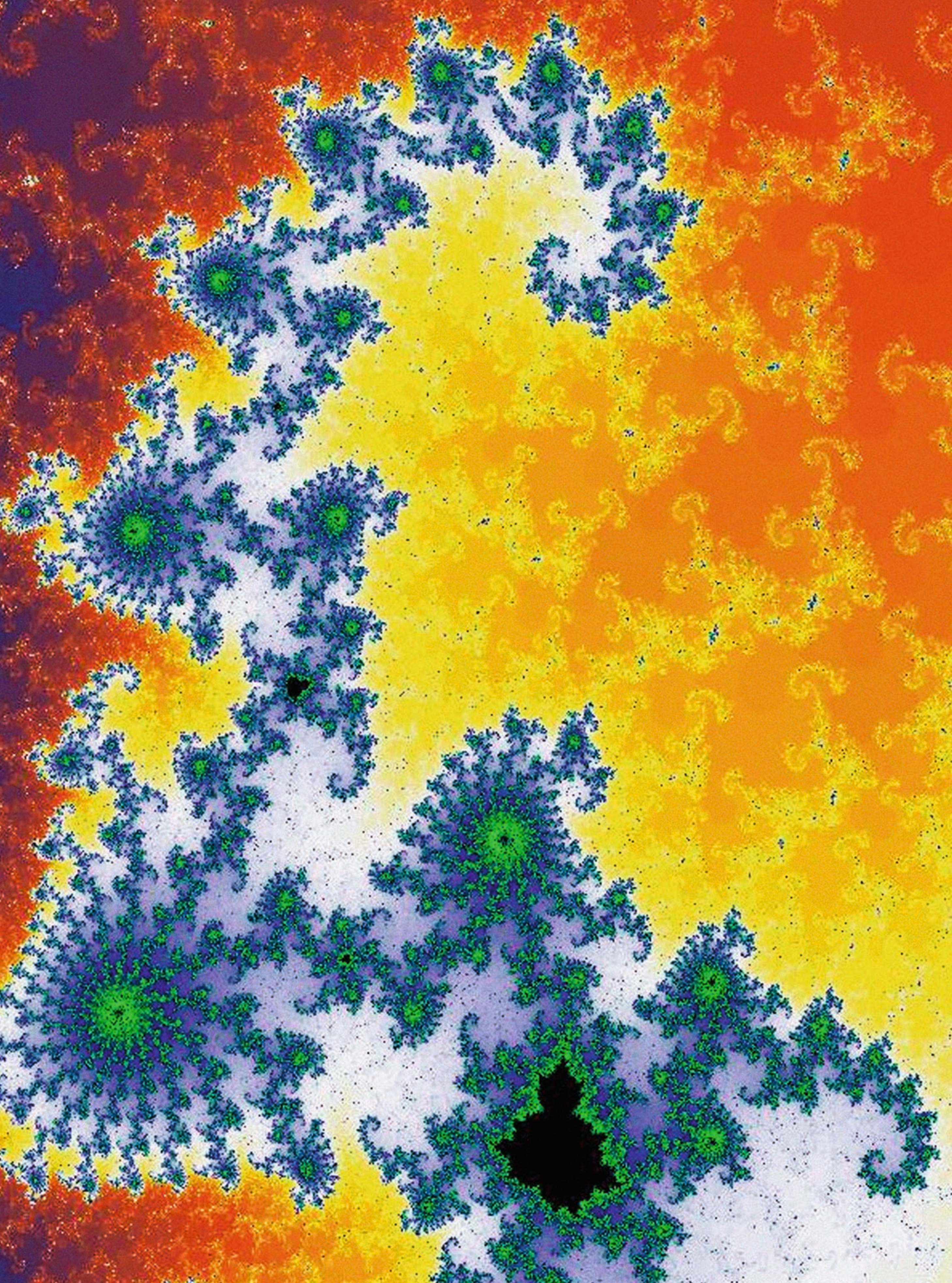
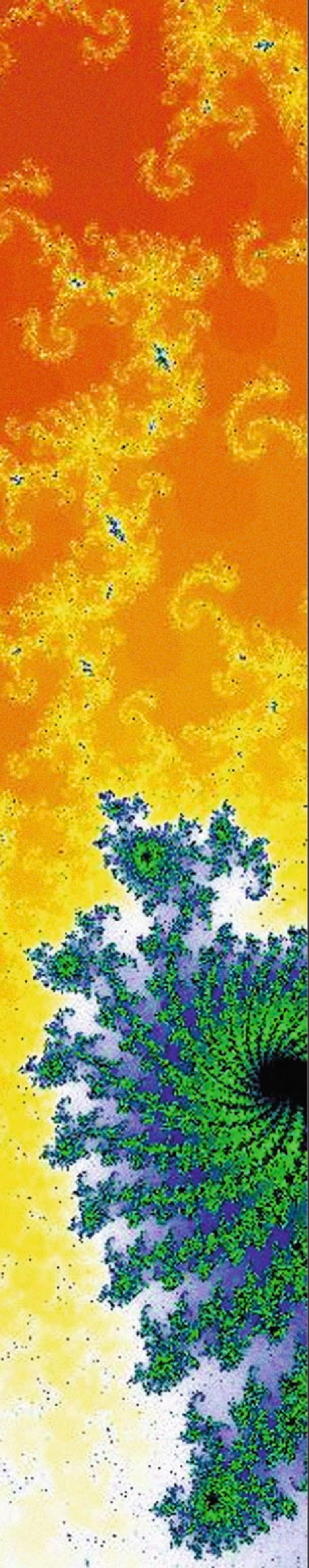
# The Colours of Infinity

## The Beauty and Power of Fractals





# The Colours of Infinity

## The Beauty and Power of Fractals

With contributions by

**Ian Stewart, Sir Arthur C. Clarke, Benoît Mandelbrot, Michael and Louisa Barnsley, Will Rood, Gary Flake and David Pennock and Nigel Lesmoir-Gordon**

Springer

Editor
Nigel Lesmoir-Gordon
Gordon Films UK
MK45 4AD Bedfordshire
United Kingdom
nigel@gordonfilms.tv

Additional material to this book can be downloaded from http://extras.springer.com
Password: [978-1-84996-485-2]
ISBN 978-1-84996-485-2 e-ISBN 978-1-84996-486-9
DOI 10.1007/978-1-84996-486-9
Springer London Dordrecht Heidelberg New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2010937321
Mathematics Classification Number (2010) 28A80



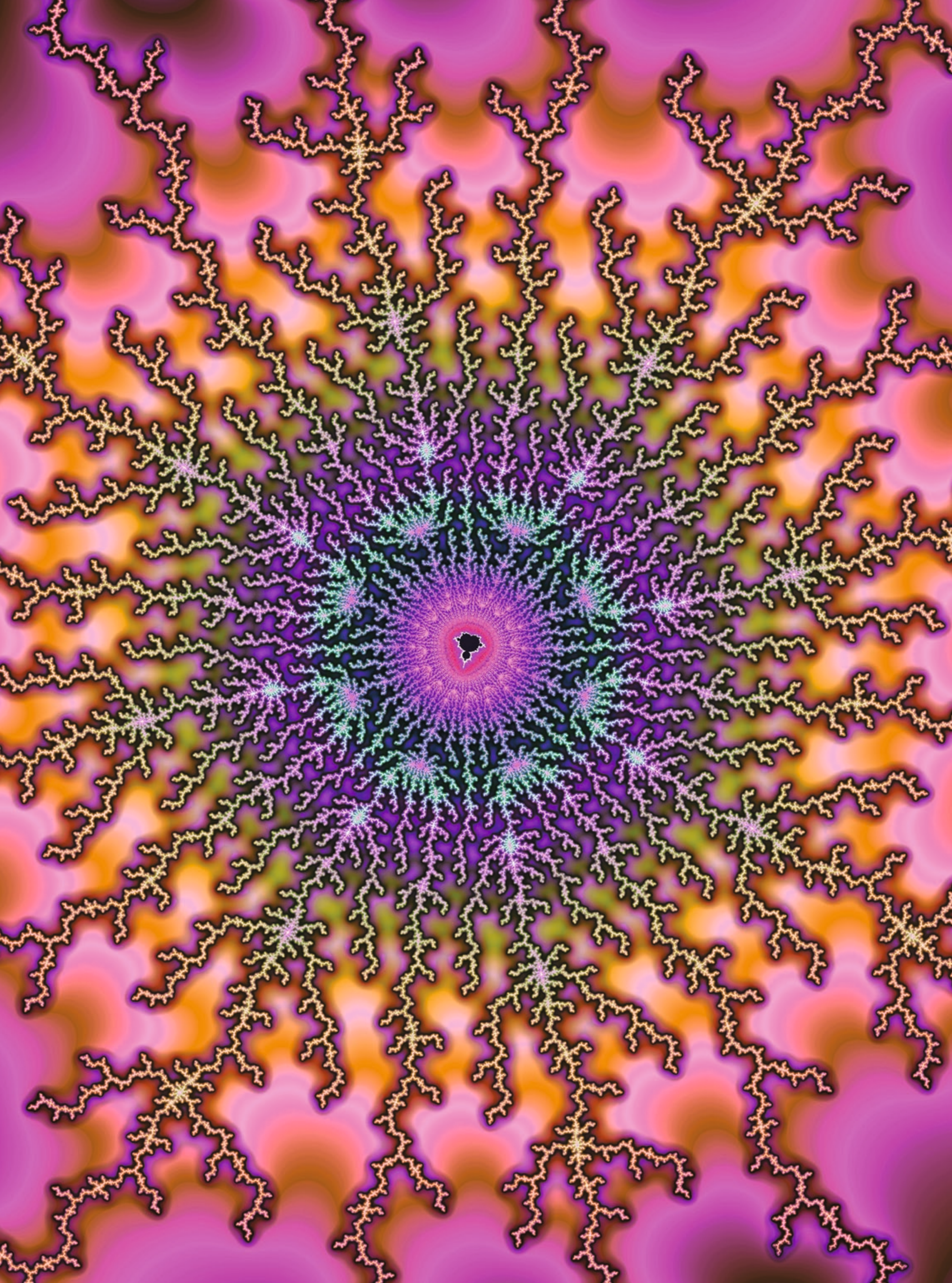
Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

This book is dedicated
to Benoît Mandelbrot

'The most beautiful thing
we can experience is the mysterious.
It is the source of all true
art and science.'

**Albert Einstein**



# Contents

# Introduction

A geometry able to include mountains and clouds now exists. I put it together in 1975, but of course it incorporates numerous pieces that have been around for a very long time. Like everything in science, this new geometry has very, very deep and long roots.

**Benoît B. Mandelbrot**

## Introduction

This enhanced and expanded edition of THE COLOURS OF INFINITY features an additional chapter on the money markets by the fractal master himself, Professor Benoît Mandelbrot. The DVD of the film associated with this book has been re-mastered especially for this edition with exquisite new fractal animations, which will take your breath away!

Driven by the curious enthusiasm that engulfs many fractalistas, in 1994, Nigel Lesmoir-Gordon overcame enormous obstacles to raise the finance for, then shoot and edit the groundbreaking TV documentary from which this book takes its name. The film has been transmitted on TV channels in over fifty countries around the world. This book is not just a celebration of the discovery of the Mandelbrot set, it also brings fractal geometry up to date with a gathering of the thoughts and enthusiasms of the foremost writers and researchers in the field.

As Ian Stewart makes clear in the opening chapter, there were antecedents for fractal geometry before 1975 when Mandelbrot gave the subject its name and began to develop the underlying theory. It took the genius of Mandelbrot, allied with the computer power available to him at IBM, to realize the practicality, beauty and fascination in the subject, and to act as its propagator through a long and influential career.

The first chapter by Benoît Mandelbrot in this book is based on a paper delivered before a Nobel Conference in Stockholm called A Geometry Able to Include Mountains and Clouds. The breadth of his vision, extending from mathematics to economics, from art to language, is extraordinary. As several of the contributors note, once you take a fractal view of the universe, you see the evidence everywhere – in water, in clouds, in trees, in art (see Rood's chapter), in the human body and in the workings of the World Wide Web (Flake and Pennock). Mandelbrot's second chapter, Fractal Financial Fluctuations looks deeply into the fractal nature of the growth and collapse of financial prices. His radically new fractal modelling techniques cast a whole new light of order into the seemingly impenetrable thicket of the financial markets.

The article by Arthur C. Clarke is a special case. Its 4,000 or so words are a lucid miniature of scientific popularization, reflecting the excitement fractal geometry induces in so many of its converts. It also, as Nigel

Lesmoir-Gordon explains in his account of how the film came to be made, offered a link between himself and Clarke, the film's anchor, and lent its name to the film project itself.

Four of the film's contributors (Stewart, Clarke, Mandelbrot and Barnsley) have chapters in the book. Rood, Flake and Pennock, as well as Nigel Lesmoir-Gordon, the film's begetter, contribute original chapters specifically for this volume.

Using a metaphor of a random soccer game, Michael Barnsley with his wife Louisa, the originators of fractal image compression technology, present the ideas of fractal transformation and colour stealing using random iteration for the first time.

Will Rood takes the animation of fractals into a new area by explaining how the M-set is coloured and then how the strange reptiles of Dutch conceptual artist M. C. Escher (1898–1972), the 'undisputed master of tessellated art', can be mapped onto the exterior of quadratic fractals, allowing the creation of tessellation with fractal limits.

Gary Flake and David Pennock propose an 'optimistic and realistic' interpretation of the NFL ('no free lunch') theory as a key to understanding the current state of the World Wide Web and how it will evolve over time. Given its huge traffic and lack of central authority, the Web could have been infinitely complex, but it is in fact exceedingly regular; and this regularity can be exploited to make more effective algorithms for finding information on the Web.

*The Colours of Infinity* brings together all the leading names in the fractal geometry field. Between them the contributors have published at least 200 books under their own names, and in collaboration. You will feel in their articles an ease with communicating sometimes difficult abstract concepts and an urge to share the powerful meanings their insights into the world of fractals have for all of us. In terms of positive energy and commitment to the subject they are a persuasive community.

The last chapter of this collection is unusual in that it sets out the full shooting script of the film, with audio and spoken word alongside. This may well prove invaluable source material in, for example, the educational use of the film, which has gradually increased over the decade or so since the film's release.

*The Colours of Infinity*, the movie, made with so much evident pleasure, is approaching cult status and now gains a new lease of life by being coupled with this stimulating collection, expanding the film's concerns still further.

The soundtrack of the DVD, with Pink Floyd's David Gilmour's soaring guitar almost an aural fractal in its own right, is totally accessible, as are Will Rood's beautifully coloured animations of the fractals. The music and the images together have become club and garage favourites, and it is easy to understand why. Is it too far fetched to see in this harmonious matching of sound and image a tribute to the way Stanley Kubrick handled them in the Stargate sequence of his science fiction masterpiece *2001: A Space Odyssey*? – a powerful link back to Arthur C. Clarke.

> *One of the many strange thoughts that the M-set generates is this. In principle, it could have been discovered as soon as the human race learned to count. In practice, since even a low magnification image may involve billions of calculations, there was no way in which it could even be glimpsed before computers were invented.*

**Sir Arthur C. Clarke**

# Contributors

**Arthur C. Clarke**
Professor Physical Research Laboratory, Ahmedabad, Gujarat India

**Benoît Mandelbrot (Emeritus)**
Stirling Professor Department of Mathematics, Yale University, New Haven, CT, USA

**Dr. David Pennock**
Research Scientist at Yahoo!, Pasadena, CA, USA

**Dr. Gary Flake**
President Yahoo! Research Labs, Sunnyvale, CA, USA

**Ian Stewart (Emeritus)**
Professor Department of Mathematics, Warwick University, Coventry, Warwickshire, UK

**Michael Barnsley**
Professor Mathematical Sciences Institute, Australian National University, Canberra, Australia

**Nigel Lesmoir-Gordon**
Director Gordon Films, Clophill, Bedford, UK

**Will Rood**
Freelance Software Designer and Filmmaker London, UK

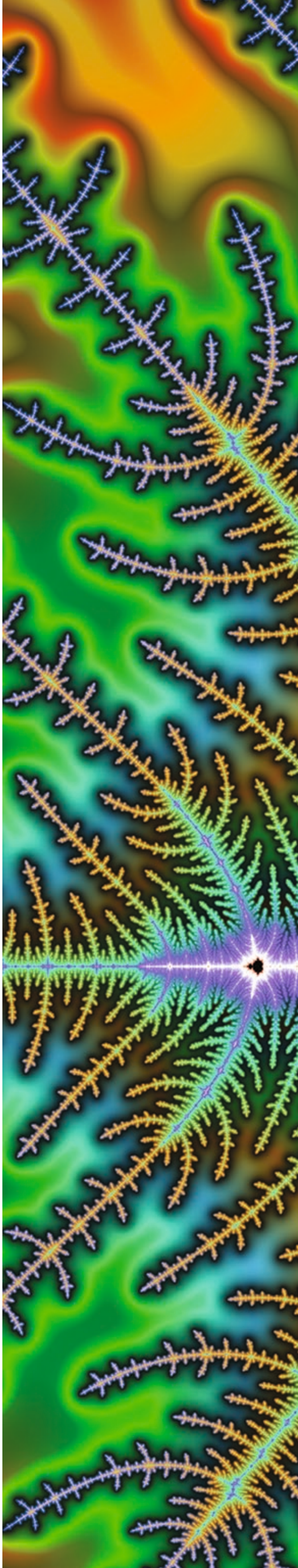
# 1 The Nature of Fractal Geometry

**Ian Stewart**

Fractals are more than just stunning visual effects – they open up new ways to model nature and allow us to quantify terms like ‘irregular’, ‘rough’ and ‘complicated’, writes mathematician Ian Stewart. His chapter does a service to the non-specialist reader in giving a largely non-technical introduction to fractal geometry in the context of mathematical traditions and its commercial applications. Stewart shows both how concepts like fractal dimension have a lengthy prehistory and also how Mandelbrot brought to the subject a systematic treatment, uniting theory and application. Mandelbrot’s most important contribution to fractal geometry, Stewart suggests, ‘was the realization that there was a subject’.

Ian Stewart is Emeritus Professor of Mathematics at Warwick University. In 1995 he was awarded the Royal Society’s Michael Faraday Medal for furthering the public understanding of science. He was elected a Fellow of the Royal Society in 2001, and won the Public Understanding of Science and Technology Award of the American Association for the Advancement of Science in 2002.

He is the author of over 60 books including *Nature’s Numbers*, *Does God Play Dice?*, *Figments of Reality*, *The Magical Maze*, *Life’s Other Secret*, *What Shape is a Snowflake?*, *Evolving the Alien*, the best-sellers *The Science of Discworld I and II* (with Terry Pratchett and Jack Cohen), and the US best-seller *Flatterland*. He has also written a science fiction novel, *Wheelers* (with Jack Cohen).



N. Lesmoir-Gordon (ed.), *The Colours of Infinity: The Beauty and Power of Fractals*,
DOI 10.1007/978-1-84996-486-9_1, © Springer-Verlag London Limited 2010

## *The universe is full of fractals. Indeed it may even be one.*

Thirty years ago, no one had heard of fractals. The concept existed, but the name was not coined until about 1975. Today, almost everyone has heard of fractals, and probably has a mug or a T-shirt or a poster somewhere around the house with one of the remarkable, intricate computer images that the word brings to mind. The importance of fractals, however, goes well beyond their visual attractiveness. What makes them so useful in today's scientific research is that they have opened up entirely new ways to model nature. They give scientists a powerful tool with which to understand processes and structures hitherto described merely as 'irregular', 'intermittent', 'rough', or 'complicated'.

What is a fractal? As a first, broad-brush description: it is a geometric form that possesses detailed structure on a wide range of scales. Think of the rocky slopes of a mountain, the proliferating fronds of a fern, and the fluffy outline of a cloud. These are physical objects: 'fractal' is a mathematical concept, and it relates to the real world in the same manner that 'sphere' relates to the shape of the Earth and 'spiral' relates to the shape of a snail shell. A mathematical fractal idealizes the intricacy of rocks and clouds: it has detailed structure on all scales. However much it is magnified, it does not 'flatten out' into a simple shape like a line or a plane.

Mathematical objects are idealized models of certain features of the real world; they are not real things, and they do not correspond exactly to real things. The Earth is not a perfect sphere; even allowing for its bulging equator, it is not a perfect ellipsoid either, even though many astronomy and earth science textbooks describe it that way. It has mountain ranges that give it a rough surface, unlike the infinite smoothness of the mathematical ideal. However, this type of inaccuracy does not stop scientists modelling the Earth as a sphere. In fact, the great advantage of a sphere as a model, for many purposes, is that it does not represent the intricacies of the real planet exactly. If it did, it would be no more use than a map of New York that is the same size as New York, with every traffic-light, doorstep, and cat rendered in perfect detail. A map must be simpler than the territory.

Models are tailored to suit particular objectives. If the objective is to understand mountain-building, then it is pointless to assume that the Earth is a smooth sphere. But if the objective is the long-term behaviour of the solar system, then a sphere is entirely acceptable, and a 'point mass' – even further from physical reality, since it assumes the Earth's diameter is zero – may well be better. In the same way, a mathematical fractal has detailed structure on scales so fine that they subdivide atoms – indeed, on scales finer than the Planck Length, at which level the universe becomes lumpy instead of smooth and 'distance' makes no sense. This discrepancy with the real world does not make fractals useless or irrelevant. As with the sphere and the map, what matters is the extent to which the model illuminates reality, not the extent to which it copies reality.

Fractals make it possible to quantify terms like 'irregular', 'intermittent', 'rough', and 'complicated'. How rough? 1.59 rough or 2.71 rough? Fractal geometry gives such statements a meaning, and makes it possible to test them in experiments. Mathematics provides a number, associated with each fractal, called its fractal dimension. The dimension reflects, among other things, the scaling properties of the fractal – how its structure changes when it is magnified. Unlike the traditional smooth curves and surfaces of much mathematical physics and applied mathematics, the dimension of a fractal need not be a whole number. It can, for example, be 1.59 or 2.71.

In fact, the difference between the fractal dimension of a geometric shape and its dimension in the usual 'topological' sense of mathematics provides a quantitative measure of just how rough the fractal is.

The notion of a fractal was brought to scientific prominence by Benoît Mandelbrot in 1975, and promoted in his book Fractals: Form, Chance, and Dimension of 1977. A revised edition appeared in 1982 under the title The Fractal Geometry of Nature. The term 'fractal' was introduced by Mandelbrot, but many of the subject's concepts – notably fractal dimension – have a lengthy prehistory. Mandelbrot's contributions to the subject have been many, but the most important was the realization that there was a subject. Mathematicians had studied spaces of non-integer dimension long before Mandelbrot; scientists had observed scaling laws and self-similarities in natural phenomena. But a systematic treatment, uniting theory and application, was lacking.

Now, some thirty years later, the theory that was stimulated by Mandelbrot's insight is thriving. A glance through the leading scientific journals, such as Nature and Science, will make it clear that fractals have become a standard technique of scientific modelling in a wide variety of areas. The mere existence of fractal structures immediately suggests a wide range of physical and mathematical questions, by directing our attention away from the classical obsession with smooth curves and surfaces. What happens to light waves passing through a medium whose refractive index is fractally distributed? Reflected in a fractal mirror? What sounds will a drum make if it has a fractal boundary? Traditional methods have little to say about such questions.

## The importance of fractals

Are they important? Undoubtedly. Turbulence in the atmosphere makes it difficult for Earth-based telescopes to produce accurate images of stars; a turbulent atmosphere is well modelled by a fractal distribution of the refractive index. Light bouncing off the ocean, with its myriad waves on many scales, closely resembles reflection from a fractal mirror. And the way trees absorb energy from the wind is closely related to the 'vibrational modes' of a fractal – and it is such modes that create the sound of a drum. The natural world provides an inexhaustible supply of important problems in fractal physics. Already, technological and commercial advances have stemmed from such questions – for example, a compact antenna for mobile phones, new ways to analyse the movements of the stock market, and efficient methods to compress the data in computer images, squeezing more pictures onto a CD.

Once our eyes have been opened to the fact that fractal objects possess a distinctive character and structure, and are not just irregular or random, it becomes obvious that the universe is full of fractals. Indeed, it may even be one. Fractals teach us not to confuse complexity with irregularity, and they open our eyes to new possibilities. Fractals represent an entire new regime of mathematical modelling, which science is just beginning to explore.

*The term 'fractal' was introduced by Mandelbrot, but many of the subject's oncepts – notably fractal dimension – have a lengthy prehistory. Mandelbrot's contributions to the subject have been many, but the most important was the realization that there was a subject.*

## Gallery of monsters

### The prehistory of fractals

The prehistory of fractals goes back over a hundred years, to when mathematicians began thinking about new kinds of curves and surfaces, totally different from the shapes typically studied in classical geometry. The classical shapes are lines and planes, cones and spheres, curves and surfaces – and, except for the occasional edge or corner, these curves and surfaces are smooth and very well behaved. Smoothness in effect implies that they have no interesting small-scale structure: when magnified sufficiently, they appear flat and featureless. This absence of structure on small scales is crucial to classical 'limiting' analysis – the time-honoured methods of the calculus, which go back to Isaac Newton and Gottfried Leibniz. The very methodology of the calculus, the central technique of physics for more than two centuries, is to approximate a curve by its tangent line, a surface by its tangent plane. This approach simply will not work on a highly irregular curve or surface.

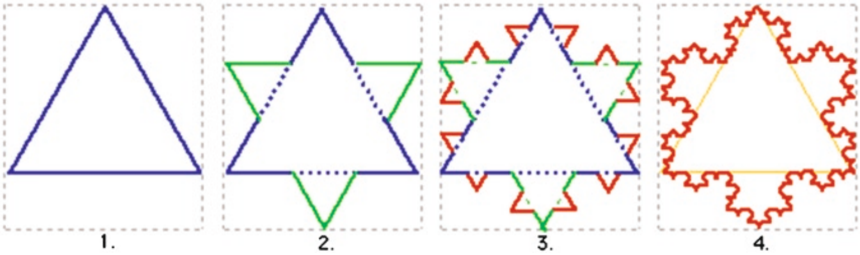
Nevertheless, we can imagine highly irregular curves. Originally these were seen as 'pathological' objects whose purpose was to exhibit the limitations of analysis. They were counter-examples, serving to remind us that the capacity of mathematics for nastiness is unbounded. The pure mathematician's motto is Murphy's Law: 'Anything that can go wrong, will go wrong.' And the wise mathematician or scientist always wants to know what can go wrong. Often this is a starting-point for finding new ways for things to go right.

For example, during the eighteenth and nineteenth centuries it was widely assumed that any continuous curve must have a well-defined tangent (that is, any continuously varying quantity must have a well-defined instantaneous rate of change) at 'almost' any point. The only exceptions were the corners, where the curve makes an abrupt change of direction. However, in a lecture to the Berlin Academy in 1872, Karl Weierstrass showed that this is not true. It is, in fact, about as false as it is possible to get. He described a class of curves that are continuous, but have no points where the tangent is well defined. The basic idea is to add together infinitely many increasingly tiny 'wiggles'. The resulting curve is continuous – it has no gaps – but it wiggles so rapidly that there is no sensible way to construct a tangent. Anywhere.

Again, in 1890 Giuseppe Peano constructed a curve that passes through every point of the interior of a unit square. This curve demonstrated the complete inadequacy of the common idea of 'dimension' as the number of (continuously varying) parameters needed to specify a point. Peano's curve takes a square, with its two dimensions and standard parametrization by two coordinates (north–south and east–west), and reparametrizes it by a single variable: how far you have to go along Peano's curve in order to hit a given point.

In 1906 Helge von Koch gave an example of a curve of infinite length that bounds a finite area: the snowflake. (Fig. 1.1) It is constructed by starting with an equilateral triangle, and erecting on each side a smaller triangle, one-third the size. This construction is repeated to infinity. Like Weierstrass's curve, the snowflake is continuous but has no tangent. A similar repetitive process occurs in the construction of one of the simplest and most fundamental pathological sets of all: the Cantor Set, named for Georg Cantor who used it in 1883 (although it was known to Henry Smith in 1875). It is constructed by repeated deletions of the middle third of an interval. (Fig. 1.2)

The mathematical community – even leading figures – found it hard to come to terms with these unsettling discoveries. Henri Poincaré dismissed



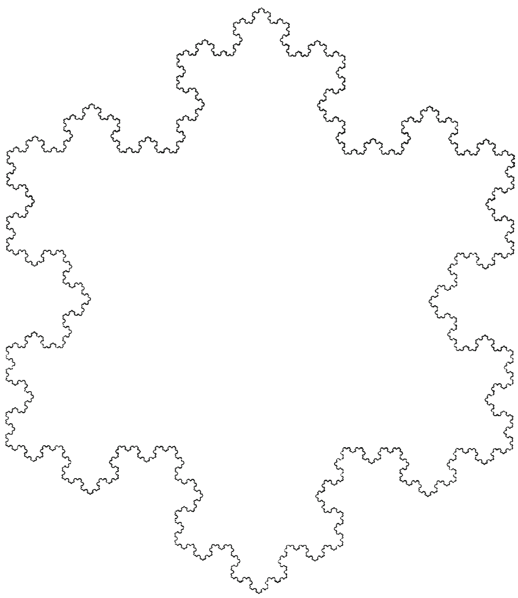



Fig. 1.1 To a casual observer this is a snowflake, but mathematically it is a classic fractal shape, constructed out of one equilateral triangle, with the middle third of each side removed and new equilateral triangles drawn out to the edge, their middle third removed, smaller triangles drawn out in turn, and so on.

Fig. 1.2 The Cantor Set: first developed in 1883, it is constructed by repeated deletions of the middle third of an interval.

them as ‘a gallery of monsters’, and Charles Hermite deplored what he called a ‘lamentable plague of functions with no derivatives’. More recently Jean Dieudonné wrote: ‘Some mathematical objects, like the Peano curve, are totally non-intuitive … extravagant.’ But Dieudonné was not suggesting they lacked interest, just that they were difficult to wrap your head round.

It is only fair to add that the undue proliferation of such sets, without any clear purpose in mind, can easily become an exercise in futility. So Poincaré and Hermite did have some basis for their complaints. But as time passed, most mathematicians came to accept that these sets play a legitimate, indeed crucial, role in mathematics: they demonstrate that there are limits to the applicability of classical analysis. In fact, this realization stimulated the development of new kinds of non-classical analysis, which turned out to be important in their own right. Indeed by 1900 the great German mathematician David Hilbert could refer to the whole area as a ‘paradise’ without causing ructions. Nonetheless, many mathematicians were perfectly prepared to operate within the classical limits. They saw the ‘pathologies’ as ‘artificial’ objects, unlikely to be of any importance in the study of Nature.

Nature, however, had other ideas.

## How long is the coast of Britain?

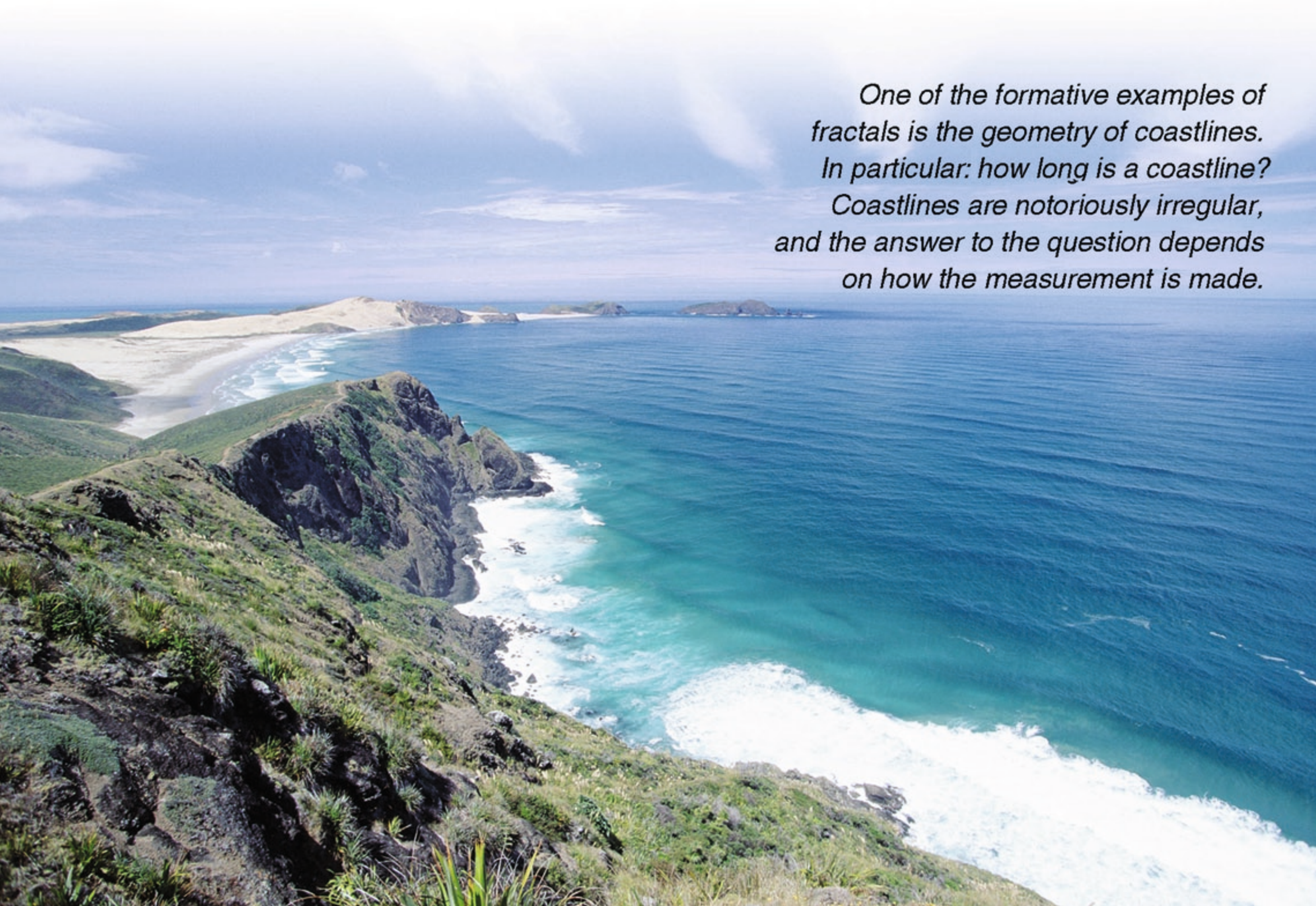
### The fractal geometry of coastlines

One of the formative examples of fractals is the geometry of coastlines. In particular: how long is a coastline? Coastlines are notoriously irregular, and the answer to the question depends on how the measurement is made. The simplest method is to take a fixed finite length x and move along the coastline in steps of length x. Adding these steps together gives a total length L(x). (Fig. 1.3)

Fig. 1.3 Mapping a coastline: the actual length depends on how many steps of length x one takes. If x = 1 km the length will be considerably less than if the length were 1 m; and this will be far less than steps of 1 cm; and so to infinity.

If the coastline is smooth, in the rigorous mathematical sense, then when x is small enough, the coastline is very close to a straight line. For a straight line, the value of L(x) tends to a definite limit L as x tends to zero, and that limit is the length of the straight line in the usual sense. It follows that if the coastline is a smooth curve, L(x) also tends to a definite limit L as x tends to zero, and that limit is the length of the curve in the usual sense. In other words, if x is small enough, L(x) is an approximation to the total length that is close enough on the scale of the model chosen.

What actually happens, with real coastlines, is quite different. Small bays of diameter smaller than x are missed by the stepping procedure. Although reducing the value of x must in some sense improve



the approximation, by 'noticing' ever smaller bays, there will still remain irregularities on some scale smaller than x, at least until we get down to molecular proportions where the whole exercise becomes meaningless. Because coastlines are fractal, the value of L(x) grows without limit, and the length is infinite.

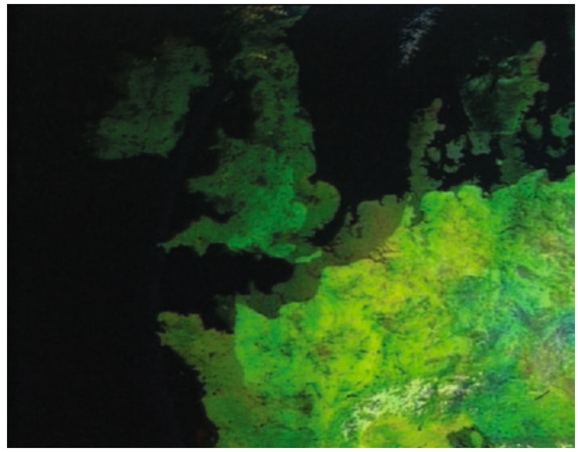
In the absence of a finite limiting value, it is often useful to study how a quantity tends to infinity. Is the growth rate fast and explosive, or slow and steady? In other words, what is the 'asymptotic' (when a curve tends towards but never reaches a straight line) behaviour? Lewis Fry Richardson once made an empirical study of the asymptotic problem, for real coastlines, and found an excellent empirical law: $L(x) \sim kx^{1-D}$ for certain constants k and D. The value of D is much the same for most coastlines on planet Earth, presumably for geological reasons, and in particular $D \sim 1.25$ for the coast of Britain.

To gain an intuitive feeling for what this result means, compare Britain to a snowflake curve. The construction of the snowflake is too regular to correspond to a real coastline, but as far as the main feature – structure on all scales – goes, it's not bad.

For simplicity, measure its length using values

$x = 1, {}^1/_3, {}^1/_9, {}^1/_{27}$, and so on.
Then $L(1) = 1$, $L({}^1/_3) = {}^4/_3$, $L({}^1/_9) = ({}^4/_3)^2$, $L({}^1/_{27}) = ({}^4/_3)^3$, and so on.
In general $L(({}^1/_3)^n) = ({}^4/_3)^n$. Let $x = ({}^1/_3)^n$, and note that ${}^4/_3 = ({}^1/_3)^{1-D}$ where $D = {}^{\log 4}/_{\log 3}$.
Then $L(x) = x^{1-D}$ and $D = 1.2618$.
This is very close to the empirical estimate $D = 1.25$ for the coastline of Britain.

I am not claiming that this implies that Britain is a snowflake. The snowflake curve's geometry is much too regular. Nevertheless, we may interpret the above calculation in the following terms. Suppose a real coastline has the same statistical distribution of bays and promontories, sub-bays and subpromontories, as does the snowflake curve. Then the value of L(x)



should follow the same asymptotic law as for the snowflake, and thus lead to the same D. If the statistical distribution is similar to that of the snowflake, but not quite the same, then the constant D should change slightly. So we conclude that the coastline of Britain has pretty much the same 'roughness' as the snowflake – but is maybe just a tad smoother.

The combinatorial regularity of the snowflake is essentially a scaling law. If a small section of the curve is suitably magnified, then it looks exactly like some larger section of the original. The constant D describes, in a quantitative manner, the precise scaling required. Here, if four copies of a segment of the curve are suitably assembled, the result has exactly the same shape as the segment, but is three times as large. The value $^{\log 4}/_{\log 3}$ of D is built from those two numbers. This property is called self-similarity. The same idea holds for coastlines, but now the scaling affects the statistics, not the curve as such. Instead of asking that a magnified version of a section of coastline should be exactly the same as the original, we ask that it should be a plausible picture of a coastline on the same scale as the original. Or, to put it another way: if you are presented with a map of a coastline, without any other markings and with no indication of the scale, then there will be no way to determine the scale just by studying the map.

Innumerable other natural phenomena exhibit structure on a wide range of scales, connected by suitable scaling laws. For instance, the bark of a tree, the ripples on the ocean, vortices in a turbulent fluid, landscapes, the inner surface of the lung, the holes in a sponge, the surface of a soap flake. Therefore we expect there to be some regime of mathematical modelling in which the 'pathological' curves and surfaces that were so despised by the classical mathematicians find natural application to the real world. Since scaling laws appear to be fundamental to the whole enterprise, the initial emphasis should be on understanding what they have to tell us. And the first thing they tell us extends the usual notion of 'dimension' in a radical way.

## Fractal dimension

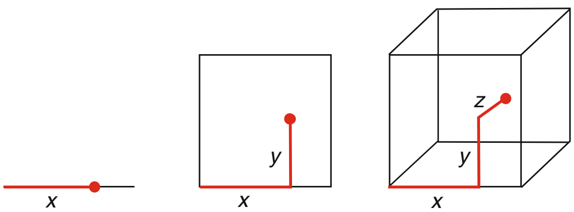
It turns out that the number D introduced above may be interpreted as a dimension. This may seem a rather curious idea, since the usual notion of dimension is always a whole number, but there are plenty of precedents in mathematics. The concept 'number', for example, originated in counting – one sheep, two sheep, three sheep. In this context, half a sheep makes no sense. But in the butcher's shop – or, less grimly, at the moneylender's, where a person might own a half share in a sheep – the extension of the number system to fractions is natural. Again, we are used to the idea that the nth power of a number is obtained by multiplying n copies of that number – so that the fifth power of 3, for example, is $3^5 = 3 \times 3 \times 3 \times 3 \times 3 = 243$. What, then, is the halfth power? What you get by multiplying half a copy of a number by itself? That makes little sense, but the halfth power makes excellent sense: it is the square root. Multiply the halfth power by itself, and you get back the first power – the original number. Twice a half is one – easy.

In fact, the generalization of dimension that occurs in fractal geometry is reasonable from several points of view. To see why, we begin by reviewing the usual concept of dimension. (Fig. 1.4)

(a) A line segment has dimension 1, by which we mean that any point in the segment can be specified using just one coordinate, one number. The point x lies x units to the right of the left-hand end of the segment.

(b) A square has dimension 2, by which we mean that any point in the square can be specified using just two coordinates (x, y). Here x is the distance from the left-hand edge and y is the distance from the bottom edge.

Fig. 1.4 The concept of dimension in geometry: (a) a line has 1 dimension and 1 coordinate; (b) a square has 2 dimensions and 2 coordinates; (c) a cube has 3 dimensions and 3 coordinates.




(c) A cube has dimension 3, by which we mean that any point in the cube can be specified using just three coordinates (x, y, z). Here x is the distance from the left-hand face, y is the distance from the bottom face, and z is the distance from the back face.

In these examples, the dimension of the object is the number of independent directions in space that it occupies. No directions are needed for a point, so it has dimension 0. A line lies along one direction, a square lies in two (a plane), whereas a cube requires three. Similar ideas apply to curved lines and surfaces. A curve has dimension 1. The surface of a smooth object, such as a sphere or torus, has dimension 2. A solid object, such as a solid sphere or a solid torus, has dimension 3. This concept of dimension is always a whole number. A point has dimension 0, a curve has dimension 1, a surface has dimension 2, a solid has dimension 3. With a suitable act of imagination, we can go into spaces of dimension 4, 5, 6, and so on – see Abbott (1884) and its modern sequel Stewart (2001). Engineers will recognize this concept as the number of 'degrees of freedom' of a system – the number of coordinates needed to determine its state – so that space-time, with 3 space coordinates and one time coordinate, is 4-dimensional.

The dimension of even a simple system can be surprisingly large. For example, describing the position and velocity of the Moon in space requires six numbers: three position coordinates, and three components of velocity relative to those coordinates. So the 3-body system composed of the Earth, Moon, and Sun, which is basic to astronomy, is an 18-dimensional system. Each body requires 3 coordinates of position in space and a further 3 of velocity.

A more extreme case is something we all carry around with us: the human body, with its innumerable flexible joints. Look at your hand. Each finger can be bent

Above: A solar eclipse

*At a conservative estimate, the 'configuration space' for the human body – the totality of possible shapes into which it can be bent – is at least 101-dimensional. Yes, we live in space of 3 dimensions, and a space-time of 4, but the complete range of possible configurations of the human body forms* through some angle, and those angles are pretty much independent of each other. So just to describe the state of your hand, you need a 5-dimensional space of possible configurations. In fact, fingers can bend sideways (a bit) too, so 10 dimensions is a more realistic number. Your two hands and two feet now require at least 40 dimensions to capture all possible combinations of positions, and then there are your wrists, elbows, shoulders, ankles, knees, thighs ... and your head, eyelids, and waist.

At a conservative estimate, the 'configuration space' for the human body – the totality of possible shapes into which it can be bent – is at least 101-dimensional. Yes, we live in space of 3 dimensions, and a space-time of 4, but the complete range of possible configurations of the human body forms a conceptual 'space' with 101 dimensions.

This notion is called topological dimension because shapes that can be continuously deformed into each other have the same dimension. Thus a wiggly curve has the same dimension, 1, as a straight line; a wobbly surface has the same dimension, 2, as a plane. And if a shape is magnified by some scale factor – say tripled in size – then its dimension remains unchanged.

Scaling laws are more sensitive: they involve not just shape, but size. Distances are important, scale matters. What count are not topological properties, but metric ones. This extra ingredient opens up the possibility of finding an extended notion of dimension which

(a) agrees with the usual definition for smooth curves and surfaces;
(b) applies to more general spaces, such as the snowflake or the Cantor Set; and
(c) reflects metric, not topological, properties, especially behaviour under scaling.

The price we pay for such an extension, however, is that the resulting concept of dimension is forced to take non-integer values. It turns out to be a price well worth paying – imaginative ideas that take us out of our comfortable world usually are.

The simplest such generalization (there are many) is the similarity dimension. This concept is based on scaling properties; it is a little too special to be entirely satisfactory, but when it does work it is very easy to understand.

Consider a unit square. If its sides are divided into n equal parts, then it can be cut into $N = n^2$ subsquares, each similar to the original. With a similar dissection of a cube, we find that $N = n^3$; with a 4-dimensional hypercube we get $N = n^4$. And with a homely line segment, $N = n^1$. (Fig. 1.5)

The pattern is obvious: if the dimension is d, then $N = n^d$. Taking logarithms and solving, we get $d = {}^{\log N}/_{\log n}$. All perfectly reasonable, and equivalent to standard geometrical properties of these simple shapes.

So let's try a shape that is not quite so simple: the archetypal 'pathological set', the Cantor Set. Remember: to form a Cantor Set, start with a line segment, remove its middle third to get two segments each one-third the size; then repeat indefinitely. What's left is the Cantor Set. It is clear that after the initial step, we construct two separate Cantor Sets, each one third the size of the whole; the Cantor Set itself is obtained by uniting these two subsets. In other words, the Cantor Set can be broken into two pieces (N = 2) each one third as big (n = 3). By formal analogy, the dimension of the Cantor Set 'ought' to be $d = {}^{\log 2}/_{\log 3} = 0.6309$, which is not a whole number. This may seem curious, but it makes a lot of sense because:

(a) it accurately reflects the scaling properties of the Cantor Set: two copies make a set just the same shape but three times as big; and
(b) the dimension is intermediate between 0, the dimension of a finite set of points, and 1, the dimension of a curve. This agrees with the intuitive idea that the Cantor Set is rather less than a curve, since it has gaps, but is more closely clustered than a finite set of points.