, so the world has changed.

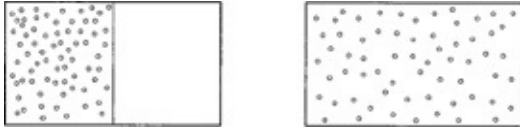


Figure 1.3 when we look at a large number of particles, such as atoms of gas in a box, the direction of time becomes obvious. when a partition is removed, gas spreads out to fill the whole box, and we have no difficulty telling which is 'before' and which is 'after', even without an arrow of time being indicated.

Extending the analogy slightly, a chess board painted in alternating black and white squares has a certain order to it. If exactly the same amount of paint is mixed up and used to paint the board a uniform grey, there is less order. You can imagine a machine delicate enough to unscramble the mixed-up molecules of paint back into pure black and pure white – but, once again, that machine cannot be 100 per cent efficient and must dissipate heat into the Universe at large as it does its job.

Entropy measures the amount of order in a system, with increasing disorder corresponding to increasing entropy. Since we know that in the real world disorder increases in any closed system (things wear out) as time passes, the inevitable increase in entropy defines a direction of time, an arrow pointing from the ordered past into the disordered future. As this process seemed inevitable and universal, the Victorian thermodynamicists envisaged an ultimate fate of the Universe in which all useful energy had been converted into heat, and everything was a bland mixture of stuff at one uniform temperature, a bleak scenario that they called the 'heat death' of the Universe.<sup>6</sup>

Life, of course, seems to defy this process, by creating order and structure out of disordered (or at least, less ordered) materials. A plant builds its structure, and may make beautiful flowers, out of carbon

dioxide, water, and a few traces of other chemicals. But it only does so with the aid of sunlight, energy from an outside source. The Earth, and in particular life on Earth, is not a closed system. It is possible to show, using the equations developed by Thomson, Clausius and their contemporaries, that anywhere in the Universe that a pocket of order appears, it always does so at the cost of *more* disorder being produced somewhere else. When we use a freezer compartment in a refrigerator to make blocks of ice, seemingly in defiance of the second law of thermodynamics, we do so only because an engine is at work pumping fluids around the refrigerator and generating more heat than is 'lost' by the freezing water. The cooling process, going on in the tubes inside the fridge, is essentially the same as the expansion and cooling that happens to the gas in our imaginary box when the partition is removed; the warming process, which goes on in the pipes outside the fridge, at the back, is essentially the same as the compression of that gas back into its original volume, making it hotter. Then the heat dissipates into the air before the fluid flows back into the pipes inside the fridge. If a refrigerator is left running in a sealed room with perfectly insulated walls, and the door of the fridge is left open, the room will get hotter, not colder, because of the work being done by the engine which drives the cooling process going on inside the fridge.

On the macroscopic scale, in accordance with laws derived by experiment and observation in exactly the approved, tried and tested scientific way, the Universe operates in an irreversible way. You can never put things back the way they used to be. But right there in our simple, classic example of thermodynamic irreversibility, entropy and the arrow of time, the box of gas with the sliding partition, we see the

apparent dichotomy between the macroscopic world and the microscopic world with stark clarity. At the level of the atoms and molecules that make up the gas (actually sub-microscopic, but nobody worries about that), every collision is, according to Newton's laws, precisely reversible, like the collisions between balls on a pool table. We can imagine opening the partition, letting the gas fill the entire box, then waving a magic wand to reverse the motion of every atom and molecule of gas. There would be nothing forbidden by Newton's laws in the pattern of reversed motion. Blindly obeying those laws, the atoms and molecules would then retrace their paths back into one half of the box, no matter how many collisions were involved along the way. Yet we never see systems acting like this in the real world, with all the gas in a room, for example, suddenly moving to one end of the room. This dichotomy between the laws that apply on large and small scales was a huge puzzle for physics at the end of the nineteenth century.

The people who tried to resolve this puzzle had to develop a new language (as we have already begun to see) and a new way of thinking about the physical world. One of the key concepts, central to the story we have to tell here, is the idea of an *attractor*. If you release gas into a box, either by sliding a partition in the way we have described, or through a hole in one wall of the box, the end result is that the gas settles down in equilibrium, spread evenly through the box. This state corresponds to the maximum entropy available to the system, and it is the same regardless of how the gas got into the box— it doesn't matter, for example, which wall of the box has the hole which allows gas in, or whereabouts on the wall the hole is made. The final, equilibrium state (which also corresponds to the state of minimum energy) is said to be an

attractor, because the system acts as if it is attracted towards that state. Once that state is reached, there is no way to tell how the system got into it – there is no history written into the final, equilibrium state.

More familiar systems can also be described in this way. If you roll a marble into a pudding basin, after a few circuits and oscillations the marble will end up stationary at the bottom of the bowl; that state is an attractor for this system. But even in such a simple system, the attractor need not be quite that simple. If the marble is rolled into the upturned brim of a sombrero, it will end up somewhere in the valley made by the upturned brim, but all the points in a circle around the floor of that valley are part of the attractor (sometimes called the 'Mexican hat attractor'), since they all correspond to the same minimum energy state. For an imaginary perfect, frictionless pendulum (a so-called ideal pendulum), swinging to and fro is an attractor; for a real pendulum, where friction dissipates energy, the swing will slow and eventually stop, with the pendulum hanging vertically; this, too, can be thought of as an attractor state. It doesn't matter how you set the pendulum swinging (left to right, forward and back, or around in circles), if you wait it will settle down into the same final state, with no record of how it got there. In the language of thermodynamics, by the time a system reaches equilibrium it has forgotten its initial conditions. All that matters is where it is now.

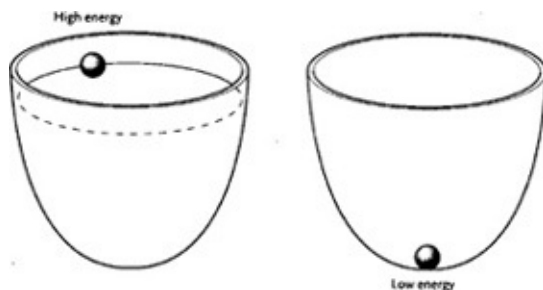


Figure 1.4 *Left to their own devices, systems (even a 'system' as simple as a marble in a mixing bowl) tend to sink to a state of minimum energy and maximum entropy – provided there is no input of energy from outside.*

But when does a system reach equilibrium? In the real world, there is no such thing as a perfect insulator, so our box of gas will always be exchanging heat with the outside world. The swinging pendulum grinding to a halt might seem a better example of a system approaching equilibrium, but even that is being jostled by molecules of air colliding with it, and even if you put it in a vacuum chamber there will be some interaction with the outside world through the string of the pendulum, which has to be fastened to something. The truth is, there is no such thing as an isolated system (except for the entire Universe), and no system is ever in perfect equilibrium. It may get very close to equilibrium – as close as you like, if you wait long enough – but the equilibrium is never literally perfect.

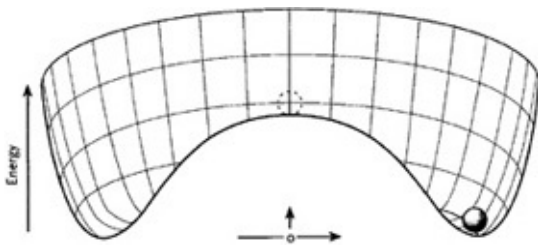


Figure 1.5 *The state that systems settle into is called an attractor. In the example shown on p. 28, the attractor is a single point at the bottom of the bowl. But an attractor can also be a spread-out region, as in this illustration. The marble on the hill is bound to roll off into the valley, but everywhere in the valley bottom is equally attractive.*

This is no mere pedantry. Borrowing an example from Ilya Prigogine and Isabelle Stengers,<sup>7</sup> if two containers are joined by a relatively narrow pipe and filled with a mixture of hydrogen gas and hydrogen



sulphide gas, then in equilibrium, with everything at the same temperature, there will be a uniform mixture of both gases in each container. If, however, one of the containers is kept slightly warmer than the other, the gas mixture begins to separate, with lighter hydrogen molecules concentrating in the hotter container, and heavier hydrogen sulphide molecules concentrating in the cooler container (the same sort of thing happens for any two gases with different molecular weights, not just for hydrogen and hydrogen sulphide). A very small deviation from equilibrium has produced order out of chaos. Away from equilibrium, a flow of energy can, under the right circumstances, create order spontaneously. This is a crucial insight into our own existence, since there is no denying that we are ordered creatures, and there is overwhelming evidence that the Universe began in a state of much less order. In general, a system that is close to equilibrium will be attracted to a state where *the rate at which entropy is being produced is a minimum* – but this insight did not become clear until well into the twentieth century. For the same reason that physicists after Newton concentrated first on solving simple problems that were tractable using his equations and laws of motion (because they were the easiest problems to solve), thermodynamicists concentrated at first on studying equilibrium systems, and developing laws and equations (such as the second law of thermodynamics) to describe them. It was this approach that led to the development of an understanding of thermodynamics in terms of probability, extending its scope under the name statistical mechanics.

A great deal of this work derived from attempts to describe mathematically the behaviour of gases, because gases are relatively simple systems involving collisions between molecules more or less in

accordance with Newton's laws, and so there was some hope of being able to solve the equations for these cases. What became known as the kinetic theory of gases (because 'kinetic' implies motion, in this case of the component molecules) was developed by several people whose ideas cross-fertilized each other in the second half of the nineteenth century. One main thread of the story starts with Clausius, who came up with the idea, in 1858, of the 'mean free path', which is the average distance travelled by a molecule between collisions in a gas at a certain temperature and pressure. He also introduced the idea that molecules have an 'effective radius' so that they behave like little hard balls with that radius. These ideas were taken up and developed by Maxwell, who went further than Clausius in allowing for a range of speeds among the molecules involved in the collisions. He showed how, with a suitable choice of all these properties, many of the observed features of the behaviour of gases could be described by relatively simple equations. In turn, Maxwell's work provided the inspiration for further developments carried out by Ludwig Boltzmann in Germany, who was the one who brought everything together in the first version of statistical mechanics. On the other side of the Atlantic, in the United States, Josiah Willard Gibbs also made major contributions to the development of this new branch of science.

We don't need to go into all the details of statistical mechanics here, but we can give you an insight into the key concept of probability, as applied by Boltzmann, using a very simple example (a mental 'model') related to the gas in a box model used earlier. The puzzle, remember, is why, since it is not forbidden by the Newtonian laws of mechanics, we never see all the gas move into one half of the box. What is so special

about the state in which gas is divided more or less evenly to fill the box? To get a feel for what is going on, imagine first that we have just two particles to put in the box. They move about in accordance with Newton's laws, bouncing off one another and ricocheting off the walls of the box. If we take a snapshot of the inside of the box at any instant, how might the particles be arranged? There are two ways to have one particle in each half of the box – particle A on the left and B on the right, or B on the left and A on the right. And there are two ways in which both particles can be in one half of the box -both on the left, or both on the right.

So there are just four possible states for the box to be in, using this criterion. If each state is equally likely (and there is no reason to think otherwise), then there is a 25 per cent chance of seeing both particles on the left, a 25 per cent chance of seeing both on the right, and a 50 per cent ( $2 \times 25$  per cent) chance of seeing one on each side. With those odds, if you take a quick look at the box you will not be surprised to see any of these possibilities at that particular moment in time.

Now carry out the same process with four particles. To keep it simple, just concentrate on one half of the box this time. There are four ways to have one particle on the left and the rest all on the right (taking each of the particles A, B, C and D as the odd one out in turn). But there are six ways to have a pair of particles on the left-hand side (AB, AC, AD, BC, BD, CD). Already, it is slightly more likely, with odds of 4:6, that you will see equal numbers of particles in each half of the box than one particle on its own – and there is still only one way in which you can have no particles at all on the left. The odds escalate dramatically in favour of an even distribution as the number of particles increases. We'll

let you work through the numbers if you wish, but for just eight particles there is still only one way for all the ‘gas’ to be in the right-hand side of the box, but seventy ways to share out the particles evenly between the two halves. Boltzmann argued that we see gas spreading out to fill a box, rather than huddling up in one corner, because this is overwhelmingly more likely than the alternative. The alternative is not impossible, just very unlikely.

The simple sketch we have given here scarcely does justice to Boltzmann’s work, which included the derivation of a mathematical equation relating entropy to probability,<sup>8</sup> and put statistical mechanics on a proper quantitative footing. But we hope it gives you a feel for what is going on. The key point, of course, is just how many particles are involved in the collisions going on in even a small box of gas, and just how unlikely it is, as a result, that they will all end up in one end of the box. It’s no coincidence that the first real understanding of the numbers involved also emerged during the 1860s, since all of these ideas fed off each other. This thread went back to the work of the Italian Amadeo Avogadro, who suggested, in 1811, that equal volumes of gas at the same temperature and pressure contain equal numbers of particles. He was responsible for the idea that a box of any gas at a particular temperature and pressure always contains the same number of little hard spheres (we would now equate these with molecules) bouncing around and colliding with one another and the walls of the box. It doesn’t matter if the gas is oxygen, carbon dioxide, something else, or, like the air we breathe, a mixture of several gases; all that matters in this context is the number of molecules involved.

Hardly surprisingly, it took a long time for anyone to come up with

an accurate figure for the kinds of numbers of molecules required on this model – if anything, the surprise is that it didn't take longer. There are actually several ways to tackle the problem, but we'll describe just one, used by the Austrian Joseph Loschmidt in the mid-1860s to calculate the number of molecules in one cubic centimetre of gas at a temperature of 0 °C and standard atmospheric pressure (this is known as Loschmidt's number). He started from the assumption that there is a lot of empty space in a gas, and that each molecule occupies a volume specified in terms of its effective radius. The pressure could be calculated from the kinetic theory, in terms of the mean free path, the average speed of the molecules, and so on. If there were a few, large molecules in the gas, there would be a long mean free path; if there were many smaller molecules, there would be a shorter mean free path. To make the model match the way pressure changes in a real gas, only one value of the mean free path, and therefore one value for the total number of particles present, would do. The modern value for Loschmidt's number is  $2.687 \times 10^{19}$  ( $10^{19}$  is a 1 followed by 19 zeros, or 10 billion billion). And that, remember, is the number of molecules in each cubic centimetre of air at sea level on a freezing cold day – which also gives you some idea just how small atoms and molecules are. To put that in perspective, there are about 300 billion ( $300 \times 10^9$ ) bright stars in our Milky Way Galaxy, and roughly the same number of bright galaxies visible to our telescopes. So there are, multiplying this number by itself, about  $90,000 \times 10^{18}$  or  $9,000 \times 10^{19}$  bright stars in the entire visible Universe. Dividing this by Loschmidt's number, we find that in round terms 450 cubic centimetres of cold air, less than half a litre,<sup>9</sup> contains as many molecules as there are stars in the Universe. And those molecules are so close together that

the mean free path between collisions, moving at an average speed of just over 460 metres per second, is just 13 millionths of a metre.

Properly applied to such large numbers of molecules, the statistics tell us that you would have to wait for far longer than the age of the Universe to have much chance of seeing all the gas in even a matchbox-sized container congregate in one end of the box.

And yet, this still does not say that such an event is impossible. Boltzmann thought he had found a way to prove that it is impossible, and to put the arrow of time into thermodynamics in this way. But it turned out he was wrong. Boltzmann set up equations to describe the overall behaviour of a box of gas, described in terms of a proper statistical averaging of the properties of the individual molecules, rather than attempting to trace the interactions of every single molecule in the box. Putting some of this into everyday language, his equations described how the properties of the molecules of gas in the box get averaged out as time passes. For example, when the gas is first let in to the box, some molecules will be moving faster than others, which corresponds to them having a higher temperature, so that the gas in one corner of the box may be hotter than the rest of the gas. But when a fast-moving molecule collides with a slow-moving molecule, the energy of motion of the two molecules (their kinetic energy) tends to get shared out between them, so that you end up with two molecules moving at a middling speed (think of a fast-moving pool ball colliding with a stationary pool ball, without any spin on the ball). So the hotter molecules would cool down and the cooler molecules would warm up, until the box of gas was at one uniform temperature, in thermodynamic equilibrium. And this would be a swift and efficient process, since each

molecule in air at 0 °C experiences a little under four million collisions every second. The equations set up by Boltzmann clearly showed random, natural processes leading to thermodynamic equilibrium along a one-way street.

So why do we live in an orderly world? Boltzmann eventually went much further than anyone else by suggesting that the heat death of the Universe at large had already occurred, and that the Universe at large must be in equilibrium in this way. How then could a place like the Earth, far from equilibrium and bearing life, exist? His answer was that the entire region of space we can see and investigate with our telescopes (which, at the end of the nineteenth century, amounted to no more than the local region of the island of stars we call the Milky Way Galaxy) must represent a small (on the scale of the Universe), local fluctuation away from equilibrium. In our hypothetical box of gas, in equilibrium, there could, just by chance, be occasions when more molecules were moving from left to right in the box than from right to left, briefly increasing the density of gas in the right-hand half of the box before it got smoothed out again by thermodynamic processes. Boltzmann said that the visible part of the Universe was equivalent to such a temporary, local fluctuation in a much bigger cosmos. In his *Lectures on Gas Theory*,<sup>10</sup> he wrote:

We have the choice of two kinds of picture. Either we assume that the whole universe is at the present moment in a very improbable state. Or else we assume that the aeons during which this improbable state lasts, and the distance from here to Sirius [one of the nearest stars], are minute if compared with the age and size of the whole universe. In such a universe, which is in thermal equilibrium as a whole and therefore dead, relatively small regions of the size of our galaxy will be found here and there; regions (which we may call 'worlds') which deviate significantly from thermal equilibrium for relatively short stretches of those 'aeons' of time. Among these worlds the probabilities of their state (i.e. the entropy) will increase as often as they decrease. In the universe as a whole the two directions of time are indistinguishable, just as in space there is no up or down. However, just as at a certain place on the earth's surface we can call 'down' the direction towards the centre of the earth, so a living organism that finds itself in such a world at a certain period of time can define

the 'direction' of time as going from the less probable state to the more probable one (the former will be the 'past' and the latter the 'future'), and by virtue of this definition he will find that his own small region, isolated from the rest of the universe, is 'initially' always in an improbable state. It seems to me that this way of looking at things is the only one which allows us to understand the validity of the second law, and the heat death of each individual world, without invoking a unidirectional change of the entire universe from a definite initial state to a final state.

It's unfortunate that Boltzmann chose gravity as his example of a way to obtain a preferred direction in space, since, as we have mentioned, it is the influence of gravity in the Universe at large that turns some of the traditional ideas about entropy on their heads on such large scales. And it's ironic that he should discard the idea of 'a unidirectional change of the entire universe from a definite initial state to a final state', since that is a succinct summary of the currently favoured Big Bang model of the Universe. Most of all, though, Boltzmann's ideas look outdated because we can now probe far beyond Sirius (a mere 8.6 light years away) to investigate the properties of the Universe on a scale often *billion* light years, with no sign of the uniform, dead cosmos that Boltzmann invokes. But beware of laughing too much at his quaint, old-fashioned ideas. There are reputable cosmologists today who argue that our entire visible Universe is just one of many expanding bubbles in a vastly greater, perhaps infinite, more or less uniform cosmos. These ideas are not based on thermodynamic principles, but they serve as a reminder that although we may laugh at Boltzmann for being so parochial as to think 8.6 light years was a large distance, in a truly infinite cosmos even ten billion light years may be seen as the cosmic equivalent of a stroll down to the chemist's shop.

But none of this is, strictly speaking, relevant to our story. For, as we have hinted, there is a flaw in Boltzmann's argument – not in his philosophical musings, but in the mathematics he used to describe the behaviour of gases. Boltzmann started out from the assumption that



BEHAVIOUR OF GASES. Boltzmann started out from the assumption that molecules which are about to collide with one another don't know anything about each other – their motion is uncorrelated, and subject to what is sometimes called 'molecular chaos', where the term chaos is being used in the everyday sense, not in the sense that it is used later in this book. But after the collision, of course, the paths followed by the molecules are correlated, since they have exchanged energy and momentum in the collision. So Boltzmann had unconsciously built in a microscopic arrow of time to his calculations from the outset, which meant that it was hardly surprising that a macroscopic arrow of time emerged at the end of the calculation. But there is no arrow of time in Newton's laws, which are used to describe the collisions – they work equally well when the entire system is run 'backwards in time'. You can't have correlations working one way and not the other. Or, to put it another way, as Laplace would have argued, the positions and velocities of all the molecules in the gas at any instant have stored in them a 'memory' of the entire past history of the gas, so that the behaviour of every molecule is correlated with the behaviour of every other molecule.

The flaw in Boltzmann's argument was pointed out by Loschmidt in the mid-1870s, soon after the original formulation of Boltzmann's idea was put forward.<sup>11</sup> The problem was taken up and developed by Henri Poincaré (1854-1912), a French mathematician and physicist. Poincaré took Laplace's argument to its logical conclusion. He proved, using the full rigour of mathematics, that if you have a box of gas containing a definite number of particles (as many as you like, as long as it is not actually infinite) and the particles strictly obey Newton's laws of motion, then after a sufficiently long interval of time the distribution of the

particles in the box must return to its original state, with each particle moving in the same direction and at the same speed that it was moving to start with. You can make a useful analogy with shuffling a pack of cards. Whatever state you start out in (whatever order the cards are in), at first shuffling will break up the initial pattern. But if you keep shuffling long enough, and the shuffling is genuinely random, there must come a time when the pack just happens to be back in the order it started out. And since the starting time is arbitrary, this means that the pack of cards or the box of gas must repeatedly pass through every possible state, including those rare occasions when all the gas is in one end of the box, and vast numbers of occasions when it is spread out evenly through the box, but with the particles in slightly different positions. All the allowed states of the gas in the box recur repeatedly, with a characteristic periodicity called the Poincare recurrence time, or Poincare cycle time. If the entropy goes up for a while, it must inevitably come down again later on, to bring the gas back to its original state (the basis of Boltzmann's Universe). And this recurrent, cyclic behaviour is a direct result of the strict application of Newton's laws, in which past and future have equal status.

But the timescales involved are beyond comprehension. For an isolated box of gas containing  $N$  molecules, the time between recurrences, Poincare proved, will be  $10^N$  seconds. Remember that a single cubic centimetre of air at sea level and zero degrees Celsius contains more than  $10^{19}$  molecules. The time you would have to wait to see the gas in a container that size run through its cycle is *10 to the power*  $10^{19}$  seconds, and  $10^{19}$  is a 1 followed by ten billion billion zeros. The entire age of the Universe since the Big Bang is only about  $10^{17}$

seconds. If you divide 17 by  $10^{19}$ , you get some idea of just how tiny the probability is of seeing the gas in the box deviate from thermodynamic equilibrium during the entire life of the Universe to date, as long as it really is insulated from the outside world. Even if the box contained only 52 particles, the same as the number of cards in a pack, the cycle time would be  $10^{52}$  seconds, which is  $10^{35}$  times the age of the Universe.<sup>12</sup>

In the light of all this, thermodynamics, and in particular the idea of increasing entropy, became statistical ideas – in the real world, given the number of particles involved, it is overwhelmingly probable that entropy will rise, but it is possible (in the sense that it is not forbidden by the laws of physics) for entropy to fall. This was not something that all physicists found easy to live with at first. Two centuries after Newton had published the *Principia*, and seventy-five years after Laplace, they were being told that the world is not deterministic in the simplest sense after all, and that chance and probability must be taken into account when trying to describe, or calculate, the behaviour of macroscopic systems. Gradually, though, they got used to the idea, and succeeding generations were brought up to regard it as natural, and even obvious. But while these ideas were being assimilated into the mainstream of science, Poincaré was moving on to other problems. In the late 1880s, he discovered something so shocking about the workings of Newton's laws themselves that the implications were largely ignored for the best part of seventy years, the Biblical 'three score years and ten' of a human lifetime. He found, in essence, that Newton had been right to worry about the stability of the orbits of the planets, and that Laplace had been wrong to think that he had solved the puzzle. Poincaré showed that the three-body problem (let alone the N-body problem) often cannot be

solved even by the familiar approximation techniques, and that seemingly simple orbits obeying the familiar Newtonian laws of mechanics and gravity can behave in a chaotic and unpredictable (literally unpredictable) fashion.

## 2

### The Return of Chaos

There's a snag with the iterative technique used to provide 'solutions' in situations like the three-body problem. It doesn't always work, and you can't always tell in advance whether or not it is going to work. The technique used to 'solve' the relevant differential equations (which, remember, cannot be solved analytically) involves successive approximations, in which the first step in the calculation is known to give only an approximate answer, the second step (hopefully) adds in a correction to give you a more accurate approximation to reality, the third step gives an even better approximation to reality, and so on until the approximation is as good as you want it to be for whatever purpose you have in mind. But you can never get *exactly* the 'answer' that fits perfectly the behaviour of the objects in the real world that you are interested in.

What you are doing is adding up a series of numbers – in principle, an infinitely long series of numbers. Mathematicians are interested in such infinite series in their own right, regardless of their possible relevance to studies of the behaviour of things like planets in orbit around the Sun, and they know plenty of infinite series that do indeed add up in this well-behaved way to give an ever better approximation to some specific number. A good example is one of the standard ways to

work out an approximate value for the number  $\pi$ , the ratio of the circumference of a circle to its diameter. You can actually work out  $\frac{\pi}{4}$ , as accurately as you like, by adding up the series of numbers:

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \dots$$

This gives you a first approximation for  $\pi$  of 4 ( $4 \times 1$ ), which is not brilliant, a second approximation of 2.6666 ... ( $4 \times \frac{2}{3}$ ), which is slightly better, and also, significantly, lies on the other side of the 'right' answer, a third approximation of 3.46666 ... and so on. The approximations get better and better, converging on the true value of  $\pi$  from either side, in this particular case. But it is a tedious process – adding up the first million terms in the series gives a value for  $\pi$  of 3.1415937, which is only correct to the first five decimal places. Nevertheless, you can find it this way to any accuracy (any number of decimal places) you wish, if you have the patience.

But many infinite series do not converge in this way. Just adding up all the positive integers ( $1 + 2 + 3 + 4 + \dots$ ), for example, produces a series which simply gets bigger and bigger, tending towards infinity; slightly more surprisingly, at first sight, adding up all the integer fractions ( $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \dots$ ) also tends towards infinity, even though each successive term in the series is smaller than the one before. And there are series whose sum zig-zags up and down as you add more terms, like the approximation  $\frac{\pi}{4}$  but gets further and further away from some particular value instead of converging on a specific value. Take the series

$$1 - 2 + 3 - 4 + 5 \dots$$

Successive 'iterations' in this case give you the values 1, -1, 2, -2, 3, -3, and so on. The more terms you have in the series, the bigger the swings in successive 'approximations' either side of zero. And there are series which do not sum in any obvious regular way at all. All these series that do not converge on a specific finite number are said to diverge, and as well as using just ordinary numbers you can have series which involve variables, known as power series, such as the series

$$x - x^{\frac{2}{2}} + x^{\frac{3}{3}} - x^{\frac{4}{4}} \dots$$

which gives you scope, if you have the mathematical skills, to describe things like orbits in this way. Such infinite series have formed an important part of mathematics since the time of Isaac Newton, who developed techniques to manipulate them in problem-solving, and was led to develop the calculus through this work.

In some cases, it is possible to prove mathematically that a series of this kind must converge, even without calculating millions of terms in the series. In some cases, it is possible to prove that a series diverges. But in many cases, there is no way to prove whether the series will diverge or converge. In such cases, although you may be working with what seems to be a well-behaved convergent series, you cannot be sure, no matter how many terms you have added up already, that the next iteration will not produce something unexpected, which changes the answer you have calculated so far significantly – perhaps drastically.

This is exactly the situation that astronomers who used the approximation techniques to calculate the orbits of the planets in the Solar System were in in the middle of the nineteenth century. The approximations seemed to work and gave satisfactory solutions. As

Laplace had discovered, they described well-behaved orbits around the Sun, which changed only in regular and predictable ways when the complicating effects of the gravitational influences of the planets on each other were taken into account. But nobody had been able to prove that the series used in these calculations must always converge. Astronomers weren't too worried about this, as long as the approximations seemed to work; but mathematicians saw it as a challenge. Nobody really thought that the Solar System wasn't stable, but it would be nice to have proof that it was.

In the 1850s, one of the leading experts in number theory, with a special interest in the behaviour of infinite series, was Lejeune Dirichlet, a German mathematician working at the University of Göttingen. In 1858, Dirichlet told his student, Leopold Kronecker, about a new technique he had developed for solving some families of differential equations that had no analytical solutions. He also hinted, but did not spell out the details, that he had used this technique to prove that the approximate solutions in terms of infinite series for the equations describing the orbits of the planets (the equations of celestial mechanics) did converge. Unfortunately, Dirichlet died in May 1859, at the age of fifty-four, without telling anyone the details of this claimed proof. Kronecker told the rest of the mathematical world about Dirichlet's claim, but neither Kronecker nor anyone else was able to use the new techniques to prove that the relevant series really did converge. Nevertheless, Dirichlet had such a high reputation that nobody doubted that he had achieved what he had claimed to have achieved – especially, of course, since everybody wanted to believe that the orbits of the planets really were stable.



By the end of the 1880s, though, it was beginning to niggle the mathematicians that they still hadn't been able to reproduce what Dirichlet had said he had done. Then an opportunity came for something of a concerted attack on the problem. Mathematicians at the University of Stockholm, in Sweden, came up with the idea of contributing to the celebrations for the imminent sixtieth birthday of their king, Oscar II, by organizing an international competition, with a major cash prize, for a piece of original mathematical work addressing one of four deep questions set by a committee. One of those questions, inspired by Dirichlet's claim (and by the failure of anybody to reconstruct his alleged proof even after the best part of thirty years), asked for a proof that the power series involved in the calculations of planetary orbits must always converge. In plain language, the question could be summed up concisely: Is the Solar System stable? Among the entries for the competition, this particular challenge was taken up by Henri Poincaré, who in the process invented the techniques that are still used today to study the behaviour of dynamical systems.

The first key contribution, the idea of phase space, arose from the work of the Irish mathematician William Hamilton (1806–65), earlier in the nineteenth century.<sup>1</sup> Hamilton found a way to reformulate the Newtonian laws of physics in terms of two properties of particles, their position and momentum (which is equal to the mass of a particle multiplied by its velocity). He described interactions in terms of changes in momentum rather than in terms of forces. The two are exactly equivalent physically,<sup>2</sup> but this approach provided a way of describing a complete system of interacting particles in terms of a set of differential

equations (the Hamiltonian) which could be manipulated in accordance with certain rules. Position and momentum are the key variables in the description, appearing on a mathematically equal footing; but each described by its own set of differential equations, which contain information about how the relevant quantity changes as time passes. The state of a single particle at any moment in time is specified by its position and momentum.

This is where phase space comes in. On a two-dimensional graph, the position of any point can be described in terms of two numbers, the  $x$  and  $y$  coordinates, which tell us how far the point is along the graph from left to right, and how far up the graph it is from top to bottom. In a similar way, the position of a particle in ordinary three-dimensional space can be represented by three numbers, coordinates indicating its distance from some chosen zero point. It is a relatively small leap of the imagination to take this idea over into an imaginary space which represents all the possible momenta a particle may have. Since momentum is proportional to velocity, and velocity is a three-dimensional property (it involves direction, not just speed), a single point in this imaginary velocity space will represent the components of the velocity of a particle in each of three directions, which can be chosen to be at right angles to one another. The numbers can be read off along the  $x$ ,  $y$  and  $z$  axes, and the three components can be combined to indicate the overall momentum of the particle. Now comes the step that only a mathematician would take. Why not combine these two sets of information, three dimensions representing space and three dimensions describing motion? The result is an imaginary six-dimensional space, in which a single point describes *both* the position *and* the momentum of a

particle at a particular moment. This is the phase space of a single particle.

Fortunately, you don't have to be able to visualize such a phase space. The equations which describe what is going on can be extended from two, to three, to many dimensions in a relatively straightforward way. As long as you can manipulate the equations, you don't have to picture what is happening, except (if it helps) in terms of the three-dimensional equivalent, a kind of cross-section through six-dimensional phase space. This is particularly important because we can't stop here. It takes six dimensions to describe the state of a single particle. If there are two particles interacting in an otherwise empty box, you need a phase space with twelve dimensions to describe the system in this way, and so on. The state of a box of gas at room temperature and pressure can also be represented by a single point in a phase space, but now the phase space involved has six times as many dimensions as the number of particles in the box – itself a number whose enormous size we have already encountered. Of course, the vast number of differential equations involved in describing such a system cannot be solved in practical terms, even if the mathematics tells us that they are solvable in principle. But the statistics in classical statistical mechanics involve analysing the probabilities associated with the distribution of points in phase space describing a system like a box of gas in these terms. There are many more points (quantifiably more) in phase space corresponding to an even distribution of particles throughout the box, for example, than there are points representing all the particles at one end of the box, and there are far more points corresponding to an even distribution of momentum among the particles than there are points representing a

division of momenta with half the particles moving very fast and half moving very slowly. This corresponds to there being a much larger volume of phase space to hold all the relevant points in the first case, and a much smaller volume of phase space to contain the points which represent the second case.

Phase space can be thought of as like a landscape, with rolling valleys, deep potholes, hills and mountains. The Hamiltonian allows mathematicians to analyse how the overall system changes as time passes, in a general way, without solving the myriad individual differential equations. If we imagine pouring water over our phase space landscape, for example, it will flow along the valleys, taking longer to flow through large valleys than through small indentations in the landscape; it will congregate in the deep potholes, and run off from the mountain tops. The Hamiltonian tells us how real systems move ‘through’ phase space, and indicates regions they will be attracted to – the deep valleys and potholes. We can even extend the analogy to include the equivalent of evaporation and rainfall, corresponding to relatively rare paths which take particles out of the ‘river’ and on to the ‘mountain tops’, from where they inevitably flow back down. The motion of a *single* particle through phase space (its trajectory) represents the way the whole system changes with time, and the amount of time the particle spends in each part of phase space is proportional to the volume of that region of phase space. You can’t say exactly where the single particle will go, any more than you can predict the exact trajectory of a single molecule of water in a swollen river; but you can say that there is an overwhelming likelihood that it will follow a certain kind of trajectory through phase space, just as our water molecule has

little choice but to stay within the banks of the river. There is only a much smaller probability that it will follow one of the rare, thinly populated trajectories up to the mountain top and back down again.

To get back to a very simple example, for a perfect, frictionless pendulum swinging to and fro there is only one relevant dimension of real space and one dimension of velocity space, so the phase space is two-dimensional and can be drawn on a flat piece of paper. Let's choose to represent position across the page, and velocity up and down the page, with up corresponding to motion to the right and down corresponding to motion to the left. As the bob of the pendulum moves from left to right and back again, its velocity starts at zero at one end of the swing, rises to a maximum in one direction at the middle of the swing, drops to zero at the other end, then repeats the process in reverse. Combine this with the changing position in real space, and it traces a circle in phase space. If you allow for friction, the swinging pendulum will gradually slow down and stop, tracing out a spiral in phase space, with the point at the centre of the spiral being an attractor for this particular system.

This idea of thinking about phase space in terms of mountains, valleys and so on is a very simple version of topology, a subject pioneered by Poincaré and which he applied, in the context of phase space, to the problem of proving that the Solar System is stable. In essence, he turned a problem involving mechanics and dynamics into a problem involving geometry. Remember that a single point in phase space corresponds to the entire state of a system – in the case of the gas in a box, that point (or the Hamiltonian corresponding to that point) uniquely represents the position and momentum of each and every

particle of gas in the box. A trajectory through phase space describes how the state changes as time passes. But if that trajectory ever passes through a point that it has passed through before, that means the whole system has *exactly* returned to an earlier state. In that case, according to the laws of Newtonian mechanics, it must proceed from that state in exactly the same way that it did before, retracing its trajectory through phase space. For boxes of gas, this is at the heart of Poincaré's idea of cycles and recurrence; for orbits, it means that if a trajectory through a phase space representing the possible states of, say, three bodies returns to the same point in phase space that it has been to before, then the orbits themselves, no matter how complicated they may look, must repeat periodically and the bodies involved will not suddenly fly apart from one another or crash into one another. You might also guess that if the trajectory through phase space returns very nearly to a point it has passed through before, then the subsequent behaviour of the system will be very nearly the same as it was before. But, as we shall see, you should take nothing for granted.

In his competition entry, Poincaré didn't even attempt to describe topologically the behaviour of the entire Solar System, but concentrated on a geometrical representation of the phase space for the orbits associated with a very simple case, known as the 'restricted' three-body problem. In a universe containing just two gravitating bodies there is, as we have seen, no difficulty in describing their orbits. As Newton showed, they move around one another in a completely regular and periodic way, which Poincaré and his successors could describe in terms of a closed loop in phase space. Add another body, though, and as we have seen, the motion becomes too complicated to calculate analytically. The restricted

three-body problem tries to get around this by imagining two bodies more or less the same size, and a third body which is so much smaller than the other two that, for the purposes of the calculations, its gravitational effect on the larger bodies can be ignored, even though they have a powerful effect on the small body (which is sometimes called a 'dust particle'). The situation still cannot be solved analytically, but all you try to do then is calculate the orbit of the dust particle, by summing the appropriate series, within the framework of the changing positions of the other two bodies. This is an impossible situation, since the dust particle must always have some gravitational influence on its partners, but the hope of mathematicians such as Poincaré was that it might give a reasonable approximation to what is going on in the Solar System if we think about the interactions between, say, the Sun, Jupiter and the Earth, or the Sun, Earth and the Moon.<sup>3</sup>

One of Poincaré's great simplifying assumptions was to look at only a small part of the appropriate phase space – in effect, a cross-section through phase space, representing a surface through which the trajectory being investigated had to pass. This is now known as a Poincaré section. If a trajectory through phase space starts at one point on a Poincaré section and returns to that point, then no matter how complicated the behaviour of the system might have been in between these two intersections with the Poincaré section, you know that it is exactly periodic. In order to lead the judges relatively gently up to his new ideas, Poincaré started out by tackling the problem in the conventional way, which meant that he still had to solve differential equations to try to obtain the longed-for proof that orbits in the Solar System (or, at least, the orbits of a dust particle in the restricted three-body problem)

were periodic, and since the equations could not be solved analytically this meant adding up series, and he didn't quite finish the job. He found the appropriate equations, and he derived the appropriate infinite series, but he didn't actually prove that the series *must* converge, only that there were appropriate solutions that *could* converge. Then, he went on to use his new topological approach to show (as he thought) that these must be the right solutions, since the appropriate trajectories through phase space did return to cross the Poincaré section at the same point they started from.

It took Poincaré more than 200 pages to set all this out, and much of it was new to the judges. But it gave the answer they wanted, it was obviously very clever, and they had to complete the judging in time for the King's birthday on 21 January 1889. Poincaré got the prize. But when the paper was published, and other mathematicians had time to study it at length, they found that Poincaré had made a mistake. His 'proof' didn't hold up. After a burst of intense work in response to the criticisms made by his colleagues, Poincaré corrected the flaws in his reasoning and published a revised paper, now regarded as one of the all-time classics in mathematics, in 1890. To his own surprise, when he corrected his mistakes he found that his new approach proved that, in the old language, even for the orbit of the dust particle in the restricted three-body problem, the appropriate series *typically* diverge – instability is normal, and permanently stable orbits are the exception. In terms of his new geometrical approach, he found that although there are periodic orbits for which the trajectory in phase space returns repeatedly to a precise point on the Poincaré section, if the trajectory cuts the Poincaré section even a tiny distance away from that point, the system can then



follow a completely different pattern of behaviour, with trajectories taking very different routes through phase space and repeatedly crossing the Poincaré section without ever quite hitting the same point twice. There could be infinitely many crossing points on the Poincaré section, allowing the trajectory through phase space to wander in an infinitely complicated and varied way without ever returning to its starting point.<sup>4</sup> The good news is that, at least in the case of the restricted three-body problem, and more generally for the planets of the Solar System, orbits can in principle be calculated as accurately as you like by going through the laborious approximation techniques and adding up the appropriate series for a large enough number of terms, even though they are not, strictly speaking, perfectly periodic. And, as we shall see, these calculations show that there are relatively simple situations in which planets can follow essentially the same orbits for very long intervals compared with any human timescale, even though they are not strictly periodic on timescales that are long compared with, say, the age of the Sun. But the key thing to take away from Poincaré's work is the realization that under some circumstances (not all circumstances, but not necessarily rare circumstances, either) systems that start out from very nearly the same state can very rapidly evolve in entirely different ways.

There are two helpful ways of getting a feel for this. The first goes back to the analogy of phase space as a landscape through which water is flowing. The trajectory of a single molecule of water through phase space represents the changing state of an entire system, as simple as the restricted three-body problem or as complicated as the entire Universe. Imagine that there is a broad river flowing through the landscape,

representing a region in which it is highly likely that the Hamiltonian corresponding to the system will be found. At some point, the river splits and diverges into a complicated network of channels, forming a delta like the mouth of the Ganges. A single molecule moving with the flow of the river may move to the left, into one branch of the forking channels, while its neighbouring molecule moves to the right, into another branch of the network. Thereafter, they follow quite different paths. In other words, two trajectories through phase space that have run for a long way beside each other, corresponding to very nearly identical states of the system, can diverge into very different states. In a similar way, imagine a drop of water falling on to a high, knife-edge mountain range. If it falls one side of the ridge it will flow away in one direction, to a deep ocean representing a strong attractor for the system; if it falls the other side of the ridge, it will flow in the opposite direction, into another deep ocean, an equally powerful attractor for the system. But the two trajectories leading to such different final states may start out infinitesimally close together.

An example relating to the everyday world also helps. In our everyday lives, most changes occur in a linear fashion. It takes twice as much force to lift a 2-kilo bag of sugar as it does to lift a 1-kilo bag of sugar, and so on. If you take two steps down the road, you travel twice as far from where you started than you would if you took one step. But suppose walking was a non-linear process, so that, after you had taken the all-important first step, each step you took carried you twice as far as the one before. Let's say the first step takes you a metre down the road. Then the second step would cover two metres, the third four metres, and so on. With linear walking, 11 steps would take you just 11 metres down

the road. But with this kind of non-linear walking, the 11th step alone would take you a distance of 1,024 metres, which itself would be one metre more than all the previous 10 steps put together. In round terms, 11 steps would take you not 11 metres but *2 kilo* metres down the road. Non-linear things move very swiftly away from their starting point, and if two trajectories start out in non-linear fashion in slightly different directions through phase space, they diverge from one another no less rapidly. Poincaré discovered that many systems in the real world are very sensitive to their initial conditions (like the raindrop falling on the mountain ridge), and that they move away from those starting conditions in non-linear fashion.

The important point is that this limits our ability to predict the behaviour of such systems. With a linear system, if I make a small mistake in measuring, or estimating, some initial property of the system, this will carry through my calculations and lead to a small error at the end. But with a non-linear system, a small error at the beginning of the calculation can lead to a very large error at the end of the calculation. A linear system is more or less equal to the sum of its parts; a non-linear system may be either much more, or much less, than the sum of its parts. Poincaré realized that this means that in some circumstances we cannot calculate just how a system will change as time passes, because the information we have about its starting conditions is not accurate enough. In his *Science et Méthode* he wrote, in 1908:

A very small cause that escapes our notice determines a considerable effect that we cannot fail to see, and then we say that the effect is due to chance. If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation *approximately*. If that enabled us to predict the succeeding situation with *the same approximation*, that is all we require, and we should say that the phenomenon had been predicted, that it is governed by laws. But it is not always so; it may happen that small

differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon.

In other words, although Laplace was right in principle, his deterministic universe could never be predicted in practice, either forwards or backwards in time.

In the same volume, Poincaré elaborated on this theme using the specific example of weather forecasting:

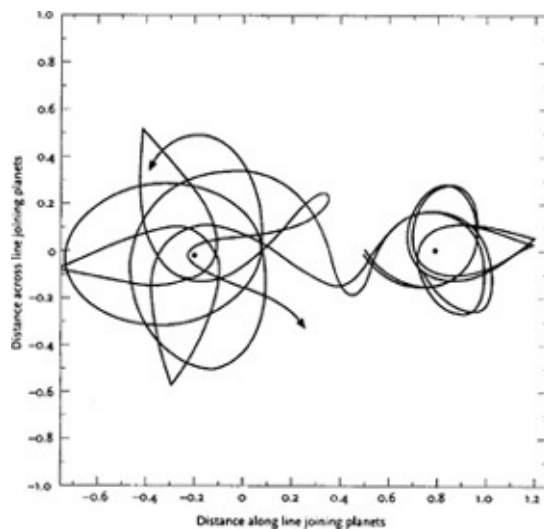


Figure 2.1 *The three-body problem. If a small ‘satellite’ is orbiting two larger ‘planets’, even a tiny change in the starting conditions of the satellite’s trajectory soon leads to a large change in its orbit. Because we can never know the starting conditions precisely, this means the orbit is unpredictable. This view shows two such trajectories, in a rotating frame of reference so that the two planets appear stationary.*

Why have meteorologists such difficulty in predicting the weather with any certainty? Why is it that showers and even storms seem to come by chance, so that many people think it quite natural to pray for rain or fine weather, though they would consider it ridiculous to ask for an eclipse by prayer? We see that great disturbances are generally produced in regions where the atmosphere is in unstable equilibrium.<sup>5</sup> The meteorologists see very well that the equilibrium is unstable, that a cyclone will be formed somewhere, but exactly where they are not in a position to say; a tenth of a degree more or less at any given point, and the cyclone will burst here and not there, and extend its ravages over districts it would otherwise have spared. If they had been aware of this tenth of a degree, they could have known of it beforehand, but the observations were neither sufficiently comprehensive nor sufficiently precise, and that is why it all seems due to the intervention of chance.

**This example can be seen with hindsight to be particularly**

This example can be seen, with hindsight, to be particularly prescient. Poincaré was far ahead of his time with his attempts to describe how orderly rules of physics can give rise to seemingly chaotic behaviour in the everyday world, and it would be many decades before these ideas were revived and became part of mainstream science. When that revival did take place, though, it came about, in no small measure, through the efforts of meteorologists to improve their weather forecasts, and to understand why those forecasts often failed. The key question is just how comprehensive and precise the initial observations have to be; and a good place to start on the problem is through the work of the English pioneer Lewis Fry Richardson, who was also a man ahead of his time, and who worked out how to do scientific weather-forecasting in the second decade of the twentieth century, long before the tools needed for the job (in particular, high-speed electronic computers) were available.

Richardson, who was born in 1881 and died in 1953, had wide-ranging interests, and was by no means a stereotypical meteorologist.<sup>6</sup> A Quaker and the son of a farmer, Richardson was superintendent of the meteorological observatory at Eskdalemuir, in Scotland, at the outbreak of the First World War. Because of his religious beliefs, he did not take part in the fighting, but volunteered as an ambulance driver, serving close to the front line. There, in his spare time, he spent hours working through tedious arithmetical calculations in an attempt to produce the first numerical weather ‘forecast’ – not really a forecast in the usual sense of the term since he had brought with him a stack of genuine weather data for one particular hour on one particular day, and was trying to use that information to calculate the broad outlines of the

weather patterns over Europe for a time six hours later than the time for which these input data applied. Even if it worked, it would take months to make a 'forecast' for six hours ahead; but what mattered to Richardson in the second decade of the twentieth century was the possibility of proving that the mathematical approximations and the laws of physics could be used in this way.

The inspiration for this project came from the work of the Norwegian Vilhelm Bjerknes, nineteen years older than Richardson, who had suggested, a decade earlier, that such a technique for predicting the weather might be possible in principle. He was the first to propose that the equations that described the weather were known well enough for forecasting to be possible, if the initial conditions for starting the calculation were known accurately enough. The idea behind this approach to weather prediction, which is at the heart of modern forecasting, is to measure the important properties of the atmosphere, such as temperature and pressure, at a grid of points over the surface of the Earth and extending upwards into the air. The closer the points are together, the more accurate is the mathematical model of the state of the atmosphere at that time. Then you apply the laws of physics to work out how the conditions at each grid point will change under the influence of its neighbours (with heat flowing from hotter points to cooler ones, wind blowing from high pressure regions to low pressure regions, convection, and so on).<sup>7</sup> The similarity to the kind of techniques needed to solve the puzzle of the orbits of the planets is obvious, and common sense tells us that you have much more hope of success if you can use a finer grid, with points closer together. Conditions at places in between the grid points are calculated simply by interpolation – taking an appropriate

average of the properties of the surrounding grid points. The expectation was that if the technique worked at all, it could be made as accurate as you liked (in principle) by using more and more points from an ever more finely meshed grid. It happened that Richardson's forecast was hopelessly inaccurate, but he was not at all surprised since he was forced to use a very coarse grid of data points (both because of the limited data available and because of the time required for the calculations), which could not realistically be expected to represent all the subtleties of the weather over Europe. What mattered at the time was that he had used the technique to calculate how weather might change, even if the results of his calculations did not match the real world.

Richardson was sufficiently enthusiastic about the technique to start work on a book, *Weather Prediction by Numerical Process*, while still serving on the Western Front. In the confusion of a rapid move caused by the fighting, the only copy of his handwritten manuscript was lost in 1917 – but it turned up, a few months later, under a heap of coal. Nobody knows how it got there, but with the manuscript restored to him Richardson was able to complete the book, which was published in 1922 and contains his most famous vision of the future of forecasting. Well aware that there is no practical point (no matter how rewarding it may be for mathematicians) in making a forecast for six hours ahead if it takes months (or even just seven hours) to do the calculations, Richardson asked his readers to envisage a 'Forecasting Factory' in which 64,000 human 'computers', each equipped with a mechanical calculator (a kind of glorified adding machine, the forerunner of the electronic pocket calculator), would each be working on one aspect of the problem,<sup>8</sup> seated in an amphitheatre like a football stadium, and all

under the direction of a kind of mathematical conductor on a stage at the centre of operations, communicating with them by flashing lights, or messages sent along pneumatic tubes. It was, of course, scientific fantasy – it would never be possible to correlate the actions of so many people in the way required. Richardson himself wrote:

Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances at a cost less than the saving to mankind due to the information gained. But that is a dream.

It was, though, a dream he came to see almost realized. Within three decades, before Richardson himself had died, the invention of the electronic computer opened up the possibility of doing essentially the same job with machines – with a *single* machine replacing all of Richardson's 64,000 human computers, and (eventually) many more besides. The first successful numerical weather forecast was made by such a machine in 1950, although at that time it still took longer for the machine to come up with the prediction than it did for the real atmosphere to evolve.

Richardson himself, no doubt aware that he had taken numerical forecasting as far as it could go in the 1920s, moved on to other things, running the physics department at Westminster Training College in London from 1920 to 1929, and then being Principal of Paisley Technical College in Glasgow until he retired in 1940. Along the way he studied psychology, and both before and after his retirement he wrote books about the psychology and causes of armed conflict. In view of this, and his Quaker beliefs, it is somewhat ironic that the huge progress in computing techniques, which saw numerical weather-forecasting becoming a reality in his lifetime, was stimulated by the work of the codebreakers in the Second World War. But although at the beginning of



the 1950s it seemed that Richardson's dream was coming true, and that all you needed for better forecasts were faster computers and more finely meshed grids of observations, ten years later one man stumbled upon the evidence that pulled the rug from under this assumption and showed that even Poincaré had been wildly optimistic with his suggestion that knowing the state of the atmosphere to an accuracy of a tenth of a degree might explain why a hurricane happens *here* rather than *there*.

In 1959 Edward Lorenz was a thirty-two-year-old mathematical meteorologist working at Massachusetts Institute of Technology (MIT). Even without being able to write down a complete set of equations to describe the state of the atmosphere at any time, he had the idea that in principle such equations could describe any state of the atmosphere, no matter how unrealistic or implausible, as starting conditions, but that if a simulation were set up, corresponding to such a state, natural processes would then drive the model atmosphere into one of a relatively small number of stable states – what we would now call attractors, although this was not the language used by meteorologists in the 1950s. Pioneers of mathematical weather-forecasting had used linear techniques, for the usual reasons that they had to start somewhere and linear equations are easier to work with than non-linear equations. Lorenz went further by developing a computer model of a very simplified version of the atmosphere (physicists sometimes call such simplifications 'toy' models) based on a set of twelve non-linear equations. The model had to be very simple because the computers available in 1959 were very simple – Lorenz used a machine the size of a large desk, which had a memory of just 4k 32-bit words, rather less powerful than the chip inside the average digital watch or washing

machine today. He was able to show that linear forecasting techniques were ‘far from perfect’<sup>9</sup> and went on to study the behaviour of his simulations in their own right.

Just as the computer power available at the time was limited, so were the output devices – no colour laser printers or flat-screen monitors. The output from Lorenz’s computer was simply a printed list of numbers, which he programmed the computer to round off to three decimal places in order to fit a complete set of twelve numbers on one line of output paper. The model moved forward in steps of six simulated hours, and simulated a day of weather in one minute. Getting really sophisticated, Lorenz then programmed the machine to print the details for just one or two variables on each line, with the distance of the symbol ‘up the page’ representing the value of the variable; he could then draw a line by hand through these points to display the changing value of the variable (the wind direction, for example) as a graph. It was at this point that serendipity struck.

Lorenz decided that he wanted to take a second look at part of a simulation he had already run, extending it further into the future, and instead of starting from the beginning and waiting until the computer caught up to the interval he was interested in, he typed in as the starting parameters the set of numbers from part-way through the old run, reading them off the paper, and then went off to have a cup of coffee while the machine chugged away. After an hour or so, during which time the machine had simulated two months of toy weather, he came back to check progress, and found that the numbers being printed bore no relation to the numbers for the corresponding ‘days’ of the earlier run. At first, he thought the computer was faulty. But when he compared

the two printouts closely, he found that the second run started out like the corresponding section of the first one but deviated more and more from it, with the errors doubling in size for every four days of the simulation – non-linear growth.

Lorenz soon realized what had happened. The numbers he had typed in were the three-decimal-place numbers from the printout; but inside the computer the numbers were calculated to six decimal places. Where Lorenz had typed 0.506, for example, the number from the first run had actually been 0.506127. The model was so sensitive to the initial conditions that the difference, one quarter of one tenth of 1 per cent, made the two runs diverge completely from one another after a short time. If the real atmosphere was as sensitive as this to its initial conditions, then there was no hope of using numerical forecasting techniques to predict the weather more than a few days ahead. Lorenz announced his discovery in a fairly low-key way at a scientific meeting held in Tokyo in 1960, and went on to develop these ideas in the following years, but the full significance of his work was not appreciated, as we shall see, until much later. In terms of weather-forecasting, Lorenz was able to show that the real atmosphere can indeed be very sensitive to initial conditions. He used a simple mathematical (numerical) model of convection, which only hints at the complexity of the real weather, but which shows a key feature of sensitivity to small changes. You can picture this in terms of our model of phase space as a landscape flooded with water – imagine two deep pools in phase space, representing equally powerful attractors, separated by a sand bar with a very shallow layer of water over it, linking the two pools. Typical trajectories in phase space will circle round and round one

or other of the pools, but some rare trajectories cross over the sand bar and into the other pool, where they circle round and round before crossing back into the first pool. The nub of Lorenz's discovery is that without knowing the precise location of a point in phase space, it is impossible to predict exactly when the trajectory passing through that point will cross over to the other side, so the flips from one state to the other seem to occur at random. Or, from a slightly different perspective, the trajectories that wander over the sand bar are very susceptible to tiny perturbations, and a small nudge from outside may be all that it takes to shift a system from a trajectory heading back into one pool on to a trajectory heading across into the other pool. It is because of this kind of process at work that the limit of accurate weather-forecasting is about ten to fourteen days, and that the real weather can also flip from one stable state to another stable state in an essentially unpredictable way.

It is the question of *precise* determination of initial conditions that Lorenz drew attention to and which lies at the heart of the modern understanding of chaos. This sensitivity of the weather (and other complex systems) to initial conditions is sometimes referred to as 'the butterfly effect', after the title of a paper that Lorenz presented to a meeting in Washington DC in 1972 – 'Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas?'<sup>10</sup> The analogy should not be taken entirely seriously, since on this scale in the real world there are so many processes involved, blurring out individual contributions, that it is unlikely that a flap of a butterfly in Brazil could ever be identified as the cause of a specific tornado in Texas (or, conversely, that a flap of another butterfly in China prevented a tornado from forming in Texas). But it does provide a striking metaphor for chaos. Coincidentally, when

the double-pooled attractor discovered by Lorenz is plotted out graphically on paper (or a computer screen), it forms a pattern rather like a butterfly's wings; it is usually known as the Lorenz attractor (although Lorenz himself calls it the Butterfly attractor), and has become the most familiar and striking visual image for chaos, even for people who don't know what the image actually represents.

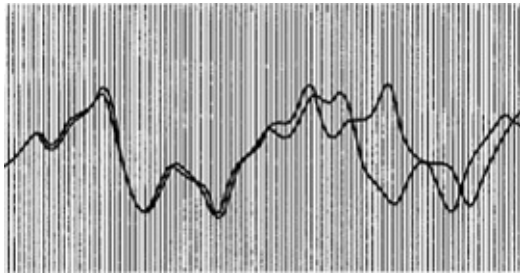


Figure 2.2 As with the orbit of the satellite in the three-body problem, Edward Lorenz found that computer 'predictions' of weather properties (such as temperature) diverge widely from almost identical starting conditions.

The weather, it turns out, is sometimes more chaotic than at other times. When meteorologists run the numerical simulations these days, they don't just take the exact data from their grid points of observations as the raw material for a single forecast. Instead, to find out if the errors and uncertainties inherent in the observations matter significantly they usually run each forecast several times with slight variations in the starting conditions. If all the forecasts come out more or less the same, they know that the overall pattern of the forecast(s) can be trusted – the weather system is, if you like, circling one of the deep pools in phase space. But sometimes when they run the same forecast with slightly different starting conditions they get widely different 'predictions' for the weather a few days ahead. This tells them that the weather is in a

chaotic state, equivalent to the trajectories wandering over the sand bar in phase space, and none of the forecasts can be trusted. This is why the forecasters on TV sometimes seem more certain about their forecasts than they are on other occasions. As one forecaster has ruefully acknowledged, 'we can predict the weather accurately provided it doesn't do anything unexpected'.<sup>11</sup>

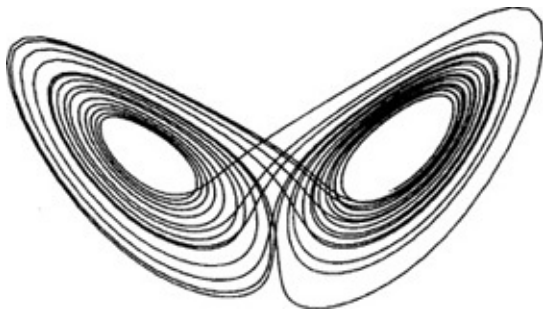


Figure 2.3 *The Lorenz ('butterfly') attractor.*

You can see chaos at work in your pocket calculator, using the technique of iteration that is so important in all this kind of numerical work. Take a simple expression,  $2x^2 - 1$ . Choose a value for  $x$  between 0 and 1, with a good few decimal places – say, 0.2468. Use this number as  $x$ , and work out the result. Then, use that number as  $x$ , and feed it back into the calculation for the next iteration. You get what looks like a string of random numbers, produced by a completely deterministic process obeying a simple law. Now try again, starting with a number that differs from your first number by just one part in the last decimal place – say, 0.2469. If you have the patience to keep stabbing away at the keys of the calculator (or the nous to write a simple computer program to do the work for you) you will find that you end up with a completely different string of random-looking numbers after the first few

iterations. Yet in both cases every number is precisely determined by the one before by the feedback process – the pattern of numbers is entirely deterministic. This is just the kind of thing that Lorenz observed – and if you try any of this on different calculators, don't be surprised if you get different 'answers', because different calculators round off numbers in their internal calculations in different ways.

You can also see a different kind of behaviour. Take the expression  $x^2 - 1$ , which looks very similar to the first expression. Now, for any value of  $x$  you start out with (between 0 and 1), after a while the pattern settles down into a stable state, oscillating between 0 and -1. It is said to be 'periodic with period two' because once the system has settled down into this pattern, wherever you start from it takes two steps to get back to where you started. And you can have periodic behaviour in other systems where it takes more steps to get back to where you started, but always the same number of steps for that particular system, wherever you start from. One simple law gives periodic behaviour, converging on an attractor; another simple law, seemingly very similar to the first law, gives 'random' behaviour that is very sensitive to initial conditions. If you have a taste for this kind of thing, you can explore further;<sup>12</sup> you'll find that some iterations converge on a single value and get stuck there (a simple attractor, a system sometimes referred to as having 'period one' because just one iterative step takes you back to where you start), some are seemingly random, and some are periodic, with period two being just a simple case hinting at much more complicated possibilities. All of these patterns of behaviour are seen in the real world, where they may apply to things as diverse as the dripping of a tap, the way populations of wild animals change, or the oscillations of the stock

market. But in essence we have already found the underlying simplicity from which chaos and complexity emerge – simple laws, non-linearity, sensitivity to initial conditions and feedback are what make the world tick. Before we move on to really complex things, though, it seems appropriate to round off the story of planetary orbits, which is where the whole thing started.

Just as with weather-forecasting, the key to understanding the long-term properties of the Solar System came with the advent of reasonably fast electronic computers capable of carrying out long numerical integrations relatively quickly. Poincaré had proved that the Solar System was in principle subject to what we now call chaos, but it was then, as it is now, quite clear that in practice planets such as the Earth follow stable orbits for a very long time – otherwise, we would not be here to wonder about such things. The way to get a handle on chaos within the Solar System was to look at systems more like the restricted three-body problem, where one small object (or a series of small objects) moves under the gravitational influence of two large objects. The largest objects in the Solar System are the Sun and the planet Jupiter, and there is a whole family of small objects under their gravitational influence, the thousands upon thousands of rocky bodies known as asteroids which orbit in a belt around the Sun between the orbit of Mars and the orbit of Jupiter.

The planets of the Solar System form two families, separated by this asteroid belt. Closer to the Sun there are four small, rocky planets (Mercury, Venus, Earth and Mars); further out there are four large, gaseous planets (Jupiter, Saturn, Uranus and Neptune).<sup>13</sup> The basic reason for this division is clear – when the planets were forming, the



heat of the young Sun drove most of the gaseous material away from the inner part of the Solar System; but in the outer regions it was cool enough for the gas to accumulate on the four giant planets. Nor is the origin of the asteroid belt much of a mystery today. Both computer models and studies of dusty discs of material around young stars suggest that planets form 'bottom up' as grains of dust stick together to make larger grains, then pieces of rock, and so on. In the region between Mars and Jupiter, however, the gravitational influence of Jupiter disrupted this process before it could go to completion, and prevented a single whole planet forming. The latest simulations suggest that by the time Jupiter grew big enough to begin to dominate this part of the Solar System gravitationally, perhaps six or seven objects the size of Mars (itself one-tenth the mass of the Earth) had formed, but under the influence of Jupiter they collided with one another violently, and instead of sticking together to make one big planet they were broken up and fragmented to form the asteroid belt, leaving Mars itself as the sole survivor.

But this detailed model of the formation of the asteroids was yet to be developed in the late 1970s, when a PhD student at Caltech, Jack Wisdom, decided to study the present-day orbits of the asteroids using a new numerical technique and the best computer available. Asteroids had only been discovered in the nineteenth century, because they are small and faint. But as the numbers observed grew, it soon became clear that their orbits are not spaced out evenly across the asteroid belt, but that there are gaps corresponding to essentially empty orbits. These are known as Kirkwood gaps, after the American astronomer Daniel Kirkwood, who drew attention to them in the late 1860s. It was clear

from the beginning what was special about these orbits – they correspond to certain so-called ‘resonances’ with the orbit of Jupiter itself. But until Wisdom (and his computer) came along, it was not at all clear how these resonant orbits were kept clear.

We have all experienced, and used, resonance when playing on a swing as a child. Resonance is a way of getting a large return for a relatively small effort, by making the effort at just the right time and pushing a system the way it ‘wants’ to go. The small amount of ‘pumping’ that we do on the swing has to be timed just right to make the swing travel through a bigger and bigger arc. In the same sort of way, Jupiter gives a nudge to an asteroid in its orbit every time the asteroid passes by Jupiter, between Jupiter and the Sun. For most orbits, these nudges are more or less random, occurring at different times in each orbit, as if you stood on a swing and jiggled about aimlessly, so there is no big effect. But resonant orbits are different. Suppose an asteroid is in an orbit around the Sun which takes exactly twice as long as the orbit of Jupiter. Then, the asteroid always gets the nudge from Jupiter at the same part of its orbit. The influence of the giant planet can add up to disturb the orbit, in the same way that by pumping the swing at the right time you can make it swing further and further out of the vertical. This wouldn’t matter much if the orbit was stable in the traditional sense, and a small perturbation simply made it wander a little way out and back again (like the long-term changes in the orbits of Jupiter and Saturn studied by Laplace). But if the orbit is one that is sensitive to perturbations, the resonance could quickly switch the orbiting asteroid from its more or less circular orbit around the Sun into a much more elliptical orbit, but one which still comes back with the same regularity

to the resonant point with Jupiter. That is exactly what Wisdom found, and published in 1982 – chaos at work in the asteroid belt, in particular for orbits close to the 3:1 resonance with Jupiter.

But even that wasn't quite the solution to the puzzle of the Kirkwood gaps, because a perturbation that can switch asteroids out of these gaps could just as easily work the other way around, switching asteroids from elliptical orbits back in to plug the gaps (remember the double-lobed attractor, like the one discovered by Lorenz). Wisdom pointed out that the solution to this piece of the puzzle came with the realization that the asteroids being ejected from the Kirkwood gaps as a result of chaos, under the influence of Jupiter, were moving into orbits that crossed the orbits of the inner planets, including the Earth. Many of these asteroids, over the long history of the Solar System, have been taken out of the relevant equations entirely, as a result of colliding with one of the inner planets, so there is no reservoir from which to draw the asteroids to plug the gaps. All of the inner planets of the Solar System (and our Moon) show on their battered faces the scars of many such impacts, and it is widely believed that one such impact on Earth, about sixty-five million years ago, brought an end to the era of the dinosaurs and opened the way for the rise of the mammals, including ourselves. This implies that we owe our existence directly to the effects of chaos at work in the asteroid belt, and carries the chilling corollary that the end of civilization may come about in the same way.

For what it is worth, though, we can take comfort from the knowledge that the end of civilization will *not* come with the Earth switching suddenly into an orbit that either takes it plunging close by the Sun or out into the depths of space. Over the years since 1982, the

computers have got better and faster, and the mathematicians have devised more cunning programs, so they have been able to look a long way into the future (or possible futures) of the Solar System. Like their counterparts in the space sciences, the researchers involved seem to delight in constructing tortuous acronyms or punning names for their projects, two of which were known as LONG-STOP (from Long-term Gravitational Study of the Outer Planets) and the Digitary Orrery (an orrery is a mechanical, sometimes clockwork, working model of the Solar System). Some of these studies look hundreds of millions of years into the future, and show that although the orbits of the planets are technically chaotic, unlike the situation for asteroids in the Kirkwood gaps there is only a vanishingly small prospect that any of the orbits of the planets proper, including that of the Earth, will change drastically during the remaining lifetime of the Sun in its present form – that is, over the next five billion years or so. But these calculations do not predict ‘the’ orbit of the Earth, or any other planet. They are sensitive to the initial conditions, although in a fairly restricted way. For example, one model, developed by Jacques Laskar at the Bureau des Longitudes in Paris, showed that the Earth will follow essentially the same orbit for at least 200 million years (the limit of that particular numerical integration). But it also showed that a difference of just 15 metres in specifying the position of the Earth at the start of the integration doesn’t stay as an ‘error’ of 15 metres at every step in the iteration but, because of non-linearity, grows into an uncertainty so big after 100 million years that we cannot say where in its orbit the Earth will be at any chosen moment that far ahead. The error has grown as big as the circumference of the Earth’s orbit, nearly 950 million kilometres. The rounding errors

inherent in any numerical integration – the choice of just how many decimal places you calculate with, whether using pencil and paper or an electronic computer – affect the outcome of the calculation in very much the same way that Lorenz’s simple weather model was affected.

One consequence of this is that if you use such a model to calculate the orbit of the Earth (or any other orbit) over, say, the next 100,000 years, and then reverse the calculation, you don’t get back to where you started. If the system is reasonably stable, you will get back somewhere near where you started, in the same corner of phase space, but perhaps all you will be able to say is that the planet concerned stays in more or less the same orbit around the Sun. The calculated trajectories are non-reversible, even though we have been brought up to think that objects obeying Newton’s laws ought to follow reversible trajectories. As long as this is just an artefact of how we do the calculations, it doesn’t change the way we think about the Universe very much. For planets like the Earth, beyond a certain number of decimal places the orbit we compute (or the trajectory through phase space) doesn’t change very much (although it *does* change) even if we do change the rounding-off process in the computer. We get a different orbit that is almost exactly the same as the previous one we calculated. Every time we run the calculation with slight modification to the initial conditions we get a similar orbit. All the orbits occupy much the same region of both real space and phase space, so that although the orbit moves about chaotically from one possibility to another, it doesn’t move outside this range of possibilities on any timescale relevant to human beings. It is a kind of restricted chaos. An analogy is sometimes made to the behaviour of the ball in a roulette wheel, after it has been thrown into the spinning wheel but

before it has settled into one of the available holes. It bounces around in an essentially chaotic fashion, but those bounces are all restrained to 'orbits' inside the rim of the wheel (assuming the croupier has done his job properly). A more homely analogy would be with someone who saves the hassle of selecting a pair of socks in the morning by only buying green socks. He knows that he can pick two socks at random from the drawer every day and they will both be green, although they are different socks with different detailed properties. Even with his eyes shut, he will never pull a red sock from the drawer.

This raises one point which needs to be addressed, even though it is very much an aside from the main thread of our story. There are people who believe that the inner planets of the Solar System, in particular Venus, have changed their orbits dramatically within the past few thousand years, and that these changes are described in various ancient myths and legends. This belief is based on a very imaginative interpretation of those ancient myths, and a very poor understanding of orbital dynamics and such fundamentals as the law of conservation of momentum; unfortunately, though, when respectable astronomers talk of chaos in the Solar System, this is grist to the mill of such fantasies, and you may see it suggested that chaos theory 'explains' or 'proves' that Venus only recently moved into its present orbit. It does nothing of the kind. The calculations can be run as well backwards in time as forwards, and they show the same kind of pattern – we cannot say *exactly* where Venus was in its orbit on, say, 4 July, 5 million BC, but we can say with great confidence what kind of orbit it was in then – one just like its present orbit (a green sock, not a red one). There is no chance at all that Venus, or any of the other planets, has experienced a dramatic orbital

change at any time during (conservatively) the past five million years, which more than covers the time since our own ancestral line split off from that of the other African apes. For sure, no human being has ever seen Venus wandering about the sky as it changed its orbit.

Before we leave the planets, though, it's worth mentioning that chaos doesn't only affect orbits. Planets spin on their axes (the Earth once every twenty-four hours), and they wobble as they spin, because of the gravitational influence of the Sun (in a similar way, a child's spinning top wobbles as it spins, because of the gravitational influence of the Earth). Resonances can occur between the period of the spin and the period of the wobble, which in the right part of phase space can lead to a sudden, chaotic change in the angle of tilt of a planet (its obliquity). This might happen, for example, if the spin of the planet gradually slows down and as a result the system enters a sensitive region of phase space. The tilt of the Earth out of the vertical (relative to a line joining the Earth and the Sun) is about 23 degrees today, and this is the cause of the seasonal cycle of weather. But it turns out that the presence of our large Moon acts as a stabilizer which prevents any wild variations in this tilt. That is not the case for Mars, or the other inner planets, where there is no large moon to act as a stabilizer (the two tiny moons of Mars, thought to be captured asteroids, are totally inadequate for the task). It happens that the average tilt of Mars is about 24 degrees, but computer simulations show that this can change dramatically over a very wide range of values, at least  $\pm 20$  degrees outside this average. There is also direct evidence that Mars has experienced dramatic changes of climate in the past, in the form of what seem to be dried-up riverbeds smallouring the now-arid surface of the Red Planet; these extreme

climatic events may well be related to epochs in the past when the frozen polar region has nodded towards the Sun, producing an extreme of summer heat, and warmed sufficiently to make the polar cap of frozen carbon dioxide or water evaporate or melt. Probably the same sort of wobble has happened on Venus and Mercury, but there are no visible traces on the surface of the planets, because on Mercury there is no atmosphere and on Venus the surface seems to have been turned over by intense volcanic activity.

This leaves the Earth, happily stable at present. In the very long term, though, as the Moon recedes from the Earth as a result of tidal forces, its influence will wane and in the distant future the tilt of our planet will also vary chaotically, perhaps suddenly increasing to as much as 90 degrees, so that in summer it will be as hot at the pole as it is now at the equator, while the other pole experiences a severe, dark winter, and six months later the situation is reversed. This would hardly be ideal for life as we know it – so although chaos in the Solar System may have been responsible for our own existence (through the impact of the asteroid that finished off the dinosaurs), it is the *absence* of chaos in the changing obliquity of the Earth that has, thanks to the presence of the Moon, allowed life on Earth to evolve under more or less stable climatic circumstances for billions of years. There is, though, another twist in the tale. The best suggestion we have for the presence of such a large Moon associated with the Earth is that it was produced in the early stages of the history of the Solar System when a Mars-sized object from the asteroid belt was diverted from its orbit through the workings of chaos and collided with the Earth, splashing molten material that then formed the Moon out into space.<sup>14</sup>



There is one last point—perhaps the most important point of all—that we want to draw to your attention before we leave the investigation of the Solar System. We pointed out a little earlier that if you use a numerical integration to calculate the orbit of a planet and then reverse the calculation, you don't get back to where you started. The calculated trajectories are non-reversible. And we said that as long as this is just an artefact of how we do the calculations, it doesn't change the way we think about the Universe very much, since Newton's laws are still reversible in principle. But how does the Universe 'do the calculations'? Like Laplace, you might imagine a perfect intelligence (living or electronic) which would be able to store exact information on all the relevant properties (such as the position and momentum of every particle), and then carry out the relevant calculations perfectly. Surely, they would then be reversible. But just how many decimal places do you need to achieve perfection? How big would the memory of our perfect computer have to be?

At first sight, that seems an impossible question. But it turns out to have a very simple answer, concerned with the nature of numbers. When most people talk about numbers, they think of integers—numbers like 1, 2, 27, 44,196, and so on. We are also used to simple fractions  $\frac{1}{2}$ ,  $\frac{1}{3}$ ,  $\frac{1}{4}$  and the like. But for most purposes in everyday life, that's as far as it goes. Even when we encounter decimals, it is usually only in terms of something like money, and then usually to just a couple of decimal places, so that we are all familiar with the idea that £17.46 means 'seventeen pounds and forty-six pence'. But there is a literal infinity of other numbers that never come into consideration in everyday life. Worse—there is an infinity of numbers in between any two numbers that

we can think of. This is fairly obvious in terms of decimals. Take the integers 1 and 2. Between 1 and 2 there must be a set of numbers with one decimal place (1.1,1.2... 1.8,1.9), and between each of those numbers there is another set of numbers with one more decimal place (1.11, 1.12, 1.13 ... 1.18, 1.19), and between each of those numbers—but you get the picture. However many decimal places there are in the two numbers you start with, the same reasoning tells you that there are infinitely many numbers in between them—even for numbers like 247.8503468295667 and 247.8503468295668, or numbers with enough decimal places to fill this entire book that differ only in the last digit, or numbers big enough to fill the entire Universe with books, that differ only in the last digit.

This wouldn't matter much if it were just a mathematical curiosity. We know, for example, that some infinities are well behaved (as in the case of some infinite series) and can be represented in simple ways. The fraction  $\frac{1}{3}$ , for example, corresponds to a decimal with an infinite number of digits—0.3333333333 ... But it can be written in a very simple, compact form, without filling the Universe with books all containing boring rows of 3s. Other kinds of decimal expression can also be represented compactly. The decimal 0.675486754867548 ..., for example, can be described by writing 'take the digits 67548 and keep repeating them for ever'. Such a compact form of describing something too big to handle conveniently if spelled out in full is sometimes called an algorithm (and might be represented either in words or by a simple mathematical expression); an expression which can be represented compactly in this way is said to be algorithmically compressible. But even the Ancient Greeks, who never came up with the idea of decimal

numbers, knew that there were some numbers which are not algorithmically compressible and cannot be written in compact form. Worse – *most* numbers cannot be written in such a simple form.

The numbers which can be written in a simple form are fractions—they are the ratios of one integer to another, like  $\frac{4}{3}$ , or  $\frac{49847}{65169}$  (even the integers are, in a sense, ratios, and could be written as  $\frac{1}{1} = 1$ ,  $\frac{2}{1} = 2$ , and so on). Because they involve ratios, the Greeks called these rational numbers. But there are other numbers, the most famous being  $\pi$ , which can never be written as the ratio of two numbers. For that reason, they are called irrational numbers, and can only be expressed as an infinite string of non-repeating decimal digits.<sup>15</sup> The Greeks, lacking decimals, didn't put it like that, but they were well aware that such irrational numbers existed. And that *most* numbers are irrational! This strikes at the heart of chaos and the Newtonian (or Laplacian) concept of reversibility, as we can see by trying to specify the exact location of a system in phase space.

The 'system' can be very small – a single particle moving around in a box, under the influence of gravity. The state of the particle is determined by its position and momentum, and Newton proved that it will behave as if all its mass were concentrated in a mathematical point at the centre of the particle. 'All' that we have to do to determine the location of the system in phase space is to specify the position of that point, and the momentum of the particle. But let's concentrate on position, simplifying the problem even more. We can simplify things yet further by assuming that the particle is moving in a straight line—perhaps, falling under the influence of gravity. Now, we just have to specify the position of the particle on that line, the simplest problem

imaginable in physics. But this turns out to be impossible, except in very rare cases. Suppose we know that the particle is somewhere between two points A and B. We need to know what *precise* fraction of the distance between A and B it has already covered. No problem if it is  $\frac{1}{3}$  of the way down the line, or  $\frac{98}{317}$ ths of the way along, or any other rational fraction. But between every pair of points on the line represented by rational fractions there are infinitely many points represented by irrational numbers, and each of those points can only be specified by an infinite string of digits, with no way to compress this into a compact form. If, for example, the particle is  $\frac{1}{\pi}$  along the line between A and B, we can represent this as accurately as we like by working out the expression to as many decimal places as we like, but we can only represent this *exactly* by writing out an infinite number of digits. And this is just for a single particle moving in a straight line! In a system that is sufficiently sensitive to initial conditions, it is always possible that, no matter how many digits we choose to work with, the entire future of the system may, as Lorenz discovered, depend significantly on the value of the next digit, the one we have in effect thrown away.

This means that a computer with an infinite memory is required to specify the state of a *single* particle. No computer can be bigger than the entire Universe, and if you define the Universe as ‘everything there is’, this means that the only system that can replicate the behaviour of the Universe in every detail is – the Universe itself. Even if, as Laplace thought, the Universe is entirely deterministic and the whole future is contained within its present state, there is no way at all to predict or know the future, except by watching the Universe evolve. Whether or not we have free will, the Universe behaves as if we have free will,

which is really all that matters. The Universe is ignorant of its own future, and is its own fastest simulator.

But what about reversibility and the arrow of time? People used to talk glibly about ‘waving a magic wand’ to reverse the motion of every particle in the Universe (or in a box of gas) to make time run backwards. But this is now seen to be impossible. Even leaving aside such subtleties as the theory of relativity, which raises profound questions about how every motion in the Universe could be reversed *simultaneously* when signals can travel no faster than the speed of light, and what is meant by ‘simultaneous’ depends on where you are observing the Universe from, it is impossible to reverse precisely the motion of even a single particle. To do that, you would (except in very rare cases where the motion is, for a split-second, described by rational numbers) first have to specify its present state to an infinite number of decimal places, in order to make the exact reversal. This is not possible in principle, not just because of human inadequacies. This seems a more than adequate bonus to receive for learning that chaos is inevitable. The Universe *cannot*, in principle, be predicted in all its detail; but, equally, time *cannot*, in principle, be reversed.

The ideas presented in this chapter are the bedrock—the deep simplicity—upon which the complexity of the Universe is based. Building from this bedrock, we can move on to see how the complexity of the Universe, and life itself, has emerged from the chaos. If you are already familiar with chaos and fractals, you can conveniently skip the next chapter, and move on to Chapter 4; but I hope you will stay with us for a brief but breathtaking ride into the realm of chaos itself, and a discussion of that related topic which has become *de rigueur* in any

investigation of chaos, the theory of fractals. This is not so much of an aside as you might think, since it takes us to the borderland of life itself; and understanding the emergence of life is the ultimate aim of our discussion.

### 3

## Chaos out of Order

The kind of chaos that we are talking about now, what scientists of the twenty-first century mean by ‘chaos’, is not the same as the kind of chaos that the Ancients referred to, or what we mean by chaos in everyday life. That kind of chaos is completely random and unpredictable even in principle. But the kind of chaos we are discussing here is completely orderly and deterministic, with one step following on from another in an unbroken chain of cause and effect which is completely predictable at every stage –in principle. It is just that in practice it is impossible to predict in detail what is going to happen more quickly than events unfold in real time. The classic example of this comes from the study of turbulence, which occurs in many places, but whose importance we can highlight in one of the simplest examples, the way the flow of water in a river changes as the current increases.

For our simple model, we can imagine a smoothly flowing river in which there is one large rock sticking out above the surface. The flowing water of the river divides around the rock, and joins up again seamlessly on the other side, so that small chips of wood floating on the water follow these ‘streamlines’. If there is rain upstream, the flow of the river will increase, and as it does so it goes through at least three distinct changes, all of which you have probably seen without thinking too

deeply about their significance. First, as the flow builds up, little whirlpools appear behind the rock. These vortices stay in the same place, and a chip of wood floating downstream may get trapped in one of them and go round and round for a long time. This is very similar to the kind of behaviour seen in phase space in the Lorenz attractor—a kind of whirlpool around which a system cycles for a long time. In phase space, this kind of attractor is known as a limit cycle, because wherever the system starts out, in the limit it will be attracted to the specific repetitive pattern of behaviour.<sup>1</sup> Using slightly sloppy (but reasonably clear) language, we might call each little whirlpool behind the rock a ‘limit cycle’.

At the next stage, as the speed of the water flowing downstream increases, vortices form behind the rock, but don’t stay there. They detach themselves (or are detached by the flow), and move away downstream, retaining their separate existence for a while before dissolving into the flow. As they do so, new vortices form behind the rock and detach in their turn. Now, a chip of wood might get caught in one of these eddies and be carried downstream still circling around in the vortex as long as the vortex lasts.

As the flow of water increases still further, the region behind the rock where the vortices survive gets smaller and smaller, with vortices forming and breaking up almost immediately to produce a choppy surface in which there are seemingly only irregular fluctuations – turbulence. Eventually, when the flow is fast enough, all trace of order in the region behind the rock disappears. No vortices form and the entire surface of the water breaks up behind the rock into unpredictable chaotic motion.



Spelled out like this, there are two key features of the route from order into chaos that turbulence highlights. First, something is changing. It seems almost too obvious to mention, but it is central to the whole story. What looks like the same system can be described in simple terms under some conditions, in terms of chaos under other conditions, and there is a complex region in between the two where interesting things happen (in this case, the 'birth' of vortices). Just one thing – one parameter – is changing in this system, the speed with which the water is flowing. Increasing the value of that one parameter past a critical point is enough to bring about the onset of chaos. Secondly, when you look in detail at the way vortices break up behind the rock during the complex, intermediate stage between order and chaos, you find something very interesting. This discovery involves painstaking attention to detail but no very technical equipment (at least, not to see the broad pattern of what is going on). Leonardo da Vinci drew attention to it half a millennium ago. He pointed out that a whirlpool that detaches from the rock and moves downstream doesn't just disappear. It breaks up inside into smaller vortices, which in turn break up into smaller vortices, eddies within eddies within eddies in what seems to be an endless process of bifurcation. The road to chaos involves what seems to be an infinite number of choices operating on an infinitely small scale – at least, it does in the case of turbulence. Can we see anything similar at work somewhere else?

Of course, the answer is 'yes'. One place you can see something similar at work, still involving flowing water, is in the dripping of a tap. If you start with the tap turned off, and open it very slightly, it is easy to produce a steady dripping, which echoes in the sink below like the

monotonous beat of a moronic drummer – drip, drip, drip, drip ... It is a rhythm with period one, in the jargon of the trade. Open the tap a little further, and it is still pretty easy (even I can do it) to see and hear the system switch into a rhythm with period two, like a slightly more skilful drummer beating out rat-tat, rat-tat, rat-tat, rat-tat... As you open the tap still further, things begin to get more interesting, and then completely messy. After period two, it's quite hard to detect the interesting stuff. With one particular tap in the house I live in, I can open it to a point where I more or less convince myself that what I am hearing is a four-time beat – rat-a-tat-tat, rat-a-tat-tat, rat-a-tat-tat, rat-a-tat-tat – that is, a rhythm with period four. But to be completely honest I'm not sure whether this isn't just because I expect to hear it, because I know that that is what is revealed in laboratory experiments under carefully controlled conditions.<sup>2</sup> The process is known as period doubling, for obvious reasons; but it cannot continue indefinitely. At some critical point (very soon, in the case of opening the tap) repeated period doublings (repeated bifurcations) produce a complicated, seemingly random pattern of dripping, as the system becomes chaotic. Open the tap still further, and the drops merge to form a smooth flow; open it further still, and the flow becomes turbulent and messy. But we don't want to go that far now, although we will return to turbulence shortly.

The best example of the period-doubling route to chaos comes from another area of science entirely, and hints at just how fundamental and widespread the implications of chaos are. There's a very simple equation, called the logistic equation, that does a rather good job of describing how the population of a species changes from one generation

to the next. To keep it very simple, we can assume that we are dealing with some kind of insect where the entire adult population dies off in winter, after laying eggs that will hatch out to provide a new generation the following spring. We start out with a population containing  $x$  individuals. Obviously, the number of individuals in the next generation (the number that survive to breed in their turn, that is) depends on the number of eggs that hatch (the birth rate), which depends on how many eggs were laid, so if on average (averaging over males and females) each insect lays  $B$  eggs, the new population will be  $Bx$ . Except that this takes no account of individuals that fail to find food and starve, and don't survive to reproduce in their turn. This death rate depends on the size of the original population – the more individuals there are, the harder it will be for each of them to get enough food. We can simplify things a little further by setting an upper limit on the population (quite sensible, if you think in terms of, say, the number of greenfly that can inhabit a single rosebush) and dividing the actual population by this number so that the value we use for  $x$  is always between 0 and 1. This trick is called renormalization. Now, in order to take account of the premature death rate, we can multiply the growth factor  $Bx$  by a new term,  $(1 - x)$ . If the population is very low (as near to zero as makes no difference), almost every individual survives and finds food, so the growth rate is almost exactly  $Bx$ ; if the population is very high,  $x$  is very nearly 1,  $(1 - x)$  is very nearly zero, and most of the population starve or fall prey to predators. In between, the population may rise or fall from one generation to the next depending on the exact value of the birth rate,  $B$ . We can see how the population changes for different values of  $B$  by iterating the expression

$$x(\text{next}) = Bx(1 - x)$$

where  $x(\text{next})$  is the population in the next generation. Multiplying out the right-hand side of the expression we have  $Bx - Bx^2$ , highlighting the fact that this process is non-linear (because of the  $x^2$  term) as well as involving feedback (through the iteration).

If  $B$  is less than 1, it means that the population fails to reproduce itself from one generation to the next. Each individual adult leaves, on average, less than one offspring, a recipe for evolutionary disaster, and the population must eventually die out, whatever the value of  $x$  you start with. Interesting things happen when  $B$  is bigger than 1, and you can investigate these interesting things for different values of  $B$  and  $x$  using a calculator or a computer, or even pencil and paper. From the 1950s onwards many ecologists did just that, using versions of the logistic equation (which can be adapted fairly easily to include the effects of prey on the population, or two populations interacting in other ways, such as competing for food) to try to find models which could provide insight into the behaviour of real populations of various species. But, to a large extent because of the limitations of the computing facilities available to them, they did not probe the most interesting possibilities, and concentrated on the more orderly implications of the equation, equivalent to hydrodynamicists ignoring turbulence and concentrating on the state of the river when eddies are forming and breaking away from behind the rock in an orderly fashion.

If  $B$  is bigger than 1 but less than 3, it turns out that there is what we would now call an attractor for this simple logistic equation. Whatever value of the population you start with (any value of  $x$  between 0 and 1),

after a sufficiently large number of generations it settles down to a steady level, corresponding to a constant population. The exact value it settles down at rises slightly as  $B$  is increased, and for values of  $B$  close to (but less than) 3 it settles down at 0.66, corresponding to  $\frac{2}{3}$  of the maximum possible population. It may start small and rise, overshooting this value and then zigzagging either side of 0.66 in alternate generations, getting closer and closer to the attractor (like the approximation for  $\pi$  mentioned at the beginning of Chapter 2), or it may start large and fall towards this value, overshooting and oscillating in the same way. But if you carry out enough iterations, it always settles down to 0.66. If you make  $B$  bigger, the oscillations get more extreme and take longer to settle down, but as long as  $B$  is less than 3 they do eventually converge on the attractor. But when  $B$  reaches 3, something different happens.

As soon as  $B$  is just a little bit bigger than 3, the pattern changes. Now, once the iteration has been carried out enough times, the population switches between two different, constant levels in alternate generations. The single attractor has split in two (bifurcated) and the period has doubled, from 1 to 2. This can be understood in terms of real populations. In one year, there is a large population which eats all the food. As a result, many individuals starve and die without reproducing. So the next generation has a small population, all of which find plenty of food and lay eggs – and so on. If you plot all this out on a graph, for values of  $B$  starting below 3 and rising above 3, you get a diagram that looks something like a tuning fork laid on its side, as the single attractor bifurcates.

All this is straightforward enough to explain in words and plot out on

diagrams, once the calculations have been done, but very tedious to calculate. Calculating all the iterations for just one value of  $B$  is boring enough, and in order to look closely at what happens near the critical value 3 you need to carry out many iterations for many slightly different values of  $B$ . The first person to probe the behaviour of the logistic equation in the kind of detail needed to show how the overall pattern of behaviour changes as  $B$  is increased was the Australian-born physicist turned ecologist Robert May, working at Princeton in the early 1970s. Then in his late thirties, May had just the right kind of background to apply ideas from physics and mathematics in biology, and it was just the right time and place to take advantage of the increasing power and speed of electronic computers. Having found the bifurcation at  $B = 3$ , the obvious next step was to increase  $B$  more and see what happened. The result was amazing. At a value of 3.4495, *each* of the two prongs in the tuning fork diagram bifurcates, producing a system that oscillates between four different populations (period four). At  $B = 3.56$ , each of these attractors splits in two, and the population jumps about between eight different levels; at 3.596, another doubling gives sixteen possible population levels, and by this time any biologist studying a real population of creatures obeying this simple deterministic rule would be hard pressed to see any order among the seemingly chaotic fluctuations of a population jumping between these levels from one generation to the next. The bifurcations, you will have noticed, get closer and closer together as the value of  $B$  is increased, and developments from May's early work show that at 3.56999 the number of attractors available to the population becomes infinite, so that anyone studying how the population changes from year to year will see genuine deterministic

chaos.

But there's more. Although it is mostly chaos for values of  $B$  greater than 3.56999, there are small ranges of values for  $B$  where order is restored, a kind of clear window among the confusion of the chaos. For example, if  $B$  is a little greater than 3.8 and a little less than 3.9, the system seems to settle down to a stable state resembling the behaviour when  $B$  is less than 3. But as the value of  $B$  is slowly increased, once again we see repeated bifurcations (period doublings), just like the pattern we see for values of  $B$  just above 3. Soon, we go through all the same stages as before, and chaos reappears. But the pattern we see is very similar indeed to the original pattern, except that it is on a smaller scale. And within this new, smaller-scale version of chaos there is a window, just like the window where we found order between 3.8 and 3.9; and within that window, the whole pattern repeats again. The repetition goes on for ever, with the same pattern of behaviour repeated on smaller and smaller scales, like a nested set of Russian dolls that goes on for ever. For obvious reasons, the patterns within patterns are said to be self-similar. In the midst of order, there is chaos; but in the midst of chaos, there is order. May realized that all of this had implications far outside the fields of ecology and biology, and drew attention to his discoveries in a paper published in *Nature*, the most widely read and interdisciplinary scientific journal, in 1976. This was just at the time when independent studies of various phenomena by different scientists were coming together to create chaos theory.

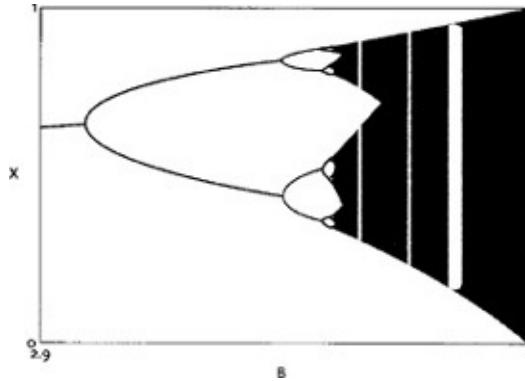


Figure 3.1a *The Feigenbaum diagram exemplifies the ‘period-doubling’ route to chaos.*

We’ve been using the term ‘chaos’ to describe this behaviour, but it was only later called that – indeed, it was while all this was going on that chaos got its name and became widely known. When Edward Lorenz had made his personal discovery of chaos in the early 1960s, he was working in the context of meteorology. Naturally enough, he reported his discovery at meetings attended by meteorologists, and published in meteorological journals – the key paper, ‘Deterministic Nonperiodic Flow’ appeared in the *Journal of the Atmospheric Sciences* in 1963. But the mathematicians, physicists and even biologists who might have picked up on the discovery didn’t read the meteorological journals, and the meteorologists who did read those journals were much more interested in how to find order among the chaos (in the everyday sense of the term) of the weather, than in looking at how to make chaos out of order. A decade later, James Yorke (1941–) was a mathematician working at an interdisciplinary institute, the Institute for Physical Science and Technology, set up at the University of Maryland specifically to try to break down some of the barriers segregating scientists into different compartments. One of Yorke’s colleagues at the University of Maryland, Alan Faller, was working in the Department of Meteorology



and had read Lorenz's 1963 paper. In a conversation with Faller, Yorke talked about his own work on non-periodicity, and Faller showed him the paper, then made photocopies which he handed out around the Institute. Yorke appreciated at once that the paper not only contained a deep truth that must have many applications outside the field of meteorology, but also that by expressing this in terms of the behaviour of a real, physical system it made the underlying mathematics more accessible to physicists. Mathematicians had been playing with numbers in ways that echoed the behaviour of the numbers in Lorenz's simple computer program for decades – but before Lorenz nobody had made a link between such mathematical abstractions and the real world. Lorenz himself did not know, at the time, how much seemingly abstract mathematical work had been carried out which could, with this insight, be translated into physical terms; but Yorke knew a man who did. On a visit to the Berkeley campus of the University of California, he passed a copy of the 1963 paper on to Stephen Smale, a topologist who had done award-winning work in his specialist field, and was also interested in dynamic systems. Smale in turn made many copies of the paper which he spread through the mathematics community. The secret was out, but it still hadn't yet received its name. That happened in 1975, when Yorke and his colleague Tien Yien Li published a paper with the title 'Period Three Implies Chaos'.

*Figure 3.1b A key feature of the Feigenbaum diagram is that it is self-similar. A small piece of the diagram blown up in scale by the right amount looks exactly like the original diagram.*

What Li and Yorke showed is that for certain families of differential equations, if there is just one solution which has period three then there

must also be an infinite number of periodic solutions, with every possible period, and in addition to that an infinite number of non-periodic solutions. This isn't quite what we mean by chaos today, and Lorenz himself prefers to call what Li and Yorke discovered 'limited chaos' because of the presence of all the non-chaotic periodic solutions also contained in the equations. It is overwhelmingly more likely that such a system will be in a periodic state, whereas in the situation that Lorenz calls 'full chaos' and most scientists today just call 'chaos', although periodic solutions exist it is overwhelmingly more likely that a system will fall into the chaotic regime. Nevertheless, although the terminology has evolved since 1975, the Li and Yorke paper is generally regarded as the one that introduced 'chaos', albeit inadvertently, as a scientific term with its modern meaning.

So by the second half of the 1970s there was a word to describe the kind of behaviour May had found in his investigation of the logistic equation. But, fascinating though all this is, it would be of little interest to anyone except mathematicians if it only applied to a simple logistic equation which, to be honest, is hardly a realistic representation of even a single biological species. Within a few years of May's breakthrough, however, in the mid-1970s Mitchell Feigenbaum, working at the Los Alamos National Laboratory in New Mexico, had shown that it had far wider implications. Feigenbaum (1945–) showed that the period-doubling route to chaos is not some special feature of the logistic equation, but that it is a product of the iterative process by which the system feeds back on itself – whether the system is an animal population, an oscillator in an electrical circuit, an oscillating chemical reaction, or even (in principle) the business cycle of the economy. What

mattered was that the systems had to be ‘self-referential’. If that were the case, then they followed the same route to chaos – not approximately, but *exactly*. Feigenbaum looked at the way the intervals between period doublings get shorter and shorter as  $B$  increases along the route to chaos studied by May. He found that there is a constant ratio in the size of one step relative to the next – to three decimal places, the ratio is 4.669:1, whether we compare the first step with the second, the second with the third, the one-hundred-and-fifth with the one-hundred-and-sixth, and so on. This pattern of behaviour is called a geometric convergence, because it takes you closer and closer to some critical point without ever quite getting there. Feigenbaum found period doubling, bifurcation and the same geometric convergence in all the self-referential systems he looked at. Not just the same kind of convergence, but exactly the same geometric convergence, with the same ratio, 4.669:1. He had found a universal truth, and received the ultimate accolade of having the mysterious number that came out of these calculations, 4.669, named after him.<sup>3</sup>

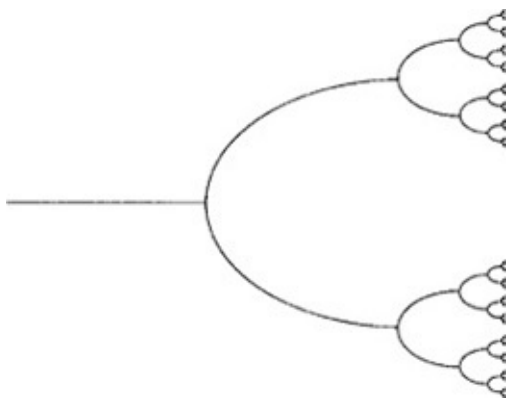


Figure 3.2a *Simplified representation of the kind of self-similar branching seen in the Feigenbaum diagram.*

There are many other examples of period doubling as a route to chaos, all with interesting things happening just at the transition from simplicity to chaos. But we don't need to go into all the details here. Instead, we can begin to see what is going on if we go back to the dripping tap and the vortices in the flowing river. The vortices also show the effect of bifurcations. Looking at what is going on in much more detail than Leonardo could, we can think of turbulence as resulting from an increasing number of regular periodic cycles added together (this was first done by the Russian physicist Lev Landau, in the 1940s). In a simple vortex, the motion corresponds to a loop around a simple attractor, the limit cycle. The next step would be to imagine a point in phase space tracing out a circle, while the centre of the circle traces out a larger circle. The resulting attractor would be a torus, a shape like the inner tube of a bicycle tyre, or an old-fashioned lifebelt. Going regularly round the small circle while the small circle goes regularly round the large circle, the point representing the state of the system follows a path like that of a curved coiled spring, or the toy known as a slinky, bent round in a circle, in a regular and predictable way.

Typically, two periodic motions in phase space interact with one another and get locked into a repeating rhythm.

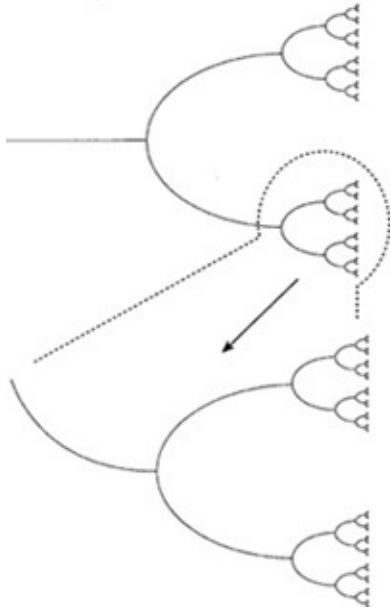


Figure 3.2b *Simplified representation of self-similarity.*

Mathematically, it is straightforward to describe further stages of increasingly complex behaviour on the road to turbulence in terms of ‘tori’ in higher numbers of dimensions. A limit cycle is a one-dimensional attractor existing in two dimensions, the surface of a torus is a two-dimensional attractor embedded in a three-dimensional phase space, and the logical next step would be a pattern of behaviour described by a three-dimensional attractor embedded in four-dimensional phase space. But the real world doesn’t go any further down this route. Turbulence occurs at the very next stage, with the point in phase space representing the state of the system staying on the two-dimensional surface of the torus but moving in a much more complicated way around the surface of the torus in a path that can never pass through the same point in phase space twice (for if it did, the system would be periodic and its behaviour would repeat) and therefore never crossing itself. Typically, in an echo of the original three-body problem, three or more periods in phase space

do not get locked into a repeating rhythm, and their combined effect is no more predictable than the orbit of the small particle in the restricted three-body problem (although it *is*, we cannot stress too strongly or too often, *completely deterministic*). The trajectory through phase space of the point describing the system would correspond to an infinitely long line wrapped in a complex way, without ever crossing itself, around a finite surface (there's actually also an additional complication, which we will deal with shortly). The Belgian David Ruelle and his Dutch colleague Floris Takens, working in Paris, dubbed such an entity a 'strange attractor', in a paper they published in 1971.

This is where fractals come into the story. Like chaos, they got their name, as we shall see, in 1975; but, also like chaos, they had been around in science for a long time before that, without their significance being fully appreciated. Entities that we would now call fractals were discovered, much to their surprise (and even horror), by mathematicians in the late nineteenth century. At the time, they were regarded as aberrations – 'monsters' that did not fit into the orderly world of mainstream mathematics. And you can see why by looking at an example which brings out their peculiarities in full force.

In 1890, Giuseppe Peano (1858–1932) published a paper in which he described how to construct a curve that completely fills a plane.<sup>4</sup> That doesn't sound too horrendous, to a non-mathematician. But think about it. A plane is a two-dimensional entity – it has length and breadth. A line is a one-dimensional entity – it has length, but no breadth. Peano showed how a line could be made to twist and turn inside a plane in such a way that it passed through every point in the plane, without ever crossing itself. A one-dimensional line completely filled a two-

dimensional plane! So how could a plane 'really' be two-dimensional, if every point in the plane lay on a single line? There's more. If you think of the plane as a square, then the Peano curve traces out the outlines of a set of smaller squares, like tiles, filling the plane; inside each smaller square, the curve traces out a similar set of 'tiles'; and so on. The pattern is self-similar, and it goes on for ever. The Peano curve is infinitely long, but it is contained within a finite area. There is a clear analogy here with the 'space filling' curve that is the attractor, wrapped around the torus in the phase space that describes a turbulent system, although none of that was known in the 1890s. The language we need to describe such entities was eventually developed by Benoit Mandelbrot, working at IBM's Thomas J. Watson Research Center in Yorktown Heights, New York, in the 1970s.

Mandelbrot, who had been born in Warsaw in 1924, had an eclectic background which served him well as the founder of what was essentially a new scientific discipline. The family moved to France in 1936, where Benoit studied in Paris after the Liberation in 1944. After spells at Caltech and the Institute for Advanced Study in Princeton, he returned briefly to France in 1955 before moving permanently to the United States and settling in Yorktown Heights in 1958. Mandelbrot realized that an object like the Peano curve could be described as having an intermediate dimension, in this case, somewhere between 1 and 2. A regular line was still a one-dimensional object, and a plane was still a two-dimensional object, but just as mathematics had to come to terms with the idea that there is an infinity of numbers between every rational number, so it would have to come to terms with the idea that there are entities with intermediate, non-integer dimensions. If the dimension of

such an entity is not an integer, then it must be fractional. In order to have a word to describe such entities, Mandelbrot tells us, ‘I coined the word fractal in 1975 from the Latin *fractus*, which describes a broken stone – broken up and irregular.’<sup>5</sup> You can just as well, though, think of it as a contraction of the word ‘fractional’, which, of course, comes from the same Latin root.

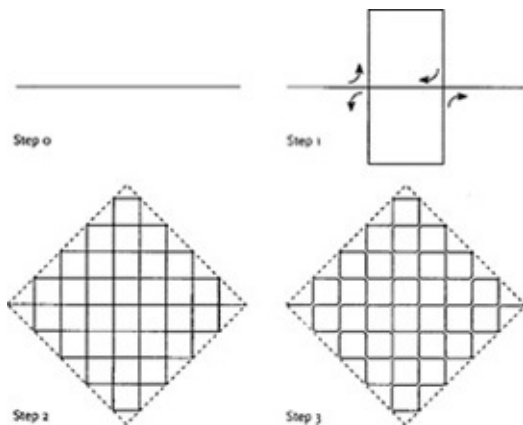


Figure 3.3 *The Peano curve – a line that thinks it is a surface.*

There are three other fractals, all of them known for decades before 1975 as mathematical scary monsters, that it is worth learning about before we get on to the relevance of all this to chaos. The oldest of these (in the sense of being the first to be discovered by mathematicians) is the Cantor set, named after the German mathematician Georg Cantor (1845–1918), who discovered it in 1883.<sup>6</sup> The Peano curve is a line that is more than a line, ‘trying’ to be a plane; the Cantor set is a line that is less than a line, ‘trying’ to be a point. It is easy to describe how to construct a Cantor set. Take a straight line, and rub out the middle third of the line (but be careful not to rub out the points exactly one-third and two-thirds of the way along the line). You are left with two lines, each one-third of the length of the original, separated by a gap, the same size



as each of these two lines. Now carry on doing this – iterate the procedure until all the lines have been rubbed out, and you are left with an infinite number of points arranged in a regular pattern with gaps between them. This is the Cantor set, a fractal with dimension intermediate between 0 (a point) and 1 (a line), a kind of ghost of the original line, like the fading grin on the face of the Cheshire cat in *Alice in Wonderland*. Simple though it is (and rather dull visually), the Cantor set turns out to be crucially important in chaos, as a summary of its key properties highlights. First, it is produced by iteration (feedback), which, as we have already seen, is one of the keys to chaos. Secondly, it is self-similar. From the second stage of our iteration (where we have four lines remaining) onwards, at each step the Cantor set is made up of two exact copies of itself, scaled down in size by a factor of one-third. And there is something else about the Cantor set. If we go back to the bifurcation diagram which represents the period-doubling route to chaos, right at the point where chaos sets in (sometimes referred to as the Feigenbaum point, the limit to which the ever-decreasing intervals between bifurcations tend), at the last step of the process before chaos sets in, all the tips of all the branches of the bifurcation tree form a Cantor set, hinting at (or rather, blazing out in neon lights) the deep relationship between fractals and chaos.

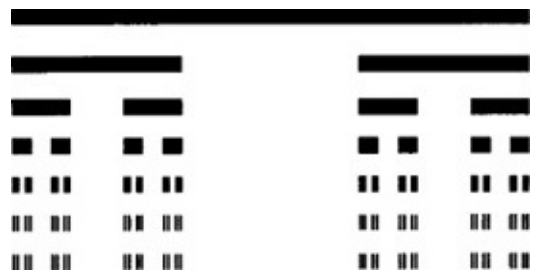


Figure 3.4 *The Cantor set. By rubbing out the middle thirds of each line as described in the text, you*

*are left with a dust containing infinitely many points but with a total length of zero.*

The Cantor set also has an important place in the history of fractals because it was one of the routes which led Benoit Mandelbrot into the work that made him famous. Mandelbrot was interested in all kinds of time-varying and space-varying phenomena and patterns as diverse as the way words are distributed in text, the distribution of large and small cities across a country, and the ups and downs of the stock market. Soon after he became an IBM researcher, he tackled a problem that was of great practical importance to the company. Then as now, information was carried from one computer to another along telephone lines, and the engineers involved in this work were plagued by noise that occurred in these lines and could scramble vital pieces of data. This didn't matter too much for the voice conversations the phone lines were originally built for – if there is a crackle on the line during a conversation you can speak a bit louder, or repeat what you were saying. But it played havoc with data transmission in those early days.<sup>7</sup> The immediate reaction of the engineers was to increase the strength of their signal, in effect shouting to drown out the noise; but they still found that from time to time a blast of noise would wipe out a piece of the data at random.

The curious thing about this noise was that although it was random and unpredictable, it seemed to come in bursts. There could be quite long intervals when transmissions went through flawlessly, followed in an unpredictable way by intervals when there were repeated bursts of noise. When Mandelbrot looked at the problem, he found that the pattern was self-similar. In a quiet interval, there was no noise at all. But within any noisy interval there were always shorter intervals with no

noise at all, and shorter intervals with bursts of noise. Within those shorter noisy intervals, the whole pattern repeated – as far as he could tell, indefinitely. And Mandelbrot found that on each scale the ratio of the length of the quiet interval to the length of the noisy interval was the same. In fact, the distribution of bursts of noise in the transmission system was the same as the distribution of the points in the Cantor set. There were practical implications for the IBM engineers. The discovery meant that there was no point in wasting money boosting the signals even more since the systems would always be noisy, so efforts would be better concentrated on developing techniques to spot errors and repeat the distorted parts of any message; it also meant that people could be taken off the fruitless task of trying to find a physical cause (like a tree branch scraping on the wires) for what were essentially random errors, and put on more productive work (including tracing sources of noise that did not fit the pattern, and therefore probably did have a physical cause). Mandelbrot had found one of the first recognized examples of chaos at work in a piece of human technology, and had linked it from the outset with fractals, although neither chaos nor fractals had yet been named.

While we were working on this book, a local news story caught our eye, and suggested that something similar to the process that first drew Mandelbrot's attention to chaos might be going on in a nearby village. The villagers were complaining that they had been affected by repeated power failures over the past year. The electrical supply company had a perfectly rational explanation for each power cut – one was caused by a swan flying into overhead power lines, another by high winds blowing a tree across the lines, a third by lightning, and so on. But the villagers

remained convinced that there must be a fault in the system. Chaos theory can now tell us that such a cluster of local disasters is bound to happen somewhere in the power supply network from time to time – but it cannot tell us where, and it cannot tell us when, so this is small comfort to the people affected.

The links between random processes, chaos and fractals can also be made clear by looking at another of the mathematical monsters, this time dating from the early twentieth century, a construction known as the Sierpinski gasket, revealed to the world of mathematics by the Polish mathematician Waclaw Sierpinski (1882–1969) in 1916. The instructions for making a Sierpinski gasket involve a simple, repetitive iteration. Take a blacked-out equilateral triangle drawn on a piece of paper (or, these days, represented on a computer screen; strictly speaking, it doesn't have to be an *equilateral* triangle, but this helps in visualizing what is going on). Now take the mid-points of each side of the triangle and join them up to make an upside-down triangle inside the first triangle. Rub out this inner triangle to leave a white upside down triangle surrounded by three right-side-up black triangles filling the original triangle. You can guess the next step – repeat this process for each of the three smaller black triangles, then for each of the nine even smaller black triangles produced, and keep going (in principle, for ever). What you are left with is the Sierpinski gasket, an entity which is clearly self-similar and which is also a fractal, with dimension somewhere between 1 and 2 (we will describe how to measure fractal dimensions shortly).

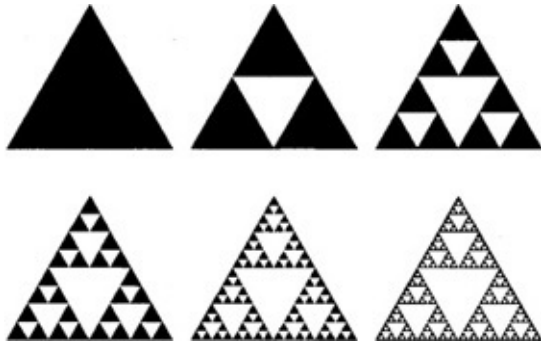


Figure 3.5 *The Sierpinski gasket.*

But there is another, equally easy way to make the Sierpinski gasket, if you have the patience. All you have to do is arm yourself with a pencil, mark three dots on a piece of paper to represent the corners of an equilateral triangle, and take an ordinary, honest die which you use to choose the numbers 1, 2 and 3 at random. Since there are six faces to the die, you can do this by saying that 4 is the same as 1, 5 is the same as 2, and 6 is the same as 3. Label the vertices of your triangle 1, 2 and 3. Now mark any point on your piece of paper as a starting point, and roll the die. If it comes up 1 or 4, mark a point on the paper halfway between the starting point and the point marked 1. If the die comes up 2 or 5, make the new mark halfway between the starting point and the point marked 2, and if the die comes up 3 or 6, make the new mark halfway between the starting point and the point marked 3. With this newly marked point as the starting point, roll the die and repeat – and repeat, and repeat, and repeat. Very slowly, after the first few points, the pattern of points that builds up on the piece of paper reveals the shape of the Sierpinski gasket. A completely random, iterative process obeying a very simple rule has built up a fractal pattern. The Sierpinski gasket is, indeed, the attractor for this particular process, and the first steps in the

iteration are not precisely on the gasket because they start at a place which is not on the gasket, and then move towards it as they are attracted by it. They are attracted by a fractal, which is what makes this a 'strange attractor', in the language introduced by Ruelle and Takens, fractals being distinctly strange compared with objects we encounter in everyday life, like triangles and elastic bands. Apart from these few initial points, though, wherever you start the game from you will end up with the Sierpinski gasket. But if you are tempted to play this game (a simple version of what is sometimes called the chaos game, where many interesting patterns can be produced from similarly simple iterative rules), two words of warning. The first is patience – it will take several hundred steps before you really see anything like the Sierpinski gasket. The second is randomness; don't be tempted to use a computer, unless you are a good enough programmer to be sure that you are using genuinely random numbers. Unfortunately, the random number generators inside computers are not always really random, although it is always possible to get round this if you have the necessary skill.

Variations on the chaos game, still based on the idea of marking points on a piece of paper (or on a computer screen, if you are sure your numbers are random) in accordance with a simple rule of this kind, repeated in a random fashion, can produce fractal images startlingly like the images of living things, including ferns and trees. For those who are interested, the mathematics involved can be found in *Chaos and Fractals*, by Heinz-Otto Peitgen, Hartmut Jürgens and Dietmar Saupe. Nobody is saying that this is precisely the way in which living things grow into complex forms. But the relevant point for us is that such seemingly complicated systems can be produced, or described, by the repeated

application of a very simple rule. The information stored in the DNA in every cell of a living entity is often referred to as containing a 'blueprint' describing how that individual being is constructed. But this is actually a rather poor analogy. A real blueprint would be very complicated – a drawing representing in detail everything in the living system and how it all joined together. A better analogy is with the recipe for baking a cake, which doesn't tell you what the final cake will look like (let alone the precise location of every raisin in the finished cake) but says 'take the following ingredients, mix well and bake for this number of minutes at that temperature'. Such a recipe is like one step in the chaos game. It is hard to see how even the wealth of DNA contained in the single cell that develops to become a human being, or a pine tree, or whatever, can contain a literal blueprint for all the complex structures involved in the final adult form. But it is much easier to see how that DNA might contain a few simple instructions along the lines of 'double in size for  $n$  steps, then divide in two, and repeat in each branch'. In the chaos game, instructions only slightly more complicated than this make complex structures like ferns by iteration. If there is an attractor for the fern shape, based on that kind of simple rule, it is no surprise that some plants grow into ferns, even though they need not be programmed to have that shape any more than a single die is programmed to make the Sierpinski gasket. It is randomness plus a simple iterative rule (or rules) that makes the complexity of the world.

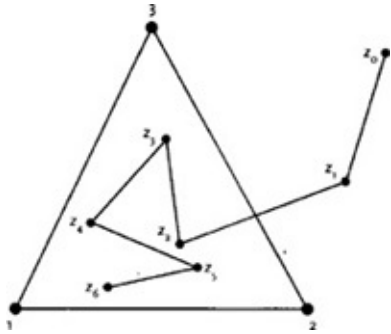


Figure 3.6a Constructing the Sierpinski gasket using the chaos game (see p. 87). The first six steps in the game.

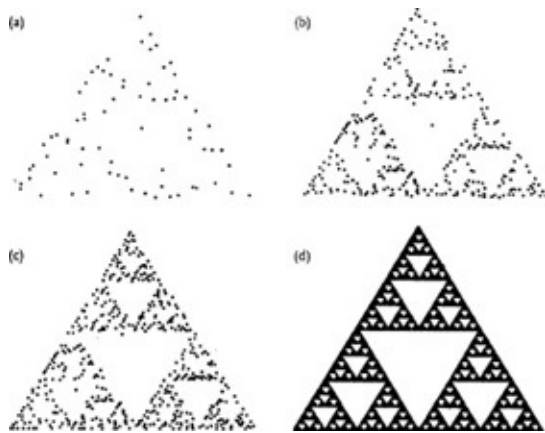


Figure 3.6b The state of play after 100, 500, 1,000 and 10,000 steps.



Figure 3.7 A variation on the chaos game produces a fern-like image.

But before we move on to look more closely at the nature of our complex world, there are still a few things worth saying about fractals



and attractors – not least, since you are bound to be curious, how you measure a fractal dimension. The simplest way to get a handle on this is through the last of the mathematical monsters from the late nineteenth century and early twentieth century that we are going to discuss in detail, the Koch curve. Apart from its intrinsic interest, the Koch curve is also important to the story because it helped to fan Mandelbrot's interest in fractals in the 1960s, and because metaphorically speaking it touches tangentially the life of a man we have already met, Lewis Fry Richardson. One of the curious things about the Koch curve, though, is that, although it is indeed a curve, the only way it can touch anything at a tangent is metaphorically, since it does not have any tangents, and in everyday language it is entirely made out of corners.

The curve was discovered, or invented, by the Swedish mathematician Helge von Koch (1870–1924), who presented it to the world in a paper published in 1904. It is easy to picture the way in which the Koch curve is constructed. Start with a straight line, and divide it into three equal parts. On the middle third of the line, construct an equilateral triangle (this time, it really must be equilateral), and then rub out the base of the triangle. You now have four equal segments of line, a flat piece on either side and an inverted V in the middle, which make up the basic starting shape for generating the Koch curve (such a starting shape is called, logically enough, the generator). Now repeat this process with each of the four straight line segments – divide into three, construct an equilateral triangle in the middle, rub out the base of the triangle. And keep on repeating this process in an endless iteration to generate the Koch curve, which is entirely made up of infinitesimally small V-shaped corners. It is also infinitely long, although the end points

of the curve are the end points of the original straight line, and in between it extends only a little way above the original line.

In a slight variation on the theme, at the generator stage you can take three of the generators (each the same size, of course) and place them along the sides of an equilateral triangle to make a six-pointed Star of David. After a few iterations, the crinkly outline of the curve resembles that of a snowflake, so this construction is sometimes called the Koch snowflake. But as the iterations proceed, the wrinkly curve also begins to resemble the structure of a rugged coastline, giving the final limiting case of this particular fractal another name, the Koch island. The Koch island has an infinitely long 'coastline' even though it is entirely contained within a circle touching the tips of the original six-pointed Star of David generator.

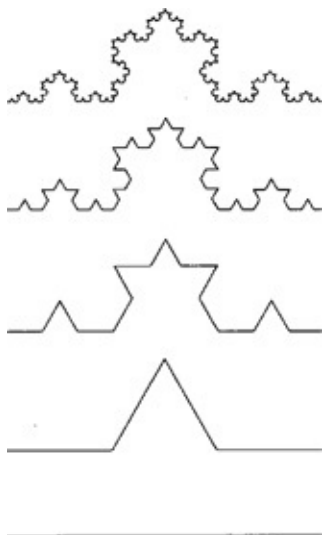


Figure 3.8 *The Koch curve (see p. 92). A ragged coastline such as that of Britain or Norway has approximately the same fractal structure as this curve.*

So are real coastlines fractal? Is the resemblance of the Koch island to a real island more than just an analogy? It was this sort of question that

really fired up Mandelbrot's investigation of fractals, and according to his own account<sup>8</sup> the trigger was pulled when he read a little-known paper by Richardson in which the sometime meteorologist puzzled over the way the lengths of the borders between Spain and Portugal and between Belgium and the Netherlands were calculated. Richardson had noticed a difference of 20 per cent in the lengths quoted by different reference sources for the same borders. The difference arises because the length you measure for such a wiggly line depends on the scale at which you are measuring it. Although there might be straight stretches along these borders, they mostly follow natural features of the landscape such as rivers and mountain ranges. Such a border might be surveyed using theodolites and all the usual surveying paraphernalia, with measuring points spaced a hundred metres apart, measuring a series of straight lines 100 metres long zigzagging across the countryside. If you were to pace out the border step by step, you would follow some of the wiggles between these points, and get another, longer value for the border's length; if you pushed a little wheel with a counter to record the number of rotations along the border, you would get a still bigger length because you would be delving into little wiggles that you stepped over in your pacing; if you took a very long piece of string and laid it carefully along the border you would get an even higher value, and so on. The point is that the borders may in principle be irregular on all scales (at least, down to the level of atoms), so the smaller the scale you look at the more wiggles you will see and the longer the border will seem.

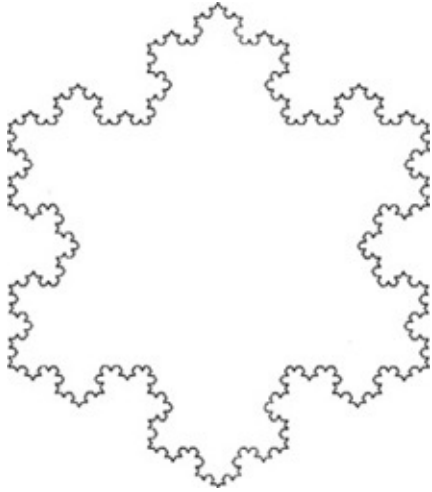


Figure 3.9 *The Koch snowflake (see p. 92).*

Mandelbrot expanded on these ideas in a paper he presented under the title ‘How long is the coast of Britain?’ which appeared in the journal *Science* in 1967. The coastline of an island has the advantage for this discussion of being essentially natural, without the long straight lines that might occur in man-made borders, like the borders between many states of the USA. The answer to Mandelbrot’s question – an answer spelled out in the paper – is that it depends what scale you measure the coastline on, and that if you go to fine enough scales the length you measure approaches infinity (only ‘approaches’ infinity, because things break down when you get to the scale of atoms). Although it is not really what Mandelbrot would soon dub a fractal, it turns out that the coastline of Britain resembles a fractal, with a dimension intermediate between 1 and 2.

So how do you measure fractal dimension? The Koch curve provides a clue. The Koch curve is self-similar. If you take any of the four segments of the Koch curve and increase it in size by a factor of three (a ‘scaling’ of 3), you get a curve that is identical to the original Koch

curve. But being self-similar doesn't necessarily make something a fractal. An ordinary cube is self-similar. If you cut out a small cube from inside the cube and scale it by the appropriate factor, it will look just like the original cube. But this will be true for *any* small cube. Self-similarity operates on any scale for many everyday objects like cubes, but only for certain specific scales for fractals; for the Koch curve, you *have* to take exactly one quarter of the original (and not any old quarter, at that) and scale it by 3 to reproduce the original. To make the distinction clear, we'll choose to scale some everyday things by the same factor, 3. If you divide a straight line into thirds, and scale one of those thirds by a factor 3, you get a line identical to the first line. The scaling is the same as the length reduction, so the dimension of a line is 1 – we divide by 3 and scale by 3. If you take a square, and divide each of its sides into thirds, you make a grid of nine small squares. So you have to take one-ninth of the original square and then scale it by a factor of 3 to get a square identical to the first one. We divide by 9 and scale by 3. The number 9 is  $3^2$  so the dimension of a square is 2. For a cube, if you divide each side up into thirds you get a set of 27 little cubes each of which will look just like the original cube if scaled by a factor of 3. Since 27 is  $3^3$  this means that a cube has dimension 3. We divide by 27 and scale by 3.

Now look again at the Koch curve. This time, we divide the length by 4 and then scale by a factor of 3. We know that  $3^1$  is 3, and  $3^2$  is 9, but what power of 3 corresponds to 4? If  $3^n = 4$ , what is  $n$ ? The answer (to four decimal places) is 1.2619, as you can easily check on a calculator. So the dimension of the Koch curve is 1.2619, between 1 and 2 as we expected, but more like a line than a plane. The same sort of reasoning

can be applied to calculate the dimension of other fractals, and even of things that are approximately fractal – the coastline of Britain, for example, turns out to have a dimension of roughly 1.3, making it slightly less like a line than the Koch curve is.<sup>9</sup>

Fractals are fascinating in their own right, and Mandelbrot's chief claim to fame is his discovery of a set that bears his name, which has a fractal structure revealed by iterating a simple mathematical expression. The difference, compared with the sets and fractals that we have been discussing, is that the expression involves what mathematicians call complex numbers, which are numbers involving the square root of  $-1$ . The key thing about complex numbers is that they are in a sense two-dimensional, whereas everyday numbers are one-dimensional. A single everyday number is all you need, for example, to specify the position of a point along a line. A single complex number is the minimum you need to specify the position of a point on what is called the complex plane (by its distance from each of two of the edges of the plane). It contains two pieces of information. The iteration that produces the Mandelbrot set involves the expression  $(Z^2 + C)$ , where  $Z$  is a variable complex number and  $C$  is a fixed complex number. In the usual way, we iterate by choosing a value of  $Z$ , feeding it into the expression, calculating the result and using that as the value of  $Z$  for the next iteration. The fractal pattern emerges when the values of  $Z$  are plotted on the complex plane. You don't have to know anything about complex numbers to appreciate that this is a very simple (but non-linear) process, similar to the iteration of the logistic equation. Yet the fractal that results is possibly the most complicated entity ever investigated by humankind, and it is not just complicated but beautiful, so that blown-up portions of the set have

achieved an iconic status as poster illustrations, and whole books have been filled with them. But such investigations take us in the opposite direction from the one we wish to follow in the rest of this book – inward into the world of mathematics rather than outward into the world of living things. The only thing we want to emphasize about the Mandelbrot set is that, once again, this extremely complicated object is indeed produced from a very simple expression by iteration. From this perspective, it is one of the simplest things ever investigated by humankind – but then, as we shall see, from this perspective so is everything else.

The key question we are addressing here is how complexity arises from simplicity. This, and the link between chaos and fractals, is clearly seen if we go back to our old friend the logistic equation, and interpret its effects in terms of topology, which introduces the additional layer of complexity that we hinted at earlier. What the logistic equation (and the kind of similar iterative processes we have been discussing) does is to transform one set of numbers into another set of numbers. If the original numbers are points on a plane, or on the surface of a sphere, or on some other surface, they may be changed, one by one, into points somewhere else in the plane, or on another plane, or some other surface. We are all familiar with this kind of process, although it is so familiar that we seldom think about what is going on. It is called mapping. The kind of map we use to find our way around a town, or to study geography, doesn't reproduce every point in the real world, but essentially a street plan is a faithful copy in miniature of the streets of the town, just shrunk down in size to fit on a piece of paper. A globe can in principle be used as a faithful map of the (roughly) spherical Earth, with no distortions.

But a map can also be a faithful copy *and* have distortions. The map of the London Underground is a good example – there are points representing every station, and lines representing all the train lines, and the map preserves the essence of the relationship between those points and lines, but it has been distorted to make it easier to interpret.

To some extent, the distortions on the Underground map are a matter of choice – you can also draw maps of the London Underground that are much more faithful copies of the original, in terms of representing every wiggle in the lines. But a map of the world drawn on a flat piece of paper *must* be distorted, because you are trying to represent the surface of a sphere on a plane. This is why the shapes of the continents on the familiar Mercator projection and the popular Peters projection (two different ways of tackling this problem) are not exactly the same, and why neither is exactly the same as the shapes of the continents on the face of the Earth itself. But in each case every single point on the surface of the Earth can be represented by a single point on the surface of the map. Indeed, the Mercator projection is also a map of the Peters projection, and vice versa; neither is intrinsically ‘better’ than the other.

The process described by the logistic equation is also a mapping. We can see this by looking at how the equation affects a set of numbers describing a straight line. Remember that the equation changes a value of  $x$  into a new value using the transformation

$$x(\text{next}) = Bx(1 - x)$$

and that thanks to renormalization we only have to consider starting values of  $x$  between 0 and 1. These starting values can represent a



straight line – if you like, think of a metre stick, one unit (1 metre) long. To keep things simple, we'll choose a simple integer value of  $B$  – say, 3 – and look at how the line is transformed by the logistic equation – how it is mapped on to a different arrangement of points.

We can see what happens by looking at the way the mapping works at points separated by one-tenth of the length of the original line. The point  $x = 0$ , obviously, maps on to the point 0. For a value of  $x$  of 0.1, a tenth of the way in from one end of the line, we get a value for  $x(\text{next})$  of  $3 \times 0.1 \times 0.9$ , or 0.27. The point  $x = 0.1$  has been mapped on to the point  $x = 0.27$ . Similarly, the point  $x = 0.2$  maps on to the point 0.48, 0.3 maps on to 0.63, 0.4 maps on to 0.72 and 0.5 maps on to 0.75. This half of the line, from 0 to 0.5, has been stretched and now covers a straight line from 0 to 0.75, demonstrating another curious property of infinity: there are an infinite number of points in each line, and a 1:1 correlation between the points in the two lines, but one is half as long again as the other.

Between 0.5 and 1, the same thing happens in reverse, as you can quickly confirm. The point 0.6 maps on to 0.72, 0.7 maps on to 0.63, 0.8 maps on to 0.48 and 0.9 maps on to 0.27. And since  $(1 - 1)$  is 0, the point  $x = 1$  maps on to the point 0. This half of the line has not only been stretched, but turned over and laid back on top of the mapping of the first half of the line. There has been a topological transformation of the line, a stretching and folding, so although the line is now 50 per cent longer, thanks to the folding it fits in to 75 per cent of the length of the original line. The link between algebra and topology is as profound an insight as the realization that an algebraic equation such as  $y = x^2$  can be represented graphically by a line on a piece of paper or a trajectory

through space – in this case, a parabola. But it is in phase space, remember, not real space that the topological transformations we are interested in are going on.

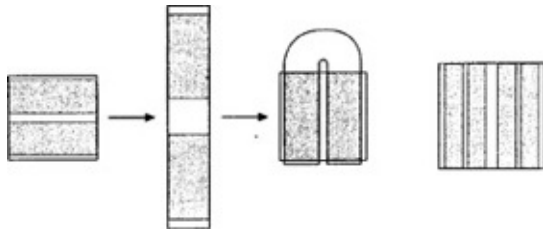


Figure 3.10 *Smale's horseshoe, produced by the Baker transformation (see below). Repeating this simple process of folding and stretching indefinitely builds up a many-layered structure, with the layers spaced like the points in the Cantor set.*

If you think in physical terms, instead of abstract mathematical lines, the stretching and folding process described by the logistic equation produces something like a horseshoe shape from an original cylinder, and it is sometimes known as the Horseshoe transformation, even though the 'bend' at the end is a 180-degree hairpin, not the gentle, rounded curve of a real horseshoe. But why stop at one Horseshoe transformation? The way to make complex structures out of simple rules, we know, is repetition. So iterate the whole process – transform the 'horseshoe' in the same way, then transform the transformed horseshoe, and so on, indefinitely. This process is a bit like the way a baker kneads dough, stretching it, folding it over on itself, pummeling it flat, stretching and folding again, repeatedly. So it is also known as the Baker transformation. Whichever name you prefer, happily for the study of chaos, by the time people like Robert May began their investigations of the period-doubling route to chaos in the 1970s the topological properties of the Horseshoe transformation had already been

investigated, simply because of its intrinsic topological interest, by the brilliant theorist Stephen Smale, who, as we have seen, became acquainted with chaos through his contact with Robert Yorke. We don't need to go into the mathematics – he has done it for us. But the simple physical picture of what happens in phase space when you do carry out the Horseshoe transformation indefinitely is backed up by the full weight of the maths, and does indeed give a true insight into what is going on.

Think of the original line as an attractor in phase space, and try to visualize what happens to it with the repeated application of the Horseshoe transformation. At each stage, the length of the line is stretched, but because it is folded back on itself it takes up less space 'horizontally'. After the first step, you have two lines on top of each other (but taking up no vertical distance, since lines have no width), with one bend; the next step has four layers with three bends; the next eight layers and seven bends, and so on. The number of layers doubles each time, and the number of bends is found by doubling the previous number and adding 1. You end up with an infinitely wiggly curve (but with the wiggles arranged in a very specific way, bends within bends within bends forming a self-similar set), made up of infinitely many layers, but taking up no distance either horizontally or vertically. Taking a cross-section up the stack of this infinitely layered structure, you would find that the points are distributed like the points in the Cantor set. But where do the points next to each other in these layers come from? Because of the repeated stretching and folding process, two points on the final curve that started out close together on the original line may end up far apart in this set, while two points that started out nearly at opposite ends of the original line may end up very close together. If the

state of a system moves along the original line smoothly in one direction, passing through each point in turn, it will seem to be jumping about among the points of the Cantor set at random. This is the topology associated with the onset of chaos produced by period doubling.

This infinite layering, like an infinite flaky pastry, also occurs in the Lorenz attractor. As shown on the page (see [page 58](#)), the line representing the Lorenz attractor in phase space seems to cross itself many times – indeed, an infinite number of times. But what is really happening in phase space is that at every crossover the attractor is ‘travelling’ in a different layer of phase space, a different plane. One way to help visualize this is to think of a book with an infinite number of infinitesimally thin pages, opened flat in the middle. One lobe of the Lorenz attractor is drawn on the left-hand page, and one lobe on the right-hand page – there is a loop, or series of loops, around the relevant lobe on every page of the book. But every time the trajectory of the attractor in phase space moves across the middle of the book and on to the opposite side, it moves on to a different page. There are an infinite number of crossing places, but the line representing the attractor never cuts itself.

In both cases, the Horseshoe attractor and the Lorenz attractor, an infinite number of layers of phase space are contained within a finite volume of phase space. Both attractors are fractal – they are strange attractors. And this is just the beginning. If you do the same kind of stretching and folding to attractors that don’t start out as straight lines in phase space – such as an attractor wrapped around a torus in phase space – you end up with an even more complicated pattern of fractal chaos, but one still based on a very simple set of rules. At last we have

all the information we need about how the underlying simplicity of the world can produce complex structures, and we can begin to move upwards through the layers of complexity to apply this understanding to the emergence of life. There are still several steps along the way; but to whet your appetite, here is a hint of how all this may be relevant not only to life in general, but to human life in particular.

Remember the way we calculated the fractal dimension of the Koch curve in terms of the power, or exponent, that appears in the scaling law. Such a relationship is called a power law – 3 to the power of 2 is  $3^2$ , 3 to the power of 3 is  $3^3$ , and so on – and we had to raise 3 to the power of 1.2619 in order to get 4, the number of pieces making up the basic unit of the curve. Something as simple as the volume of an object also obeys a power law. If you have a cube with sides of length  $l$ , the volume is proportional to  $l^3$ , whatever the value of  $l$ , if you have a sphere of radius  $r$ , the volume is proportional to  $r^3$ , whatever the value of  $r$ . Volume, or size, obeys a power law with exponent 3. In the mid-1980s, researchers investigating the metabolic rates of animals of different sizes were intrigued to discover that although these also obey a power law, it is not the simple exponent 3 law that you might expect from the way the sizes of the animals scale. They compared the metabolic rate with the body mass for mammals as diverse as rats, dogs, humans and horses. The mass of the body of an animal is proportional to its volume, and as you would expect, the more mass there is the higher the metabolic rate, because there is more body burning up its food and releasing energy. But although the mass increases as a power law with exponent 3, the metabolic rate scales in line with a power law with exponent 2.25. In this respect, the animals are behaving as if their size does not scale like a

three-dimensional volume, but like something intermediate between a volume and a two-dimensional surface – specifically, like an extremely crumpled up, fractal surface. A mathematician (at least a topologist) would immediately interpret this power law as suggesting that the objects involved were fractal surfaces crumpled up within finite volumes.

Looking at bodies in more detail we find that there are very many features of living systems that are like fractals. The way that arteries and veins branch, for example, turns out to be essentially fractal, enabling blood to reach and return from every part of the body without the veins and arteries taking up so much room that there is no space for anything else. This is particularly clear in the kidneys, where veins and arteries branch in an extremely complex intertwining fashion to exchange fluids. The kidney is very much a finite, three-dimensional object; yet the veins and arteries within the kidney tend towards the infinite length of a true fractal.

Of course, the analogy breaks down at the extremes. Systems within the human kidney do not actually branch infinitely many times, just very many times; and going in the other direction, we do not find kidneys embedded within superkidneys, and so on for ever, but that each system is self-contained. Nevertheless, the similarity between many living systems and fractals is more than merely an analogy, and it explains also, for example, how the two-dimensional surface of the lungs can have a large enough area to allow carbon dioxide and oxygen to be exchanged across the surface with the rapidity required to keep you alive, even though they are contained within quite a small volume. Near-fractal self-similarity is a pervasive feature of the bodies of living

organisms. And as well as doing the job effectively, remember that the amount of DNA required to code for the production of such a system is very simple indeed, compared with the amount of information contained in a literal blueprint for a structure as complex as the kidney. It is the simplicity of the rules underlying the structure of fractals that allows living things to be complicated enough for creatures able to ask questions about the nature of the world to evolve.

What we have seen is that you can have very simple systems, in which nothing interesting happens, like a tap dripping at a steady rate with period one. And you can have chaotic systems, in which there are essentially random fluctuations, so there is no order and structure is destroyed. But in between the two you have, starting from the boring end of the spectrum, an increasing level of complexity, starting from the dripping tap with period one and getting more and more interesting until suddenly chaos sets in. So the most complex and interesting things in the Universe are happening right at the end of chaos, just before order is destroyed. It is here that we find taps (or leaky gutters) dripping with strange and exotic rhythms, whirlpools within whirlpools spinning about in wonderful patterns, and the extraordinary complexity of the kidney, or the multiply folded and refolded near-fractal surface of the cerebral cortex of the human brain. So far we have looked at order, and we have looked at chaos. It is time now to look at the edge of chaos, where complexity lives.

## 4

### The Edge of Chaos

In a sense, classical thermodynamics pretends that time does not exist. Systems are described in terms of infinitesimally small changes that would take an infinite time to shift them from one state to another. Classical thermodynamics also assumes that there is no flow of energy – specifically, heat – through a system. But as we have already seen, in the simple example of the separation of hydrogen gas from hydrogen sulphide gas by thermodiffusion, interesting things immediately begin to happen when energy is flowing through a system. No significant amount of energy flows through a corpse, but a great deal of energy flows through your body – chemical energy from the metabolism of the food you eat, powering your muscles and the rest of your bodily activity, and ultimately being dissipated as heat. Dissipation of energy is an essential feature of non-equilibrium thermodynamics, giving us another way of referring to the flow of energy through a system, as a dissipative process. These are also referred to as open systems, because (unlike the hypothetical closed box of gas beloved by the nineteenth-century thermodynamic pioneers) they are not sealed off from the world at large. It is in closed systems that we encounter time reversibility and Poincaré recurrences; in open systems we encounter irreversibility and an arrow of time.



As we saw in [Chapter 2](#), classical thermodynamics is based on a contradiction. The very term ‘dynamics’ tells us that it describes how systems change; yet ideas such as entropy were derived from calculations involving systems in equilibrium, where nothing changes. Equilibrium itself is of no intrinsic interest, because nothing happens there. But how things move towards equilibrium can be very interesting. The nearest a living thing ever gets to equilibrium is when it dies. The fact that something is dead is nowhere near as interesting as how it died – a premise on which the success of a whole genre of detective novels, the murder mysteries, is based.

When people began to study dissipative processes, it was natural that they should start by investigating systems close to equilibrium, and processes like thermodiffusion. If systems are close to equilibrium, they generally respond in a linear fashion to changes in their environment – change the temperature gradient in the thermodiffusion process by a small amount, for example, and the system will respond in a linear fashion, also changing by a small amount. The groundwork for the study of linear thermodynamics was laid by the Norwegian chemist Lars Onsager (1903-76), working at Brown University in the early 1930s (he moved on to Yale in 1933). His key contribution, singled out for special mention in the citation for his Nobel Prize (awarded in 1968), was the discovery of what are known as the reciprocal relations, which say that there is a symmetry in such systems. In the case of thermodiffusion, for example, just as a temperature gradient produces a gradient in the concentration of the mixture of gases, so, according to Onsager’s theory, setting up and maintaining a concentration gradient must produce a temperature gradient, causing a flow of heat. This was later confirmed

by experiments, and the reciprocal relations also became known as the fourth law of thermodynamics.

Onsager had given chemists a set of tools which they could use to begin the investigation of irreversible phenomena in dissipative systems, at least in systems where linear rules applied (the 'linear regime'). The person who made the most effective use of this tool kit was Ilya Prigogine, born in Moscow in 1917, who emigrated with his family to Belgium when he was twelve, and by 1945 was working at the University of Brussels. It was Prigogine who showed, as we mentioned in [Chapter 1](#), that a dissipative system in the linear regime settles down not into the state of death corresponding to maximum entropy (as it would be in equilibrium), but into a state where entropy is being produced as slowly as possible, and the dissipative activity is ticking over at a minimum rate. This is the situation in thermodiffusion when the gradient of the concentration of gases has stabilized for a particular temperature gradient, but Prigogine showed that it applies in general for systems described by linear thermodynamics. Things exist in a steady state in the linear regime; a human being, for example, can maintain his or her integrity for many years using the flow of energy (and food) through their body, although in this case the steady state eventually breaks down, for reasons which are still not understood. But the steady state of a human adult is strikingly different from the dramatic changes that occur when a new human being develops from the single cell of a fertilized egg, which is clearly a non-linear process. More of this shortly.

Over the following two decades, the work of Prigogine and his colleagues in Brussels (who became known as the 'Brussels School') focused on the attempt to provide a mathematical description of

thermodynamics for systems far from equilibrium, in the non-linear regime where small changes in the external environment can produce large changes in the system. This should perhaps best be regarded, even today, as work in progress; it is far from being a complete theory. Happily, we won't need to go into the details, or to get embroiled in the controversies surrounding this work, since the important and widely accepted aspects can be understood in terms of the behaviour of simple physical and chemical systems. But there is one key point to take on board – such systems are described by non-linear equations in situations where feedback is important.

Along the way, Prigogine has also developed his own ideas about the nature of time, and the relationship between thermodynamics and the arrow of time; it is for these ideas that he is probably best known outside professional scientific circles, and that is the only reason that we mention them. The puzzle he sets out to solve in this work is how order – things like us – can have arisen out of chaos – *not* deterministic chaos, but essentially what the Ancient Greeks meant by chaos, the smooth distribution of gases that existed when the Universe was young. But we shall not go into this aspect of his work, because it seems to most physicists to be a blind alley, since there is a natural way to explain why there is an arrow of time in the Universe, and in particular, as we shall now explain, how ordered systems such as ourselves can have arisen from the almost uniform state of disorder that emerged from the Big Bang.

The best place to begin to develop an understanding of the interesting things that can happen in open systems far from equilibrium, in the non-linear regime, is with a phenomenon first described by the

Frenchman Henri Bénard in 1900. Because it was later investigated by the English physicist Lord Rayleigh, it is variously known as the Bénard instability or as the Rayleigh-Bénard instability. The most effective way to see what is going on is to heat a flat pan containing a shallow layer of silicone oil carefully from underneath, under controlled conditions in the lab. The oil layer has to be only a millimetre or so deep, and the heating has to be uniform across the flat bottom of the container. It can be done using a saucepan on an ordinary kitchen stove, although we would not recommend anyone to try heating thin layers of oil in this way; something similar can sometimes be seen going on in a pot of thick soup as it is slowly brought to the boil. But the whole point of the experiment we are about to describe is that we don't want to heat the oil so much that it boils in a chaotic and messy fashion. We are interested in what happens just before the onset of chaos. Since the oil itself is clear, to make what is going on more easy to see it is usually sprinkled with aluminium dust (not something we recommend with the soup).

What goes on, of course, is convection. The hot fluid at the bottom of the layer expands, becomes less dense, and tries to rise up through the colder, denser fluid above. If the heating isn't uniform, this poses no puzzle. The fluid simply rises where it is hottest, subsides somewhere else, and flows back to the hot spot to make a circulating convection cell. But in the uniformly heated fluid, there is no special place where this can begin to happen. There is a uniform symmetry across the bottom of the fluid, and it all 'wants' to rise at once. At first, nothing at all happens. Then, as the temperature is gradually increased underneath while kept cool on the top (so that the temperature gradient gets steeper), at a critical point the symmetry is broken and the uniform

surface of the fluid is suddenly broken up into a pattern of tiny hexagonal convection cells. The point at which this happens depends on the temperature difference between the top and bottom of the fluid – the temperature gradient – and the cells that form in this way are about as big across as the oil is deep. In each cell, warm fluid rises in the centre, moves to the edges, and descends at the boundary with the neighbouring cell. Because the surface of the fluid is open to the air, surface tension effects are important, and these encourage outward movement of the fluid at the surface, pulling it away from the centre of each cell and producing a dimpled effect with the middle of the cell lower than the edges, even though fluid is rising here. The situation studied by Rayleigh was slightly different, with a thin layer of fluid trapped between two plates, a warm one underneath and a cold one on top. This makes both the physics and the mathematics much simpler, since surface tension doesn't have to be taken into account, and the pattern of behaviour that is revealed neatly links the Rayleigh-Bénard instability with the onset of deterministic chaos.

In this simplified version of Bénard convection, several steps on the road to chaos can be seen. As convection begins, it does so in the form of long sausages, so that the surface of the fluid takes on a stripy appearance. These sausages have a roughly square cross-section, and in neighbouring rolls the convection goes respectively clockwise or anticlockwise, as it must do or the convection would not work smoothly. But there is no way to tell in advance whether any particular point in the fluid will become part of a clockwise convective roll or an anticlockwise convective roll. The exact arrangement is determined at random. There is a bifurcation, in which the system has gone from a

single state to a pair of states. As the temperature gradient increases,<sup>1</sup> at another critical point the pattern abruptly changes into what is called a bimodal form, in which there are two sets of convective rolls operating simultaneously at right angles to each other. This produces a pattern of roughly square convective cells as seen from above, like a check shirt. The number of possible states for the components of the fluid has doubled again, from 2 to 4. This is another example of bifurcation at work, a variation on the theme of the period-doubling route to chaos (although none of that was known to Rayleigh, of course), and if the temperature gradient (or Rayleigh Number) is increased far enough, you end up with a turbulent form of convection in which short-lived patterns appear, break up and re-form endlessly. But the most interesting stable patterns appear right at the edge of chaos, and for a thin layer of oil heated in an open pan, thanks to surface tension the specific interesting pattern that appears is a honeycomb arrangement of hexagons. This is happening far from equilibrium, thanks to energy flowing through an open system and being dissipated. This is the secret of the existence of order in the Universe, and specifically the secret of life. Whenever you are confronted by the existence of order, and especially by the existence of life, and you are wondering how the world got to be like that, you can always say to yourself 'Remember the hexagonal cells of the Bénard convection? It's the same thing.'

A system can only be held in an interesting state away from equilibrium if it is dissipative and open to the environment, so there has to be an outside source of energy. On Earth, that energy comes, ultimately, from the Sun. And it is understanding why the Sun shines that makes it clear why order has arisen out of chaos and how the

Universe came to have an arrow of time.

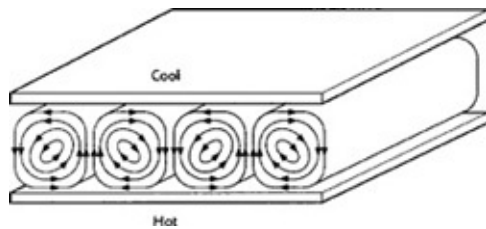


Figure 4.1 *In the version of Bénard convection studied by Rayleigh, rolls of convection with a roughly square cross-section (like a squashed Swiss roll) develop.*

It's all thanks to gravity. Gravity has a very curious property – the gravitational energy of any object that has mass is negative. It is literally less than zero, and the more compact the object is the more negative its gravitational energy is. As a bald statement of fact, this may be hard to accept. But it is easy to understand what is going on by looking at gravity in two stages. First, the gravitational force between two objects diminishes as they get further apart. It happens that it diminishes in line with an inverse square law, as Newton proved. The amount of gravitational energy stored in a collection of objects (be they atoms, or stars, or lumps of concrete, or paintings by Van Gogh) depends on the force between them. This isn't just true of gravity – if you stretch a piece of elastic, you need to pull harder and harder to stretch it more and more, and the energy stored in the elastic is related to how hard you are pulling. But to 'stretch' the distance between two (or more) gravitating objects you need to pull less and less as they get further apart. This is why it takes a huge booster rocket to put a satellite in orbit around the Earth, but once it is in orbit only a small rocket is needed to send it to the Moon, or Mars. If the objects were infinitely far apart, the total energy stored by the gravitational field linking them would be zero – essentially because the force is proportional to 1 divided by infinity

essentially because the force is proportional to 1 divided by infinity squared.

Now imagine a collection of particles (atoms, or house bricks, or whatever) spread out infinitely far apart and given a tiny nudge so they start to move together under the influence of gravity. As they do so, each of the particles gains kinetic energy, and moves faster and faster as it falls. This energy comes from the gravitational field, in just the same way that if you tie the ends of a piece of elastic to two pebbles, pull them apart and then let go, the pebbles will rush together as the energy stored in the elastic is converted into energy of motion (kinetic energy). Similarly, the kinetic energy of the particles falling together has come from the gravitational field; but the gravitational field started out with zero energy. So now it has less than zero – negative energy. By the time the particles come together to make a star, the gravitational field associated with that star has a lot of negative energy. How much? You may find the answer surprising, but if so you are in very good company indeed.

I have told this anecdote before,<sup>2</sup> but there is no better way to illustrate the point, and it would be foolish to assume that everyone reading this has already read my earlier books. It dates back to the 1940s, when physicists were already aware of what might be called the negativity of gravity, but when most of them were mainly occupied with war work – not least, in the US, the Manhattan Project. One physicist based in America, who wasn't allowed to work on the most secret projects because he was a Russian by birth, was George Gamow (1904–68), an ebullient character with a wide range of interests who, among other things, developed the first model of the Big Bang origin of the



Universe. His contribution to the war effort was as a consultant to the Bureau of Ordnance in the US Navy Department at Washington, DC, and one of his duties was to travel up to Princeton every fortnight with a briefcase of papers for Albert Einstein. There was nothing very secret about the papers – they described ideas for inventions that might help to win the war, sent in to the Navy by well-meaning but usually misguided citizens, and the Navy wanted Einstein’s advice on whether any of them might be worth pursuing. The papers included such ideas as the suggestion that the Atlantic Ocean might be frozen to prevent submarines operating, so Einstein’s skill at sorting wheat from chaff was not unduly tested, although as a former Patent Officer in Switzerland he was rather good at spotting the flaws in more subtly unworkable proposals.

One day, while walking through Princeton with Einstein from his home to the Institute for Advanced Study, Gamow casually mentioned an idea which had been tossed out some time before by one of his colleagues, the quantum physicist Pascual Jordan. It was just a wild idea Jordan had come up with, the sort of thing to enliven a coffee-time discussion, or a walk through Princeton. Jordan had calculated that for any material object, if the mass were concentrated at a point, then the negative energy of the associated gravitational field would be  $-mc^2$  and would therefore exactly cancel out the positive mass-energy of the star. In other words, as Gamow put it to Einstein, a star might be created out of nothing at all. ‘Einstein stopped in his tracks,’ Gamow tells us,<sup>3</sup> ‘and, since we were crossing a street, several cars had to stop to avoid running us down.’

It's worth emphasizing that this is no mere analogy, or a result of the way we choose to measure things like energy. It is a fundamental truth about the way the Universe works that gravitational fields have negative energy, and that for matter concentrated at a point this negative energy exactly cancels out the mass-energy of the matter. The idea that stopped Einstein in his tracks while crossing a road made little impact at the time, but forty years later it became a cornerstone of the idea that the entire Universe could have appeared in this way out of nothing at all, as a bubble of mass-energy with an associated gravitational field carrying an equal and opposite amount of energy, so that the total energy of the Universe is zero, and with the tiny bubble getting blown up into what became the Universe we see around us now by a process called inflation. All this is discussed in my book *In Search of the Big Bang*, but the only point that matters here is that we know that the Universe started out from the Big Bang in a very uniform state. The details of what went on in the Big Bang are still a matter for debate, but direct observations of the Universe tell us what it was like very shortly after the beginning.

The most important fact about the Universe today is that it is expanding. All the stars we see in the night sky are part of a disc-shaped system, a galaxy called the Milky Way, which contains hundreds of billions of stars. This is just one among hundreds of billions of galaxies, and galaxies are grouped together in clusters (held together by gravity) which move apart from one another as the Universe expands. The expansion is caused by space itself stretching between the clusters of galaxies, not by the clusters moving through space, and is explained in complete detail by Einstein's general theory of relativity (indeed, it was predicted by the theory). If things are getting further apart now, then

obviously they were closer together in the past, and if we go back far enough into the past, everything was on top of everything else in a single lump, the Big Bang. By measuring the present expansion rate of the Universe and using the general theory of relativity we can pin down the time that has elapsed since the Big Bang, to around fourteen billion years.

One of the great things about astronomy is that because light takes a finite time to travel through space, we see things far away as they were long ago. A galaxy ten million light years away, for example, is seen by light that left it ten million years ago. And if you set 'the beginning' as time zero and work forwards, the furthest back in time we can 'see' with our detectors is to an epoch when the Universe was about 300,000 years old, and consisted of a very nearly uniform sea of hot gas (strictly speaking, a plasma), at about the same temperature as the surface of the Sun today – roughly 6,000 °C.

The radiation from that hot fireball has cooled as it has aged and the Universe has expanded (just as hot gas in a closed box cools if the box expands), and is detected today as a faint hiss of radio noise, in the microwave part of the electromagnetic spectrum, with a temperature of just under *minus* 270 °C (or 2.7 degrees above the absolute zero of temperature). Very slight variations in the temperature of this 'cosmic microwave background radiation' coming from different parts of the sky tell us that the hot gas from which the galaxies and stars were formed was then very nearly, but not quite, uniformly distributed through space. In some places, the gas was a little more dense than in other places. The denser regions would then have attracted more material to themselves through gravity, increasing the irregularity of the Universe and

eventually leading to the situation we see today, with matter concentrated in galaxies of bright stars and with a lot of dark space in between them.

This is a crucially different insight from the one we gain by studying a box of gas here on Earth. Where the effects of gravity can be ignored (that is, the effects of the gravitational forces of the particles in the box acting on each other can be ignored), the state of maximum entropy for a box of gas is with the gas spread uniformly through the box at a uniform temperature. But when the gravitational forces between the individual particles that make up the gas cannot be ignored, as is the case for large clouds of gas and dust in space, gravity can pull things together in clumps, making more order and *at the same time* reducing the entropy. As Paul Davies has put it, 'gravitationally induced instability is a source of information'.<sup>4</sup> More information implies less entropy, so you can also regard this as implying that as the information 'comes out' of the gravitational field in a collapsing cloud of gas, the gravitational field is swallowing entropy, to go with its negative energy. It is the negative energy of the gravitational field that makes it possible for it to swallow up entropy in this way, explaining why the Universe is not in thermodynamic equilibrium today.

But let's leave the big picture of the Universe at large and focus back on the emergence of life on Earth. What matters is that gravity has been responsible for a kind of domino effect, not on a small scale but more like the giant arrays of dominoes sometimes toppled to break records or raise money for charity. In a step by step process it has created increasing amounts of organization in the Universe, right down (or up) to the level of beings intelligent enough to speculate about how it all

happened. Without going into details of how stars work,<sup>5</sup> we can see that stars only exist because clumps of gas have been drawn together by gravity and got hot enough inside (thanks to the kinetic energy gained by their constituent particles from the gravitational field) for nuclear fusion reactions to take place. This has led to a situation in which the star and its surroundings are not in thermodynamic equilibrium. Instead, we have a hot star sitting in cold space, so that energy pours out of the star in an attempt to equalize the temperature inside and out. The thermodynamic arrow of time points in the direction where the future corresponds to energy flowing out of the star; and it is no coincidence that the thermodynamic future is also the direction of time away from the Big Bang, because of the role of gravity in all this. Ultimately, it is gravity that tells the arrow of time which way to point.

A planet like the Earth is bathed in the flow of energy from a star, which makes the whole surface of the planet an open, dissipative system. All life on the surface of the Earth makes use of this energy to maintain itself far from equilibrium, on the edge of chaos.<sup>6</sup> Plants get their energy directly from sunlight through photosynthesis; grazers get their energy from plants; carnivores get their energy from other animals. But it is all originally from the Sun, and it is all, originally, thanks to gravity. But the way in which systems organize themselves, using this flow of energy, into what seem to be complex forms, is really quite simple. We can see this clearly by following the lead of the brilliant mathematician Alan Turing (1912–54), who had the chutzpah to attempt to explain, more than half a century ago, what happens in the most complicated process we know of, the development of an embryo from a single living cell. Turing was way ahead of his time, and the importance of his work in

this field was not recognized until long after his death; but with hindsight it is the logical place to begin the next stage of our story.

Turing, who was born in Paddington, London, on 23 June 1912, is best known as a cryptographer, the leading member of the team at Bletchley Park in Buckinghamshire, which cracked the German codes (including the famous Enigma code) during the Second World War. By then he was deeply interested in the possibility of creating an artificial intelligence, a ‘universal computer’ (today sometimes referred to as a Turing machine) that could solve any problem. It was through his interest in how human intelligence developed that he started thinking about embryonic development. He would certainly have achieved even more but for the fact that he was harassed by the authorities for being a practising homosexual (then still illegal in Britain) and killed himself by eating an apple he had dipped in cyanide on 7 June 1954, a couple of weeks short of his forty-second birthday. Turing had graduated from King’s College, Cambridge, in 1934, and worked there for two years before taking his PhD at Princeton, returning to King’s in 1938. In 1936 he wrote a paper, ‘On Computable Numbers’, which introduced the idea of the Turing machine. At the time this was purely an imaginary device, a ‘thought experiment’ designed to demonstrate the logical structure of a hypothetical universal computer, but the principles that Turing spelled out in that paper form the basis of all modern computers, as well as providing insight into the world of self-organizing complex systems.

Turing’s imaginary machine was supplied with a long paper tape (in principle, infinitely long), divided into squares in which there were symbols or numbers which could be read, and, if appropriate, erased and replaced with another symbol. Turing was thinking of a paper tape,

written on in pencil, which could be erased and rewritten as required. Today, the image conjured up is of a magnetic tape, like the tape in an audio cassette, or even the hard disk of a computer, or its solid-state read/write random access memory (RAM). They are all logically equivalent to one another for this purpose. Reading the number in one square would tell the machine whether to move forwards or back along the tape (through the memory), how to calculate a new number from the information on the tape, and when to rewrite the contents of a square. For example, the symbol in one square might instruct the machine to add the numbers in the next two squares and write the answer in the second of those two squares. In this way, if the instructions on the tape were specific enough, the machine could calculate anything. Starting at one end of the tape, or at some specified square, the machine would plod its way along, sometimes going backwards and forwards over the same stretch of tape, writing and rewriting symbols in the squares, changing its internal state in accordance with the instructions it read (for example changing from 'add' to 'divide'), and eventually finishing at the other end of the tape, or at some square specified by the instructions, having written a message of its own, the answer to the question posed by the original message. Depending on which symbol there is in a certain square, it always knows what to do next. Turing proved that there could exist a *universal* machine of this kind (now called a universal Turing machine) which could solve any problem that could be expressed in the appropriate symbolic language.<sup>7</sup> It would be a machine which, in Turing's own words:

... can be made to do the work of any special-purpose machine, that is to say to carry out any piece of computing, if a tape bearing suitable 'instructions' is inserted into it.

This is so familiar today to even the most naive computer user as a description of how computers work, with the programmed software in its memory playing the part of the tape and the hardware of the computer representing Turing's machine, that it is hard to grasp how fundamental a breakthrough this was in 1936. Quite apart from the practical value of these ideas, soon to be realized under the pressures of the wartime emergency, Turing's paper raised profound issues that are relevant to an understanding of life and other complex systems. In many cases, the advantage of a Turing machine is that the instructions telling it how to carry out a computation (the algorithm) are far more compact than the fruits of that calculation. For example, there is a reasonably short algorithm which specifies how to calculate  $\pi$ , without actually writing out the infinite string of digits representing  $\pi$ . For many purposes, the algorithm is  $\pi$ . More prosaically, the familiar 'multiply' technique is an algorithm, although in everyday terms writing, say,  $6 \times 9$  is hardly a shorthand version of writing 54. Turing showed that there are systems which cannot be 'compressed' algorithmically, and whose most compact representations are themselves – one of the key concepts we encountered, from another direction, in our earlier discussion of chaos. In particular, as we have seen, the shortest description of the Universe is the Universe itself.

By the time the Second World War broke out, Turing was already trying to build actual machines that would demonstrate these principles at work, even though they would fall some way short of being universal computing machines. In the course of the code-breaking work at Bletchley Park a secret British project did develop the first properly programmable digital computer. In the years immediately after the war,



Turing continued to work on the design of computers, first at the National Physical Laboratory (NPL) and then (from 1948 onwards) at the University of Manchester. It was the impetus to computer development provided by the code-breaking work, and particularly Turing's contribution, that made Lewis Fry Richardson's dream of numerical weather-forecasting become a reality just about in Richardson's own lifetime, and which then led Edward Lorenz to the rediscovery of chaos.

But Turing's own vision was already moving much further along. He spent the academic year 1947–8 working in Cambridge, on secondment from the NPL, and wrote a paper, never published in his lifetime, on what we would now call neural nets – an attempt to demonstrate that any sufficiently complex mechanical system could learn from experience, without actually being programmed by an outside intelligence. By 1950, settled in Manchester, he was ready to begin to apply the knowledge he had gained about mechanical systems and electronic computers to biological systems and the human brain. The jump from there to his work on how embryos develop wasn't as great as it might seem, since Turing wasn't only interested in how brains grow and form connections; his interest in the way the variety of living things develop from simple beginnings had been stimulated in his youth by reading D'Arcy Thompson's classic book *On Growth and Form*. So at the time Turing was elected as a Fellow of the Royal Society in 1951, for his contributions to computer science, he was already working on what would probably, had he lived, have been an even greater contribution to science.

Even Turing couldn't leap straight from the understanding of biology that existed at the beginning of the 1950s to a model of how the brain

itself develops its network of connections – after all, the double helix structure of DNA, the life molecule, was not determined until 1953, by Francis Crick and James Watson, working in Cambridge. Instead, he decided to tackle the fundamental problem of how structure emerges in the developing embryo from what is an almost spherical, almost featureless initial blob of cells, the blastocyst formed from the fertilized egg. In mathematical terms, the problem was one of broken symmetry, a phenomenon already familiar to physicists in other contexts (not least, Bénard convection). A good example of symmetry breaking occurs when certain kinds of magnetic substances are heated and then cooled down. Magnetic materials such as iron can be regarded as made up of a collection of tiny dipoles, like little bar magnets. Above a critical temperature, known as the Curie point (after Pierre Curie, who discovered the effect in 1895), there is enough heat energy to break any magnetic links between these dipoles, so that they can spin around and are jumbled up in a random fashion, pointing in all directions, so that there is no overall magnetic field. In magnetic terms the material can be said to be spherically symmetric, because there is no preferred magnetic direction. As the temperature drops below the Curie point (760 °C for iron), the magnetic forces between adjacent dipoles overcome the tendency of their declining heat energy to jumble them up, and the dipoles line up to produce an overall magnetic field, with a north pole at one end and a south pole at the other end. The original symmetry has been broken. Such a change is called a phase transition, and is similar to the way water freezes into ice in a phase transition at 0 °C. The concept of a phase transition also has important applications in particle physics, which we need not go into here; the relevant point is that although such

ideas had not been widely applied in biology before 1950, at that time it was natural for a mathematician moving into the theory of biological development to think in terms of symmetry breaking, and to have the mathematical tools describing the general nature of such transitions available.

In 1952, Turing published a paper which described in principle how the symmetry of an initially uniform mixture of chemicals could be spontaneously broken by the diffusion of different chemicals through the mixture. The anticipated relevance of this to biology was clear from the title of the paper, 'The chemical basis of morphogenesis',<sup>8</sup> and Turing's proposal was that something like the process that he described mathematically might actually take place in the developing embryo, to produce patterns where none existed originally.

At first sight, Turing's proposal seems utterly counter-intuitive. We expect diffusion to mix things up, and to destroy patterns, not to create patterns where none exist originally. The obvious example is the way a drop of ink placed in a glass of water spreads out to fill the glass with a uniform mixture of water and ink; it seems almost as if Turing was suggesting a reversal of the thermodynamic processes that operate on this scale, with time running backwards and a uniform mixture of water and ink separating out into a glass of clear water surrounding a single droplet of ink. But that was not the case, and the key to Turing's insight is that the pattern-forming process he described involves at least two chemicals interacting with one another.

It all depends on the process known as catalysis, whereby the presence of a particular chemical substance (the catalyst) encourages a

particular chemical reaction to take place. In some cases, the presence of a chemical compound (which we can represent by the letter A) in a mixture of chemicals encourages reactions which make more of the compound. The reaction is said to be autocatalytic, and since the more A there is, the more A is produced, we can see that this is another example of positive feedback at work in a non-linear process. On the other hand, there are chemicals that act in the opposite way, to inhibit certain chemical reactions. Logically enough, they are called inhibitors. And there is nothing to say that a single substance cannot encourage more than one chemical reaction at the same time. Turing calculated that patterns could arise in a mixture of chemicals if the catalyst A not only encouraged the production of more A, but also encouraged the formation of another compound, B, which was an inhibitor that acted to slow down the rate at which more A is produced. His crucial suggestion was that once A and B formed they would diffuse through the mixture of chemicals at different rates, so that there would be more A than B in some parts of the mixture, and more B than A in other places. In order to calculate just how much A and B there would be in different places, Turing had to use the simplest equations he could, since electronic computers were still highly limited in their abilities, and in very short supply, so he was working everything out on paper. This meant working with linear approximations to the real non-linear equations describing the situation, and these equations turn out to be very unstable, in the sense that a small error in one part of the calculations leads to a big error later on. As a result, Turing could only calculate what was going on for the simplest systems, but that was enough to hint at the possibilities. Turing himself acknowledged that a full investigation of his

ideas would have to await the development of more powerful digital computers, but in developing his ideas as best he could beyond those sketched out in his 1952 paper he showed how the competition between A and B was the key to pattern formation, and that it was essential that B must diffuse through the mixture more quickly than A, so that while the runaway production of A by the autocatalytic feedback process is always a local phenomenon, the inhibition of A by B is a widespread phenomenon. Also the rapid diffusion of B away from where it is being made means that it doesn't entirely prevent the manufacture of A at its source.

To picture what is going on, imagine the mixture of chemicals sitting quietly in a glass jar. Because of random fluctuations, there will be some spots in the liquid where there is a slightly greater concentration of A, and this will encourage the formation of both A and B at those spots. Most of B will diffuse away from these spots, and prevent any A forming in the spaces between the spots, while the autocatalytic process ensures that more A (and B) continues to be produced at the spots. (There will also be places in the original mixture where random fluctuations produce an excess of B to start with, but, of course, nothing interesting will happen there.) Now suppose that chemical A is coloured red and chemical B is coloured green. The result will be that an initially uniform, featureless jar of liquid transforms itself spontaneously into a sea of green dotted with red spots that maintain their positions in the liquid (as long as the liquid is not stirred up or sloshed around). The pattern is stable, but in this particular case it is a dynamic process, with new A and B being produced as long as there is a source of the chemicals from which they are being manufactured, and as long as there is a 'sink'

through which the end-products can be removed. In the terminology that ought to be becoming familiar by now, the pattern is stable and persistent provided that we are dealing with an open, dissipative system which is being maintained in a non-equilibrium state. Turing also described mathematically systems in which there is a changing pattern of colour rippling through the liquid, where it would be more obvious to any observer (if such systems could be replicated in real experiments) that a dynamic process was going on. Today, an autocatalytic compound such as A is called an actuator, while B is indeed known as an inhibitor; Turing himself, though, didn't use these terms, and referred to B as a 'poison', which now has chilling echoes of his own death.<sup>9</sup> Although it may seem far removed from the development of an embryo (let alone a brain), the essential point about Turing's discovery was that it provided a natural chemical way in which symmetry could be broken to spontaneously create patterns in an initially uniform system – if there were real chemical systems that behaved in this way.

Intriguing though Turing's ideas were, although his paper is seen as being of seminal importance to theoretical biology today, in the 1950s and through most of the 1960s it attracted little interest among chemists and biologists, precisely because nobody knew of any real chemical system which behaved in the way that this mathematical model described. Nobody, that is, except one person, the Russian biochemist Boris Belousov, and he, not being a reader of the *Philosophical Transactions of the Royal Society*, didn't know about Turing's work, just as Turing never learned about Belousov's work before his own untimely death. At the beginning of the 1950s, Belousov was working at the Soviet Ministry of Health, and interested in the way glucose is broken

down in the body to release energy. He was already in his early fifties, an unusually advanced age for any scientist to make a major new discovery, and had a background of work in a military laboratory, about which little is recorded, except that he reached the rank of *Combrig*, roughly equivalent to a Colonel in the army and an unusually high distinction for a chemist, before retiring from this work after the Second World War. Like many other metabolic processes, the breakdown of glucose that Belousov was interested in is facilitated by the action of enzymes, different kinds of protein molecules which act as catalysts for different steps in the appropriate suite of chemical reactions. Belousov concocted a mixture of chemicals which he thought would mimic at least some features of this process, and was utterly astonished when the solution in front of him kept changing from being clear and colourless to yellow and back again, with a regular, repeating rhythm. It was as if he had sat down with a glass of red wine, only to see the colour disappear from the wine, then reappear, not once but many times, as if by magic. This seemed to fly in the face of the second law of thermodynamics, as it was understood at the time. It would be entirely reasonable for the liquid to change from clear to yellow, if the yellow represented a more stable state with higher entropy. And it would be entirely reasonable for the liquid to change from yellow to clear, if clear represented a more stable state with higher entropy. But *both* states couldn't be at higher entropy than the other! It was as if, using the original nineteenth-century ideas about the relationship between thermodynamics and time, the arrow of time itself kept reversing, flipping backwards and forwards within the fluid.

Belousov might have been less astonished if he had known about

some earlier work, which to an extent foreshadowed both his own experiments and Turing's mathematical models. Back in 1910, another mathematical modeller, the Austrian Alfred Lotka (1880–1949), had come up with a mathematical description of a hypothetical chemical system which oscillated in this way, first producing an excess of one compound then reversing to produce an excess of another compound, then reversing again, and so on.<sup>10</sup> In a neat example of how simple processes involving feedback can often be described by the same laws under what may seem at first sight to be very different circumstances, in the 1930s the Italian Vito Volterra (1860-1940) showed that Lotka's equations worked quite well as a description of the way fish populations change when there is an interaction between one prey species and one predator species, following a boom and bust cycle as first one species and then the other flourishes. And as early as 1921 the Canadian-born chemist William Bray (1879-1946), then working at the University of California, Berkeley, had found that a chemical reaction involving hydrogen peroxide and iodate produced a mixture of iodine and oxygen in which the proportions of the two products oscillated in more or less the way Lotka had described. Even though Bray referred to Lotka's model when he announced his discovery, the response of his colleagues was essentially that, since his results violated the second law of thermodynamics, there must be something wrong with his experiment, so that the 'discovery' was an artefact caused by carelessness in mixing and keeping track of the ingredients. Lotka, Volterra and Bray were all dead when Belousov encountered almost exactly the same response when he tried to publish a paper describing his findings in 1951, a year before Turing would publish his seminal paper. Belousov's results, said



the editor of the journal he submitted them to,<sup>11</sup> contravened the second law of thermodynamics, so his experimental procedure must be faulty.

Some idea of the status of the second law as something approaching Holy Writ at this time can be gleaned from a famous quotation from the writings of the British astrophysicist Arthur Eddington. In his book *The Nature of the Physical World*, published by Cambridge University Press in 1928, he wrote:

The law that entropy always increases – the second law of thermodynamics – holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations – then so much the worse for Maxwell's equations. If it is found to be contradicted by observation – well, these experimentalists do bungle things sometimes. But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.

Replace the word 'theory' by 'experiment' in the final sentence and you have essentially the response to Belousov – he was a bungling experimentalist, and should 'collapse in deepest humiliation'. I confess that I have previously quoted this remark by Eddington with approval, and without pausing to explain that although true as far as it goes, the original formulation of the second law is not the ultimate truth that he thought it was, and has to be reconsidered in non-equilibrium situations and where gravity is involved. But that was far from being widely understood in 1951.

Belousov's reaction to the rejection of his paper was perhaps what one might have expected from a man of his age and background. He felt personally insulted at the slur on his professional skill as an experimenter, and resolved not to do any more work in the field if his results were going to be dismissed in this way. One of his younger colleagues, S. E. Shnoll, tried to encourage him to persevere, but without success. After trying unsuccessfully for years to get his work published,

in 1959 Belousov managed to smuggle a two-page summary of his findings into print by attaching it to the published version of a report he had given on a completely different topic, to a symposium on radiation medicine held in Moscow the previous year. Then, he entirely abandoned this work.<sup>12</sup> The proceedings of the meeting (which were not subject to the refereeing process or approval by an editor) appeared only in Russian; they were almost unread outside the Soviet Union, and not exactly widely read within the USSR. But Shnoll maintained his interest in Belousov's work, and in the 1960s he drew it to the attention of one of his postgraduate students, Anatoly Zhabotinsky, encouraging him to follow up Belousov's lead. Just that one person in the next generation of chemists had the incentive to follow up Belousov's two-page summary, but one was all it took to grab hold of the discovery and make the scientific world pay attention.

Zhabotinsky was a graduate student at Moscow State University when he was introduced to Belousov's discovery and was intrigued enough to try the reaction for himself (it's hard not to be intrigued when your supervisor 'suggests' that you look into a particular research problem), confirming that it worked in the way Belousov had described, and going on to tinker with the ingredients until he produced a mixture which exhibited a much more dramatic colour change, from red to blue and back again. It shouldn't be a surprise that it was a student who took up the idea, since young researchers are generally less hidebound by tradition than their elders and more willing to consider the possibility that sacrosanct laws might be broken (although more often than not the laws do stand up to this pressure). Zhabotinsky described his results at an international meeting held in Prague in 1968, where western

scientists first learned about the intriguing behaviour of what came to be known as the Belousov-Zhabotinsky, or BZ, reaction. This made all the more impact because some of them were already aware of Turing's work, but had not, as we have seen, thought that it might be relevant to real chemical systems. Unlike Turing, Belousov lived long enough to see his discovery taken up in this way, but he died in 1970, before the full importance of such reactions was appreciated.

Hardly surprisingly, one of the first people to pick up on Zhabotinsky's work and develop a theoretical model to describe the kind of oscillations seen in the BZ reaction was Ilya Prigogine, who had met Turing in England in 1952, shortly after Turing had written his paper on the chemistry of making patterns, and had discussed that work with him. Working now in Brussels with his colleague René Lefever, and jumping off from Turing's work, before the end of 1968 Prigogine had come up with a model involving two chemical substances which are converted into two other chemical substances in a multi-step process that involves two short-lived intermediate products. The model became known as the Brusselator; we don't need to go into the details of how it works (although they are only marginally more complicated than Turing's model of how to make spots), but the important point is that the reactions involve feedback and non-linearity. If we imagine that the products of the series of reactions are respectively red and blue, then the Brusselator tells us that as long as the mixture is kept in a dissipative state far enough away from equilibrium, with a constant supply of raw materials being added and the end-products being drained away, it will change regularly from red to blue and back, without settling down to the uniform purple colour that would be expected from a naive faith in the

second law of thermodynamics. The whole process is, indeed, consistent with the understanding of how the second law has to be modified in far from equilibrium conditions already developed by Prigogine and his colleagues.

In the 1970s, progress was made both with the modelling and with the investigation of real chemical systems in which structure is spontaneously created by self-organization. On the experimental side, chemists soon found ways to make waves of colour travel through chemical mixtures – in the BZ reaction itself, in a shallow dish of the appropriate chemical stew, it is possible to produce concentric circles and spirals of red and blue which travel outwards from their source. Among the huge variety of patterns developed in similar experiments over the following decades, in the 1990s chemists at last found a way to produce *stationary* patterns of spots just like the ones originally described by Turing. The detailed chemistry of the BZ reaction was investigated in the early 1970s by a team at the University of Oregon, who identified at least thirty separate chemical species involved in the interlocking series of chemical reactions that produce the overall colour changes, including some short-lived intermediaries like those in the Brusselator. This led them, in 1974, to come up with a model describing the key steps in the BZ process in terms of just six kinds of chemical substance, interacting with one another in five distinct steps, including the all-important influence of autocatalysis. The difference between this model and both Turing's model and the Brusselator was that whereas the latter dealt with hypothetical substances labelled A, B and so on, the Oregon team's model dealt with real chemical compounds involved in actual chemical reactions. The model became known as the Oregonator. Again, we don't

need to go into details; the point is that what seems to be a complicated pattern of self-organization can be explained in terms of a few simple interactions.

But there's more. It is actually possible to set up the BZ mixture in an unchanging, uniform state – if you wait long enough without adding any new ingredients, it will eventually settle down all by itself. Now, when you add some more reactants, the oscillatory behaviour we have described will begin. There has been a bifurcation, with the system changing from a period one state to a period two state. You can guess what is coming next. If you gradually increase the rate at which new ingredients are flowing into the system and 'waste products' are being removed, at a critical threshold the oscillatory pattern becomes more complicated, exhibiting a double rhythm, as the system bifurcates further into a period four state. Keep increasing the flow of reactants, and the period-doubling cascade familiar from the dripping tap and other examples we have mentioned appears in all its glory, the periodic pattern becomes less obvious, and the system tips over the edge in chaos (in this case, actually rather quickly, with everything after period four happening in a rush). All the interesting things we have been describing, notably self-organization and the spontaneous appearance of patterns out of uniform systems, occur on the edge of chaos. All of this can be described in the language of phase space, limit cycles and attractors, just as in the earlier examples we discussed. There is even evidence for the existence of strange attractors associated with the trajectories in phase space describing the evolution of a BZ reaction.<sup>13</sup> But although it is good to see how all of this relates to the story so far, we seem to have wandered a long way from Turing's hope of providing insight into the

process of morphogenesis, the aspect of embryology concerned with the development of pattern and form in the growing embryo. Not so; although his ideas have yet to be proved as a key contribution to the overall development of the embryo (though work is still very much in progress on this), there is one area in particular where they have been spectacularly, and highly visibly, successful.

This triumph for the Turing mechanism concerns the way markings such as stripes and spots develop on the skin and coats of mammals, and more generally how patterns form on the surfaces of other animals. A major contribution to this investigation was made by James Murray, initially at the University of Oxford and later of the University of Washington, Seattle, who summed up many of his discoveries in a highly readable article in *Scientific American* in 1988, under the title 'How the leopard gets its spots'; <sup>14</sup> a more technical but less readable full account of this work can be found in his book *Mathematical Biology*. Murray found that not only the spots of the leopard but also the stripes on a zebra, the blotches on a giraffe, and even the absence of patterning on the coat of a mouse or an elephant, can all be explained by the same simple process, involving diffusion of activator and inhibitor chemicals across the surface of the developing embryo at a key stage during its growth. Nobody has yet proved that this definitely is the way the patterning process works, but it is certainly true that the kind of patterns formed are the ones that would form if such a patterning process were at work. The idea has great appeal, not least because of its simplicity. At one level, in terms of the DNA code describing the construction of a single individual body, storing the information which says, in effect, 'release these two chemicals at this stage of development' takes up much

less room (less memory, using the computer analogy) than a literal blueprint describing precisely the exact location of every spot and stripe on the adult body. And at another level, having one simple mechanism that explains how and why patterns of different kinds appear on the bodies of different animals, and also why some animals have no pattern at all, is much more parsimonious than having a different blueprint to describe each different kind of pattern on each different kind of animal. Finally, as we shall see, the kind of simple mechanism originally proposed by Turing and discussed in detail by Murray and his contemporaries offers important insights into the mechanisms of evolution. It is not quite *always* true in science that the simplest solution to a problem is certain to be the right one, but this approach, known as Ockham's Razor,<sup>15</sup> has proved an extremely reliable rule of thumb in most circumstances, and it is certainly always advisable to choose the simplest solution unless there are overwhelming reasons not to. In this case, the Turing process is the simplest solution to the puzzle.

The patterns we see on the surfaces of mammals are either skin colours or colours that the hairs of the pelt have picked up as they grow from a particular region of skin. Either way, it is the presence of something in the skin that determines the colour. There are surprisingly few of these colours – black, white, brown, and a range of orangey-yellow colours – with just about the full range expressed in the markings of a tortoiseshell cat. The colours depend on the presence or absence of two pigments produced by cells in the skin, with the intensity of the colour depending on how much of each pigment is present – eumelanin, which gives a black or brown colour, and pheomelanin, which gives a yellow or orange colour (the absence of either melanin leaves the hair or

skin white). But what decides whether certain cells are 'switched on' to produce one or the other kind of melanin, and in what quantity?

Murray's triumph was to show that even though we do not yet know exactly what chemicals are involved, the patterns we see on real, living animals are exactly (and *only*) the patterns that would be produced by Turing reactions involving diffusion of an activator and an inhibitor across the surface of the growing embryo early in its development, within a few weeks of conception (in the zebra, for example, there is evidence that patterns in the skin are laid down about 21-35 days after conception, with an overall gestation period of 360 days). If the presence of one of these chemicals then also switched on the ability of a cell to produce melanin, the result would be that a pattern equivalent to the pattern seen in the shallow-dish versions of the BZ reaction would be invisibly imprinted on those cells, but would only show up later in life when some other chemical trigger (or the growth of hair) sent a message to begin making melanin – a message that would be received by every cell in the skin, but would only be acted upon by those cells that had been primed during the Turing reaction.

So, without worrying too much about the biochemical processes involved, 'all' that Murray had to do was to develop a mathematical model that could be used to predict the way patterns would form as a result of the Turing reaction on surfaces shaped like the surfaces of mammalian embryos at different stages of their development. Because the process involves waves travelling across (or through) the surfaces, both the size and shape of the surface affect the pattern produced by the reaction. As Murray points out, the situation is superficially rather like the way in which the sounds produced by the stretched skin of a drum



depend on the size and shape of the drumskin, because different-sized sound waves (that is, ones with different wavelengths, corresponding to different musical notes) fit neatly into different-sized skins (there is actually a rather close mathematical analogy between the two systems, although the physical processes involved are very different). He found that if the surface is very small at the time the Turing process is triggered, no pattern can form at all. There is not enough room for the mechanism to get to work, or, if you prefer to think in those terms, the 'wavelength' associated with the reaction is bigger than the size of the skin, so the pattern cannot be seen (it would be like trying to paint fine details on a small canvas using a paint-roller designed for use on walls). At the other extreme, where relatively large surfaces are involved, the interactions become too complicated to allow any overall patterns to emerge. It is as if very many different conversations are going on in a room at once, producing an overall effect which is just a uniform hubbub of noise. There is, in fact, a very fine scale 'patterning' possible on large surfaces, and if you look closely enough you would see that not every hair on the surface of, say, an elephant has precisely the same colour, but from a distance the elephant looks a uniform colour, just as in a pointillist painting there is a very fine structure of paint spots which merge into a uniform colour when seen from a little distance away (and just as, in a crowded room full of chattering people, we can pick out what our neighbour is saying in spite of the background noise). So, according to the model, both very small and very large mammals should have unpatterned surfaces, which is exactly what we see in nature. What happens in between these extremes?

Starting small and working upwards in size, it turns out that the first

kind of pattern that can form is one with rather broad bands, then stripes, followed by spots, followed by large blotches separated by narrow strips forming the boundaries between the blotches, while for larger surfaces still the blotches merge into a uniform colour. The overall patterns produced closely resemble the range of patterns seen in nature, from the leopard's spots to the stripes on a tiger or a zebra, to the blotchy markings on a giraffe. As Murray sums up in his book:

We see that there is a striking similarity between the patterns that the model generates and those found on a wide variety of animals. Even with the restrictions we imposed on the parameters for our simulations the wealth of possible patterns is remarkable. The patterns depend strongly on the geometry and scale of the reaction domain although later growth may distort the initial pattern... It is an appealing idea that a single mechanism could generate all of the observed mammalian coat patterns.

But, of course, an animal doesn't *have* to have a pattern, even if there is a biochemical mechanism which allows it to have a pattern. The mechanism can always be switched off, and there are very obvious evolutionary reasons why a polar bear, say, should have a uniform white colour. But where patterns have evolved, a very neat example of the correlation between the kind of pattern and the size of the surface available can be seen in the tails of many members of the cat family. For tails which are more or less cylindrical, the patterns seen on them may be either spots or circular bands, stripes around the tail. But for tails which taper down at the end, like that of the jaguar, even if the base of the tail is covered in spots, the tip of the tail is marked with striped bands, in line with the prediction of the model that bands always form on smaller areas and spots are produced on larger areas.



Figure 4.2 *The patterns produced on the body of an animal by self-organized chemical processes during the development of an embryo (see below) are related to the size of the animal. The chemical diffusion processes produce stripes on smaller areas, and spots in larger areas.*

One of the key features of the model, though, is that the kind of pattern that forms over the surface of an animal does not depend on the size and shape of the adult, but on the size and shape of the embryo at the time the Turing process is at work. Clearly there is some correlation with the size of the adult, because from very soon after conception elephant embryos tend to be larger at the same stage of development than mouse embryos; but the significance of the embryo size is beautifully highlighted by the differences in the stripes of the two kinds of zebra, *Equus burchelli* and *Equus grevyi*. The former has fewer and broader stripes than the latter, making them distinctly different when seen alongside each other, even though the adults are roughly the same size. By counting the number of stripes in each case, and taking account of the way the pattern had been distorted by the growth of the animal, in the 1970s J. B. L. Bard showed that the pattern seen on *burchelli* must be laid down on the embryo when it is twenty-one days old, while the pattern seen on *grevyi* must be laid down on the embryo when it is five weeks old. This was known before Murray came up with his

mathematical model of the Turing effect, but the differences exactly correspond to the predictions of the model, with broader stripes corresponding to an earlier diffusion of the activator and inhibitor across the surface of the embryo. The dramatic way in which the genetics and the environment ('nature' and 'nurture') combine in this way was highlighted early in the year 2002, when news was released of the birth of the first cloned kitten. Because the pattern of colours on multicoloured animals results from a combination of their genetic inheritance and events in the womb (including, for example, the amount of nutrition received by the developing embryo), the kitten did not have exactly the same pattern of markings in her fur as her mother, even though the two animals had identical DNA.

This also brings us on to the significance of all this to our understanding of evolution. The visible differences between the patterns in the two species of zebra are produced simply by changing the time at which the Turing effect is at work in the embryo. As far as we know there is no evolutionary advantage in either pattern in this particular case (not every feature of anatomy has to be adaptive). But if there were an advantage in having narrower (or broader) stripes, perhaps in terms of providing better camouflage, it is easy to see how natural variations from one individual to another could provide the raw material to respond to the selection pressure and shift the whole population of one species of zebra in that direction, without any change in anything except the *timing* of a particular event during the development of the embryo – one of the smallest 'mutations' imaginable. We shall have more to say about evolution – much more – in the rest of this book. But to set the scene, it's worth highlighting a few other examples of the way in which

external form seems to be under the control of simple chemical processes. We went into some detail on Murray's model of how mammals get their markings, both because it was one of the first of this kind of model and because it follows on naturally from our discussion of the Turing mechanism and the BZ reaction; but there will be no need to go into quite so much detail for these other models, since they operate on very similar principles.

Without pre-empting our later discussion of evolution, we should, perhaps, first briefly recap the Darwinian process of natural selection, because there is so much confusion about the relationship between evolution and natural selection. Evolution is a fact, a process seen at work both in living organisms today and in the fossil record.<sup>16</sup> In the same way, the way apples fall from trees and the Moon orbits around the Earth are facts. The movement of apples and the Moon (and other things) is described by a theory of gravity – Newton's theory is adequate for most human purposes, but Einstein's theory is appropriate if you are dealing with extreme phenomena such as collapsed stars. Evolution is described by a theory, Darwin's theory of natural selection, which is adequate for most purposes on Earth today, but which, like Newton's theory of gravity, has been modified to take account of other things which were unknown to its originator. Darwin's theory is to evolution what Newton's theory is to gravity, but we do not yet have an evolutionary equivalent of Einstein's theory of gravity, a complete theory of evolution going beyond Darwin's theory to incorporate all the facts known today, although various suggestions have been made about how Darwin's theory needs to be modified. The essence of Darwin's theory remains at the core of our understanding of evolution. It says that

offspring-resemble their parents, but imperfectly, so that there are variations from one generation to the next and among individuals in each generation. The individuals that are best suited to their environment (best 'fitted', in the sense of the way a piece fits into a jigsaw puzzle) do best in terms of finding food, mating and (crucially) reproducing, so they leave more offspring in the next generation. Because those offspring resemble their parents, they will inherit the particular attributes that made their parents successful, but perhaps with slight variations. And so on, in succeeding generations.

This is a process of natural selection, which selects the fittest (in the jigsaw puzzle sense, not the athletic sense, although the two may go together) in each generation. In the classic example, if it is an advantage to be able to get at juicy leaves in the tree tops, individual grazers with long necks will do better in the struggle for survival and raise more offspring while individuals with shorter necks will get less food and find it harder to raise their offspring. Provided that the tendency for long or short necks is inherited, but there are still variations among individuals (some necks are longer or shorter than the average, whatever the average may be), over many generations the selection pressure favouring longer necks will lead to the evolution of a creature like the giraffe. Variety isn't just the spice of life; it is at the very heart of how life works. Darwinian theory explains beautifully why species should be superbly fitted for their ecological niches, like the fit of a key in a lock. But it is important to appreciate that it also allows scope for features to exist that are evolutionarily neutral, in terms of selection, conferring neither an advantage nor a disadvantage on the individuals that possess them. These include variations on a theme like the stripy patterns seen

on the shells of some ordinary garden snails, where individual members of the same species can look much more different from one another than an individual *burchelli* zebra does from an individual *grevyi*, even though they are members of different species. But that's enough about the basics of evolution to be going on with.

Following on from Murray's work on stripes and spots formed by the Turing mechanism, many more of nature's patterns have been investigated in the same sort of way, both by Murray and by others. One of the most relevant of these investigations to the story we have been telling was carried out by Hans Meinhardt, who works at the Max Planck Institute for Developmental Biology, in Tübingen, and his colleague André Koch. Using a similar approach to Murray, but based on the mechanism of the BZ reaction instead of the Turing reaction, they found that in their mathematical model very lifelike patterns (including those corresponding to the spots on a leopard) could be produced when the production of an actuator is triggered at random places on the skin of the embryo at the appropriate moment during its development. The advantage of this particular model is that it is able to produce more complicated patterns, even though the underlying chemistry is still very simple. Patterns on the shells of marine creatures also match the patterns we would expect to be produced by some chemical processes involving actuator and inhibitor compounds, and many biologists believe that they may have found a species where they can see the process at work. In the angel fish *Pomacanthus imperator* the adult has parallel stripes which run from head to tail along the fish. As the fish grows, new stripes form so that the individual stripes stay the same size, but the spacing between the stripes also stays the same. The new stripes develop from forks in

some of the earlier stripes, branching like a single railway track forking at a set of points to become two parallel tracks. In the 1990s, Shigeru Kondo and Rihito Asahi, working at Kyoto University, developed a mathematical model which reproduced exactly this pattern of behaviour, using the Turing mechanism. This suggests that the Turing process itself is still going on in these adult fish, rather than being a one-off event that happened during embryonic development, and raises the hope that the actual chemicals involved in the process might soon be identified.

Similar models have been used to mimic the patterns produced on the wings of butterflies, one of many such applications, but one we pick out because of its relevance to what is coming later. Murray studied the mechanisms for producing many features of these wing patterns, including the formation of large spots that bear a superficial resemblance to eyes (these are thought to be selected by evolution because a predator taking a quick look at the butterfly might see, not a juicy tit-bit of food, but the eyes of a larger creature staring back at it). One aspect of this work showed how easy it is to make such patterns from the basic chemistry, without any need for a complicated blueprint stored in the genes of the butterfly, and therefore how likely it is that such a feature could indeed be selected by evolution. But in this particular case, the model also showed that a key feature of the eyespot pattern, its radius, changes gradually if the conditions under which the relevant chemistry is going on change gradually. For example, butterflies that develop at one temperature will have eyespots of a certain size, while butterflies that develop at another temperature will have eyespots of a different size, and there will be a continuous, smooth correlation between the size of the spots and the temperature. This is only worth



commenting on because it is not the only pattern of behaviour that can result from the kind of underlying chemistry we have been discussing. In some processes, small changes in the environment in which the reactions are taking place have little effect until a critical point is reached, and then the process suddenly switches into a new phase of operation. One example, cited by Murray, of a process which is sensitive in this way, is the development of the vertebrate limb. If the biochemical processes involved in the stage of development where fingers are growing are disrupted slightly, the result is not a slightly larger or smaller hand, but a hand with a sixth finger. This can arise naturally through a mutation (a small change in the DNA code of a gene, perhaps resulting from a copying error) which causes a slight change in development; the altered DNA will then be passed on down the generations, unless the mutation is harmful. This is why such traits tend to run in families, including that of Anne Boleyn, one of the wives of Henry VIII, who was born with six fingers on one hand, although one was very quickly cut off. In another example famous among developmental biologists, a Boston man had a kind of double hand, with no thumb and seven fingers arranged in two sets, one of three and one of four, either side of where the thumb 'ought' to have been. Similar structures can also be produced by grafting cells from one limb bud on to another.<sup>17</sup> And they can be described mathematically by the appropriate models – models which describe processes such as the Turing mechanism that occur in dissipative systems at the edge of chaos. The point to take hold of is that sometimes small changes in the environment, or small mutations, can have big effects on the body form that develops. This is one of the relatively new things that were unknown to Darwin, and which give us more insight into how

evolution works. But remember that under other circumstances there is gradual change. As Murray says:

Depending on the mechanism and the specific patterning feature we focus on, we can have gradual or discontinuous change inform ... it is clear that to understand how evolution takes place we must understand the morphogenetic processes involved.

In fact, this is as far as we wish to delve into morphogenesis. All we want to carry away from this discussion is the knowledge that in morphogenesis and in evolutionary biology sometimes small effects produce small changes, and sometimes small effects produce big changes – and that this can be understood, in principle, in terms of pattern-formation models involving very simple chemistry. It turns out that small, random changes can also have either big or small effects in the world at large, in particular when we are dealing with dissipative systems at the edge of chaos – and that understanding how and why this happens will help us to a better understanding of the emergence of both life itself and intelligence.

## 5

### **Earthquakes, Extinctions and Emergence**

When other people hear scientists refer to 'complex systems' this sometimes 'creates a barrier, since to many people 'complex' means 'complicated', and there is an automatic assumption that if a system is complicated it will be difficult to understand. Neither assumption is necessarily correct. A complex system is really just a system that is made up of several simpler components interacting with one another. As we have seen, the great triumphs of science since the time of Galileo and Newton have largely been achieved by breaking complex systems down into their simple components and studying the way the simple components behave (if necessary, as a first approximation, taking the extra step of pretending that the components are even simpler than they really are). In the classic example of the success of this approach to understanding the world, much of chemistry can be understood in terms of a model in which the simple components are atoms, and for these purposes it scarcely matters what the nuclei of those atoms are composed of. Moving up a level, the laws which describe the behaviour of carbon dioxide gas trapped in a box can be understood in terms of roughly spherical molecules bouncing off one another and the walls of their container, and it scarcely matters that each of those molecules is made up of one carbon atom and two oxygen atoms linked together. Both systems are complex. in the scientific sense. but easy to understand.

And the other key to understanding, as these examples highlight, is choosing the right simpler components to analyse; a good choice will give you a model with widespread applications, just as the atomic model applies to all of chemistry, not just to the chemistry of carbon and oxygen, and the 'bouncing ball' model of gases applies to all gases, not just to carbon dioxide.

At a more abstract level, the same underlying principle applies to what mathematicians choose to call complex numbers. The name has frightened off many a student, but complex numbers are really very simple, and contain only two components, scarcely justifying the use of the term 'complex' at all. The two components of a complex number are themselves everyday numbers, which are distinguished from each other because one of them is multiplied by a universal constant labelled  $i$ . So whereas an everyday number can be represented by a single letter (say,  $X$ ), a complex number is represented by a pair of letters (say,  $A + iB$ ). It happens that  $i$  is the square root of  $-1$ , so that  $i \times i = -1$ , but that doesn't really matter. What matters is that there is a fairly simple set of rules which tell you how to manipulate complex numbers – what happens when you multiply one complex number by another, or add two of them together, and so on. These rules really are simple – much simpler, for example, than the rules of chess. But using them opens up a whole new world of mathematics, which turns out to have widespread applications in physics, for example in describing the behaviour of alternating electric current and in the wave equations of quantum mechanics.

But there's a more homely example of the simplicity of complexity. The two simplest 'machines' of all are the wheel and the lever. A toothed

cogwheel, like the gearwheels of a racing bicycle, is in effect a combination of lever and wheel.<sup>1</sup> A single wheel – even a single gearwheel – is not a complex object. But a racing bicycle, which is essentially just a collection of wheels and levers, is a complex object, within the scientific meaning of the term – even though its individual component parts, and the way they interact with one another, are easy to understand. And this highlights the other important feature of complexity, as the term is used in science today – the importance of the way things interact with one another. A heap of wheels and levers would not in itself be a complex system even if the heap consisted of all the pieces needed to make a racing bike. The simple pieces have to be connected together in the right way, so that they interact with one another to produce something that is greater than the sum of its parts. And that's complexity, founded upon deep simplicity.

When scientists are confronted by complexity, their instinctive reaction is to try to understand it by looking at the appropriate simpler components and the way they interact with one another. Then they hope to find a simple law (or laws) which applies to the system they are studying. If all goes well, it will turn out that this law also applies to a wider sample of complex systems (as with the atomic model of chemistry, or the way the laws of cogwheels apply both to bicycles and chronometers), and that they have discovered a deep truth about the way the world works. The method has worked for over 300 years as a guide to the behaviour of systems close to equilibrium. Now it is being applied to dissipative systems on the edge of chaos – and what better terrestrial example could there be of a system in which large amounts of energy are dissipated than an earthquake?

One of the most natural questions to ask about earthquakes is how often earthquakes of different sizes occur. Apart from its intrinsic interest, this has great practical relevance if you live in an earthquake zone, or if you represent an insurance company trying to decide what premiums to charge for earthquake insurance. There are lots of ways in which earthquakes might be distributed through time. Most earthquakes might be very large, releasing lots of energy which then takes a long time to accumulate once again. Or they might all be small, releasing energy almost continuously, so that there is never enough to make a big 'quake. There could be some typical size for an earthquake, with both bigger and smaller events relatively rare (which is the way the heights of people are distributed, around some average value). Or they could be completely random. There is no point in guessing; the only way to find out is to look at all the records of earthquake, and add up how many of each size have occurred. Appropriately, the first person to do this was Charles Richter (1900–85), who introduced the eponymous scale now widely used to measure the intensity of earthquakes.<sup>2</sup>

The Richter scale is logarithmic, so that an increase of one unit on the scale corresponds to an increase in the amount of energy released by a factor of 30; a magnitude 2 earthquake is 30 times as powerful as a magnitude 1 earthquake, a magnitude 3 earthquake is 30 times more powerful than a magnitude 2 earthquake (and therefore 900 times more powerful than a magnitude 1 earthquake), and so on. Although the name attached to the scale is Richter's alone, he worked it out, at the beginning of the 1930s, with his colleague Beno Gutenberg (1889–1960), and in the middle of the 1950s the same team turned their attention to the investigation of the frequency of earthquakes of different

sizes. The team looked at records of earthquakes worldwide, and combined them in 'bins' corresponding to steps of half a magnitude on the Richter scale – so all the earthquakes with magnitude between 5 and 5.5 went in one bin, all those between 5.5 and 6 in the next bin, and so on. Remembering that the Richter scale itself is logarithmic, in order to compare like with like they then took the logarithm in each of these numbers.<sup>3</sup> When they plotted a graph showing the logarithm of the number of earthquakes in each bin in relation to the magnitude itself (a so-called 'log-log plot'<sup>4</sup>), they found that it made a straight line. There are very many small earthquakes, very few large earthquakes, and the number in between lies, for any magnitude you choose, on the straight line joining those two extreme points. This means that the number of earthquakes of each magnitude obeys a power law – for every 1,000 earthquakes of magnitude 5 there are roughly 100 earthquakes of magnitude 6, 10 earthquakes of magnitude 7, and so on. This is now known as the Gutenberg-Richter law; it is a classic example of a simple law underlying what looks at first sight to be a complex system. But what exactly does it mean, and does it have any widespread applications?

First, it's worth stressing just how powerful a law of nature this is. An earthquake of magnitude 8, a little smaller than the famous San Francisco earthquake of 1906, is 20 billion times more energetic than an earthquake of magnitude 1, which corresponds to the kind of tremor you feel indoors when a heavy lorry passes by in the street outside. Yet the same simple law applies across this vast range of energies. Clearly, it is telling us something fundamental about how the world works. Part of that something is related to what we learned about chaos and fractals in

Chapter 3, where power laws also showed up. Indeed, if you choose to, you can take one of the archetypal examples of a fractal in nature, the length of a coastline, and express it in just the same way. A particularly fine example is provided by the coast of Norway,<sup>5</sup> where large fjords branch into smaller fjords, which branch into smaller fjords, and so on. To describe this in terms of the fractal dimension of the coastline, we imagine taking a very detailed map of the coast (or even the actual coast!) and covering it with a grid of identical squares. If the square is big enough, one will be sufficient to cover the whole of Norway. But as you make the squares smaller, you need more and more of them to cover every twisting inlet in the coastline. It's pretty obvious that if you make the squares half as big, you will need more than twice as many to do the job. The rate at which the number of squares you need increases, compared with the rate at which the size of each square decreases, gives you the fractal dimension of the coastline. In this case the fractal dimension turns out to be 1.52. You find this by plotting the logarithm of the length of coast measured at each level (for each size of square) against the logarithm of the size of each square. You get a straight line, and the fractal dimension is the gradient slope of this straight line graph.

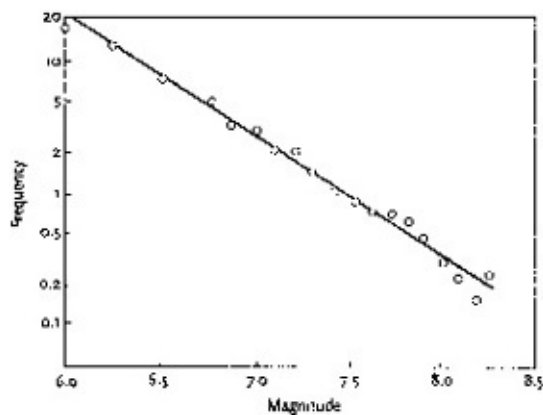




Figure 5.1 *The Gutenberg-Richter law relates the magnitudes of earthquakes to the frequency with which they occur. It is an example of a power law demonstrating  $j$  behaviour.*

Apart from the size of the gradient (the steepness of the line on the graph), this is exactly the same kind of power law as the Gutenberg-Richter law. There seems to be a connection between the fractal geometry of the coast of Norway (and other coastlines) and the frequency with which earthquakes of different sizes occur worldwide (and, indeed, in individual earthquake zones around the globe). Part of that connection lies in a property of fractals that we are already familiar with – they are scale-free. That is why this coastline plot obeys a power law, and what we can infer is that the occurrence of earthquakes is also scale invariant. Interesting though this is in terms of satisfying our curiosity, what it means in practical terms is that there is essentially no difference between a large earthquake and a small earthquake except for their size. You do *not* need to invoke some special, rare and peculiar physical effect in order to explain why large earthquakes occur – they just do. Large earthquakes occur more rarely than small earthquakes, but they are produced by essentially the same physical process as small earthquakes – and the power law tells you this even if you have no idea what that single physical cause of earthquakes is.

Apart from the size of the gradient (the steepness of the line on the graph), this is exactly the same kind of power law as the Gutenberg-Richter law. There seems to be a connection between the fractal geometry of the coast of Norway (and other coastlines) and the frequency with which earthquakes of different sizes occur worldwide (and, indeed, in individual earthquake zones around the globe). Part of

that connection lies in a property of fractals that we are already familiar with – they are scale-free. That is why this coastline plot obeys a power law, and what we can infer is that the occurrence of earthquakes is also scale invariant. Interesting though this is in terms of satisfying our curiosity, what it means in practical terms is that there is essentially no difference between a large earthquake and a small earthquake except for their size. You do *not* need to invoke some special, rare and peculiar physical effect in order to explain why large earthquakes occur – they just do. Large earthquakes occur more rarely than small earthquakes, but they are produced by essentially the same physical process as small earthquakes – and the power law tells you this even if you have no idea what that single physical cause of earthquakes is.

Figure 5.2a *The length of the coastline of Norway can be estimated by covering it in a grid of squares, as described in the text. The smaller the squares, the longer the measured coastline. When this relationship is plotted in a log-log graph, it produces a straight line, following a power law (see p. 142). In this case, the slope of the line (its power) is a measure of the fractal dimension of the coast of Norway, which turns out to be 1.52, almost exactly midway between a line (1) and an area (2).*

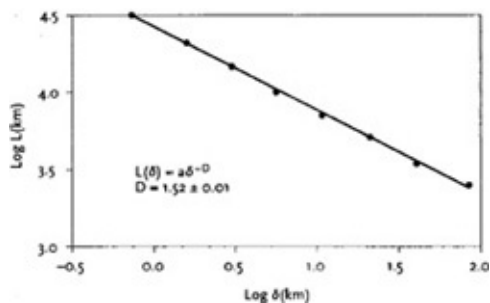


Figure 5.2b (See caption on p. 141.)

This is quite different from the idea, still widely held by both lay people and even some geologists, that large earthquakes require big triggers. That idea would suggest that large earthquakes (like San

Francisco in 1906) occur after a large amount of strain has built up in the Earth's crust, which then 'gives' at a weak point. One comforting implication of this would be that after one big earthquake you would wait a long time before the next one struck in the same place. The statistics, and the power law in particular, say otherwise. Both large and small earthquakes occur at random, but (and this is very important) with different frequencies. An earthquake of any size could happen at any time in an earthquake zone, and just as the possibility of a tossed coin coming down heads is still 1 in 2 even if the last three tosses have all come down tails (assuming the coin is properly balanced), so the chances of another earthquake like the event of 1906 were no less (and no more) in 1907 than they were in 1905. Power laws always mean that the thing being described by the law is scale invariant, so that earthquakes of any size are governed by exactly the same rules.

There's a nice example of scale invariance at work in a simple physical system, which is no less apposite from being familiar from many accounts – Mark Buchanan expresses it particularly neatly in his book *Ubiquity*, where he boldly attempts to apply these ideas to what he calls 'the true science of history'. Imagine an experiment in which frozen potatoes are fired against a solid wall and shattered into pieces. This really has been done, because the way the frozen potatoes break up is very similar to the way in which lumps of rock break up when they collide, and observations of this shattering process can provide insight into, for example, the way lumps of rock in space collide with one another and fragment to form the asteroids that orbit around our Sun in a belt between the orbits of Mars and Jupiter. The fragments of potato come in many different sizes, with lots of tiny pieces, a few large chunks,

and some in between. It is straight forward to allocate these pieces to bins, according to how much they weigh, in the way that Gutenberg and Richter allocated earthquakes into bins according to their magnitude. At first, ignore all the really tiny fragments, keeping them to one side, and bin all the rest. Then you can plot a graph of the number of fragments in each bin against the weight corresponding to that bin. You get a power law. Now you take a look at the tiny fragments of potato you ignored before, and use some kind of finer technology to measure their weights and divide them among bins in the same way, but on a smaller scale. You get the same power law. In experiments carried out at the University of Southern Denmark in the early 1990s, the researchers found that the same power law applied to chunks of potato analysed in this way ranging in size from 10 grams down to one-thousandth of a gram. It means that if you were the size of an ant, and you crawled about among the potato pieces, the kind of landscape you would see would be exactly the same (statistically) as if you were the size of a ladybird crawling about among the debris. It would not be like comparing, say, the rolling hills of Sussex with the spiky peaks of the Himalayan mountains. The 'landscape' provided by the potato pieces looks the same on all scales.

It happens that the landscape of the Moon is also scale invariant, in terms of the numbers of craters of different sizes that scar its surface. Since these craters were produced by the impact of different-sized lumps of rock from space, and those lumps of rock were themselves almost certainly produced by collisions between larger asteroids that break up in accordance with the same laws that govern the break-up of frozen potatoes, this is no real surprise; but it extends the scale of this

particular power law up from potatoes to planetoids.

There's another way of describing all of these kinds of variations (really, as we have seen, different versions of the *same* kind of variation). Large 'events' (earthquakes, stretches of coastline, pieces of potato) are more rare, and you can express this by saying that the frequency of an event is equal to 1 divided by some power of its size. Turning this around, you can say that the size of an event is proportional to 1 over some power of its frequency,  $f$ ; since the exact power is not particularly important, in general this is called '1 over  $f$  noise', and written as  $f^{-1}$  noise. It may seem a strange way of talking about earthquakes, but the reason for doing so is that  $f^{-1}$  noise turns up in naturally varying systems of many kinds, and had already been studied mathematically before people like earthquake experts got in on the act. For our purposes, the terms 'power law behaviour' and ' $f^{-1}$  noise' are synonymous.

This kind of noise contains variations on all scales, ranging from rapid flickering up to slow pulsation, all superimposed on one another. As someone who trained in astronomy, to me the classic example of  $f^{-1}$  noise is the variation in the light from a quasar, which covers the range of fluctuations from flickering on a scale of minutes up to pulsations over periods of years and decades;<sup>6</sup> the same kind of pattern is seen in the light from some stars. When the flickering of such an object is plotted out as a graph, with brightness plotted against the time (astronomers call this a 'light curve'), you get a jagged line looking like the profile of a mountain range, with ups and downs occurring on all scales. This is  $f^{-1}$  noise. At one extreme, the direct contrast with  $f^{-1}$  noise would be what is called white noise, which is completely random, and at the other extreme the contrast would be with a pure signal containing

just one frequency, like a single musical note. As the names suggest, these phenomena were first discussed in acoustics, where the signals being investigated are sound waves; white noise is the boring hiss of static you get from an AM radio tuned away from any station, a single frequency 'noise' would be a pure (but equally boring) tone at one note, and  $\frac{1}{f}$  noise (also sometimes referred to as pink noise) sounds interesting to the human ear. The sound of music has a  $\frac{1}{f}$  structure, as has the sound of the spoken word (including the hubbub of noise at a party). In a nutshell,  $\frac{1}{f}$  noise contains information.

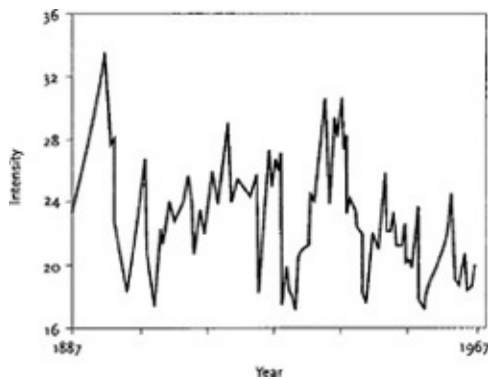


Figure 5.3 Changes in the brightness of light from a quasar, observed over the interval from 1887 to 1967, provide an example of  $\frac{1}{f}$  noise.

But a few words of caution. Whenever new discoveries are made in science there is a bandwagon effect, with people trying to explain everything in terms of their new enthusiasm. One over  $f$  noise has become a hot topic in science, and its enthusiasts rightly use it to probe everything they can, which will either give new insights or reveal the limitations of this field (or both). But  $\frac{1}{f}$  noise is clearly not the only effect at work in all time-varying phenomena that lie between white noise and a single frequency. Take the rise in global average

temperatures that has taken place since the middle of the nineteenth century, for example. This looks very like the kind of light curve of interest in quasar astronomy, with an overall rising trend but many jagged peaks and valleys. Analysis shows that these fluctuations are, indeed,  $\frac{1}{f}$  noise. It could be that the long-term rising trend is just part of his overall pattern, a very long-term fluctuation of the noise. But in this case the long-term trend has occurred exactly at the time when human activities have been releasing gases that trap heat into the atmosphere, and when the short-term fluctuations are averaged out, the size of this rise, over one and a half centuries, exactly matches the warming that would be expected from the amount of these greenhouse gases added to the atmosphere. To deny the link between cause and effect would make about as much sense as to deny that an earthquake that happened in the Nevada desert just after a nuclear bomb was exploded beneath the surface was caused by the bomb, and to claim that it was just part of the background  $\frac{1}{f}$  noise.

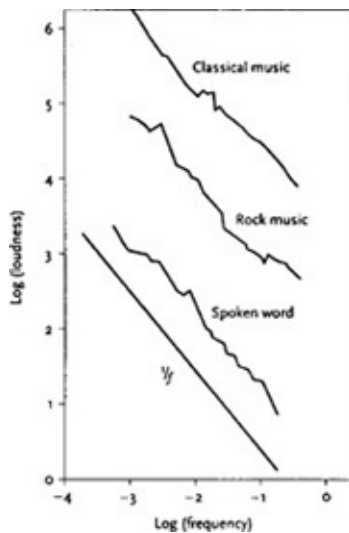


Figure 5.4 Using the standard log-log graph, classical music, rock and the spoken word all display the characteristics of  $\frac{1}{f}$  noise, because they all contain information.

The conclusion is that although there is a lot of noise in the weather system, and this is largely  $\frac{1}{f}$  noise (which is an interesting and important discovery in itself), there is also a genuine upward trend in temperature linked to human activities. The particularly interesting thing about the fact that the noise in the weather system is  $\frac{1}{f}$  noise is, of course, the same as the interesting thing about earthquakes obeying a power law. It means that events of any size can happen at any time, but that big events (in this case, big departures from average temperature) are rare. Traditionally, insurance companies and civic planners have worked out the odds against extreme events on an historical basis. They look at how often a particular type and size of event (say, a severe drought) has occurred in the past, and guess that it will happen with the same frequency in future. They even talk about things like a ‘once in a hundred years’ drought. But this is only half the story. The  $\frac{1}{f}$  signal tells us that such events actually obey a power law. And that means that if you get a ‘once in a hundred years’ drought one summer, the odds of getting a second drought of the same severity the following year are exactly the same as if there had not been a drought that year. It is *not* safe to assume that after one ‘once in a hundred years’ drought it will be a hundred years before the next one happens. Equally, though,  $\frac{1}{f}$  noise implies that a record-breaking drought could be followed by a record-breaking freeze. Next year might be the coldest ever recorded, without this in any way invalidating the greenhouse effect hypothesis, thanks to a random fluctuation of the natural weather system. That is why it is so important to look at long-term trends, not just one-off events.

There are many other examples of power laws and  $\frac{1}{f}$  noise in nature, and no need to catalogue them all here. But there are two that we would



like to mention, before we get on to our main theme of the relevance of all this to the emergence of life, since they strike much closer to home than things like the light curves of quasars and the statistics of long-term temperature trends. Most of the readers of this book will have lived in a large town or city at some time in their life, even if they do not now. Way back in the 1940s, George Zipf of Harvard University studied the populations of cities around the world. Pretty obviously, there are a few cities with very large populations, and many more cities with smaller populations. Rather less obviously at the time (although you have probably guessed what is coming), when the number of cities with each population is binned and plotted against the population on a log-log graph, you get a straight line. Both for the world as a whole and for different regions of the globe, the way people congregate in cities obeys a power law – and the same pattern is as true today as it was in the 1940s. We all make our decisions about where to live independently, exercising free will; yet somehow we are subject to the same laws that govern the flickering light of quasars and the occurrence of earthquakes. Clearly, we are on the trail of a deep truth; and we can begin to understand what is going on by looking at simpler systems, starting with one that also involves people.

Although many people live in cities, not all cities are affected by earthquakes, which remain an abstract phenomenon for most of us; but everyone who lives in a city has experienced a traffic jam, and traffic jams, it turns out, also obey a power law. The kind of traffic jams we are talking about are the ones that occur on the open highway, on unobstructed dual-carriageway roads, seemingly for no reason at all. Any driver (indeed, any passenger who isn't asleep) knows that when the

traffic is light it flows along smoothly, but that once the amount of traffic on the road gets too great, traffic jams occur even though there seems to be no reason why everybody cannot move along smoothly at, say, fifty miles an hour. The problem is that if one vehicle brakes for any reason (perhaps because the driver has simply got too close to the vehicle in front), there is a knock-on effect as everybody behind slows, and it is (fortunately for road safety) easier to slow down than it is to speed up. This kind of behaviour has been studied by researchers at the University of Duisberg, who, in the early 1990s, developed a computer model to put it on a mathematical footing.<sup>7</sup> The model doesn't purport to describe the real world of multilane highways, but (like all good models) is sufficient to give an insight into what is going on. In the model there is just one line of vehicles moving down a single-track road, with no overtaking. All vehicles ('cars') are the same size, and the speed of a car is specified in terms of how many car lengths it can move between one step in the computer simulation and the next. If a car is moving so fast that it will hit the car in front it has to slow down; if there is a large gap, it will speed up (up to a specified speed limit) to close on the car in front. Finally, the ability to slow down and speed up is quantified, with it being easier to brake than to accelerate. All the model has is a line of cars in a row travelling in a straight line and obeying these rules. Yet it shows all the behaviour of real traffic jams, with freely flowing cars when traffic is light, some traffic jams when traffic is heavy, and more or less gridlock once a certain density of vehicles is reached. What's more, both in the models and in studies of real traffic flows we find a power law – the number of jams of different sizes (measured in terms of the number of vehicles involved) obeys the

same kind of law as the number of earthquakes of different sizes. Just like real traffic jams, the jams in the model move backwards through the traffic, as more cars brake behind the obstruction while those in front speed away, and just like real traffic jams they ‘dissolve’ as the average speed of the traffic is slowed by the jam. As well as there being big jams and small jams, there are jams within jams, in a fractal pattern. It is all a clearcut example of  $\frac{1}{2}$  noise at work.

There are two practical lessons to be drawn. First, you don’t need an unusually large ‘trigger’ (such as a crash) to cause a large jam;<sup>8</sup> any size jam can be triggered by even the smallest disturbance, such as one car getting too close to the car in front and braking sharply. Nobody ever hits anybody, there is no obvious cause visible to most of the people involved (and the car that did cause the jam has probably sped off into the distance, its driver unaware of what is happening behind), but hundreds of people may be inconvenienced. Secondly, if the density of traffic increases, it can be kept flowing more freely by limiting its top speed, because this reduces the impact of the difference between the time taken to accelerate and the time taken to brake. It really is true that if everyone obeys the speed limits on crowded motorways they will all get to their destination sooner than if they try to drive faster.

There’s another area of human activities where power laws and  $\frac{1}{2}$  noise may be relevant, which has received a great deal of attention recently. This is economics, and in particular the economics of the stock market. We don’t wish to go into this in any detail, because it would take us too far from our main theme of the emergence of life and its place in the Universe.<sup>9</sup> But the link is certainly worth a mention. In the 1960s, shortly before he turned his attention to fractals, Benoit

Mandelbrot looked at the way the prices of commodities such as steel and cotton had varied on the New York stock exchange, and found that the price fluctuations follow a power law, like  $\frac{1}{2}$  noise. The implication is that the economy is a system that follows the same laws as the occurrence of earthquakes (or traffic jams), and that large events (in this case, things like stock market crashes, such as the crash of October 1987) can happen out of the blue, as a consequence of small triggers. This was not a discovery welcomed by economists, who like to think that economies can be controlled by government action, for example by adjusting interest rates. But if stock market fluctuations really do follow a power law, a small adjustment of interest rates could (if only rarely) be the trigger for a large swing in the market, or might have no effect at all.

The point to bear in mind, though, is the same one that applies to earthquakes, traffic jams and the like. Overall the pattern may be  $\frac{1}{2}$  noise, but sometimes human intervention on a large enough scale will have an effect – there is little doubt, for example, that the Second World War provided a boost to the US economy that pulled it out of depression. The question is, where between a  $\frac{1}{4}$  per cent adjustment in interest rates and a world war does government intervention become important?

Part of the answer to that question probably lies in work largely stimulated by the Irish-born Brian Arthur, working first (in the late 1970s) at the International Institute for Applied Systems Analysis (IIASA), a think-tank based in Austria, and later (in the 1980s) at Stanford University. Classical economics is based on principles very similar to those of classical thermodynamics, and deals with systems (economies) that are close to equilibrium. This involves a concept known as ‘diminishing returns’, which is a form of negative feedback. Crudely

speaking, it says that although you might make a fortune by inventing something new and selling it, once everyone has it it becomes harder to sell any more and your profits decline. Arthur, who had a background in engineering, realized that the opposite could be the case – that you could have positive feedback, or increasing returns. Once you corner the market, everybody has to buy your product, and you make huge piles of money – the obvious example today is provided by Bill Gates and Microsoft. And Arthur realized that you don't even have to have the best product. Apple computers, it is widely acknowledged, are technically superior to PCs based on the Microsoft architecture (this was certainly true in the 1980s). But they were not marketed so effectively at first, and if something becomes an industry standard, the 'me too' effect ensures that it generates huge sales, and huge profits.

All of this makes perfect sense to an engineer or a physicist, but classically trained economists are still fighting a desperate rearguard action against such heretical ideas. We are not concerned with that debate now, but in developing the analogy between the way classical thermodynamics differs from non-equilibrium thermodynamics, and the way classical economics differs from what we might as well call non-equilibrium economics. Modern economists like Arthur are dealing with dynamic, changing systems, in which positive feedbacks are involved and through which a form of energy (in this case, money) flows. Put like that, the story is familiar. Thanks to feedback in particular, economies are actually self-organizing systems on the edge of chaos, with all that that implies. The situation is just rather more complicated, and it is harder to see the wood for the trees, because we are on the inside of the economy looking out, and human beings are an integral part of the

system they are trying to analyse. Nevertheless, even at the simplest level there is no room to doubt that stock market fluctuations behave as  $\frac{1}{2}$  noise, and anyone who believes that governments are in control of economies is living in cloud-cuckoo-land.

But let's get back to simpler things – things like the death of the dinosaurs, and the meaning of life. The dinosaurs died out<sup>10</sup> about 65 million years ago, as part of a mass extinction of life on Earth so clearcut in the geological record that it is used to mark the end of one period of geological time, the Cretaceous, and the beginning of another, the Tertiary. Because the initial 'C' has already been used in a similar context in connection with the Cambrian period, this marker is usually referred to as the K-T boundary, from *Kreide*, the German name for the period. It wasn't just the dinosaurs that were affected, although they are the ones that feature most strongly in popular accounts of the disaster. About 70 per cent of all species alive on Earth at the end of the Cretaceous had disappeared at the beginning of the Tertiary, making this truly a 'mass' extinction and explaining why geologists and palaeontologists use the K-T boundary as a major marker in the fossil record. Given the difficulties of interpreting evidence from that long ago, and the slowness with which geological strata accumulate, all we can really say about the speed of the extinction is that it took place in less than a few tens of thousand or a hundred thousand years, not spread over many millions of years; but that counts as an abrupt change in terms of the geological timescale.

The obvious questions this raises are the same as the obvious questions following a big earthquake – why did it happen, could it happen again, and if so, when? In the case of the K–T event, there is a

very good candidate for the trigger which caused the extinction to occur 65 million years (Myr) ago, rather than, say, 60 Myr ago or 55 Myr ago. The remains of a huge-impact crater with just the right age have been discovered beneath what is now the Yucatan Peninsula of Mexico, and 65 Myr-old strata from around the world contain traces of the heavy metal iridium, which is rare in the crust of the Earth but is known to be a component of some kinds of meteorite. The implication is that a largish rock from space, perhaps 10 kilometres across, struck the Earth and caused worldwide devastation. The layer of iridium is so thin that it must have been laid down in less than 10,000 years (maybe much less), consistent with the idea that the whole K-T event was triggered more or less instantaneously by a blow from space.

There is no trouble explaining how this could have happened. The kinetic energy involved in such an impact would be equivalent to the explosion of about a thousand million megatonnes of TNT, throwing debris into space in the form of large lumps which would travel on ballistic trajectories (just like Intercontinental Ballistic Missiles) and re-enter the atmosphere all around the globe, spreading heat worldwide and raising the temperature everywhere. There would be a heating effect of 10 kilowatts for every square metre of the Earth's surface for several hours, graphically described by Jay Melosh of the University of Arizona as 'comparable to that of a domestic oven set at "broil" '.<sup>11</sup> Following that, tiny particles of dust thrown into the upper atmosphere would spread around the globe and, combined with smoke from all the fires started by the 'broiling', block out sunlight, causing plants that depend on photosynthesis to die, and temporarily freezing the planet. There is no need to go into more detail here, although it is worth emphasizing

that the whole scenario has been carefully worked out and set in its geological context, and is not some flaky half-baked idea; that said, all we want to do is to make the point that impacts from space can be bad news for life on Earth.<sup>12</sup> But this is not the whole story. First, there is evidence that on other occasions, notably some 35 Myr ago, the Earth has been struck by comparable impacts from space without there being an extinction on the scale of the K-T event. The same size triggers do not all cause the same size events. Secondly, there is also some evidence that the dinosaurs and other species were already in decline in the last couple of million years of the Cretaceous. Palaeontologists still argue about the interpretation of this evidence, and it is far from clear that this would have been a terminal decline (the dinosaurs had experienced several ups and downs during the 150 million years in which they roamed the Earth), while some say that they were not in decline at all. But it is certainly possible that the Yucatan impact (which definitely did happen 65 Myr ago, at the K-T boundary) may have been a 'last straw' which triggered widespread extinctions among a web of life on Earth which was already under stress for other reasons (such as, perhaps, climate change caused by continental drift). The point is that any single event might be a special case, and doesn't on its own tell you anything much about the underlying cause of similar events, or likelihood of their recurrence, any more than studying a single earthquake tells you much about earthquakes in general and how often they occur. We need to look at the patterns of extinctions throughout geological history in order to get a proper feel for what is going on, and to find out whether the death of the dinosaurs was a special event or just one of those things.

The K-T event is actually just one of five similar catastrophes (as far



as life on Earth at the time was concerned) which are together referred to by geologists as the 'Big Five' – and it is by no means the biggest. Each is used as a marker between geological periods, and they have all occurred during the past 600 million years. The reason for this focus on the relatively recent geological past is that it was only around that time that living creatures first developed features, such as shells, which could be easily fossilized, leaving traces that can be recognized in strata studied today. Before that time (in the long geological era known as the Precambrian) life flourished in the oceans for almost four billion years, in the form of soft-bodied, single-celled creatures that left us little to study. But around 600-590 million years ago, at the beginning of the geological period known as the Cambrian (hence, Precambrian; the Cambrian is the first Period of the Palaeozoic Era), there was an explosion of life into different varieties as multicellular forms, and then living creatures with hard bits evolved. Fairly obviously, we know more and more about the changing patterns of life on Earth the closer we come to the present day, and the kind of knowledge we are interested in here, involving large extinctions of life, only begins to become clear after the end of the Precambrian.

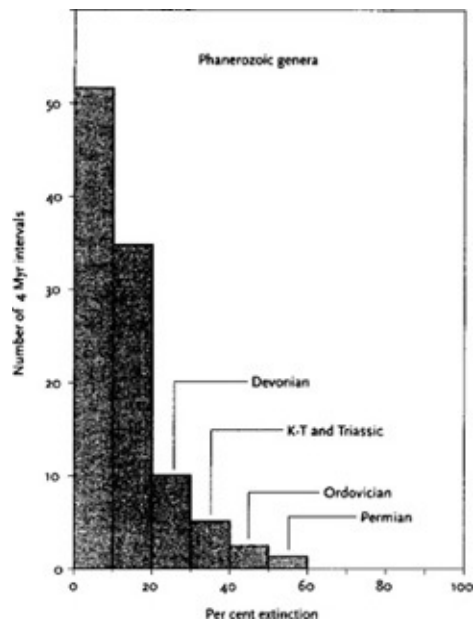
Taking them chronologically, the Big Five extinctions occurred about 440 million years ago (marking the boundary between the Ordovician and Silurian periods), 360 Myr ago (between the Devonian and the Carboniferous), 250 Myr ago (between the Permian and the Triassic), 215 Myr ago (at the Triassic-Jurassic boundary) and 65 Myr ago (at the K-T boundary). There are plenty of other extinctions in the fossil record, and they are also often used as markers in the geological 'calendar' – for example, the boundary between the Jurassic and the Cretaceous, some

145 Myr ago, is also defined by an extinction. But these five are the biggest. The biggest of them all is the event that occurred about 250 Myr ago, at the end of the Permian. It wiped out at least 80 per cent and possibly as much as 95 per cent of all species alive on Earth at the time, both on land and in the oceans, and it did so within a span of less than 10,000 years. Altogether it is estimated that just over a third of all the species that have ever lived on Earth have died out in mass extinctions. But since it is also estimated that 99 per cent of all species that ever lived on Earth have gone extinct, that means twice as many have disappeared in lesser events. The intriguing question is whether the mass extinctions really are special events, different in character from the lesser extinctions, or whether they are the same kind of event writ large – are extinctions of life on Earth, like earthquakes and all those other phenomena we have discussed, scale free? The honest answer is ‘we don’t know’. But there is enough evidence around to suggest that this is a very real possibility.

Most of this evidence is thanks to the painstaking work of Jack Sepkoski of the University of Chicago. He has put together a huge database on extinctions, combining information on thousands and thousands of different species from the published records of other researchers available (to anyone with the patience to carry out the task) in a good scientific library. Of course, there’s a little more to it than that. You have to know the species you are studying well (Sepkoski focused on marine invertebrates), and you have to have the nous to distinguish real extinctions from gaps caused by the difficulty of obtaining fossil remains, while also making allowance for the fact that it is easier to find remains from more recent strata, and so on. After all this work, Sepkoski

was able to plot a graph showing how the level of extinctions occurring in each four-million-year-long interval has fluctuated over the past 600 million years. The overall fluctuations of the graph look fairly random, with long intervals of relatively little variation, just a few species going extinct from time to time, punctuated by dramatic events in which large numbers of extinctions occur – the mass extinctions we have already met.

Figure 5.5 Jack Sepkoski's data plotted as a histogram to show the number of intervals, each four million years long, during which the indicated amount of extinction occurred. Just over 90 per cent of the extinctions occur in just under 50 per cent of the intervals. This is not a uniform distribution.



The percentages on the plot shown here, taken from a paper by Sepkoski published in 1993, do not precisely match the percentages mentioned earlier because they refer to families, not species, and to only some families (the marine invertebrates), not all life on Earth; but the match is still pretty good, and shows, for example, that the ‘death of the dinosaurs’ was also the ‘death of the marine invertebrates’. The question is, what kind of randomness is this, if it really is random? It turns out to

be a power law – our old friend,  $\frac{1}{f}$  noise. One of the demonstrations of this was made by David Raup, who also worked at the University of Chicago, and ‘binned’ Sepkoski’s data in the usual way, adding up the number of 4-Myr intervals in which up to 10 per cent of the relevant genera went extinct, the number of 4-Myr intervals in which between 10 and 20 per cent of the genera went extinct, and so on. The resulting histogram looks just like the equivalent histogram for earthquakes. Raup has also analysed the evidence in the form of a histogram in which the life span of a genus revealed in the fossil record is compared with the number of genera having that life span. Again, the pattern obeys a power law.<sup>13</sup> Exactly the same pattern has since been found in a much larger fossil database, analysed by Michael Boulter of the University of East London, and his colleagues.

Now, it doesn’t seem likely that all extinctions of life on Earth have been caused by impacts from space. What the fossil record seems to be telling us is that extinctions happen on all scales, all the time, and that (like earthquakes) an extinction of any size can happen at any time. Some extinctions might be triggered by meteorite impacts; some might be triggered by Ice Ages. But the other lesson from what we know about power laws and  $\frac{1}{f}$  noise is that you do not need a large trigger in order to set off a large event. An extinction of any size might be set off by a trigger of any size. What matters is that we are dealing with a complex system – life on Earth – which is self-organizing, feeding off a flow of energy, and exists at the edge of chaos. We are almost ready to focus down on what all this means for life itself. But there is one last step that will prove an important aid to understanding the emergence and evolution of life. In the best scientific tradition, having found a universal

law – the power law – that applies to a wide category of complicated (and complex) situations, we want to strip away everything except the essentials, and look deep to find a simple model which goes to the heart of the story and can be used to provide insight into the mystery of complex systems that are built on this deep simplicity. Just as insight into the complexities of gravity starts with understanding the simplest gravitating system, a single object in orbit around another object, before moving upward to provide insight into the workings of the Solar System and the way galaxies move in the Universe at large, so, it turns out, insight into the complexities of the web of life can start with studies of a much simpler system – a pile of sand sitting on a table.

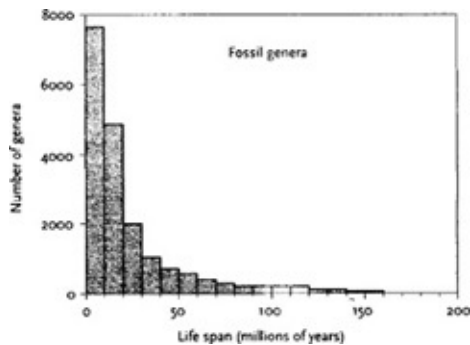


Figure 5.6a David Raup's presentation of the number of genera which survive for a particular lifetime, plotted as a histogram and as a log-log graph (see p. 157). Lifetimes of genera obey a power law, with power close to 2.

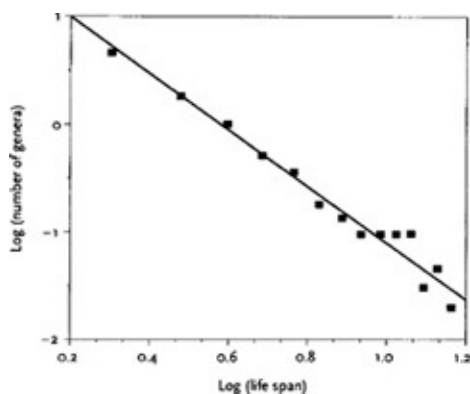


Figure 5.6b (See caption on p. 156).

In the second half of the 1980s, the Danish-born physicist Per Bak, then working at the Brookhaven National Laboratory on Long Island, New York, became interested in the complex behaviour of systems on the edge of chaos.<sup>14</sup> He came at this interest through his work in fundamental physics, the details of which need not concern us here, but soon appreciated that he was dealing with a phenomenon of very broad, if not universal, importance. Bak realized that power laws and  $1/f$  noise are always associated with large systems that are made up of many component parts<sup>15</sup> – what we have been calling complex systems. He also appreciated, of course, the importance of the systems being open, with energy supplied from outside, if they were to arrive at the edge of chaos, a state that came to be known as self-organized criticality. As Bak's interest grew, he began to look for such systems outside the world of fundamental physics, and became aware of things like the Gutenberg-Richter law for earthquakes and Zipf's work on city sizes. The instinctive response of a physicist to the realization that a general law is at work is to find a simple model to describe what is going on, and, working with two younger colleagues (Chao Tang and Kurt Wiesenfeld), Bak came up with the now-famous sandpile model.

At first, this was just a conceptual model, which could be described in words and modelled on a computer, but did not involve real piles of sand – that came later. The key to the model, and to the insight it provides into the behaviour of complex systems, lies in the difference between a pile of sand sitting quietly on a table in equilibrium (just about as boring as the equilibrium systems of classical thermodynamics)

and the same pile of sand, sitting on the same table, on to which new grains of sand are being dropped, one at a time.<sup>16</sup> Our everyday experience, and memories of playing on the beach, tell us in outline what will happen. The sand will pile up in a heap, until the slope of the heap reaches some critical value. Then, adding more sand will cause a landslide, or a series of landslides, making the sandpile slump down. Adding more sand makes the process repeat, until sand covers the entire table and dribbles over the edge when landslides occur. In this state, on average the amount of sand in the pile stays the same, with the same amount falling off the edge of the table as is being added from above. The system is in a state of self-organized criticality, feeding off a flow of energy carried by the new sand grains being dropped on to the pile. And, just as in the way real earthquakes of any size can be triggered by stimuli of the same size, adding a single grain of sand may cause one large avalanche, or a series of small avalanches, or just leave the new grain of sand delicately balanced on the pile. But the pile always remains close to critical.

Bak and Tang built on this idea to provide a clue to the workings of earthquakes and the relevance of the power law. They developed a computer model simulating the behaviour of an earthquake zone like the San Andreas Fault, where two blocks of the Earth's crust are sliding past one another in a stop/go fashion. As we have mentioned, the traditional model of earthquakes said that when the fault sticks, strain builds up until a lot of energy is released in one go, essentially resetting the strain to zero. But in the model developed by Bak and Tang, just as in the sandpile, the strain builds up and the rocks deform until they are on the brink of slipping, in a self-organized critical state. Then, the earthquake

zone undergoes repeated slips on all scales – slips which specifically do not ‘release all the energy’ in one go but which release just enough to maintain the fault in the critical state. Unlike the traditional model, this explains why earthquakes obey a power law.

The computer models of ‘sandpiles’ developed by Bak and his colleagues specified things like the roundness of the grains being added to the pile, their stickiness, and so on. They looked at what happens if grains are added at random over the surface of the pile, and at what happens if they are always dropped from the same place. But you don’t need a computer (or a sandpile) to play the sandpile game and get insight into one of the most fundamental laws affecting non-equilibrium systems, including living systems. In his book *How Nature Works*, Bak describes how anybody can get an insight into the sandpile model by using ordinary child’s building blocks.<sup>17</sup> Using a grid of squares (a chess board would be ideal) you can place blocks on the squares to a maximum height of, say, three blocks, at random, so that on some squares there are no blocks, on some one, on others two, and on some three (draughts pieces work just as well, if you can get hold of enough of them, or coins). We choose an arbitrary rule which says that once a stack reaches four blocks high it goes critical, and all of the four blocks (corresponding to grains of sand) in that stack are taken off, with one of them added to each of the four squares at the sides of the critical square, or allowed to ‘fall off’ the side of the board as appropriate. If this makes one of those squares go critical, repeat the process. Now, add blocks one at a time, at random, to the grid (you can choose the square to add a block to by rolling dice, or using a random number generator on a computer<sup>18</sup>). At each step, move the blocks in accordance with the rules.



Watch how the system moves towards the critical state, with ‘avalanches’ on all scales and ‘sand’ being pushed off the edge of the grid as new material is added from above. In the self-organized critical state, both in the computer models and this ‘model of a model’ the avalanches obey a power law.

Or, of course, you could try something similar with a real pile of sand. The only snag is that ordinary sand doesn’t behave in quite the same way as the simple model, because the sand is not very sticky and each grain has a lot of inertia, so the avalanches are hard to stop. It still obeys a power law, but the fine scale of smaller avalanches is harder to pick out. It turns out that a much better approximation to the ‘ideal’ sandpile of the computer modellers is provided by heaps of long-grain rice, because the long grains of the rice provide more friction, enabling them to stick in interesting configurations for longer rather than slumping at once in an all-or-nothing avalanche. The experiments using rice were first carried out by Vidar Frette and colleagues at the University of Oslo in Norway. To make things even simpler, the long-grain rice used in the experiments was fed into a gap between two vertical panes of glass (like a double-glazed window), making the experiment essentially two-dimensional, just one rice grain thick (another example of the physicist’s urge to keep things simple), and allowing the behaviour of the rice to be recorded on video as grains were dropped into the gap one at a time. The sizes of rice avalanches occurring in the system at the critical state, analysed the usual way, follow the now-anticipated power law. Even better, analysis of a snapshot of the pattern made by the edge of the rice pile whenever it is stuck between avalanches shows that, like the coastline of Norway, it is

fractal. But one of the most important insights provided by the Norwegian team came when they coloured some of the rice grains being dropped in through the slot at the top of the 'window', so that they could follow the motion of individual grains. It turns out that the grains being dropped in do not always, as you might have guessed, slide down the surface and become part of the next wave of avalanches. Instead, any individual grain might burrow deep inside the pile before emerging, much later, in an avalanche. Any grain can stay in the pile for any length of time. Conversely, no grains stay in the pile all the time. Even the most deeply buried grains eventually work their way to the surface and get swept up in avalanches. Just how and why this happens is still not fully understood, but it is clearly telling us that even in this extremely simple example of self-organized criticality, stripped to its bare bones, the whole system has an important influence on every component part, and every component part has an important influence on the whole system. There are no components that sit around and, essentially, do nothing.

The idea of colouring in certain 'grains' also proved fruitful in the computer version of the sandpile game, which has retained its name even though, with hindsight, it could more accurately be dubbed the rice-grain game. Where the insights in the rice game came from looking at the pile of rice from the side, this new insight from the sandpile game came from looking at the computer simulation of the sandpile from above – one of the advantages of a computer simulation, of course, is that you can 'look' at it from any perspective, and even take a cross-section through it if you choose. With this point of view, the Brookhaven team then instructed the computer to colour-code the grains of sand visible on the surface of the pile, with green for grains that lie on flat or

gently sloping regions where there is little or no likelihood of an avalanche, and red for regions that lie on steep slopes, close to the critical point where an avalanche is triggered (on our chess-board version of the game, this would be equivalent to colouring all the stacks containing three blocks red, and all the others green). The experiment (all inside the computer, of course) starts out with a scattering of sand grains, all coloured green, spread across the table. As the pile grows, eventually a few red spots appear. As it grows further, towards the critical state, the red spots grow and merge into threads that run like a network across the surface of the sandpile. When there are only a few isolated red spots, even if a new grain falling on the pile lands in one of those spots, it can only have a local effect, causing a small-scale, local rearrangement of the sand grains in the pile. But when the network is dense enough, a new grain falling on one red spot can cause a disturbance which affects other red spots nearby, in a domino effect which might trigger a small rearrangement of the sand or a series of avalanches involving most of the surface of the sandpile. What matters is the density of the network of critical points in the system.<sup>19</sup> But there is something else of equal importance. Like earthquake zones, the system does not reset to zero, no matter how complicated the network. In the self-organized critical state, the fall of a single extra grain of sand on to the network may trigger a large-scale rearrangement of the sand; but after this rearrangement has occurred there will still be a complex network of critical states, it will just be arranged in a different pattern. *Plus ça change, plus c'est la même chose.*

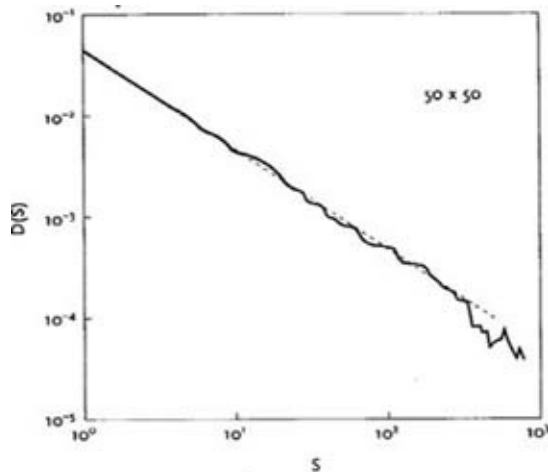


Figure 5.7 *The number of avalanches of different sizes occurring in the sandpile model (see p. 160) also obeys a power law, producing a straight line on a log-log graph.*

We are ready for the next step, and the first real insight into the emergence of life. Having stripped the study of complexity down to its bare essentials, and found the same deep truth underpinning things as diverse as earthquakes, the stock market and the movement of human populations, we discover that it is all built on networks, interconnections between the simple parts that make up a complex system. The importance of such networks in general, and their specific importance for the emergence of life, has been investigated by Stuart Kauffman, of the Santa Fe Institute in New Mexico.

Kauffman has a striking analogy which brings out both what we mean by complexity and the importance of networks in emerging complex systems. He asks us to imagine a large number of ordinary buttons, perhaps 10,000 of them, spread out across a floor. Individual buttons can be connected in pairs by thread, but at the start of the thought-experiment there are no such connections. This is definitely not a complicated system in the everyday use of the term. But the way you connect the buttons turns it into a complex system. You start with just

the buttons on the ground and a supply of thread, but with no button connected to any other button. Choose a pair of buttons at random and tie them together with a single thread, leaving them in their original positions on the floor. Repeat the process a few times, and if you happen to choose a button that is already connected to another button, don't worry about it; just use the thread to connect it to another button as well. After you have done this a few times, there will be a small amount of structure in the collection of buttons. Increasingly, you will find that from time to time you do indeed choose a button that is already connected to another button; sometimes, you will even choose a button that already has two connections, and you will link it to yet a third component of what has become a growing network of connections.

The way to tell how interesting this network has become is to pick up a few buttons, one at a time, and count the number of connections they each have with other buttons. Each connected cluster of buttons is an example of what is known as a component in a network; the buttons are examples of nodes, points that connections are connected to. The number of buttons in the largest cluster (the size of the largest component) is a measure of how complex the system has become. The size of the largest cluster grows slowly at first, more or less in a linear fashion, as the number of threads connecting pairs of buttons is increased; because most buttons don't have many connections, already there is only a small chance that each new connection will add another button or two to the existing largest cluster. But when the number of threads approaches and then exceeds half the number of buttons, the size of the largest cluster increases extremely rapidly (essentially exponentially) as each new thread is added, because with most of the

buttons now in clusters there is a good chance that each new connection will link one smaller cluster (not just an individual button) to the existing largest cluster. Very quickly, a single supercluster forms, a network in which the great majority of the buttons are linked in one component. Then the growth rate tails off, as adding more threads usually just increases the connections between buttons that are already connected and only occasionally ties in one of the few remaining outsiders to the supercluster. Although the network has stopped changing very much as new connections are made, it is by now, without doubt, a complex system.

You can see this sort of effect at work yourself, by actually playing with buttons (fifty or so might be a more sensible number to choose than 10,000), but the pattern of behaviour is brought out very clearly using a simple computer model of such a system. The important point is that as the number of connections in the network exceeds half the number of nodes, it switches very quickly from one rather boring state (a lot of buttons with a few connections between them) to another stable state with a lot more structure, but in which there is little scope for further change. This is a simple example of the phenomenon known as a phase transition, which Kauffman likens to the way in which water undergoes a phase transition as it freezes into ice. In the network, complexity has emerged naturally from a very simple system just by adding more connections, with the interesting activity associated with the changeover occurring all at once, right at the phase transition, where the system is switching from one state to another as a result of a very few final connections being made.

Of course, the connections do not have to be physical pieces of

thread. In the sandpile model, the connections are determined by the laws of physics that are involved—the law of gravity, pulling grains of sand down the pile, the laws of friction, encouraging sand grains to stay put, and so on. What matters is that, as the pile gets steeper, more connections are made between individual grains, because the movement of one grain increasingly affects the stability of its neighbours. In a shallow pile of sand, removing (or adding) one grain doesn't affect the neighbours very much. At the critical angle, removing one grain may represent taking away the piece of the puzzle that, like the keystone in, the arch of a bridge, was preventing many other pieces from moving downwards, and when they slide many other grains are affected, and so on. Making the angle of the sandpile steeper has the same effect, in terms of increasing the complexity of the network, as adding more connections between buttons. In a similar way, as we have seen, an asteroid orbiting between Mars and Jupiter is not just under the individual influence of the Sun, but is part of a complex network of connections, made by gravity, involving (in principle) every other gravitating object in the Universe, but even at a simpler level involving all of the major planets in the Solar System, as well as the Sun. That's what makes its behaviour unpredictable.

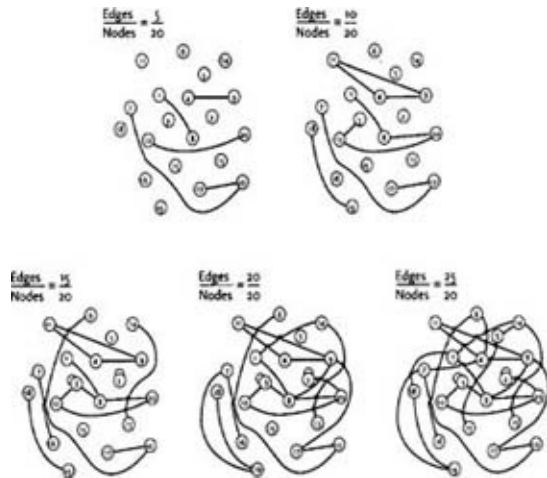


Figure 5.8a Stuart Kauffman's 'button' model of how networks form (see p. 163).

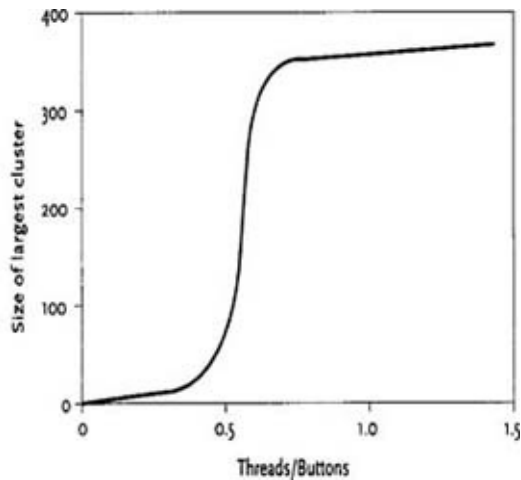


Figure 5.8b As the number of connections increases, there is a sharp transition from the state in which there are many 'loose' buttons and a few connections to the state in which almost every button is part of the network.

The reason Kauffman was interested in networks was because he is interested in the most important emergent phenomenon of them all – how life emerged from non-life. What was it that happened in the chemical 'primordial soup', either on Earth or somewhere else, to make some of those chemicals link up into self-replicating systems? Kauffman (1939–) studied at Dartmouth College in New Hampshire, and then from



1961 to 1963 at the University of Oxford on a Marshall Scholarship. Although he saw himself as a philosopher when he started on this course, in Oxford philosophy was taught alongside a great deal of biology, in order to equip students to address questions such as the relationship between the images of the world that we carry in our heads and the physical way in which our eyes and nervous systems interpret information from the outside world. Increasingly disillusioned by the navel-gazing aspects of philosophy, and intrigued by the biology, Kauffman switched to medicine, studying to be a doctor at the University of California, San Francisco. It was there that he developed his interest in how genes work within the cell and how life itself emerged, officially becoming a theoretical biologist when he moved to the University of Chicago in 1969, after completing his medical studies. But his interest in the fundamentals of the emergence of life lay fallow until the 1980s, simply because he was too busy on other work (first in Chicago, then at the National Institutes of Health, then at the University of Pennsylvania) to pick up the threads. The key work was eventually published in 1986, and its core idea can be summed up very simply in terms of the kind of networks we have been discussing; although he was still on the staff at the University of Pennsylvania, where he was teaching, Kauffman's research base became the Santa Fe Institute, and his ideas were developed in collaboration with colleagues there, very much in an interdisciplinary environment where there was a buzz about ideas such as chaos, complexity and self-organized criticality.

All you have to do is to imagine that in the primordial chemical broth there were some substances that acted as catalysts for the formation of other substances, like the catalytic processes at work in the

BZ reaction. Chemical A catalyses the formation of chemical B. It's hard to see, given the variety of chemical raw materials around,<sup>20</sup> how this could not be the case, and even if the encouragement A gave to the formation of B was only small, it would still increase the concentration of B in the mix. Now suppose that the presence of B encourages the formation of C, C acts as a catalyst for the formation of D, and so on. Somewhere down the line, one of the compounds involved catalyses the formation of A, and you have a self-sustaining loop of interactions, which in effect feeds off the raw material available and makes more of the compounds in the loop, with the aid of energy from sunlight or the heat from volcanic vents. There can also be other interactions, as we saw in our earlier examples of self-sustaining chemical interactions – D might act as a catalyst for A, as well as for E, and there could be inhibitors involved as well, again like the BZ reaction. It is quite easy to see how a network of connections between the chemicals in the broth can arise, an autocatalytic network which sustains itself. This, argues Kauffman, is the way life arose – as a phase transition in a chemical system involving a sufficient number of connections between the nodes (individual chemical compounds) of the network. A crucial, and persuasive, feature of this argument is that like the emergence of a supercluster in the button and thread network, this is an all-or-nothing event. If the network is insufficiently connected, there is no life; but add one or two more connections and life becomes not only possible but inevitable. You do not have to build a long chain of unlikely chemical events one after the other in order for life to emerge, and there are no half-and-half states where you are not quite sure whether the system is dead or alive.

There is a great deal more to the argument, involving details of real

chemical interactions, and the need to consider how the raw materials could have got sufficiently close together for the series of interactions to occur – one possibility is that they were concentrated in a thin layer of material, either a smear of liquid across a surface, or in a layer bound to a clay surface. We should emphasize that these ideas are both speculative and controversial; not everybody is persuaded by the argument, and there are other ideas about how life might have got started. But not everybody is convinced by those other ideas either, and nobody knows for sure just what did occur when life emerged from non-life. Even in Kauffman's model, many details remain to be explored. But we find the overall package of ideas persuasive, not least since it fits the emergence of life into the same set of complex systems based on simple laws that we find so often elsewhere. As Kauffman puts it, 'life crystallizes at a critical molecular diversity because catalytic closure itself crystallizes'.<sup>21</sup> Whatever the *origins* of life, though, ideas involving networks, connections and self-organized criticality provide powerful new insights into the way life operates once it has emerged, which we shall begin to discuss in the next chapter. But before we do so, there is another of Kauffman's speculations that is worth describing. Although it is no less controversial than his ideas on the emergence of life, it offers a mind-blowing example of the extent to which the complexity of living things (the most complex things in the entire Universe) may very well be based on deep simplicity.

Kauffman was – and is – interested in the way cells work, at the level of genes providing the instructions which operate what is sometimes loosely called the machinery of the cell. The instructions are ultimately coded in DNA, the long molecules which make up the genes themselves;

but both the machinery and the structure of the body are actually made of proteins. Things like hair and fingernails, as well as muscle, are forms of protein, and so are things like haemoglobin, which carries oxygen around in your blood, and enzymes, which are essentially biological catalysts that promote chemical reactions important for life. Proteins themselves are large molecules made up of sub-units called amino acids, which is why the discovery that amino acids exist in the kind of interstellar clouds from which the Sun and its family of planets formed is so intriguing. The genetic code in DNA contains instructions for making proteins, and those proteins then carry out the work of life. But there is another step in the process. When a gene is activated (just how and why this happens is beyond the scope of the present discussion), the relevant piece of information is first copied into a piece of very similar molecule called RNA. Then, the machinery of the cell reads the RNA and acts on the instructions to make the appropriate protein. This two-step process is probably telling us something about how life originated, and it is at least possible that RNA was 'invented' before DNA. In Kauffman's scenario, the 'crystallization' of life occurred at the protein level, in a chemical soup rich in amino acids, where the first autocatalytic networks of life emerged; the model easily accommodates the possibility that RNA was involved at an early stage, and that later the evolutionary pressures associated with competition between different autocatalytic networks could have led to the system we see today. But we digress.

The relevant points which underpin Kauffman's investigation of the way cells work are that genes act to control the machinery of the cell, and (crucially) that genes can affect each other, with one gene turning another one on or off. <sup>22</sup> His investigation of the machinery of cells in

terms of networks goes back to his unofficial interest in networks as a medical student, but came to fruition in his work at the Santa Fe Institute in the late 1980s. At the time Kauffman carried out this work, it was thought that there were about 100,000 different genes in human DNA – the human genome. Since then the human genome project has shown that this was an overestimate, and suggests there are only about a third as many genes that specify what it is to be a human being. All of those genes are present in the DNA of every cell in the human body, but they are not all active in every cell in the body. Different kinds of cell are specialized for different tasks, so that liver cells, for example, do quite different things from the cells in muscles. This differentiation of cells into specialized forms occurs during the development of the embryo, and understanding the process of development and cell specialization is one of the most important areas of biological research. But however the process happens, in the fully developed human being there are some 256 different kinds of specialized cell. In each case, only the appropriate bits of DNA (the appropriate genes) ever get ‘switched on’ during the normal course of life, so that a liver cell is only ever a liver cell. But all the rest of the genetic information is still there, as has been dramatically demonstrated in cloning, where the DNA from a specialized cell is transferred into an egg cell, which then develops into a new adult that is a replica of the one the DNA came from.

The network involved in running a cell can be described in terms of a system with one node for each gene, and connections linking the genes like threads linking the buttons in the earlier model. With between 30,000 and 100,000 genes involved, you might guess that the problem of describing the behaviour of such a network would be too complicated

to solve, even using the best modern computers. Kauffman uses an analogy in which the nodes in the network are lightbulbs, wired together at random. In such a system, all of the bulbs might be on, or they might all be off. And in between those two extremes there are a vast number of different states in which different combinations of bulbs are on or off. To be precise, if there are  $N$  such bulbs, then since each one can exist in either of two states (on or off), the number of states for the whole system is  $2^N$ . With one bulb, there are two states; with two, there are four; with three, there are eight; and so on. With just 100 lightbulbs wired together at random, there are  $2^{100}$  different possible states, and  $2^{100}$  is roughly  $10^{30}$ . It is impossible to put such a large number in any sort of human perspective, but in so far as the comparison may help, there have only been about  $4 \times 10^{17}$  seconds since the Big Bang in which the Universe was born, and remember that  $10^{30}$  is not 13 times bigger than  $10^{17}$ , but  $10^{13}$  times bigger (that is, ten thousand billion times bigger).

Undaunted, though, Kauffman and his colleagues investigated the behaviour of such large networks in computer models, with what seems like the added complication of feedback, in which whether or not one particular lightbulb is turned on depends on the state of the lightbulbs it is wired up to. They apply the rules of Boolean logic (or Boolean algebra), <sup>23</sup>so that, for example, if a particular bulb is wired up to two other bulbs, then it might only be on if both the other bulbs are on (an AND operation), or it might only be on if one or other of the other bulbs is alight (an OR operation), and so on. In these investigations they looked at the overall patterns of behaviour for networks consisting of different numbers of bulbs, with different numbers of connections

attached to each node, and with different Boolean operator functions assigned at random to each node. Hardly surprisingly, the investigation took years.

What the team was looking for were stable patterns resulting from simple rules. Any of these systems has a finite number of states, even if the number is ludicrously large. If you start the system in some randomly chosen state, with some lightbulbs on and some off, the lights will flicker and change in accordance with the particular set of Boolean rules that has been applied, and if you are lucky it will settle into a repeating pattern, following its own trail endlessly around the same cycle of states (called, logically enough from the context, a state cycle, but essentially the same as the limit cycles we discussed earlier). The state cycle might be short enough for you to notice the pattern, or it might involve so many changes that you could wait for the lifetime of the Universe and never notice any repetition. Or the system might freeze into a state with some bulbs on and some off, and stay unblinking for ever. The state cycles act as attractors for the system, and in some cases they turn out to be very powerful attractors, so that whatever state you start the system in it very quickly moves towards one of these state cycles. That may be the key to cellular life.

Kauffman and his colleagues found that if there is just one input per bulb, however many nodes you have and whatever rules of Boolean logic you apply, nothing interesting happens. The system either gets into a very short state cycle, or freezes up. If the number of connections to each node is greater than 2, chaos reigns. There are attractors, but the attractors have huge numbers of states, and they are also very sensitive, in the sense that a small change (perhaps switching just one of the

lightbulbs on or off) will shift the system in favour of another attractor – very much like the classic Butterfly Effect. The *only* systems that behave in a way that is both complicated enough to be interesting and stable enough to be understood are those in which each node is connected to exactly two other nodes. Interesting things happen at the edge of chaos, and feedback is an essential ingredient in what makes them interesting. In those systems alone, each state cycle has a length equal to the square root of the number of nodes. Out of all of the  $2^{100}$  possibilities in the system we just described, under these conditions it will settle down into a state cycle which has just 10 steps in it (10 being the square root of 100), and the pattern will be obvious to anyone watching the system. Even in a system with 100,000 nodes, a typical state cycle will be just 317 steps long. Even better, these state cycles are powerful attractors, and if the system is set up in any random state, it quickly moves towards one of these cycles and settles down, without being sensitive to small disturbances. You can set up a system which has  $2^{100,000}$  different states (about  $10^{30,000}$  states), start it in any randomly chosen pattern, and almost literally in the blink of an eye it will settle down into a stable cycle, repeatedly visiting just 317 of those states in the same regular order. <sup>24</sup>

There's more. The same kind of statistics apply to the number of different state cycles that can exist in such a system, although only in a more approximate way. In the kind of systems we have just described there should be a number of different attractors approximately equal to the square root of the number of nodes. With 100,000 different nodes, there will be about 317 different attractors. With 30,000 nodes, there will be about 173 different attractors. There are between 30,000 and



100,000 genes in the human genome, and there are 256 different kinds of cell in the human body. Could it be that each cell type represents a particular state cycle for the human genome, in which specific genes are turned on and others off? To test this possibility, Kauffman has compared the number of genes and the number of cell types in different organisms. Bacteria have only one or two different kinds of cell, yeasts maybe three, the fruit fly 60, and so on. The number of cell types increases in proportion to the square root of the amount of DNA present in the organism, and even though genetic blue prints have not yet been obtained for every species, it seems a reasonable rule of thumb that the number of genes is proportional to the amount of DNA. The number of cell types does increase as the square root of the number of genes! The implication is that the same rules that govern sandpiles and networks of lightbulbs are at work in the cells of our bodies. In round numbers, there are a couple of hundred types of human cell because there are a couple of hundred state cycles which are attractors in a network of tens of thousands of genes interacting with one another in accordance with the rules of Boolean logic. In each type of cell, most of the genes are frozen in the 'off' state, but a minority of genes twinkle away, turning each other on and off in a pattern which runs through a cycle a couple of hundred steps long, during which it operates the chemical machinery of the cell in accordance with the kind of cell it is, before repeating. And that's it – the secret of life at the level of cells.

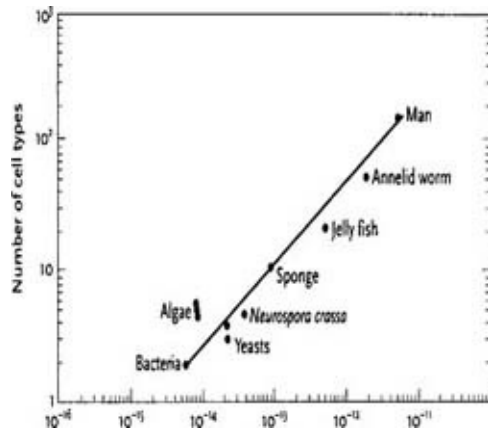


Figure 5.9 A log–log graph comparing the amount of DNA in a cell with the number of different cell types for various organisms. This closely follows a straight line with slope 0.5 (a square-root law, or a power of  $1/2$ ), as predicted by Stuart Kauffman on the assumption that the number of structural and regulatory genes is proportional to the amount of DNA in the cell (see pp. 172–3).

This is still work in progress, and much more needs to be done to turn these provocative and persuasive ideas into a complete theory of cell differentiation and development. But the story so far hangs together well, suggesting strongly that even creatures like us, the most complicated systems in the Universe, are built upon very simple rules. All the apparent complexity of an interacting system of tens of thousands of genes boiling down to a few hundred possible states, thanks to the deep simplicity of the world. As Kauffman says, ‘we are the natural expression of a deeper order’. It is time to move on to see how that deeper order affects the interactions not just between genes and other chemicals, but between different kinds of living organisms in the world at large.

## 6

### The Facts of Life

Evolution is a fact, just as the elliptical shape of the orbit of a planet around the Sun is a fact. There is ample evidence of evolution at work, transforming one species into another, both in the fossil record and from studies of present-day life on Earth.<sup>1</sup> The theory of natural selection, arrived at independently by Charles Darwin and Alfred Russel Wallace in the second half of the nineteenth century, is a model which offers an explanation of why evolution happens, just as the theory of gravity, arrived at by Isaac Newton in the second half of the seventeenth century, is a model which offers an explanation of why planets follow elliptical orbits. Neither theory, or model, is the last word on its subject. Newton's theory was improved on at the beginning of the twentieth century by Albert Einstein, who found a more complete model of how gravity works – the general theory of relativity. Einstein's theory is more complete than Newton's theory because it applies, with great precision, in situations where Newton's theory is an inadequate description of the Universe we live in – most notably under conditions where there is an intense gravitational field. But the fact that Einstein's theory is right does not mean that Newton's theory is wrong. Newtonian gravitational theory is still entirely adequate for describing (or calculating) the fall of an apple from a tree on Earth, say, or the trajectory of a spacecraft travelling

between the Earth and the Moon. It's just that under some circumstances you need something more than Newtonian gravitational theory in order to get the full picture.

By the end of the twentieth century, it was also clear that under some circumstances you need something more than the Darwin–Wallace theory of natural selection in order to get the full picture of how evolution works – most notably when there are many different species interacting in an ecosystem. But this does not mean that the Darwin–Wallace model is wrong, and that model is still entirely adequate for describing some evolutionary situations, especially those which involve only a few species interacting with each other and their physical environment. We shall refer to this as ‘Darwinian evolution’, following a tradition that goes back to Wallace himself, who was the first to acknowledge Darwin as the primary discoverer of the theory, not least because Darwin had taken the crucial step of presenting it to a wide audience (in his book *On the Origin of Species by Means of Natural Selection*, first published in 1859). In order to appreciate the significance of the recent work, which has provided a more complete picture of how evolution works by including the complex web of interactions involved in an ecological network made up of many interacting species, we first need to get a clear idea of what Darwinian evolution is all about.

The essence of Darwinian evolution can be summed up in three steps, forming a chain of logic so straightforward that it is almost a tautology, which led Thomas Henry Huxley, a contemporary of Darwin and a champion of the idea of natural selection, to comment when he first encountered the idea, ‘How extremely stupid not to have thought of that myself.’<sup>2</sup> First, offspring resemble their parents. Characteristics are

inherited, passed on from one generation to the next. Secondly, the process (whatever it is) that copies characteristics from one generation to the next is not perfect, so there are slight (sometimes, more than slight) variations between individuals in each generation – even between siblings. All people have noses, for example, but some people have larger noses than other people do. A large nose may be a family characteristic, but two brothers within that family will not have exactly the same size nose, even if each of them has a nose that is larger than yours. Finally—and this is the crucial point that both Darwin and Wallace seized upon—there are more individuals born in each generation than survive to become adults and reproduce in their turn.<sup>3</sup> But *why* do some individuals survive to reproduce while others do not? The great insight, which led Huxley to his famous comment, is that the individuals that survive are the ones which fit in best to their surroundings. They are the ones who are good at finding food, finding (or attracting) a mate, and avoiding being eaten by something else. This is what is meant by ‘survival of the fittest’. Not necessarily the most physically fit and athletic individuals (although that might be part of the story in some cases), but the ones who fit in best to their environment (including the ‘environment’ provided by the presence of other species), like the fit of a piece in a jigsaw puzzle, or the fit of a key in a lock. Not the ones who *are* fit, the ones who *do* fit.

Now, go back to the beginning of the chain of logic. The individuals who fit in best survive and reproduce. So they pass on their characteristics to the next generation, including those characteristics that made them successful in the first place. As long as the environment stays the same (a very big ‘if’, as we shall see) this process will fine-tune the

members of a species to fit ever more exquisitely into their ecological niches. Indeed, part of the observational evidence that led Darwin to the idea of natural selection (though it took rather a long time for the penny to drop) came from his studies of finches from different islands in the Galapagos archipelago. Although the birds on different islands are all members of the same family, on one island, where the seeds that the birds feed on are tough and hard to crack open, the finches have short, powerful beaks. On another island, where the food that is available has to be probed out of nooks and crannies, their cousins have longer, more slender beaks. Each is well adapted to life on its own island, but if birds from one island were taken to the other island and released they really would struggle to survive.

The struggle, we should emphasize, is not between the finches and other species – not even with the plants on whose seeds they feed. There is a competition for food *between members of the same species*. In this example, if you are a bird that lives on an island where there are only tough seeds to eat, having a stronger beak gives you an advantage over other birds. It means you can get more food, and that will help you to raise a larger brood of chicks, who will inherit your stout beak. Birds with slightly less stout beaks will get correspondingly less food, have less success in raising offspring, and in the next generation, on average, beaks will be stouter than in the previous generation.<sup>4</sup> Similarly, a rabbit who runs away from a fox is not competing with the fox, in evolutionary terms. It is competing with other rabbits. The ones who run fastest, and zigzag most unpredictably, will survive and breed; the ones who don't will get eaten. So, in succeeding generations, rabbits get better and better at running quickly but erratically. The fox, of course, is in

competition with other foxes – the fox that runs fastest, or uses most cunning to catch its prey, will survive and raise more offspring (who inherit those characteristics) than a slow, stupid fox. Getting back to our example of variations in nose size, you can imagine that a change in the environment might place people with some kind of nose at a disadvantage—some kind of pollution might make it advantageous to have a big nose. Babies with small noses might then be more likely to die in infancy, and remember that because of imperfections in the process of copying characteristics from one generation to the next, even siblings can have different sizes of nose. Families with big noses will fit into the new environment better than families with small noses, but even within those families, siblings with bigger noses will do better than their brothers and sisters.<sup>5</sup> Over many generations the average size of the nose in the population will increase.

But we don't need to deal in hypothetical examples alone, because just this kind of effect has been observed, and measured, among the very finch populations that helped to open Darwin's eyes to what was going on. In the early 1970s, a husband and wife team from Princeton University, Peter and Rosemary Grant, began a long programme studying the birds of Daphne Island, in the Galapagos. The island is home to a species of ground finch, *Geospiza fortis*, which uses a strong beak to crack seeds. In 1977, there was a severe drought on the island, in which more than a thousand of the 1,200 or so finches being monitored in the study died. They died because many of the plants they depended on for food died, so there was fierce competition among the birds for food. The finches that survived were all large birds with big beaks that were able to eat seeds that the smaller birds could not crack

open. As the population recovered after the drought, with all the new birds descended from these survivors, the Grants' team measured the beaks of the new population of finches and found that they were, on average, 4 per cent larger than the average before the drought.

This is the essence of Darwinian evolution, and it is very good at explaining relatively stable (at most, slowly evolving) situations involving only a few species, as in the classic example of life on the Galápagos Islands. We can take this line of argument to its logical limit by considering how this kind of natural selection can produce a complete standoff, or balance, between two possible ways of life, resulting from what is called an evolutionarily stable strategy, or ESS.

The insight comes from the application of the mathematical rules of games theory to evolutionary biology. In spite of its cosy name, and although it can be applied to games such as chess or draughts, 'games theory' has received a great deal of attention, particularly during the Cold War, because of its relevance to war 'games', hypothetical scenarios in which the implications of different military and political strategies are assessed. It turns out that these techniques are relevant to evolutionary races as well as to arms races; one of the leading proponents of this approach to the study of evolution is John Maynard Smith of the University of Sussex, from whom the following example comes.

He asks us to imagine a population of a single species of an animal in which each member of the population acts in one of two characteristic ways. Some are 'hawks', which means that they behave aggressively in disputes with other members of their species over food, and some are 'doves', which means that they behave more submissively in such



arguments. Real doves, as Maynard Smith acknowledges, are really rather aggressive birds, but since they are often used as symbols of peace the choice of name makes sense in this context. When a hawk finds a piece of food, and another member of the species is present, it will always fight for the food. If the opponent fights back, one of them will eventually give in, short of it being a fight to the death, leaving the food for the victor. When a dove finds a piece of food and another member of the species is present, it will never start a fight. If it is attacked, it will run away at once; if it is not attacked, it will make a threatening display, but when two birds are making such displays at each other, one will eventually give up and withdraw more or less gracefully from the confrontation. The whole point of games theory, though, is to put numbers into such scenarios, so that different strategies can be compared in a qualitative way, and for the sake of an example we can say that the value of the food to the animal is 50 units. If it eats the food, it scores 50 points. If it runs away, it gets nothing, but it costs nothing to run away. If it fights for the food, it may be injured, and lose 100 points, or it may win and gain the 50-point prize; if it makes a threat display before running away, or before eating the food, that costs it 10 points.

We can study the implications by starting out with a population consisting entirely of doves. In any confrontation, each individual loses 10 points for making its display, but one of them gains 50 points (40 points overall) by eating the food. On average, everybody gets 15 points in each confrontation (40, minus 10, divided by 2). Nobody gets hurt, everybody gets to eat. As near to Utopia as this model permits.

But now imagine that because of a mutation a single hawk arises in the population. In every confrontation, the hawk gets the food—50 points

every time, at no cost at all. As long as there are only a few hawks, most other encounters are between dove and dove, with an average reward of a measly 15 points. Clearly, the hawk will thrive and produce lots of offspring who share its hawkish tendencies.

What happens, though, if everybody is a hawk? Absolute disaster. Every confrontation leads to a fight, the winner gains 50 points while the loser loses 100 points, and the average is a lousy -25. Such a population would soon go extinct unless there was plenty of food around that could be picked up without a fight. But imagine that in such a situation a single mutant dove arises. It never gets in a fight, so in any 'conflict' it scores zero, but it too gets the free food that is available. The dove will do better than the hawks, and dove characteristics will spread among the population.

Neither extreme is stable, and from either end there is an evolution towards the middle. The point of balance is actually reached, for this particular choice of numbers, when there are five doves for every seven hawks. In that situation, it turns out that each individual gets 6.25 points per conflict, on average. The first important point to notice is that there is a stable situation – an ESS – and that evolution takes the population there. Equally significantly, though, this is not the best of all possible worlds. An average score of 6.25 is markedly less than the average score of 15 in the all dove society. The *stable* scenario is not necessarily the *best* scenario, the one that would be chosen by a logical collection of individuals acting in their own mutual best interests. And there is a more subtle but intriguing point to note – the situation is the same if there are five doves for every seven hawks in the population, or if each individual member of the population acts like a dove five-twelfths of the time and

like a hawk seven-twelfths of the time, choosing which strategy to employ in each encounter at random, without planning it in advance.

The ESS approach has proved useful in evolutionary biology in providing insights into the real patterns of behaviour seen in some populations. But, as its very name suggests, it deals with stable situations, in which there is essentially no change, the evolutionary equivalent of classical thermodynamics. These situations involve a small number of interacting components (in this case, the hawk strategy is one component, the dove strategy is the other component, and they are connected by a feedback involving one link). There are slightly more complex examples of ESS at work, but they are all like the simple networks of a few lightbulbs, linked in accordance with the rules of Boolean logic, which freeze into a single state, or cycle around a very limited and boring number of states. But there is a much more interesting possibility, in which an ecosystem that seems to be in a stable state is actually constantly changing, being kept in balance only because both (or all) of the components are evolving as fast as they can in order to keep up with the others. This has very clear analogies with the nuclear arms race of the Cold War, which we do not intend to dwell on here; in evolutionary biology, it is usually known as the 'Red Queen effect', from the character in Lewis Carroll's *Through the Looking-Glass*, who has to run as fast as she can in order to stay in the same place.<sup>6</sup>

The best way to see what the Red Queen hypothesis is all about is to consider an example – Kauffman is particularly fond of describing this in terms of an imaginary species of frog that feeds on an imaginary species of fly, and we will follow his example. There are lots of ways in which the frogs, who want to eat flies, and the flies, who want to avoid being

eaten, interact. Frogs might evolve longer tongues, for fly-catching purposes; flies might evolve faster flight, to escape. Flies might evolve an unpleasant taste, or even excrete poisons that damage the frogs, and so on. We'll pick one (hypothetical) possibility. If a frog has a particularly sticky tongue, it will find it easier to catch flies. But if flies have particularly slippery bodies, they will find it easier to escape, even if the tongue touches them. Imagine a stable situation in which a certain number of frogs live on a pond and eat a certain proportion of the flies around them each year. This is a kind of equilibrium, but it is unstable, not an ESS, as you can see by imagining that because of a mutation (or even just through the natural variations between individuals) a frog develops an extra sticky tongue. It will do well, compared with other frogs, and genes for extra sticky tongues will spread through the frog population. At first, a larger proportion of flies gets eaten. But the ones who don't get eaten will be the more slippery ones, so genes for extra slipperiness will spread through the fly population. After a while, there will be the same number of frogs on the pond as before, and the same proportion of flies will be eaten each year. It looks as if nothing has changed – but the frogs have got stickier tongues, and the flies have got more slippery bodies.

In the real world, this kind of thing is happening all the time, in every generation, not just as a result of mutations. In each generation there are some frogs with stickier tongues than others, and some flies with slipperier bodies than others, so step by tiny step, as the generations pass, the trend towards stickier tongues and more slippery bodies is ratcheted forward. There is no obvious end to this – frogs get better and better at catching flies (one way or another) while flies get

better and better at escaping frogs. Each has to run as hard as it can to stay in the same place.

Of course, in the real world there are more than two species interacting with one another. To some extent, this just means that there is a lot more running on the spot going on. But, just as with the connections within networks of lightbulbs, there are important differences between networks, arising from the way the species (which are now thought of as nodes in a network) are connected. If what happens to one species affects every other species directly, it is impossible for even a pseudo-stable Red Queen situation to develop. There is utter (but deterministic) chaos in the system, because even small changes have large and in practice unpredictable effects. But the real world isn't like that. In real life, what happens to one species affects only a few other species directly, although what then happens to those species may affect others, in a widening ripple of interactions. This is very much like what happens in the sparsely connected network of lightbulbs, or in the sandpile game where moving one grain of sand only directly affects its immediate neighbours, but the way they move affects their neighbours, and so on.

Another example from something approximating real life may help to make the point. Foxes eat rabbits, so what happens to foxes (if the population of foxes goes up or down, for any reason) affects rabbits, and what happens to rabbits affects foxes, directly. They are linked nodes in the network. But foxes do not eat grass, so at this level what happens to foxes does not affect grass. The fox and grass nodes are not directly connected to one another. But rabbits do eat grass – those two nodes are connected – so what happens to foxes, by affecting rabbits, in turn has

an effect on grass. More foxes means (other things being equal) fewer rabbits, giving the grass a chance to grow longer. And this in turn will affect other species that feed on grass, and so on and so on. Getting back to Kauffman's frogs, if the frogs themselves (or their tadpoles) are eaten by fish that live in the pond, and the fish are eaten by bears who live in the woods nearby, if the fly population crashes (perhaps because people have begun spraying the area with insecticide), the bears will go hungry. Real ecological networks, of course, are more complex than this, with some (but not too many) interconnections – the fish might eat flies as well as frogs, for example. It is interesting, though, that such systems were referred to colloquially as 'the web of life' even before network theory was applied to them by people like Kauffman.

What's exciting about all this is that it seems that a process of co-evolution, in which all the species involved in a network change together when one of them changes, will naturally push complex ecosystems from either extreme towards the interesting region of self-organized criticality, in the phase transition on the edge of chaos. If a group of organisms is locked into a stable strategy, a mutation involving one of the species is likely to open up the network, allowing it to evolve. Evolution by natural selection will ensure that a change which is detrimental to the species involved will be washed away over the generations; but a change that is beneficial will spread, and by spreading it will open up more networks, pushing the system towards the edge of chaos. On the other side of the phase transition, in the chaotic regime, the same thing happens in reverse. With the rules of the game of life changing in every generation, any group of individuals that manages to insulate itself from the chaos to some extent, by reducing the number of

its connections to the outside world, will get a chance to evolve, by natural selection, into a state that takes advantage of the opportunities provided. The logic of this hand-waving argument is backed up by numerical simulations and experiments with clusters of lightbulbs wired up in different ways, carried out by Kauffman and his colleagues. Simply by each individual species acting to maximize its own evolutionary fitness, the whole ecosystem evolves towards the edge of chaos. This is exactly analogous to the sandpile model, in which the addition (or removal) of a single species from the network corresponds to the addition (or removal) of a single sand grain from the pile; it is also essentially the same as the lightbulb model, with the addition or removal of a single species corresponding to flipping a single lightbulb on or off. Complex systems naturally evolve towards the phase transition at the edge of chaos.

All of this translates beautifully into a language that evolutionary biologists were already familiar with, and extends that model in ways that were undreamed of by its pioneers, but which ought to be familiar from what we have learned about the way classical thermodynamics was extended by the work of people such as Ilya Prigogine. The model involves what are known as fitness landscapes, and stems from the work of R. A. Fisher (1890–1962; he was knighted in 1952) at the Rothamsted Experimental Station, a British research centre, in the 1920s. Fisher was the first person to put evolutionary biology on a mathematical footing, calculating how the distribution of genes in a population changes as a result of natural selection operating on individual members of the population. But Fisher, like the classical thermodynamicists, was only interested in equilibrium situations (a reasonable enough first step),

identifying the mathematical best possible solution to each evolutionary problem.

This approach was developed further by the geneticist Sewall Wright (1889–1988), working at Yale University in the early 1950s. The analogy he introduced is with a landscape, where hills represent successful evolutionary strategies ('good genes', or more accurately, good packages of genes working together), and valleys represent unsuccessful evolutionary strategies ('bad genes'). Any particular individual can be represented by a single point on the landscape, which is really a kind of phase space, with the hills as attractors. But because members of a species are not identical to one another (one of the key points, remember, about Darwinian evolution), a species will be represented on the landscape by a little cluster of points, like a little flock of sheep. The sheep who are a bit higher up the slope will do better in each generation, leaving most offspring, while the ones who are lower down the slope will do worse, leaving fewer offspring. So as the generations pass, the little flock moves up the slope until it reaches the peak, where it stays.

This approach to understanding evolution is ideal when we are considering what happens to a single species interacting with only a few other species in a stable environment – precisely the conditions that have produced the variety of finches seen by Darwin on the Galápagos. It is thought that these are all descended from a few birds (perhaps only a single breeding pair) who found their way to the islands some time in the past. These birds were all originally members of a single species, only poorly adapted to eat the food they found on the islands – a little flock sitting in a valley between the hills of the fitness landscape. Some



members of the flock evolved by developing longer beaks, moving up one particular hill in the landscape, while others evolved stouter beaks, moving up a different hill, until they became separate species. The process comes to an end when, in one case, the beak becomes so long and thin that it cannot be used to probe for food without snapping off, or, on the other hand, when the beak becomes so unwieldy that it cannot be used to manipulate the seeds.

The fitness landscape, underpinned by Fisher's mathematics, provides a beautifully simple image; even the mathematics is quite straightforward. But there is a problem. Suppose the peak the species has evolved to is just a foothill, and there are much higher peaks (more evolutionarily advantageous states) that it could be on nearby. There is no way for the species to cross the intervening valley floor and climb those peaks, because to do so it would first have to evolve towards a worse (less well fitted) state than the one it is in already. In effect, it has evolved up a blind alley, from which there is no escape. Fisher's landscape is not only a static one, it is one in which species can only achieve a local equilibrium, and have to stay there.

The Red Queen effect is one example of how interactions between species change that simple picture. When one species changes, either through mutation or, in the most extreme case, through extinction, all the species it interacts with are affected. The effect is as if the heights of the corresponding peaks in the landscape have changed. Individual species still evolve towards the peaks nearest to them on the landscape, but because of co-evolution the position and size of those peaks change all the time. In the static regime, nothing changes; in the chaotic regime, the landscape changes so rapidly that nothing interesting can ever get

time to happen. On the edge of chaos, the landscape is always changing, but usually slowly, and this keeps opening up new evolutionary possibilities for individual species, and groups of species, within the ecosystem. It is the endlessly changing landscape, a kind of rubber landscape which deforms like a slack trampoline as the species 'walk about' on the rubber surface,<sup>7</sup> that provides the opportunity for endless evolution. In our frog and fly example, the arrival of humans with cans of insecticide knocks down the peak the flies were sitting on and leaves them deep in a valley; but if they do not all go extinct they may well evolve resistance to the insecticide, moving up another peak as the population recovers from its setback. Note that this is not a new peak – it was there all the time, corresponding to insecticide-resistance. But the peak got higher once there was insecticide around, because that made the ability to resist insecticide more useful. Even so, the flies could not cross over to it until their previous peak collapsed. Unlike Fisher's scenario, there is no single 'best place' for a species to settle down in on the evolutionary landscape; all it can do is move towards a better place whenever the opportunity arises, even though what is better in one generation might be worse in a few generations' time.

Note, though, that this is no longer solely about individual species doing the moving towards a better state. John Holland, of the University of Michigan, Ann Arbor, has drawn attention to the importance of sub-units of the ecosystem, made up of individual species working together in groups. Indeed, this image of groups of individuals working together applies at all levels from genes, where one group of genes may work as a unit to control one aspect of the workings of the cell, to cells, where one group of cells work together as an organ such as the liver, right up

through to species, where several species may be linked together in an interacting network, somewhere between an ESS and total independence. As Holland has put it, 'The cut and try of evolution isn't just to build a good animal, but to find good building blocks that can be put together to make many good animals.'<sup>8</sup> His favourite analogy is with the way in which 'Identikit' pictures used to be put together, before the days of computers, to make an impression of the face of a criminal suspect, based on eye-witness descriptions. From a choice of ten different facial features (hairline shape, size of nose, and so on), each coming in ten varieties, it is surprisingly quick and easy to produce a 'picture' resembling any chosen individual – surprising, that is, until you remember that you are dealing with  $10^{10}$  combinations, ten billion different faces, rather more than the number of people alive on Earth today. Holland might have gone even further, pointing out that the cut and try of evolution also operates to find 'good building blocks' in the form of interacting networks of species that operate in some ways as distinct evolutionary units. Co-evolution, not just evolution.

Although he did not express it that way, this was what Leigh Van Valen had discovered, through his studies of marine fossils, which led him to come up with the idea of the Red Queen effect in 1973. He found that any genus within a group (for example, the bony fishes) has a constant probability of going extinct and disappearing from the fossil record. It doesn't matter how long, or how short, a time the genus has existed already, the chances of it going extinct in any chosen interval of geological time are identical. The same applies at other levels in the web of life. In this sense, species get neither better at surviving (in spite of evolution) nor worse as time passes; they die off at random. Van Valen's

interpretation of this evidence was that the struggle for survival is always just as hard as it used to be. Since that struggle involves individuals interacting with one another, this must mean that everyone gets more efficient all the time, until something happens to remove one species (or genus) from the scene. This leads to a reshuffling of species among the available ecological niches, followed by a new arms race involving a lot more running on the spot. All species must evolve as fast as they can just to keep up with the other species – as much a ‘Keeping Up With The Joneses’ effect as the Red Queen effect. One result of all this interaction, which has greatly excited some evolutionary biologists, is that successful species actually get better at evolving as time passes – a mutation which enables a species to adapt more quickly to its changing environment is itself likely to be successful and spread through a population. An example of this has been the evolution of sex, which is a way of reproduction that speeds up evolution (by mixing genetic material from two parents in one offspring), giving large, slow-breeding multi-celled creatures like ourselves a chance to keep up in the arms race in competition with small, swiftly reproducing organisms such as bacteria and parasites.<sup>9</sup>

We should pause here and explain our use of some evolutionary terminology, which probably annoys serious evolutionary biologists but which ought to be clear to everybody and acceptable outside the halls of academe. In much of this discussion we talk about species going extinct, and the evolution of species as a whole. Of course, we are well aware that evolution operates at the level of the individual and that in the struggle for survival it is the least fit individuals that fail (by leaving no descendants). But if that happens to enough individual members of a

species, then the species goes extinct. It seems reasonable to talk about species evolving as a whole in this way, even though we know that the underlying mechanism is individual selection, and even to talk about an individual species 'mutating' so that there is a change at one node in the ecological network, which ripples out to affect other species. After all, we talk about a particular football team competing for a trophy as if it were a single entity – 'City knocked out of the cup' – even though we know that the team is made up of eleven players at any one time, many more than that over the course of the season, and relies also on the skills of the manager and other non-playing staff.

You can see why it is a reasonable shorthand to describe evolutionary changes in the geological record in terms of whole species 'mutating' or going extinct, rather than focusing on selection at the level of individuals, by looking at a very simple example of the way in which evolutionary changes affect individual members of a species. This happens in ways which are far too small to be noticed in a human lifetime, but can lead to effects which show up as sudden and dramatic changes in the fossil record, with one species being replaced by another in the geological equivalent of overnight. Our favourite example of how this might work involves a hypothetical mouse which slowly evolves into something like an elephant. Imagine that a baby mouse grew up to be  $\frac{1}{20,000}$ th bigger than its parents, and that the same increase in size occurred in every succeeding generation. Then it would take 20,000 generations to 'evolve' supermouse, the size of present-day elephants. Mice breed very quickly, and elephants breed very slowly, so for the purpose of this hypothetical example we can say that each generation takes five years to achieve maturity, somewhere between the actual times for mice and

elephants. So the change from mouse to supermouse would occupy 100,000 years. In terms of the fossil timescale, that is instantaneous. Differentially dating fossils that have been laid down just 100,000 years apart is almost impossible, and if this kind of evolutionary change were taking place in a species, what we would see in the fossil record is one layer of strata with mouse-sized remains, and the next layer of strata with elephant-sized remains, with nothing in between.

In the real world, such a dramatic change would not occur in isolation; other species would be affected by the evolution of supermouse (assuming that such a change could happen at all), and we would see a whole suite of changes occurring ‘instantaneously’ in the fossil record. But all of this would be a result of slow changes from one generation to the next, taking place in line with the rules of Darwinian evolution.

It is important to stress this, because there has been a lively debate among the experts surrounding what is called by evolutionary theorists ‘punctuated equilibrium’, a pattern seen in the fossil record when there are long intervals with little or no evolutionary change (except for species becoming ever more finely fitted to their ecological niches through the Red Queen effect) ‘punctuated’ by short (sometimes, seemingly instantaneous) intervals of dramatic change in which some species die out and others evolve into new forms. This is sometimes presented, even by those who should know better, as conflicting with the ideas of Darwinian evolution, which are then referred to as ‘gradualism’ to emphasize the contrast. But there is no conflict. Even during the intervals of seemingly rapid change (the punctuation marks of evolution), at an individual level what is going on is still Darwinian

evolution, evolution by natural selection and survival of the fittest. It's just that the rules about what makes a species fit have changed. And change happens only slowly from one generation to the next – a mouse does not give birth to an elephant, or the other way round. Indeed, what happened to the Galápagos finches that Darwin studied, giving him a key insight into how evolution works, is an example of punctuated equilibrium. The finches that arrived on the islands were superbly adapted to a lifestyle that was appropriate on the mainland. They evolved rapidly (by geological standards, though only with tiny changes from one generation to the next) to fit the ecological niches available on the Galapagos, and once they had adapted in this way they then changed only slightly as they became fine-tuned to their surroundings. Darwin would have been thoroughly at home with the idea of punctuated equilibrium, and would have found his ideas totally unthreatened by it.<sup>10</sup>

The sandpile analogy helps to make this clear. When the sandpile is in a self-organized critical state, it experiences sudden landslides, interspersed by longer intervals of relative calm, while the sand grains are building up. But each sand grain is always subject to the same laws of physics – Newton's laws of motion, friction, gravity and so on. There is no fundamental difference between the two situations, except one of timescale. If grains are dropped on the pile every second, but you only look at the sandpile every half-hour, it will never look the same twice, and will appear to be in a constant state of change. If you look at it every microsecond, almost every time you look you will see the sand grains in almost exactly the same positions they were in last time you looked, and you will think that the pile only changes gradually. If the

situation among the living species of an ecological network stays much the same for a hundred million years, then changes occur over a million years or so, that is punctuated equilibrium, even though the species involved in the changes might not be aware of any dramatic changes going on around them. If the intervals of change lasted for a hundred million years, and the intervals of calm for one million years, we might well talk about ‘punctuated chaos’, but it would also obey the same evolutionary rules applied to sparsely connected networks. The point is that sparsely connected networks naturally evolve to such a state, with intervals of equilibrium punctuated by intervals of change, although our understanding of such networks is as yet far too inadequate to be able to say whether there is anything special about the particular balance between punctuation and equilibrium that seems to have been struck by the network of life on Earth. We do, though, have a good idea of the last ingredient needed to drive that living network to the phase transition on the edge of chaos – a small, but important, point that Kauffman missed, but which Per Bak and his colleagues laid bare in the middle of the 1990s.

In all the earlier work of this kind, using computer models to simulate the evolution of networks of interacting species, the researchers had chosen a species at random to ‘mutate’, and watched the way the resulting ripples spread through the network like an avalanche of sand grains in the sandpile model. This gave simulations that looked superficially very like the sandpile model, but which did not actually evolve to the edge of chaos. The breakthrough came when Kim Sneppen, from the Niels Bohr Institute in Denmark, visited Bak at Brookhaven in the United States in 1993. Sneppen was interested in interactions



between surfaces – an example Bak uses is the way spilled coffee spreads through a paper napkin, with the interface being the changing boundary between the wet part of the napkin and the dry part of the napkin. But Sneppen's calculations apply to a wide range of phenomena in physics; he found that changes in the boundary, or interface, are most likely to occur at places where the value of some appropriate parameter is either at its largest or at its smallest – a situation called 'extremal dynamics'. In the particular (almost trivial) example of the spilled coffee, the liquid spreads more easily through the napkin at locations where there are the widest pores allowing the liquid to flow more easily. In discussions between Bak and Sneppen about the failure of attempts to get simulations of ecological networks to evolve to the phase transition by selecting species at random for mutation, the two of them realized that it might be more appropriate to apply a form of extremal dynamics here as well. After all, even in the archetypal sandpile model, new avalanches do not begin at random locations but at the point where the slope of the sand is steepest; and earthquakes always begin at the point where the strain that has built up first exceeds some critical threshold, even if the extent of the resulting earthquake (like the extent of the resulting avalanche in the sandpile model) is then governed by other factors.

The extreme cases in the ecological networks that people such as Kauffman and Bak had been studying are the species that are best fitted to their environment, and the species that are least fitted to their environment. What would happen if Bak's team re-ran the simulations, but selected only the least fit species to change at each step of the simulation? This would be pure Darwinian evolution – extinction of the least fit. But it doesn't have to be extinction, it could just be change. The

least well-adapted species in any real ecology are not just going to sit around being badly fitted; they will either die out, or they will evolve into a better fit with their surroundings (like Darwin's finches). The best-fitted species will change little (except through the Red Queen effect) unless something happens to their surroundings – a frog that is superbly adapted to eating flies will be in big trouble, for example, if the flies go extinct. The least fit species would be the ones where any change which didn't make them go extinct would be almost certain to be an improvement, sending them further up the foothills of the evolutionary landscape. If such a species is only on a low hill to start with, it is also relatively easy for it to evolve off that hill, across a valley and up a higher hill nearby. It all made sense; and it made even more sense when the computer simulations were run.

The model that Sneppen and Bak came up with is stunningly simple. It just takes a thousand 'species' and represents each of them by a single random number, between 0 and 1, to indicate its Darwinian fitness. A high number corresponds to a species that fits in well with the whole network, and sits on a high peak in the fitness landscape. A low number corresponds to a species that really is struggling to survive, and sits unsteadily on a low hill in the fitness landscape. In this sense, the amount of 'fitness' corresponds to the size of the peak in the landscape. In order to take account of the sparse connections between the nodes in the network, each of the resulting numbers was then randomly linked with two others so that each 'species' interacts with just two others (as if the frogs only eat flies, and are themselves only eaten by fish). Then the model was programmed to remove the least fit member of the population, *and its two partners*, and replace them with three 'new'

species with fitnesses assigned at random. These need not be thought of as literally being new species; the point is that the fitness of the least fit member of the network changes at random, and this causes the fitness of each of its partners – you might think of them as its neighbours, in the ecological network – to change at random, no matter how fit they were to start with. It is the rubber landscape at work. Keep repeating this simple action iteratively in the computer and watch what happens.

At first, having set the network up at random, there are nodes with values of fitness covering the entire range from 0 to 1. The least fit ‘species’ will therefore have a very low value of fitness. But when the least fit species is removed and replaced by another with a fitness assigned at random, it is highly likely that the new species will have a higher fitness, because a random number between 0 and 1 is likely to be higher than the very low number associated with the old species. Simply by removing the ‘least fit’ species and its two neighbours in each iteration, and replacing them by new species with fitness assigned at random, these computer simulations very quickly raise the overall fitness of the entire network, until all the nodes are associated with relatively high peaks. Specifically, for the rules we have described here, the system settles down with all the species having a fitness of at least two-thirds. In real life, this corresponds to a stable ecological network, with each individual species occupying a high peak in the fitness landscape. Such a situation ought to last for a long time. But what happens if, and when, one of these species does go extinct, or suffers a loss of fitness caused by a mutation, affecting the landscape of its two neighbours? Now, the only way is down. Replacing three species with relatively high fitness values by species with fitness assigned at random is more likely than not to

produce a drop in the overall fitness of this little cluster of nodes. This means that these three new (or mutated) species, and their neighbours, are the ones most likely to be affected by the subsequent iterations, producing a ripple of extinctions (or at least, fitness reductions) that spreads out from the original site of the change. Remember that a fitness reduction does not necessarily mean that a species has changed; it is the fitness landscape that changes, so that a species superbly adapted to one situation finds itself in a different situation. A kind of avalanche, or earthquake, runs through the system, before it adjusts and builds back up to the state in which all the fitness levels are above two-thirds and the system can settle down for another quiet interval. The model naturally produces a situation in which intervals of stability are separated by mass extinctions, even though the same rules apply during the quiet intervals and the extinctions. It should be no surprise to learn that the result is described by a power law, meaning that the extinctions are self-similar, although the exact value of the power law (the steepness of the slope in the power law graph) doesn't quite match the power law of actual extinctions in the fossil record.

Critics of this work have argued that the models are simply too simple to be realistic, or to provide any insight into the complexities of real ecosystems and the way they evolve. That is hardly surprising, given how simple the approaches used by Kauffman, Bak and their colleagues are. But such criticism misses the point that the simple models do give a good description of the real world. If these models failed to produce patterns of extinctions described by power laws, nobody would take them seriously. The test of a good model is not how simple it is, but how well it provides insight into real systems. In atomic physics, for example,

it is at first sight an almost ridiculous simplification to treat atoms as being like miniature solar systems, with electrons 'in orbit' around a central nucleus. We know that atoms are more complicated than this. But this simple model, developed by the Dane Niels Bohr in the second decade of the twentieth century, actually works very well in predicting the exact wavelengths of the lines seen in the spectra of different elements. It is therefore a good model, even though we know atoms are not 'really' like that. A model that could predict not only that extinctions obey a power law, but also the right value of the power in that law, would be just as valuable in its own way as the Bohr model of the atom, and would also be telling us a deep truth about the underlying simplicity of the world. More complicated features of the behaviour of atoms are described by adding more detail to the simple Bohr model, without throwing away the insight that it provides. The fact that such simple models as those of Kauffman and Bak immediately reproduce the essential feature of real extinctions, their scale invariance, suggests in a similar way that the researchers are on the right track, and that they ought to be able to produce a closer match between the models and the real world by adding in a few more details to make the models more realistic. In fact, it turns out that adding just one more detail is enough to get exactly the right power law.

This is still very much work in progress, and it is probably fair to say that we understand about as much (or as little) about real evolutionary networks in the first decade of the twenty-first century as Bohr did about real atoms in the second decade of the twentieth century. But one of the most intriguing refinements to this kind of model came right at the end of the twentieth century, in 1999, when Luis Amaral of MIT and Martin

Meyer of Boston University introduced that one crucial (and clearly relevant) detail to the model. They added in a simple representation of another deep truth about the real world – some species eat other species. Amaral and Meyer added predators and prey to the game devised by Bak and Sneppen. Again, this is a simplified version of the way the real world operates, but it produces a dramatic improvement in the match between model and reality.

In this variation on the theme, there are six layers in the food chain. Individual species on each level are allowed (indeed, required) to feed on several different species from the level below, and on each level there are a thousand ecological niches, equivalent to nodes in the Bak-Sneppen version of the game. The game starts with most of these ecological niches empty, and a few species randomly assigned to niches in the lowest level of the chain. The rules, applied as usual in each iteration, go like this. First, there is a small probability that each species in existence at the beginning of the iterative step will split in two (like the splitting of ancestral finches into different species on the Galápagos), producing a new species that is randomly assigned to one of the empty niches in the same level or on the one above or below that of the parent species. Each species that emerges in this way is randomly assigned a few species from the level below which are its prey. Secondly, at each iteration a few of the species in the lowest level, chosen at random, go extinct. The difference from the Bak-Sneppen model is that when a species disappears in this way it doesn't affect neighbours that are otherwise its equals, but the predators that feed on it. Predators who lose more of their food supply do less well than those who lose less of their food supply, and in extreme cases all of the prey of one predator may disappear, so that the

predator itself goes extinct, affecting predators in the next layer up the food chain. In a pattern that has by now become familiar, with just these simple rules at work and random starting conditions, the game organizes itself into a state on the edge of chaos, a critical state in which the disappearance of just one prey species at the bottom of the food chain (perhaps simply because they have all been eaten by the predators in the level above) can sometimes trigger a wave of extinction rippling upwards through the layers of predators and prey. Not only do we once again find that the pattern of these waves is scale invariant and obeys a power law; this time, the exact value of the power in that law is the same as the power of the law involved in real extinctions of life on Earth. Interacting species in an ecological network move towards a self-organized critical state in which mass extinctions obeying a power law will occur without the help of any outside intervention at all.

And yet, we know that there are outside interventions – the classic example is the meteorite that struck the Earth at the time of the death of the dinosaurs. Even if, for the sake of argument, we said that this was not the primary cause of the extinctions at the K-T boundary, it certainly must have had some effect on life on Earth. Working out how such an extra influence might affect the pattern of extinctions is clearly important, but adds another layer of complexity to the calculations. As usual, the first instinct of the modellers is to strip away everything they don't need from the models and start with the simplest possible new model to begin to understand the 'new' phenomenon. In this case, Mark Newman of Cornell University decided in the mid-1990s to investigate a model in which outside influences are the *only* cause of extinctions, ignoring the effects of interactions between species on the rubber

landscape of fitness. The hope is that by looking both at extinctions caused by interactions between species in the absence of external shocks, and at extinctions caused by external shocks without interactions between species, we might see which version is closer to reality, and how the two effects might combine.

The basic set-up in Newman's model is very similar to the basic set-up in the Bak-Sneppen model. The first difference is that instead of just the single 'least fit' species being removed, the model chooses a fitness level (between 0 and 1) at random, and removes *all* species with less than this fitness in one go – equivalent to an external event, such as a meteorite striking the Earth, which takes out all but the most well-adapted species. The empty niches in the ecology are then filled in by new species with fitness (as ever, between 0 and 1) chosen at random. Then, to take account of the fact that the rules of the game have changed (the evolutionary landscape has shifted), a few of the survivors, the fittest species in the original ecological network, are taken out and replaced by others with fitness chosen at random. This is equivalent to altering the fitness of the original species in those niches. As one final option in setting the rules for each particular run of the computer model, it was possible to select in advance what kind of external shocks the researchers wanted to have – there might be a relatively high proportion of small shocks that only took out the least fit species, with just a few big shocks that wiped out nearly everything in the network; or the model could be set up the other way around, or with any statistical mixture of large and small external shocks, although at each iteration the model chose the size of the next shock at random in accordance with the specified statistical rules. The surprising and dramatic result was that



for all but a few extreme versions of this game the system settles down into a critical state in which the pattern of extinctions obeys a power law similar to the one we see in the pattern of extinctions in the actual fossil record. <sup>11</sup>

One way to think of this is that just as tension builds up in an earthquake zone until something has to give (but you can't say in advance how big the 'give' will be), it is as if gradual Darwinian evolution builds up more and more tension in the ecological network, through the Red Queen effect, until something snaps, and the whole network (or perhaps just part of it) collapses. There is no way to predict in advance whether the next wave of extinctions to sweep across the ecological network will be large or small, just as there is no way to predict whether the next earthquake to strike San Francisco will be large or small. Living and non-living systems follow the same deep imperative.

Once again, the model is very simple, and we cannot pretend that it describes perfectly what happens to a real ecological network when disaster strikes from outside. But what it is telling us is that the power law pattern of extinctions is a very robust feature of the way shocks affect such networks. It tells us, first, that the kind of external influence being applied is essentially irrelevant, as long as it kills things. Frequent small shocks, occasional large shocks, or any mixture of the two (including occasional small shocks) produce the same pattern. So a mixture of different natural occurrences – meteorites striking the Earth, bursts of intense volcanic activity, Ice Ages, or whatever – each with their own pattern, will combine to produce exactly the same picture of power law extinctions. Even better, as we have already seen, evolutionary changes in the ecological network alone, *without* any

external shocks, also produce the same power law pattern. All of this explains why the actual pattern we see in the record of fossil remains distributed through time is so simple. All of the different effects produce the same pattern! We may never be able to say whether the events at the K-T boundary were primarily a result of meteorite impact, or of evolutionary change, or major volcanic activity, but only that some mixture of these (and possibly other) effects provided the trigger for that wave of extinctions. As in Agatha Christie's *Murder on the Orient Express*, everyone may be guilty.

From the perspective of biologists, or geologists, or astronomers, each hoping to find proof of the validity of their own pet hypotheses, this might be frustrating; but from the point of view of someone trying to understand how complex networks of interacting life-forms evolve, it is highly interesting, and even encouraging. It suggests that in spite of the extreme simplicity of all these models we have indeed discovered a fundamental truth about the way nature works, a truth which applies to any system driven away from equilibrium, including the particular systems we are interested in, living systems. Over an extremely wide range of possibilities, whatever conditions you start out with and whatever shocks you apply to the living systems (external or internal, or both), you arrive at the self-organized critical state on the edge of chaos, where even a small trigger can, on occasion, produce a very large change in the system as a whole. Life really is like that.

As Newman's work emphasizes, when the physical environment changes, the fitness landscape changes. But life itself also changes the fitness landscape, and the fate of individual components of life on Earth at any particular moment in history depends on the way the two

influences interact. A beautiful example is provided by the greatest of all mass extinctions, the one that occurred at the end of the Permian. We have seen that a small trigger could have been responsible even for that great wave of extinctions; but this would only have been possible if most of the species around at the time were indeed connected in one ecological web. Today, if a disaster of some kind struck North America (and remember that in this context a disaster might be caused by a single species going extinct and triggering a wave of extinctions), the effects might ripple out across both North and South America. But it is hard to see how they would be of much concern to the species that lived in Australia, Africa or Eurasia. That is one reason why people invoke truly global events, such as the aftermath of a giant meteorite impact, in an attempt to explain global mass extinctions. But at the end of the Permian, almost all of the land mass of our planet was combined in one supercontinent, Pangea. That provided the opportunity for essentially all of life on the surface of the Earth and in the shallow seas bordering that supercontinent to be part of a single network. A shock affecting one part of that network (a small shock from outside, or a species going extinct as a result of evolutionary pressures alone) could have rippled out to engulf almost everything alive at the time. There is no proof that this is what happened then. There never could be. But that is not the point. The point is that 250 million years ago, at the end of the Permian, a small trigger *could* have caused the extinction of most species around at that time, whereas a comparable small trigger affecting one continent today could not produce such a large global wave of extinctions. The physical environment and the biological environment are interconnected, much more subtly than it seems at first sight.

But could the connection operate the other way around? Could the biological environment of the Earth also affect the physical environment, so that the two together form part of a single network obeying the same deeply simple set of rules? Such a speculation is the kernel of the 'Gaia hypothesis' put forward by the independent British scientist Jim Lovelock. Lovelock's insight, which grew into the Gaia hypothesis, dates back to 1965, and predates much of the work on complexity and emergence that we have been describing. But taking Lovelock's ideas on board now extends the study of chaos, complexity and emergence to encompass the whole of our planet, and to provide hints at the way life may have evolved on worlds beyond our Solar System.

## 7

### Life Beyond

The big difference between the way we have been looking at life so far, and the way we are going to look at life from now on, is that previously we were looking at life from the inside out, while now we are going to look at life from the outside in. You can study the way molecules like amino acids and DNA interact inside cells, and the way cells work together to make up a body, and get a lot of insight into how life works. Or you can watch a human being (or a dog, or a jellyfish) going about their daily life, and get a different perspective on what life is all about.

The shift in perspective, which provides new insights into the nature of life on Earth as a whole, came about as a result of two things: a spectacular image, and the work of one man, both of them linked to the exploration of space. The image was a photograph taken by the Apollo astronauts, showing the Earth as our home in space, a single blue-white oasis of life surrounded by a black desert. The man was Jim Lovelock, who came up with the idea that living and non-living components of the terrestrial environment interact in a network which maintains stable conditions suitable for life on Earth. His insight was directly based on the principles of thermodynamics and the differences between equilibrium and non-equilibrium systems that we have already discussed, even though, because of Lovelock's background, that insight came quite independently from the work we have already described

came quite independently from the work we have already described concerning the application of non-equilibrium thermodynamics to situations involving systems at the edge of chaos.

Lovelock was born in 1919, and after leaving the University of Manchester in 1941 with a degree in chemistry he spent two decades in medical research, developing his skills as a 'hands on' experimenter who designed and built much of his own equipment. It was only at the beginning of the 1960s that he took the first steps towards the independence for which he later became famous, eventually being funded by the income from the scientific instruments he invented, and free to do research on anything he liked, although he had always been independently minded and non-conformist even by the sometimes eccentric standards of scientific researchers.<sup>1</sup> The first steps towards that independence involved working for NASA and the Jet Propulsion Laboratory (JPL) as a consultant on the design of instruments for spacecraft intended to land on the Moon and Mars and analyse the samples of surface material they found there. After contributing all he could to the chemical side of the programme, Lovelock gravitated towards the engineers designing the hardware, becoming, as a person who could speak both languages, a link between them and the biologists making plans to search for evidence of life on Mars. Most readers of this book probably do not remember the primitive state of ordinary domestic electronic equipment in the 1960s, when a decent living could be made as a TV repair man. The thought of launching that kind of equipment into space, landing it on Mars, and expecting it to work and send back useful information to Earth using a transmitter with a power of just 100 watts (that of an ordinary light-bulb), seemed fantastic to most people.

As Lovelock puts it, the engineering required was about as far in advance of the engineering used for a domestic TV set of the time as that TV was in advance of the engineering of Roman times. It did work; but there was a tight limit on how much information could be sent back by the feeble transmitter, which made deciding just what to measure – if the lander made it to the surface of Mars – a matter of urgent priority.

The key insight which led him to the concept of Gaia came, Lovelock recalls, in September 1965. By that time, he was based back in England, but still a regular visitor to JPL as a consultant.<sup>2</sup> On one of those visits, in 1964, he sat in on a meeting of a team discussing the kind of experiments to include on a lander in order to search for traces of life on Mars. He was struck by the fact that all of the ideas being discussed were based on the assumption of finding the proverbial ‘life as we know it’ – as Lovelock says, the experiments would have been good at finding life in the Mojave desert, just to the east of the JPL base near Los Angeles, but nobody seemed to be considering the possibility that life on Mars might be completely different from life even under extreme conditions on Earth. He suggested that what was needed was an experiment that could look for the general attributes of life, rather than specific kinds of life, and, when challenged to explain what kind of experiment could do that, replied that what was needed was an experiment to look for entropy reduction. As Lovelock appreciated, and as we have seen earlier in this book, living systems characteristically bring local order to systems, making entropy ‘run backwards’ as long as they have an external source of energy to feed on. In the best traditions of the ‘can do’ ethos of those early days of space exploration, Lovelock was given a couple of days to come up with a practicable idea for an experiment to

measure entropy reduction, in effect being told to 'put up, or shut up'. And once his mind was concentrated on the problem, it proved disarmingly easy to solve. The best way to look for entropy reduction processes at work on Mars would be to measure the chemical composition of its atmosphere. If there were no life on Mars, the gases in the atmosphere would be in a state of thermodynamic and chemical equilibrium, dominated by stable compounds such as carbon dioxide. If there were life on Mars, then the waste products from life processes would be dumped in the Martian atmosphere, providing reactive gases such as methane and oxygen, which would lower the entropy of the atmosphere.

There were other possibilities, some of which Lovelock later spelled out in a paper in the journal *Nature*, including the possibility of listening for and analysing sounds in the Martian atmosphere. As we have seen, the sounds made by living things contain information (negative entropy), and are characterized as  $\frac{1}{T}$  'noise', quite unlike the white noise of random fluctuations. Whether it was the Martian equivalent of bird song, or chirping crickets, or Mozart, a simple analysis of sound patterns could reveal the presence of life. These, and other ideas, intrigued the planners at JPL (even though many of the biologists remained unimpressed). NASA planners were equally impressed, and to Lovelock's astonishment he became Acting Chief Scientist in charge of developing such physical life-detection experiments for a proposed Mars mission. It was too good to be true. In September 1965, funding for the proposed mission failed to get the approval of the US Congress, and the less ambitious Viking project, that did land instruments on Mars in 1975, ended up with the standard package of biological experiments, some of



them based on Lovelock's designs – the biologists were willing to make use of his technical expertise, even if they distrusted his biological insight. When asked how he felt about seeing so much effort go into a project where the outcome was a foregone conclusion, he said that it felt like 'designing a robot to look for signs of life in the Sahara desert and equipping it with a fishing rod'. We know that fish exist and are examples of living systems, so finding a fish in the Sahara would be proof of life in the desert. But the negative results from the biology experiments on the Viking landers actually revealed little or nothing about the presence of life on Mars except that it is not the kind of life found in the Mojave desert, just as the failure of a fishing expedition to the Sahara proves only that there are no fish in the desert, not that there is no life there. To this day, the best evidence that there is no life on Mars comes not from space probes, but from spectroscopic studies of the composition of the Martian atmosphere – spectroscopic studies which, coincidentally, were also first reported in September 1965, ten years before the Viking landers carried their superbly designed but useless 'fishing rods' down to the Martian surface.

It was that same month that Lovelock happened to be at JPL when news came in that astronomers at the French Pic du Midi Observatory had obtained detailed spectroscopic information about the atmosphere of Mars from their analysis of light from the planet in the infrared part of the spectrum. Spectroscopy is the process of analysing the light from an object by splitting it up into the rainbow spectrum, using a prism or some other device, and looking at the lines in the spectrum produced by different atoms or molecules in the object being studied. These lines resemble the pattern of stripes in a bar code, and are just as individual;

one particular pattern will unambiguously reveal the presence of iron, say, while others (more relevant in the present context) are associated with oxygen, or methane, or carbon dioxide. The French observations showed for the first time that the atmosphere of Mars is almost entirely composed of carbon dioxide, with only traces of other gases present. It is a stable, unreactive atmosphere in a state of high entropy, corresponding to thermodynamic equilibrium – and as we have seen, nothing interesting happens in thermodynamic equilibrium. The very month that Lovelock's hopes of sending an entropy-measuring experiment to Mars were dashed, these ground-based observations showed that there was no need for the experiment. The relevant measurements had been made for him, and showed unambiguously that Mars is a dead planet today (they say nothing, of course, about what conditions might have been like on Mars millions, or billions, of years ago).

It was the arrival of the news from France that set Lovelock thinking about the contrast between the Martian atmosphere, dominated by carbon dioxide, and the terrestrial atmosphere, dominated by nitrogen but containing about 21 per cent oxygen, one of the most reactive gases, and only traces of unreactive carbon dioxide. Nitrogen in the atmosphere, Lovelock knew, is constantly involved in reactions with oxygen in the atmosphere (a kind of slow burning), ultimately to make nitric acid, which dissolves in the sea to make stable nitrates, which are broken down by bacteria (using the energy of sunlight) to return nitrogen to the air. It came to Lovelock, 'suddenly, just like a flash of enlightenment', that for the atmosphere of the Earth to stay in this seemingly stable state for hundreds of millions of years 'something must be regulating the atmosphere and so keeping it at its constant

composition. Moreover, if most of the gases came from living organisms, then life at the surface must be doing the regulation.’<sup>3</sup> Without pausing to think it through further, he blurted out the insight to his companions at the time, a NASA colleague called Dian Hitchcock, and another visitor to JPL, the astronomer Carl Sagan. This was the seed from which the idea of Gaia, the Earth as a self-regulating system, grew (the name was suggested, incidentally, by one of Lovelock’s neighbours in England, the writer William Golding).

To put it in a slightly different perspective, you might ask what would happen to all of the highly reactive oxygen in the Earth’s atmosphere if it were not constantly being renewed by the action of life. Take all life away from the planet, and within a very short time all the oxygen would be locked up in stable chemical compounds such as nitrates, carbon dioxide, water, iron oxides and silicate rocks. To be precise, without the intervention of life, all the oxygen in the atmosphere would get locked away in less than ten million years. That is a measure of how sensitive the seemingly stable physical environment of our planet is to the presence (or absence) of life. This isn’t too worrying on a human timescale – the popular myth that if the Amazon rainforest disappeared tomorrow we would all suffocate is far from the truth – but ten million years represents only about 0.2 per cent of the life of the Earth to date. If you are an astronomer who happens to observe a planet like the Earth and notice that it has an oxygen-rich atmosphere, that either means you are witnessing a rare, transient event on that planet, or that the atmosphere is being maintained in a state far from equilibrium.

The idea that life may be part of a self-regulating system that determines the physical nature of the surface of the Earth today (at least

in the 'life zone', a thin layer from the bottom of the ocean to the top of the troposphere, about 15 kilometres above our heads) met with a hostile reception from biologists initially, and still has its opponents, many of them put off by what they (wrongly) see as its mystic, quasi-religious overtones. There is also a mystic, quasi-religious Gaia movement (about as irritating to Lovelock as the Tolkien Society was to J. R. R. Tolkien), which is based on a misunderstanding of what Lovelock and his colleagues were saying. And then there are those who completely misunderstand what the whole thing is about. My copy of the CD version of the *Encyclopaedia Britannica*, which really ought to know better, says that 'the Gaia hypothesis is highly controversial because it intimates that individual species (e.g. ancient anaerobic bacteria) might sacrifice themselves for the benefit of all living things'. It certainly does not! That sentence makes about as much sense as saying that Darwinian theory is highly controversial because it intimates that rabbits sacrifice themselves for the good of foxes. So perhaps we should spell out that Lovelock is *not* saying that Gaia is God, or that Mother Earth is looking after us, or that one species makes sacrifices for the sake of the whole. He has simply found a way of describing all of the processes in which life is involved on Earth, including many processes that have traditionally been thought of as non-living physical processes, as part of a complex network of interactions, a self-regulating (or self-organizing) system, which has evolved to an interesting but critical state in which balance may be maintained for very long intervals by any human standard, but in which sudden fluctuations away from equilibrium (analogous to the punctuated equilibrium of biological evolution) can occur. In the language of the previous chapter, he is saying that the

behaviour of life on Earth alters the physical landscape (where ‘physical’ includes things like the composition of the atmosphere) as well as the biological landscape, and that both these changes affect the overall fitness landscape, with feedback a key component of the interactions. In the overall context of the ideas discussed in this book, the concept of Gaia is a logical and straightforward development from what has gone before; the surprise, which indicates the power of Lovelock’s intuition, is that he began to develop the concept of Gaia before many of the ideas we have discussed here had themselves become respectable science, so that it is only with hindsight that we can now see how it all fits together, and think (echoing Thomas Henry Huxley’s comment when he first learned from Darwin of the theory of evolution by natural selection) ‘how extremely stupid not to have thought of that myself’.

There is no need to detail the entire story of how Gaia became respectable science, but with the benefit of that hindsight we can pick out two examples of Gaia at work, one a theoretical model and the other taken from the real world, which bring out the ways in which self-regulation results from the interaction between biological and physical components of a living planet. The first of these, a model called Daisyworld, is particularly apposite since it builds directly from, and solves, a puzzle presented to Lovelock by Sagan soon after he had his moment of revelation at JPL; it is also a neat example of the main theme of our book, the idea of emergence, with the whole being greater than the sum of its parts. And it is, says Lovelock, ‘my proudest invention’.

The puzzle that Daisyworld solves is known to astronomers as the ‘faint young Sun paradox’, although it was really only ever a puzzle, not a paradox, and thanks to Lovelock it is no longer even a puzzle. The

puzzle stems from the fact that astronomers are able to say with a great deal of confidence that the Sun gave off much less heat when it was young than it does today. They know this by combining information about nuclear interactions obtained in experiments here on Earth with computer simulations of the conditions inside stars, and comparing the results of their calculations with information about the energy output and composition of stars of different sizes and ages, obtained by spectroscopy. This is one of the great (and largely unsung) achievements of twentieth-century physics, but we do not have room to go into the details here, and you will simply have to accept for now that astronomers do know how stars work.<sup>4</sup> The relevant point is that we can say with confidence that when the Solar System was young the Sun was between 25 and 30 per cent cooler than it is today – or, to put it the other way around, since it settled down as a stable star the energy output of the Sun has increased by between 33 and 43 per cent. The Solar System had settled down into more or less its present configuration by about 4.5 billion years ago, and we know, from fossil evidence in the oldest surviving rocks on the surface of the Earth, that both liquid water and life existed on the surface of our planet by 4 billion years ago. The puzzle is why the increase in heat output of the Sun, roughly 40 per cent over four billion years, did not boil the surface of the planet dry and wipe it clean of life.

There is no problem in explaining why the Earth was not a frozen ball of ice when the Sun was cool. We now know that the atmosphere of Venus, like that of Mars, is dominated by carbon dioxide, and carbon dioxide, along with water vapour, is a major constituent of the gases released by volcanic activity. There is no reason to think that the

atmosphere of the early Earth was any different from the atmospheres of its two nearest planetary neighbours, and an atmosphere rich in carbon dioxide would be good at trapping heat from the Sun near the surface of the planet, and keeping it warm through the so-called greenhouse effect. Although it is not the main reason why the air in a greenhouse stays warm (that is largely because the glass roof of the greenhouse stops the hot air rising by convection and escaping), the greenhouse effect is easy to understand. Visible light from the Sun passes through gases like carbon dioxide (and, indeed, through oxygen and nitrogen) without being absorbed, and warms the surface of the Earth. That warm surface re-radiates its energy back towards space, but at longer wavelengths, in the infrared part of the spectrum. Carbon dioxide (and also, as it happens, water vapour, but not nitrogen or oxygen) absorbs some of this infrared radiation, and then re-radiates the energy it has absorbed in all directions, with some going back down towards the surface and warming the planet, and some escaping into space. The overall result is that, viewed from the outside, there is a dip in the spectrum of light from the surface of the planet at the wavelengths where it is being absorbed in the atmosphere. An astronomer on Mars, armed with a suitably sensitive telescope and spectrometer, could tell that there was a trace of carbon dioxide in the atmosphere of the Earth by measuring this characteristic infrared radiation, just as the Pic du Midi team identified carbon dioxide in the atmosphere of Mars. But there is a much smaller proportion of carbon dioxide in the atmosphere of the Earth than in that of Mars.

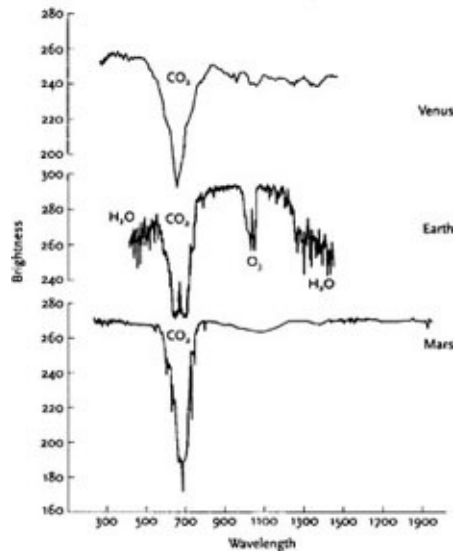


Figure 7.1 A comparison of the spectra of Venus, Earth and Mars immediately shows which of the three planets is in a non-equilibrium state and a likely home for life.

The power of the greenhouse effect can be seen by contrasting the average temperature at the surface of the Earth today with the average temperature of the airless Moon, which is essentially at the same distance from the Sun that we are. On the Moon, the average surface temperature is  $-18\text{ }^{\circ}\text{C}$ ; on Earth, the average surface temperature is  $15\text{ }^{\circ}\text{C}$ . The difference, an increase of  $33\text{ }^{\circ}\text{C}$ , is all down to the presence of just 0.035 per cent carbon dioxide in the atmosphere, plus water vapour and traces of things like methane, which are also greenhouse gases. More than 99 per cent of our atmosphere – the nitrogen and the oxygen – does not contribute to the greenhouse effect at all. If so little in the way of greenhouse gases today can have such a big effect, it is easy to see why the temperature of the young Earth never fell below freezing even when the Sun was faint. But the Earth hasn't just failed to freeze when it was young; it has failed to fry as it has aged, and has actually maintained a remarkably uniform temperature for billions of years as the



Sun has grown hotter.

It is quite easy to think of ways in which the temperature of the planet could have kept constant thanks to changes in the composition of the atmosphere, and people such as Carl Sagan had produced various hand-waving arguments along these lines before Lovelock came up with the notion of Gaia. But what those arguments lacked was any justification, except the desire of the theorists to maintain a constant temperature on Earth. What natural process could lead to such stability? Nobody knew. So was it just blind chance? Whatever the reasons, what you need to do, of course, is to steadily reduce the amount of greenhouse gases in the atmosphere as the Earth ages, to compensate for the increasing heat from the Sun. It is straightforward to see in a general way how this might have happened. The first photosynthesizing life-forms on Earth (those 'ancient anaerobic bacteria' that the *Britannica* refers to) would have taken carbon dioxide out of the air and used it in building their bodies, but they would have released methane into the air in its place, another greenhouse gas but with different infrared-absorbing properties from carbon dioxide. When the bacteria are more active, the balance shifts in favour of methane; when they are less active the balance shifts in favour of carbon dioxide.

The key to beginning to understand how this might work in nature was to introduce feedback into the calculations. In a simple model which took account of the increasing heat output of the Sun, Lovelock was able to show that by allowing bacteria to grow at a maximum rate when the temperature is 25 °C, but less vigorously at higher or lower temperatures, and not at all when the temperature fell below 0 °C or rose above 50 °C, the temperature could be held steady for the first

billion years or so of the history of the Earth. Then, other processes kicked in, notably the emergence of forms of life that released oxygen into the air, where it would react with methane to remove that component from the network, and the gradual decline of carbon dioxide concentrations over the aeons. It can all be made to work in a plausible fashion, but critics of the approach pointed out that it depends a lot on hindsight, and that it was far from clear where the links between living species and the physical environment operated – it seemed to many as if the constancy of the Earth’s temperature might just have been a matter of luck, not a matter of life manipulating the physical environment (through natural feedbacks, not consciously) for its own benefit. This is where Daisyworld comes in.

Daisyworld was initially developed by Lovelock and his colleagues during the early 1980s, and has since (perhaps appropriately) taken on a life of its own, with variations on the theme being developed by several people, and even being included in the computer game ‘SimEarth’ in the 1990s. It was in December 1981 that Lovelock first came up with the idea, in response to criticisms of the early versions of the Gaia hypothesis. He presented it at a conference in the Netherlands in 1982, then enlisted the help of Andrew Watson, one of his former PhD students at Reading, to collaborate on a more mathematical version of the model, which was published in the journal *Tellus* in 1983. Daisyworld starts out as a planet like the Earth, but lifeless, orbiting a star like the Sun at the same distance the Earth orbits the Sun. In the simplest versions of the model, the surface of the planet is chiefly land, to provide somewhere for daisies to grow, and the composition of the atmosphere is kept constant, so there is a constant greenhouse effect. Daisies come in two

colours, white or black, and thrive when the temperature is 20 °C. They do proportionately less well as the temperature falls below this optimum, and cannot grow below 5 °C, and proportionately less well as the temperature rises above the optimum, and cannot grow above 40 °C. The model is set running with the temperature of the model Sun slowly increasing in the way that the temperature of the real Sun did when it was young. Once the temperature at the equator of the model Earth reaches 5 °C daisy seeds of both varieties are scattered across its surface and left to their own devices – with the proviso that they breed true, with white daisies always producing white offspring and black daisies always producing black offspring.

As anyone who has ever got into a black car that has been parked in the sunlight on a summer's day knows, dark-coloured objects absorb solar heat more effectively than light-coloured objects. So a patch of black daisies will absorb heat and warm the little area of ground they inhabit, while a patch of white daisies will reflect away heat and cool the ground beneath them. While Daisyworld is cool, this means that black daisies have an advantage, since they warm their surroundings closer to the optimum temperature and thrive. In succeeding generations, black daisies spread across the surface of the planet at the expense of white daisies, so that the whole planet becomes more efficient at absorbing heat from the Sun, and its temperature increases even faster than it would simply as a result of the temperature of the Sun increasing. But once the temperature passes 20 °C at any place on the surface of the model Earth, it is the white daisies that are at an advantage, because by cooling the surface they shift things back towards the optimum temperature. Even though the temperature of the Sun

continues to increase, because white daisies now spread at the expense of black daisies, the temperature of the planet hovers around 20 °C until all the planet is covered in white daisies. Then, as the temperature of the Sun keeps going up the daisies have an increasingly difficult time, until the temperature reaches 40 °C and they all die off. The entire range of solar output covered by this version of the model is from 60 per cent to 140 per cent of the present output of our Sun.

The overall effect is that for a long period of time, even though the heat output of the model Sun is increasing, the temperature of the model Earth not only remains constant, but remains at the optimum temperature for life – without any conscious planning on the part of the daisies, and with no hint that one kind of daisy is ‘sacrificing itself for the benefit of all living things. Both varieties only ever act in their own self-interest. But can such a very simple system really be representative of how nature actually works? One criticism of the model was that it might be susceptible to ‘cheating’. Black and white daisies have to put some of their energy into making the pigment that affects the temperature of their surroundings, and you might guess that colourless daisies could thrive by invading Daisyworld and reaping the benefits of the ‘effort’ put in by the other two varieties. So Lovelock set up a model in which colourless daisies were added to the mix, and both black and white daisies were charged a ‘colour tax’ of 1 per cent to allow for making the pigment, reducing their growth efficiency. As before, black daisies flourish when solar output is low, and white daisies flourish when solar output is high, but now the colourless daisies do best when the solar output is just right to make the temperature on ‘Earth’ close to optimum without any help, the conditions when the cost of the

pigmentation gives neither of the coloured daisies a selective advantage. But the overall effect on temperature is still the same; the temperature of the model Earth stays close to optimum while the temperature of the model Sun increases dramatically.

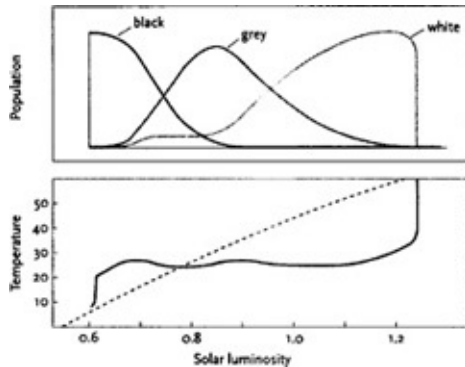


Figure 7.2 In a Daisyworld model with three species of daisy (see pp. 211–12) and a Sun that is steadily increasing in brightness, the populations of the different daisies change as time passes (top), but the temperature at the surface of the planet stays almost constant (bottom) even while the luminosity of the Sun increases dramatically. The broken line shows how the temperature would increase without the influence of the daisies.

Another criticism of the model, even with this refinement, is that it does not allow the daisies to evolve. So Lovelock made a variation on the theme, starting out with grey daisies, which have no effect on temperature, but allowing them to mutate at random to slightly darker or slightly lighter shades in each generation. Natural selection ensured that, once again, dark daisies flourished when the solar output was low, light daisies flourished when the solar output was high, and the temperature on Earth stayed almost the same for a very long time.

Daisyworld can be refined to take account of many factors, including many more varieties of daisy, the addition of animals that eat daisies, and even predators that eat the grazers. This approach links the physical

environment of the planet to the kind of predator-prey relationships described long ago by Lotka and Volterra (see [Chapter 4](#)), which is entirely appropriate since Alfred Lotka, as it happens, was one of the first people (possibly *the* first person) to appreciate that taking account of the changing relationship between living species and their physical environment ought actually to make our understanding of evolution more straightforward. In 1925 he wrote, near the beginning of his classic book *Elements of Physical Biology*:

It is customary to discuss the 'evolution of a species of organisms'. As we proceed we shall see many reasons why we should constantly take in view the evolution, as a whole, of the system (organism plus environment). It may appear at first sight as if this should prove a more complicated problem than the consideration of a part only of the system. But it will become apparent, as we proceed, that *the physical laws governing evolution in all probability take on a simpler form when referred to the system as a whole than to any part thereof.*<sup>5</sup>

If that isn't a cornerstone of Gaia theory I don't know what is – and it dates from more than three-quarters of a century ago. After combining the Lotka-Volterra model with Daisyworld, to give a scenario in which daisies are eaten by rabbits and rabbits are eaten by foxes, Lovelock then tested the robustness of his model world by introducing four catastrophes into the model – plagues which each wiped out 30 per cent of the daisy population. With all of this going on, as the figure on page 214 shows, the daisies still manage to maintain the temperature of the planet in the life zone, with only short-lived blips caused by the plagues – a form of punctuated equilibrium.

Like all models, Daisyworld is only intended to show what is possible, not to say how the world 'really is'. Clearly, the temperature of the Earth is not controlled by daisies, but by the changing infrared-absorbing properties of the atmosphere. But Daisyworld shows how a little feedback can allow (or require) life to regulate the temperature of a

planet. If the Earth got too cold for life to flourish, carbon dioxide would build up in the atmosphere and increase the greenhouse effect; as life flourishes, carbon dioxide is drawn down out of the atmosphere and a runaway greenhouse effect is avoided. What happens when all the carbon dioxide is gone and the Sun keeps getting hotter is an open question – perhaps there will be a final catastrophe for life like the eventual thermal runaway in the simple Daisyworld models, or perhaps some other mechanism will kick in to keep the runaway warming at bay, possibly by increasing cloud cover and reflecting away more of the incoming solar energy. That possibility is suggested by our other example of the way Gaia works, an example from the real world, and one which marked the turning point when Gaia became a theory rather than a mere hypothesis. For that reason, although Daisyworld is Lovelock’s proudest achievement, he regards the discovery of a mechanism linking cloud cover and biological activity in the oceans as his most important discovery.

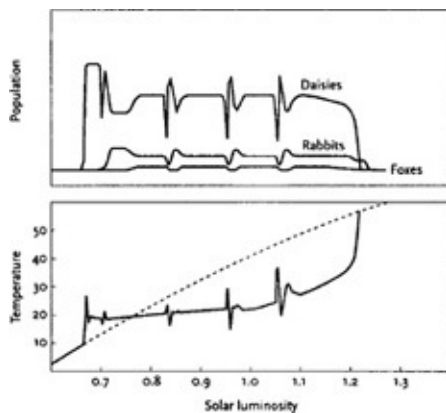


Figure 7.3 A variation on the Daisyworld theme in which rabbits eat the daisies, foxes eat rabbits, and there are four plagues that each kill 30 per cent of the daisies. All the while, the brightness of the Sun increases; yet the system proves rugged enough to keep the temperature at the surface of the planet within comfortable limits.

The difference between an hypothesis and a theory is that while an hypothesis is an idea about how things might work, it has not been tested by experiment and observation. If an hypothesis makes a prediction (or better, a series of predictions) about the way a new experiment will turn out, or what new observations will uncover, and if that prediction is borne out by events, then the hypothesis becomes a theory. Newton's theory of gravity, for example, makes predictions about the orbits of comets, and Edmond Halley used the theory to predict the return of the comet that now bears his name. It is thanks to the success of such predictions that we talk about Newton's theory of gravity, rather than his hypothesis. The distinction is often blurred; but what happened with Gaia in the second half of the 1980s is a clear example of an hypothesis coming of age and becoming a theory.

Very early on in the development of his ideas about Gaia (indeed, before William Golding even gave it that name), in the early 1970s Lovelock was interested in the way sulphur is transmitted from the oceans to the land. Sulphur is an essential ingredient of life, but is constantly being lost from the continents in the run-off from rivers, in the form of sulphate compounds. Without some mechanism to return sulphur to the land, life on land would soon be in difficulties. At the time, the conventional wisdom was that sulphur was emitted into the air from the oceans in the form of hydrogen sulphide, the noxious gas responsible for the 'bad eggs' smell of schoolboy stink-bombs. This is not typically the smell you associate with the sea, and in any case, as Lovelock was well aware, hydrogen sulphide is easily broken down by reactions involving oxygen dissolved in sea-water. But he also knew that two decades earlier researchers at the University of Leeds had found that



many marine organisms release sulphur in the form of a compound known as dimethyl sulphide; and as a chemist of the old school he also knew that dimethyl sulphide, in weak concentrations in the air, has a fresh smell, like that of 'fresh fish straight from the sea, but not part of the smell of fresh fresh-water fish'. Curious to find out if dimethyl sulphide (DMS for short) really could be the main carrier of sulphur from the oceans to the land, Lovelock designed and built a sensitive instrument to measure traces of DMS in the air, and arranged for it to be carried on the research vessel *Shackleton* on one of its routine voyages from Britain to Antarctica and back. This voyage, completed in 1972, is now part of scientific history, because another of Lovelock's instruments carried on the trip revealed for the first time the way in which chlorofluorocarbons, gases produced by human activities and now known to be responsible for destroying ozone in the stratosphere, had spread through the atmosphere of our planet. But that is another story. Lovelock found the DMS, just as he had anticipated, but it was not for another ten years, until the early 1980s, that further observations and measurements showed that there is indeed enough DMS being produced in the oceans and getting into the atmosphere each year to fall in rain over land and replace the losses from sulphate run-off.

But the microscopic marine organisms that make the DMS do not do so because they 'want' to help life-forms that live on land. Like all living things, they are only interested in maximizing their own chances of survival. The kind of organisms that release DMS, marine algae, have to fight a constant battle to prevent sodium chloride (common salt) in seawater from penetrating the membranes of their cells and disrupting the chemistry of life within. One way to keep the salt out is to build up a

suitable pressure inside the cell, using a non-toxic compound that has no adverse effect on the life processes involving DNA, RNA and amino acids. Many marine algae use a compound called dimethylsulphonio propionate, built around sulphur, for this purpose. It has all the required chemical properties, and it is convenient for them to use sulphur as the key ingredient because (thanks to all that sulphate washed down off the land) there is plenty of sulphur in the sea. When the algae die or are eaten, the dimethylsulphonio propionate gets broken down, releasing DMS into the air. What, though, has all this got to do with the Gaian self-regulation of climate on Earth?

In 1986, on a visit to the University of Washington, Seattle, Lovelock was surprised to learn, from the atmospheric scientist Robert Charlson, that nobody knew how clouds formed over the oceans. It is easy enough to make rain. When warm, wet air rises by convection and cools, at a suitable altitude the moisture will condense out as large drops of water, which fall straight back down. But in order to make clouds you need to encourage the water molecules in the atmosphere to form tiny droplets, which float in the air. They can only do this if there are even tinier 'seeds' (sometimes called cloud condensation nuclei, or CCN) around which the water molecules can collect. Over land, there are always plenty of these seeds, some produced by windblown dust, others by organic activity, and even some from the pollution of the atmosphere caused by human activities. Charlson told Lovelock that samples of air from above the Pacific Ocean showed an abundance of suitable seeds, in the form of droplets of sulphuric acid and ammonium sulphate. But he and his colleagues had no idea where the sulphuric acid and the ammonium sulphate came from, until he heard Lovelock describing the

sulphur-recycling process involving DMS. DMS could be oxidized in the air to make the cloud condensation nuclei.

The importance of this contribution to the global network is clear. Clouds reflect away so much of the incoming solar energy that without any cloud cover the average surface temperature on Earth would be 35 °C, a full 20 °C higher than it actually is. Since oceans cover about 70 per cent of the surface of our planet, and the dark ocean waters are particularly good at absorbing heat from the Sun, without clouds over the oceans the world would be uncomfortably warm (an *average* of 35 °C, worldwide, would encompass much higher temperatures in the tropics). The implication of all this is that microscopic life-forms in the oceans play a key role in controlling the climate of the Earth. In a natural feedback process, if the algae became more active, cloud cover over the oceans would increase, there would be less sunlight available for photosynthesis, and biological activity would decline; but if biological activity declines, less DMS is released by the algae, fewer clouds form, so there is more sunlight available for photosynthesis and life thrives. It is just like the kind of self-organizing feedback we see at work in Daisyworld, and the tendrils of the network extend into many aspects of life on Earth.

We can give you some idea of the complexity of those links, and you can find a lot more detail in the various books about Gaia written by Jim Lovelock. A key point is that the open ocean, far from land, is essentially a desert, sparsely populated by life compared with the rich waters of the continental shelf. The reason is that far away from land there is a shortage of nutrients on which life can feed. Close to the continents, by contrast, there is always plenty of stuff for life to feed on, being carried

down to the sea by rivers. By acting as cloud condensation nuclei, DMS produced by marine algae can help the oceanic deserts to bloom (at least a little) in two ways. First, extra cloud cover has a direct effect on the local weather, making the winds more vigorous. This acts to stir up the surface layers of the sea, bringing up nutrients from deeper layers, below the zone in which photosynthesis occurs. Secondly (and probably more importantly), the clouds and rain affect windblown dust from the continents, which is carried high in the atmosphere over even the most remote oceans of the world. Dust from the Sahara, for example, has been found over the West Indies, and dust from the heartland of the Asian continent routinely gets carried across the Pacific to Hawaii. This dust is rich in nutrient minerals that are essential for life, but the dust particles do not have the right physical properties to act as cloud condensation nuclei. Without clouds, they would simply stay in the high atmosphere, because they would not be affected by the kind of cloudless rain produced by evaporation alone. But in the clouds produced as a result of the presence of DMS in the atmosphere, the rainfall washes the dust out of the air and into the oceans, where it can be used by the very algae that make the DMS. So the link between oceans and land resulting from the production of DMS by marine algae operates as a two-way street, to the benefit of life in both places. Much of the sulphur in the D M S gets rained out over land, where it acts as a fertilizer for land-based life; but some is responsible for the clouds and rain which deposit nutrients from the land on to the oceans and act as fertilizer for ocean-based life. But nowhere is there any suggestion that life in one place is acting in a self-sacrificing manner for the benefit of life in the other place. Each does what is best for itself. This is just the kind of interaction that we have

seen to be important in self-organizing networks.

But even this only explains an existing situation. The great thing about the DMS work is that it also makes a prediction about what the world would be like when a change is made to the overall situation, and that prediction matches our understanding of what the world was like during the latest series of Ice Ages.

Over the past few million years, Ice Ages have followed a fairly regular rhythm, in which Ice Ages roughly 100,000 years long are separated by intervals of relative warmth, called Interglacials, which last for about 10,000 to 15,000 years. We live in an Interglacial that began about 10,000 years ago. The detailed pattern of these Ice Age changes exactly matches the detailed pattern of the way in which changes in the balance of the heat between the seasons occur as a result of the way the Earth tilts and wobbles as it orbits the Sun. The total amount of heat received from the Sun each year is constant, but sometimes there is a stronger contrast between the seasons (hot summers and cold winters), and sometimes there is less contrast (cool summers and mild winters). The pattern seems to be that Interglacials only begin when summers in the Northern Hemisphere (where most of the land is today) are at their hottest.<sup>6</sup> But these changes are not in themselves large enough to explain the change from Ice Age to Interglacial and back again. There must be some other process which is triggered by the astronomical rhythms, and which then acts as a positive feedback to enhance the astronomical influence. The obvious candidate is the amount of carbon dioxide in the air.

We know that when there is less carbon dioxide in the air the

greenhouse effect is reduced and the Earth cools – that lies at the very heart of Gaia theory. And we also know that during the most recent Ice Age, not only was the overall concentration of carbon dioxide in the air lower than it is today, but that the concentration fluctuated over the millennia exactly in line with variations in the average temperature of the globe. We know because there are exact records of both the temperature variations and the carbon dioxide fluctuations available, from long ice cores drilled through the ice sheets of Greenland and Antarctica. There are now several of these, but the archetypal example comes from the Russian Antarctic base Vostok, where researchers have obtained a continuous core of ice covering more than 160,000 years, enough to reveal details of the entire latest Ice Age. Drilling began in 1980, and produced a core 2.2 kilometres long, containing ice laid down, year by year, in layers of snow that were then turned into ice by the weight of new snow falling on top of them. The ice at the bottom of the core was produced from snow which fell more than 160,000 years ago, and has been compressed so much that at that level a year's accumulation of snow has been squeezed into a layer of ice one centimetre thick.

Successive layers of ice in the core can be dated by standard geological techniques. As well as the ice itself, the core contains bubbles of air trapped when the snowflakes were compressed into ice, and the contents of those air bubbles can be analysed to reveal how the composition of the atmosphere has changed over the latest Ice Age—Interglacial cycle. At the same time, the water from the ice itself can be analysed to tell us how the temperature has changed over the same interval. There are several ways to do this, all using variations on the

same technique. For example, water molecules contain hydrogen, and hydrogen comes in two varieties (two isotopes), the common form and so-called heavy hydrogen, or deuterium. Heavy hydrogen is literally heavier than common hydrogen, but the two forms are the same chemically. So water molecules come in two forms, some heavier than the common variety. The heavier molecules of water evaporate less readily than common water, but freeze out of water vapour in the air more easily; the exact proportion of each variety falling in snow depends on the average temperature in that part of the world at the time, and so the temperature long ago can be inferred from measurements of the proportion of heavy water present in the ice from deep down the core. Similar techniques involving different isotopes of oxygen, obtained both from ice cores and the shells of long-dead organisms buried in deep ocean sediments, give a broader picture of how global temperatures were changing over the same period. The records show that during the coldest period both of the latest Ice Age and the Ice Age before that temperatures worldwide averaged 9 °C lower than today, while during the warmest years of the Interglacial separating those two Ice Ages, temperatures were 2 °C higher than they are today. At the end of each of the latest two Ice Ages, the concentration of carbon dioxide in the atmosphere rose from 190 parts per million to 280 parts per million, an increase of 47 per cent, with a comparable decrease at the beginning of the latest Ice Age. And, as the figure above shows, carbon dioxide fluctuations and temperature fluctuations marched in step right through the interval covered by the Vostok core. The question is, what triggered the change in carbon dioxide concentration which contributed so much to the Ice Age?

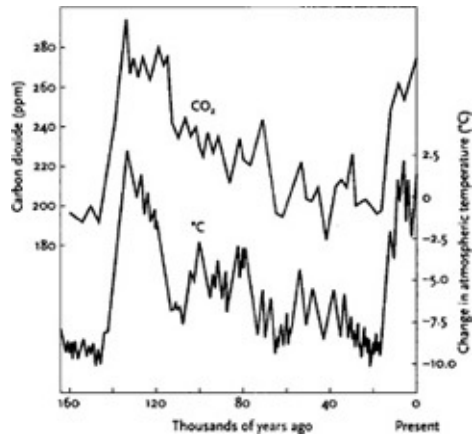


Figure 7.4 As the Earth shifts into or out of an interglacial, temperature variations and changes in the concentration of carbon dioxide in the air move in step. (Data from the Vostok core.)

The obvious possibility is that an increase in the biological activity of the oceans, which cover two-thirds of the surface of the globe, would draw carbon dioxide out of the air and lock it up in the form of carbon compounds in the shells of those photosynthesizing organisms (forms of plankton) which fall to the floor of the ocean when those organisms die. In the 1980s, John Martin and Steve Fitzwater, of the Moss Landing Marine Laboratories in California, followed up a hunch by Martin that iron might play a key role in fertilizing the oceans. The waters of the Antarctic Ocean and the sub-Arctic Pacific Ocean are, they knew, rich in important plant nutrients such as phosphates and nitrates. These nutrients are essential for plant growth, and are usually swiftly taken up by plants; but the plankton that live in those areas were just not spreading and taking them up. Clearly, the plankton were lacking something else, which prevented them from growing and taking advantage of those nutrients. Iron seemed a good bet, because iron is a crucial component of chlorophyll, the green pigment responsible for absorbing the light used in photosynthesis. Sure enough, when Martin



and Fitzwater took samples of water from the Northeast Pacific and added iron (in the form of dissolved iron compounds), there was a rapid growth of plankton, which took up the iron and the other nutrients available in the water. The studies showed that what is stopping the plankton from increasing in the cold, nutrient-rich waters at high latitudes is indeed a shortage of iron. Since then, these laboratory-scale experiments have been followed up by large-scale trials at sea, where iron compounds have been added to the oceans themselves, producing sometimes dramatic blooms of plankton. There are even suggestions that such plankton blooms might be encouraged on a large scale, to take carbon dioxide out of the air and reduce the impact of the anthropogenic greenhouse effect,<sup>7</sup> but that has nothing to do with our present story.

The relevant point is that windblown dust from the continents could provide a vital source of iron to fertilize the oceans in these areas, and that, because there is less rainfall when the Earth cools, the amount of such iron-bearing dust is likely to increase when some external influence (such as the astronomical rhythms) tilts the heat balance of the Earth in favour of an Ice Age.

The scenario is that the onset of cooling, triggered by the astronomical influences (or even by some other effect), leads to drier conditions over land at high altitudes, because when the world cools there is less evaporation from the oceans and therefore less rainfall. Dry conditions allow windblown dust from the continents, carrying iron compounds, to spread over the oceans in high latitudes, where (thanks to the DMS from marine algae), it is rained out, and fertilizes the seas. The increase in biological activity that results draws carbon dioxide down out of the air, reducing the greenhouse effect and encouraging

more cooling, thereby allowing more dust to be blown off the land and to fertilize the oceans. The process stops when all the free phosphate and nitrate has been used up. At the same time, with more DMS being produced, there will be more clouds over the oceans, which will reflect away more of the incoming solar energy, and so encourage the cooling. On this picture it is easy to see how the world can go from an Interglacial into a full Ice Age as a result of a fairly small external trigger. Getting out of the Ice Age might then occur when the astronomical influences conspire to produce maximum heating at high northern latitudes, melting back some of the ice and damping down the dry land around the fringes of the ice cap. With less windblown dust as a result, there will be less iron for the plankton, and the feedback will go into reverse, with life in the oceans going into decline as the ice caps retreat more and more, and life on land flourishes as the continents become wetter.

This is only a simple cartoon version of what goes on, but it is enough to highlight the prediction made by the DMS model. If Ice Ages are associated with an increase in the biological activity of the oceans in this way, there must be a lot more DMS getting into the air during an Ice Age. One of the products of DMS in the air is methanesulphonic acid, or MSA. So a prediction of the Gaian model developed by Lovelock and his colleagues is that there should be more MSA in the atmosphere, and therefore more MSA falling in rain or snow, during an Ice Age. The ice cores show that there was, in fact, between two and five times as much MSA falling in snow over Antarctica each year during the latest Ice Age as there is today. The same ice cores also show that there was more dust falling with the snow then than there is today. Everything fits. This is

compelling evidence that the biological and physical components of our planet are part of a single network which operates in a self-organized way (Lovelock uses the expression self-regulating) to maintain conditions that are broadly suitable for the existence of life, but which undergoes fluctuations on all scales (including Ice Age-Interglacial rhythms, and mass extinctions) analogous to the fluctuations that occur in the sandpile model. In a real sense, the Earth is home to a single living network, and the existence of that network (Gaia) would be easily apparent to any intelligent life on Mars capable of applying the Lovelock test and looking for signs of entropy reduction.

Neither NASA nor anyone else ever took the Lovelock test seriously enough to apply it to the search for life in the Solar System; but the idea is now being taken seriously enough to form the basis of the search for life beyond the Solar System. This search has been given impetus by the realization that if there is life beyond, it is very likely to be life more or less as we know it, made from the same fundamental (and simple) building blocks working together in the same kind of networks that we see here on Earth.

The story of life in the Universe is another example of surface complexity built upon foundations of deep simplicity. There is now compelling evidence that the Universe as we know it emerged from a hot dense state (the Big Bang) some 14 billion years ago.<sup>8</sup> The basic building blocks that emerged from the Big Bang were hydrogen and helium, almost exactly in the proportions 3:1. All of the other chemical elements (except for tiny traces of a few very light elements, such as lithium) have been manufactured inside stars and scattered through space when those stars swell up and eject material (in some cases they explode) in the later

stages of their lives. A star like the Sun generates heat by converting hydrogen into helium in its heart; in other stars the key processes involve successive fusions of helium nuclei together. Since each helium nucleus is a unit containing four 'nucleons' (two protons and two neutrons), denoted by the shorthand helium-4, this means that elements whose nuclei are made up of multiples of four nucleons are relatively common in the Universe, with the exception of unstable beryllium-8. In particular, carbon-12 and oxygen-16 are produced in early stages of this process, and it happens that nitrogen-14, although not containing a whole number of helium-4 nuclei, is produced as a by-product of a series of interactions involving oxygen and carbon nuclei which operate in stars a little more massive than our Sun. As a result, these are by far the most common elements apart from hydrogen and helium. Since helium is an inert gas which does not react chemically, this means that the four most common reactive elements in the Universe are carbon, hydrogen, oxygen and nitrogen, collectively known by the acronym CHON.<sup>9</sup> It is no coincidence that the four chemical elements which overwhelmingly contribute to the composition of living things on Earth are carbon, hydrogen, oxygen and nitrogen. Carbon plays the key role in life, because a single carbon atom is able to combine chemically with as many as four other atoms at once (including other carbon atoms, which may themselves be linked to yet more carbon atoms in rings and chains), so that it has an unusually rich chemistry. In the science-fiction world of space operas such as *Star Trek*, it is common for our kind of life to be referred to as 'carbon-based life-forms', with the implication that there might be other forms of life. There *might* be; but all the evidence from astronomy suggests that it is overwhelmingly more likely that life

beyond is also based on CHON.

Much of that evidence comes from spectroscopic analysis of the material present in clouds of gas and dust in space – the kind of clouds from which planetary systems like our Solar System form, but which are themselves laced with ejected material from previous generations of stars. These clouds contain many compounds built around carbon atoms, and carbon is so important for life as we know it that carbon compounds are generally referred to as ‘organic’ compounds. The compounds detected in interstellar clouds include fairly simple substances such as methane and carbon dioxide, but also much more complex organic material including formaldehyde, ethyl alcohol, and even at least one amino acid, glycine. This is an eye-opening discovery, because any material that exists in interstellar clouds is likely to have been present in the cloud from which our Solar System formed, about five billion years ago. Inspired by these observations, two teams of scientists carried out experiments here on Earth in which the kind of raw materials known to exist in interstellar space, in the form of mixtures of various ices, were kept in sealed containers under the kind of conditions that exist in space, exposed to ultraviolet radiation at temperatures below 15 Kelvin (less than *minus* 258 degrees on the Celsius scale). Their results were announced in the spring of 2002. Starting with a mixture of water, methanol, ammonia and hydrogen cyanide, one team found that three amino acids (glycine, serine and alanine) arose spontaneously in the containers. In the other experiment, using a slightly different mixture of ingredients, no less than sixteen amino acids and several other organic compounds were produced under the conditions that exist between the stars.<sup>10</sup> To put that in perspective, the proteins of all living things on

Earth are composed of various combinations of just twenty amino acids. All the evidence suggests that this kind of material would have rained down upon the young planets in the early stages of the formation of the planetary system, deposited by comets swept up by the gravitational influence of the growing planets. As we saw earlier, a broth of amino acids has the capacity to organize itself into a network with all the properties of life. The implication is that amino acids formed over very long periods of time in the depths of space (using energy from starlight) will be brought down to the surface of *any* young planet like the Earth. Some may be too hot for life, others too cold. But some, like the Earth itself (and like Baby Bear's porridge in the Goldilocks story) will be just right. There, in some 'warm little pond', to borrow Charles Darwin's expression, they get the opportunity to organize themselves into living systems.

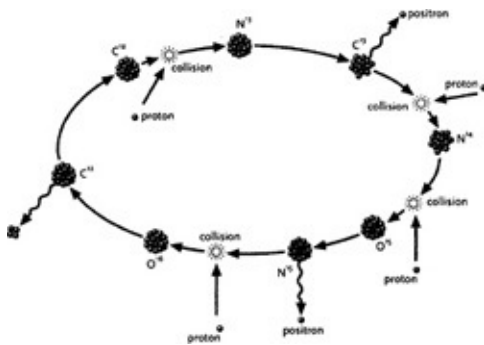


Figure 7.5 *The cycle of nuclear interactions, going on inside stars slightly more massive than our Sun, which links carbon, nitrogen and oxygen— together with hydrogen, the most important elements of life.*

Until we knew that there were other planetary systems in our Milky Way Galaxy, even this possibility could not be taken as evidence that planetary life is common in the Milky Way. But now, more than a hundred stars in our part of the Galaxy are known to have planets in orbit around them. Almost all the planets discovered so far are gas giants

orbit around them. Almost all the planets discovered so far are gas giants like Jupiter and Saturn (as you would expect, large planets have been found first, because they are easier to detect than small planets), but it seems difficult to avoid the conclusion that if other Jupiters exist then other Earths probably exist as well.

The relationship between stars, which convert simpler elements into things like CHON and spew those materials out into space, and the clouds of gas and dust in space, which collapse to form new stars, has been investigated by Lee Smolin of the University of Waterloo, Ontario. Our home in space, the Milky Way Galaxy, is one of several hundred billion similar 'islands' scattered across the visible Universe, and it seems to be completely typical – average in size, chemical composition, and so on. The Milky Way forms a flat disc, about a hundred thousand light years across, made up of a few hundred billion stars which orbit around the centre of the disc. The Sun (an entirely undistinguished member of this crowd of hundreds of billions of stars) orbits about two-thirds of the way out from the centre. At the centre of the Milky Way there is a bulge of stars, so that from the outside it would look rather like a huge fried egg, with the bulge representing the yolk. But the way in which the disc rotates reveals that all of the bright material that makes up the visible Milky Way Galaxy is held in the gravitational grip of about ten times as much dark matter, spread out in a halo around the Milky Way and extending far beyond the edge of the disc of bright stars. Finding out just what this dark matter is is of key interest to astronomers, but plays no further part in our story.

Many disc galaxies are marked by streamers of bright stars which spiral outwards from the centre, giving them the alternative name spiral

galaxies. It is easy to study the patterns made by these spiral arms, as they are called, because galaxies are relatively close to one another, compared to their size. The nearest comparable spiral galaxy to the Milky Way, the Andromeda galaxy, is a couple of million light years away from us; it sounds a great distance, but the Andromeda galaxy is so big (a little bigger than the Milky Way Galaxy) that even at that distance it covers a patch of sky as big as the Moon, as seen from Earth, and can just be seen with the naked eye on a clear, moonless night far away from cities and other sources of light pollution.<sup>11</sup> From mapping the positions of the stars in the Milky Way, astronomers know that our own Galaxy is also a spiral – and they know, as we have hinted, that there is a lot of gas and dust in clouds between the stars, so that a proper understanding of what a spiral galaxy is and how it works requires an understanding of the way energy and matter are exchanged, in a two-way process, between the stars and this interstellar medium. We also need to appreciate the timescales that are important to a galaxy. Our Milky Way is already about ten billion years old, and the Sun takes about 250 million years to orbit the centre once. The processes which exchange matter and energy between interstellar clouds and stars may seem slow by human standards, yet still be rapid by the standards of the Galaxy itself.

Another key point is that stars themselves come in different sizes – more importantly, different masses. The bigger (more massive) a star is, the more vigorously it has to use up its supply of nuclear fuel (converting hydrogen into helium, and so on) in order to hold itself up against its own weight. This makes it very bright, but subject to an early burnout. Although the Sun has a lifetime of about ten billion years in



this stable phase of its life, a star with twice as much mass can only hold itself up for a quarter as long, and a star with thirty times as much mass as the Sun will live for only about ten million years, shining as brightly as 30,000 suns, before its nuclear fuel is exhausted. Then it will collapse inward, releasing a huge amount of gravitational energy, which turns it inside out and blasts the entire star apart, as a supernova. But such dramatic events are rare, because such massive stars are rare; the processes by which stars form seem to favour the production of large numbers of smallish stars more or less like the Sun. Indeed, on average there are just a couple of supernova explosions per century in a galaxy like the Milky Way – but to put that in a suitable context for galactic processes, that still adds up to 20,000 or so every million years.

The spiral arms which are such a prominent feature of galaxies like our own are visible because they are lined with hot, massive stars which shine brightly. This means that they are also young stars, since there are no old massive stars. Since it typically takes a star a hundred million years or so to orbit around the Galaxy once, and the bright stars that edge spiral arms shine brightly for only a few million, maybe ten million, years, it follows that these stars formed more or less where we see them now. The bright spiral arms do not indicate regions where stars are particularly common, but only regions where stars are particularly bright – Smolin says they stand out ‘like lights on a Christmas tree’.

There is no mystery about how the spiral pattern is maintained. It is all thanks to feedback. The giant clouds from which new stars form may contain as much as a million times the mass of the Sun when they begin to collapse to make stars. Each collapsing cloud produces not a single large star, but a whole cluster of them, as well as many lesser stars. As

the bright stars light up, the energy of their starlight (especially in the ultraviolet part of the spectrum) makes a bubble within the cloud, and tends to stop any more stars forming. But after the massive stars have run through their lifecycles and exploded, as well as seeding the interstellar material with elements of many kinds, the blast wave gives a squeeze to the nearby interstellar clouds and makes them begin to collapse. Such waves from different supernovae, criss-crossing one another, interact to sweep up interstellar material and make new clouds of collapsing gas and dust, which give birth to more stars and supernovae, in a classic example of a self-sustaining interaction involving an input of energy (from supernovae) and feedback. Computer simulations show that there is a favoured density of cloud for the self-sustaining process to continue, and that the feedbacks naturally move the process towards these optimum conditions. If the cloud is too dense, its inner region will collapse very quickly to form a few large stars which swiftly run through their lifecycles and blow the cloud apart before many stars can form. This means that the next generation of stars is born from a thinner cloud, because there have been few supernovae to sweep the material up into dense patches. If it is thinner than the optimum density, many stars will be born, and there will be many supernova explosions, which produce many Shockwaves that sweep the interstellar material up into denser clouds. From either side, the feedbacks operate to maintain a roughly constant balance between the density of the clouds and the number of supernovae (and sun-like stars) being produced in each generation. The spiral pattern itself results from the fact that the galaxy is rotating, and that it is held in the gravitational grip of all the dark matter – analogous to the way the spiral pattern

made by cream poured into your coffee results from the fact that it is rotating and held in the grip of the dark coffee. In fact, the spiral pattern moves around the galaxy at a speed of about 30 kilometres per second, while the stars and clouds of gas and dust that make up the bulk of the disc of the galaxy move at speeds of about 250 kilometres per second, overtaking the spiral arms and getting squeezed as they pass through them. Every hundred million years or so, *everything* in the disc of the galaxy is subjected to this squeeze twice, once on each side of the galaxy, as it orbits the centre.<sup>12</sup>

Away from all this activity, in the rest of a galaxy like the Milky Way there are all the quieter, longer-lived stars like the Sun, and at least 15 per cent as much material in the form of gas and dust in the disc of the galaxy as there is in the form of stars. This interstellar material comes in different varieties. There are the clouds of cold gas and dust, rich in interesting molecules and dubbed giant molecular clouds, from which new stars (and planets) form. There are clouds of what we would think of as 'normal' gas, made of atoms and molecules of things like hydrogen, and perhaps as warm as the room you are sitting in. And there are regions that have been superheated by the energy from stellar explosions, so that electrons have been ripped from their atoms to form an electrically charged plasma. There is also a great variety of densities within the interstellar medium. At its thinnest, the stuff between the stars is so sparse that there is just one atom in every thousand cubic centimetres of space; at its most dense, clouds that are about to give birth to new stars and planets may contain a million atoms per cubic centimetre. That is still very dilute compared with, say, the air that we breathe, where each cubic centimetre contains more than ten billion

billion molecules, but a difference of a factor of a billion in density is still a spectacular contrast. The point which Smolin and a few other researchers picked up on at the end of the 1990s is that in all these respects – composition, temperature and density – the interstellar medium is far from being uniform. It is most emphatically not in equilibrium, and it seems that it is being maintained far from equilibrium by processes associated with the production of the spiral pattern.

This means that the Milky Way (like other spiral galaxies) is a region of entropy reduction. It is a self-organizing system maintained far from equilibrium by a flow of energy through the system, and, as we have seen, by feedback. At this level, the Galaxy passes the Lovelock test for life, and Smolin has argued that galaxies should be regarded as living systems. A more cautious approach is to point out that the Lovelock test constitutes what is called a ‘necessary but not sufficient’ criterion for the existence of life. If a system is in thermodynamic equilibrium – if it fails the Lovelock test – then we know for sure that it is dead. If it is alive, it must produce an entropy reduction and pass the Lovelock test. But a system might produce negative entropy without being alive, as in the case of gravitational collapse that we discussed earlier. There is no clear boundary between living and non-living things, from this perspective. A sandpile on which grains are dribbled from above maintains itself in a critical state, feeding off outside energy, but is certainly not alive; a human being certainly is alive; people still argue about whether Gaia can be regarded as a single living system; as for galaxies, this kind of investigation has only just begun, and I wouldn’t like to bet on how the consensus will turn out. The very fact that the boundary between life

and non-life is blurred, and where to draw the boundary is a subject for discussion, is, though, an important discovery. It helps to make clear that there is nothing unusual about life in the context of the way the Universe works. As we have seen, it is natural for simple systems to organize themselves into networks at the edge of chaos, and once they do so it is natural for life to emerge wherever there is a suitable 'warm little pond'. It is part of a more or less continuous process, with no sudden leap where life begins. From that perspective, the most important thing that science could achieve would be to discover at least one other planet where life has emerged. Thanks to Lovelock's insight about the nature of life, we are on the edge of being able to do just that, and telescopes capable of detecting other Gaias could be launched into space within the next twenty or thirty years.

There are two stages to discovering other Gaias. First, we have to detect Earth-sized planets in orbit around other stars; then, we have to analyse the atmospheres of those planets for evidence of entropy-reducing processes at work. The first 'extra-Solar' planets were detected using Doppler techniques, which reveal tiny changes in the motion of the stars those planets orbit around. The effect, named after the nineteenth-century physicist Christian Doppler, shifts the position of lines in the spectrum of light from an object by an amount which depends on how fast the object is moving relative to the observer. To put these kinds of observations in perspective, the gravitational pull of Jupiter, tugging on the Sun, produces a velocity shift of about 12.5 metres per second, and moves the Sun bodily (relative to the centre of mass of the Solar System) across a distance of 800,000 kilometres, just over half the solar diameter, as the Sun and Jupiter orbit around their

mutual centre of mass. The speed of this motion is comparable to that of an Olympic 100-metres sprinter, and to an observer outside the Solar System this produces a tiny to-and-fro shift in the exact position of the lines in the spectrum of light emitted by the Sun, by the Doppler effect. This is the kind of shift that has been detected in the light from scores of stars in our neighbourhood, showing that they are orbited by Jupiter-like companions. For comparison, the Earth induces a velocity shift of just 1 metre per second (a gentle walking pace) in the Sun as it orbits around it, and moves the Sun by only 450 kilometres, relative to the centre of mass of the Solar System. We do not yet have the technology to measure such a small effect at distances corresponding to those of even the nearer stars, which is why no other Earths have yet been detected by this method.

There are other techniques which might work to identify small planets, and which are sometimes in the news. For example, if the planet happens to pass directly in front of its parent star (an occultation, or transit), producing a regular dimming of the light from that star. Statistically, because the orbits of extra-Solar planets could be inclined in any direction relative to us, only 1 per cent of those planets will be in orbits suitable for us to see occultations, and in any case each transit lasts for just a few hours (once a year for a planet in an Earth-like orbit; once every eleven years for a planet in a Jupiter-like orbit). Nevertheless, there are plans to launch satellites into space in the next few years (perhaps as early as 2005) in order to monitor large numbers of stars to search for such occultations. If 100,000 stars are studied, and 1,000 of them show transits, the statistics would imply that virtually every Sun-like star is accompanied by planets. But although all searches

of this kind are invaluable, the Doppler technique is the one that can be applied most generally to the search for other Earths. And however other Earths are found, the next stage in the search for other Gaias will be the same.

The best immediate prospect of finding Earth-like (at least, Earth-sized) planets in large numbers is a NASA satellite called SIM, from the name Space Interferometry Mission. This could be launched within the next couple of years (by 2005), and will use a technique known as interferometry, in which data from several small telescopes are combined to mimic the observing power of one much larger telescope. If all goes well, SIM will be able to measure the positions of stars so accurately that it will detect the wobbles caused by any Earth-like planets orbiting any of the 200 stars closest to the Sun, and any Jupiter-like planets out to a distance of 3,000 light years from the Sun. Around the end of the first decade of the twenty-first century (again, if all goes well), the European Space Agency will launch a satellite with the slightly misleading name GAIA, whose main mission is not actually to look for other Gaias but to map the positions of the billion brightest objects in the sky. Because GAIA will be looking at so many stars, it will not look at each one of them very often or for very long, so it will not be able to detect wobbles caused by Earth-sized planets; but it will be able to detect Jupiter-sized planets in orbits with periods longer than about a couple of years. If planets are as common as the first indications from ground-based telescopes suggest, then within ten years we should have identified tens of thousands of extra-Solar planetary systems in our neighbourhood of the Milky Way Galaxy. But these would still only be indirect observations, and in order to take the spectra of some of those

planets a further technological leap is needed.

Both NASA and ESA are working on next-generation projects that could use that technology to make those observations by 2030. Because of the huge cost of any such mission, however, it seems likely that at some stage the efforts will be merged into one truly global project. It would be a collaboration among all the relevant experts on Earth to search for evidence that they (and we) are not alone in the Universe – Gaia as a whole looking for other Gaias. The ESA project has been known as Project Darwin, but is also referred to, more prosaically, as the Infrared Space Interferometer (IRSI); the NASA equivalent is known as the Terrestrial Planet Finder, or TPF. But both would work on exactly the same principles.

Surprising though it may seem, especially after seeing those images from space of the Earth as a bright blue and white ball against a dark background, visible light does not offer the best prospects for direct detection of other Earths. There are two reasons. First, the visible light from a planet like the Earth is essentially the reflected light from its parent star, so it is not only relatively weak, but very difficult to pick out at astronomical distances against the glare of the light from the star itself. Secondly, terrestrial planets are actually brightest in the infrared part of the electromagnetic spectrum, because of the way the energy they absorb from the Sun is re-radiated in the infrared part of the spectrum, with wavelengths longer than those of visible light. At a wavelength of a few microns, Earth is the brightest planet in the Solar System, and would stand out as a striking object to any suitably sensitive infrared telescope in our stellar neighbourhood. The catch is that because infrared radiation is absorbed by the very gases in the Earth's



atmosphere, such as carbon dioxide and water vapour, that we are interested in finding, the telescope to search for other Earths will have to be placed in deep space, far away from any potential sources of pollution. It will also have to be very sensitive, which means very large. Which is why we are talking about an expensive, international project that will take decades to come to fruition.

The plan is for a much more powerful interferometer-based telescope than SIM. Like all interferometers, this will require that the signals from the different telescopes be added together with great accuracy. When the technique was originally developed on the ground, using radio telescopes, the signals were brought together using wires; now, it is routine to use laser beams in this kind of work, and that technology will be essential for the planet finder telescope, whatever name it eventually flies under. Each of the two proposals now on the drawing board is for a group of six satellites (six telescopes) that will fly in formation, at least 100 metres apart, at the corners of a hexagon. The relative positions of the spacecraft will have to be measured to an accuracy of better than one millimetre (again, using lasers), while the information from the telescopes is combined into one signal in a central master satellite, and beamed back to Earth, perhaps 600 million kilometres away. Such an instrument will be able to detect the infrared emissions from Earth-like planets (or rather, planets like Earth, Venus and Mars, in comparable orbits to those of Earth, Venus and Mars) orbiting around a few hundred stars out to a distance of about fifty light years. The next step is to take the spectra of those planets – exactly the step that Jim Lovelock realized was the best way to search for life on Mars.

All of this takes time, and the more planets you look at in each stage

of the search the less time you can spend looking at each planet. So, for a planned six-year mission, the first two years would be spent looking at as many stars as possible to find any sign of infrared emission from planets. Over the next two years, about eighty of the best candidates would be studied in more detail, with about 200 hours of observing time devoted to each target, to search for the prominent spectral features in the infrared associated with carbon dioxide, and the slightly less prominent features associated with water vapour. The presence of these gases alone is not a sign of life, but it would be a sign of the existence of planets that were terrestrial in the sense of having an atmosphere like Venus and Mars, while water in particular would indicate a probable home for life. From those eighty candidates, in the final stage of the mission twenty targets would be selected to study for about 800 hours each, the time required to bring out traces of the most exciting potential feature of this part of the spectrum – infrared features of oxygen. These would not be from the kind of ordinary, di-atomic molecules of oxygen ( $O_2$ ) that we breathe, because di-atomic oxygen does not emit or absorb in this part of the spectrum. But any terrestrial planet which, like the Earth, has an oxygen-rich atmosphere will also have an ozone layer, produced by the action of the light from its parent star on di-atomic oxygen to make tri-atomic ozone ( $O^3$ ). And ozone produces a sharp feature in the infrared spectrum, as it happens smack between the features associated with carbon dioxide and water vapour. The presence of this feature alone (or the presence of features indicating other active compounds, such as methane) would be enough to tell us that the atmosphere of a planet orbiting a star several tens of light years away is not in thermodynamic equilibrium, and that entropy reduction processes

– in other words, life – are at work on its surface. And all without any space probe, let alone a human being, even leaving the Solar System.

This is the most striking example we know of the deep simplicity which underpins the Universe. The most complex things in the known Universe are living creatures, such as ourselves. These complex systems are made from the most common raw materials known to exist in galaxies like the Milky Way. In the form of amino acids, those raw materials naturally assemble themselves into self-organizing systems, where simple underlying causes can produce surface complexity, as in the case of the leopard and its spots. Finally, in order to detect the presence of these most complex of universal systems you do not need any sophisticated tests to distinguish living from non-living, but only the simplest techniques (albeit aided by highly advanced technology) to identify the presence of one of the simplest compounds in the Universe, oxygen. Chaos and complexity combine to make the Universe a very orderly place, just right for life-forms like us. As Stuart Kauffman has put it, we are ‘at home in the Universe’. But it isn’t that the Universe has been designed for our benefit. Rather, we are made in the image of the Universe itself.

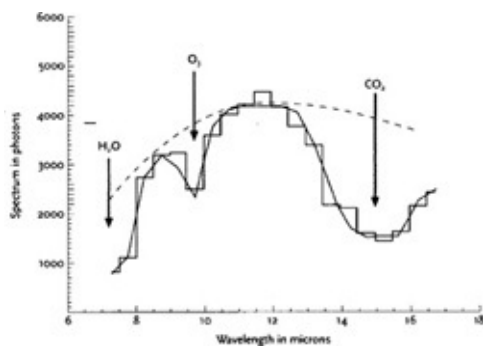


Figure 7.6 A simulation of what the spectrum of the Earth would look like to the Darwin telescope at a distance of thirty light years. If instruments like Darwin find spectra like this from planets orbiting other

*stars, it will show that entropy-reducing processes are at work on those planets, and indicate the probable presence of life beyond the Solar System.*

## Further Reading\*

Robert Alexrod, *The Evolution of Cooperation* (Basic Books, New York, 1984) Per Bak, *How Nature Works*(Oxford University Press, Oxford, 1997)

Philip Ball, *The Self-Made Tapestry*(Oxford University Press, Oxford, 1999) Michael Boulter, *Extinction*(Fourth Estate, London, 2002)

Mark Buchanan, *Ubiquity*(Weidenfeld & Nicolson, London, 2000)

Lewis Carroll, *Through the Looking-Glass and What Alice Found There* (Macmillan, London, 1871) Charles Darwin, *On the Origin of Species by Means of Natural Selection* (Penguin, London, 1968; reprint of the first edition of 1859) Paul Davies, *The Fifth Miracle* (Allen Lane/Penguin Press, London, 1998) Stillman Drake, *Galileo at Work* (Dover, New York, 1978)

Ivar Ekeland, *The Broken Dice* (University of Chicago Press, Illinois, 1993) J. K. Fauvel, R. Flood, M. Shortland and R. Wilson (eds), *Let Newton Be!* (Oxford University Press, Oxford, 1988) J. Feder, *Fractals* (Plenum, New York, 1988)

Michael Field and Martin Golubitsky, *Symmetry in Chaos* (Oxford University Press, Oxford, 1992) R. A. Fisher, *The Genetical Theory of Natural Selection*(Clarendon Press, Oxford, 1929) George Gamow, *My World Line* (Viking, New York, 1970)

Leon Glass and Michael Mackey, *From Clocks to Chaos* (Princeton University Press, 1988) James Gleick, *Chaos* (Heinemann, London, 1988)

John Gribbin, *In Search of the Big Bang* (Penguin, London, 1998) John Gribbin, *Almost Everyone's Guide to Science* (Weidenfeld & Nicolson, London, 1998) John Gribbin, *Stardust* (Allen Lane/Penguin Press, London, 2000)

John Gribbin, *Science: A History* (Allen Lane/Penguin Press, London, 2002) John and Mary Gribbin, *Fire on Earth* (Simon & Schuster, London, 1996) Mary Gribbin and John Gribbin, *Chaos and Uncertainty* (Hodder Children's Books, London, 1999) Nina Hall (ed.), *The New Scientist Guide to Chaos* (Penguin, London, 1992) Bai-Lin Hao (ed.), *Chaos* (World Scientific, Singapore, 1984)

Andrew Hodges, *Alan Turing* (Burnett Books, London, 1983)

John Holland, *Hidden Order: How Adaptation Builds Complexity* (Addison-Wesley, Reading, Mass., 1995) John Holland, *Emergence* (Oxford University Press, Oxford, 1998)

Stuart Kauffman, *The Origins of Order* (Oxford University Press, New York, 1993) Stuart Kauffman, *At Home in the Universe* (Oxford University Press, New York, 1995) Pierre Laplace, *A Philosophical Essay on Probabilities* (Dover, New York, 1951; translation of the original *Essai philosophique sur les probabilités*, first published in English in 1819) Roger Lewin, *Complexity* (Macmillan, New York, 1993)

Edward Lorenz, *The Essence of Chaos* (University of Washington Press, 1993) Alfred Lotka, *Elements of Mathematical Biology* (Dover, New York, 1956; originally published in 1925 as *Elements of Physical Biology*)

James Lovelock, *The Ages of Gaia* (Oxford University Press, Oxford, 1988) James Lovelock, *Homage to Gaia* (Oxford University Press, Oxford,

2000) Benoit Mandelbrot, *The Fractal Geometry of Nature*(W. H. Freeman, New York, 1983; original edition published in 1977) Lynn Margulis, *The Symbiotic Planet* (Weidenfeld & Nicolson, London, 1998) Jacques Monod, *Chance and Necessity*(Knopf, New York, 1971)

James Murray, *Mathematical Biology*(Springer, New York, 1993; corrected second edition of volume first published in 1989) Heinz Pagels, *The Dreams of Reason*(Simon & Schuster, New York, 1988) Heinz-Otto Peitgen and Peter Richter, *The Beauty of Fractals*(Springer, New York, 1986) Heinz-Otto Peitgen, Hartmut Jurgens and Dietmar Saupe, *Chaos and Fractals*(Springer, New York, 1992) Ivars Peterson, *Newton's Clock* (W. H. Freeman, New York, 1993)

Henri Poincare, *Science et Methode* (Flammarion, Paris, 1908; English translation available from Dover, New York, 1952, under the title *Science and Method*) Ilya Prigogine, *The End of Certainty*(The Free Press, New York, 1997)

Ilya Prigogine and Isabelle Stengers, *Order out of Chaos*(Heinemann, London, 1984) David Raup, *Extinction: Bad Genes or Bad Luck?* (Norton, New York, 1991) Lewis Fry Richardson, *Weather Prediction by Numerical Process* (Dover, New York, 1965; reprint of the original 1922 edition) David Ruelle, *Chaotic Evolution and Strange Attractors* (Cambridge University Press, Cambridge, 1989) David Ruelle, *Chance and Chaos* (Princeton University Press, Princeton, 1991) M. Schroeder, *Fractals, Chaos, Power Laws*(W. H. Freeman, New York, 1991) S. K. Scott, *Chemical Chaos*(Oxford University Press, Oxford, 1991)

John Maynard Smith, *Evolution and the Theory of Games* (Cambridge

University Press, Cambridge, 1982) Lee Smolin, *The Life of the Cosmos* (Oxford University Press, New York, 1997) Ian Stewart, *Does God Play Dice?*(Penguin, London, 1990)

D'Arcy Thompson, *On Growth and Form* (Cambridge University Press, Cambridge, 1917) Mitchell Waldrop, *Complexity*(Simon & Schuster, New York, 1992)

Jonathan Weiner, *The Beak of the Finch*(Jonathan Cape, London, 1994)

Michael White and John Gribbin, *Darwin: A Life in Science*(Simon & Schuster, London, 1995) Shu-yu Zhang, *Bibliography on Chaos* (World Scientific, Singapore, 1991) Carl Zimmer, *Evolution*(Heinemann, London, 2002)

George Zipf, *Human Behavior and the Principle of Least Effort*(Addison-Wesley, Cambridge, Mass., 1949)



\* Several of my own books are included in this list, not because they are necessarily the best books on the relevant topic, but because they deal in complementary fashion with issues mentioned in passing in this book.

# Index

Figures in italics refer to illustrations

acceleration [7–9](#)

acoustics, power law behaviour [145, 146](#)

algebra

    Boolean logic [171](#)

    topology [99](#)

algorithms [65, 65](#)

Amaral, Luis [194](#)

amino acids [167n](#)

    interstellar clouds [224–6](#)

    self-organizing system [234](#)

*see also* [proteins](#)

Andromeda galaxy [227](#)

angel fish, reaction diffusion patterns [133–4](#)

approximation technique [39–41](#)

arrow of time [11, 18, 26, 106](#)

    direction [18, 20–21](#)

    entropy [26](#)

gravity [113](#)

irreversibility [27–8](#)

open systems [104](#)

*see also* [thermodynamics](#)

Arthur, Brian [150](#)

Asahi, Rihito [134](#)

asteroids

formation [60–61](#)

Kirkwood gaps [61–2](#)

resonant orbits [61](#)

atomic physics, Bohr model [193–4](#)

attractors [29](#)

bifurcation [74–5](#)

limit cycle [82](#)

Lorenz attractor [57, 58, 58, 100–101](#)

state cycles [170–71](#)

strange attractors [82, 89, 100–101](#)

BZ reaction [126](#)

thermodynamics [28–30](#)

torus [81–2](#)

Avogadro, Amadeo [32–3](#)

Avogadro's constant [32–3](#)

Bak, Per [158–60](#), [190–91](#)

*How Nature Works* [159–60](#)

Baker transformation, Smale's horseshoe [99–100](#), [99](#)

Bard, Jonathan B. L. [131](#)

Belousov, Boris [121–3](#)

Belousov-Zhabotinsky (BZ) reaction [124–6](#), [167](#)

Bénard, Henri [106–8](#), [209](#)

Bénard instability *see* [Rayleigh, Lord: Rayleigh-Bénard instability](#)

bifurcations *see* [period doubling](#)

Big Bang theory [111–12](#), [223](#)

Big Five catastrophes [153–4](#)

Bjerknes, Vilhelm [52–3](#)

Bohr, Niels [193–4](#)

Boltzmann, Ludwig [21](#), [31–5](#)

Boltzmann's constant [31–2](#)

*Lectures on Gas Theory* [35](#)

Boole, George [171n](#)

Boolean algebra [171](#)

Bray, William [122](#)

Brusselator model, reaction diffusion patterns [124–5](#)

Brussels School [106](#)

Buchanan, Mark [143](#)

*Ubiquity* [143](#), [14n](#)

Butterfly attractor *see* [Lorenz attractor](#)

butterfly effect [57](#)

butterfly wings, reaction diffusion patterns [134](#)

calculus [10–11](#)

Cantor, George

Cantor set [84–6](#), [85](#)

carbon, carbon-based life forms [224](#)

carbon dioxide

and the greenhouse effect [207](#)

and infrared emissions [207–8](#), [234](#)

and Interglacials [218–21](#), [220](#)

and Martian atmosphere [203–4](#)

Carroll, Lewis [181](#)

catalysis [118–19](#)

autocatalytic process [119–20](#)

emergence of life [167](#)

celestial mechanics, equations [41](#)

cells

differentiation and development [169–73](#)

DNA and gene proportionality [173](#), [173](#)

embryogenesis [170](#)

gene networks [169–70](#)

lightbulb model [170–72](#)

state cycles [172–3](#)

chaos

butterfly effect [57](#)

and fractals [85–7](#), [89](#)

initial condition determination [56–7](#)

molecular [36](#)

period-doubling route [72](#)

pocket calculator model [58–9](#)

random process [86–7](#)

restricted, Earth orbital changes [64](#)

turbulence model [70–71](#)

Universe complexity [39–69](#)

chaos game [89](#), [90](#)

and DNA [89](#), [91](#)

fractal images [89](#), [90](#), [91](#), [91](#)

Charlson, Robert [216](#)

chemical clocks [122](#)

chlorofluorocarbons, atmospheric spread [215](#)

CHON (carbon, hydrogen, oxygen and nitrogen) [223–4](#)

Clausius, Rudolf [25](#), [31](#)

cloning

DNA [170](#)

nature and nurture [131](#)

cloud condensation nuclei

cloud formation [216](#)

ocean deserts [217](#)

cloud cover

Earth's temperature [213–17](#)

formation [216](#)

ocean biological activity association [214–15](#)

coastline and border measurements

fractal dimensions [139–42](#), [141](#), [142](#)

Koch island [93–5](#)

power law [141](#), [142](#), [142](#)

complex numbers [96–7](#), [136–7](#)

universal constant 137

complex system networks 163

button model 163–5, 165

components 163–4

defined 136

and emergence of life 166–74

and genes 169–70

lightbulb model 170–72

phase transition 164–5, 167–8, 184

power law association 158

racing bike model 137

sandpile model 158–62, 162, 164–6

concept of diminishing returns 150–51

convection

bimodal form 108

hexagonal cells 108

Rayleigh number 108

Rayleigh-Bénard instability 106–9, 109

convergence, geometric 79–80

Copernicus, Nicolaus 6

*De Revolutionibus Orbium Coelestrum* 6



cosmic microwave background radiation [112](#)

Cretaceous-Triassic boundary *see* [K-T event](#)

Curie, Pierre

Curie point [117](#)

Daisyworld [206–14](#), [212](#), [214](#)

catastrophes [213](#)

grey daisies [212](#)

predator-prey relationships [212–13](#)

dark matter, Milky Way galaxy [226](#), [229](#)

Darwin, Charles [175](#)

Darwinian evolution [176–9](#)

Darwinism [132–3](#)

*On the Origin of Species by Means of Natural Selection* [176](#)

digitary orrery, Solar System [62](#)

dimethyl sulphide (DMS)

and cloud condensation nuclei [216](#)

and Ice Ages [222](#)

and sulphur recycling [215–16](#)

dinosaurs, decline and extinction [151](#), [153](#)

Dirichlet, Lejeune [41](#)

dissipative processes [104–5](#)

DNA [89](#), [91](#)

cloning [170](#)

gene proportionality [173](#), [173](#)

proteins [169](#)

Doppler, Christian [231](#)

Doppler effect, planetary systems detection [230–32](#)

dust particle, restricted three-body problem [45–6](#), [50](#)

Earth

cloud cover [213–17](#)

obliquity changes [65–6](#)

orbital changes [62–4](#)

self-regulation [204–6](#), [222–3](#)

solar energy [113–14](#)

spectrum [235](#)

surface temperature [206–9](#)

earthquakes

extinctions and emergence [136–74](#)

frequencies and sizes [138](#), [139](#), [142](#)

Gutenberg-Richter law [139](#), [140](#), [140](#)

power law 139

Richter scale 138–9

scale invariance 142

and self-organized criticality 159–63

economies

concept of diminishing returns 150–51

feedback system 150–51

human intervention 150

power law behaviour 149–51

Eddington, Arthur 122–3

*The Nature of the Physical World* 122–3

Einstein, Albert 20, 175

electromagnetism 18–19

Faraday effect 19

Maxwell's equations 19–21

quantum effects 20

embryogenesis 114–35

cell differentiation 170

embryo size 130–31

and Turing 118–21, 124, 126–8

emergence of life

complex system networks 166–74

energy

conservation and dissipation 24–5

Joule heating 24

kinetic 31

entropy 26, 28

and Earth-like planet detection 202, 234

probability relationship 32

spiral galaxies 229–30

thermodynamics 25–8

*see also* [arrow of time](#)

environment changes

biological and physical 198–9

fitness landscapes 198

equations of celestial mechanics 41

equilibrium

thermodynamic 29–30, 34–5, 203–4, 208

eumelanin, skin colours 127–8

evolutionarily stable strategy (ESS) 179–81

evolutionary biology 175–99

Bak-Sneppen model 192–3

co-evolution [183–7](#)

Identikit model [186–7](#)

Darwinism [132–3](#)

Evolution of sex [187](#)

games theory [179–81](#)

mammalian coat patterns [126–35](#), [130](#)

predators and prey [194–5](#)

Red Queen effect [181–2](#)

supermouse model [188–9](#)

extinction patterns [151–7](#), [155](#), [156](#), [157](#), [187–8](#)

Big Five catastrophes [153–4](#)

outside influences [195–6](#)

Permian period [198](#)

as power laws [156](#), [193–7](#)

scale invariance [154](#), [193–4](#)

extremal dynamics [191](#)

faint young Sun paradox [206–7](#)

Faller, Alan [78](#)

Faraday, Michael [19](#)

feedback systems

concept of diminishing returns [150–51](#)

economies [150–51](#)

galaxies [228–9](#), [230](#)

linear/non-linear regimes [124–5](#)

Feigenbaum, Mitchell [79](#)

Feigenbaum diagram [76](#), [77](#), [80](#)

Feigenbaum point [85](#)

Feigenbaum systems [79–80](#)

Feigenbaum's number [80](#), [80n](#)

finches *see* [Galápagos finches](#)

Fisher, R. A. [184](#)

fitness landscapes [184–6](#), [192–3](#)

influenced by life [198](#)

Red Queen effect [185–6](#)

Fitzwater, Steve [221](#)

Fourier, Joseph [23](#)

fractals [82–103](#)

Cantor set [84–6](#)

coastline and border dimensions [139–42](#), [141](#), [142](#)

dimension measurements [92](#), [95–6](#)

living systems [102–3](#)

Mandelbrot set [96–7](#)

relationship with chaos [85](#)

strange attractors [82, 89, 100–101](#)

Frette, Vidar [160](#)

friction [7–8](#)

fur patterns and colours *see* [mammalian coat patterns](#)

Gaia hypothesis [198–9, 206–14, 223, 230](#)

Daisyworld model [206–14](#)

*Encyclopedia Britannica* entry [205](#)

GAIA satellite [232](#)

Galápagos finches

Darwinian evolution [177, 178–9, 185](#)

And punctuated equilibrium [189, 189n](#)

galaxies

Andromeda [226–7](#)

dark matter [226, 229](#)

and feedback [228–9, 230](#)

as living systems [230](#)

Milky Way [226–8](#)

as self-organizing systems [228–30](#)

spiral [226–30](#)

Galileo Galilei

and gravity and acceleration [6–8](#)

games theory [179–81](#)

Gamow, George [110](#)

genetic engineering [169n](#)

genetic theory of natural selection [184–5](#)

geological periods, Big Five catastrophes [154](#)

giant molecular clouds [229](#)

Gibbs, Josiah Williard [31](#)

Gleick, James [2](#)

global warming

and power law behaviour [145–6](#)

Golding, William [204](#)

Grant, Peter and Rosemary [178–9](#)

gravity [175](#)

and acceleration [7–9](#)

arrow of time [113](#)

constant of gravity [11](#)

general theory of relativity [175](#)

inverse square law [11, 12](#)

negative energy [109–13](#)



Newton's law of gravitation [8–10](#)

solar energy [113–14](#)

universal law [12](#)

greenhouse effect [146](#), [207–10](#)

and Daisyworld model [210](#)

Gutenberg, Reno [138](#)

Gutenberg-Richter law [139](#), [140](#), [140](#)

Hamilton, William [42](#)

Hamiltonian equations [42](#); and phase space [42–9](#)

heat death of the universe [26](#), [27](#), [34](#)

helium, Big Bang [223](#)

Holland, John [186](#)

homodinic tangle [47](#)

Horseshoe transformation [99–101](#)

Huxley, Thomas Henry [176](#), [206](#)

Ice Ages [218–22](#)

astronomical theory [218](#), [221](#)

Interglacials [218–21](#), [220](#)

infrared emissions

and carbon dioxide [207–8](#), [208](#), [234](#)

and oxygen [234](#)

from planets [207–8](#), [208](#), [233](#), [234](#)

tri-atomic ozone [234](#)

infrared space interferometer (IRSI) [232](#)

inheritance, Darwinian evolution [176](#)

initial conditions, precise determination [56–7](#)

interferometry [232](#), [233](#)

Interglacials [218–21](#)

and carbon dioxide [218–22](#), [220](#)

inverse square law [11](#), [12](#)

iron

and ocean fertilization [221](#)

iteration [14](#), [16](#), [58–9](#)

Mandelbrot set [96–7](#)

period-doubling bifurcations [79](#)

pocket calculator model [58–9](#)

Sierpinski gasket [87–8](#), [88](#), [89](#)

Jordan, Pascual [111](#)

Joule, James, Joule heating [24](#)

Jupiter

Jupiter-sized planets, GAIA project [232](#)

orbit calculation [15–16](#)

resonant asteroid orbits [61](#)

Jürgens, Hartmut

*Chaos and Fractals* [89](#)

K-T event [155](#)

dinosaur extinction [151](#)

fossil record marker [151](#)

meteorite impact [152–3](#)

Kauffman, Stuart [163–74](#), [181](#), [183–4](#)

Kelvin, Lord (formerly William Thomson)

Thomson effect [25](#)

Kepler, Johannes [6](#), [12n](#)

kinetic energy, stars [113](#)

Kirkwood, Daniel

Kirkwood gaps [61–2](#)

Koch, André [133–4](#)

Koch curve/island/snowflake [92–6](#), [93](#), [94](#)

Kondo, Shigeru [134](#)

Kronecker, Leopold [41](#)

Landau, Lev [80](#)

Laplace, Pierre [15–18](#)

*Essai philosophique sur les probabilités* [17–18](#)

laser beam technology, planet detection [233](#)

Laskar, Jacques [63](#)

law of conservation of momentum [64](#)

*Lectures on Gas Theory* (Boltzmann) [35](#)

Lefever, René [124](#)

Leibniz, Wilhelm [10](#)

Leonardo da Vinci

whirlpool turbulence model [71–2](#)

Lewin, Roger [3](#)

Li, Tien Yien [78](#)

‘Period Three Implies Chaos’ [78](#)

light, speed of [19–20](#)

light curve [145](#)

limit cycles [70–71](#), [80](#), [82](#)

linear/non-linear regimes [105–6](#)

feedback system [124–5](#)

steady state [105–6](#)

logarithms

Guten-Richter law [140](#)

log-log plot [139](#)

Richter scale [138–9](#)

logistic equation

Baker transformation [99](#)

Horseshoe transformation [99](#)

mapping [97–9](#)

population changes [72–4](#)

LONG-STOP (Long-term

Gravitational Study of the Outer Planets) [62](#)

Lorenz, Edward [54–7](#), [57](#)

chaos discovery [76–9](#)

Lorenz attractor [57](#), [58](#), [58](#), [100–101](#)

Loschmidt, Joseph [33](#), [36](#)

Loschmidt's number [33](#)

Lotka, Alfred [121–2](#), [212–13](#)

*Elements of Physical Biology* [213](#)

Lovelock, Jim [3](#), [199](#), [200–235](#)

mammalian coat patterns [126–35](#), [130](#)

surface size [129–30](#)

Mandelbrot, Benoit [83–4](#), [86–7](#), [92](#), [93](#), [94–5](#), [96](#)

economic systems and power laws [149–50](#)

‘How long is the coast of Britain?’ [95](#)

Mandelbrot set [96–7](#)

mapping, logistic equation [97–9](#)

Mars

atmospheric composition [203–4](#)

life-detection measurements [201–4](#), [208](#)

obliquity changes [65](#)

thermodynamic equilibrium [203–4](#)

Viking project [203](#)

Martin, John [221](#)

mathematical techniques

approximations [39–40](#)

deterministic [14](#)

power series [40](#)

*see also named techniques*

Maxwell, James Clerk [19–21](#)

Maxwell's equations [19–21](#)

May, Robert [75–6](#), [99](#)

Maynard Smith, John [179](#)

Meinhardt, Hans [133–4](#)

melanin, eumelanin and phaeomelanin [127–8](#)

Melosh, Jay [152](#)

Mercator projection [97](#)

Mercury, obliquity changes [65](#)

metabolic rates, power law [101–2](#)

meteorite impact, K–T event [152–3](#)

methane, greenhouse gas [209](#)

methanesulphonic acid (MSA), Ice Age [222](#)

Meyer, Martin [194](#)

Milky Way

    spiral galaxy [226–8](#)

    spiral pattern formation [229](#)

molecular chaos [36](#)

momentum

    Hamiltonian equation [42](#)

    law of conservation [64](#)

Moon

    landscape, scale invariance [143–4](#)

    orbit calculation [46](#)

    surface temperature [208–9](#)

morphogenesis [118—35](#)

motion, Newton's laws [12–17](#)

mouse, supermouse model [188–9](#)

Murray, James [126–7](#)

‘How the leopard got its spots’ [126](#)

*Mathematical Biology* [126](#)

mutation

Darwinian evolution, [176](#)

evolution [187—8](#)

Nagel, K [148n](#)

natural selection *see* [evolutionary biology](#)

neural nets [117](#)

Newman, Mark [195](#)

Newton, Isaac [175](#)

law of gravitation [8–10](#)

laws of motion [12–17](#)

*Philosophiae Naturalis Principia Mathematica* [8–10](#)

Norway, coastline fractal dimensions [139-42,141,142](#)

ocean desert, self-organizing system [217–18](#)

ocean fertilization, iron [221](#)

Ockham’s Razor [127](#)



Onsager, Lars 105

open systems

arrow of time 104

self-organized criticality 158

orbital mechanics 13–14

approximation technique 41

elliptical paths 6,11,12

three-body problem 14–17, 17, 38, 50

Oregonator model 125

oxygen, tri-atomic ozone 234

ozone, tri-atomic 234

Peano, Giuseppe 82–4

Peano curve 83–4, 84

Petitgen, Heinz-Otto

*Chaos and Fractals* 89

pendulum behaviour 6–7

phase space 44–5

Period Three Implies Chaos' (Li and Yorke) 78

period-doubling bifurcations

Baker transformation 99–100

convection 108–9

Feigenbaum diagram 76, 77, 80

Horseshoe transformation 99–101

iteration 79

population changes 74–6

reaction diffusion patterns 126

repeated 72

self-similarity 76–7, 77, 80, So, 81 81

periodic behaviour 59

Permian period, mass extinctions 155, 198

Peters projection 97

phaeomelanin, skin colours 127–8

phase space 42–9

Hamiltonian equation 42–9

landscape model 44, 48, 56

limit cycle 70–71, 80

pendulum behaviour 44–5

planetary orbits 45–6

system location 68–9

phase transition 118

complex system networks 164–5, 167–8

Pi

approximate value [39–40](#)

irrational number [65–8](#)

pink noise (one-over- $f$  noise) [144–7](#), [144](#), [146](#), [156](#)

planetary orbits [50](#)

changes [64](#)

convergence [41–2](#)

non-reversible trajectories [63–4](#), [66](#)

phase space [45–6](#)

Poincaré section [46–8](#)

*see also named planets*

planets

infrared emissions [207–8](#), [208](#), [233](#), [234](#)

occupations [231](#)

planetary system detection [202](#), [230–34](#)

terrestrial planet finder (TPF) [232](#)

*see also named planets*

plankton, iron deficiency [221](#)

pocket calculator, chaos model [58–9](#)

Poincaré, Henri [36–7](#), [42](#), [45–50](#)

Poincaré recurrence time [37](#)

Poincaré section [46–50](#); homoclinic tangle [47](#)

*Science et Méthode* 49–50

population changes, logistic equation 72–4

population studies, power law behaviour 147–8

power law behaviour 101–2

coastlines and borders 141, 142, 142

earthquakes 139, 140

economies 149–51

extinction patterns 156, 156–7, 193–7

global warming 145–6

long-grain rice model 160–63

music and spoken word 145, 146

planetary life detection 202

population studies 147–8

quasars 144–5, 145

traffic jams 148–9

weather systems 57, 145–7

white and pink noise 145

power series 40

Prigogine, Ilya 2, 30, 105–6

Brusselator 124–5

Order out of Chaos 2

Project Darwin (infrared space interferometer) [232](#)

proteins, DNA [169](#)

punctuated equilibrium [189](#)

connected networks [190](#)

Daisyworld [213](#)

sandpile model [190](#)

quasars, power law behaviour [144–5](#), [145](#)

random process chaos [86–7](#)

rational/irrational numbers [65–8](#)

Raup, David [156](#), [156](#)

Rayleigh, Lord [106](#)

Rayleigh number, convection [108](#)

Rayleigh-Bénard instability [106–9](#), [109](#)

reaction diffusion patterns [118–35](#)

angel fish [133–4](#)

Brusselator model [124–5](#)

butterfly wings [134](#)

BZ model [124–6](#)

Oregonator model [125](#)

period-doubling bifurcations [126](#)

reciprocal relations [105](#)

Red Queen effect

evolutionary biology [181–2](#)

fitness landscapes [185–6](#)

frog model [181–2](#)

relativity

general theory [112](#), [175](#)

special theory [20](#)

renormalization [73–4](#)

resonance, defined [61](#)

Richardson, Lewis Fry [51–4](#), [93–4](#)

*Weather Prediction by Numerical Process* [53](#)

Richter, Charles [138–9](#)

Richter scale [138–9](#)

Ruelle, David [82](#)

Rumford, Count (formerly Benjamin Thompson) [24](#)

San Andreas Fault *see* [earthquakes](#)

sandpile model [162](#)

complex system networks [164–6](#)

and long-grain rice model [160–63](#)

colour coding [161–3](#)

power law [158–62](#)

punctuated equilibrium [190](#)

self-organized criticality [159](#)

wooden cube model [159–60](#)

Saturn, orbit calculation [15–16](#)

Saupe, Dietmar

*Chaos and Fractals* [89](#)

scale invariance

earthquakes [142](#)

extinction patterns [193–4](#), [154](#)

frozen potatoes model [143](#)

lunar landscape [143–4](#)

scaling law, fractal dimension [95–6](#)

self-organized criticality

co-evolution [183](#)

earthquakes [159–63](#)

networks of interactions [195](#), [197–8](#)

open systems [158](#)

sandpile model [159–63](#), [162](#)

self-organizing systems

amino acids [234](#)

Earth's climate [217](#)

galaxies [228–30](#)

ocean desert [217–18](#)

self-regulating system, Earth [204–6](#), [222–3](#)

self-similarity

period-doubling [76–7](#), [77](#), [80](#), [81](#)

simplified models [80](#), [81](#)

Sepkoski, Jack [154–6](#), [155](#)

sex, biological evolution [187](#)

*Shackleton* voyage, DMS

measurements [215](#)

Sierpinski, Waclaw [87](#)

Sierpinski gasket [87–90](#), [88](#), [89](#)

SIM (space interferometry mission) [232](#), [233](#)

skin colours, [128–9](#)

*see also* [mammalian coat patterns](#)

Smale, Stephen [78](#), [99–100](#)

horseshoe attractor, Baker

transformation [99–100](#), [99](#)

Smith, Henry [84n](#)

Smolin, Lee [226](#)

Sneppen, Kim [190–91](#)



## Solar System

asteroid-belt formation [60–61](#)

planetary obliquity changes [65](#)

*see also named planets*

spectroscopy, Martian atmosphere [203](#)

stable ecological network [192–3](#)

## stars

bright [227–8](#)

GAIA project [232](#)

interstellar dust and new stars [226](#)

kinetic energy [113](#)

mass [227](#)

nuclear reactions [225](#), [225](#)

spiral galaxies [226–9](#)

supernova [227](#)

state cycles [170–72](#)

cells [172–3](#)

feedback system [172](#)

statistical mechanics, thermodynamics [22](#), [30–32](#)

Stengers, Isabelle [2](#), [30](#)

*Order out of Chaos* [2](#)

Stevin, Simon [9](#)

stock market fluctuations, power law behaviour [149–51](#)

Sun

Earth's temperature [206–9](#)

faint young Sun paradox [206–7](#)

solar energy [113–14](#)

'survival of the fittest', defined [176–7](#)

symmetry breaking [117](#)

Takens, Floris [82](#)

Tang, Chao [158](#)

terrestrial planet finder (TPF) [232](#)

thermodynamics [21–34](#)

energy dissipation [104–5](#)

entropy concept [25–8](#)

equilibrium [29–30,111–13](#)

first law [24](#)

Fourier's universal law [23](#)

fourth law [105](#)

open systems [104](#)

phase transition [118](#)

second law [24–8, 30](#)

contraventions [122](#)

statistical mechanics [22](#), [30–32](#)

unstable equilibrium [51](#)

*see also* [arrow of time](#)

Thompson, Benjamin (later Count Rumford) [24](#)

Thompson, D'Arcy [117](#)

*On Growth and Form* [117](#)

Thomson, William (later Lord Kelvin), Thomson effect [25](#)

three-body problem [14–17](#), [17](#), [50](#)

approximations [39–40](#)

restricted [45–6](#), [47–8](#), [50](#), [60](#)

*Through the Looking Glass* (Carroll) [181](#)

time

reversibility [69](#)

*see also* [arrow of time](#)

topology [45](#)

algebra [99](#)

period doubling [100](#)

phase space [45](#)

torus [80–82](#)

traffic jams, power law behaviour [148–9](#)

turbulence

chaos model [70–71](#)

periodic cycles [80–82](#)

Turing, Alan [114–21](#)

and morphogenesis [118–21,124, 126–8](#)

Turing machine [114–16](#)

universal constant, complex numbers [137](#)

Universe

Copernican model [6](#)

earth-centred model [6](#)

equilibrium state [111–13](#)

heat death [26, 27, 34](#)

van Helmont, Joannes (formerly Jan) [21](#)

*Ortus Medicinae* [21](#)

Van Valen, Leigh 18m, [187](#)

Venus

Infrared emissions [208](#)

obliquity changes [65](#)

orbital changes [64](#)

vertebrate limb, biochemical processes [134–5](#)

Viking project, Mars landing [203](#)

Volterra, Vito [122](#)

volume, power law [101–2](#)

von Koch, Helge [92](#)

Vostok, Ice Age temperature variation research [219–20](#), [220](#)

Wallace, Alfred Russel [175](#)

water, heavy, Ice Age temperatures [219](#)

Watson, Andrew [230](#)

weather forecasting [54–5](#)

- computing techniques [54–5](#)

- convection numerical model [56](#)

- initial conditions [57–8](#), [59](#)

- numerical [50–55](#)

- synoptic forecast [52n](#)

weather systems, power law behaviour [145–7](#)

white noise [145](#)

Wiesenfeld, Kurt [158](#)

William of Ockham [127](#)

Wisdom, Jack [61](#)

Yorke, James [78](#)

‘Period Three Implies Chaos’ [78](#)

Yorke, Robert [99](#)

Yucatan impact, K-T event [152–3](#)

zebra, coat patterns [130–31](#)

Zhabotinsky, Anatoly [123–4](#)

Zipf, George [147](#)

1. Quoted by, for example, Roger Lewin, in *Complexity*.

1. There is no evidence, by the way, that Galileo himself ever carried out such an experiment from the leaning tower in Pisa, a myth which seems to have originated in the colourful biography of Galileo written by Vincenzo Viviani, who was only seventeen when he became Galileo's hero-worshipping scribe after the great man went blind in old age.



2. A single copy of one of Archimedes' manuscripts, lost for centuries, was translated late in the twentieth century and revealed that the great Greek mathematician had developed the essence of this kind of integration more than two millennia ago. This was, of course, completely unknown to Newton and Leibniz, but puts their squabble over priority in a different perspective.

3. It's perhaps worth pointing out that only as 1609, when Johann kepler realized that something must be holding the planets in their orbits around the sun, he had called it the 'Holy Sprit Force' and nobody laughed.

4. The same is true if three moving balls collide with one another simultaneously, but that is even harder to arrange.

5. We use billion in the conventional scientific sense, meaning thousand million.

6. This gloomy prognosis has now been superseded. The fact that the Universe is expanding (discovered only at the end of the 1920s) alters the whole context of such forecasts, and the realization that gravity effectively has negative energy, dating from the 1940s, essentially rules out the kind of heat death imagined by the Victorians. These intriguing ideas are discussed in my book *In the Beginning* (Penguin, 1994) and by Paul Davies in *The Cosmic Blueprint* (Heinemann, 1987); more of this later.

7. *Entre le temps et l'éternité*, Fayard, Paris, 1988. Prigogine died in 2003, just as this book went to press.

8. He was so proud of this equation,  $S = k \log P$ , that it is carved on his gravestone.

9. There are a thousand cc in a litre, not a hundred!



10. Available in a reprint from University of California Press, Berkeley, 1964; originally published in German in two volumes, 1896 and 1898.

11. It was because of such criticism that Boltzmann later came up with the idea of our world as a local fluctuation in a dead Universe, in an attempt to salvage something from the wreckage.

12. For the cycle time to equal the age of the Universe, of course, the box would have to contain just 17 particles.

1. The term 'phase' is used for historical reasons and has no significance today except as a name.

2. For the mathematically inclined, the force is the rate of change of momentum, or the first derivative of momentum.

3. This approach was, in fact, used with some success late in the nineteenth century by the American astronomer George Hill to calculate the Moon's orbit.

4. It's actually even more complicated than this. Remember that phase space involves higher numbers of dimensions than we are used to. So a Poincaré section is not (generally) a simple two-dimensional surface, but a multi-dimensional surface. This allows an even more complicated tangle of infinite possibilities, called a homoclinic tangle. But for our purposes it's enough to know that the possibilities are indeed infinite.

5. An unstable equilibrium is like a pencil balanced on its point. Poincaré is saying that in such a situation we know the pencil will fall over, but we cannot predict which direction it will fall in.



6. He was also an uncle of the actor Ralph Richardson, though that has no relevance to our story.

7. Contrast this approach with the traditional 'synoptic' forecast, which is based on studying weather patterns on a chart and using a basic understanding of the physics involved plus a gut feeling that develops with experience to guess how the highs and lows will move over the next few hours.

8. They could each be working on the data for one grid point, perhaps.

9. See Edward Lorenz, *The Essence of Chaos*.

10. The paper is reproduced in his book *The Essence of Chaos*.

11. Quoted by Ian Stewart, in *Does God Play Dice?*

12. The contribution by Franco Vivaldi to *The New Scientist Guide to Chaos* (edited by Nina Hall) will give you some ideas.

13. For historical reasons, the object known as Pluto is also still usually called a planet, although in fact it is simply the largest member of a belt of icy, frozen objects beyond the orbit of Neptune, debris left over from the formation of the Solar System.



14. See John and Mary Gribbin, *Fire on Earth*.

15. The rational expression  $\frac{22}{7}$ , often used as an approximation to  $\pi$ , is just that—an *approximation*; it is not really  $\pi$ .

1. Strictly speaking, a limit cycle would be a closed loop in phase space, but for our purposes we can think of it as a single lobe of the Lorenz attractor.

2. Intriguingly, though, just after writing this section I was out shopping after a heavy shower of rain, and while waiting in a queue at the hardware store I noticed a drip from a leaky gutter outside the window, where I could clearly both see the sets of four falling drops and hear their impact on the pavement below making exactly this rat-a-tat-*tat* rhythm!

3. Feigenbaum's number is actually, like  $\pi$  and most other numbers, irrational. To a few more decimal places, it can be written 4.6692016090.

4. Just a year later, David Hilbert (1862–1943) made the same discovery.

5. See his contribution to *The New Scientist Guide to Chaos*, edited by Nina Hall.

6. Unknown to Cantor, this set had already been discovered by the clever but rather self-effacing Dublin-born mathematician Henry Smith (1826–83), in 1875. But with Smith dead and his discovery virtually unknown, it was Cantor's name that got attached to the set.



7. That kind of problem is largely overcome with digital transmission of data, but this was still in the days of analogue signals.

8. Quoted by James Gleick in *Chaos*.

9. There are actually slightly different ways to measure fractal dimension, which give slightly different 'answers', so you may see a slightly different value quoted somewhere else. But we don't want to go into these subtleties here, and will leave them for the mathematicians. What matters to us is that here is indeed at least one fairly obvious way to measure fractal dimensions.

1. Strictly speaking, as the value of a parameter called the Rayleigh Number increases. This number depends not only on the temperature gradient but also on the properties of the fluid.

2. See *In Search of the Big Bang*.

3. See George Gamow, *My World Line*.

4. See Paul Davies, *The Fifth Miracle*.

5. A topic covered in my book *Stardust*.



6. Energy also comes from within the Earth, chiefly as a result of the decay of radioactive elements in the Earth's core. This radioactive material was produced in previous generations of stars, and spread through space when those stars exploded, becoming part of the interstellar cloud from which the Solar System formed. So this energy source, too, ultimately owes its origin to gravity. Life forms that feed off this energy, which escapes through hot vents in the ocean floor, may do so entirely independently of the energy from sunlight, but they are as much a product of gravity as we are.

7. Technically, Turing proved that such a machine could solve any computable problem, which means one that can be specified using a suitable algorithm; he also proved that some problems are not computable, which is of profound interest to mathematicians and is related to the idea that some things are not algorithmically compressible.

8. *Philosophical Transactions of the Royal Society*, 1952, volume B237, p. 37; this is now regarded as one of the most influential papers in the whole field of theoretical biology.

9. Turing seems to have had an obsession with poison. His biographer Andrew Hodges describes how Turing went to see the movie *Snow White and the Seven Dwarfs* in Cambridge in 1938, and was very taken ‘with the scene where the Wicked Witch dangled an apple on a string into a boiling brew of poison, muttering: “Dip the apple in the brew. Let the Sleeping Death seep through.” ’ Apparently, Turing was fond of chanting the couplet ‘over and over again’ long before he suited the action to the rhyme.

10. Such oscillating systems are known today as 'chemical clocks' because of the regularity of their rhythms; but this regularity is only relative, and they are not accurate enough to use as real clocks.

11. Acting on the advice of an outside expert, a 'referee' used, as with most learned journals, to vet the suitability of papers for publication.

12. An English translation of the original rejected paper was eventually published in 1985, in *Oscillations and Travelling Waves in Chemical Systems*, edited by R. J. Field and M. Burger, Wiley, New York.

13. This was found in 1983 by a team at the University of Texas, Austin, which included Harry Swinney, who was later one of the first people to make the ‘Turing spot’ pattern.



14. Volume 258, number 3, p. 80.

15. Named after the English logician and philosopher William of Ockham (c.1285-1349), who said 'entities ought not to be multiplied except of necessity'.

16. My favourite examples of evolution at work today are the way 'superbugs' have evolved resistance to antibiotics, and insect pests have evolved resistance to insecticides.

17. Important experiments along these lines, using chicks, were carried out by Lewis Wolpert and his colleagues in London.

1. As is a cog in a clockwork chronometer, which would serve as an equally valid example in the following argument.

2. The name 'Richter scale' is actually applied today to modern versions of the scale, which differ slightly from Richter's original, but the differences are not important for the discussion here.

3. 'Logarithm' is another term that frightens people who have an aversion to mathematics (usually caused by bad teaching). It is simply a way of expressing numbers in terms of powers of ten, or powers of some other chosen number. For example, 100 is  $10^2$  so the logarithm of 100 is 2. That is absolutely all there is to it, and all you need to remember is that logarithms are ways of expressing numbers as powers.

4. Plotting the logarithm, instead of the simple number, up the side of the graph, just means that instead of equal steps up the axis being labelled 1, 2,3 and so on, they are labelled 1,10, 100 and so on.



5. See J. Feder, *Fractals*

6. And probably longer, but quasars were only discovered in the early 1960s and photographic records of objects now known to be quasars only go back to the 1880s.

7. K. Nagel and M. Paczuski, 'Emergent Traffic Jams', *Physical Review*, 1995, E51, p. 2909.

8. Although, of course, some jams are caused by crashes, just as some earthquakes are caused by underground nuclear tests, and some global warming is caused by human activities.

9. But if you want to know more about this aspect of the application of power laws, we recommend Mitchell Waldrop's book *Complexity*, or Mark Buchanan's *Ubiquity*.

10. Not all of them; birds are thought to be descended from one or more dinosaur lines.

11. *Nature*, 1990, vol. 343, p. 251.

12. If you want more detail, see our book *Fire on Earth*.



13. We don't need to worry too much about the subtleties of biological nomenclature, but, for the record, a species is the lowest level of the classification scheme; a species is a member of a genus, and a genus is a member of a family.

14. Bak (1947–2002) died just as this book was being completed.

15. Actually, this is almost a tautology. In this context, 'large' really means 'being made up of many parts', rather than having to do with the literal physical size of a system.

16. One at a time, because we are interested in systems that are in a kind of steady state, like the living human body, only being disturbed slightly as they feed off a steady flow of energy. Things are different if you dump a whole bucketful of sand on the pile, just as what happens to life on Earth would be different if, instead of a steady flow of energy from the Sun, all of the Sun's energy were released in one vast explosion.

17. He suggests Lego, but then he was a Dane. Wooden cubes work better.

18. For this toy model, it doesn't matter if the numbers the computer generates are not truly random.

19. It's also intriguing, to an astronomer, that this network pattern resembles the network pattern of galaxies scattered across the Universe; but I have no evidence that this is a significant similarity.

20. Modern investigations of the Universe at large have revealed that the kind of clouds in space from which planetary systems like our own form contain a great variety of molecules, up to and including amino acids, that would have been brought down to the surface of the Earth by comets when the Solar System was young, seeding our planet, if not with life itself then certainly with the raw materials of life. More of this later.



21. See Stuart Kauffnan, *At Home in the Universe*.

22. This is a major reason why genetic engineering will never be as simple as some accounts would have you believe. Even if it is possible to modify a single gene to, say, enhance the capability of the lungs (which is itself pretty unlikely), by changing that one gene we would also be changing, in as yet unpredictable ways, the behaviour of all the genes in the same network.

23. Named after George Boole (1815–64), the English mathematician who developed the idea. Boolean algebra is important in many programming languages.

24. There is one other parameter, to do with the choice of Boolean rules, that affects these systems, but this is a kind of fine-tuning and does not affect the fundamental point we are making.

1. See, for example, *Evolution* by Carl Zimmer, or *The Beak of the Finch* by Jonathan Weiner.

2. See Carl Zimmer, *Evolution*, p. xii.

3. The same argument applies to plants, of course. We just describe it in terms of animal life since that emphasizes that it applies to us as well.

4. We are confident that everyone who reads this is at least aware of the idea that the physical characteristics of a body (its phenotype) are determined by the set of genes within its cells (the genotype), and that changes in the phenotype occur because of changes in the genotype, which can happen, for example, as a result of copying errors when DNA is replicated, and are sometimes known as mutations. If you want to remind yourself how all this works, almost any book by Richard Dawkins will help, or you might try *In Search of the Double Helix* by John Gribbin.



5. 'Better' means 'leave more descendants'; this is the only criterion of evolutionary success.

6. The Red Queen terminology was introduced by evolutionary biologist Leigh Van Valen of the University of Chicago in the early 1970s.

7. Actually, this appealing image is upside-down compared with Sewall Wright's landscapes, where *peaks* are the places where you expect to find species, not dips; but you get the picture.

8. Quoted in Mitchell Waldrop, *Complexity*.

9. See *The Mating Game* by John Gribbin and Jeremy Cherfas (Penguin, 2001).

10. On the appropriate timescale, you can also regard the events associated with the drought of 1977 on the Galápagos as an example of punctuated equilibrium.

11. Our mathematical friends insist that we point out that this is not a ‘critical state’ according to their strict definition of the term; but it does correspond closely to the kind of critical states we have been discussing, and the subtlety does not affect the point we are making.

1. Lovelock's delightful autobiography, *Homage to Gaia*, describes his background with sometimes painful but always complete honesty.



2. Although Lovelock became an independent researcher, working from his home in the English countryside, in the 1970s he also became an honorary Visiting Professor at the University of Reading. This unpaid post provided him with what he called ‘a respectable cover’, which helped to get papers published in journals whose editors were suspicious of missives sent from private addresses; it also gave him contact with other researchers and students who he could bounce his ideas off.

3. See Lovelock's *Homage to Gaia*.

4. The full story is told in my book *Stardust*.

5. Our emphasis.

6. Details of the astronomical theory of Ice Ages can be found in our book *Ice Age* (Penguin, 2001).

7. Although this is a pretty obvious idea, it was first proposed in print by me, in *Nature* in 1988 (vol. 331, p. 570). I have so few bright ideas that I hope I will be forgiven for the immodesty involved in pointing this out!

8. The details of that compelling evidence can be found in *The Birth of Time* by John Gribbin (Phoenix, 1999).

9. Apart from the overwhelming amount of hydrogen in our Solar System, for every 100 atoms of oxygen there are 57 atoms of carbon and 13 atoms of nitrogen, with silicon, the next most common element, trailing in with half the abundance of nitrogen. But everything except hydrogen and helium together makes up only 0.9 per cent of the mass of the Solar System.



10. See *Nature*(2002), vol. 416, pp. 401 and 403.

11. Roughly speaking, the distance to the Andromeda galaxy is about 20 times its diameter. If the nearest star to the Sun were 20 solar diameters away, it would be orbiting the Sun at a distance of just 30 million kilometres, well within the orbit of Mercury!

12. At least, in a simple spiral galaxy with two arms. Some have more complicated structure, but discussion of these subtleties is beyond the scope of the present book.