Complexity

Partial Course Notes

by

Rajesh R. Parwani

Abbreviated Fifth Edition: 2005/18

Preface

These are part of the unedited lecture notes for a Complexity course that was taught between 2001 and 2010 at NUS, Singapore.

These notes supplement material in the book 'Simplicity in Complexity' (SIC) described at www.simplicitysg.net/books/sic.

The SIC book covers the following topics: the scientific method, model building, dynamical systems, agent-based modelling, emergence, self-organisation, complex networks, measures of complexity, and applications in the social and natural sciences.

Rajesh Parwani
Singapore
enquiry@simplicitysg.net

Contents

Chapter 1

Introduction

1.1 Overview

Complexity refers to the study of **complex systems**, of which there is no uniformly accepted definition because, well, they are complex. Roughly speaking, one says that a system is complex if it consists of many interacting components (sub-units) and if it exhibits behaviour that is interesting but at the same time not an obvious consequence of the known interaction among the sub-units.

That sounds very vague, especially the use of words like "interesting" and "obvious", but it reflects an evolutionary perspective. For example, a hundred years ago one might have described the study of how a substance changes under heat (phase transitions) as a difficult and interesting problem that required one to deal with systems with a large number of interacting components (atoms). However by now very powerful tools, such as thermodynamics and statistical mechancis, have been developed to deal with such *equilibrium* systems leading to impressive quantitative agreement between theory and experiment. Though such systems are not commonly referred to as complex, they still provide valuable examples and concepts that have been used in complexity studies.

Current interest has shifted to **dynamical** systems that are (generally) **out-of-equilibrium** and thus highly **non-linear**. Such sytems actually form the bulk of natural phenomena but for which the theoretical tools are as yet poorly developed. **Some examples of such complex systems or phenomena are: The economy, the stock-market, the weather, ant colonies, earthquakes, traffic jams, living organisms, ecosystems, turbulence, epidemics, the immune system, river networks, land-**

slides, zebra stripes, sea-shell patterns, and heartbeats.

There is no single "Theory of Complexity", and it is unlikely that there will ever be one. Rather *one hopes that apparently different complex systems can be grouped according to some common features that they have, so that intuition and insight gained in studying one can be transferred to another.* Thus one of the main aims of complexity studies is to develop concepts, principles and tools that allow one to describe features common to varied complex systems. This leads to exciting **interdisciplinary** studies because it turns out that ideas developed to handle complex systems in the physical sciences have relevance also for systems in the biological and social sciences, and *vice versa*!

What are some of the characteristics of complex systems? One often quoted concept is that of **emergence**, which refers to the appearance of laws, patterns or order through the cooperative effects of the sub-units of a complex system. Thus the emergent phenomena or laws are not an intrinsic property of the sub-units but rather something that is a property of the system as a whole. Simple examples are those of "temperature" and the "gas laws": At the individual microscopic level, none of those make any sense, but they are features of a large system. More sophisticated examples are of "intelligence" and "conciousness" – where do they come from ?

Sometimes one sees the phrase "the whole is more than the sum of its parts", as a definition of emergence. This again reflects the non-linearity of the system, whereby the output is not proportional to the input, small changes can give rise to large effects, and the non-obvious results that can be produced in a large system.

It is important to realise that the universe consists of many hierarchial levels of complexity linked to each other. Each level has its own emergent patterns and laws: As one goes down from galaxies, solar-systems, planets, ecosystems, organisms, organs, cells, and atoms to quarks, different effective laws emerge. However these laws would not be useful if there was not some degree of **universality**, that is, one hopes that at each level of complexity the same laws apply to varied systems rather than each following its own tune. It is the apparent universality of the laws of physics, for example, that makes the world comprehensible and gives us faith in its ultimate simplicity. For example, at the atomic level weird quantum mechanics rules, but larger systems are well described according to Newtonian laws, while engineers often use empirical rules, and so do the social scientists.

It appears that nature has chosen to be economical (or is that an illusion on our part?), so that the branching of trees or the air-passages in our lungs, the shape of coastlines or clouds, the form of cauliflower or a mountain range, can be described by **fractal** geometry: Such shapes are self-similar over a wide range of scales, thus implying **scale-invariance**, whose hallmark is the appearance of ”power-laws”. In an equilibrium system scale-invariance naturally appears at the **critical point** of a **second-order phase transition**, such as that between the liquid and vapour phases of water. However natural systems are out-of-equilibrium and the common appearance of fractals and power-laws in such systems is not as well understaood. **Self-organised criticality** is the idea that many out-of-equilibrium systems naturally organise themselves, without external tuning or prodding, into a state which is at the threshold between complete disorder and complete order: That is, the system arranges itself into a critical state, and so displays scale-invariance and power-laws.

Living systems are the most complex examples one can think of and it is remarkable how such systems tend in their development towards greater order, organisation and complexity, in contrast to the **arrow of time** dictated by the **Second Law of Thermodynamics**. Of course there is no conflict as the increase in disorder and **entropy** required by the Second Law refers to closed equilibrium systems. Living systems are neither closed nor in equilibrium, but rather use an inflow of energy to drive processes that increase their order (thus decreasing their entropy), and dissipate heat and other waste products that lead to an overall increase in entropy of the universe. One can say that organisms are **dissipative structures**, and have a tendency towards **self-organisation** and **pattern formation**.

Ant-colonies are classic examples of self-organisation. Without a leader (the queen is actually an egg-laying machine) orchestrating everything, and without any of the ants having taken a course in engineering or social science, each ant seems to do its own thing, following a few simple rules that determines its interaction with its environment or its ant-mates. Yet, an incredibly complex and organised society emerges from such an interaction of the many ants. Ant-colonies display remarkable **adaptation** to changing circumstances, using both **feedback** mechanisms and **parallel** analysis of options. In recent years social and computer scientists have taken a keen interest in studying ant colony behaviour in order to help solve problems in their own fields.

Not all systems in nature appear organised or have some pattern to them. Indeed many seem disorderly or ruled by random events. However some of that randomness might only be on the surface. **Chaos** refers to the property of some non-linear dynamical systems whereby they become extremely sensitive to initial conditions and display long-term aperiodic behaviour that seems unpredictable. Though chaotic behaviour might appear essentially random, there is actually hidden order, apparent only in "phase-space" rather in ordinary space. Furthermore, many chaotic systems show universality in their approach to chaos, giving one some predictive power. Thus discovering that some random-like events are actually chaotic means one has uncovered a simple determinstic basis for the system and so enabled its understandability.

Often one encounters debates between **reductionism** and **holism**. Reductionists like to get right down to the bottom, meaning they are interested in the basic sub-units that make up the whole and believe that that is where all that is of interest lies, the whole itself being just a complicated and uninteresting consequence of the fundamental laws applied to a large system. In short, knowing the microscopic explains all to the extreme reductionist. Particle physicist are such, and in more recent times some molecular biologists involved with genomics are another example. While it is undoubtedly true that knowledge of the microcomponents of a system and the basic interactions among those is essential for us to progress, it is also a fact that such knowledge by itself is insufficient to predict all the diversity and novelty that can arise in a large system. (Take for example the task of predicting superconductivity from Schrodinger's equation – it is a problem that required much effort *after* the fact–one knew what to look for. Similarly knowing the whole genome code is not going to predict for us every feature of an organism or a society).

The problem of precisely deducing the whole (of a large system) from its parts is at least two-fold. Firstly it is a computational problem. Problems with a large number of degrees of freedom are too complicated for exact solutions, and for systems far from equilibrium, as complex systems are, they are also not solvable by the probabilistic averaging methods used for equilibrium systems. In recent years the growth of computer power at low cost has produced the first tool that allows large systems to be simulated or solved numerically. However this brings the second problem: Often one does not have full knowledge of the fundamental dynamics, or the initial conditions, or the problem is still too complicated to be handled directly even by computers.

Often what is required is some guesswork or intuition to reduce the actual problem to a simpler **model** which can then be tested on a computer. **Computer simulations** of simplified models let one test assumptions quickly, and when the results appear similar to the real world one can take it as *plausible* validity of the model and the assumptions. Qualitative similarities of course do not constitute a proof, because other models with different assumptions might give similar results, but at least the insight gained helps one to make further guesses and tests in a particular direction rather than being lost in a mess of detail. In fact one of the most important lessons computer simulations have taught us is that a large system with very simple local rules can give rise to collective behaviour of great complexity and variety, showing on the one hand that complex phenomena need not require complicated rules, but at the same time reminding us how difficult it is (without computers) to deduce the emergent behaviour from the sub-units and their interactions.

Thus studying the whole is as interesting as studying its parts, as novel structures and emergent laws arise at each level of complexity. The condensed matter physicist studying superconductivity is not going to be replaced by the string theorist, and neither is the ecologist going to be become obsolete because of the molecular geneticist. Explaining dynamic patterns, order and emergent laws of a complex system by understanding the organising principles among the sub-units is what might be called holism, the counterpoint to reductionism.

1.2 Examples

Let us briefly look at examples to illustrate some of the points above. These are just apetisers, we will leave details for the main course.

1.2.1 Schooling of Fish

Try out the applet at Ref.[16]. Does it not look like a very realistic simulation of fish swimming? The motion of each individual fish is not scripted right from the beginning but rather each individual follows just three simple local rules: cohesion, alignment and separation. Each of the rules is sensible from the biological perspective and so the model is plausible. What is remarkable of course is how the realistic and complicated *collective* behaviour emerges from the few simple local rules. There is no leader and none of the individuals has a global plan or perspective, (the motion is not orchestrated from the beginning).

This example is an example of *self-organisation.* Other examples are the herding behaviour of humans, say for example in the stock-market, and the alignment of magnetic spins to form a ferromagnet. Many such examples will be studied later.

1.2.2 Bacterial Colonies

Look at the picture of a bacterial colony. It shows a branching structure, which has the property that if one zooms into any region, that part looks similar to the whole. The bacterial colony is an example of a random *fractal.* Exact fractals appear the same at different magnification scales while random fractals appear only statistically similar at different magnification scales.

Fractals are ubiquitous in Nature. Another example is the branching network of air-passages in the human lung. The advantages of such a structure are an increase in surface to volume ratio which maximises functional efficiency while minimising material and space costs.

The word "fractal" itself means more than just self-similarity at different scales. It also implies a *fractal dimension* which we will discuss later.

1.2.3 Forest Fires

Look at the figures which show the number of fires as a function of the area burnt, in different regions of the United States of America and Australia. On a log-log plot one sees that the data is well approximated by straight lines, meaning that the number of fires as a function of area is a power law:

$$N \propto A^{-\alpha} \tag{1.1}$$

with $\alpha \sim 1.3 - 1.5$. It is important to note that the straight line fits are for a wide range of the data (one can always fit a straight line to a small range), and for different geographical regions. This suggests a *universality* in the phenomena that requires a explanation.

Power laws are observed in many other natural phenomena such as earthquakes and solar flare activity. It has been suggested that these phenomena are examples of *self-organised criticality*, that is, the systems are attracted to a state which is between that of total order and total disorder. The word "critical" is borrowed from well-studied thermal equilibrium systems that undergo second-order phase transitions at critical points and display power laws. However in the case of forest fires (or earthquakes etc.), the systems are far from equilibrium and the power-law behaviour, that is criticality, does not require fine-tuning — it is self-organised.

Given the complexity fo the actual systems, it is impossible at present to study those systems from first principles. Rather one studies simplified models to check whether power-laws emerge naturally.

Power laws imply a self-similarity at different scales, so it is natural to suppose that self-organised criticality might be the common dynamical mechanism behind the wide occurence of fractal structures in nature.

Power laws also appear in social contexts, for example the frequency of occurence of words in a literary text, and the magnitude of wars. We will come back to these examples later in the course.

1.2.4 The Double Pendulum

Most of you are familiar with the simple pendulum: A small heavy object suspended at the end of a thin light string and set into oscillation. For small oscillations (and in the absence of air friction) the motion is that of a "harmonic oscillator": That is, periodic motion with a period proportional to the sqaure-root of the length of the pendulum.

For larger oscillations, the motion of the simple pendulum is still periodic but no longer given by a simple formula. Indeed, for large oscillations the equations governing the motion fo the pendulum are non-linear in contrast to the linear equations for small oscillations. However while the equations are nonlinear, the motion is still regular and predictable.

A double pendulum consists of two simple pendulums in tandem: One attaches a single pendulum to the end of another! The equations of otion are again non-linear for large oscillations but now the motion becomes quite irregular and very sensitive to the initial conditions. This kind of behaviour is the hallmark of *chaos*. See the simulation of the double pendulum in the references.

Chaos occurs in many nonlinear systems and it implies that even systems with a few degrees of freedom, and hence naively simple, can show complicated behaviour which is essentially unpredictable on long time scales. However chaos is very different from randomnes: The former arises in perfectly deterministic systems while the later is intrinsically nondeterministic, and the distinction between the two at the practical level can be seen by looking at the "phase space" of the system, as we shall see later.

1.2.5 The Leopard Spots

How did the spots on a leopard, the stripes of the zebra or tiger, or the patterns on sea-shells, form? Is there some simpel general framework which can explain these beautiful patterns? Yes, these patterns are called Turing structures, named after the Bristish mathematician who came up with a model to explain such structures.

Some chemical systems, when maintained far from equilibrium display oscillatory behaviour. When the chemical waves of such a reaction are allowed to diffuse through a medium at different rates andthe resulting pattern stabilised, one obtains Turing structures.

Later in this course we will look at many other interesting far from equilibrium systems, some of them showing cyclical behaviour.

1.3 Summary

The aim of the course **USSC3001: Complexity** is two fold:

(1) To provide you, the student, with a relatively gentle introduction to the concepts mentioned above so that you can continue on your own in greater depth if you so desire.

(2) To broaden your horizons by introducing you to the interdisciplinary nature of complex systems studies, which exemplifies in a concrete sense the often quoted "unity of knowledge".

As in all popular fields, one often finds in descriptions of "Complexity" misleading or hyped statements, and metaphorical deviations. So some caution and cynicism is required in filtering the raw data from a search, especially over the Internet.

In conclusion, explore the exercises below and hope for some enlightenment.

1.4 Exercises

1. Read as many of the articles in Refs[8-11] and then
   (a) Engage in a critical discussion with your colleagues,
   (b) Next, reread the Introduction and try to identify the keywords highlighted there with those in the above mentioned articles.

2. Try to find examples from your daily life or systems around you that illustrate the keywords used in the introduction.

3. Do you think there are other characteristics of complex systems that have not been mentioned above ? You might want to search for examples in Refs.[6-7].

4. To get some ideas for your project,
   (a) Read the guidelines and some suggestions in the course outline on the course webpage.

(b) For more ideas, look at Refs.[12-15], the other references mentioned below, some links on the "Useful Web Resouces" page of the coures webpage, and also the books mentioned in the references. The books are available in the NUS library and some even in the National Library branches.
(c) Come discuss your potential project ideas with me and ask for suggestions. (**Remember: Your project must be approved by me before you start!**)

1.5 References

1. Web of Life, by F. Capra
2. How Nature Works, by P. Bak
3. The Self-Made Tapestry, by P. Ball
4. The Computational Beauty of Nature, by G. Flake
5. Butterfly Economics, by P. Omerod
6. The Santa Fe Institute at http://www.santafe.edu/
7. The Center for the Study of Complex Systems at Michigan at http://www.pscs.umich.edu/complexity.html
8. From Cells To Whales: Universal Scaling Laws In Biology at http://www.sciencedaily.com/releases/1999/02/990202072253.htm
9. Nature's Bottom Line by Eric Roston at http://www.time.com/time/personal/article/0,9171,1101010528-127256,00.html
10. Inspiration for optimization from social insect behaviour by E. Bonabeau, M. Dorigo and G. Theraulaz at http://www.nature.com/cgi-taf/DynaPage.taf?file =/nature/journal/v406/n6791/abs/406039a0_fs.html
11. A review on self-organized criticality at http://xxx.lanl.gov/abs/cond-mat/9906077
12. Complexity (journal) at http://www3.interscience.wiley.com/cgi-bin/jtoc?ID=38804n

13. Wars and Chaos at
http://americanscientist.org/articles/95articles/Saperstein-full.html
14. Web of Human Sexual Contacts at
http://cps-www.bu.edu/∼amaral/Sex_partners/Content_sex.html
15. Some references to complexity at http://www.calresco.org/applicat.htm
16. A Fish schooling applet at
http://www.codepuppies.com/∼steve/aqua.html
17. The double pendulum applet at
http://www.maths.tcd.ie/∼plynch/SwingingSpring/doublependulum.html

Chapter 2

Fractals: Geometrical Complexity from Simple rules

How does one describe the skyline of a mountain range, the outline of a tree or a coastline, or the spatial structure of a snowflake? A description in terms of regular geometric shapes of Euclidean geometry, formed from straight lines and smooth curves, appears inadequate to capture the intricate structure of such objects. Benoit Mandelbrot introduced the word **fractal** to describe such "shapes made of parts similar to the whole in some way". For a start we may adopt the following definition.

Definition of Fractal (informal): A fractal is a shape that appears self-similar on multiple spatial scales, that is, any piece of it looks like the whole after a change of scale (magnification).

Figure 2.1 A fern.

Figure from *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation*. Copyright © 1998–2000 by Gary William Flake. All rights reserved. Permission granted for educational, scholarly, and personal use provided that this notice remains intact and unaltered. No part of this work may be reproduced for commercial purposes without prior written permission from the MIT Press.

The technical term that describes self-similarity of shapes under change of observation scale is **scale-invariance**. Systems that are scale-invariant do not have any characteristic length, that is a typical or mean length. For example, if one observes an aerial photograph of a coastline, it is difficult to guess the actual size of the features unless some man-made objects are also visible. This is because, as mentioned above, coastlines are self-similar on a wide range of scales, or approximately scale-invariant, while man-made products have a natural characteristic length: A car has a characteristic length of about 5 m, a house about 10m and Man has a characteristic length of 2m. Objects that have a characteristic length scale look different at different magnifications: for example, the fingers, the arm, and the torso of a human body are not self-similar shapes.

2.1 Mathematical Fractals

Natural fractals are self-similar only over a limited range. Also sometimes the self-similarity is only statistical rather than exact. In this section we will study mathematical fractals that are self-similar at all scales and are exact. Such fractals are generated by iteration, that is, by repeating a procedure a number of times. The number of times (steps) the procedure is performed can be kept track of by an integer k.

2.1.1 The Koch Curve and Snowflake

This fractal is generated by iteration as follows. The *initiator*, the initial or $k = 0$ step, is a unit line element. The first step, $k = 1$, called the *generator*, involves removing the middle one-third of the unit line and replacing it with two line segments each one-third in length as shown in the figure. The figure now contains four equal line segments. In the next step, $k = 2$, each of the four line segments is replaced by the (scaled) generator, leading to a figure with 16 segments. The procedure is repeated endlessly, $k \to \infty$, to generate the Koch curve.

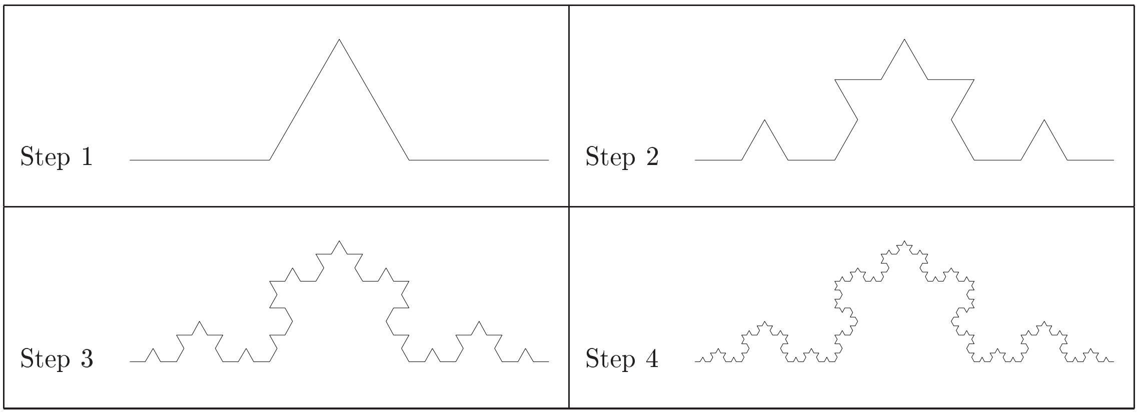


Figure 2.2 The first few steps in constructing the Koch curve

Note that the Koch curve, with its exact self-similarity at all scales, is obtained only after an *infinite* number of iterations, at the $k = \infty$ step. The figures leading up to the Koch curve are called *prefractals* and at large k these already look like the final fractal because of the finite resolution of our eyes.

How long is the Koch curve? Clearly at each step in its generation, the length increases by a factor of 4/3 as a line segment of one-third unit is replaced by two of equal length. Therefore the length of the prefractals diverges as the number of steps increases, leading to an infinite length for the Koch curve! The reason for this is of course obvious since the Koch curve is not smooth but infinitely 'kinky' (In fact the curve consists entirely of corners !)

Three Koch curves can be fitted together to form a Koch snowflake. (Try it. Does the result look like a real snowflake ?) Alternatively, one can start with an equilateral triangle and apply the generator of the Koch curve repeatedly to each line segment. Again, the Koch snowflake has an infinite length, but its area is bounded by that of the circle that circumscribes the original triangle.

The Koch snowflake, with its infinite length but finite area is an unusual object from the viewpoint of Euclidean geometry where objects occupy finite lengths, areas and volume in finite space. In order to characterize such objects we have to generalize our notion of dimension.

2.2 Dimensions

We live in a three-dimensional world: That is, we need three coordinates to specify the location of any point. The coordinate system we use locally might be the orthogonal (right-angular) Cartesian grid with x-y-z axes, or more practically for a global description, a location of a point in terms of latitude, longitude and altitude. When we restrict our attention to a subset of the world we can often make do with a description in terms of fewer coordinates. For example, a point on the surface of a table can be described in terms of two coordinates.

Sometimes a description in terms of fewer coordinates is a useful and economical approximation. Consider a ball of string lying on top of a table. From a very large distance it might appear to someone as a point object, its location in space given by three coordinates: Far away it is a point and so zero dimensional. When one approaches closer it will look disc-like and two-dimensional. Even closer up the three dimensional structure of the ball will become apparent. Finally one realizes that the ball is made of string, a one-dimensional object, and so any point on the ball of string can be located by tracing out a single coordinate along the string!

In the example above we have used two definitions of dimensions. Firstly there is the Euclidean dimension (D_e): The number of coordinates required to specify an object. Secondly there is the Topological dimension (D_t), which, roughly speaking, is a measure of the intrinsic dimension of the object. For example, a thin string has topological dimension one but when it is spread out in space, as in the ball, it has a Euclidean dimension of three.

Topology is often called "rubber" geometry because it deals only with the qualitative shape of an object. If the object is imagined to be made of rubber then by stretching (but without tearing) it can be deformed into another topologically equivalent object. Thus, a curve of any shape is topologically equivalent to a straight line, and has a topological dimension of one.

The Euclidean and topological dimensions are always integral. For characterizing fractals it is useful to introduce another definition called the *similarity dimension* which is often fractional. To motivate this definition, consider first a unit Euclidean line, square and cube, each divided into N equal self-similar parts of linear dimension s.

For the line, since $Ns = 1$, each smaller part has a length $s = 1/N$. For the square, $Ns^2 = 1$, so each smaller square has length $s = 1/N^{0.5}$, while since $Ns^3 = 1$ for the cube, each smaller cube has length $s = 1/N^{1/3}$. The values of s above are called scale factors.

Definition of Similarity Dimension: If an object of unit size contains N self-similar copies of itself of size s, then its similarity dimension D_s is determined by the equation

$$Ns^{D_s} = 1\,. \tag{2.1}$$

For the Euclidean figures above, $D_s = 1$ for the line, $D_s = 2$ for the square and $D_s = 3$ for the cube. These numbers are identical to the topological and Euclidean Dimensions for these figures. Let us rewrite the above equation in the form

$$D_s = \frac{\log(N)}{\log(1/s)}\,. \tag{2.2}$$

Now we can find the similarity dimension of the Koch curve. At each observation scale, the curve contains 4 self-similar copies of itself of size $s = 1/3$, so

$$D_s = \frac{\log(4)}{\log(3)} = 1.2618... \tag{2.3}$$

Thus the similarity dimension of a Koch curve is larger than its topological dimension which is one, but smaller than its Euclidean dimension of two. Since D_s for a Koch curve is larger than that for a line but smaller than that for area, one can roughly say that the Koch curve is more than a line but not quite a plane. We are now in a position to appreciate the following formal definition of a fractal:

Definition of Fractal (formal): A fractal is an object whose similarity dimension is larger than its topological dimension.

An equivalent way of thinking about equation (4.18) and fractal dimensions is as follows. Imagine measuring everthing relative to the length of a measuring stick. So if a smooth curve can be covered using N units of that stick of measure s, the length of the curve can be estimated as Ns. If a stick half in length, $s/2$, is used instead, one would need $2N$ units to cover the same smooth curve, and the final length of the curve will come out to be the same. For a fractal curve the above is no longer true. As the size of the measuring stick is changed, the total length of the curve changes, as we saw for the Koch curve. Historically this fact was observed by Richardson: He found that the length of some borders between countries seemed to increase when the length of the measuring instrument was reduced. Can you explain in physical terms what is happening in this case ?

Here is another example of a fractal. Mandelbrot discovered that noise in telephone lines is clustered and can be modeled as a Cantor set. The Cantor set is generated as follows: The initiator is a unit line element. The generator involves removing the middle one-third of the unit line. After this first step the figure consists of two line segments, each one-third in length. The procedure is repeated endlessly, each time removing the middle one-third of the remaining line segments.

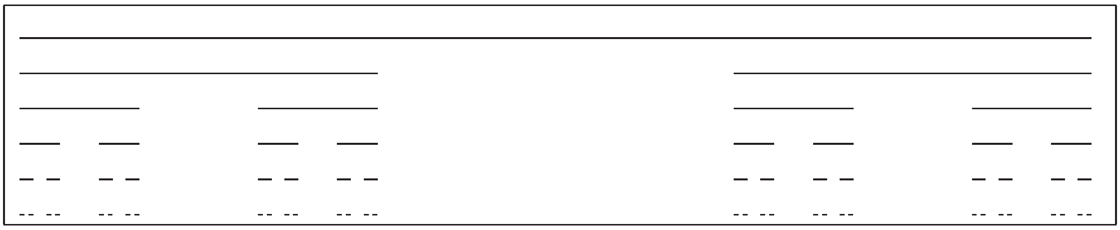


Figure 2.3a The first few steps in constructing the Cantor set

Figure from *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation*. Copyright © 1998–2000 by Gary William Flake. All rights reserved. Permission granted for educational, scholarly, and personal use provided that this notice remains intact and unaltered. No part of this work may be reproduced for commercial purposes without prior written permission from the MIT Press.

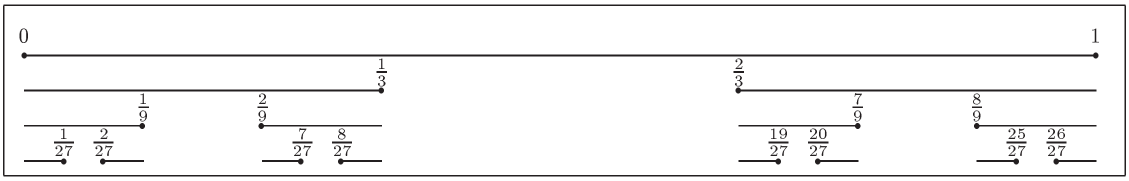


Figure 2.3. The Cantor set with points labeled

Figure from *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation*. Copyright © 1998–2000 by Gary William Flake. All rights reserved. Permission granted for educational, scholarly, and personal use provided that this notice remains intact and unaltered. No part of this work may be reproduced for commercial purposes without prior written permission from the MIT Press.

The Cantor set has an infinite number of points but it is of width zero (technically one says it is of measure zero). The Cantor set is in some sense the opposite of the Koch curve: In the generation of the Cantor set, at each step (prefractal) the line segment was made shorter by one-third while in the Koch curve it was made longer by one-third. Therefore one would expect that the Cantor set is in fact 'less than a line', just as the Koch curve was considered 'more than a line'. This is confirmed by computing its similarity dimension. What are the topological, Euclidean and similarity dimensions of the Cantor set?

2.3 Random Fractals

Natural objects do not contain identical scaled down copies within themselves and so are not exact fractals as described above. However, natural objects can often be classified as random fractals, meaning that each smaller part of it is statistically similar to the whole. Random fractals can be generated by modifying the iteration process of the last section to include a probabilistic element. Consider the generation of a random Koch curve. The initiator and generator are as before, but in the following steps ($k = 2$ onwards), the prefractals are obtained by replacing each line segment with the generator in such a way that the triangle of the generator points randomly (for example, determined by a coin toss) to either side of the original line. The figure for the final fractal shape looks very irregular compared to the exact Koch curve but is closer to the shape of natural objects such as coastlines.

Just as for exact fractals, one can introduce a dimension to characterize random fractals. One example is the *box counting dimension.* In this method, space is divided into equal sized cubes (or squares if the figure lies on a plane) of linear dimension r. One then counts the number of cells, $N(r)$ that are needed to cover the given shape. If

$$N(r) \propto r^{-D_b} \tag{2.4}$$

as the length r is changed, one says that the distribution of points is D_b-dimensional. This definition obviously agrees with the Euclidean dimension for straight lines and planes but gives fractional values for more complicated shapes such as coastlines. Note that the equation above is of the same form as that which comes from the definition of the self-similarity dimension mentioned above.

2.4 Occurrence and uses of Fractals

Consider the Amazon river system. The box-counting method can be used to determine its fractal dimension, giving $D_b = 1.85$. By comparison the fractal dimension of the Nile river is about 1.4 [Takayasu, Section, 2.1.2].

The branching of a tree looks like that of a river and its fractal dimension lies between 1.3 and 1.8, with a mean of 1.5. [D. Morse, et. Al., Nature, 314 (1985) 731].

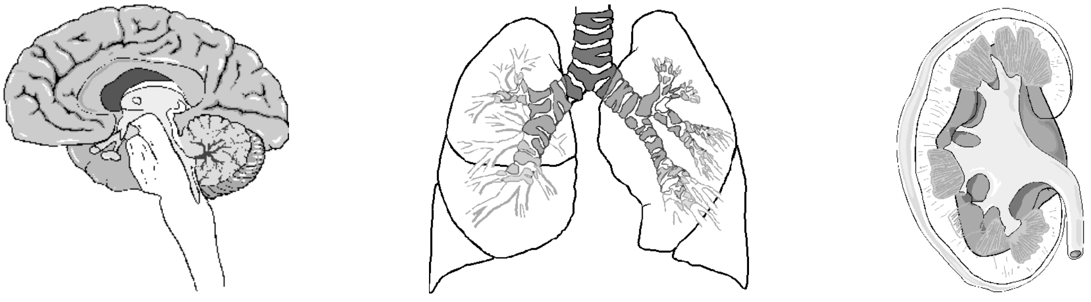
The shape of lightning discharges is similar to that of rivers. The fractal dimension in this case is about 1.7 [L. Niemeyer, et.al. Phys. Rev. Lett., 52 (1984), 1033].

Consider the diameter distribution of blood vessels in a bat's wings. If $N(r)$ Is the number of vessels thicker than r, then $N(r) \propto r^{-2.3}$, [Takayasu, Sect.2.2.1].

Given a time-varying signal, one may consider its power spectrum, that is, the Fourier transform of the intensity. The power spectrum indicates the relative magnitude of different frequency components in the signal. Empirically it is found that fluctuations (or noise) in many man-made and natural systems have a power spectrum of the form f^{-a} where $a \sim 1$. This is generically referred to as a "$1/f$" law.

Note that if $a = 0$ then all frequencies would have the same magnitude (the fluctuations are random) and one has 'white' noise. For a close to one, which is what is observed, the lower frequencies dominate and one has "pink noise". The $1/f$ noise has been observed in electrical circuits, voltage fluctuation of nerve cells, heart beats and even in music. This universal occurrence, which cannot be due to random fluctuations that give rise to white noise, demands a simple explanation but there is currently none (see however the later chapter on self-organised criticality). The $1/f$ noise is of course fractal because of its self-similarity. Finally, if $a \geq 2$ one has "brown noise" which for audio signals sounds dull and is not observed in nature. Thus it seems that the naturally occuring and interesting "$1/f$ noise" is poised between the randomness of white noise and the dullness of brown noise.

Look at the pictures of branching air passages in the lung, the branching of blood vessels in the human body, or the folds on the surface of the brain.



(a) (b) (c)

Figure 2.5 Naturally occurring fractals in the human body: (a) brain, (b) lungs, (c) kidney

Why should nature design such fractal structure ? Recall from the example of the Koch curve that one can accomodate long lengths in small areas. Thus nature apparently maximises functional efficiency while using minimum space by adopting fractal structures.

Fractals images have been used for image creation in science-fiction movies (e.g. Star Wars) and also for data compression (If you are interested, check out the work of Lindenmayer and Barnsley and explore the related fractals using the FRACTINT software). In case you have not guessed it yet, Fig.(2.1) is not an image of a real fern but a computer simulation! Here is another simulation that looks natural:

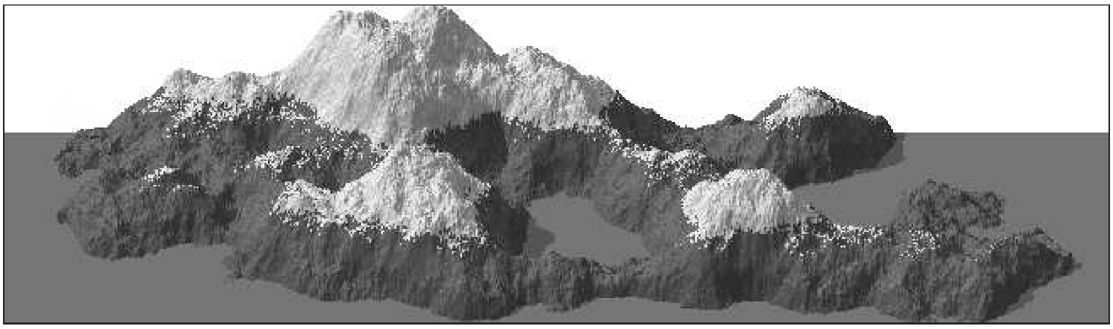
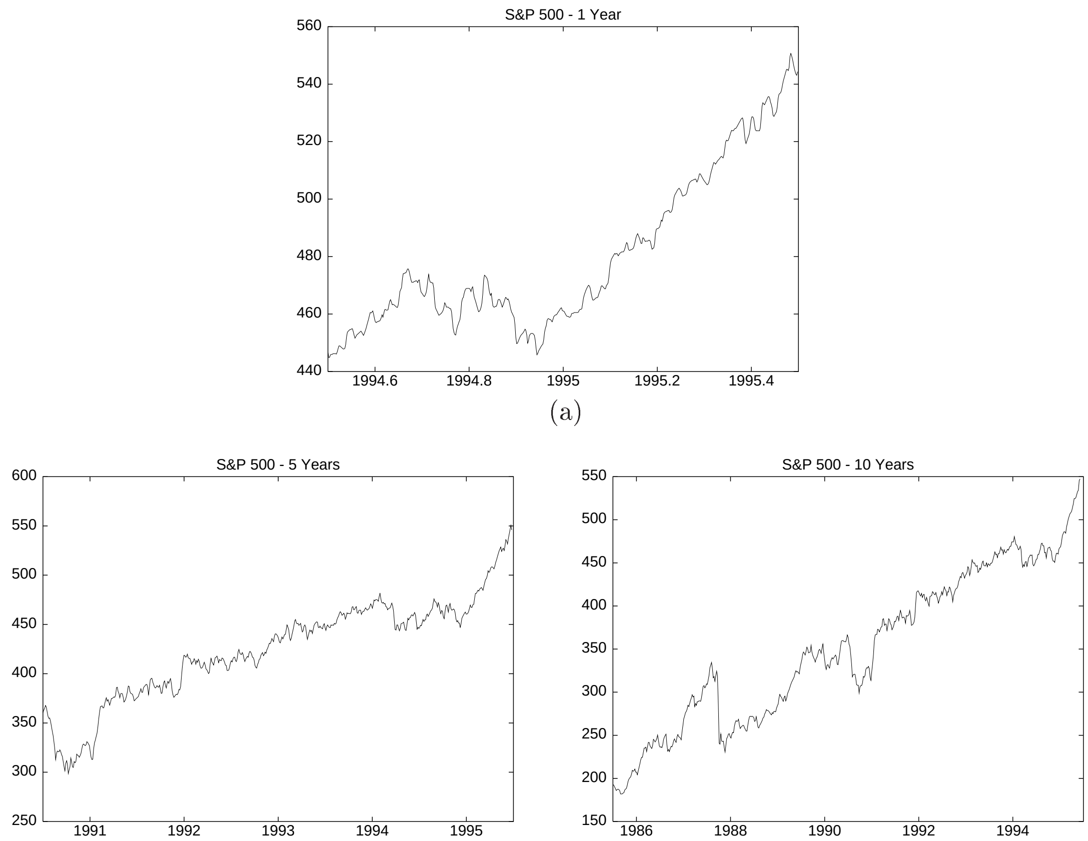


Figure 2.6 Fractal mountain generated by `xmountains` (an X Windows program written by Steven Booth)

The graphs shown below plot the actual S&P 500 stock index for one year, five year, and ten year spans. Do the time-series look self-similar to you ?



(a)

(b)

(c)

Figure 2.7 The S&P 500 stock index shown on various time scales. (a) one year, (b) five years, (c) ten years

Now look at Mandelbrot's paper on finance [Mandelbrot, Scientific American Feb 1999, pg.73] . Some of the figures there are of real data but some have been artificially generated using fractal concepts. Can you tell them apart? These last two examples show that time series in finance are generated by complicated processes that are apparently self-similar on many scales.

2.5 Dynamical generation of some random fractals: Diffusion Limited Aggregation

Some random fractals, such as the clusters describing a bacterial colony, can be generated by a physically motivated model called 'diffusion limited aggregation' (DLA). Consider for simplicity the formation of such a cluster in the plane, with the initial (seed) particle located at the origin. Other particles are then released far from the origin, at random locations, and allowed to diffuse: Mathematically this is done using an algorithm such as a 'random walk' to simulate the diffusion process. When the diffusing particle encounters the seed particle it is made to stick to it. The process is repeated with other diffusing particles, leading to the formation of a cluster. As the cluster forms, there is a greater probability for particles to stick to the ends than to penetrate the interior. Hence this leads to the formation of a branch-like structure emanating from the origin.

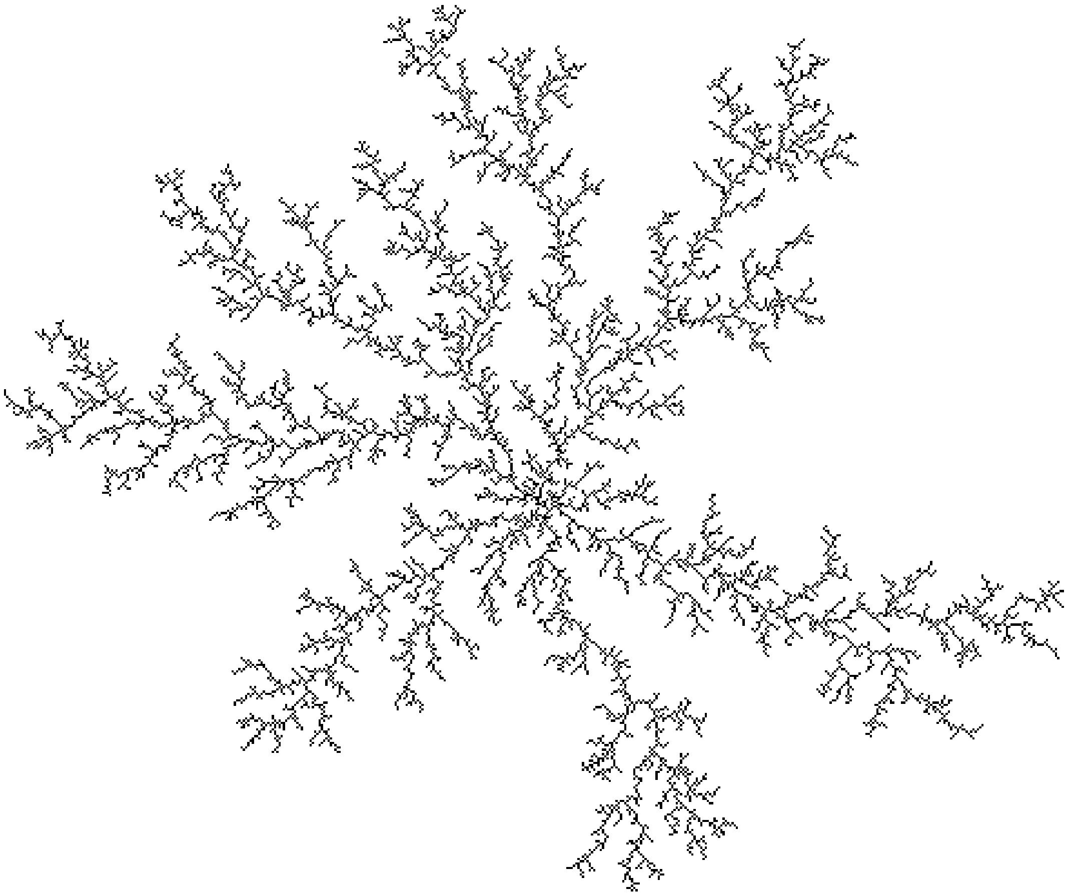


Figure 2.4 An example of growth from diffusion limited aggregation

Some other examples of DLA are in the growth of crystals (e.g. snowflakes) and coral reefs. (Note: A 'random walk', colloquially referred to as the 'drunken man's walk' is a path generated by a random process. Consider a two dimensional random walk starting at the origin. The location of the the next step is generated, for example, by two random numbers which give respectively the direction (angle) and length of the walk. The random process is repeated at the following time intervals. See the Exercises.)

Another dynamical 'explanation' of the ubiquitous occurrence of fractals and power laws in nature is the idea of 'self-organised criticality' that we will discuss later in the course.

2.6 Scaling Laws in Biology

Several power laws are known in biology. For example, it had been known for decades that the metabolic rate of an organism, B, scaled with the mass of the organism, M, like $B \propto M^{3/4}$. This relation was found to be valid over many orders of magnitude, from a tiny mouse to an elephant. Not only was the universality of the exponent a surprise, so was its value. Different naive arguments would have suggested values of 1 or 2/3. Although some suggestions were made over the years to explain the exponent value of 3/4, these were restricted to certain animal types. A general model was presented only quite recently by Brown, Enquist and West [10].

In the BEW model, three assumptions (supported to varying degree by empirical evidence) were made concerning the transport network of nutrients in an organism: (1) The network is a volume-filling hierarchical branching system (fractal), (2) The smallest branch is of the same size in different animals, (3) The energy required to transport materials through the network is to be minimised.

The remarkable simplicity of the model, and its explanation of the 3/4 exponent makes it very attractive. Several follow up investigations have been done, see 11. It is interesting to note that the numerical value of 3/4 arises in the model through the expression $\frac{d}{d+1}$ where $d = 3$ is the dimension of space 12.

2.7 Fractals and Art

Jackson Pollock was an abstract painter with an unusual (even by abstract standards!) style: His paintings were huge, typically 2-4 metres each side, and he painted by dripping or streaking paint from a can using a large brush or stick. His paintings have a “natural” look to them, similar to undergrowth in a forest or a thick bush, and this might be the reason they appear pleasing to many.

Richard Taylor analysed Pollock's paintings, concluding that they were fractal (!), with a dimension D that increased from close to 1 to about 1.72 during the evolution of the artists technique.

2.8 Summary

Fractals are shapes that are self-similar on multiple scales (spatial or temporal), and have (in general) fractional dimensions. Natural fractals are of course self-similar only over a limited range, and the similarity is often statistical rather than exact. Nature apparently chooses fractal structures to optimize functional efficiency given limited resources. (An interesting recent article is Ref[5].

Physical growth mechanisms, such as DLA, are reasonable explanations for some natural fractals, while for fractals in biological systems some algorithms have been proposed (such as those of Lindenmeyer and Barnsley) that give realistic pictures.

Though natural fractals have intricate structures, it is remarkable that one is able to 'explain', at least qualitatively, the complex geometry of nature using computer simulation of models obeying simple rules. Thus in this case one might say that the *emergence* of complex patterns and structures can be understood as arising from simple underlying causes. However note that to capture the depth of self-similarity, one has to iterate the algorithms over many steps, and it is often not possible to guess what the final structure will look like until the numerous steps have been performed.

Self-similarity implies that a system is scale-invariant, or equivalently, it means the absence of a charactersitic length scale. Scale-invariance will be revisited when we study phase transitions and self-organised criticality.

Look out for fractals later in the course when we study chaos, phase transitions and self-organised criticality. In the meantime, read about some technological apllications of fractals at Ref.[**??**].

2.9 Exercises

1. Explain why fractals in nature are self-similar only over a limited range. Is there an upper limit or lower limit or both?

2. Many movies use miniature models in their special effects. Discuss how the human brain is tricked into thinking that those models are life sized and how this relates to the concept of characteristic length and/or self-similarity.

3. Look at natural objects (or pictures of them) that are frequently referred to as fractal. Do you notice the self-similarity on multiple length scales?

4. Can you draw a unique tangent at any point of a Koch curve? Why not? What is the mathematical terminology to describe such curves?

5. Suppose the initiator of the Koch curve was 1cm long. Approximately how long would the Koch curve be after 100 iterations? If the curve was a piece of thread that could be stretched to its full length, could you use it to tie the Earth to its Moon ?

6. Take a piece of paper 2cm by 2cm. Make a hole in it big enough for you to push a glass of water through it. (Hint: This question is not really about fractals but is used to illustrate that you can have long lengths in small areas as in the Koch snowflake).

7. What is the area of the Koch snowflake if the initiator was an equilateral triangle of unit length on each side?

8. What are the topological and Euclidean dimensions of the Koch curve?

9. Starting from the definition of self-similarity dimension, derive the explicit expression for the self-similarity dimension of a fractal. What is the base of logarithms used in the equation?

10. Calculate the self-similarity dimensions of the Koch curve using the scales $s = 1/9$ and $s = 1/27$.

11. (a) Give examples of objects that are self-similar over a wide range of scale but which are not fractals.
    (b) Give an example of an object which is a fractal but does not have fractional dimensions.

12. Show that the length of the initiator that remains once a Cantor set is formed is zero. Convince yourself that the Cantor set is a fractal, by showing that it is self-similar and also by computing its similarity dimensions and comparing that with its Euclidean and topological dimensions.

13. (a) Determine the area and perimeter of the Sierpinski carpet.
    (b) Determine the volume and area of the Menger sponge and compare the result with that of a real sponge.
    (c) Determine the topological, Euclidean and self-similar dimensions of the quadratic Koch curve (Minkowski sausage), the Sierpinski carpet, and the Menger Sponge.
    (d) In what way are the Cantor set, Sierpisnki carpet and Menger sponge related ?

14. Search for two examples of fractals in your environment that are almost exact, and two examples of those that are random.

15. Do you think the fractal (self-similarity) dimension of a two-dimensional projection of a tree is more or less than 2? Why?

16. Look at the pictures of branching air passages in the lung, the branching of blood vessels in the human body, or the folds on the surface of the brain.
    (a) Suggest reasons why nature would have designed such fractal structures, that is, what is to be gained by adopting such a structure? Also suggest reasons for the fractal structures of trees. (Hint: See exercise (13) above).
    (b) Where are the instructions for growth in biological systems encoded?

17. Measure the fractal dimension of a long coastline given in your Atlas using the box-counting method.

18. The coastline in the above exercise may be approximated by a set of straight lines each of fixed length, and its total length estimated by the sum (the "structured walk technique"). Determine the number $N(r)$ of straight-line segments of length r that are required to approximate the figure. Plot the results on a log-log plot and determine the fractal dimension of the coastline by this method. Compare your result with the box-counting method.

19. There exist many online web simulations of random walks. Explore some of the one and two dimensional versions. See for example Ref.[7].

20. Explore the DLA simulation at Ref.[8]

21. Optional: Look at the graphs in Mandelbrots paper on finance [Mandelbrot, Scientific American Feb 1999, pg.73]. Can you make out which of the graphs represent real data and which fractal simulations? Can one make money out of this?

22. Widen your horizons by browsing through the articles in Refs[5, 6].

2.10 References

1. The Computational Beauty of Nature, by Gary William Flake.
   See also the books homepage at http://mitpress.mit.edu/books/FLAOH/cbnhtml/home.html

2. Some of the figures (**those with a copyright notice in smallprint in the figure caption**) used in this chapter were taken from the website of Ref.[1] while some other figures used in the lectures were taken from Refs[3,4].
3. Fractals and Chaos, by P.A. Addison.
4. Fractals in the Physical Sciences, by H. Takayasu.
5. Some frontier interdisciplinary research in fractals that might have solved a long-standing mystery of nature is reported in
   http://www.sciam.com/explorations/1999/061499power and also
   http://www.santafe.edu/sfi/publications/Bulletins/bulletin-summer97/feature.html
6. Fractals in art in
   http://www.sciencenews.org/sn_arc99/9_18_99/mathland.htm
7. Some random walk simulations on the web are
   (a) http://stp.clarku.edu/simulations/one-dimensional-walk.html
   (b) http://math.furman.edu/ dcs/java/rw.html
8. A DLA simulation is at http://apricot.ap.polyu.edu.hk/ lam/dla/dla.html
9. Mandelbrots article on multifractals in finance is in Scientific American Feb 1999, pg.73.
10. Brown, Enquist and West, ”A General Model for the Origin of Allometric Scaling Laws in Biology” , Science 276 (1997): 122-126.
11. Brown and West (eds), Scaling in Biology, Oxford University Press (2000).
12. Chee Jing Yee and Hong Wenxian, ”Size and Scaling”, in Proceedings of the 8th Youth Science Conference, 2002, Singapore.

Chapter 3

Chaos

I accept chaos. I am not sure that it accepts me.
Bob Dylan

3.1 Introduction

When the meteorologist Edward Lorenz was studying a simplified model of the weather he discovered that tiny differences in the initial conditions led rapidly to very different final results. This came as a surprise to everyone as it was always assumed that small errors in any dynamical system would lead to small corrections. Indeed in practice we never have the exact information about initial conditions and we always deal in approximations of some sort. The sensitive dependence on initial conditions came to be called the *butterfly effect* : A butterfly flapping its wings in Aruba might completely change the weather in Bali!

The fact that a completely deterministic system could lead to results that were essentially unpredictable came to be called chaos. A necessary condition for chaos is that the equations for the system are non-linear as errors in linear systems remain small if they were initially small. As most of the dynamical systems in real life are described by nonlinear equations, it is expected that chaos will be commonplace. However one must remember that even for the simple Lorenz model chaos appears only for some values of the control parameter (a free parameter that occurs in the equations), and so real life nonlinear systems might not always be chaotic. That is, nonlinearity is not a sufficient condition for chaos.

3.2 Dynamical Systems and Iterative Maps

Whereas a real dynamical system, such as the motion of the planets, is described by differential equations and continuous time, it is often convenient to consider simpler mathematical models, called difference equations, where the system evolves through a set of discrete time steps. The simplest models of difference equations are the iterative maps because the future value of some variable at time $t = (n + 1)$ depends only on its value at the present time $t = n$ (here n is an integer). An iterative map is of the form

$$x_{n+1} = f(x_n)\,, \tag{3.1}$$

where $f(x)$ is called the *mapping function*. Thus starting with some initial value x_0 of the variable, its next value x_1 is obtained by evaluating $f(x_0)$, which itself becomes the input to evaluate x_2 and so on. Thus the time-evolution of the discrete variable x is obtained by repeated iteration of the mapping function. Note that this mathematical iteration corresponds in physical terms to a feedback process.

3.3 A Simple Example of an Iterative Map

3.3.1 An Intuitive Analysis

Consider the map

$$x_{n+1} = x_n^2\,, \tag{3.2}$$

which corresponds to the mapping function $f(x) = x^2$. It is easy to guess, or check with a calculator, what happens as one repeatedly squares a number $x_0 \geq 0$: If the initial value is greater than 1, repeated iterations give values that increase without bound. On the other hand, an initial value of exactly 1 remains at 1, while any intial value less than 1 will converge to 0.

If we restrict our discussion to finite values, then this example has two **fixed points**, one at $x^* = 0$ and another at $x^* = 1$. If one starts the iteration at exactly any of these two values one remains there, and so the name 'fixed points'. The fixed point at the origin is called **stable** because a small deviation (perturbation) from an intial value of zero causes future iterations to be attracted back to the origin. One says that the *basin of attraction*, (that is the range of values that are attracted to a fixed point), of the fixed point at the origin is $[0, 1)$. By contrast the fixed point at $x^* = 1$ is **unstable**: any small perturbation away from that value give future values that move further away from that point.

A mechanical analogy of the two fixed points, stable and unstable, is a ball lying at the bottom of a valley and one lying on a hill top. Both positions of the ball are points of static equilibrium, but one position is clearly stable while the other not. Furthermore the stability of the ball at the bottom of the valley is limited to perturbations that do not cause the ball to go over a nearby peak, and hence in general the basin of attraction of a point of stable equilibrium is of finite extent.

3.3.2 A Graphical Analysis

It is useful in general to have a geometrical picture, called a **state space** or **phase space** plot, which shows how the next iterate x_{n+1} depends on the current value x_n. This is done as follows. First draw the curve $y = f(x)$ to represent the right hand side of Eq.(3.1) . On the same graph plot the line $y = x$ to represent the iteration process $x_{n+1} = f(x_n)$. Fig(3.1a) illustrates the method for the case $f(x) = x^2$. Starting with any x_0, for example $x_0 = 0.8$, the next value $x_1 = f(x_0)$ is obtained by drawing a vertical line from $x_0 = 0.8$ to the curve $y = x^2$. Since $x_1 = f(x_0)$ is now needed as the next input to the iterative map, graphically this is achieved by moving horizontally to the $y = x$ line and then moving vertically to the 'square' curve to give x_2. The procedure is repeated to produce a generic 'cobweb' diagram. For this example Fig(3.1a) shows a step-wise approach to the fixed point at the origin.

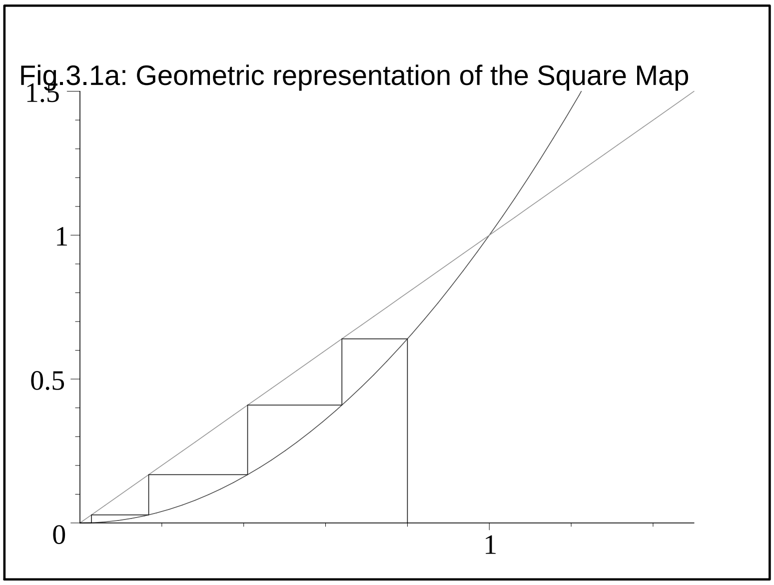


Fig.3.1a: Geometric representation of the Square Map

It is easy to see from this diagram that for the example considered, the origin is an attractive (stable) fixed point with a basin of attraction $[0, 1)$, while as Fig(3.1b) shows, the point $x^* = 1$ is an unstable fixed point.

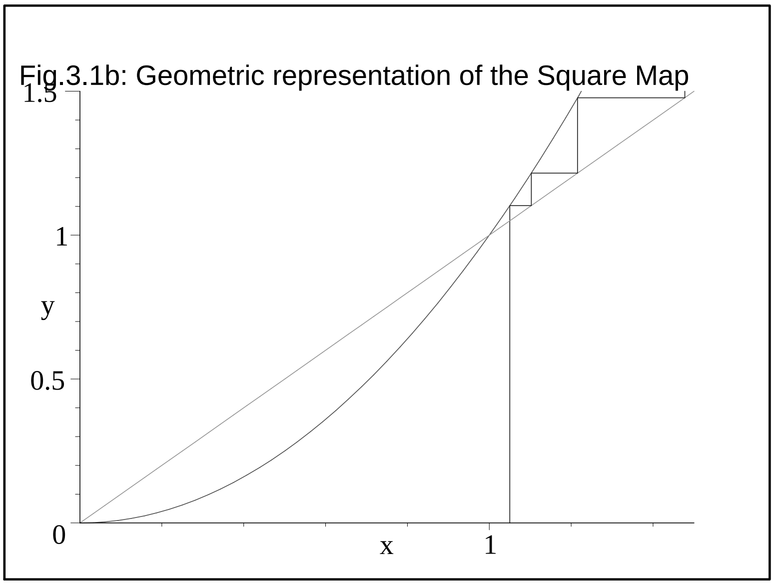


Fig.3.1b: Geometric representation of the Square Map

Since the fixed points above correspond to a situation where the future values of x at *each* time step remain the same, they are called *period-one fixed points.* This terminology is introduced as we will later generalise the concept of fixed points. Since period-one fixed points correspond to solutions of $x_{n+1} = x_n$, graphically these can be found from the intersection of the mapping function $y = f(x)$ with the staright line $y = x$. For the case $f(x) = x^2$, the two intersections are at the origin and at $x = 1$ as we already know.

3.3.3 The Analytical Approach

Analytically, period one fixed points are determined by solving the equation $x_{n+1} = x_n$. Denoting the solution as x^* and using Eq.(3.1), one obtains

$$x^* = f(x^*)\,. \tag{3.3}$$

For the mapping function $f(x) = x^2$ the solutions are $x^* = 0, 1$ as expected.

The stability of fixed points can also be analysed analytically. Consider a nearby point $x_n = x^* + \epsilon$ with ϵ a small parameter. Then using a Taylor expansion and to leading order in ϵ, one obtains

$$\begin{aligned} x_{n+1} &= f(x^* + \epsilon) && (3.4)\\ &= f(x^*) + \epsilon\,\frac{df}{dx}|_{x^*} && (3.5)\\ &= x^* + (x_n - x^*)\,\frac{df}{dx}|_{x^*}\,, && (3.6) \end{aligned}$$

and hence

$$|x_{n+1} - x^*| = |x_n - x^*|\;|f'(x^*)|\,. \tag{3.7}$$

Therefore if $|f'(x^*)| < 1$ an iteration brings the starting point x_n closer to the fixed point, implying that the fixed point is stable, at least locally. In other words, *a period-one fixed point is stable if the slope of the mapping function at that point has magnitude less than one.* Similarly, if the slope of the mapping function at a period-one fixed point is greater than one, the point is unstable. Finally, if the slope is exactly equal to one the fixed point is locally **neutral**, meaning that small perturbations lead to another equilibrium position (the mechanical analogy here is that of a ball lying on a flat surface).

For the mapping function $f(x) = x^2$, we had determined a fixed point at $x^* = 0$ and another at $x^* = 1$. Since $f'(0) = 0$, this implies that the origin is a stable fixed point, while since $f'(1) = 2$ the other fixed point is unstable. These results of course agree with the earlier qualitative discussion.

3.4 A Discrete Model of Population Growth

Difference equations and iterative maps occur naturally in mathematical biology. For example, an important problem is how the population size of a particular species varies from one generation to another. Let P_n be the population of a species at time n (corresponding say to the n-th generation), and P_{n+1} the population at time $n+1$. Then the change in population size during the time interval is given in the simplest model of population growth by

$$P_{n+1} - P_n = kP_n \, , \tag{3.8}$$

where k is the growth rate. That is, the population is assumed to increase in one time interval by an amount proportional to its value at the beginning of that interval. However, this would imply a population that increases without bound as time increases. A more realistic model is obtained if the rate of growth is not a constant but depends on P_n itself in such a way that it decreases when the population becomes too large and the species runs out of food and/or space. Therefore we put,

$$k = b(c - P_n) \, , \tag{3.9}$$

where b and c are positive constants. Clearly when $P_n = c$, $k = 0$ and the population has reached a maximum value. Substituting equation (3.9) into (3.8) gives the Verhulst model (also associated with the biologist Robert May),

$$P_{n+1} = P_n + bcP_n - bP_n^2 \, . \tag{3.10}$$

The last term in (3.10) is nonlinear and in addition it provides *negative feedback* as compared to the *positive feedback* due to the second term. By a change of variables the equation can be written in the simplified and conventional form

$$x_{n+1} = 4rx_n(1 - x_n) \, , \tag{3.11}$$

where the symbol P has been replaced with a new variable x that can be interpreted as the fractional population. Equation (3.11) is known as the logistic map and the constant r is called the control parameter (here interpreted as the reproduction rate). We wish to study the properties of the logistic map as it is evolved forward in time. This evolution is obtained simply by repeated interation of an intial value x_0 and the sequence of values taken by the successive x is called an **orbit**.

3.4.1 Varying the control parameter in the logistic map

Let us study how the orbits of the logistic map depend on the control parameter that is restricted to $0 \leq r \leq 1$. That range of r is selected because it keeps all future values of x between 0 and 1 if initially it was between those values.

If $r = 0$ then the population does not reproduce at all and becomes immediately extinct at the next time step. For $0 < r \leq 1/4$, it is easy to see that irrespective of the initial population, the population will monotonically decrease as time progresses, and will eventually become extinct. That is, the origin is the only fixed point for $0 \leq r \leq 1/4$ and it is attractive (stable) for all x_0. Figure(3.2) shows the state plot for $r = 0.2$ and $x_0 = 0.6$.

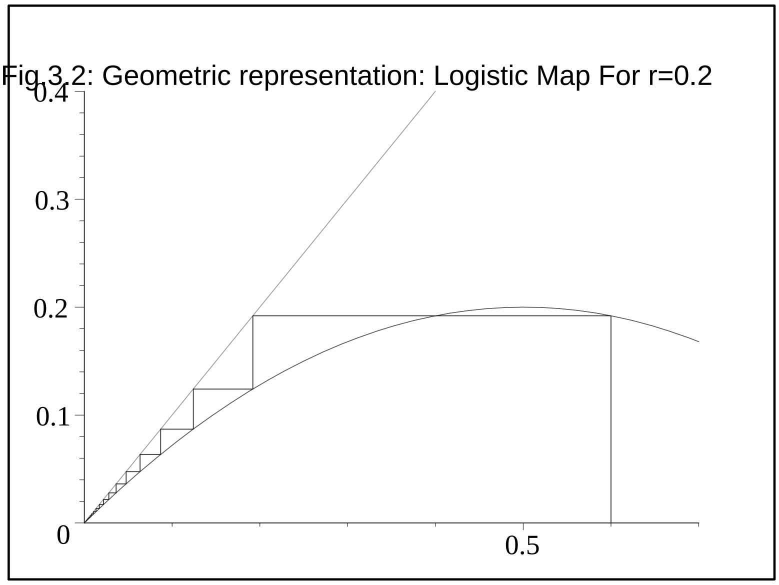
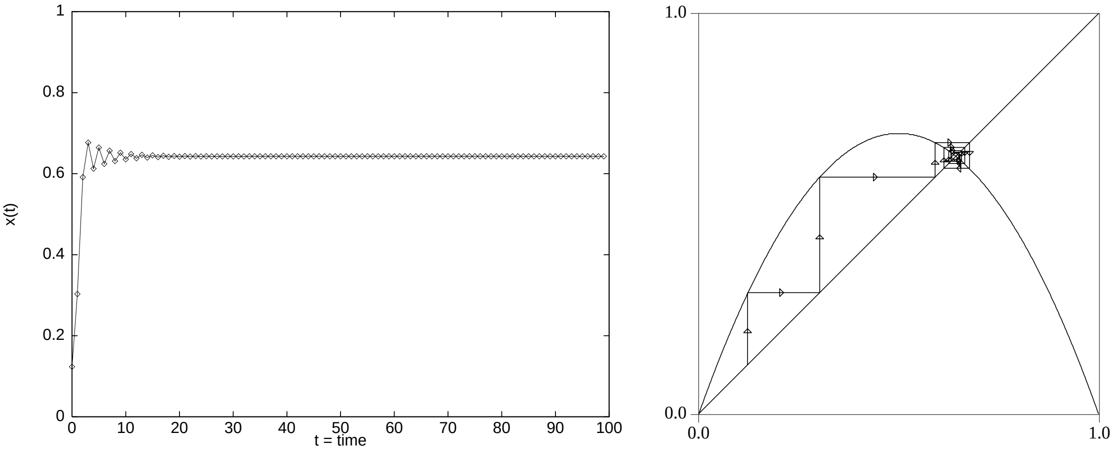


Fig.3.2: Geometric representation: Logistic Map For r=0.2

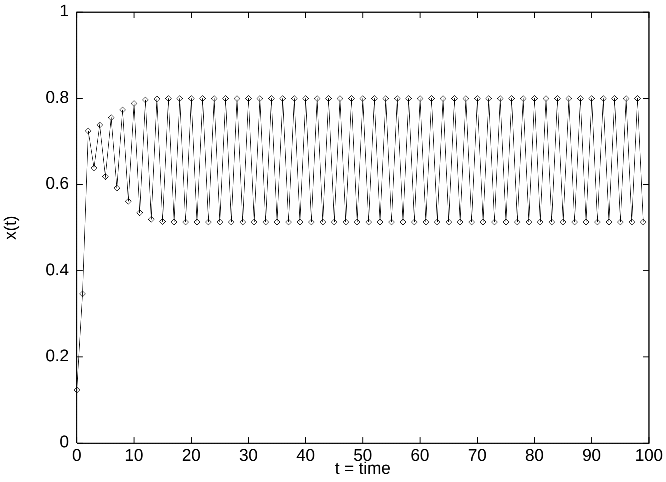
Consider now a larger value of r, say $r = 0.7$. For an intial value $x_0 = 0.1$, the corresponding time series is shown in Figure (3.3a). The initial behaviour shown in that figure is called the transient orbit, while the final sequence to which the orbit tends to is called the post-transient orbit. It is the **post-transient** orbit which is of interest. As the state-space plot in Fig(3.3b) confirms, the fixed point at the origin is now unstable, and the system is attracted to a new stable period-one fixed point. *Thus we have our first example of a phenomena whereby as the control parameter in an iterative* *map is changed, new fixed points can appear, and the stability properties of old fixed points can change.* Indeed, for $r > 1/4$ the new fixed point is at $x^* = 1 - 1/4r$ and this point is stable for $1/4 < r < 3/4$.



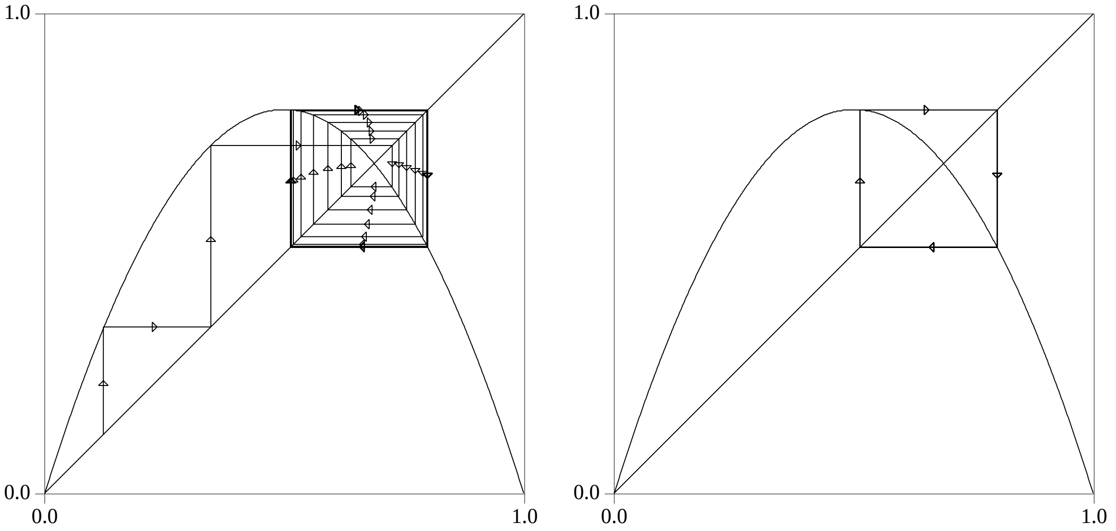
(a) (b)

Figure 3.3 Logistic map with $r = \frac{7}{10}$: (a) The time series quickly stabilizes to a fixed point. (b) The state space of the same system shows how subsequent steps of the system get pulled into the fixed point.

Next consider a value of r slightly above $3/4$, say at $r = 0.8$. We already know from the above discussion that the period-one fixed point at $x^* = 22/32$ will be unstable. In fact as Fig.(3.4) shows, the final orbit does not converge to any fixed point but rather alternates between two points. This orbit is called a period two orbit (or period two limit cycle), and the two points between which the orbit oscillates are sometimes labelled as period-two fixed points (or period two attractors) because they satisfy the the relation $x_{n+2} = x_n$.



(a)

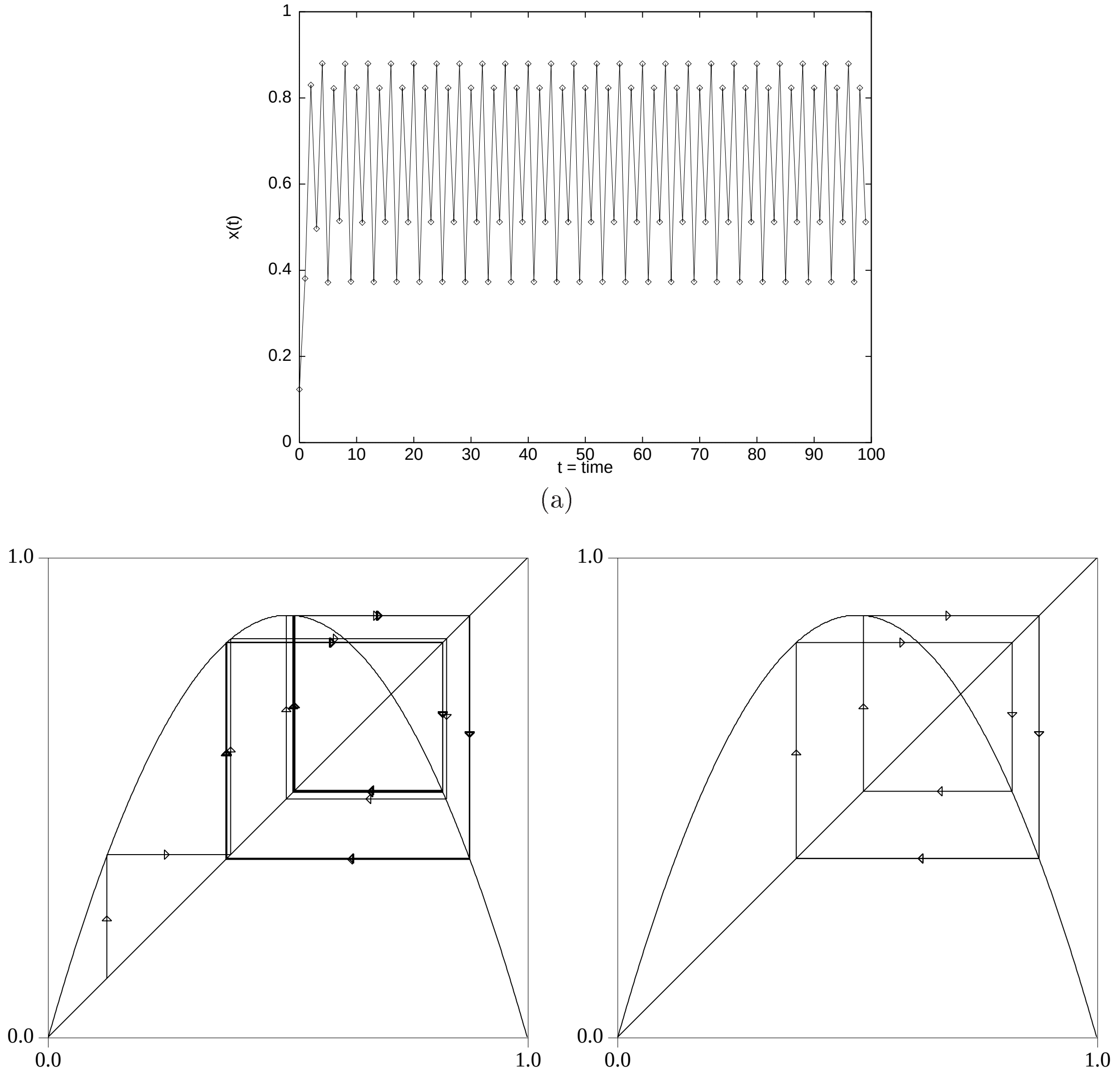


(b) (c)

Figure 3.4 Logistic map with $r = \frac{8}{10}$: (a) The time series quickly stabilizes to a period-2 limit cycle. (b) The state space of the same system shows how subsequent steps of the system get pulled into the limit cycle. (c) The state space of the same system but with only the converged values for x_t plotted, so as to clearly show the limit cycle's location.

As r is increased further, the system will at some point undergo a period-

4 cycle. Fig.(3.5) shows this for $r = 0.88$.



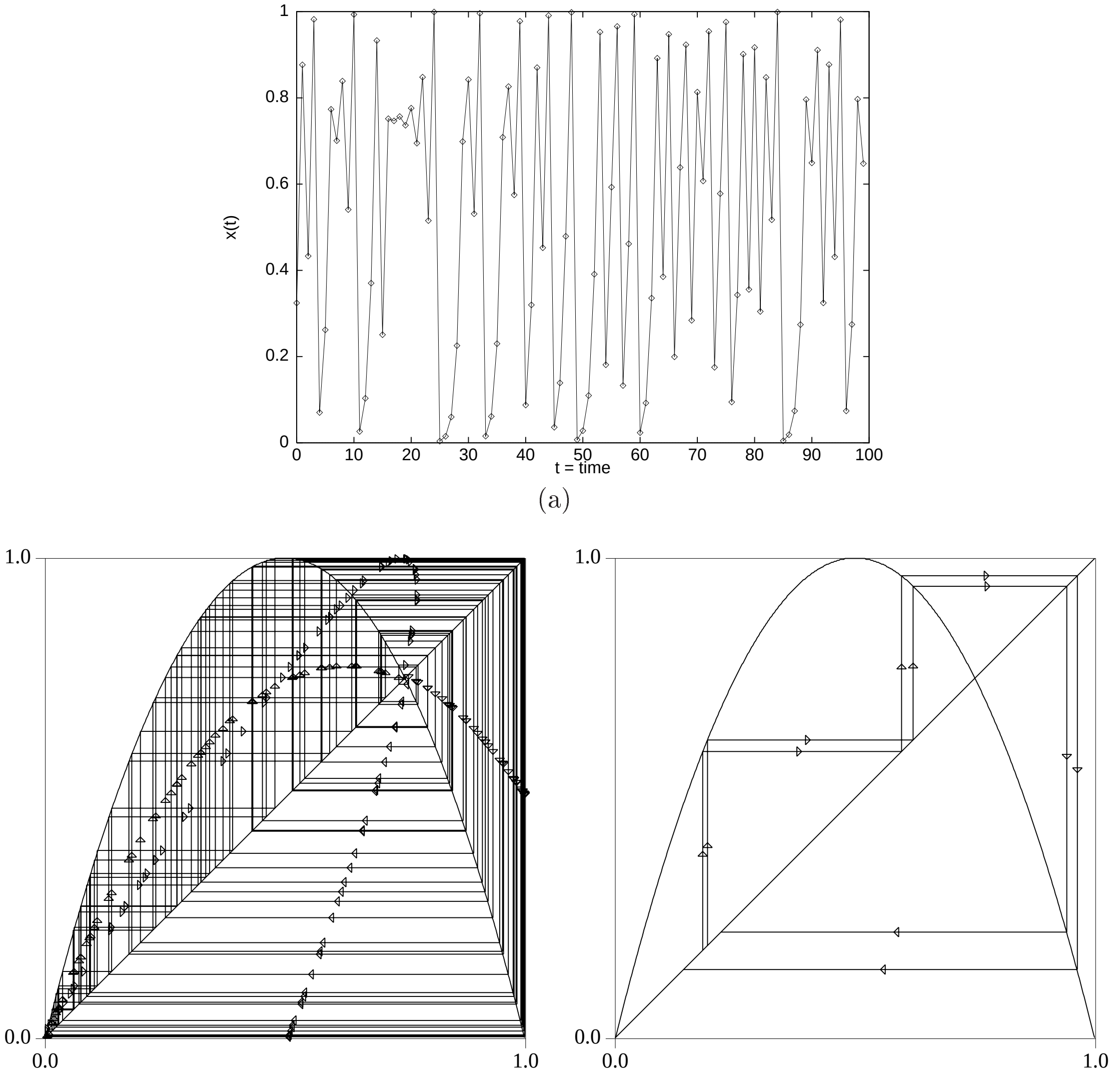
(a)

(b)

(c)

Figure 3.5 Logistic map with $r = \frac{88}{100}$: (a) The time series quickly stabilizes to a period-4 limit cycle. (b) The state space of the same system. (c) The state space of the same system but with only the converged values for x_t plotted.

These period doublings continue as r is increased until at $r = 0.8924865....$ the system shows a completely **aperiodic orbit**. Such orbits are termed 'chaotic', and though there are no fixed points now, the system is said to have a chaotic or **strange attractor**. Fig(3.6a) shows the time series for $r = 1$ which is in the chaotic regime. Fig(3.6c) shows how in the chaotic regime, small differences in the initial conditions amplify quickly – this was the **sensitive dependence on initial conditions** mentioned in the introduction as a defining property of chaos.



(a)

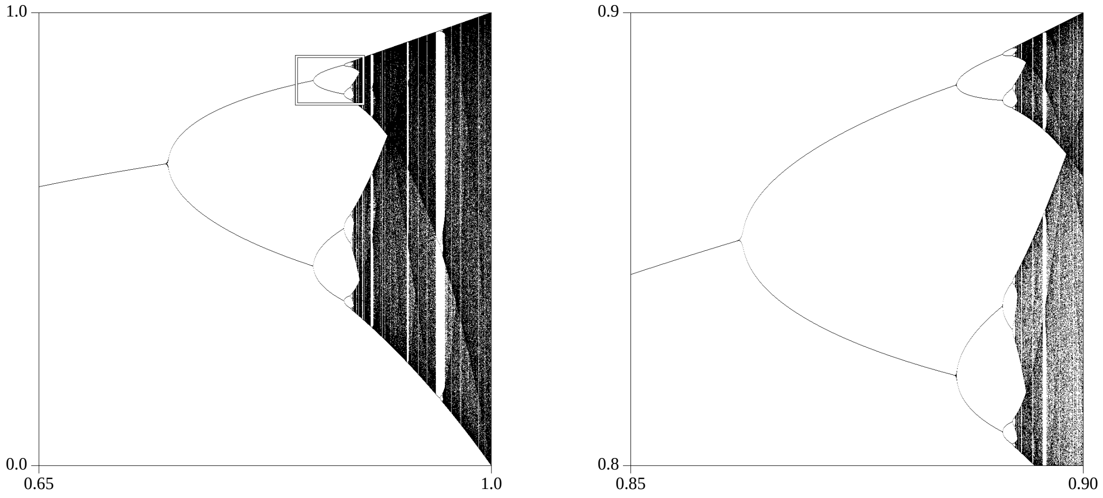
(b)

(c)

Figure 3.6 Logistic map with $r = 1$: (a) The time series is chaotic and has the appearance of noise. (b) The state space of the same system, which illustrates how the system's trajectory visits every local region. (c) The state space of the same system with only four steps plotted, so as to show how small differences turn into larger differences.

3.4.2 Bifurcation Diagrams

Since changes in the value of the control parameter r produce attractors with different periods, it is useful to plot the post-transient solutions against r. This is shown in the bifurcation diagram of Fig.(3.7a). For $r < 1/4$ all orbits decay to the attractive fixed point at the origin. As r crosses $1/4$ a new fixed point appears that is attractive for $1/4 < r < 3/4$, while the one at the origin becomes unstable. At $r = 3/4$ there is a period doubling *bifurcation* because for $3/4 < r < 0.86237...$, period-2 attractors exist. These period doubling bifurcations continue as r crosses various thresholds, leading to period 4, 8 etc attractors: In each case, as a particular threshold is crossed, new period-2^k attractive fixed points emerge, while the old period-2^{k-1} fixed points become unstable. This is illustrated symbolically in Fig.(3.8) where the dotted line represents a fixed point that become unstable as a threshold is crossed, and as two new stable fixed points emerge (the two prongs of the pitchfork) . Finally at $r = r_c = 0.8924865....$ an infinite period results (strange attractor) leading to chaos.



(a)

(b)

Figure 3.7 Bifurcation diagrams for the logistic map: (a) This image has values of r such that fixed points, limit cycles, and chaos are all visible. (b) This image shows the detail of the boxed section of (a).

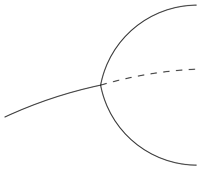


Figure 3.8 A single bifurcation

As r is increased beyond r_c, windows sometimes open up in the diagram as periodic attractors recur at some points. Notably, now one begins to see periodic attractors at odd integers. As r is further increased, these odd periodic attractors themselves undergo period doubling bifurcations to chaos. Indeed, it has been proven (Yorke) that period three orbits imply chaos. For the logistic map, there is a period three window near $0.9571.. < r < 0.9603...$

One feature of the bifurcation diagram that should be noted is its self-similarity. Fig.(3.7b) shows a magnified view of the boxed region in Fig(3.7a): It looks very much like the original whole. In fact such approximate self-similarity continues as further bifurcations take place. The self-similarity is also important in establishing some universality properties of the bifurcation process, as discussed below.

The above discussion has been for the critical values of the control parameter. To find the period 2^{k-1} fixed points, one needs to solve the equation

$$x_n = x_{n+k}. \tag{3.12}$$

3.4.3 The Feigenbaum Constant and Universality

Let us denote the critical value of r at which the logistic map bifurcates into a period-2^k orbit as a_k, so that for $a_k < r < a_{k+1}$ the map has a stable period k orbit. Look at the bifurcation diagram in Fig.(3.7) and Fig.(3.9). Notice how the distance, $a_k - a_{k-1}$, between period doublings decreases as the control parameter is increased.

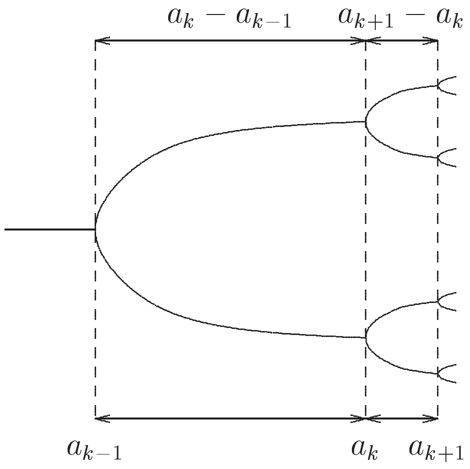


Figure 3.9 Detail of a bifurcation diagram to show the source of the Feigenbaum constant

---

The first few values of a_k are

$$\begin{aligned}
a_1 &= 0.75\,, \\
a_2 &= 0.862372... \\
a_3 &= 0.886022... \\
a_4 &= 0.891101... \\
a_5 &= 0.892189... \\
a_6 &= 0.892423... \\
a_7 &= 0.892472... \\
a_8 &= 0.892483....
\end{aligned}$$

Mitchell Feigenbaum noticed that succesive differences appear to converge geometrically (see Fig.(3.9)) and that the ratio of successive separations tends to a constant as k goes to infinity,

$$\delta \equiv \lim_{k->\infty} \frac{a_k - a_{k-1}}{a_{k+1} - a_k} = 4.669201609.... \tag{3.13}$$

Although the discussion so far has been strictly limited to the logistic map, *the constant δ is the same for other smooth one-dimensional maps with a single hump.* This is an example of **universality**, a concept which we will encounter more of when studying phase transitions in the next chapter. In general systems fall into different universality classes, so that systems within each class have the same behaviour. For the present discussion, one says that all 'unimodal' (smooth, concave downwards, with a single hump) maps belong to the same universality class, that is bifurcate at a rate leading to the universal Feigenbaum constant δ. The actual proof of this statement is quite involved, but briefly stated, it uses the concept of the renormalisation group that was developed to deal with critical phenomena in statistical mechanics.

Although Eq.(3.18) is defined in the limit $k \to \infty$, it can be used to estimate the point at which a system becomes chaotic, a_∞, if the first few bifurcations of the system are known.

3.5 Experimental Tests

Careful experiments have revealed period-doubling bifurcations in a number of real dynamical systems of different types: hydrodynamic, electronic, laser, chemical and acoustic. Details may be found in the book by Cvitanovic. Here we will examine the first of such measurements, the Rayleigh-Benard convection experiment of Libchaber. In that experiment, liquid mercury that was confined in a small box was heated from below to set up a temperature gradient measured in terms of a dimensionless quantity called the Rayleigh number, R. At low values of R the heat flow is by conduction, but as R exceeds some critical value R_c, there is bulk motion of the fluid (convection), leading to the formation of a pattern of cylindrical rolls as the hot fluid rises on one side and cooler fluid descends on another. The temperature is measured at a fixed point of a roll. For R only slightly above R_c, the rolls are straight and the temperature constant. But as the heat flow is increased, an instability occurs leading to wave propagation along the roll and oscillations in the measured temperature. Further increases in the heat flow lead to period doublings in the measured temperature oscillations. From the first few period doublings the Feigenbaum number was estimated to be 4.4 ± 0.1, remarkably close to the theoretical value 4.67. The other experiments quoted above give similarly close values.

There is a puzzle here. The Feigenbaum constant was derived for unimodal one-dimensional maps. Those maps contain none of the physical details of the real experiments which are dynamical systems with many degrees of freedom evolving in continuos time and three dimensions. So why is there any agreement between the predictions of the one-dimensional maps and the real experiments ? In order to understand how that is possible, it is useful to consider first the simple dynamical system of the next section.

3.6 The Rossler System

A model dynamical system which is simpler than the Lorenz model was proposed by Rossler. The system consists of three coupled first-order differential equations

$$\begin{aligned} \dot{x} &= -(z+y) && (3.14) \\ \dot{y} &= x+ay && (3.15) \\ \dot{z} &= b+xz-cz && (3.16) \end{aligned}$$

The system has three degrees of freedom represented by the dynamical variables x, y and z. Notice the nonlinear term xz in the last equation. There are also three parameters a, b and c. By fixing two of the parameters and varying the third one can study the approach to chaos. Let us fix $a = b = 0.2$ and treat c as the control parameter. Instead of plotting the time series for the system, it is useful to consider phase portraits, that is by plotting one degree of freedom against another. Figure(3.11) shows the $x - y$ phase portraits (actually only the post-transient orbits are shown) for various values of c. One sees clearly period one, period two and period four trajectories: The period here refers to the number of times the cycle goes around before closing. These period doubling sequences lead eventually a chaotic state shown in Figure(3.11d).

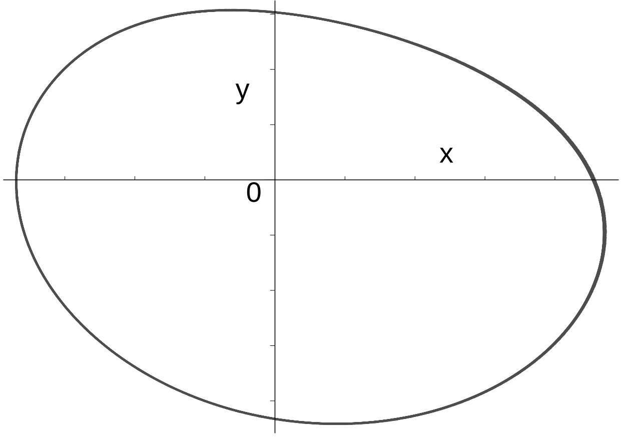


Fig.(3.11a): Rossler attractor for c=2.5

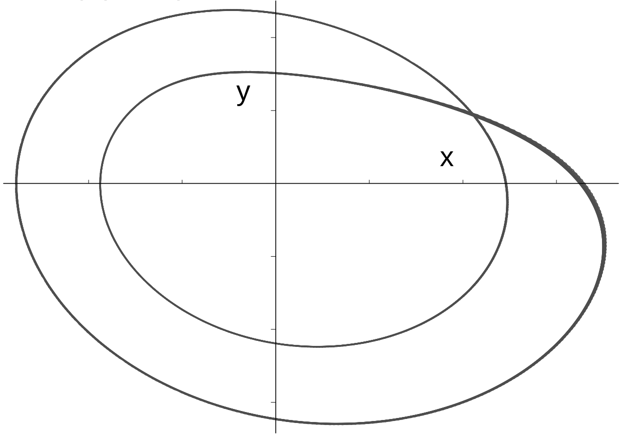


Fig.(3.11b): Rossler attractor for c=3.5

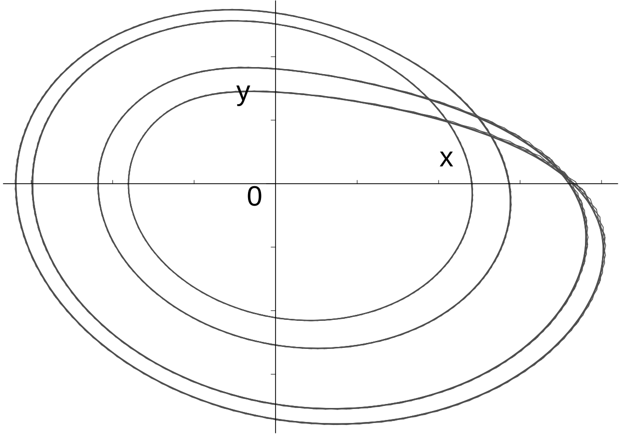


Fig.(3.11c): Rossler attractor for c=4

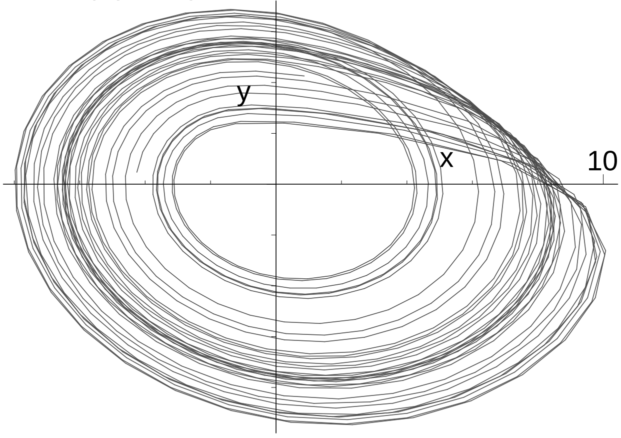


Fig.(3.11d): Rossler attractor for c=5

The behaviour of the z variable in the chaotic state ($c = 5$) is illuminated by looking at the time series shown in Figure(3.12a). From that one sees that the z coordinate is small most of the time, meaning that the trajectories mostly are close to the $x - y$ plane. However occasionally there are large spikes in the z time series leading to a jump in the z coordinate. For reference the time series for the x variable is shown in Fig(3.12b).

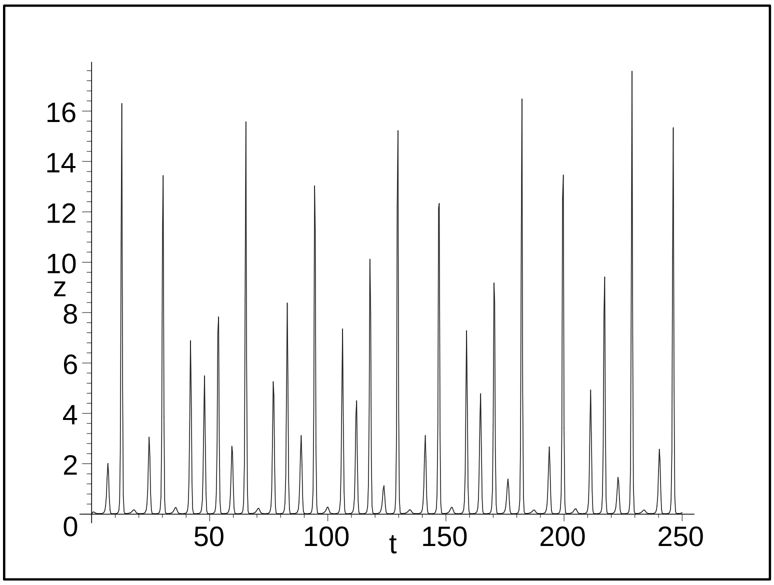


Fig.(3.12a)

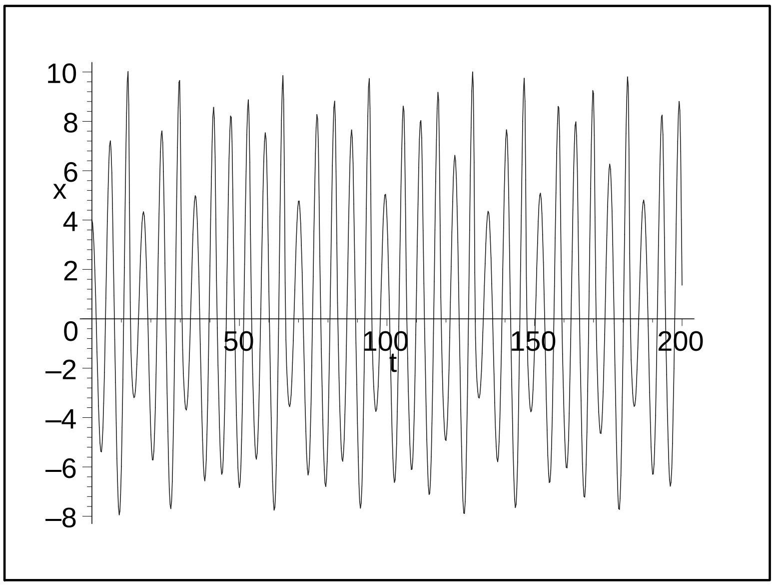


Fig.(3.12b)

The full three dimensional phase portrait for the chaotic state shown in Figure (3.13). This is the strange attractor for the system. Notice that the attractor lies in a bounded region of phase space, and its structure, though complicated, is far from random. Indeed strange attractors have been shown to have a fractal structure, that is self-similarity at different scales! This is our first indication that deterministic chaos is different from noise or randomness.

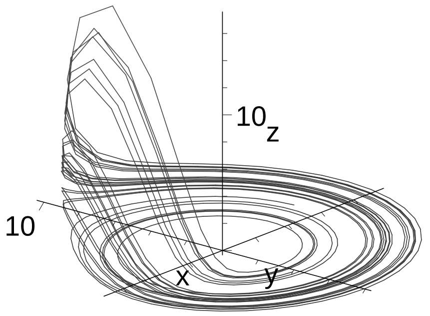


Fig.(3.13): Rossler attractor for c=5

There is an apparent paradox about chaotic states. We noted that they were characterized by a sensitive dependence on initial conditions, which leads nearby trajectories to diverge exponentially . On the other hand the phase portrait is bounded: This is why the strange attractor justifies its attractor label. So how are trajectories repelled and attracted at the same time ? This is done by a process of stretching of nearby trajectories on short time scales, followed by a process of folding at longer scales. The folding process is evident in Figs(3.11d) and (3.13): Follow some trajectory in the $z = 0$ plane, it will soon be lifted up at a spike of the z-component, and is then folded back into the $z = 0$ plane of the attractor at another point, to continue its evolution. The stretching and folding processes mix nearby trajectories in the attractor and results in a decorrelation of closeby initial states.

3.7 The Lorenz map

Now let us compare the results for the Rossler system with those of the one-dimensional map. To do that we will plot the **Lorenz map** for the Rossler system. The Lorenz map (not to be confused with the 'Lorenz system'!) is a way of obtaining a one-dimensional map from a time-series of a dynamical system. For the Rossler system in the chaotic state, say $c = 5$, the time series for the x variable is shown in Fig(3.12b). Label the successive local maxima of that graph as x_n, that is the n-th maxima is at x_n. The Lorenz map is then a plot of x_{n+1} versus x_n for various n. The Lorenz map for the Rossler system is shown below, Fig.(3.14).

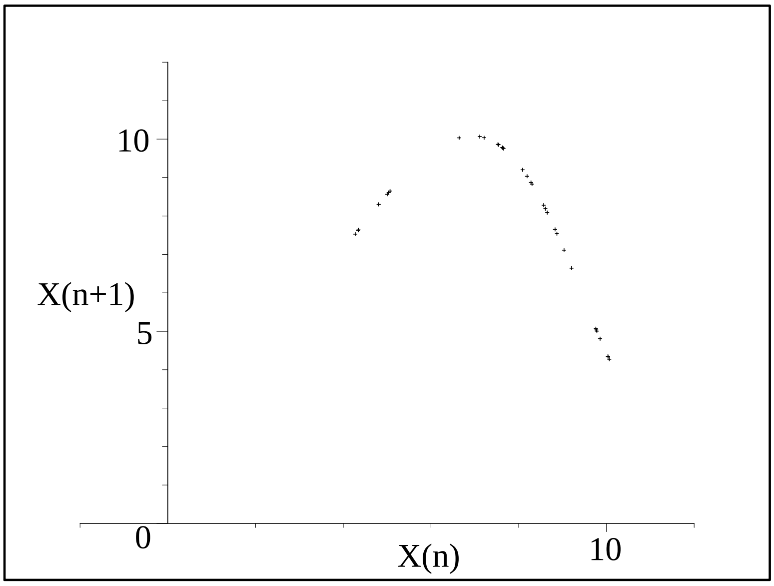


Fig.(3.14)

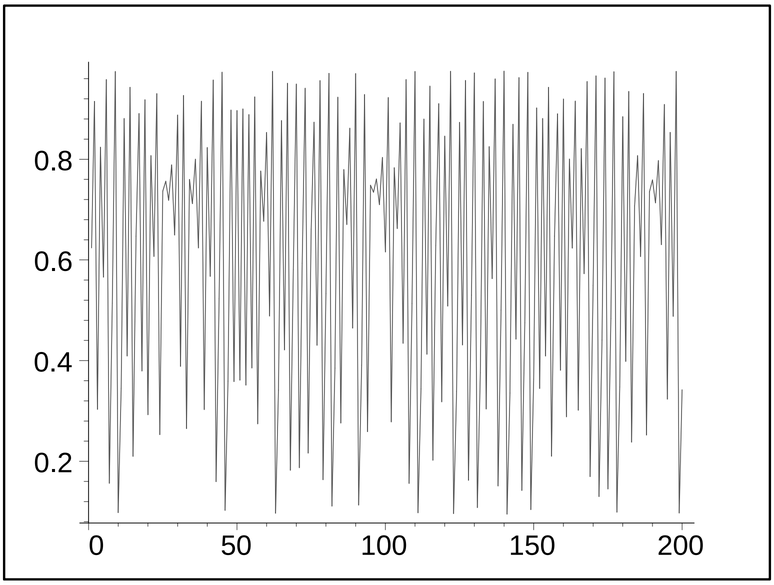
The points lie on an almost one-dimensional unimodal curve. Thus there should be an approximate relationship of the form $x_{n+1} = f(x_n)$ with f a unimodal map. From Feigenbaum's work it then follows that the Rossler system will also show a period-doubling route to chaos and the same Feigenbaum constant, as is indeed the case !

However not all systems have a one-dimensional Lorenz map, since that requires the system's strange attractor to be almost flat. In physical terms this requirse that the system have only one or two degrees of freedom that are dominant. In fact the experiments quoted above had that feature. In other words, the universality theory is developed only for low-dimensional chaotic systems.

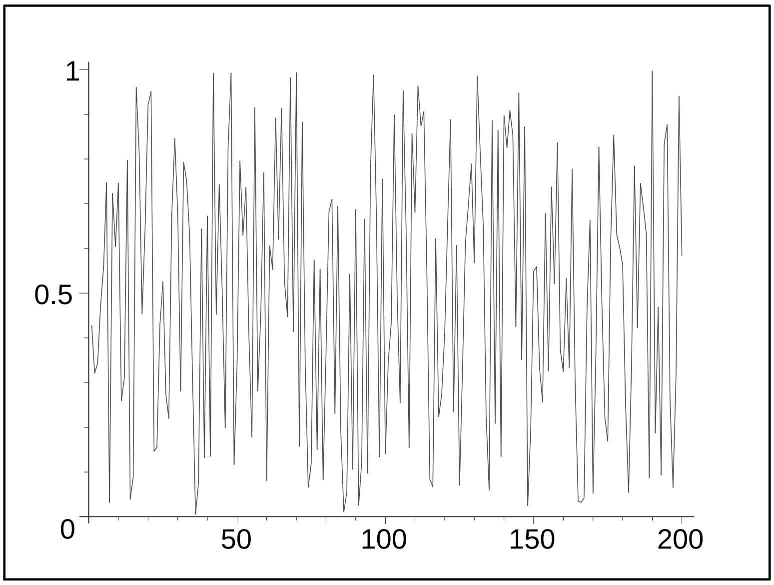
3.8 Noise versus Chaos

Look again at Fig.(3.12b). If you were not told that the time-series was from a deterministic system, would you have been able to decide if the figure was due to noise (random events) or to deterministic chaos ? This is an interesting issue as one would like to know whether many apparently noisy time-series in our daily lives (such as sunspots or brainwaves) are actually the result of deterministic chaos and hence whether there are relatively simple laws underlying those phenomena.

If there is an underlying strange attractor to the time series then one can attempt to show this using the idea of **attractor reconstruction**. Suppose one is given a time series $x(n)$ obtained by sampling points at regular time intervals. For a deterministic dynamics, The simplest assumption then is that the value of $x(n+1)$ depends on its value at the previous time, $x(n)$ Thus one plots $x(n+1)$ against $x(n)$ to see if a pattern emerges.



Fig(3.15a)



Fig(3.15b)

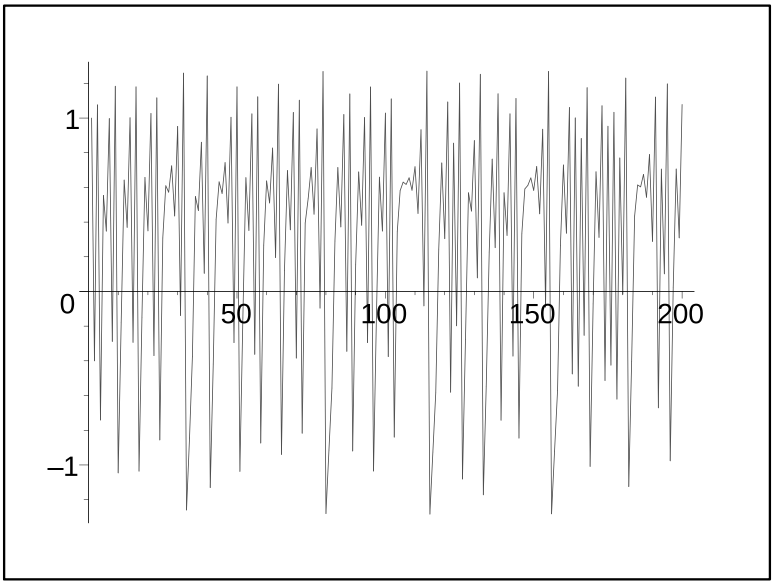


Fig.(3.15c)

Fig(3.16a) Fig.(3.16b) Fig.(3.16c)

For the attractor reconstruction to be succesful there are two issues that one has to address. Firstly, the *embedding dimension* must be large enough for the attractor to become visible. In the above examples we plotted the vector $(x(n), x(n+1))$, meaning that we investigated a two-dimensional embedding space. Since the dimension of the Lorenz or Rossler strange attractors is three, a similar procedure for the latter cases would require an embedding space of at least three dimensions, so we would need to plot for example $(x(n), x(n+1), x(n+2))$.

The second issue is the optimal *time delay* to be used for the above reconstruction. In the mentioned examples, the time delay was one. More generally one can consider other values, so that for the Rossler system one could plot $(x(n), x(n+N), x(n+N+1))$, corresponding to a time delay of N. The value of N might be determined by trial.

There have been several attempts to discover determinstic chaos in stock-market data and to so profit from that knowledge. You can read about the incredible and entertaining adventures of some physicists in this endeavour in Ref.[6].

3.9 Summary

If a dynamical system becomes chaotic at some parameters of physical interest then that is usually an unwelcome occurrence. Firstly, it means that the system in the chaotic state is extremely sensitive to initial conditions, and hence the errors inherent in real-life measurements of finite precision. The extreme sensitivity implies a lack of practical predictability of the long-term behaviour of the system even though the equations of motion are completely deterministic !

Nevertheless long-term unpredictability due to chaos does not preclude short-term predictability and usefulness. Furthermore, one might be able to exploit the sensitive dependence of chaotic systems to initial conditions by controlling those systems: very small perturbations to the system can be used to completely change its long-term behaviour. Some attempts in this direction are currently underway, for example, for control of cardiac chaos (See Ref.[10]). However chaos might be useful in some cases: It has been reported that epileptic seizures are the result of the electrical brain waves syncronising, while their natural state is "chaotic"!

It is very important to note that chaos arising in deterministic systems is quite different from noise in random (stochastic) systems. This is because one still sees order and patterns in the phase space plots for chaotic systems which is absent for random systems. In fact one can say that chaos (in deterministic systems) is order camouflaged as disorder ! An important technique in distinguishing whether a given time-series has a deterministic underlying basis is through the method of attractor reconstruction.

For low-dimenisonal systems that are at least approximately equivalent to a unimodal one-dimensional map, the approach to chaos is through period-doubling bifurcations. The Feigenbaum constant summarizes the universality of this class of systems. Chaos in higher dimensional systems with many degress of freedom can occur in different ways from that in low-dimensional systems.

In the previous chapter we saw how the geometric complexity of nature could be described by simple algorithms and models that generate fractal structures. That is, complex patterns and natural order can arise from a simple underlying basis. In this chapter we see the flip side : very simple deterministic dynamical models can lead to behaviour that *looks* random. However the apparent disorder is actually chaotic rather than random as the order is hidden now in the phase-space of the system. What this suggests is that some of the apparent irregularity observed in nature or man-made systems might be ammeanable to deterministic modelling instead of appealing to random processes.

3.10 Exercises

1. Practise plotting the state-space plots for the map $x_{n+1} = x_n^2$ for *different* initial values of x_0. Use your own software or that in Ref.[1] or a modified form of the MAPLE program mf35.mws from Ref.[3].

2. Consider the one-dimensional map $x_{n+1} = \sqrt{x_n}$ for $x_0 \geq 0$.
   (a) Explore the above with a calculator, or guess its general behaviour. Where are the fixed points, are they stable, and what is their basin of attraction ?
   (b) Confirm your results in (a) analytically.
   (c) Confirm your results above with a graphical analysis using the same software as you used in the previous problem.

3. (a) Give one example of positive feedback and one of negative feedback from your personal experience.
   (b) From your experience, give one example of a linear change and one example of nonlinear dynamics or change.

4. With regards to the logistic map,
   (a) Complete the steps leading to its derivation in the text.

(b) Show that if the control parameter is restricted to $0 \leq r \leq 1$, then any initial $0 \leq x_0 \leq 1$ leads to a value of x_n that is bounded between the same values.
(c) Show alegbraically that for a reproduction rate $0 \leq r \leq 1/4$, the population of species will eventually (at infinite time) become extinct no matter what the initial population is.
(d) Illustrate the result in (c) above graphically using a state-space plot.
(e) Confirm the results in (4c,4d) analytically by showing that for $0 \leq r < 1/4$, the only period-one fixed point of the system is the 'trivial' one at the origin and that it is stable (atractive). What happens when r is exactly 1/4?

5. With regards to the logistic map,
(a) Show analytically that for $r > 1/4$ the trivial fixed point at the origin becomes unstable.
(b) Show analytically that for $r > 1/4$ there is a period-one fixed point at $x_f = 1 - 1/4r$ and that this point is stable for $1/4 < r < 3/4$.
(c) Verify the result in 5(b) graphically, using your favourite software, for different values of r and x_0.

6. (a) With regards to the logistic map, show that as r is increased beyond 3/4 the fixed point at $x_f = 1 - 1/(4r)$ becomes unstable.
(b). Explain why the values of the two new period two fixed points that result for $r > 3/4$ are obtained by solving the equation

$$x_{n+2} = x_n. \tag{3.17}$$

(c) Write out the above equation explicitly in terms of the single variable x_n. Without actually solving it, explain why two of the roots of the equation must be $x_f = 0$ and $x_f = 1 - 1/4r$, the two fixed points that already exist for lower values of r.
(d) Verify explicitly that $x_f = 0$ and $x_f = 1 - 1/4r$ are roots of the above equation.
(e)Find the two new roots of the above equation. Show that these roots exist only for $r > 3/4$.
(f)What is the condition for stability of period-two fixed points ? Show that the two new roots in (e) above represent stable (attracting) fixed points.
(g) Convince yourself of the period-2 orbits using your favourite software.

7. (a) Use the critical values of the control parameter for the logistic map, a_k, given in the text to evaluate
$$d_k \equiv \lim_{k->\infty} \frac{a_k - a_{k-1}}{a_{k+1} - a_k} \tag{3.18}$$
for $2 \leq k \leq 7$.

   (b) Compare the result of (7a) with the Feigenbaum constant δ.
   (c) Use the values of a_6, a_7 and the Feigenbaum constant δ to estimate the value of a_8. Compare the result with the known value.

8. Use your favourite software (see for example the Maple file mf36.mws of Ref.[3]) to explore the bifurcation diagram for the logistic map. Look closely at the periodicity windows that open up in the chaotic region and convince yourself that period three bifurcations exist there.

9. Explore the sine-map, $x_{n+1} = r\sin(\pi x_n)$ for $0 \leq x_0 \leq 1,\ 0 \leq r \leq 1$, by plotting the state space plots and the bifurcation diagram. Compare with the logistic map.

10. Explore the map $x_{n+1} = \text{frac}(2x_n)$. Here 'frac' means 'keep only the noninteger part'.
    (a) Make a state-space plot.
    (b) Determine the fixed points.
    (c) Show that there are infinitely many periodic and aperiodic orbits.
    (d) Show that the map displays sensitive dependence on initial conditions.
    Hint: It is useful in this problem to consider the binary representation of numbers.

11. Use your favourite software (e.g. the maple file ross.mws) to explore the Rossler system (time series and phase plots) for different values of the control parameter c. Note that in the Maple file you can easily view different perspectives of three-dimensional plots (especially that of the strange attractor) by clicking the mouse on the plot and dragging it. Note also that by plotting only later points of the iteration, the transient part fo the orbit is neglected. Practise plotting also the transient parts fo the orbits to see the approach to the attractors.

12. (a) Repeat the above excercise for the Lorenz system using your favourite software (for example the Maple file mf10.mws of Ref.[3]).

(b) Obtain the Lorenz map for the Lorenz system using its $z(t)$ time series.
(c) Reconstruct the Lorenz attractor using one of its time-series, (using for example the file mf42.mws of Ref.[3]).

13. Give an example of a situation where Chaos might be useful. Inform yourself about some applications of chaos as mentioned in, for example, Ref.[10].

14. Check out the resources in Refs.[7, 8, 9] for self-study.

3.11 References

1. The Computational Beauty of Nature, by Gary William Flake. See also the book's homepage at http://mitpress.mit.edu/books/FLAOH/cbnhtml/home.html where you can obtain the source code for many of the plotting programs and simulations used in Flake's book.

2. Fractals and Chaos, by P.A. Addison.

3. Nonlinear Physics with Maple, by R.H. Enns and G.C. McGuire, 2nd Edition. This book comes with a CD containing useful Maple files.

4. Nonliear Dynamics and Chaos, by S.H. Strogatz.

5. The figures used in this chapter were taken from the website of Ref.[1] (those with copyright notice attached below the figure caption) or generated using the Maple software in the CD that comes with Ref.[3]. Some other figures used in the lectures were taken from Ref[4].

6. The Predictors, by T. Bass.

7. The freeware FRACTINT, easily available on the internet, contains software for plotting strange attractors and bifurcation diagrams of many chaotic systems, and also of fractals.

8. A pedagogical website on non-linear dynamics is at http://www.apmaths.uwo.ca/ bfraser/version1/nonlinearlab.html

9. An introductory web-book on Chaos is at http://hypertextbook.com/chaos/

10. Some engineering applications of Chaos are mentioned in http://www.ornl.gov/etd/etdchaos.htm

Chapter 4

Equilibrium Systems

As mentioned in the introductory chapter, the complex systems one is interested in "complexity studies" are usually those that consist of a large collection of interacting sub-units and that are *out-of-equilibrium.* Such systems, or rather simplified models of the realistic systems, have become accessible to study in the last two decades due to the explosive growth in computer power at decreasing costs. We will study such out-of-equilibrium complex systems in subsequent chapters, but here we will focus our attention on simpler equilibrium systems.

However even a large system in equilibrium is not easy to describe and understand analytically. Nonetheless, beginning already In the nineteenth century brilliant minds applied themselves to the problem and developed the fields of thermodynamics, kinetic theory and statistical mechanics. Some of the concepts that have been developed for the description of phenomena in equilibrium systems turn out to be relevant also for the study of more realistic out-of-equilibrium complex systems that occur in nature. Indeed, the study of equilibrium systems gives a concrete and quantitative illustration of the concepts of *entropy*, *emergent laws* and *universality.*

4.1 Thermodynamics and Kinetic Theory

Thermodynamics is the study of the macroscopic physical properties of a large collection of particles. That is, instead of attempting to describe the individual motion of the microscopic particles, one tries to make statements about the properties of the system as a whole. One says that a system has reached **thermodynamic equilibrium** when the macroscopic properties do not change with time. The word *system* used here refers to the part of the world that one is interested in studying. Hence the world is conveniently divided into the system, and an exterior or *environment*.

Note: Historically, the fields of kinetic theory and statistical mechanics were developed even before there was direct proof about the reality of atoms. Some indirect evidence that fluids consist of fast moving molecules was provided by observations in 1827 by the botanist Robert Brown who noticed the jiggling motion of pollen grains suspended in water. After many careful experiments he came to conclude that the motion was real and existent even for inorganic suspensions. It came to accepted by scientists that this motion of the suspended particle was due to the constant and random bombardment it suffered from all sides by molecules of the liquid. A quantitative explanation of the motion was later provided by Einstein.

4.1.1 State Variables

We are all familiar with the concept of temperature. The **Zeroth Law of Thermodynamics** formalizes our intuition and experience as follows:
If a system A is in equilibrium with system B (that is, has no exchange of heat with it), and if system B is in equilibrium with system C, then A is in equilibrium with C.
This law allows us to associate a quantity called temperature to each system in thermal equilibrium, so that two systems in equilibrium have the same temperature. The thermometer is a device that uses the Zeroth Law in a quantitative and practical way.

Note: We will use the absolute temperature scale where units are measured in Kelvins. The conversion from Celsius, t to Kelvin T is : $T = t + 273.15$.

In addition to the temperature, one may need more thermodynamic parameters, called **state variables**, to completely characterize the state of the system. For example, for a gas these are the pressure P and volume V. Variables such as temperature and pressure that are independent of the size of the system are called *intensive*, while those such as the volume are called *extensive*. The parameters that can be used to describe a system are not all independent but related by an **equation of state**. For an ideal gas one has the equation

$$PV = NkT \tag{4.1}$$

where $k = 1.38 \times 10^{-23}$ Joules/K is called Boltzmanns constant, and N is the number of molecules. As you must have learnt in school, an ideal gas is the universal limiting description of real gases when their density is very low and the temperature high.

In general, the equation of state of a real substance is more complicated. It is usual to plot the equation of state as a function of its parameters. One useful curve follows by keeping V constant and representing the equation of state on a $P-T$ plot as shown in the figure for a generic substance. The lines mark boundaries between the different phases of the substance, where changes occur in the physical properties of the substance.

4.1.2 The Ideal Gas

Since thermodynamics is a phenomenological description of macroscopic matter, it does not take into account the underlying atomic structure. The thermodynamic relations, such as the ideal gas law, are therefore often approximations that are good when one is talking about the average properties of a large system, that is, a large number of atoms. Indeed the various thermodynamic relations can be considered as examples of **emergent laws**: generalities about the system that are apparent only at the macroscopic scale but are not obvious or existent at the microscopic level. Indeed even the concept of temperature may be considered a macroscopic emergent feature that is ill-defined for a system with only a few atoms.

The above discussion can be illustrated for the case of the equation of state of an ideal gas that allows a simple derivation starting from an atomistic description. Consider atoms of a gas confined to a box which has a frictionless and perfectly reflecting piston at one end. Let the area of the piston be A the volume of the box V (see figure), and the number of atoms be N. As the atoms rebound from the piston, they will impart momentum to it and cause it to move. To prevent the piston from moving one must therefore apply a force to it to balance that imparted by the atoms. Let this external balancing force be F. We now calculate the force F in terms of the parameters of the gas. Now Newtons law states that

$$F = \text{rate of change of momentum}, \tag{4.2}$$

Hence in a short time interval dt,

$$\text{Force} \times dt = \text{change of momentum}. \tag{4.3}$$

Consider an atom of mass m and velocity v, with x-component v_x hitting the piston. It rebounds with an x-component $-v_x$, and so the change in momentum of the atom is $-2mv_x$. Now for atoms with a velocity v_x to hit the piston in the time interval dt, they must be a distance $v_x dt$ away. Assuming that the atoms are uniformly distributed, then there are $n = N/V$

atoms per unit volume and so $nAv_x\ dt$ atoms that will hit the piston in that time interval.

Hence the net change in momentum is $-2nmAv_x^2\ dt$, and so the pressure F/A imparted by the piston on the gas is $-2nmv_x^2$ and so by Newtons Third Law the pressure of the gas is

$$P = 2nmv_x^2\,. \tag{4.4}$$

Since the atoms in the gas have different velocities, the above result must be averaged. Let $< v_x^2 >$ be the average of the squared x-component of the velocities. This includes averages over both positive and negative velocities and so taking half of this (to represent atoms moving towards the piston) and substituting into the above equation we get

$$P = nm < v_x^2 > \ . \tag{4.5}$$

Now

$$v^2 = v_x^2 + v_y^2 + v_z^2\,, \tag{4.6}$$

and so

$$< v^2 > = < v_x^2 > + < v_y^2 > + < v_z^2 > \ . \tag{4.7}$$

But on the average the motion in the three directions is equivalent and so

$$< v_x^2 > = < v_y^2 > = < v_z^2 >\,, \tag{4.8}$$

and hence

$$PV = \frac{2N}{3} < mv^2/2 > \ . \tag{4.9}$$

For a monatomic gas $< mv^2/2 >$ is the average kinetic energy of an atom and so the quantity $N < mv^2/2 >$ represents the total internal energy U of the gas. So

$$PV = \frac{2}{3}U \tag{4.10}$$

for a monatomic gas.

The above is an elementary kinetic theory derivation of an ideal gas law. Note however that temperature is absent from the equation. On comparing the theoretical result (4.9) with the empirical equation (4.1), a relation is obtained between the macroscopic concept of temperature and the mean microscopic kinetic energy

$$< mv^2/2 >= \frac{3}{2}kT. \tag{4.11}$$

Thus at least in this simple case one has succeeded in deriving an empirical law (the ideal gas law) for a large system starting from the underlying microscopic dynamics. The derivation is complete only after associating some of the emergent macroscopic parameters with averages of microscopic quantities. In more complicated situations, especially in the case of out- of-equilibrium complex systems as defined in the introduction, one will not be able to make such simple or direct associations between emergent laws and patterns and the underlying microscopic theory.

Note: A more realistic equation for gases, valid at higher densities is the Van der Waals equation of state $(V - b)(P + a/V^2) = RT$ where a and b are constants characteristic of the substance studied.

4.2 Statistical Mechanics

The kinetic theory method can in principle be generalized to more complicated systems and even to out of equilibrium situations since once the forces among the constituents are known one simply writes down the dynamical equations and explores their consequences. The problem is that the equations are often too complicated and too many (for a large system) to be solved for practical cases even using the most powerful computers. Furthermore one rarely knows the initial conditions to plug into the dynamical equations.

An alternative procedure called statistical mechanics allows various properties of equilibrium systems to be deduced using probabilistic methods. This is now described. Consider an isolated classical system consisting a large number N of molecules in a large molecular volume V. Typically N and V are numerically of order 10^{23}. If the position q_i and momentum p_i of each molecule is known at a particular instant, then the subsequent dynamics of the system is uniquely determined by the known laws of molecular dynamics. Thus the state of the system is completely specified by $2 \times 3 \times N = 6N$ variables summarized as (p, q). It is useful to represent the state of the entire system as a point in $6N$ dimensional phase space. Of course this point traces out a path as the system evolves in time. However since the system is isolated, its energy is constant (conserved) and so the path is restricted to lie on a fixed energy surface.

Now, *we are interested not in the detailed time evolution of the state but rather its macroscopic properties in thermodynamic equilibrium.* Let $f(q(t), p(t))$ be some physical quantity written as a function of the state of the system. Then in experiment one actually measures the time average of that quantity over a period T that is very long compared to the mean molecular collision time.

$$< f >= \frac{1}{T} \int_0^T f(q,p)\ dt\,. \tag{4.12}$$

From a theoretical viewpoint the above formula is inconvenient and can be replaced by an equivalent expression after appealing to the

Ergodic Theorem: After a sufficiently long time, any representative point of the system will cover the entire accessible phase space.

Therefore instead of the time average as in (4.12), one can rewrite the physical observable $< f >$ as an average over phase space:

$$< f >= \int f(q,p)\ \rho(q,p)\ dq\ dp \tag{4.13}$$

where $\rho(q,p)$ is called the density function and represents the probability to find the system in a state with coordinates between q and $q + dq$ and momenta between p and $p + dp$. Thus regions of phase space of high density are frequented more often than regions of low density. (By assuming equilibrium, the density function does not depend explicitly on time).

Let the accessible phase space be divided into cells of size $\delta p \times \delta q = \Delta$, so that cells smaller than size Δ are not distinguished. Define $\Gamma(E)$ as the total number of phase space cells, or different states. Then the *entropy* of the system is defined by

$$S(E,V) = k \log \Gamma(E)\,, \tag{4.14}$$

where k is Boltzmanns constant. That is, *entropy is a measure of the total number of different microscopic states the macroscopic system can exist in.* Thus it is a measure of our lack of precise knowledge of the system, or equivalently the amount of disorder in the system: The larger the number of microstates for a particular macrostate, the greater our ignorance of the underlying microstate, and also the larger the amount of disorder in the system (a system with a limited number of possible microstates is more restricted or ordered).

4.3 The Second Law

Second Law of Thermodynamics: The entropy of a thermally isolated system never decreases.

Thermally isolated systems are also called closed systems: There is no exchange of energy or matter with the environment. The above statement of the second law can be shown to be equivalent to other statements found in books (e.g. the Kelvin and Clausius statements) that summarize some empirical facts about nature. Processes in closed systems for which the entropy remains constant are called *reversible* while those for which the entropy increases are called *irreversible*.

To see how the above terms correspond to our intuitive use of them, consider the following example: An isolated box contains an equal number of two types of molecules, say 'white' and 'black'. Suppose one had the extreme case where all the white molecules were on the left and all the black molecules on the right. Of course this is an unnatural situation and soon, due to collisions, the molecules will totally mix. The first situation corresponds to a state of maximum order and minimum uncertainty about the microstate while the second situation corresponds to large disorder (randomness) and maximum ignorance on our part about the microstate of the system.

One can compute the entropy of the ordered and disorderd states: Clearly there are many more ways to form the disordered state than the ordered state and thus from the definition of entropy, the entropy of the disordered state is much higher than the entropy of the ordered sate. Since we know that the gases will mix, the system goes from the ordered state to the disordered state, increasing its entropy in accordance with the second law. We also know that if the number of molecules is very large, it is extremely unlikely that the gases will revert to the totally separated state at some future time. Thus the increase in entropy and disorder in the system is for all practical purposes irreversible, in agreement with the definition above.

Note that the irreversibility is not due to the underlying fundamental laws (e.g. Newtons laws are reversible) but a result of the system going from an unlikely ordered state to a more probable disordered state, and the fact that for large systems (large number of molecules), the probability of the system reverting to the ordered state being negligible. In the above example, the probability of any one molecule being on the left half of the box is 1/2. If there are N molecules, the probability that all of them are on the left is $(1/2)^N$. Even for N as small as 100 this works out to be about 10^{-30}, an infinitesimal quantity ! For macroscopic materials, N is of the order of 10^{23}, and so the resulting probability is even lower. **Thus the Second Law is actually a statement about average behaviour that becomes overwhelmingly likely in a very large system, meaning that exceptions will be unobservable in all practical situations.**

What is the largest isolated system? The universe of course! Thus the second law of thermodynamics states that the entropy of the universe never decreases. In fact looking around we see that most of the changes in the universe lead to an increase in randomness and are irreversible. Thus the entropy of the universe seems to be increasing! Thermodynamically, one says that time flows in the direction of increasing entropy, or that the **arrow of time** is in the direction of increasing entropy of the universe.

4.4 The First Law

We have discussed the Zeroth and Second Laws. What is the First law?

The first law of thermodynamics is actually a statement of the conservation of energy. The experiments of Benjamin Thompson, James Joule and others demonstrated that heat was a form of energy and that it could be produced by doing mechanical work, and also conversely could be transformed into mechanical energy. Let dQ be the net amount of heat absorbed *by* a system in an infinitesimal transformation and dW the amount of work done *by* the system (note the sign conventions!). Then the first law states that the quantity dU defined by

$$dU = dQ - dW \tag{4.15}$$

is the same for all transformations that lead the system from its initial to final state, that is $\int dU$ is independent of the path of integration in state space and only depends on the end points. This property is not true individually of dQ and dW since the amount of heat absorbed or the work done depends also on the actual process, that is, path in state space. The quantity U is called the internal energy of the system (Recall its appearance in the ideal gas law (4.10) derived using kinetic theory).

4.5 Entropy for Open Systems

We would like to derive a thermodynamic relation for entropy (as opposed to the statistical definition given above) in general systems which are *not necessarily closed*, that is which allow an exchange of heat with the surroundings. Firstly note that *for systems which are not closed, the defintion of reversibility is more general : It means that the system retraces its path in time when the external conditions are reversed in time.* So consider now an infinitesimal reversible transformations. We can then write (recall the kinetic theory discussion again, work done = force × distance = $P \times A \times$

distance.)

$$dW = PdV\,. \tag{4.16}$$

Thus for a general infinitesimal reversible transformation, we have from the first law

$$dQ = dU + PdV \tag{4.17}$$

On the other hand, the entropy S can be treated as a function of U and V (see the statistical definition given earlier), which under an infinitesimal reversible transformation changes by

$$dS = \left(\frac{\partial S}{\partial U}\right)_V dU + \left(\frac{\partial S}{\partial V}\right)_U dV \tag{4.18}$$

Define (this will be checked for consistency later!),

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T} \tag{4.19}$$

and

$$\left(\frac{\partial S}{\partial V}\right)_U = \frac{P}{T} \tag{4.20}$$

Then Eq.(4.18) becomes

$$TdS = dU + PdV \tag{4.21}$$

Which then implies on comparing with the first law (4.17) that *for an infinitesimal reversible transformation (in a general, not necessarily closed system), the change in entropy is given by*

$$dS = \frac{dQ}{T} \tag{4.22}$$

One can use (4.22) to compute the the difference in entropy between two states A and B. Choose a reversible path joining the two states and integrate the above equation to get

$$S(A) - S(B) = \int_B^A \frac{dQ}{T} \tag{4.23}$$

Note: S is a state function, so its value is fixed once the independent parameters, say U and V are fixed. On the other hand, to calculate it using the formula above, one must choose a reversible path since for a general irreversible path the quantity calculated depends on the path taken and is not the entropy difference between the two states. (It is an assumption of classical thermodynamics that one can always find a reversible transformation between any two states connected by an irreversible one).

To illustrate this point, consider two different isothermal expansions of an ideal gas from volume V_1 to volume V_2.
1. Reversible isothermal expansion.

$$\Delta S_{gas} = \int \frac{dQ}{T} = \frac{1}{T}\int dQ\,. \tag{4.24}$$

Now for an ideal gas $U = U(T)$ only (see section(1)), hence in the expansion $\Delta U = 0$ and so from the first law the heat absorbed by the gas equals the work done. So

$$\Delta S_{gas} = \frac{1}{T}\int p\ dV = N\ k\int \frac{dV}{V} = N\ k\log\left(\frac{V2}{V1}\right)\,. \tag{4.25}$$

The entropy change of the heat reservoir is

$$\Delta S_{reservoir} = -\Delta\left(\frac{Q}{T}\right) = -\Delta\ S_{gas} \tag{4.26}$$

Is it surprising that the entropy of the reservoir has decreased ? No, there is nothing forbidding it since it is not an isolated system. Indeed if one looks at the change of entropy of the combined larger system of gas plus reservoir, it is zero, in agreement with the second law (for isolated systems). Note that reversibility in the above system can be achieved by storing the work done by the gas on expansion in a spring attached to the piston, so that this can be used to compress the gas thus reversing the expansion.

2. Free Expansion. Since the intial and final states (volume and temperature) of the gas are the same as before for the reversible case, the change in entropy of the gas is as before (because S is a state function, and anyway its calculation requires choosing a reversible path joining the initial and final states).

Since as before $\Delta U = 0$ for the gas and as no work is done in a free expansion, $\Delta W = 0$ also, implying that no heat is exchanged between the gas and its reservoir. Hence the entropy change of the reservoir is zero. Thus there is a net increase of entropy of the combined gas plus reservoir system. Furthermore the expansion is clearly irreversible (Why?)

4.6 Phase Transitions

We have seen a generic phase diagram earlier. It showed that under certain conditions (along the boundary lines), one could have the substance existing simultaneously under more than one state (for example, liquid and gas). That is, the system along the lines is not homogeneous. Away from that line of course the state of the system is homogeneous and is called a *phase* of the system. When one crosses a line, by tuning the parameters, the system undergoes a phase transition, often accompanied by an absorption or release of heat into the surroundings (e.g. consider melting of ice or the boiling of water). In phase transitions the entropies of the respective phases before and after are often different: Such transitions are called *first-order* phase transitions, while those for which the entropy is continuous across the boundary are called **second-order**. For first-order phase transitions, one defines the latent heat L as that which is exchanged with the surroundings,

$$L = T\ \Delta S. \tag{4.27}$$

4.6.1 Second Order Transitions

Thermodynamic properties of systems near a second order phase transition are of great interest because one observes great simplicity and universality in their behaviour. A ferromagnetic system is an example of one that displays a second-order phase transition and will be used below for detailed discussion. At low temperature, an appreciable fraction of the atomic spins (the "atomic magnets") in some metals (e.g. iron) become spontaneously polarized in the same direction, thus giving rise to a net measurable magnetic field. However as the temperature is raised, the spins become randomized due to thermal agitation and the magnetization is lost. The critical temperature at which the change occurs is called in this system the Curie point. (For iron this is $1044K$).

Symmetry

We are all familiar with the colloquial use of the word symmetry to describe objects: Something that looks the same from more than one side. Mathematically one defines an object as being symmetrical if it looks the same after one has done “something” to it. For example a circle looks the same even after it has been rotated about its centre by any amount. A square on the other hand looks the same only when rotated by multiples of ninety degrees. Thus different objects can have different amounts of symmetry. Though symmetry is important, the breaking of symmetry is even more relevant in the real world: Most systems we see do not correspond to cases of high symmetry eventhough the basic laws might be. Thus one has to understand how and why this symmetry is broken. “Self-organisation”, which is a form of emergence that we will look at in later chapters, is actually a case of symmetry being spontaneously broken.

The spatial symmetries of the system in the two phases are different. In the high-temperature phase the system is disordered, with no net magnetization but with complete rotational symmetry (isotropy). At low temperature, the system becomes ordered and the net magnetization defines a preferred direction in space. The material thus becomes anisotropic at low temperature, breaking rotational symmetry. *The low-temperature ordered phase is therefore less symmetrical* and to describe it fully one needs to introduce an extra parameter called the **order parameter**, which in this case is just the magnetisation vector $\vec{M}$. For simplicity, we will work below with a single component M, called the scalar magnetization (you can think of this as the magnetization in the z-direction of a planar ferromagnetic system). When $M = 0$ one is in the disordered high-temperature phase while for $M \neq 0$, one is in the low-temperature ordered phase with net magnetization.

(A more familiar example is the ice to liquid water phase transition. The molecules in liquid water are in a disordered state and the system is homogeneous and isotropic. As water is cooled it freezes into ice whose properties are no longer isotropic as the molecules are now arranged in a regular lattice which defines fixed directions. That is, the solid state is *more ordered* but has *less symmetry* than the liquid state. Unfortunately this liquid-solid transition is an example of a first order phase transition, while our interest below is in continuous or second-order phase transitions. The liquid-vapour transition on the other hand does have a second-order phase transition at the critical point as we will discuss briefly later.)

When one discusses ferromagnetic systems, often it is in the presence of some external magnetic field $\mathcal{H}$, which actually defines the direction of M from one of the many otherwise equivalent possibilities. The susceptibility χ measures the change of the system to a change in the external field and is given by

$$\chi \equiv \frac{1}{V}\frac{\partial M}{\partial \mathcal{H}}. \tag{4.28}$$

Note: Just as the pressure P and volume V were thermodynamic parameters for a gas, the relevant parameters for a ferromagnetic system are $\mathcal{H}$ and M. In fact $\mathcal{H}$ plays a role analogous to the intensive parameter P and $-M$ (note the minus sign) plays a role similar to the extensive parameter V. For example the work done *by* the ferromagnetic system in an infinitesimal change is given by

$$dW = -\mathcal{H}\, dM\,. \tag{4.29}$$

4.7 Correlation Function and Critical Exponents

Let the order parameter M be written as a volume integral (in d dimensions) over an order-parameter density $m(\vec{r})$. Then

$$M \ = < \int d^d r \ m(\vec{r}) > \tag{4.30}$$

where $<>$ represents the statistical average by which one obtains thermodynamic functions and $d^d r$ is the volume element in d dimensions. A useful quantity that can now be defined is the correlation function

$$\Gamma(\vec{r}) \ = < m(\vec{r}) m(0) > - < m(\vec{r}) >< m(0) > . \tag{4.31}$$

It measures how the value of the order parameter at one point is correlated to its value at some other point. If Γ decreases very fast with distance, then far away points are relatively uncorrelated and the system is dominated by its microscopic structure and short-ranged forces. On the other hand, a slow decrease of Γ would imply that faraway points have a large degree of correlation or influence on each other. The system thus becomes organsied at a macroscopic level with the possibility of new structure beyond the obvious one dictated by the short-ranged microscopic forces. As we shall see below, this possibility does actually occur.

Usually, near the critical point ($T \to T_c$), the correlation function can be written in the form

$$\Gamma(r) \to r^{-p} \ \exp(-r/\xi) \tag{4.32}$$

where ξ is the **correlation length**. The correlation length is a measure of the range over which fluctuations in one region of space are correlated with (influence) those in another region. Two points which are separated by a distance larger than the correlation length will each have fluctuations which are relatively independent, that is, uncorrelated.

Experimentally, the correlation length is found to diverge at the critical point. Thus near the critical point, the correlation length may be written as

$$\xi \sim |t|^{-\nu} , \tag{4.33}$$

where $t = \frac{T-T_c}{T_c}$. The divergence of the correlation length at the critical point means that very far points become correlated. In other words, the long-wavelength fluctuations dominate. *Thus the system near a second-order phase transition loses memory of its microscopic structure and begins to display new long-range macroscopic correlations.*

Exactly at the critical point, the correlation function (4.32) therefore displays a *power law* behaviour $\sim r^{-p}$ with

$$p \equiv d - 2 + \eta \tag{4.34}$$

Here d represents the effective space dimensionality of the system. The quantities ν and η are examples of what is known as **critical exponents**. Experiments, supported by renormalization group theory, have shown that systems undergoing second-order phase transitions can be grouped into universality classes. Within each universality class, very different systems with widely different critical temperatures, have approximately the same critical exponents. The reason for this is precisely the loss of memory mentioned above, so that systems with different microscopic structures can give rise to the same long-range behaviour.

In addition to the above two critical exponents, there are four more, $\alpha, \beta, \gamma, \delta$ defined by

Heat Capacity: $C \sim |t|^{-\alpha}$

Order Parameter: $\frac{M}{V} \sim |t|^{\beta}$

Susceptibility: $\chi \sim |t|^{-\gamma}$

Equation of state (at t=0): $\frac{M}{V} \sim |\mathcal{H}|^{1/\delta}$.

Here M/V refers to the magnetisation density. The first three exponents refer to the case $\mathcal{H} = 0$. Note that in general the thermodynamic quantities near the critical temperature may contain both finite and singular parts (or a part with singular derivatives). It is the singular dominant parts to which one refers to above with the $\sim$ symbol above (the symbol also implies a proportionality constant).

4.7.1 Critical Opalescence

As mentioned above, at the critical point the correlation length diverges so that the correlation function obeys a power-law. This of course implies that the system has no characteristic length at this point and is scale-invariant, exhibiting fluctuations of all length scales.

Consider liquid water heated in a sealed container. As the temperature is raised, more of the liquid will vaporise, quickly increasing the density of the gas phase. At the same time the density of the liquid phase decreases, so that eventually a point is reached (about 647K and 218 atm) where the two phases have the same density. Near this point one finds in the fluid domains of liquid-like and vapour-like phases of all sizes (that is, scale-invariant fluctuations between the two phases). In particular once there are liquid drops of the same size as the wavelength of visible light, strong scattering takes place giving the fluid a cloudy appearance: this is called *critical opalescence*.

Experimentally it is easier to observe critical opalescence in binary fluid mixtures that show a second-order phase transition with regard to their mixing. See pictures in Ref[4].

4.8 Scaling Laws

Using the scaling hypothesis as described in the appendix, four scaling relations can be obtained among the six critical coefficients:

Josephson: $\nu d = 2 - \alpha \quad ,$

Rushbrooke: $\alpha + 2\beta + \gamma = 2 \quad ,$

Widom : $\gamma = \beta(\delta - 1) \quad ,$

Fisher: $\gamma = \nu(2 - \eta) \; .$

In the table the experimental values of the scaling relations are compared with the theoretical predictions. The agreement is very good! Thus near a second order phase transition, as the correlation length diverges, a system loses memory of the underlying microscopic structure so that **different systems** obey the same **universal relations**. These relations may be considered as further examples of *emergent laws.*

4.8.1 The Ising Model

Also shown in the table are results for the Ising model. The Ising [8]model is a very simple model of ferromagnetic systems. The model in d-dimensions consists of a periodic lattice (hexagonal, cubic etc) with fixed lattice points. Attached to each lattice site i are spin variables s_i which take either the value $+1$ or 1. These spin variables represent the atomic magnets. Each spin is allowed to interact with only its nearest neighbours or an external magnetic field. The spin-spin interaction is such that it is energetically favourable for neighbouring spins to align. Therefore at low temperatures, where thermal fluctuations are small, one might expect that the thermodynamically stable state will correspond to a state of spontaneous magnetization (even in the absence of an external magnetic field). That is, although the spins only interact with their neighbours, the net result can be a cooperative state, in which far-away spins become correlated. Thus one sees the emergence of long ranged (macroscopic) correlations and order at low enough temperatures even though the microscopic model has only short-ranged interactions.

Quantitatively, it is remarkable that such a crude model gives theoretical predictions for the critical exponents and scaling relations that agree with experiments on real systems. This again illustrates the power of universality near the critical point: the microscopic (short-distance) information of the system, whether real or theoretical, is washed out, leading to common and similar macroscopic properties for systems within each universality class. (Universality classes differentiate between systems in different dimensions with different underlying symmetries).

Although the Ising model is simple to state, the computation of its thermodynamic properties is extremely involved and for $d = 3$ requires numerical effort using a computer. Nevertheless many other models, as simple as the Ising model but different from it have been used to model real systems and to compute and compare their critical properties. This is not only a test of universality but gives insight into properties of varied systems.

4.9 Percolation

Percolation is a geometric analog of thermal phase transitions that is interesting on its own. Consider a square lattice, large enough (ideally infinite) so that we may ignore boundary effects in our discussion. Let each site of the lattice be empty or occupied with a probablity p: The occupation of the sites is decided by a random process, so the sites are independent of each other. Now for $p = 0$, all sites are empty while for $p = 1$ all sites are occupied. Define a *cluster* as a group of nearest neighbour sites that are occupied. As p increases from 0, a critical point $p = p_c$ is reached when a large cluster is formed stretching from one edge of the lattice to the opposite edge.

The value p_c is called the percolation threshold and at this point there is a significant change in the properties of the lattice. For example, if sites represent pores in a rock, and being occupied means the pores are open, then at the percolation threshold water can seep through from one end of the rock to the other. There are other physical problems that can be studied with a percolation model, such as forest fires, or conductivity of a random network.

Since percolation is a random process each simulation on a lattice for fixed p will give rise to different clusters of varied sizes and one must discuss statistical properties of relevant quantities (such as cluster size) obtained after an averaging. It is found that near the percolation threshold the physically interesting quantities diverge and show power-law behaviour similar to that near the crtical point of a second-order phase transition, with p playing the role analogous to temperature. Therefore for the percolation problem one can again define critical exponents and show their universality (that is, independence from underlying lattice type).

At the critical point, the structure of the clusters becomes fractal, that is, there are clusters of all scales and the self-similarity dimension is fractal. This is perhaps not very surprising since at the critical point the properties of the system become scale-invariant and obey power-laws.

4.10 Summary

Statistical mechanics supplements the deterministic fundamental laws with probabilistic tools in order to obtain effective descriptions of equilibrium macroscopic systems consisting of a large number of particles. This gives rise to *emergent* or effective laws that are not apparent at the microscopic level. The more useful of these laws are those that show *universality*, that is are independent of the microscopic details of theory. This was emphasized already for the case of the ideal gas law.

For equilibrium systems, the concept of entropy summarizes via the Second Law the tendency of isolated systems to move towards greater disorder. As we saw from the statistical definition of entropy, the Second Law itself is an emergent law valid for large systems: Indeed the Second Law suggests macroscopic irreversibility eventhough the microscopic dynamics is reversible. For systems that are not isolated but with some state-variables kept constant, the free-energies F and G determine the equilibrium states through extremal principles.

More usually quoted examples of universal emergent laws are those that arise near the critical point of a second-order phase transition. Near the critical point the correlation length diverges so that the system becomes scale-invariant, meaning that its properties become insensitive to the microscopic structure and display some universality.

Percolation is a geometric analog of thermal systems that shows behaviour similar to that of second order phase transitions.

It should be noted that because of universality near the critical point, the very simple models (whether for thermal systems or percolation) give results for critical properties that are in quantitative agreement with experiments even though the real microscopic dynamics of the experimental systems might be much more involved. This will not always be the case for other complex systems that we will study, where usually the agreement between models and reality will be qualitative at best. Nevertheless the virtue of the models even in those cases is that they highlight the important features that enable one to obtain crucial insight that might otherwise be lost in a mass of detail if the system was represented and studied by more accurate equations.

4.11 Appendix

Although one can do so more rigorously, simply using the mapping descibed above Eq.[4.29] allows us to obtain the relation

$$M = -\left(\frac{\partial G}{\partial \mathcal{H}}\right)_T \tag{4.35}$$

Where G is now the Gibbs free energy of the ferromagnetic system with $\mathcal{H}$ and M the relevant parameters. Another quantity of interest is the heat capacity given by

$$C_{\mathcal{H}} = \left(\frac{\partial Q}{\partial T}\right)_{\mathcal{H}} = -T\ \left(\frac{\partial^2 G}{\partial T^2}\right)_{\mathcal{H}} \tag{4.36}$$

Under an infinitesimal reversible transformation one has, $dG = -SdT + VdP$, where $G = U - TS + PV$ is the Gibbs free energy. Thus one deduces the relations

$$S = -\left(\frac{\partial G}{\partial T}\right)_P \tag{4.37}$$

$$V = \left(\frac{\partial G}{\partial P}\right)_T \tag{4.38}$$

$$C_P = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_P \tag{4.39}$$

These relations will be used in the next subsection.

4.11.1 The Scaling Hypothesis

If a quantity is dimensionless (for example, fractional volume), then it remains unchanged when the length scale is changed. On the other hand, any dimensional quantity (for example, volume) must be expressed in terms of some unit of length and it will change when that unit is changed.

In general a system will have an (or more than one) intrinsic length scale, such as the mean distance between nearest lattice points (atoms) in a crystal. The **scaling hypothesis** posits that near the critical point the correlation length, ξ, is the only characteristic length scale in terms of which all other quantities with dimensions of length are to be measured. Using this assumption one can derive a number of scaling laws which can be compared with experiment. (A more rigorous derivation of scale invariance and critical exponents requires renormalization group theory which is beyond the level of this course. Ken Wislon received the Nobel prize for that development [5]).

Let us first determine the length dimensions of various quantities. The Gibbs free energy has the same dimension as energy so G/kT is dimensionless. The Gibbs free energy per unit volume $g = G/kTV$ therefore has the dimension (length)$^{-d}$. One writes this statement as

$$[g] = L^{-d} \tag{4.40}$$

Note: Although the real world is three dimensional, systems can be studied under conditions which effectively reduces their diemsnionality to two or one. Thus d is the effective dimensionality of the system.

The correlation function has by definition the length dimension $2-d-\eta$.

$$[\Gamma] = L^{2-d-\eta}\,. \tag{4.41}$$

Therefore it follows from the definition of Γ that

$$[\frac{M}{V}] = L^{(2-d-\eta)/2}. \tag{4.42}$$

Hence from (4.35), the dimension of the magnetic field is

$$[\mathcal{H}/kT] = L^{(\eta-2-d)/2} \tag{4.43}$$

and then from Eq.(4.28) the dimension of χ is

$$[kT\chi] = L^{2-\eta}\,. \tag{4.44}$$

Now since the scaling hypothesis states that ξ is the only characteristic scale near the transition temperature, one replaces L in the formulae above by ξ and uses also the definition that $\xi \sim t^{-\nu}$ to obtain the critical exponents. Comparing those with the definitions given one obtains four relations:

$$\begin{aligned} 2-\alpha &= \nu d && (4.45)\\ \beta &= -\nu(2-d-\eta)/2 && (4.46)\\ \gamma &= \nu(2-\eta) && (4.47)\\ \beta\delta &= \nu(2+d-\eta)/2\,. && (4.48) \end{aligned}$$

combinations of these relations result in the four named scaling laws described above.

4.12 Exercises

1. (a) How does one know where the solid, liquid and gas regions are in an unlabled $P-T$ plot of a substance?
   (b) What is the significance of the triple and critical points?
   (c) How does the $P-T$ plot for water differ from that of a generic substance?
   (d) Do you know why the "anomalous" $P-T$ property of water is wonderfully important ?

2. Show that for a thermodynamic system with parameters P, V and T, the heat capacities are
   (a) $C_V \equiv \left(\frac{\partial Q}{\partial T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V$.
   (b) $C_P \equiv \left(\frac{\partial Q}{\partial T}\right)_P = \left(\frac{\partial H}{\partial T}\right)_P$, where $H = U + PV$ is called the **enthalpy** (not to be confused with the entropy!).
   (c) Obtain C_v and C_P for an ideal gas. Why is C_p larger than C_v ?

3. Use the definitions (4.19) and (4.20) together with the equation of state of an ideal gas to determine the entropy of an ideal gas.

4. Show that heat conduction, along a metal bar between two heat reservoirs at temperatures T_1 and T_2 leads to an increase of entropy of the whole system. The result again conforms to our experience that heat flow is an irreversible process in the sense that it flows from the hot to cold object and never the other way.

5. Diffusion is a familiar physical example of the Second Law in action. Give some examples of diffusive processes. Where else in this course have you learnt about diffusion in a useful model of nature?

6. The Second Law can be stated in many equivalent ways. One macroscopic statement of the Second Law is as follows (**Clausius**): *In any cyclic transformation throughout which temperature is defined, one has*

$$\oint \frac{dQ}{T} \leq 0, \tag{4.49}$$

where the integral is over one cycle. The equality holds for reversible transformations. Show that as a consequence of this statement
(a) For any transformation,

$$\int_A^B \frac{dQ}{T} \leq S(B) - S(A) \tag{4.50}$$

(b) The entropy of a thermally isolated system never decreases (i.e. recover the original statement of the second law in the text).

7. Two useful state functions are the Helmhotz free energy F, and the Gibbs free energy G, defined by $F = U - TS$ and $G = F + PV$. Show using (4.50) that
(a) For a mechanically isolated system kept at constant temperature, the Helmhotz free energy never increases (thus the state of equilibrium of a mechanically isolated system at constant temperature is one of minimum F.) Does this statement reduce to something familiar in the limit of zero temperature ?
(b) For a system kept at constant temperature and pressure, the Gibbs potential never increases (thus the state of equilibrium of a system kept at constant temperature and pressure is the state of minimum Gibbs free energy.). Convince yourself that a limiting case of this reproduces (a).

(c) The above are examples of extremum principles, applicable to more general situations than the first version of the Second Law mentioned in the text (which applied only to isolated states, U and V constant). These principles determine the thermodynamically stable equilibrium states. Convince yourself that in limiting cases the above extremum principles are equivalent to the first version of the Second Law stated in the text.

8. (a)Show that both sides of equation (4.27) have units of heat energy. (b) Show that melting and boiling result in an increase of entropy of the system.

9. (a) Water vapour, where the atoms are free to move more randomly, clearly corresponds to a state of larger entropy than the state of liquid water. So why does liquid water at any fixed temperature remain liquid and not spontaneously vaporize to increase its entropy ? (b) The above example highlights why things do not spontaneously 'disintegrate' or 'decay', as suggested by misinterpretations of the Second Law. Can you now explain why paper does not undergo spontaneous combustion into a disordered mess ? (*Hint:* Check out Ref.[3])

10. (a) Notice that near the crtitical point of a second order phase transition one obtains power law behaviour of relevant quantities. Where else in this course have you come across power laws ? (b) Compare the concepts of "characteristic length" and "scale invariance" used here with their use earlier in the course. (c) Compare the universality in second-order phase transitions with its earlier usage in the course.

11. There are many Ising Model simulations freely available on the web. (a) Run the program at Ref.[6] for various temperatures, trying both hot and cold initialisations, and looking at various plots. Do the results appear reasonable ? (b) Try also the smaller online simulation at Ref.[7].

12. (a)Explore the two dimensional percolation online simulation package at Ref.[9]. Try various lattice sizes and probability values. Check for the cluster sizes. (b) Estimate the critical probability and compare with the theoretical value 0.5928 (Remember that the crtical probability refers to a lattice

of infinite extent. In practise this means you should explore larger and larger lattices and try to find upper and lower bounds on your p_c).

4.13 References

1. Statistical Mechanics, by Kerson Huang.
2. The Feynman Lectures on Physics, by Feynman, Leighton and Sands.
3. A website explaining ”Why things dont go wrong more often” is at http://www.secondlaw.com/default.htm
4. A web demonstration of critical opalescence is at http://www.physicsofmatter.com/CriticalOpal/OpalFrame.html
5. Wilson at http://www.physics.ohio-state.edu/ kgw/kgw.html
6. A Ising model simulation package is at at http://bartok.ucsc.edu/peter/java/ising/keep/ising.html .
7. A small online Ising model simulation is at http://www.phy.syr.edu/courses/ijmp_c/Ising.html
8. Ising at http://www.bradley.edu/las/phy/ising.html
9. An online percolation simulation package is at http://kzoo.edu/ jant/p705/Percolate.html
10. Percolation Theory, by D. Stauffer and A. Aharony.

Chapter 5

Self-Organised Criticality

When we studied equilibrium systems, we saw that many thermodynamic relations or laws could be derived by averaging over the unobserved microstates of the system. The resulting relations were thus emergent laws that are incredibly accurate for large systems. It is important to note, for contrasting with what follows below, that the macrosocpic laws derived in chapter(4) describe **each** macroscopic state. On the other hand, many complex systems in nature display regularities in a statistical description of their observable *macrostates*.

5.1 Power Laws in Nature

An example of a statistical macrosopic relation is the distribution of the magnitude of earthquakes. If $N(E)$ is the annual mean number of earthquakes (in a zone or worldwide) of size E ($\sim$ energy released), then empirically one finds over a wide range,

$$N(E) = E^{-b} \tag{5.1}$$

with the constant $b \sim 1$. The relation (5.1) is called the Gutenberg-Richter law and is obviously a statistical relation for observables – it does not specify when an earthquake of some magnitude will occur but only what the mean distribution in their magnitude is.

The Gutenberg-Ricter law is a power-law and is therefore scale-invariant – a change of scale in M can be absorbed in a normalisation constant, leaving the form of the law invariant. The scale-invariance of the law implies a scale-invariance in the phenomena itself: earthquakes happen on all scales and there is no typical or mean magnitude!

There are many other natural phenomena which exhibit power laws *over a wide range of the parameters*: Volcanic activity, solar-flares, charge released during lightning events, length of streams in river networks, forest fires, and even the extinction rate of biological species! Some of these power laws refer to spatial scale-free structures, or fractals, while some others refer to temporal events and are examples of the ubiquitous "one-over-f " phenomena (see chapter 2).

Can the frequent appearance of such power laws in complex systems be explained in a simple way? Note that the systems mentioned above are examples of dissipative structures, with a slow but constant *inflow of energy and its eventual dissipation.* The systems are clearly out of equilibrium, since we know that equilibrium systems tend towards uniformity rather than complexity. On the other hand the abovementioned systems display scale-free behaviour similar to that exhibited by equilibrium systems near a critical point of a second-order phase transition.

However while the critical point in equilibrium systems is reached only for some specific value of an external parameter, such as temperature, for the dissipative structures above the scale free behaviour appears to be robust and *does not seem to require any fine-tuning.* Bak and collaborators proposed that many dissipative complex systems naturally self-organise to a critical state, with the consequent scale-free fluctuations giving rise to power laws. In short, the proposal is that **self-organised criticality** is the natural state of large complex dissipative systems, relatively independent of initial conditions. It is important to note that while the critical state in an equilibrium second-order phase transition is unstable (slight perturbations move the system away from it), the critical state of self-organised systems is *stable*: systems are continually attracted to it!

The idea that many complex systems are in a self-organised critical state is intuitively appealing because it is natural to associate complexity with a state that is balanced at the edge between total order and total disorder (sometimes loosely referred to as the "edge of chaos"). Far from the critical point, one typically has a very ordered phase on one side and a greatly disordered phase on the other side. It is only at the critical point that one has large correlations among the different parts of a large system, thus making it possible to have novel emergent properties, and in particular scale-free phenomena.

In addition to the examples mentioned above, self-organised criticality has also been proposed to apply to economics, traffic jams, forest fires and even the brain!

5.2 Models

How does one test the idea of self-organised criticality? It is quite hopeless to solve the complete dynamical equations for the relevant systems and try to see if they do self-organise to the critical state. Rather one makes highly simplified models to test the idea. If the models display self-organised criticality *and are robust to various changes in detail at the microscopic level*, then one can take that as supporting the plausibility of the proposal.

A simple model of landslides or avalanches is that of a two-dimensional sand-pile. On a grid, define the local slope at each site (square on the grid) by an integer Z. The simulation procedure is as follows. First choose a critical value Z_c and populate the grid with a starting configuration with each site having a random value less than Z_c. The updating rules are then:

1. Choose a random site and increase Z by one (add sand).
2. If Z exceeds critical value, reduce Z by four units and redistribute one unit to each of four neighbours.
3. Check if Z exceeds the critical value at any of the neighbours and continue redistribution process until the avalanche stops.
4. Count the total number of "topplings" involved in that avalanche.
5. Go to step (1)

Thus a number of avalanches of different sizes (number of topplings) are generated. By plotting the number of avalanches against their size for a large lattice (sand that goes out of the boundary is lost) one finds a power law over a large range of the parameters. A power law is also obtained for the distribution of lifetimes of the avalanches.

What is interesting is that *adding a single grain (step 1) can initiate an avalanche of any size, from one that involves only a few grains to one that involves almost the whole pile*. That is, the system shows scale-invariance. (Notice that at the critical state the output is not proportional to the input, that is, the system is highly nonlinear and gives rise to "non-obvious" effects)

After an avalanche has reduced the slope everywhere to below the critical value, the *slow* adding of send again brings the pile to another critical state and more avalanches. That is, the system continually self-organises to the critical state from which scale-free avalanches occur.

Although the effect of adding each particular grain is almost impossible to guess, the statistical distribution of avalanches according to an approximate inverse power law implies that small avalanches are more frequent than the larger ones. Note that a crucial feature of the models, which is a reflection of the real systems, is that the external process that drives the system (the inflow), occurs much slower than the faster internal reorganization processes that result in dissipation.

Many other models have been investigated showing self-organised criticality. It has thus been suggested that self-organised criticality might not only be the reason for the diverse power laws in nature but also the *dynamical mechanism* behind fractal geometry and one-over-f temporal phenomena (see chapter 2).

5.3 Experiments

Do any real systems behave like the simple computer model above? Although many natural complex phenomena seem to show power laws, the computer models are such extreme caricatures of the physical situations that, though they show SOC, it is unclear whether SOC would occur in real systems similarly tested. That is, one would like to perform controlled experiments on real dynamical systems to test the idea that they are attracted to a critical state.

Some initial experiments with real sandpiles showed power-law behaviour but others did not. It was soon realized that the reason was probably because real sand is not like the computer particles: Real sand particles are difficult to stop once they start rolling, so that large avalanches are the norm and the dynamics is dominated by individual motion of grains over long distances. By contrast the computer model in the last section implicitly assumed ideal particles of low inertia (so that little kinetic energy was accumulated as it rolled) with power laws resulting from *collective* dynamics.

Thus one is led to design an experiment with particles that do not readily roll or slip over one another. Long-grained rice was used in an experiment performed at the Univeristy of Oslo as it has low inertia and its large-aspect ratio ensured that it did not roll easily. The results of the experiment did indeed show power laws.

The moral of this is that while SOC is indeed a physical phenomena, and not just an artifact of simple computer models, it might not necessarily be an automatic state of all complex systems.

5.4 Life

As we have seen above, while the approach to the critical state can take a long time, once the system achieves crticality small perturbations can trigger avalanches of all sizes. That is, long periods of "quiet" can be interrupted by large events. Interestingly, prior to the idea of SOC, the paleontologists Gould and Eldredge had proposed the idea of "punctuated equilibrium" to describe how evolution seemed to occur: Long periods of stasis punctuated by extinction and evolution events of all sizes. Clearly the independently proposed idea of SOC is similar and supposedly of applicability to other complex systems.

Different species compete for resources, forming a large and complex evolving system. Thus it seems possible that the SOC scenario is realized for evolution. Indeed the extinction rate for the last 500 million years, from data collected by Sepkoski, does seem to follow a power law. What is more, such power law behaviour is also seen in various computer simulations of evolution, such as Ray's "Tierra" model and the simple Bak-Sneppen model.

The Bak-Sneppen model consists of a one dimensional cellular automaton with N species placed on a circle with nearest neighbour interactions. In this model, evolution is approximated to act at the level of a species rather than at the level of the individual; thus each species is represented by a single fitness level $0 < f < 1$. At each discrete time step, the species with the lowest fitness level is made extinct and is replaced by a new species with a randomly assigned fitness. In addition, the two neighbouring species are also replaced by new species with randomly assigned fitness values. The frequency power spectrum of the changes experienced by any given species in time is found to follow a power law.

An interesting feature of the Bak-Sneppen model is evolution by the mechanism of "elimination of the least fit" rather than the popular folklore of "survival of the fittest". This extremal dynamics whereby the weakest link in a complex network is removed, or breaks down, is probably a realistic modelling of most natural phenomena.

Indeed, Chialvo and Bak [3] have used extremal dynamics in their model of a brain that learns from mistakes. In their neural network model, whenever mistakes are made, all the synapses that contributed to that decision are punished. This is in contrast to usual Hebbian rules where neurons that perform well are strengthened. Chiavlo and Bak show that their model is able to achieve quick learing of new patterns because their network is closer to a "critical state" rather than the sub-critical (highly stable) states of traditional models.

5.5 War and Peace

L.F. Richardson's pacifist leanings motivated him to search for the dynamics and reasons for wars. We have already mentioned the Richardson model of the arms race earlier. Richardson also found that the statistical distribution relating the number of wars to a given intensity, defined as the number of battle dead, followed a power-law. Richardson considered 82 conflicts between 1820 and 1929, and the power-law trend has been confirmed by Levy who investigated 119 wars between 1495 to 1973.

Although Richardson hoped that his mathematical studies of deadly conflicts would lead to a way of reducing aggression, he apparently came to the conclusion that war was an intrinsic feature of mankind.

In [8] the authors argue that the outbreaks of war are analgous to the "avalanches" of all sizes that occur in a complex system that continually self-organises to a critical state.

If indeed world order behaves as a self-organsied critical system, relatively unaffected in the long term by efforts to moderate and change behaviour, then the conclusion appears pessimistic for the human species. Will humans ever learn from the past or are we doomed to repeat the mistakes of the past? Will the human species be the first to annihilate itself ? (see the exercises).

5.6 Zipf's Law

In 1932 George Zipf published an empirical rule describing a statistical regularity in the distribution of words used in any large text.

Let us rank (r) words in a given text according to the frequency (f) of their occurence. So, the most frequently occuring word would be given the rank $r = 1$, the next most frequently occuring word, $r = 2$ and so on. Zipf's law states that

$$f(r) \sim r^{-\alpha} \tag{5.2}$$

with $\alpha \sim 1$. This power law has been tested over a large volume of literature and also different languages and is found to be accurate for words whose rank is not too low or too high! [10]. It has been argued [11] that for larger collection of texts, there are significant deviations from a single power law and that at large ranks there is a transition to a second power law regime.

It is still unclear as to how one may explain the law within a model of human behaviour, or as to how one may use the generalised law to characterise and differentiate human writing from other froms of text. However it has ben argued [12] that the distribution of a word in text can be used to distinguish whether it is a noun or verb!

Zipf also noted that a similar power law applied to the sizes of cities. That is, if the largest city in a country was assigned the rank $r = 1$, the second largest $r = 2$, then a power law as above is found. This has been found to be true for many countries but not for those with central planning. Thus one can hope that the explanation of the power-law of city size distribution can be found in the self-organising properties of a society. In reference [13] the authors consider a model of city formation and find that if the individuals in the model interact pairwise then Zipf's law emerges for larger cities. Other models have also been considered as discussed in that reference.

5.7 Summary

Self-organised criticality combines the ideas of self-organisation in dissipative structures with those of critical phenomena in equilibrium second-order phase transitions, in an attempt to explain the frequent appearance of scale-free statistical laws that summarize events in many complex systems.

It has been suggested that SOC might be the mechanism behind the Gaia hypothesis of Lovelock, where the earth is treated as a single organism far from equilibrium. More ambitiously there is an attempt by Smolin to explain the complexity of the universe also in this approach.

As yet, a comprehensive analytical understanding of SOC, comparable to that of critical phenomena in equilibrium second-order phase transitions, is lacking but some work has been done (see, for example, Ref.[2] and references therein).

5.8 Exercises

1. Try the sandpile applet at Ref.[7], varying the parameters and the conditions, and compare what you see with what has been described above.

2. (a) Explain where dissipation occurs in the sand-pile model.
   (b) Explain why the sand must be added slowly to the sandpile. What would happen otherwise ?
   (c) In actual simulations, the straight line fit on a log-log plot does not hold for very large avalanche sizes. Explain why.

3. Read the article [3] and discuss the applications of the self-organised criticality concept.

4. Read the article [8].
   (a) What evidence is there that forest fires are an example of a self-organised critical system?
   (b) What similarities are there between forest fires and wars ?
   (c) What in your opinion is the best way to prevent conflicts from getting very large and destructive ?
   (d) Can you justify your answer in (c) in the face of claims of self-organised criticality of the system ?
   (e) Check out Ref.[9] and others similar in spirit.

5. Check out the websites below of Zipf's law.
   (a) Find other situations where Zipf's law applies.
   (b) Discuss possible human factors that could be the drving force behind the laws.
   (c) Is Zips's law an example of self-organised criticality ?

5.9 References

1. How Nature Works, by P .Bak

2. Self-Organized Criticality, by H.J. Jensen, Cambridge lecture notes in physics,1998.

3. Self-Organization of Complex Systems, by M. Paczuski and P.Bak, at http://xxx.lanl.gov/abs/cond-mat/9906077

4. The Self-Made Tapestry, by P. Ball

5. Gaia, by J.E. Lovelock.

6. Life of the Cosmos, by L. Smolin.

7. A sandpile applets at
   http://pauli.uni-muenster.de/ special/java/sandpile/ver1.1/ and
   http://cmth.phy.bnl.gov/ maslov/Sandpile.htm

8. Fractality and self-organised criticality of wars, by D. Roberts and D. Turcotte, in Fractals, Vol.6 (1998) 351-357
   http://www.worldscinet.com/journals/fractals/64/robe.html .

9. The Richardson Institute for Conflict and Peace Research at http://www.lancs.ac.uk/depts/richinst/home/

10. Some websites on Zip's law are at http://linkage.rockefeller.edu/wli/zipf/ and http://www.cut-the-knot.com/do_you_know/zipfLaw.html

11. Beyond the Zipf-Mandelbrot law in quantitative linguistics, by M. Montemurro, http://xxx.lanl.gov/abs/cond-mat/0104066.

12. See the article at http://www.nature.com/nsu/010927/010927-1.html

13. Interacting individuals Leading to Zipf's Law, by M. Marsili and Y. Zhang, http://xxx.lanl.gov/abs/cond-mat/98011289.