An Introduction to
the Dynamics of Unimodal Maps

Trieste

August 1988

Colin Sparrow

Department of Pure Mathematics and

Mathematical Statistics,

Mill Lane,

Cambridge, CB2 1SB,

England.

The style of these notes will not suit everyone; I have deliberately avoided doing things in the pure mathematical style, with lists of definitions and theorems you are meant to already know. There are lots of words and relatively few symbols, and I have been quite happy to skip details or whole proofs where they would take up a lot of space and prevent me getting to interesting results as quickly as I would like. The development is probably not one that would be preferred by someone hoping to introduce you to the most modern techniques in the subject, who might well prefer to get more quickly to the heart of the matter. Rather, I have tended to assume that you know almost nothing and tried to build up a picture of a subject in easy intuitive steps. Some details have been included enclosed in parentheses, **[   ]**; these can be omitted on first reading. These choices have some inevitable bad consequences; if you already know a lot you will probably be infuriated by an occasional lack of precision, and the lack of an index may make it hard to find the definition you want, though I hope the use of **bold type** will have helped. But I do hope that for some of you, this approach will make for easier understanding and will lead you to want to know more.

None of the results below are my own, and I have therefore felt free to treat the subject as an established field where it is no longer necessary to give credit to everyone who has made contributions. The books and papers listed in the bibliography are more scholarly, and contain many references.

1. Introduction

We will study the dynamics of certain maps of the interval to itself. In particular, we are interested in <u>continuous unimodal maps</u>, $f:[-1,1]\to[-1,1]$ satisfying $f(-1)=f(1)=-1$. We say a map f is unimodal (has

one mode) if it is strictly increasing on an interval $[-1,c]$ and strictly decreasing on an interval $[c,1]$ We usually call the point c the <u>critical point</u> of f. We want to understand the properties of orbits of such maps:- that is, given an initial condition $x_0 \in [-1,1]$, we generate the <u>orbit</u> of x_0

$$\mathcal{O}(x_0) = \{x_0, x_1, x_2, x_3, x_4, x_5, \ldots ; \quad x_{i+1} = f(x_i), i \geq 0\}$$

by repeatedly iterating the map f. Fig. 1 below shows examples of such orbits for particular choices of map f. Notice that it is easy to generate the orbits geometrically. Given a choice of x_0 we find $f(x_0)$ by moving
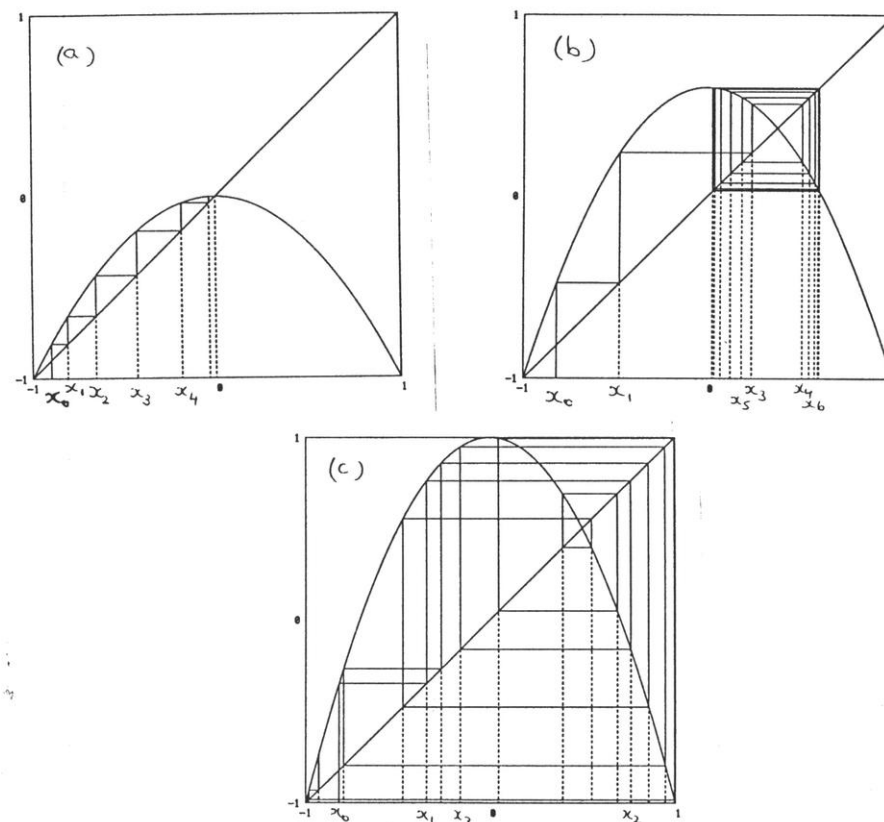

Fig.1   Orbits of the unimodal map $f(x) = \mu - 1 - \mu x^2$ for

(a) $\mu=1.0$,   (b) $\mu=1.6$,   (c) $\mu=2.0$

vertically up to the graph of f. We then move horizontally across to the diagonal to obtain an x-value of $x_1 = f(x_0)$, and then by moving vertically to the graph of f again we find $f(x_1) = x_2$. We can then proceed, moving alternately horizontally to the diagonal and vertically to the graph of f, to generate the sequence of points x_i which makes up the orbit $\mathcal{O}(x_0)$. In case (a) the orbit looks fairly simple; the points x_i are tending towards a fixed point $x^* = f(x^*)$ where the graph of f intersects the diagonal. In case (b), it seems that the orbit is tending towards a periodic orbit consisting of two points y_1 and y_2 such that $f(y_1) = y_2$ and $f(y_2) = y_1$. In case (c), the behaviour looks more complicated and it is not clear that the sequence of points in the orbit is settling down to any kind of regular behaviour at all. In no case are we particularly interested in the exact sequence of values x_i, but we will be interested to see how much we can understand about the topological properties of the orbits (such as whether they tend to an attracting fixed point, or periodic orbit, or not), how these differ for different choices of initial condition x_0, how they change as the map f changes, and other related questions. But first there are a few more introductory remarks to make which will set our study into a historical and mathematical context.

(i) Historical

It was an interest in unimodal maps in the 1970's that was responsible for kindling much of the modern interest in Nonlinear Dynamics, particularly amongst applied mathematicians, physicists, and other scientists. Unimodal maps are simple and yet provide examples of dynamical behaviour typical of much more complicated systems; for example, they can, as in case (c) above (Fig. 1), behave "chaotically" (we will return to this later). This, and other interesting forms of behaviour, can easily be observed by anyone with a computer, however small. In fact, many interesting results about these maps, and many new ideas in Dynamical Systems more generally, were motivated by and only obtained after extensive numerical simulations on maps like ours. Thousands of papers have been published on unimodal maps in the last fifteen years, and more are still published every year.

(ii) Completeness and robustness of the results

For the pure mathematician (and for the many others who have become pure mathematicians whilst studying unimodal maps), unimodal maps present a collection of problems ranging from the easy to the very sophisticated. A few of the latter remain unresolved, but an almost complete understanding is now available. Perhaps the most remarkable fact is that so much of the understanding applies equally to almost any unimodal map (or family of such maps) and it is not usually necessary to study many separate cases. This is true of both topological and metric properties, and the discovery of certain 'universal' constants governing some of the behaviour of these maps has inspired a completely new approach (renormalisation theory, see section 3 below) to the dynamics of many systems.

(iii) Relevance to other systems

The one-dimensionality of unimodal maps is crucial to our understanding of them. Nonetheless, various "almost" one-dimensional systems have very similar behaviour. An example is dissipative

diffeomorphisms such as the Henon map, or diffeomorphisms arising as return maps in dissipative ordinary differential equations. Mathematicians understand that there are essential differences between the one-dimensional case and the others, but there are some strong similarities when the dissipation is very great, and an understanding of the one-dimensional case seems to be a prerequisite for understanding many of the types of behaviour occurring in higher-dimensional systems.

(iv) Applications

The maps are almost too simple to be taken seriously as models for any complicated real-world system. Nonetheless, as a first approximation they can be useful, and they have been cited in the literature of many subjects to show that simple deterministic models can produce random-looking or chaotic behaviour without the inclusion of random terms representing the unknown influences of effects external to the model. One of the papers responsible for exciting the interest of the scientific community at large in unimodal maps, gave as an example the population of fruit flies in a cage with a constant food supply. If we write the population of flies on day t as N_t then, as a first approximation (ignoring the fact that N_t must be integer) we can write,

$$N_{t+1} = f(N_t)$$

where f is a unimodal map; if N_t is small the food supply is adequate and the population increases, but if N_t becomes too large competition forces the population to decrease.

2. The quadratic family

We will begin by trying to understand the simpler features of behaviour which occur in a particular family of unimodal maps which depends on one parameter. The family we will study is the quadratic family (or logistic family), $f_\mu(x) = \mu-1-\mu x^2$, $0<\mu\leq2$. If μ is chosen so that $\mu \in (0,2]$, each f_μ is a unimodal map $f_\mu:[-1,1]\to[-1,1]$ with a critical point at 0. Fig. 1 showed examples of this family for three different μ-values (see figure caption). (This family is sometimes written, after a simple linear change of variables, as $g_r(x) = rx(1-x)$, $g:[0,1]\to[0,1]$, $r \in (0,4]$.)

We will be interested in the dynamics of f_μ for each μ in the interval (0,2], and also in the way in which our results change as μ changes. Definitions will be introduced as we go along.

Preliminary definitions

x is a fixed point of f_μ if $f_\mu(x)=x$. In Fig. 1(a) the point x=0 is a fixed point of $f_{1.0}$. A set of p points $\{x_1, x_2, ..., x_p\}$ is a periodic orbit of f_μ if $x_{i+1}=f_\mu(x_i)$ for i=1,2,....,p and $x_1=f_\mu(x_p)$. The orbit is said to be of least period p if p is the smallest integer such that the orbit is a periodic orbit of period p. Fig. 1(b) showed an orbit tending towards a periodic orbit of least period 2. When there is no ambiguity we sometimes write just f instead of f_μ. We write the nth iterate of f as f^n, so $x_n = f^n(x_1)$ = f(f (f (......f(x_1)))). A point x on a periodic orbit of period p is clearly a fixed point of f^p.

We will say that a fixed point x^* is an attractor if there is an

interval U containing x^* such that $y \in U \Rightarrow f^n(y) \to x^*$. If x^* is in the interior of U we say that x^* is a __two-sided attractor__; otherwise x^* is a __one-sided attractor__. The fixed point at $x^*=0$ in Fig. 1(a) is a (two-sided) attractor; in fact, in that map we can choose $U=(-1,1)$ and all orbits tend towards x^*. It is clear that a fixed point is a (two-sided) attractor if $|f'(x^*)|<1$ where $f'(x)$ is the derivative df/dx of f with respect to x. In this case we say that f is a __hyperbolic__ attractor, and maps C^1 close to f also have hyperbolic attracting fixed points. If $|f'(x^*)|>1$ we say that x^* is a hyperbolic __repellor__. In the unusual cases where $f'(x^*)=\pm1$, the fixed point may attract orbits from one side and repel them from the other, and the details of the behaviour depend on the second or higher derivatives; note, however, that we have chosen our definitions so that if orbits are attracted from one side and repelled from the other then the fixed point x^* is called a (one-sided) attractor. (If these results are not obvious to you, try drawing an orbit near to a fixed point where f has slopes of absolute magnitude greater than or less than 1. Or, more mathematically, expand f to first order in a Taylor series about x^* and compare $|f(x^*+\delta) - x^*|$ with $|\delta|$.)

A periodic orbit $\{x_1, x_2, ..., x_p\}$ is said to be an attractor if x_1 is an attractor for f^p. The orbit of period 2 in Fig. 1(b) is an attractor, and attracts all nearby orbits. The conditions on the derivative of f^p at x_1,

$|df^p/dx| \leq 1$, can also be written, using the chain rule, as

$|f'(x_1)f'(x_2).......f'(x_p)| \leq 1$. (Proof; exercise.)   Notice that if 0 is a point on the orbit, $x_i = 0$ for some $1 \leq i \leq p$, then $f'(0)=0$ implies the orbit is an attractor. In such cases we say the orbit is __superstable__.
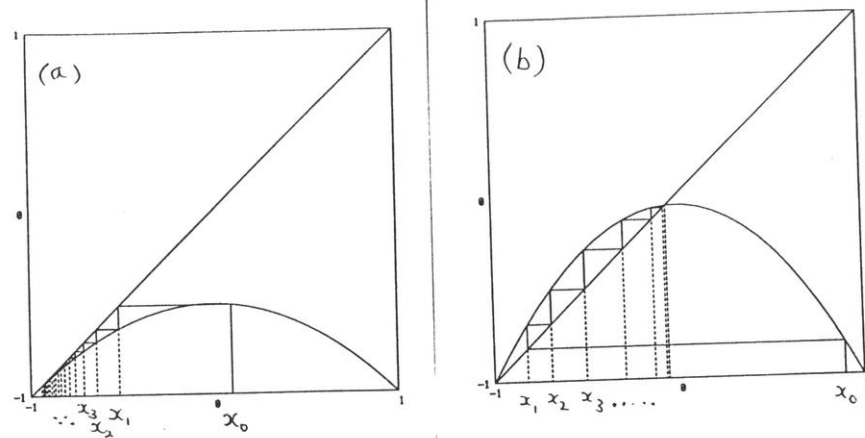
Fig.2 Orbits for (a) $\mu=0.49$; -1 is an attracting fixed point,

(b) $\mu=0.95$, -1 is repelling and x_μ^* is attracting.

(i) $0<\mu\leq0.5$

In this parameter range all orbits tend towards the attracting fixed point at $x=0$. This is easy to see geometrically – see Fig. 2(a) – and easy to prove. (Proof; exercise.)

(ii) $0.5<\mu\leq1$

When $\mu=0.5$, a __bifurcation__ occurs and a new fixed point appears in the interval $[-1,1]$ at $x_\mu^* = (\mu-1)/\mu$. We say that a bifurcation occurs at a particular μ-value if the topological nature of the orbits changes at that μ-value. For all $\mu>0.5$ the fixed point at $x=-1$ is a repellor. The fixed point x_μ^*, however, is an attractor in this parameter range. See Fig. 2(b). When $\mu=1$ the fixed point is at $x=0$ and is superstable, as shown in Fig. 1(a). It is easy to prove that all orbits other than those starting at $x_0=-1$ or 1 tend towards the attracting fixed point in this parameter range. (Proof; exercise.) It will be useful later if we notice now that the behaviour we

have just described for $\mu=1$ implies that the graph of $f_1^n(x)$ looks like Fig. 3 for large n; as n increases, the plateau looks increasingly flat since for any $x \neq -1, 1$, $f_1^n(x) \to 0$ as $n \to \infty$.
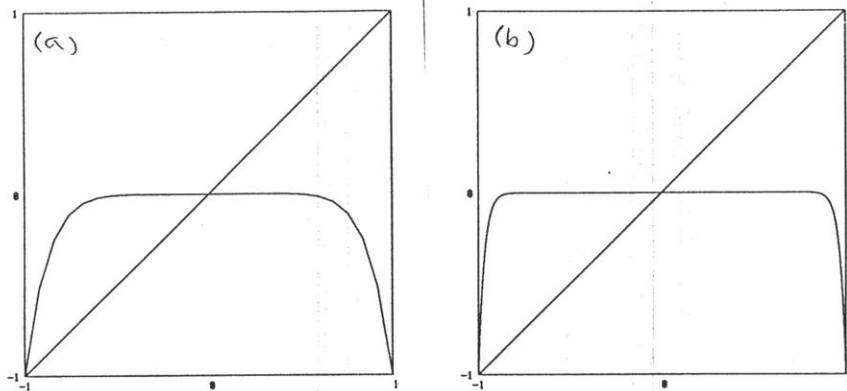

Fig. 3 $f^n(x)$, $\mu=1$ for (a) n=3,   (b) n=5

(iii) $1 < \mu < 1.5$

In this parameter range the fixed point $x_\mu{}^*$ is still an attractor and all orbits except those started at -1 or 1 are still attracted towards it, but it now lies in $x > 0$, the derivative of f there is negative, and orbits approach it in an oscillatory fashion, successive points, x_i, x_{i+1} lying on opposite sides of $x_\mu{}^*$. See Fig. 4(a). It is relatively easy to prove that this description of the behaviour in this parameter range is correct. (Proof; exercise.)

(iv) $\mu = 1.5$

At $\mu = 1.5$, the derivative $f'(x_\mu{}^*)$ decreases through -1 and there is a <u>period-doubling bifurcation</u>. In this bifurcation the fixed point $x_\mu{}^*$
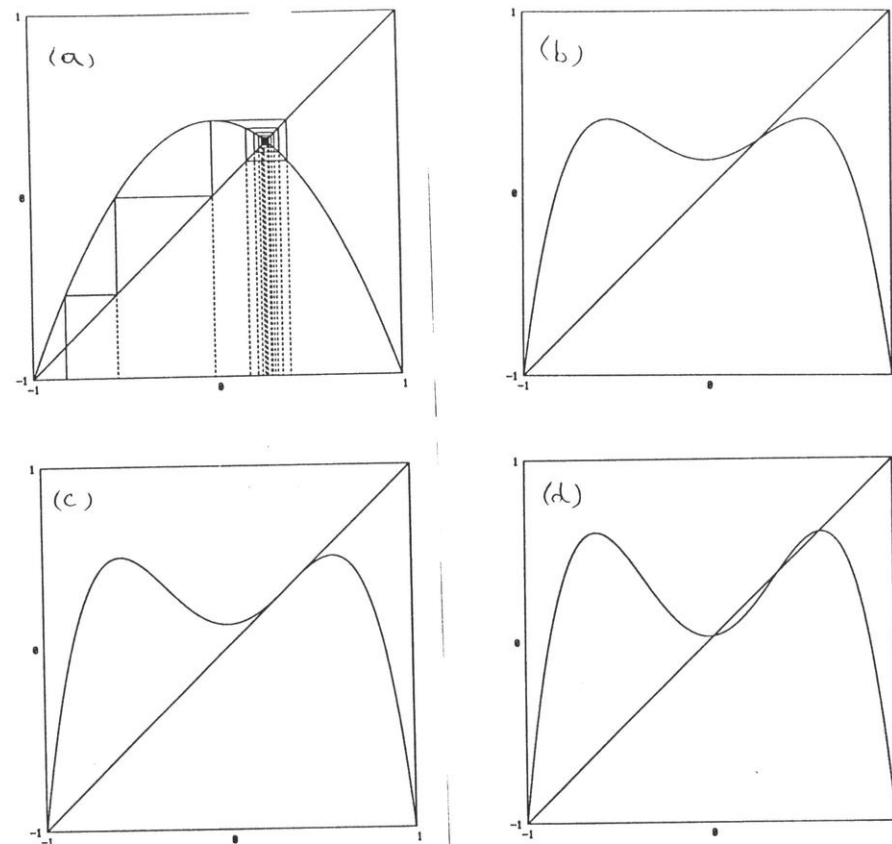
Fig. 4.   (a) For $\mu=1.4$ the fixed point is attracting, and (b) f^2 has a single fixed point. At $\mu=1.5$, f' decreases through -1 at the fixed point, and (c) the slope of f^2 increases through 1. For $\mu=1.6$, (d), f^2 has two new attracting fixed points. The behaviour of f at this parameter value was shown in Fig. 1(b).

becomes a repellor, and an attracting periodic orbit of period 2 appears for $\mu > 1.5$. The behaviour for μ just greater than 1.5 ($\mu=1.6$) was shown in Fig. 1(b). To see why this period 2 orbit appears we look at f^2. Figs. 4(b), (c) and (d) show f^2 for $\mu < 1.5$, $\mu = 1.5$ and $\mu > 1.5$. We can see that two new fixed points of f^2 appear as μ increases; since they are not also fixed

points of f, they must lie on a periodic orbit of period 2. Figs. 4(d) and 1(b) show f^2 and f at the same parameter value (μ=1.6).

**[Bifurcations involving fixed points occur at parameter values where $|f'(x^*)|$=1, i.e. where the fixed point is <u>non-hyperbolic</u>. We have already seen one such bifurcation, at μ=0.5 and x=-1, but that we did not consider it fully because to understand completely what happens at that bifurcation (commonly called a <u>transcritical bifurcation</u>) we would have needed to consider the behaviour and existence of fixed points outside the interval [-1,1]. We will, however, now consider the period-doubling bifurcation, which occurs often in our family of maps and which, in the absence of special symmetries or restrictions on the maps, is the bifurcation that occurs in any family of maps when the derivative at a fixed point decreases through -1. To be more precise, for a map $h_\lambda(x)$ satisfying:-

(1) $h_\lambda(0)$=0 for all λ (there is a fixed point at x=0)

(2) $h'_0(0) = -1$ (there is a bifurcation at λ=0)

(3) $f'''_0{}^2(0) \neq 0$ (the third derivative with respect to x of f^2 at 0 is

non-zero at λ=0; notice that the first and second derivatives automatically equal zero if (1) and (2) are satisfied - (proof; exercise))

and

(4) $\frac{d}{d\lambda} f_\lambda{}^2(x) \neq 0$ when x=0,λ=0

then:

there exists a continuous function p(x): U $\rightarrow$ V, where U and V are neighbourhoods of x=0 and λ=0 respectively, such that $\forall$ x$\in$U, $f_{p(x)}{}^2(x) = x$, $f_{p(x)}(x) \neq x$ if x$\neq$0. Furthermore, p(0)=p'(0)=0 but p''(0)$\neq$0.

Figs. 5(a) and 5(b) are two <u>bifurcation diagrams</u> showing the position of fixed points or periodic points for various values of λ, as given

by the two possibilities (p^*>0 or p^*<0) allowed by the theorem; the curve of period 2 points is the function p giving parameter λ in terms of x. Thus, in a typical map where conditions (3) and (4) are satisfied at the bifurcation, then on one side of the bifurcation, say λ<0, there is just a fixed point, whereas on the other side, say λ>0, there is a fixed point and a periodic orbit of period 2. The period 2 orbit will be an attractor if exists on the side of λ=0 for which the fixed point is a repellor (Fig. 5(a)). For the quadratic map, after a linear change of variables and a rescaling of the parameter, this theorem applies at μ=1.5, $x_\mu{}^* = {}^1/_3$. Furthermore, it is possible to show that for any map with negative <u>Schwarzian derivative</u>,

$$Sf = \left\{ {}^{f'''}/_{f'} \right\} - 1.5 \left\{ {}^{f''}/_{f'} \right\}^2 < 0,$$ only the possibility shown in Figure 5(a) (which is known as the <u>supercritical</u> case) can occur - i.e. the period 2 orbit produced in a period-doubling bifurcation in the quadratic family (for which f and f^n for n$\geq$2 all have negative Schwarzian) will always be an attractor. The proof of the period-doubling theorem can be found in standard text-books (e.g. Devaney, 1986). The Schwarzian derivative will appear again below, and the proof of the remarks about maps with negative Schwarzian are not difficult once you have a lemma from section (ix).]**
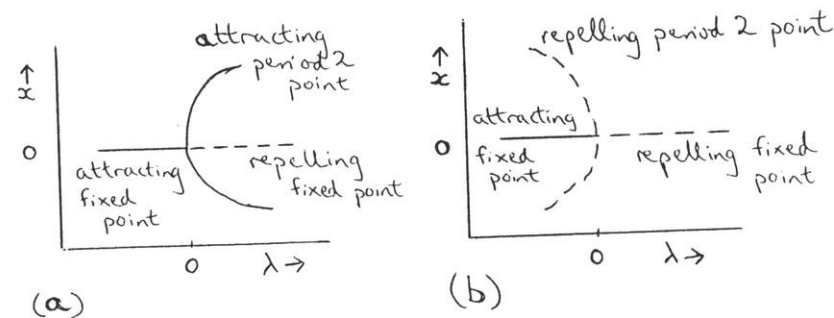

Fig.5. The position of fixed and periodic points against λ for a period-doubling bifurcation. (a) supercritical, (b) subcritical.

(v) $1.5 < \mu < \mu_\infty \approx 1.7849$

The behaviour in this parameter interval is more difficult (if not impossible) to calculate analytically and so we will just rely on the results of numerical experiments in this section. It is not expected that you should see why the behaviour we are about to describe should occur, just that you believe (or check on a computer) that it does. (We will come to more sophisticated arguments about this interval later on.)

If we increase μ above $\mu=1.5$, the period 2 orbit is an attractor until $\mu \approx 1.725$ when it becomes a repellor and an attracting period 4 orbit appears. This is another period-doubling bifurcation, but this time it is a fixed point of f^2 which period-doubles as the slope (of f^2) decreases through -1 as μ increases. Figs. 6(a)-(c) show f, f^2 and f^4 at $\mu=1.75$, just above the bifurcation value. Each of the diagrams in Fig. 6 shows the behaviour of different iterates of the map at the same parameter value; if a point is on a period 4 orbit of f, it is on a period 2 orbit of f^2 and is a fixed point of f^4. Notice that the regions in each of the dotted boxes in Figs. 6(b) and (c) look just like scaled down versions of f and f^2 just after the first period-doubling at $\mu=1.5$ (cf Figs. 1(b) and 4(d)), though in the box which includes $x=0$ everything is upside down; we will return to this point later.

Numerical experiments show that as μ increases further, more period-doubling bifurcations occur. At $\mu \approx 1.775$ a period eight orbit appears (and the period 4 orbit becomes a repellor), at $\mu \approx 1.782$ a period sixteen orbit appears, and there are further bifurcations to orbits of periods 32, 64 etc. Fig. 7 shows an attracting period thirty-two orbit at $\mu=1.7845$. Careful numerical experiments indicate that the sequence continues as far as the accuracy of the computer will allow, with bifurcation values μ_i tending to $\mu_\infty \approx 1.7849..$ as $i \to \infty$, and attracting orbits of period 2^i appearing at each μ_i and becoming repellors at μ_{i+1}.
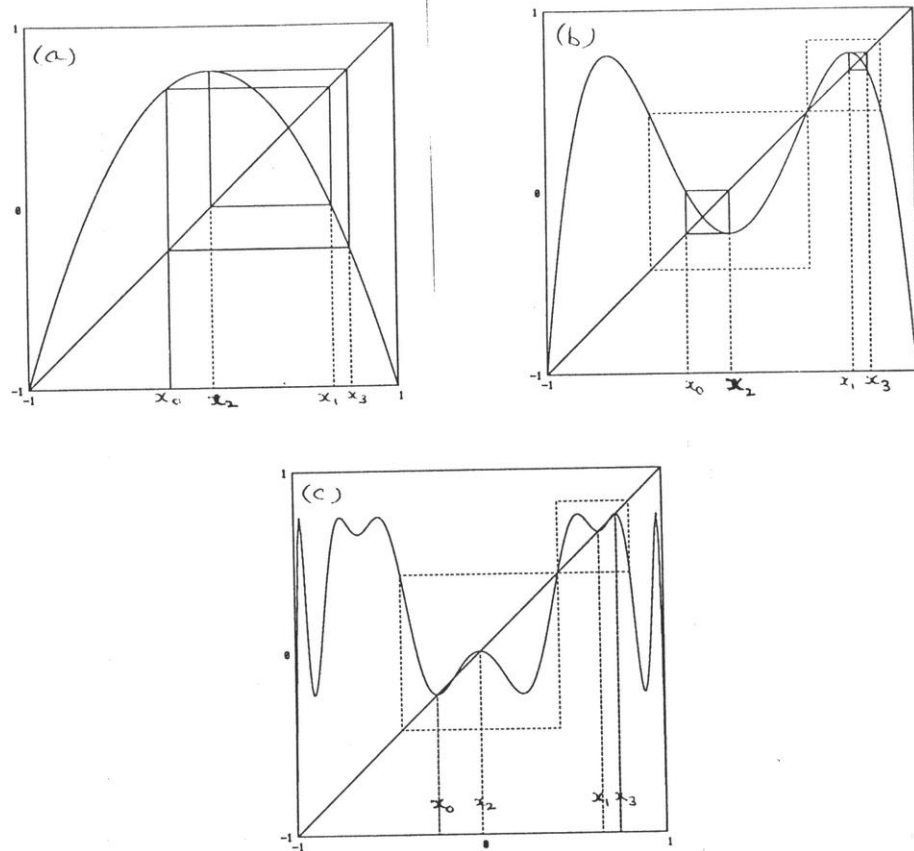
Fig.6. (a) f, (b) f^2, and (c) f^4 at $\mu=1.75$. In (a) we see an attracting period 4 orbit, in (b) the corresponding attracting period 2 orbits, and in (c) four attracting fixed points. Note that in each dotted box in (b), f^2 is a unimodal map.

Furthermore, the μ_i satisfy $(\mu_i - \mu_{i-1})/(\mu_{i+1} - \mu_i) \to 4.6692016...$ as $i \to \infty$. This last fact was only discovered in the late 70's during numerical experiments on a pocket calculator (by Feigenbaum), and, as we shall see later, is the more remarkable because the same number $\delta = 4.669..$ will appear often in our study and elsewhere in the study of dynamical systems.

We refer to the parameter value μ_∞ as the <u>accumulation of period-doubling</u>.

Fig. 8(a) shows the behaviour of an orbit at $\mu \approx \mu_\infty$; it is not easily distinguished at this resolution from Fig. 7. Fig. 8(b) shows two orbits of f^2 at the same μ-value. Once again, notice how much the two dotted boxes in Fig. 8(b) resemble Fig. 8(a).
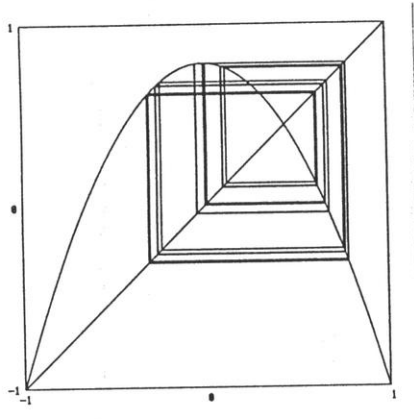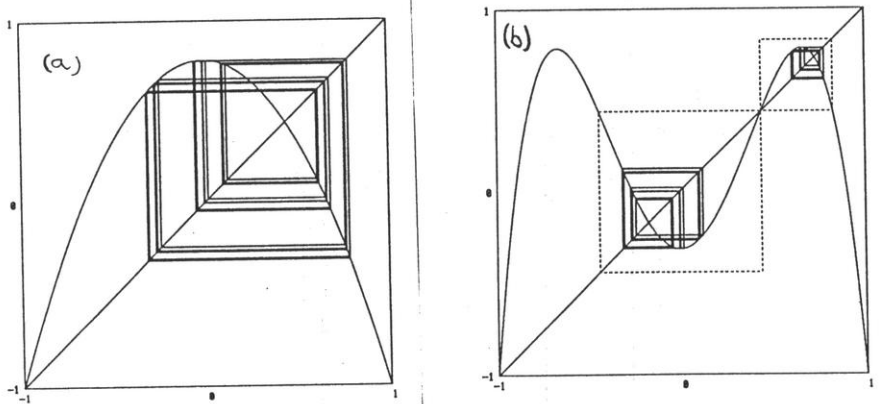

Fig.7.   An attracting period 32 orbit at $\mu=1.7845$.

Fig.8.   Orbits of (a) f and (b) f^2 at $\mu \approx \mu_\infty \approx 1.7849..$ Notice how each of the boxes in (b) resembles all of (a).

(vi) $\mu=\mu_\infty \approx 1.7849..$

The behaviour at $\mu = \mu_\infty$ is a little harder to describe and to understand than anything we have met so far. There certainly does seem to be an attractor of some sort since for any initial condition x_0 we choose in $(-1,1)$, after allowing a few iterations for transient behaviour to die out, we find dynamics like that shown in Fig. 8(a). Some authors try to describe this attractor as being a periodic orbit of period 2^∞, but this is rather misleading. We will attempt a more precise description.

If we look closely at the dynamics at $\mu=\mu_\infty$, we see that for any $n \in \mathbb{Z}^+$ we can find 2^n closed intervals, $I^{(n)}_1$, $I^{(n)}_2$, ..., $I^{(n)}_{2^n}$ such that f maps points cyclically from one interval to the next, so $f(I^{(n)}_i)=f(I^{(n)}_{i+1})$ and $f(I^{(n)}_{2^n})=I^{(n)}_1$ as shown in Fig. 9. The repelling periodic orbits of periods 2^{n-1}, 2^{n-2}, ..., 2, 1 (each of which was attracting at some lower μ-value, and which continue to exist according to the period-doubling theorem despite becoming repelling in period-doubling bifurcations as μ increased to μ_∞) sit in the gaps between these intervals, and so any orbit not actually on one of these periodic orbits is eventually pushed into one of the intervals $I^{(n)}_i$ and then continues to cycle round the intervals $I^{(n)}$ forever. The 2^{n+1} intervals $I^{(n+1)}$ and the repelling orbit of period 2^n are contained inside the intervals $I^{(n)}$, as illustrated in Fig. 9, so we can consider the set $I^\infty = \{x: x \in I^{(n)} \ \forall \ n \in \mathbb{Z}^+\}$. This will be closed and non-empty (since it is the intersection of a nested sequence of closed and non-empty sets), and invariant under f (by construction). In fact, it is possible (but not easy) to prove that the length of the intervals $I^{(n)}_i$ tends to zero as $n \to \infty$, so I^∞ is a Cantor set (see below). The set I^∞ is not an attractor in the sense defined for periodic orbits, since arbitarily close to any point of I^∞ we can find a point on a repelling periodic orbit (and this orbit does not tend to I^∞), but this fact

only reflects the difficuly we have in finding a satisfactory definition of an attractor. I^∞ certainly satisfies many other properties that one usually wants in an attractor. For instance, all initial conditions except those on the repelling periodic orbits have orbits which tend to I^∞, so I^∞ attracts an open dense set of initial conditions with full Lebesque measure. Also, I^∞ is transitive, meaning that it has a dense orbit, and so it needs to be considered as a single object which cannot be sensibly decomposed into separate attractors. In fact, every orbit in I^∞ is dense in I^∞. We refer to I^∞ and the dynamics of f on I^∞ as an <u>infinite register shift</u> or as an <u>adding machine</u>, for reasons explained further below. Finally, notice that two orbits in I^∞ which start close together (say in the same interval $I^{(n)}_i$ for some n and i) will stay close together (cycling round together in the intervals $I^{(n)}_j$), and this behaviour will be contrasted with sensitive dependence on initial conditions which is defined in (vii) below.
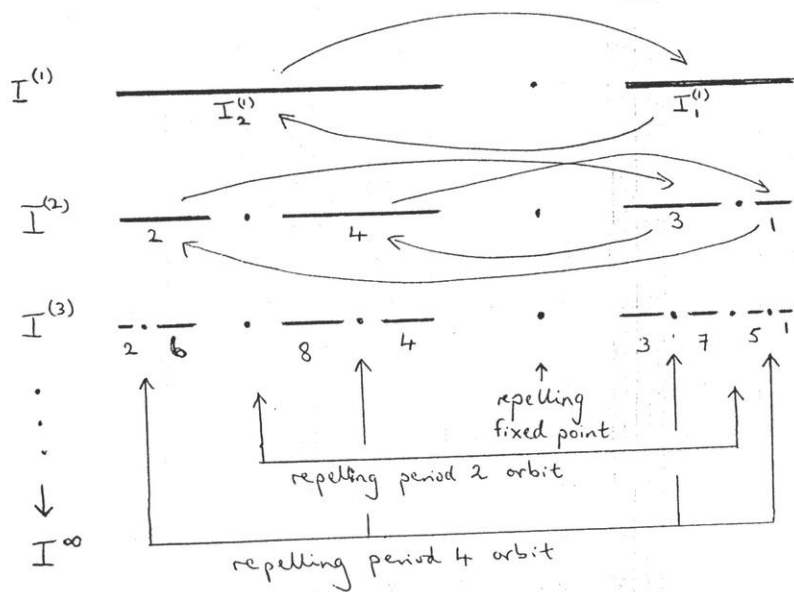

Fig.9.   Construction of the set I^∞, the attractor at $\mu=\mu_\infty$.

**[A Cantor set, C_∞, is constructed in a standard way by taking a closed interval, C_0, and then inductively defining sets C_i for $i \geq 0$ by removing one or more open intervals (without leaving any isolated points, and in such a way that the remaining closed intervals have length tending to zero as $i \to \infty$) from each of the closed intervals remaining in C_{i-1}. C_∞ is the set of points $\{x : x \in C_i \; \forall \; i \geq 0\}$. In the case of I^∞ above, we remove an open interval containing a point of the repelling orbit of period 2^i from each remaining closed interval at each stage. It is not obvious that the length of the closed intervals, which we called $I^{(n)}_i$, have length tending to zero as $n \to \infty$, but this can be proved.

As just described, the Cantor set is merely a geometric object; we also have to consider the dynamics of f on it. If we always label the rightmost interval in $I^{(n)}$ with $I^{(n)}_1$, as in Fig. 9, then it is reasonably easy to see that we can obtain a consistent description of each point in $x \in I^\infty$ by writing an infinite sequence of symbols $...s_4 s_3 s_2 s_1$ where each s_i is 0 or 1 and the binary number $s_n s_{n-1} .... s_2 s_1 = i$ tells us which interval $I^{(n)}_i$ contains x at the nth level of our construction. There is, in fact, a one-to-one correspondence between such sequences, infinite to the left, and points of I^∞. Now, given x with symbol sequence $....s_4 s_3 s_2 s_1$, f(x) will have symbol sequence $....s_4 s_3 s_2 s_1 + 1$, where addition is done in the usual way, carrying to the left as required. This explains the choice of the expressions <u>adding machine</u> or <u>infinite register shift</u> to describe the set I^∞ and its dynamics under f.]**

The behaviour described above occurs only for $\mu=\mu_\infty$. Before considering the general behaviour in $\mu>\mu_\infty$, it is advantageous to look at the particular parameter value $\mu=2$.

(vii) $\mu=2$

The behaviour of the map at $\mu=2$ is easier to understand because at this parameter value we can actually solve the map. Writing $(x_0+1)/2 = \sin^2\theta_0$, $0 \le \theta_0 \le \pi/2$ we obtain $(x_n+1)/2 = \sin^2(2^n\theta_0)$ as a solution for $x_n = f^n(x_0)$. We can immediately see that there can be no attracting periodic orbits at this parameter value, since if we consider two close together initial points x_0 and $y_0 = x_0 + \delta$ and set $(x_0+1)/2 = \sin^2\theta_0$ and $(y_0+1)/2 = \sin^2\varphi_0$ (so that θ_0 and φ_0 are close, $\varphi_0 = \theta_0 + \epsilon$) we obtain $(x_n+1)/2 = \sin^2(2^n\theta_0)$ and $(y_n+1)/2 = \sin^2\{2^n(\theta_0+\epsilon)\}$. As $2^n\epsilon$ becomes comparable with $\pi/2$, x_n and y_n will get far apart in $[-1,1]$; thus orbits started close together move apart and, in particular, orbits started close to a periodic orbit do not get steadily closer to it, and so no periodic orbit is attracting.

This behaviour, which was illustrated in Fig.1(c), is known as _sensitive dependence on initial conditions_ (s.d.o.i.c.); a term which is frequently used synonomously with the term _chaotic_. The phrase is self-explanatory; the behaviour of orbits depends sensitively on the initial condition so that small changes in the initial condition cause large changes in the behaviour of the orbit some iterates later. Notice also that the perturbation grows very rapidly, doubling in size on each iteration; this exponential divergence of orbits is typical of dynamical systems with s.d.o.i.c.

We can deduce more from the formula above. For example, there are a countable infinity of repelling periodic orbits, given by those θ_0 values, $\theta_0 = \alpha\pi$, for which $2^n\alpha \mod 1$ is periodic as $n\to\infty$. (If we write α in binary,

this will be all α's with periodic binary expansions.) These periodic orbits are densely distributed through the whole interval $[-1,1]$. In addition, there will be uncountably many aperiodic orbits (corresponding to all those α's with binary expansions that are non-periodic), and, in particular, many of these are dense in the whole interval $[-1,1]$. There is clearly no sensible sense in which we want to consider any subset of the interval $[-1,1]$ as an attractor in this case; all reasonable definitions of attractor will lead us to consider the whole interval as the only attracting set.

The dynamics at this parameter value is actually very well understood; one of the more interesting aspects of the behaviour concerns probabilistic questions about the distribution of iterates on the interval. We are not going to go at all deeply into this question in these notes, but it is easy to show, (using the formulae above), that the probability density function $d(x) = 1/\pi\sqrt{\{(1+x)(1-x)\}}$ defines an invariant distribution for f. In other words, if an initial condition x_0 is chosen at random from this distribution, the first iterate $x_1 = f(x_0)$ is distributed according to the same distribution. In Fig. 10(a) we have illustrated the distribution of points on a single orbit; it clearly has a similar form to the function $d(x)$. In Fig. 10(b) we have illustrated the cumulative distribution for the same orbit, and also plotted the cumulative density function $\int_0^x d(y)\,dy = 2/\pi \arcsin\sqrt{\{(x+1)/2\}}$. The similarity is striking, and reflects the ergodic nature of the dynamics - for almost all initial conditions we expect the distribution of points on the orbit to tend towards $d(x)$ as the length of the orbit tends to infinity.

Finally, in (ii) above we looked at the behaviour of f^n for large n when $\mu=1$. It will help our arguments below if we now do the same thing for $\mu=2$. Consideration of the formula for x_n allows us to show that f^n has

the form shown in Fig. 11, with 2^n fixed points and 2^{n-1} maxima where $f^n(x)=1$ alternating with $2^{n-1}+1$ minima where $f^n(x)=-1$. (The maxima and minima are given by the 2^n+1 values of x satisfying $(x+1)/2 = \sin^2 \theta$ and $\theta = i\pi/2^{n+1}$, $i=0,1,2,...,2^n$.)
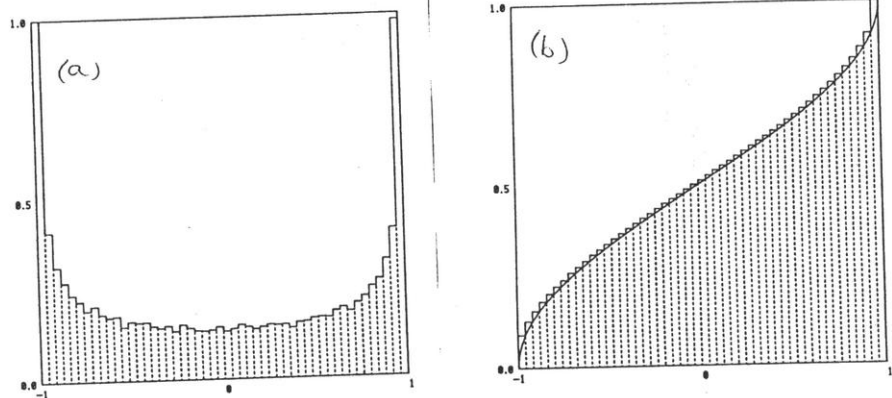

Fig. 10. (a) The distribution of 50000 points on one orbit when $\mu=2$. (b) Cumulative distribution for the same orbit and the theoretically predicticted cumulative distribution.
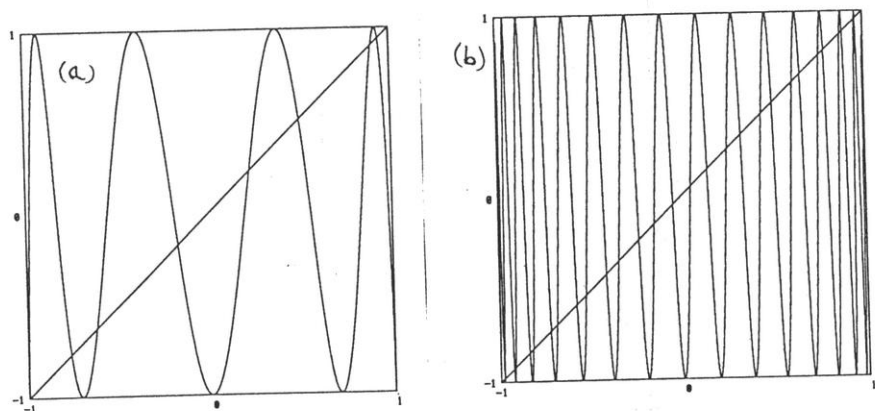

Fig. 11. (a) f^3 and (b) f^5 at $\mu=2$.

(viii) $\mu_\infty < \mu < 2$.

There are many different approaches to this interval of parameter values, the only one which we have not yet examined. The approach we will adopt here is a combination of reporting the results of numerical experiments, and making simple arguments connected with the fact that the higher iterates of f, f_μ^n, depend continuously on μ. More results will be described in subsequent sections. Let us first consider the third iterate, $f_\mu^3(x)$.

We already know (Fig. 3 and Fig. 11) that f^3 changes from a map with 2 fixed points when $\mu=1$ to a map with 8 fixed points when $\mu=2$. We need to consider how new fixed points appear as μ increases and f_μ^3 changes continuously with μ? Remembering that fixed points of f^3 that are not also fixed points of f must appear in multiples of three (since each must lie on a periodic orbit of period 3, and each such orbit has three points on it, each of which is a fixed point of f^3) the only way in which they can appear is shown in the transition between Figs. 12(a) and 12(b), which shows the actual transition between μ-values $\mu=1.91$ and $\mu=1.915$. Figs. 12(c) and 12(e) show a blow up of a region in these figures; Fig. 12(d) shows an intermediate diagram. We can see that at the bifurcation value (Fig. 12(d)), where there is a fixed point with slope = +1, we have a single fixed point, whereas we have either no fixed points (Fig. 12(c)), or two, one of which is an attractor and the other of which is a repellor (Fig. 12(e)), on either side of the bifurcation value. This picture is repeated at each of the three near tangencies of f^3 with the diagonal in Figs. 12(a) and (b), so a total of six new fixed points are created in the bifurcation. Such a bifurcation is known as a <u>saddle-node bifurcation</u> and is another of the common bifurcations occurring in many maps and dynamical systems.
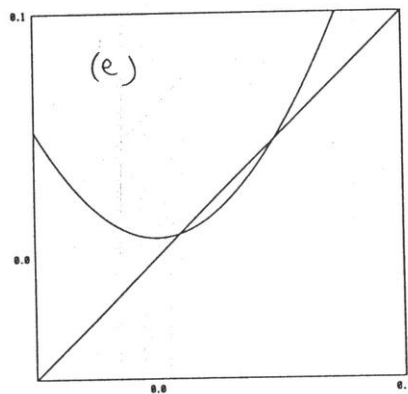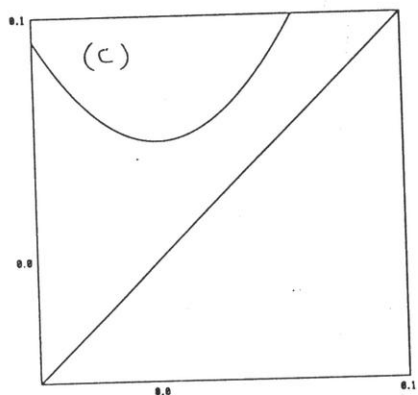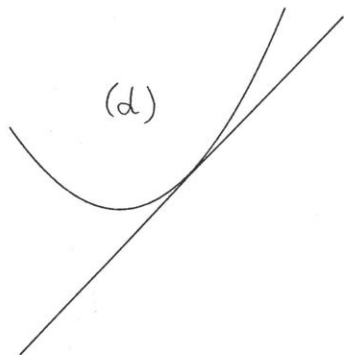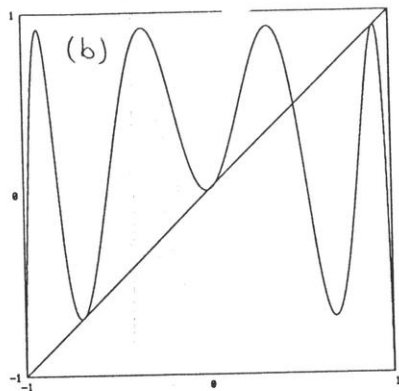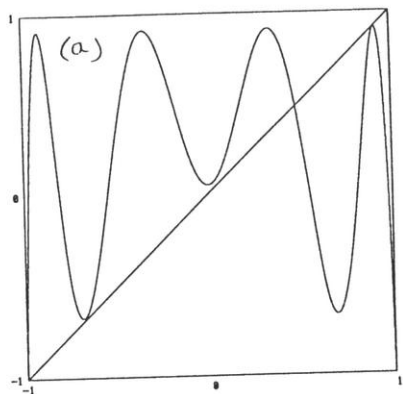
(a)

(b)

(d)

(c)

(e)

Fig.12. f^3 near a saddle-node bifurcation. (a) and (c), $\mu=1.91$;

(b) and (e), $\mu=1.915$; (d) $\mu=1.9141$.

**[Suppose that a C^2 function $f_\mu(x)$ depending smoothly on a parameter μ has a fixed point, x^*, when $\mu=0$, satisfying:-

(i) $f_0(x^*)= x^*$

(ii) $f'_0(x^*) = 1$

(iii) $f''_0(x^*) \neq 0$,

and　(iv) $\frac{d}{d\mu} f_\mu(x) \neq 0$　　　　at $\mu=0$, $x=x^*$,

then

there are neighbourhoods U of $x=x^*$ and V of $\mu=0$ and a continuous function $p:U\to V$ such that $f_{p(x)}(x)=x$ for all $x\in U$, $p(x^*)=0$, $p'(x^*)=0$, and $p''(x^*)\neq 0$.

In other words, there will be two fixed points for $\mu>0$ and none for $\mu<0$, or vice versa. This theorem can be proved by a straightforward application of the implicit function theorem to the function $g_x(\mu)=f_\mu(x)-x$.]**
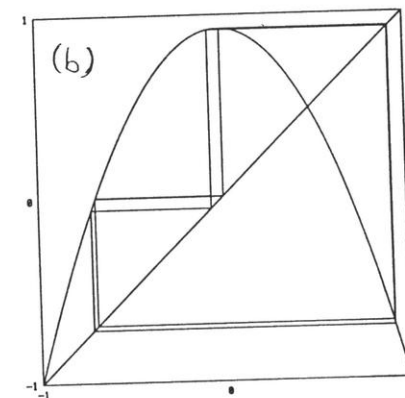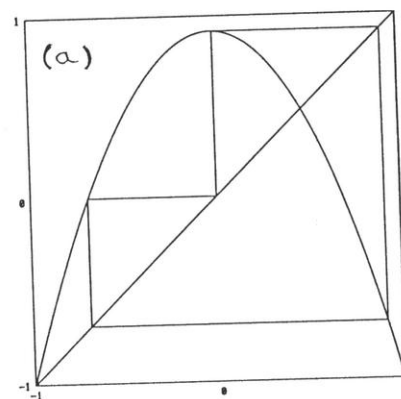

(a)

(b)

Fig. 13. (a) An attracting period 3 orbit at $\mu =1.915$;

(b) an attracting period 6 orbit at $\mu = 1.923$.

Numerical simulation of the map for the μ-value used in Figs. 12(b) and (e) does indeed show an attracting period 3 orbit. See Fig. 13(a). As μ is increased, this becomes a repellor at $\mu \approx 1.921$ and an attracting period 6 orbit appears (Fig. 13(b)). This change is a period-doubling bifurcation, which occurs as the slope of f^3 at its fixed points decreases through -1. It may come as no surprise that as μ is increased further, numerical experiments show an infinite sequence of period-doubling bifurcations in which an attracting orbit of period 3.2^i becomes repelling and an attracting orbit of period 3.2^{i+1} appears, at parameter values which accumulate at some μ-value near 1.924 and where the bifurcation values, μ_i, satisfy $(\mu_i - \mu_{i-1})/(\mu_{i+1} - \mu_i) \to 4.669....$ as $i \to \infty$. The whole range of parameter values, $1.914.. < \mu < 1.924..$, from saddle-node bifurcation to accumulation of period-doubling, is often referred to as a period-doubling window or period-doubling cascade.

A similar argument can be applied to the changes in f^n for each value of $n \geq 3$ as μ varies between 1 and 2. In each case, we will find that the number of fixed points of f^n increases as μ increases, and in each case we will be able to argue that many of these fixed points are created in (usually more than one) saddle-node bifurcations. (In the case where n is prime, every fixed point of f^n which is not a fixed point of f must lie on a periodic orbit of period n. There are 2^n-2 such fixed points when μ=2 and none when μ=1. Thus, there must be $(2^n-2)/n$ periodic orbits created, and since they are created in pairs there must be $(2^n-2)/2n$ saddle-node bifurcation values. If n is composite a more careful calculation is needed.) This shows that there are infinite number of saddle-node bifurcations occurring between the two parameter values μ=1 and μ=2. In each of these bifurcations a collection of new pairs of fixed points appears for some iterate f^n of f, each such pair consists of an attracting and a repelling fixed point of f^n and these correspond to an attracting and a

repelling periodic orbit of period n for f; numerical experiment shows that each of the attracting orbits undergoes a whole cascade of period-doublings in which stable orbits of period $n.2^i$ appear at parameter values μ_i accumulating at the expected rate. Thus we expect there to be infinitely many complete period-doubling windows in the interval under consideration. Several orbits from these windows are shown in Fig. 14.
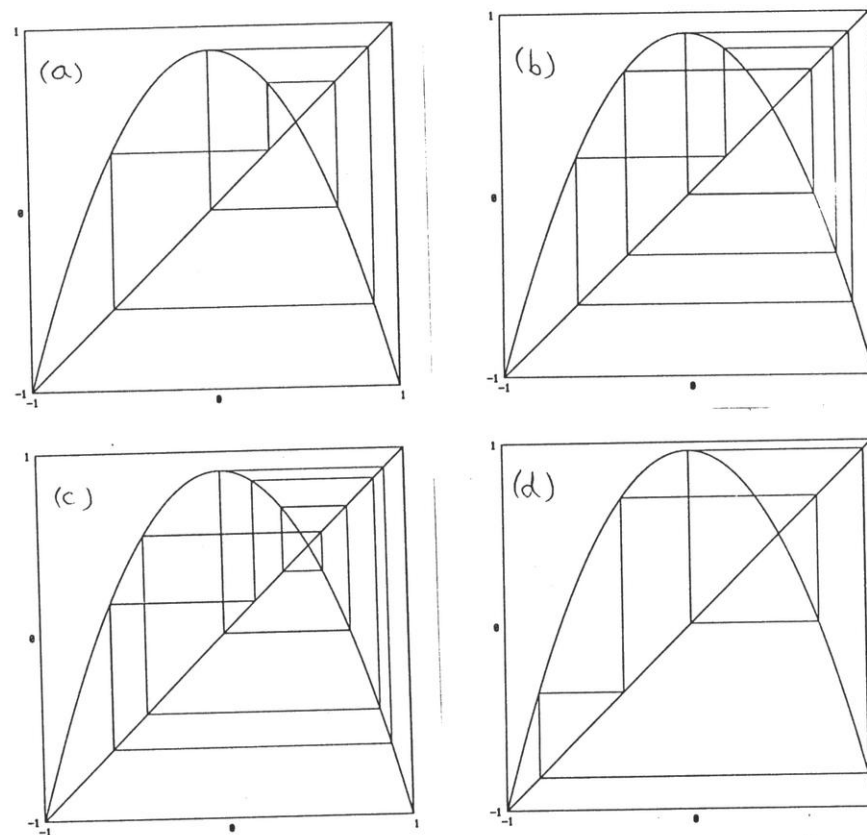

Fig. 14. Attracting periodic orbits. (a) μ=1.87, period 5; (b) μ=1.8872, pd=7; (c) μ=1.8929, pd=9; (d) μ=1.953, pd=5. Note that the orbits in (a) and (d) have differently ordered points; they are created in different saddle-node bifurcations.

It may already be difficult to imagine how all the behaviour just described can be sensibly organised; more details follow. First, however, note that there are also parameter values in $\mu > \mu_\infty$ where the behaviour of the map is similar to that which occurs when $\mu = 2$ and where we have sensitive dependence on initial conditions. Figs. 15(a)-(c) show the distribution of points on an orbit for three different parameter values; as
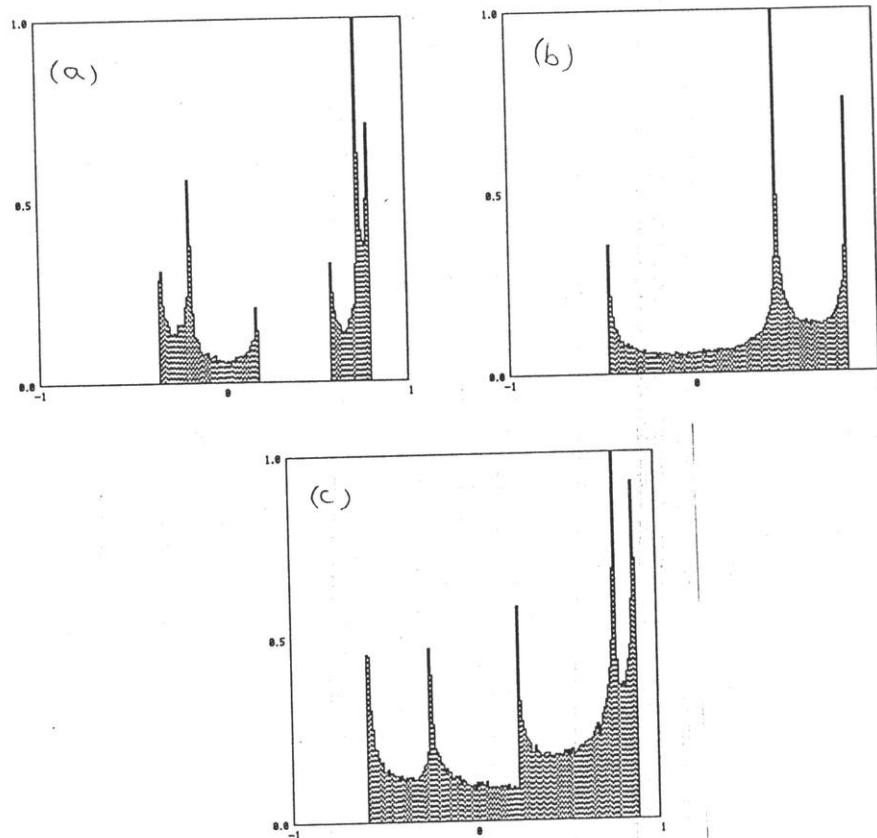

Fig. 15. Distribution of points on a single orbit: (a) $\mu = 1.7965$;
(b) $\mu = 1.839286$; (c) $\mu = 1.882475$. [cf Fig. 17.]

when $\mu = 2$ these distributions are invariant, but notice that they only have support (non-zero value) on a finite collection of sub-itervals of $[-1,1]$, rather than on the whole interval.

Fig. 16 shows a numerically computed bifurcation diagram for $1.5 < \mu < 2$. The procedure used to generate this figure is to plot the 400 points x_i, $2000 \le i \le 2400$, from the orbit of $x_0 = 0$, for 256 μ-values in $1.5 < \mu < 2$. For μ-values where there is an attracting orbit of low period, all the plotted points lie on this orbit, so only the points on the orbit appear. Otherwise, we see 400 points scattered over the 'attractor', whatever that may be. The parameter values at which the various figures in these notes were computed are marked. Notice that only relatively few of the period-doubling windows are observable, and most of these cover only an extremely short interval of μ-values; though we know there are infinitely many windows, most either involve orbits of such high period that we cannot distinguish between them and more complicated behaviour, or exist in such a short μ-interval that they are entirely missed by our procedure.
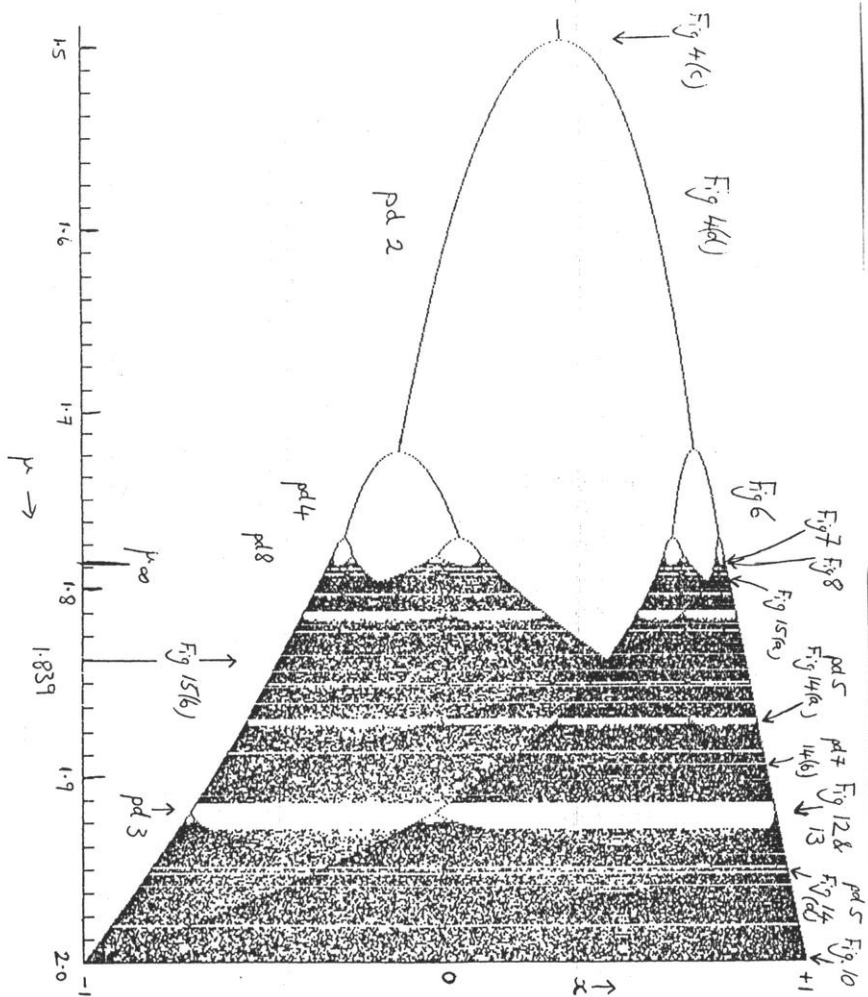
F I G. 16

Fig.16. An approximation to the attracting set for μ-values
1.5<μ<2.0. 400 points (x_{2000} to x_{2400}) on the orbit of 0 are shown.

(ix) Singer's Theorem

In the last section we argued that there would be infinitely many saddle-node bifurcations producing attracting period orbits. Each of these attracting orbits will exist for some interval of μ-values. However, our argument proceeded by considering different iterates of f separately, and there was nothing in it to suggest that two or more attracting periodic orbits (possibly of different periods) could not co-exist at the same μ-value. The task of organising the information we have generated so far will be considerably simplified if we can prove that there is at most one attracting periodic orbit at any particular μ-value. Such a result is available as a simple corollary of <u>Singer's Theorem</u> which can be stated for our purposes as:-

If f_μ has an attracting periodic orbit, the orbit of 0 is attracted to it.

The corollary then follows since the orbit of 0 cannot be attracted to two different orbits simultaneously. A proof of the theorem follows.

**[The proof of this theorem uses the fact that f has negative Schwarzian derivative, $Sf = \left\{ f'''/f' \right\} - 1.5 \left\{ f''/f' \right\}^2 < 0$. First notice that if Sf and Sg are both negative, so is S(f(g)). (Proof; exercise). In particular, $Sf^n < 0$ for all $n \geq 1$. Also, note the following lemma:-

Lemma If f satisfies Sf<0, f' has no positive local minima or negative local maxima.

Proof of lemma Suppose f' has a positive local minimum at x. Then f'(x)>0, f''(x)=0 (since we are at a turning point of f'), and f'''(x)≥0 (since the turning point is a minimum). But this implies Sf(x)≥0, contrary to assumption, so the lemma is proved.

Now, let us first show that an attracting fixed point of a map satisfying Sg<0 attracts the orbit of a critical point, c, of g (so g'(c)=0).

Let p be an attracting fixed point. Let B(p), the <u>basin of attraction of p</u>, be B(p)={y: $g^n(y) \to p$ as $n \to \infty$}. Let IB(p), the <u>immediate basin</u> of p, be the maximal connected component of the basin, B(p), containing p. Since p is attracting, IB(p) is non-empty. In general, IB(p) could be infinite, but for our quadratic map, f, and all its iterates, f^n, we have –1 and 1 $\notin$ B(p), so we consider only the case where IB(p) has left and right end-points, l and r. Continuity of g and the maximality of IB(p) imply that g(l)=l or r and g(r)=l or r. (Proof; exercise). Our aim is to show that IB(p) includes a critical point of g, and there are three cases to consider. These are (i) g(l)=g(r)=l or r; (ii) g(l)=l, g(r)=r; and (iii) g(l)=r, g(r)=l. In case (i), the Mean Value Theorem implies the existence of a critical point $c \in (l,r)$ s.t. g'(c)=0. In case (ii) the Mean Value Theorem implies the existence of a point $d \in (l,r)$ for which g'(d)=1. If g' is negative anywhere in (l,r), then the Intermediate Value Theorem implies $\exists$ $c \in (l,r)$ s.t. g'(c)=0. So we need only show that we cannot have g'(x)>0 for all $x \in (l,r)$. If p is not an end-point then g'(p)≤1 and g'(x)<1 for some x near p (since p is an attractor) and since l and r are repelling (at least on the appropriate side), g'(l) and g'(r)≥1. This implies the existence of a local minimum of g', which, by the lemma, must be negative, so we cannot have g'(x)>0 for all x. If p is at an end-point, g'(p)=1 and g'(x)<1 for some $x \in (l,r)$ near p. (Otherwise p would not be attracting.) We still have g'≥1 at the other end-point. Again, this implies the existence of a local minimum of g', as in the case just considered. Thus, in case (ii), we cannot have g'(x)>0 $\forall$ $x \in (l,r)$. In case (iii), we consider the second iterate of g, $h=g^2$, h is in case (ii) above, and so (l,r) contains a critical point c of h. The definition of IB(p) means that $c \in (l,r) \Rightarrow g(c) \in (l,r)$, and since h'(c)=0 $\Rightarrow$ g'(c)g'(g(c))=0, either c or g(c) is a critical point of g in (l,r). Thus, in all three cases, there is a critical point, c, of g in (l,r), and the orbit of c is attracted to p.

Now, to deal with the case of an attracting periodic orbit of period n

in one of our maps f, we consider the map $g=f^n$. We pick one point x_1 on the orbit, and the argument above shows that the IB(x_1) under $g=f^n$ contains a critical point, c, of f^n. Using the chain rule, $df^n/dx(c)=0 \Rightarrow$ $f'(c)f'(f(c))f'(f^2(c))...f'(f^{n-1}(c))=0$, so, for some i , $f'(f^i(c))=0$, and so $f^i(c)=0$, the critical point of f. Clearly, the orbit of 0 is attracted to the periodic orbit, since some iterate of 0 falls into IB(x_1).]**

Our use of the orbit of 0 in the statement of Singer's Theorem, is actually a very weak one, though the result is already quite useful. The study of <u>kneading theory</u>, or <u>kneading sequences</u>, takes these and similar ideas much further, describing the dynamics more-or-less completely in terms of an infinite sequence of symbols which tells us whether successive iterates of 0 lie to the left, on, or to the right of 0. We may have time to mention this theory briefly in the lectures, but it is outside the scope of these notes. It is useful for us, however, to note the following theorem, which can be proved using kneading theory. Further details can be found in the articles listed in the bibliography, particularly Van Strien (1988) and the references therein.

<u>Theorem</u>

For each parameter value μ, 0<μ≤2 either;

(a) there is an attracting periodic orbit and the orbit of 0 is attracted to it. In this case, an open dense set of initial conditions also have orbits attracted to the periodic orbit. Or,

(b) there is an attracting Cantor set (infinite register shift) which includes the critical point and its orbit, and which attracts an open dense set of initial conditions, or

(c) the map has sensitive dependence on initial conditions, and an open dense set of initial conditions has orbits attracted towards an

invariant set consisting of a finite union of closed intervals on which
there is an invariant probability distribution.

This theorem includes the results of many workers and goes some way beyond anything we have attempted to prove in these notes. (Even so, the statement of the theorem is fairly weak; many more details about the types of possible behaviour are known.) It is worth including, though, since it reassures us that (a) we have seen examples of each of the possible types of behaviour, and (b) each type of behaviour occurs on its own, and if it occurs it attracts an open dense set of initial conditions. The result is also very reasonable and relatively easy to accept without proof; if there are to be attracting orbits, or attracting Cantor sets, most of the contraction will occur near to the critical point, so the orbit of the critical point should give us all the information we need about attracting orbits and attracting sets. Notice also that if the orbit of the critical point is not attracted to an attracting periodic orbit or to an attracting Cantor set then we are in case (c); it is clear that there will be infinitely many parameter values where this occurs, for example where the orbit of 0 hits a repelling fixed point or repelling periodic orbit. Example are shown in Fig. 17, and the μ-values used in these figures are the same as those used in Fig. 15. Fairly trivial alterations to the theorem above make it true for much more general one-dimensional maps than we are considering, so the types of behaviour described in these notes should be considered very typical for general continuous maps.

(x) Sarkovskii's Theorem

We have seen above that we expect there to be infinitely many period-doubling cascades, occurring one at a time, as μ increases between
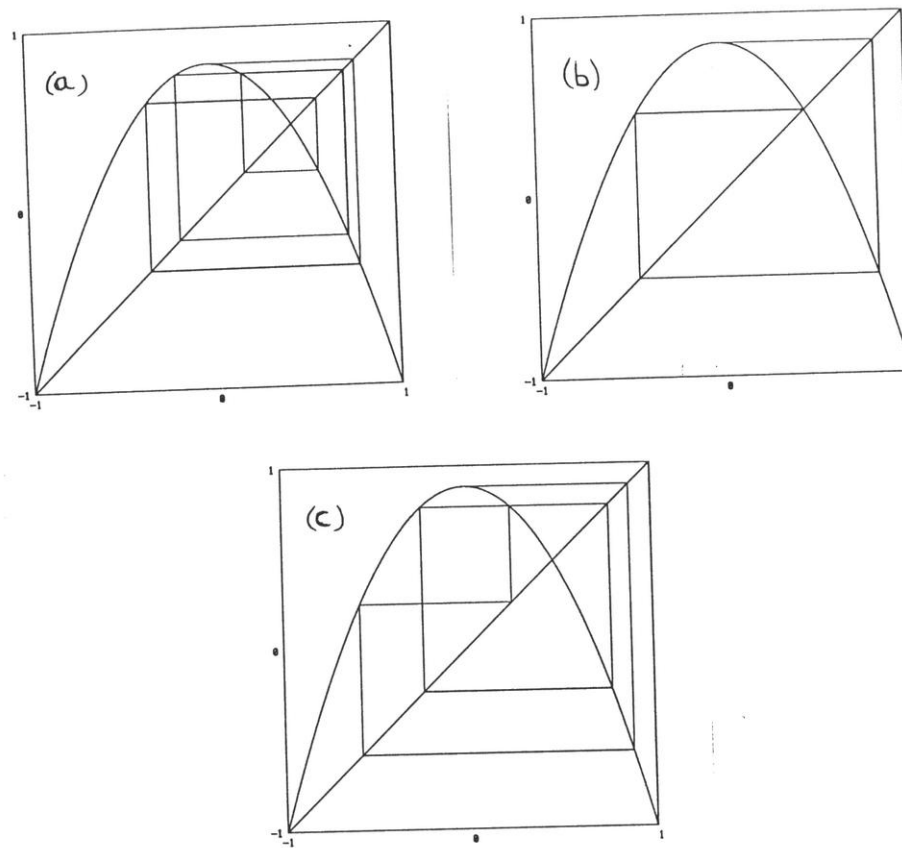
Fig.17.   The orbit of 0 for the three μ-values used in Fig.15.

μ_∞ and 2. Sarkovskii's Theorem tells us something about the order in which these occur. The Theorem itself applies to any continuous map of an interval to itself, and states:-

If a continuous map $f: \mathbb{R} \to \mathbb{R}$ has a periodic orbit of period q it also has periodic orbits of all periods p<q where the order < is given by:

$$2 < 4 < 8 < \ldots < 2^n < \ldots < 2^n.5 < 2^n.3 < \ldots < 4.5 < 4.3$$
$$< \ldots < 2.7 < 2.5 < 2.3 < \ldots < 11 < 9 < 7 < 5 < 3.$$

In particular, if there is an orbit of period 3 there are also orbits of all periods. Notice that the Theorem says nothing about the stability of orbits; we already know that at most one of those orbits existing at a particular parameter value is attracting.

The Theorem helps us to work out the order of period-doubling cascades. Suppose we consider a μ-value at which we have an attracting period 3 orbit. We know from the Theorem that there is a repelling orbit of period 7 (and all other periods) at this μ-value, and this must have been created in a saddle-node bifurcation at some lower μ-value. (Remember that the orbits involved in period-doubling cascades persist beyond the end of the cascade, though they are all repelling.)

This conclusion is certainly valid, but rather weak. In fact there are, as calculated in (viii) above, 9 saddle-node bifurcations involving period 7 orbits, and some of these occur before the period 3 cascade and some after. More sophisticated versions of the theorem are available which put all of these into order (distinguishing between the various period 7 orbits by looking at the order of the points of the orbit on the line), and the same results can also be obtained from kneading theory.

(xi)   Additional remarks

We now have a fairly complete picture of what happens as μ increases in $0 < \mu \leq 2$. Before looking, in the next section, at a technique which helps us to understand a little more about why it happens, and that the complicated sequences of behaviour are in many senses very repetitive, it is worth giving one or two more results which may help us to organise what we know so far. First, we now know that the sequence of bifurcations occurs monotonically as μ increases. In other words, each of the bifurcations we have discussed occurs only once, and, for example, we

do not have situations where two orbits are created in one saddle-node bifurcation, destroyed in a subsequent one, and then created again in a third. Second, we know that the μ-intervals in which period-doubling cascades occur are dense in $[1.5, 2]$, so any interval of μ-values includes intervals of values where there are attracting periodic orbits. Both of these results are hard to prove, and seem to require that one studies a complex version of the map $f : \mathbb{C} \to \mathbb{C}$. Despite the second result, we also know that the Lebesque measure of the set of μ-values for which there is no attracting orbit, is positive, so that the probability of randomly choosing such a μ-value is greater than zero. (This probability actually tends to 1 if we consider intervals $(2-\varepsilon, 2)$ as $\varepsilon \to 0$.) This is another hard result.

To clarify the ideas here, let us consider the question, 'What is the first thing that occurs as μ increases above μ_∞?'. This is a little like asking what is the first rational after 0, since if we consider any small interval $(\mu_\infty, \mu_\infty + \varepsilon)$, we will find:–

(i) infinitely many complete period-doubling cascades (involving periodic orbits with periods of the form $2^n.k$, where $n \to \infty$ as $\varepsilon \to 0$ – see Fig.16 where it is clear that all the behaviour cycles round 2^n intervals for $\mu > \mu_\infty$ as well as for $\mu < \mu_\infty$, though in the $>$ case the behaviour within the intervals is more complicated);

(ii) infinitely many μ-values where there is an attracting Cantor set (infinite register shift);

(iii) uncountably many parameter values where the map displays sensitive dependence on initial conditions; the probability of choosing one of these parameter values at random is greater than 0.

3. Renormalisation

Fig. 16, a condensed version of which is shown in Fig. 18(a), showed us the way in which the attracting set of f_μ developed as μ increased from 1.5 to 2. One of its more remarkable properties is self-similarity; if we look closely at the boxed region in Fig. 18(a) it looks very similar to the whole figure, except it is squashed and upside down. Fig. 18(b) shows the boxed region turned over and scaled up to be the same size as the original; it is remarkably similar to Fig. 18(a). We will be able to see why this occurs (and to argue that it should occur) in a moment, but first let us consider some of the implications of this result.

Let us call the whole of Fig. 18(a) B_1 and the boxed region B_2; we are assuming that B_2 contains a scaled, slightly distorted and upside down copy of B_1. This implies that inside B_2 there is another smaller box which we shall call B_4 which contains a scaled, upside down and slightly distorted version of B_2. B_4 is, therefore, a scaled, right way up and slightly distorted version of B_1. We can proceed with this argument, generating smaller and smaller boxes, B_8, B_{16}, B_{32}, etc., each of which is contained in the previous one and each of which contains a distorted copy of B_1, the copies appearing alternately the right and the wrong way up. The boxes will converge on μ_∞ in the μ-direction and on $x=0$ in the x-direction. Now, providing the distortion does not grow, and numerical experiments indicate that on the contrary, as we go to smaller and smaller boxes the pictures look more and more similar (after scaling up to the original size), and are actually tending towards some limiting picture B_∞, we can make various deductions that tie in with things we have already observed. For example,
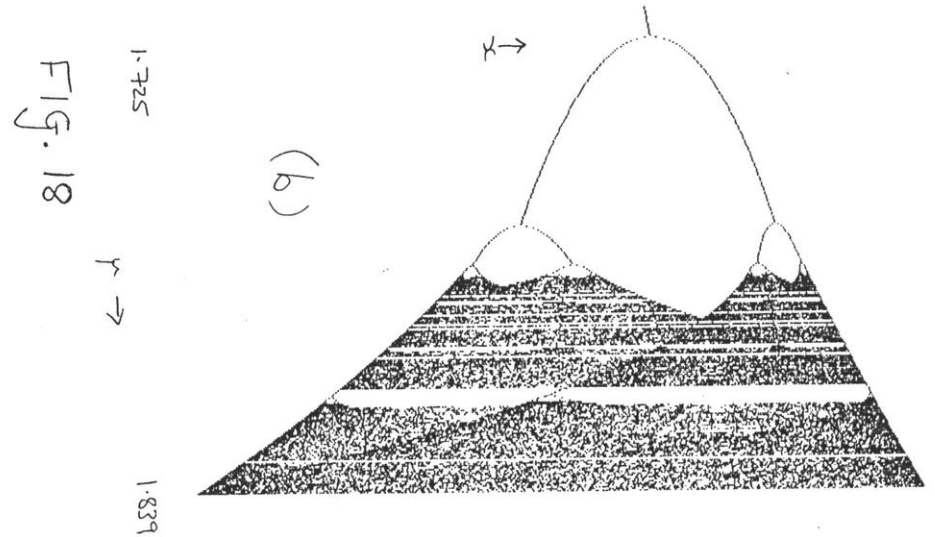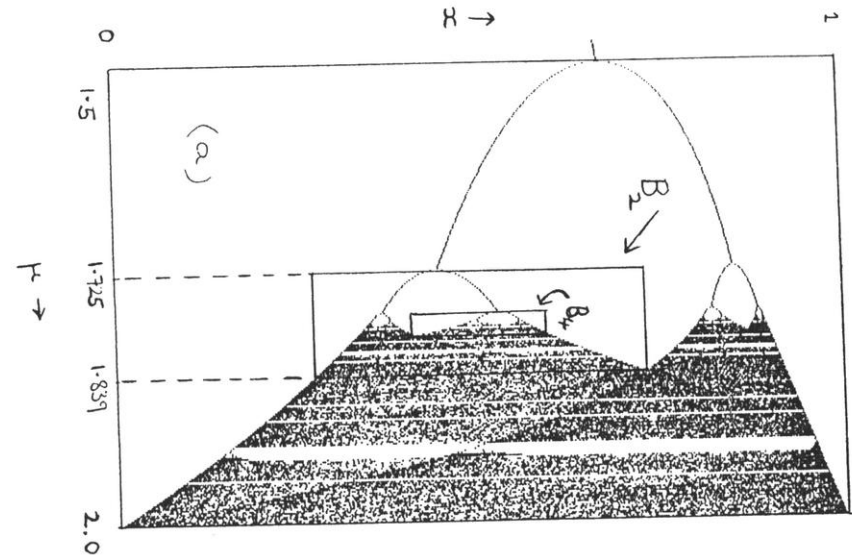
Fig.18. (a) $1.5<\mu<2.0$, as in Fig.16. (b) the box B_2 scaled up to be the same size as the original.

given that we see a period-doubling bifurcation (the left-most one) in B_1, we know there will be one in B_2, B_4, ... etc., so there will be an infinite sequence of period-doublings, and the ratio of distances between the period-doublings will tend towards some constant which is the ratio of the distances between the first two period-doubling bifurcations in the limiting picture B_∞ (and which happens to equal 4.669..). We will also expect, for example, that the sizes of the intervals used in the description of the infinite register shift (Cantor set) attractor which occurs at μ_∞, the accumulation of period-doublings, will get smaller at some limiting rate α equal to the limiting scaling necessary in the x-direction; this is also observed numerically, with $\alpha \approx 0.399..$. As an example of a different type, the existence of a period 3 window in B_1 implies the existence of period 3.2^i windows in B_i, so these windows will therefore accumulate on μ_∞ from above.

**[Notice that we can make similar arguments for sequences of boxes which do not accumulate on period-doublings. For example, in Fig. 18 we could have chosen a much smaller box around the central piece of the period 3 window as B_2; this also contains a scaled down copy of B_1. (This may be clearer in Fig. 16.) This sequence could be continued as above, with a sequence of smaller and smaller boxes converging on some point which is not at the accumulation of a sequence of period-doubling bifurcations. Once again, the picture inside the boxes converges to some limiting picture, but the rate of convergence of the parameter values involved is a different from the usual 4.669... . Also, it is possible to argue that the map will have a Cantor set attractor (infinite register shift) at the parameter value obtained by taking the limit of the left-hand (or right-hand) sides of the boxes, but that this will not be quite like the one

obtained after period-doublings. In the case illustrated, as each level of the Cantor set is constucted each interval is divided into three sub-intervals (rather than two), and other examples can be constructed with different numbers of sub-divisions on each level. We will not consider these cases any further here, but the renormalisation theory we are about to discuss can be generalised to cope with them.]**

It would be nice at this point to understand two things about the sort of arguments we have been making in this section. These are:-

1) Why should there be any similarity between B_1 and B_2?

2) Why, as we take an infinite limit of boxes within boxes, do the pictures within the boxes tend, after scaling up to the original size, towards some limiting picture B_∞, rather than getting more and more distroted?

The answer to the first of these questions is easy enough to understand intuitively. (We will not attempt a proof, though if we had developed a kneading theory the proof would not be particularly difficult.) Let us consider the development of f between $\mu=0.5$ and $\mu=2.0$ and the development of f^2 between $\mu=1.5$ and the μ-value, $\mu \approx 1.839$ at the right of box B_2. See Fig. 19. The boxed region in Figs. 19(c) and (d) look just like the whole of Figs. 19(a) and (b), except slightly distorted and upside down. Now, very few of the arguments that we have used above about the series of bifurcations occuring for f actually depended on the fact that f was the quadratic map; rather, they depended on the negative Schwarzian derivative and simple topological arguments about the development of the unimodal map as it went from a point where the fixed point at the left became repelling ($\mu=0.5$) and where the second iterate of the critical point hits the left-hand repelling point ($\mu=2$). We call such a family a _full family_. Now, f^2, restricted to the box of Figs. 19, is a full family between
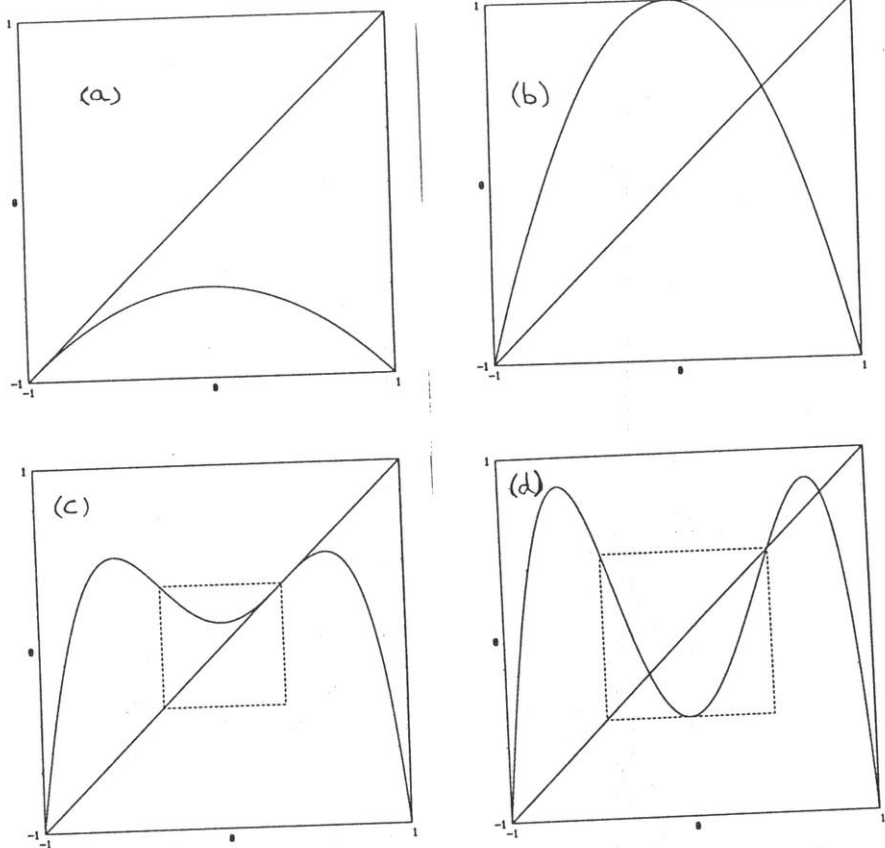
Fig.19. (a) and (b); the change in f between μ=0.5 and μ=2.0;
(c) and (d); the change in f^2 between μ=1.5 and μ=1.839.

μ=1.5 and μ≈1.839 so we expect to see the same sequence of bifurcations for f^2 in the interval 1.5<μ<1.839 that we saw for f in the interval 0.5<μ<2. In particular, we can now continue, looking at the behaviour of f^4 on a smaller interval of x-values and in a smaller μ-interval in which f^4 acts as a full family of unimodal maps, and obtain the same sequence again. This procedure parallels our looking at smaller and smaller boxes B_2i almost exactly (except that we have chosen the left-hand edge of our

boxes to be the first period-doubling bifurcation rather than the parameter value at which the left-hand fixed point becomes repelling) which explains why the pictures in the successively smaller boxes should look topologically the same.

It is much less obvious why the limiting picture does not become more and more distorted, or why the scalings (the amounts by which we have to decrease the μ-interval and x-interval under consideration at each step) tend to limits. The reason for this is the existence of a special map f^* and constant α such that f^* is a fixed point of the <u>renormalisation operator</u> $\mathcal{R}(f)(x) = \alpha^{-1} f^2(\alpha x)$, and the properties of $\mathcal{R}$ near to f^*. This can be explained, very loosely, as follows.

**[The operator $\mathcal{R}$ defines a new map $\mathcal{R}(f)$ on [-1,1] by taking the second iterate of f on a subinterval [α,-α] of [-1,1] and then rescaling by a factor α^{-1}. This is obviously related to the procedure we have been describing above, where we looked at the second iterate of f on a smaller interval, αx, and then blow the picture up by α^{-1} to look the same size as the original one. It is clear that $\mathcal{R}(f)$ has an orbit of period 2^n if and only if f has an orbit of period 2^{n+1}, so, in general, $\mathcal{R}(f)$ will have orbits of lower period than f. However, if f is a map at the accumulation of period-doubling, so that it has orbits of periods 2^n for all n≥0, $\mathcal{R}(f)$ will also have orbits of all periods 2^n. Thus, it is clear that any fixed point f^* of $\mathcal{R}$ must be at the accumulation of period-doubling. Now, in the appropriate space of functions f, it turns out (though it is very hard to prove!) that there is a unique value of α for which $\mathcal{R}$ has a fixed point f^*, which is itself unique, and that the spectrum of $\mathcal{R}$ at f^* has one eigenvalue outside the unit disc, $\delta \approx 4.669...$, and all the rest of the spectrum is inside the unit disc. If we had continued the process illustrated in Fig. 19, computing families of maps formed from higher and higher iterates of f on smaller and smaller x and μ-intervals, the map f^* would be the member of

our limiting family at the accumulation of period-doubling. This suggests that given a map at the accumulation of period-doubling, like our quadratic map at $\mu=\mu_\infty$, as we repeatedly apply $\mathscr{R}$ to it (which is equivalent to considering higher and higher iterates), we will move towards the fixed point f^* along the stable manifold of f^* under the influence of the bits of the spectrum inside the unit disc, and the x-scaling we need to use will tend towards the constant α. However, if we look at a map just off the accumulation of period-doubling, such as our quadratic f at $\mu=\mu_i$ where it only has orbits of period up to 2^i, $\mathscr{R}(f)$ will have orbits of half the period and will be further from f^* by a factor governed by the unstable bit of the spectrum, $\delta \approx 4.669...$. This explains why the distances $|\mu_i - \mu_\infty|$ increase by a factor of approximately 4.669.. as i decreases.

The description above is not entirely satisfactory; in particular, it avoids the problem that for maps other than f^* one wishes to define the renormalisation operator with an x-scaling that depends on f (to ensure that $\mathscr{R}(f)$ is a unimodal map on $[-1,1]$ mapping -1 and 1 to -1). Nonetheless, I hope it helps to explain what is becoming an increasingly important approach to various kinds of dynamical system. The most remarkable feature is the single spectral value outside the unit disc. This ensures that the stable manifold of f^* is codimension 1 in the space of functions, and given that it also seems to be quite large, most one parameter families of functions intersect it; thus period-doubling cascades are common and the limiting behaviour as the accumulation of period-doubling is reached is always very similar.]**

4. Other families of unimodal maps, and higher-dimensional systems

We have already noted, in the section above, that we expect certain families of maps to have behaviour very similar to that of the quadratic family. Indeed, in our renormalisation argument we used the argument (without proof) that the full families of maps with negative Schwarzian derivative, which arise as higher iterates of f_μ, display the same series of bifurcations as the quadratic family. In fact, almost any full family displays the same sequence and has similar behaviour. Much recent work on these maps tries to get away from the restriction to maps with negative Schwarzian (since this is considered a much too restrictive and fairly unnatural condition), and to replace it with a restriction to maps with good behaviour in a neighbourhood of the critical point (which usually just means non-zero second-derivative) and not too much distortion elsewhere (a precise definition of distortion is outside the scope of these notes). It seems that it is possible to prove that maps satisfying these weaker conditions will have similar behaviour to the quadratic family, and so we should not be surprised to see the same sequences of bifurcations, and the smae scaling behaviour near accumulations of period-doubling, whichever one-dimensional maps we study (providing we avoid ones with unusual behaviour near the critical point).

Rather more remarkably, many of the properties described above, and in particular the universal scalings near accumulation points of period-doubling cascades, seem to be observed in systems of dimension higher than one. These systems (e.g. two-dimensional diffeomorphisms or three-dimensional differential equations) are relatively badly understood in comparison with the one-dimensional maps, largely because in general there is nothing corresponding to the unique critical point, and therefore

nothing equivalent to Singer's Theorem ensuring that only one thing happens at once. (Crude numerical experiments often provide sequences of bifurcations which look very similar to those of the one-dimensional maps, but we know that we can and do get situations where, for example, there are infinitely many attracting periodic orbits co-existing at the same parameter value.) Finding higher-dimensional versions of renormalisation theory to account for the occurrence of the same universal constants in these systems will probably keep mathematicians busy for some time to come.

5. Concluding remarks

These notes (and lectures) provide only a brief introduction to the theory of unimodal maps. Anyone interested in learning more about them can find a great deal more information in some of the references listed below. Perhaps the most obvious omission in our treatment, as mentioned several times already, has been the theory of kneading sequences - this is a particularly powerful tool which is useful in the study of many different types of one-dimensional map (including those with many critical points and/or discontinuities) and more general dynamical systems as well. It is unfortunate that the theory is not at its most elegant in the study of unimodal maps, so it would have taken up rather too much of the time available to develop it carefully enough to help (rather than hinder) our understanding of these maps.

Another omission has been the study of the measure-theoretic properties of the behaviour we have studied. The omission is justified on the grounds of simplicity, but the problems need to be addressed if applications are considered important - after all, when we pick an initial condition or a parameter value at random, what we really want to know is

the probability that a certain behaviour will be observed, or that a certain behaviour is likely to occur with probablility one. This is a measure-theoretic problem - knowing that a set of parameter values is dense (say) tells us nothing about the probability of picking a member of that set in a particular experiment.

6. Brief Bibliography

Devaney, R. L. *An Introduction to Chaotic Dynamical Systems*, Benjamin/Cummings, 1986. This is an excellent introduction to the study of unimodal maps and other simple dynamical systems.

Collet, P. and Eckmann, J.-P., *Iterated maps of the interval as Dynamical Systems*, Birkhauser, Boston, 1980. A harder book containing much more information than Devaney.

van Strien, S., 'Smooth Dynamics on the Interval (with an emphasis on quadratic-like maps)', in *New Directions in Dynamical Systems*, ed. T. Bedford and J.Swift, London Mathematical Society Lecture Note Series 127, 1988. A remarkably clear exposition of much recent work but you need to have some idea of what the problems are before you begin.

The last two of these three contain extensive bibliographies (and, in their turn, refer you to other papers with even more extensive bibliographies).