Multiscale Analysis of Complex Time Series: Integration of Chaos and Random Fractal Theory, and Beyond
by Jianbo Gao, Yinhe Cao, Wen-wen Tung and Jing Hu

CHAPTER 2

OVERVIEW OF FRACTAL AND CHAOS THEORIES

The formal treatment of fractal theory will not begin until Chapter 5, and chaos theory will not be treated in depth until Chapter 13. In this chapter, we devote two sections to fractal and chaos theories to provide readers with a general overview of these theories without getting into the technical details.

2.1 PRELUDE TO FRACTAL GEOMETRY

Euclidean geometry is about lines, planes, triangles, squares, cones, spheres, etc. The common feature of these different objects is regularity: none of them is irregular. Now let us ask a question: Are clouds spheres, mountains cones, and islands circles? The answer is obviously no. In pursuing answers to such questions, Mandelbrot has created a new branch of science — fractal geometry.

For now, we shall be satisfied with an intuitive definition of a fractal: a set that shows irregular but self-similar features on many or all scales. Self-similarity means that part of an object is similar to other parts or to the whole. That is, if we view an irregular object with a microscope, whether we enlarge the object by 10 times or by 100 times or even by 1000 times, we always find similar objects. To understand this better, let us imagine that we were observing a patch of white cloud drifting away in the sky. Our eyes were rather motionless: we were staring more or less in the same direction. After a while, the part of the cloud we saw drifted away, and we were viewing a different part of the cloud. Nevertheless, our feeling remained more or less the same.

Mathematically, a fractal is characterized by a power-law relation, which translates into a linear relation in the log-log scale. In Chapter 5, we shall explain why a power-law relation implies self-similarity. For now, let us again resort to imagination. We are walking down a wild, jagged mountain trail or coastline. We would like to know the distance covered by our route. Suppose our ruler has a length of ϵ, which could be our step size, and different hikers may have different step sizes — a person riding a horse has a huge step size, while a group of people with a little child must have a tiny step size. The length of our route is

$$L = N(\epsilon) \cdot \epsilon \,, \tag{2.1}$$

where $N(\epsilon)$ is the number of intervals needed to cover our route. It is most remarkable that typically $N(\epsilon)$ scales with ϵ in a power-law manner,

$$N(\epsilon) \sim \epsilon^{-D}, \;\; \epsilon \rightarrow 0 \,, \tag{2.2}$$

with D being a noninteger, $1 < D < 2$. Such a nonintegral D is often called the fractal dimension to emphasize the fragmented and irregular characteristics of the object under study.

Let us now try to understand the meaning of the nonintegral D. For this purpose, let us consider how length, area, and volume are measured. A common method of measuring a length, a surface area, or a volume is to cover it with intervals, squares, or cubes whose length, area, or volume is taken as the unit of measurement. These unit intervals, squares, and volumes are called unit boxes. Suppose, for instance, that we have a line whose length is 1. We want to cover it by intervals (boxes) whose length is ϵ. It is clear that we need $N(\epsilon) \sim \epsilon^{-1}$ boxes to completely cover the line. Similarly, if we want to cover an area or volume by boxes with linear length ϵ, we would need $N(\epsilon) \sim \epsilon^{-2}$ to cover the area, or $N(\epsilon) \sim \epsilon^{-3}$ boxes for the volume. Such D is called the topological dimension and takes on a value of 1 for a line, 2 for an area, and 3 for a volume. For isolated points, D is zero. That is why a point, a line, an area, and a volume are called $0-D$, $1-D$, $2-D$, and $3-D$ objects, respectively.

Now let us examine the consequence of $1 < D < 2$ for a jagged mountain trail. It is clear that the length of our route increases as ϵ becomes smaller, i.e., when $\epsilon \rightarrow 0$, $L \rightarrow \infty$. To be more concrete, let us visualize a race between the hare and the tortoise on a fractal trail with $D = 1.25$. Assume that the length of the average step taken by the hare is 16 times that taken by the tortoise. Then we have

$$L_{\text{hare}} = \frac{1}{2} L_{\text{tortoise}} \,.$$

That is, the tortoise has to run twice the distance of the hare! Put differently, if you were walking along a wild mountain trail or coastline and tired, slowing down

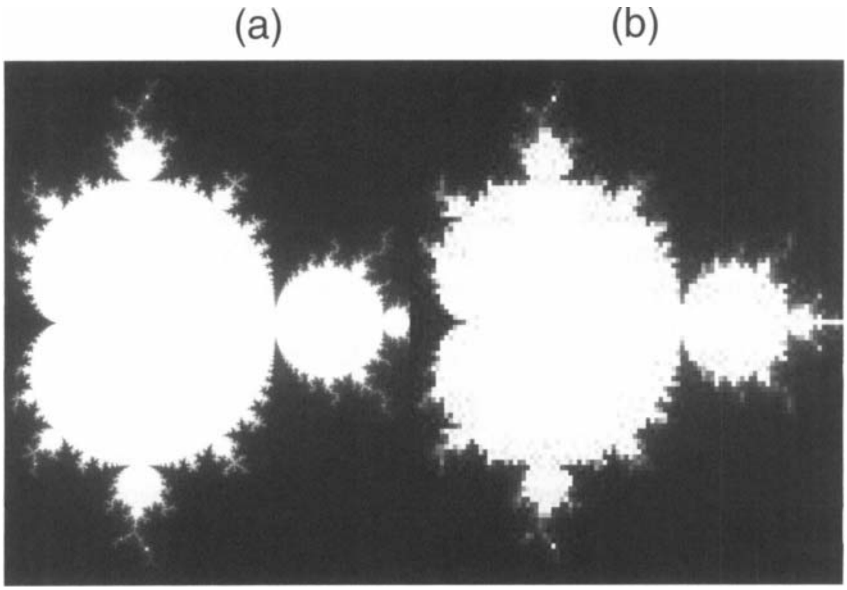

Figure 2.1. Mandelbrot set

your pace and shrinking your steps, then you were in trouble. It certainly would be worse if you also got lost.

To better appreciate the open-endedness of the concept of a fractal, let us say a few words about the path associated with one's desire to achieve a certain nontrivial goal. This can be viewed as an extension of the above discussion.

A highly motivated person trying to achieve a grand goal often desires to be a lucky person such that the route to success is smooth and full of excitement. Typically, however, the opposite is true: the route is full of frustrations and failures, and one could then lose confidence and even become depressed. That is, the path might be a bumpy fractal one. A better strategy to achieve one's goal is perhaps to stop whining, accept the fact that often the most wanted help or environment may not be there, and take steps as large as possible to finish the journey.

Now, a word for outside observers: if you could not help but offer advice, expect your own fractal-like path of frustration. This is because advice is often drawn from past painful struggles, and is thus incompatible with a mind seeking lesser complexity or difficulty, or simply cheers or fun.

Now back to signal processing. In practice, random fractals are more useful than fractal geometry. We shall devote a lot of time to a discussion of random fractal theories. For the moment, to stimulate readers' curiosity as well as to echo Mandelbrot's observation that clouds are not spheres, nor mountains cones, nor islands circles, we design a simple pattern recognition problem related to the sensational Mandelbrot set.

The Mandelbrot set is the set of the complex C such that for $Z_0 = 0$, Z_n defined by the following simple mapping,

$$Z_{n+1} = Z_n^2 + C, \tag{2.3}$$

remains bounded. Fig. 2.1 shows an example of C in the square region $[-0.5, 1.5] \times [-1.2, 1.2]$, where (a) has a resolution of 800×800 and (b) has a resolution of 100×100. In other words, (b) is a downsampled version of (a). Having a lower resolution, some features in (b) are blurred. *Now we ask*: Can we recover (a) from (b), fully or partly, by using conventional pattern recognition methods to classify the values of C that are close to the low-resolution Mandelbrot set of Fig. 2.1(b)? In one of the author's (J.B.'s) pattern recognition class, Mr. Jason Johnson took such a challenge and tried a few classification methods. But none of them worked. Now that fractal phenomena have been found to be ubiquitous, it is time for us to think seriously about how conventional stochastic pattern recognition methods can be integrated with fractal-based methods so that broader classes of pattern recognition problems can be tackled with higher accuracy.

2.2 PRELUDE TO CHAOS THEORY

Imagine that we are observing an aperiodic, highly irregular time series. Can a signal arise from a deterministic system that can be characterized by only a very few state variables instead of a random system with infinite numbers of degrees of freedom? A chaotic system is capable of just that. This discovery has such far-reaching implications in science and engineering that sometimes chaos theory is considered one of the three most revolutionary scientific theories of the twentieth century, along with relativity and quantum mechanics.

At the center of chaos theory is the concept of sensitive dependence on initial conditions: a very minor disturbance in initial conditions leads to entirely different outcomes. An often used metaphor illustrating this point is that sunny weather in New York could be replaced by rainy weather sometime in the future after a butterfly flaps its wings in Boston. Such a feature contrasts sharply with the traditional view, largely based on our experience with linear systems, that small disturbances (or causes) can only generate proportional effects, and that in order for the degree of randomness to increase, the number of degrees of freedom has to be infinite.

Mathematically, the property of sensitive dependence on initial conditions can be characterized by an exponential divergence between nearby trajectories in the phase space. Let $d(0)$ be the small separation between two arbitrary trajectories at time 0, and let $d(t)$ be the separation between them at time t. Then, for true low-dimensional deterministic chaos, we have

$$d(t) \sim d(0)e^{\lambda_1 t} \tag{2.4}$$

where λ_1 is called the largest positive Lyapunov exponent.

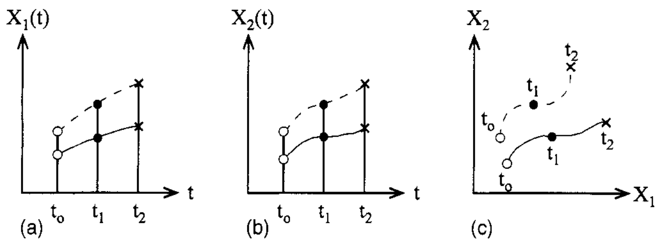

Figure 2.2. Schematic illustrating the concept of phase space.

Another fundamental property of a chaotic attractor is that it is an attractor — the trajectories in the phase space are bounded. The incessant stretching due to exponential divergence between nearby trajectories, and folding due to boundedness of the attractor, then cause the chaotic attractor to be a fractal, characterized by Eq. (2.2), where $N(\epsilon)$ represents the (minimal) number of boxes, of linear length not greater than ϵ, needed to cover the attractor in the phase space. Typically, D is a nonintegral number called the box-counting dimension of the attractor.

At this point, it is worthwhile to take time to discuss a phase space (or state space) and transformation in the phase space. Let us assume that a system is fully characterized by two state variables, X_1 and X_2. When monitoring the motion of the system, one can plot out the waveforms for $X_1(t)$ and $X_2(t)$, as shown in Figs. 2.2(a,b). Alternatively, one can monitor the trajectory defined by $(X_1(t), X_2(t))$, where t appears as an implicit parameter (Fig. 2.2(c)). The space spanned by X_1 and X_2 is called the phase space (or state space). It could represent position and velocity, for example.

The introduction of phase space enables one to study the dynamics of a complicated system geometrically. For example, a globally stable fixed point solution is represented as a single point in the phase space; solutions with arbitrary initial conditions will all converge to it sooner or later. Similarly, a globally stable limit cycle is represented as a closed loop in the phase space; again, solutions with arbitrary initial conditions will all converge to it.

To make readers unfamiliar with the concept of phase space even more comfortable, we note that this concept is frequently used in daily life. As an example, let us imagine that we were driving to our friend's home for a party. On our way, there was a traffic jam, and our car got stuck. Afraid of being late, we decided to call our friend. How would we describe our situation to her? Usually, we would tell her where we got stuck and how quickly or slowly we were driving, but not the signal waveforms shown in Figs. 2.2(a,b).

We can now talk about the transformations in phase space. For this, we ask readers to imagine how a patch of dust would be swept across the sky on a very windy day: if the dust originally resembled the face of a person, after being swept the face would get twisted badly. As another example, let us consider the fish transformation shown

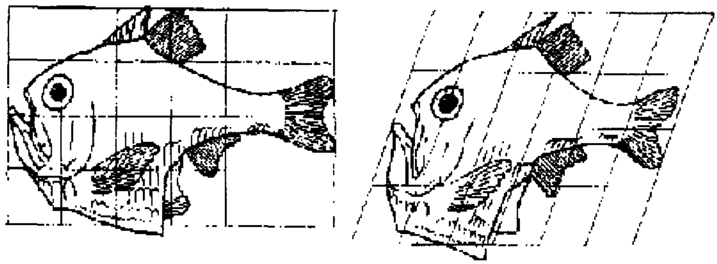

Figure 2.3. To appreciate a chaotic transformation, one can imagine that the head and tail of the fish gets mixed up.

in Fig. 2.3. In his famous book *Growth and Form*, Sir D'Arcy Thompson found that simple plane transformations can bring two different types of fishes together. With awesome intuition, he then concluded that different fishes have the same origin and that this is a vivid demonstration of Darwin's theory of evolution. To appreciate a chaotic transformation, imagine that the head and tail of the fish get mixed up, just as a "dust face" gets twisted by a gusty wind. In exercise 3, we shall make these ideas more concrete.

To appreciate more concretely the concept of sensitive dependence on initial conditions, let us consider the map on a circle,

$$x_{n+1} = 2x_n \mod 1 , \tag{2.5}$$

where x is positive, and mod 1 means that only the fractional part of $2x_n$ will be retained as x_{n+1}. This map can also be viewed as a Bernoulli shift, or binary shift. Suppose that we represent an initial condition x_0 in binary

$$x_0 = 0.a_1a_2a_3 \cdots = \sum_{j=1}^{\infty} 2^{-j} a_j, \tag{2.6}$$

where each of the digits a_j is either 1 or 0. Then

$$x_1 = 0.a_2a_3a_4 \cdots ,$$

$$x_2 = 0.a_3a_4a_5 \cdots ,$$

and so on. Thus, a digit that is initially far to the right of the decimal point, say the 40th digit (corresponding to $2^{-40} \approx 10^{-12}$), and hence has only a very minor role in determining the initial value of x_0, eventually becomes the first and the most important digit. In other words, a small change in the initial condition makes a large change in x_n.

Since chaotic as well as simple, regular motions have been observed almost everywhere, we have to ask a fundamental question: Can a chaotic motion arise from a regular one and vice versa? The answer is yes, and lies in the study of bifurcations and routes to chaos. Here, the key concept is that the dynamics of a system are controlled by one or a few parameters. When the parameters are changed, the behavior of the system may undergo qualitative changes. The parameter values where such qualitative changes occur are called bifurcation points.

To understand the idea better, let us reflect on graduate student life. In the initial spinning period, which could last for a few years, a student may be learning relevant materials diligently, but has few new ideas and thus is not productive at all. At some point, the student suddenly feels a call to work out something interesting. Beyond that point, he or she becomes quite productive, and the professor can now relax a bit. Here, the parameter we could vaguely identify would be the maturity of the student.

As another example, let us consider cooperation between two companies. When the profit is small, the companies cooperate well and both make profits. After a while, the market has grown so considerably that they become competitors. At a certain point, in order to dominate the market, they decide to sacrifice profit. After a while, the loss makes them seek cooperation again, and so on. It is clear that such a process is highly nonlinear.

To be more concrete, let us now consider the logistic map

$$x_{n+1} = f(x_n) = rx_n(1 - x_n), \quad 0 \leq x_n \leq 1, \tag{2.7}$$

where $0 \leq r \leq 4$ is the control parameter. Let us now fix r to be 2 and iterate the map starting from $x_0 = 0.3$. With a calculator or computer, we quickly find that x_n becomes arbitrarily close to 0.5 for large n. Now if we start from $x_0 = 0.5$, we find that $x_1 = x_2 = \cdots = 0.5$. This means that 0.5 is a stable fixed point solution. For now, let us not bother about the formal analysis of the stability of fixed point solutions and simple bifurcations (which can be found in Chapter 12), but instead adopt a simple simulation approach: For any allowable fixed r, we arbitrarily choose an initial value for x_0, and iterate Eq. (2.7). After throwing away sufficiently many iterations so that the solution of the map has converged to some attractor, we retain, say, 100 iterations and plot those 100 points against each r. When the map has a globally attracting fixed point solution, then after the transients die out, the recorded values of x_n all become the same. One then only observes a single point for that specific r. When the solution is periodic with period m, then one observes m distinct points for that specific r. When the motion becomes chaotic, one observes as many distinct points as one records (100 in our example). Figure 2.4(a) shows that the logistic map undergoes a period-doubling bifurcation to chaos.

In fact, Fig. 2.4(a) contains more structures than one could comprehend at a casual glance. For example, if one magnifies the small rectangular region in Fig. 2.4(a), one obtains Fig. 2.4(b). To have hands-on experience with such self-similar features, readers new to chaos theory are strongly encouraged to write a simple code to reproduce Figs. 2.4(a) and 2.4(b).

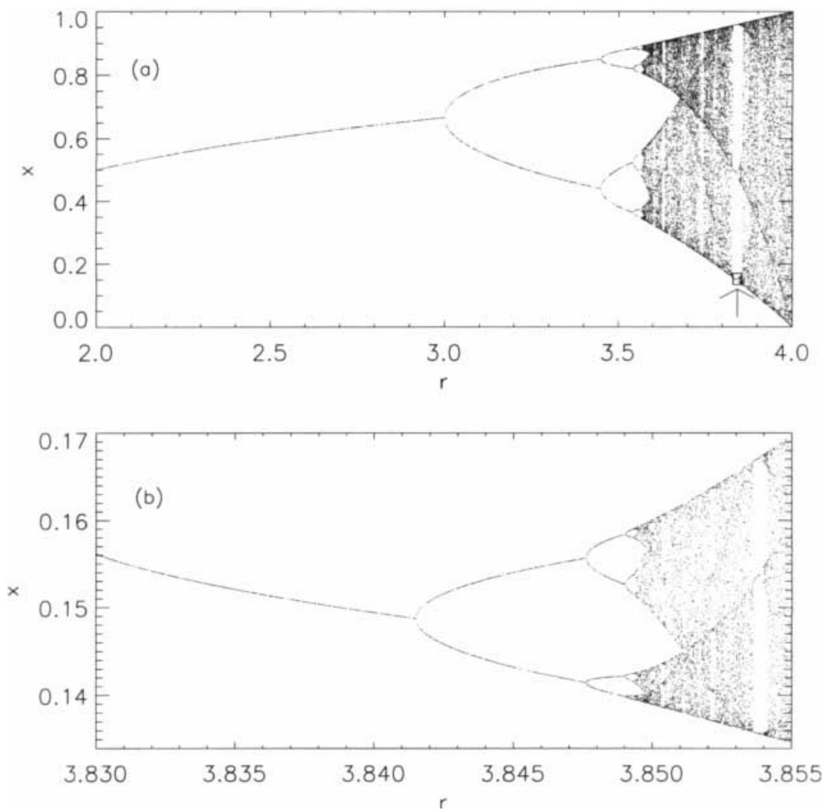

Figure 2.4. Bifurcation diagram for the logistic map; (b) is an enlargement of the little rectangular box indicated by the arrow in (a).

While one might think that the first period-2 bifurcation in Fig. 2.4(a) is too simple to be interesting, we note that the worst case operation of a noisy NAND gate can be described by a simple map, which upon transformation can be made equivalent to the logistic map and that the first period-2 bifurcation point gives the error threshold value for the noisy NAND gate to function reliably. We shall have more to say on this in Chapter 12.

Period-doubling bifurcation to chaos is one of the most famous routes to chaos identified so far. It has been observed in many different fields. An important universal quantity associated with this route is the limit of the ratio of the differences between successive bifurcation parameters,

$$\delta = \lim_{k\to\infty} \frac{r_k - r_{k-1}}{r_{k+1} - r_k} = 4.669201\cdots .$$

Other well-known routes to chaos include the quasi-periodicity route and the intermittency route. The quasi-periodicity route to chaos occurs when a critical parameter is varied, the motion becomes periodic with one basic periodicity, quasi-periodic with two or more basic periodicities, and suddenly the motion becomes chaotic. Recently, it has been found that this route underlies the complicated Internet transport dynamics. The third route to chaos, intermittency, refers to the state of a system operating between smooth and erratic modes, depending on the variation of a key parameter. This route to chaos may also be very relevant to transportations on the Internet.

2.3 BIBLIOGRAPHIC NOTES

Motivated newcomers to the field of chaos and fractal theories should find James Gleick's best-seller *Chaos* [192] and Mandelbrot's classic book *The Fractal Geometry of Nature* [294] entertaining and inspirational. Sir D'Arcy Thompson's book is [434]. For the three routes to chaos, we refer to [131, 351, 379] as well as to the comprehensive book by Edward Ott [330]. Readers interested in chaotic Internet transportations are referred to [163, 164, 363]. An interesting paper on route flap storms is [274], while [41] is on pathological network behavior. Finally, [114] is a classic book on pattern classification.

2.4 WARMUP EXERCISES

After finishing the following three problems, you will be enlightened: "Aha, they are much simpler than I thought." With this, all intimidations shall be gone.

1. Write a simple code (say, using Matlab) to generate the Mandelbrot set.

2. Reproduce the bifurcation diagram for the logistic map and explore self-similarities.

3. We now resume the discussion of the distorted face. To be concrete, let us consider the chaotic Henon map,

$$\begin{aligned} x_{n+1} &= 1 - a x_n^2 + y_n, \\ y_{n+1} &= b x_n, \end{aligned} \tag{2.8}$$

   with $a = 1.4$ and $b = 0.3$. Now sample a unit circle centered at the origin by, say, 1000 or 10,000 points, treat them as initial conditions, and iterate the Henon map. Plot out images of the circle at, say, iteration 10, 100, 1000, etc. Can you still see a "trace" of the circle when the number of iterations increases?

   You could perform the same operation on a "face" represented by five circles: one big, representing the head, and four small, representing two eyes, the nose, and the mouth.

   A densely sampled circle should soon converge to the famous Henon attractor. What do you get for the Henon attractor?

Now assume that we start from just one arbitrary initial condition and iterate the map many times. After throwing away sufficiently many points (called transient points), we record, say, 1000 or 10,000 points. Plot out those points as dots (i.e., do not connect them by lines). Do you obtain the same pattern that you got earlier? Can you explain why they are the same?