PMA 324 Chaos 2009–2010

One was the face of Nature; if a face:
Rather a rude and indigested mass:
A lifeless lump, unfashion'd, and unfram'd,
Of jarring seeds; and justly Chaos nam'd.

Ovid *Metamorphoses* Book 1, Garth translation.

0 Introduction to chaos course

Warning: these notes are incomplete! They are intended to ensure that you have, after the lectures, an accurate written record of the details of proofs *et cetera*, but they do not contain either diagrams or any record of computer demonstrations, both of which are essential to the course. You will need to take notes which include these elements.

0.1 Administration

The course is two lectures per week, Wednesdays at 10 in Lecture Room 9 and Thursdays at 4 in Lecture Room F24. Set work will be given approximately once a week and solutions distributed one week later; set work will be collected and marked approximately one week in three.

0.2 Books

1. D. Gulick,"Encounters with chaos" (McGraw-Hill Education – Europe, 1992) hardback ISBN 0-07-025203-3 £65.99 (The nearest book to the course, it is also recommended for the 'Fractals' module PMA 343, but unfortunately it seems to be available for purchase only in hardback and in the library.) ML (5 copies), 531.3(G)

2. R. L. Devaney,"An introduction to chaotic dynamical systems" second edition (Westview Press, Oxford, 2003) paperback £29.99, ISBN 0-8133-4085-3, 335pp. (Covers the Chaos module only, in greater depth.) ML (5 copies), 531.3(D)

3. Saber N. Elaydi,"Discrete chaos" (CRC Press, 1999) paperback £39.99, ISBN 1584880023 (About half the book (160 pages) is relevant to this course; a further 40 pages on Fractals would be useful for the PMA443 Fractals module; the remaining 100 pages on chaos in two-dimensional systems is interesting but outside the scope of this course) ML (2 copies) 517.9 (E)

4. R. M. Crownover,"Introduction to fractals and chaos" (Jones and Bartlett, 1995) £39.99, ISBN 0867204648 (Both Chaos and Fractals, but with a less than complete coverage of the Chaos course.) ML (4 copies loan), 531.3 (C)

5. P. Cvitanovic, (ed.) "Universality in chaos" 2/e (Institute of Physics Publishing,1996) (a reprint selection). ML (1 copy) 531.3(U) ISBN 0852742606 £35.00

6. J. Gleick "Chaos" (Vintage, 1997) ISBN 0749386061 £7.99 (a popular introduction, with some defects, but great enthusiasm!) ML (2 copies), 531.3(G)

7. I. N. Stewart "Does God play dice? The new mathematics of chaos." (Penguin, 2004) ML (5 copies) 531.3(S), ISBN 0140256024 £9.99 (£6.99 + postage at Amazon)

8. Leonard Smith "Chaos: A Very Short Introduction" (Oxford U. P., 2007) ISBN 0192853783 £7.99 (£4.79 + postage at Amazon). (As the title says, a very short introduction, but a useful overview of the subject.)

9. R. L. Devaney and L. Keen "Chaos and fractals: the mathematics behind the computer graphics" (O.U.P./American Mathematical Society, Proceedings of Symposia in Applied Mathematics vol.39, 1989) ISBN 0-8218-0137-6 £21.00 (A collection of articles introducing the subject; technical in places.) ML (1 non-loan copy) 3 PER 510.5

10. W. de Melo and S. van Strein, "One-dimensional dynamics" (Springer, 1993) [advanced] ML (1 non-loan copy) 3 PER 510.5

0.3 Dynamical systems

The best way to start this lecture course would be to define exactly what we mean by "chaos". In a classical branch of mathematics such a basic definition should cause no problems. However, this is a new subject and, though there is no shortage of precise definitions, a clear consensus is yet to emerge. The same applied to the term 'fractal' which is so often bracketed with 'chaos'. The novelty of the subject is not the only reason for this. The subject spans a continuous spectrum of approaches from the pure mathematics to the pure engineering and the latter tend to have rather little patience with the pure mathematician's search for precise definitions. They can tell at a glance if a motion is chaotic or a set is fractal. Perhaps our best bet is to look at some examples.

"Chaos", whatever it is precisely, is a property of a "dynamical system", so we begin with that idea. When one thinks of dynamical systems, one thinks first of physical systems which are "dynamic". To take one example: the solar system.

We can think of the solar system's "state" at any one time as being given by the positions and velocities of each of the 9 planets relative to the centre of mass of the whole system. Thus its state is a sequence of 54 real numbers; i.e. a point in 54-dimensional space $\mathbb{R}^{54}$. The inevitable evolution of the system is a "flow" in $\mathbb{R}^{54}$; i.e. for every $t \in \mathbb{R}^+$, we have a continuous function ϕ_t of the "phase space" $\mathbb{R}^{54}$ to itself such that $\phi_t(\phi_s(x)) = \phi_{t+s}(x)$. The point $\phi_t(x) \in \mathbb{R}^{54}$ represents the state of the system at a time t after it is in state x. Such families of functions are "continuous dynamical systems". They are, in a sense the "integrals" of differential equations. Thus dynamical systems theory is the modern theory of (generally *nonlinear*) differential equations.

We may simplify the picture by "strobing" the system: looking at it only at a regular sequence of *discrete* times — say at midnight each day. Now time t is an integer. Indeed, writing ϕ for ϕ_1, we see that ϕ_t is just $\phi \circ \phi \circ \ldots \phi$ (t times), which we shall henceforth write ϕ^t. Thus the theory of "discrete dynamical systems" is just the study of iterations of continuous maps.

In the above example, the transition to discrete dynamics did not affect the phase space, but there is a more important construction where it does: the "Poincaré map".

Suppose we observe the solar system not once a day but about once a year, at the precise moments when the earth's x-coordinate is zero. (For this system, these moments are, to a good approximation, one year apart; but they are not exactly so, owing to perturbations of the earth's orbit. The difference in the principle is crucial.) We have a "return map", or "Poincaré map", $\psi : \mathbb{R}^{54} \to \mathbb{R}^{54}$ such that if the state at one observing time is x, the state at the next one will be $\psi(x)$. However, because all the observations are made when the earth's x-coordinate is zero, we may ignore this coordinate and view ψ as a map : $\mathbb{R}^{53} \to \mathbb{R}^{53}$. Thus we have replaced the continuous dynamical system by a discrete one on a space of dimension lower by one.

In part, then, the case for studying discrete systems is that continuous systems can be reduced to discrete ones (with some simplification). The other argument for preferring to study the discrete case is just that it is simpler to work with a single map ϕ rather than with a whole family of maps $\{\phi_t : t \in \mathbb{R}^+\}$. Moreover, because rapid strobing of a continuous system will produce a discrete system with similar behaviour, discrete systems show most of the interesting phenomena found in the continuous case.

Generally, interesting dynamics are associated with "folding" maps, of which the simplest is an inverted parabola. Consider the maps

$$\phi(x) := F_\mu(x) := \mu x(1 - x),$$

for values of μ in the range $1 \leqslant \mu \leqslant 4$, so a chapter of the course will be devoted to this.

Generally one may find chaos in:

1. discrete dynamical systems in $\geqslant 1$ dimension;
2. discrete dynamical systems which iterate invertible maps in $\geqslant 2$ dimensions;
3. continuous (invertible) dynamical systems in $\geqslant 3$ dimensions.

We shall investigate various aspects of dynamical systems as exemplified in the iteration of one-dimensional maps. We shall discuss fixed and periodic points, but we shall be most interested in the regions of "chaos". We shall look at some ways of formalizing the notion of "chaos", and we shall consider in detail the way that the dynamics of the quadratic family F_μ depends on the parameter μ. We shall conclude with a look at complex quadratic maps and the weird designs—Mandelbrot and Julia sets—to which they give rise.

1 Orbits and periodic points

Let $f : X \to X$ be a function on a set X. We shall write

$$f^n(x) = \underbrace{f(f(\ldots(f}_{n}(x))\ldots));$$

that is, $f^0(x) = x$ and $f^{n+1}(x) = f(f^n(x))$ $(n = 0, 1, 2, \ldots)$.

We shall use this notation f^n only for the nth iterate of f, throughout the course. (Thus $\sin^2(x)$ will, *in this course only* mean $\sin(\sin(x))$. We shall have to write its more usual meaning as $(\sin(x))^2$.)

Note that iterates obey the usual power laws:

$$\begin{aligned} f^{n+m}(x) &= f^n(f^m(x)), \\ f^{nm}(x) &= (f^n)^m(x). \end{aligned}$$

If f is bijective, we write f^{-n} for $(f^{-1})^n$. It is easy to show that the above power laws still apply.

Definition 1.1 The *(forward) orbit* of $x \in X$ is $O_f^+(x)$, or just $O(x)$,

$$O(x) = \{x, f(x), f(f(x)), \ldots\} = \{f^n(x) : n = 0, 1, 2, \ldots\}.$$

Example 1.2 Let $f : [0,1] \to [0,1] : x \mapsto x^2$. Then

$$O^+(\frac{1}{2}) = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{16}, \ldots\}.$$

Definition 1.3 We say that a point $x \in X$ is a *fixed point* of f and write $x \in \mathrm{Fix}(f)$ iff $f(x) = x$.

e.g. in Example 1.2, $\mathrm{Fix}(f) = \{0, 1\}$.

Definition 1.4 We say that a point $x \in X$ is a *periodic point* of f of period n iff $f^n(x) = x$. The least such $n > 0$ is the *order* (or *prime period*) of x. (I don't like the term "prime period", because it suggests that it is a prime number, which need not be true.)

Proposition 1.5 *The order of a periodic point divides every period.*

Proof. Let x be periodic for f with order a and let k be any period of x. Let $k = qa+r$ with $0 \leqslant r < a$. Then

$$\begin{aligned} x &= f^k(x) \\ &= f^r((f^a)^q(x)) \\ &= f^r(x), \text{ since } f^a(x) = x. \end{aligned}$$

Since a is, by definition the smallest positive integer with $f^a(x) = x$, we must have $r = 0$, i.e. a divides k. $\Diamond$

If x is periodic of order m, then

$$O(x) = \{x, f(x), f^2(x), \dots, f^{m-1}(x)\}$$

and all the points of this orbit are periodic, order m.

We write

$$\mathrm{Per}_n(f) := \{x : x \text{ is periodic of period } n\};$$
$$\mathrm{Per}(f) := \bigcup_{n=1}^{\infty} \mathrm{Per}_n(f).$$

Note that $\mathrm{Per}_1(f) = \mathrm{Fix}(f)$, and, generally, $\mathrm{Per}_n(f) = \mathrm{Fix}(f^n)$.

Example 1.6 Let $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2 - 1$. Solving a quadratic gives

$$\mathrm{Fix}(f) = \left\{ \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2} \right\}.$$

The periodic points of period 2 are roots of a quartic. They must include the fixed points and, by inspection, $\{0, -1\}$ is an orbit of order 2, so

$$\mathrm{Per}_2(f) = \left\{ \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}, 0, -1 \right\}.$$

Further analysis of the dynamics of f shows that there are no more periodic points: $\mathrm{Per}(f) = \mathrm{Per}_2(f)$.

[PICTURE]

Example 1.7 Let $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ denote the unit circle in the complex plane and let $f : z \mapsto z^2$. (Writing $z = \exp i\theta$, makes $f(z) = \exp(i2\theta)$, so we could call this the 'angle-doubling map'.) Then $f^n(z) = z^{2^n}$ so $z = \exp i\theta$ is periodic period n iff $\exp(2^n i\theta) = \exp i\theta$, i.e. iff $2^n\theta = \theta + 2k\pi$, for some $k \in \mathbb{Z}$; i.e. iff

$$\theta = 2k\pi/(2^n - 1)$$

for some $0 \leqslant k < 2^n - 1$. Thus the periodic points are the $(2^n - 1)$th roots of unity.

In this case, $\mathrm{Per}(f)$ is *dense* in S^1, i.e. there are periodic points in every nontrivial interval of S^1.

Example 1.8 Define $g : [0,1] \to [0,1]$ by

$$g(x) = \begin{cases} x + \frac{1}{4} & (0 \leqslant x < \frac{3}{4}) \\ x - \frac{3}{4} & (\frac{3}{4} \leqslant x \leqslant 1). \end{cases}$$

The points of [0,1) are all periodic of order 4 (periods 4,8,12,...). The point 1 is not periodic, but is "eventually periodic"

Definition 1.9 A point x is *eventually periodic* of period n iff there exists N such that $f^N(x)$ is periodic of period n

[PICTURE]

Note that x is eventually periodic iff there exist non-negative integers $a \neq b$ such that $f^a(x) = f^b(x)$.

If f is injective, all eventually periodic points are periodic; (easy exercise)

In Example 1.6, the point $(\sqrt{5} - 1)/2$ is eventually fixed, but not periodic. The point 1 is eventually periodic of order 2, but not periodic. It is an interesting exercise to find (roughly) where the other eventually periodic points lie. There are infinitely many of them.

In Example 1.7, $f^n(\exp i\theta) = 1$ iff $\theta = 2k\pi/2^n$ for some integer k, so the set of all eventually fixed points is

$$\{\exp(ki\pi/2^{n-1}) : n \in \mathbb{Z}, 0 \leqslant k < 2^n\}.$$

There are likewise many more eventually periodic points.

2 Analysis in $\mathbb{R}^n$.

We recall that the *distance* $d(x, y)$ between two points $x = (x_1, \ldots, x_n)$ and $x = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ is defined in the usual way by

$$d(x, y) = \sqrt{\sum_{i-1}^{n} |x_i - y_i|^2}.$$

We say that a sequence $(x_k)_{k=1}^{\infty}$ in $\mathbb{R}^n$ converges to a point $x \in \mathbb{R}^n$ if and only if $d(x_k, x) \to 0$ as $k \to \infty$; i.e. for all $\varepsilon > 0$ there is a positive integer K such that for all $k \geqslant K$ we have $d(x_k, x) < \varepsilon$.

We write

$$B(x; \varepsilon) = \{y \in \mathbb{R}^n : d(x, y) < \varepsilon\}.$$

When discussing analysis in $\mathbb{C}$, we use the distance function

$$d(z, w) = |z - w| \qquad (z, w \in \mathbb{C}$$

and this is equivalent to regarding $\mathbb{C}$ as $\mathbb{R}^2$.

Closures; open and closed sets; connectedness

If $X \subseteq \mathbb{R}^n$, the closure $\overline{X}$ of X is the set of all limits of sequences of points in X. Equivalently[1], $\overline{X}$ is the set of all points y such that for all $\varepsilon > 0$ there is a point $x \in X$ with $d(x, y) < \varepsilon$.

If $D \subseteq X \subseteq \mathbb{R}^n$ and $\overline{D} \supseteq X$, we say that D is *dense* in X.

If $X = \overline{X}$, we say the set X is *closed.* A set $X \subseteq \mathbb{R}^n$ is *open* if $\mathbb{R}^n \setminus X$ is closed; equivalently, X is open iff for every $x \in X$ there is an $\varepsilon > 0$ with $B(x; \varepsilon) \subseteq X$.

A set $X \subseteq \mathbb{R}^n$ is *disconnected* if it may be written as a disjoint union $X = X_1 \cup X_2$ such that no sequence in X_1 converges to a point of X_2 and no sequence in X_2 converges to a point of X_1. A set is *connected* iff it is not disconnected.

Continuity

Let $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ and $f : X \to Y$. Then f is said to be *continuous* if whenever $x_n \to x$ in X we have $f(x_n) \to f(x)$ in Y. Equivalently, f is continuous iff for all $x \in X$ and all $\varepsilon > 0$ there exists $\delta > 0$ such that $d(f(x), f(y)) < \varepsilon$ whenever $y \in X$ with $d(x, y) < \delta$.

If $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ and $f : X \to Y$ is a bijection, and both f and f^{-1} are continuous, then we say that f is a *homeomorphism* and that the sets X, Y are *homeomorphic.* [Do not confuse 'hom**e**omorphisms' with the 'homomorphisms' of group theory.]

If two sets are homeomorphic then they share any property that can be described purely in terms of convergence of sequences. For example, connectedness is such a property. If X is a circle and Y a (perimeter of a) square, it is easy to produce a homomorphism between X and Y. These sets share the property that they are connected and remain connected if any one point is removed. [Proving this rigorously is non-trivial.] The set $[0, 1] \subseteq \mathbb{R}$ does not have this property and therefore is not homeomorphic to X and Y.

Fact A function $f : \mathbb{R} \to \mathbb{R}$ is a homeomorphism iff either:

(a) f is continuous and strictly increasing and $f(x) \to +\infty$ as $x \to +\infty$ and $f(x) \to -\infty$ as $x \to -\infty$; or

(b) f is continuous and strictly decreasing and $f(x) \to -\infty$ as $x \to +\infty$ and $f(x) \to +\infty$ as $x \to -\infty$.

[1] Here, and in the discussion of continuity that follows, we take a cavalier attitude to the foundations of mathematics and assume (a weak form of) the Axiom of Choice in order to say that the sequential definition and the definition using epsilons are equivalent. Non-believers in the Axiom of Choice take the epsilon definition and refer to the other as 'sequential closure', 'sequential continuity', etc. As for those extremists whose minds have been so warped by constructivism that they believe all functions continuous...

3 WEAKLY Attracting and repelling points — Graphical iteration

In order to discuss the idea of a sequence $x, f(x), f^2(x), \ldots$ converging, we need to have a some idea of distance between points of the space on which f is acting. From now on, f will be a continuous function acting on a subset X on some Euclidean space $\mathbb{R}^N$. In some places, where we need to talk about f^{-1}, we shall require f to be a *homeomorphism*; i.e. that both f and f^1 be continuous. We shall write $d(x, y)$ for the Euclidean distance between two points $x, y \in \mathbb{R}^N$.

Actually, all the discussion in this section can be carried out in the more general context of 'metric spaces' as those of you taking the Metric Spaces course will readily recognize.

Definition 3.1 Let p be a fixed point of f. A point $x \in X$ is *(forward) asymptotic* to p if $f^i(x) \to p$ as $i \to \infty$. In this case we write $x \in W^s(p)$.

Let p be a periodic point of f of order m. A point $x \in X$ is *(forward) asymptotic* to p if $(f^m)^i(x) = f^{mi}(x) \to p$ as $i \to \infty$. Again, we write $x \in W^s(p)$.

We call $W^s(p)$ the *stable set* or *basin of attraction* of p.

The following proposition tells us that we may apply the definition only knowing a period, rather than the precise order of p.

Proposition 3.2 *Let p be periodic of order m and let n be any period of p. Then $f^{ni}(x) \to p$ if and only if $f^{mi}(x) \to p$.*

Proof. We know that m divides n, say $n = ma$. If $f^{mi}(x) \to p$, then $f^{ni}(x) = (f^m)^{ia}(x) \to p$, because subsequences of convergent sequences converge to the same limit.

Conversely, suppose $(f^m)^{ia}(x) = f^{ni}(x) \to p$. Then

$$f^{(ia+k)m}(x) = f^{km}(f^{iam}(x)) \to f^{km}(p) = p \quad (0 \leqslant k < a),$$

using the continuity of f. Thus the sequence $(f^{jm})_{j=1}^{\infty}$ is a union of a subsequences, each convergent to p. Therefore $f^{jm} \to p$ as $j \to \infty$. $\Diamond$

Proposition 3.3 *Let f be a continuous function acting on X. If $f^i(x) \to p$ as $i \to \infty$, then p is a fixed point of f.*

Proof. If $f^i(x) \to p$ as $i \to \infty$, then $f^{i+1}(x) \to p$ as $i \to \infty$; but $f^{i+1}(x) = f(f^i(x)) \to f(p)$, since f is continuous. Therefore $f(p) = p$. $\Diamond$

Example 3.4 Let $f : \mathbb{R} \to \mathbb{R} : x \mapsto -x^3$. Then $\mathrm{Fix}(f) = \{0\}$, (the only real root of $-x^3 = x$), and $\mathrm{Per}_2(f) = \{-1, 0, +1\}$, (by solving $-(-x^3)^3 = x$). In fact $\mathrm{Per}(f) = \mathrm{Per}_2(f)$; this can be shown directly by solving the equation $f^n(x) = x$, i.e.

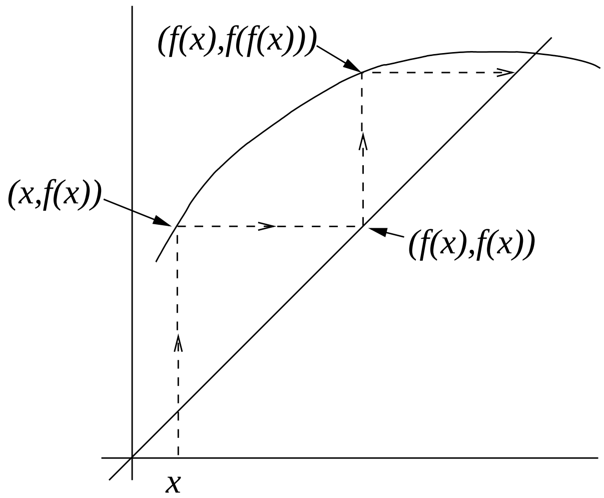
$$(-1)^n x^{3^n} = x,$$

but an alternative approach is given in Example 5.3 below.

$$\begin{aligned} W^s(0) &= (-1, 1) \\ W^s(1) &= \{1\}, \\ W^s(-1) &= \{-1\}. \end{aligned}$$

Graphical Analysis

Graphical analysis is a very useful means of visualizing the sequence of iterates of a point under a given function. Draw the graph of the function, f, say. Starting at a point with first coordinate x, draw a line vertically to the graph, reaching a point $(x, f(x))$; then draw horizontally from there to the diagonal $y = x$, reaching a point $(f(x), f(x))$; then vertically to the graph at $(f(x), f(f(x))) = (f(x), f^2(x))$; thence to the diagonal at $(f^2(x), f^2(x))$; and so on: $(f^2(x), f^3(x))$, $(f^3(x), f^3(x))$, $(f^3(x), f^4(x))$, $\ldots$. This enables us to track the iterates $f^n(x)$.



Definition 3.5 A periodic point p of order m is said to be *weakly attracting* if $W^s(p)$ contains an open ball $B(p,\delta) = \{x \in X : d(x,p) < \delta\}$ for some $\delta > 0$; i.e. for all $x \in B(p,\delta)$, $f^{mk}(x) \to p$ as $k \to \infty$.

A periodic point p of order m is said to be *weakly repelling* there is an open ball $U = B(p,\delta)$ such that for all $x \in U \setminus \{p\}$, there is a positive integer k with $f^{mk}(x) \notin U$. It need not follow that $f^{mk}(x)$ does not tend to p: for example, in Example 1.7, (the map $f : S^1 \to S^1 : z \mapsto z^2$), the fixed point 1 is weakly repelling, but every point $z = \exp(2^{-n}\pi i)$ $(n = 1, 2, 3, \ldots)$ (and indeed all $\exp i\theta$ for which the binary expansion of θ/π terminates) have $f^N(z) = 1$ for all sufficiently large N. Every U contains infinitely many such points z.

4 Maps on the real line and the interval: Hyperbolic points

Until further notice, let f be continuously differentiable : $[0,1] \to [0,1]$ or $\mathbb{R} \to \mathbb{R}$.

Definition 4.1 A fixed point p of f is said to be *hyperbolic* iff $|f'(p)| \neq 1$. We call $f'(p)$ the *multiplier* of p. For a periodic point p of order n, p is *hyperbolic* if $|(f^n)'(p)| \neq 1$.

Example 4.2

1. $f(x) = x$ and $f(x) = -x$ are maps whose fixed points are NOT hyperbolic. Such maps are atypical.
2. $f(x) = (x^3 + x)/2$ has fixed points -1,0,1: $f'(0) = 1/2, f'(\pm 1) = 2$ All three fixed points are hyperbolic.
3. $f(x) = -(x^3 + x)/2$ has fixed point 0: $f'(0) = -1/2$. The points +1,-1 form a periodic orbit of order 2. Then $(f^2)'(\pm 1) = f'(1)f'(-1) = (-2)(-2) = 4$, by the chain rule. All periodic points are hyperbolic.

Note: if p is periodic for f, order m, then, by the Chain Rule,

$$\begin{aligned}
(f^m)'(p) &= f'(f^{m-1}(p))\,(f^{m-1})'(p) \\
&= f'(f^{m-1}(p))\,f'(f^{m-2}(p))\,(f^{m-2})'(p) \\
\cdots &= f'(f^{m-1}(p))\,f'(f^{m-2}(p))\ldots f'(f(p))\,f'(p) \\
&= f'(p_m)\,f'(p_{m-1})\ldots f'(p_2)\,f'(p_1),
\end{aligned}$$

where $\{p_1, p_2, \ldots, p_m\}$ is the orbit of p. In particular,

$$(f^m)'(p) = (f^m)'(p_i)$$

for any p_i in the orbit.

Theorem 4.3 *Attracting fixed points are weakly attracting. That is, if p is a fixed point of f such that $|f'(p)| < 1$ then there is an interval $(p - \delta, p + \delta)$ around p, for some $\delta > 0$, such that, for all $x \in (p - \delta, p + \delta)$ we have $f^n(x) \to p$ as $n \to \infty$.*

Corollary 4.4 *Attracting periodic points are weakly attracting: if p is periodic order n and $|(f^n)'(p)| < 1$ then there is an interval $(p - \delta, p + \delta)$ around p, for some $\delta > 0$, such that, for all $x \in (p - \delta, p + \delta)$ we have $f^{nk}(x) \to p$ as $k \to \infty$.*

The corollary follows immediately by applying (4.3) to the function f^n, having noted that f^n is continuously differentiable since f is.

Proof of Theorem. Let $k = (1 + |f'(p)|)/2$ and $\varepsilon = (1 - |f'(p)|)/2$.

Since f' is continuous, $|f'|$ is continuous, so there exists $\delta > 0$ such that

$$|f'(p)| - \varepsilon < |f'(y)| < |f'(p)| + \varepsilon = k < 1$$

for all $y \in (p - \delta, p + \delta)$.

By the Mean Value Theorem, if $0 < |x - p| < \delta$, then there exists y between x and p such that

$$\left| \frac{f(x) - f(p)}{x - p} \right| = |f'(y)| < k.$$

So $|f(x) - f(p)| \leqslant k|x - p|$ for all x in $(p - \delta, p + \delta)$,
so

so $$|f(x) - p| \leqslant k|x - p| < \delta \qquad (x \in (p - \delta, p + \delta)),$$

so $$|f^2(x) - p| \leqslant k|f(x) - p| \leqslant k^2|x - p| < \delta \qquad (x \in (p - \delta, p + \delta)),$$

et cetera:

$$|f^n(x) - p| \leqslant k^n|x - p| \qquad (x \in (p - \delta, p + \delta),\ n = 1, 2, 3, \ldots).$$

So $f^n(x) \to p$, for all x in $(p - \delta, p + \delta)$; i.e. $(p - \delta, p + \delta)$ is contained in $W^s(p)$. $\Diamond$

Definition 4.5 If p is a periodic point of order n and $|(f^n)'(p)| < 1$, we say p is an *attracting* periodic point. If p is a periodic point of order n and $|(f^n)'(p)| > 1$, we say p is a *repelling* periodic point.

The latter part of the definition is justified by:

Theorem 4.6 *If p is a repelling periodic point of f of order n, then there is an interval $U = (p - \delta, p + \delta)$ of p such that, for every x in $U \backslash \{p\}$, there is a positive integer k with $f^{nk}(x)$ outside U.*

Proof. Left as an exercise. $\Diamond$

For non-hyperbolic points many different behaviours are possible.

For a non-hyperbolic periodic point p, it is natural to say that p is *weakly attracting* if the conclusion of Theorem 4.3 holds, and *weakly repelling* if the conclusion of Theorem 4.6 holds. This explains why we introduced this terminology in the previous chapter.

Non-hyperbolic points are atypical, you do not expect to find them for a function selected "at random". However, if you sweep through a whole family of maps, you should not be surprised to find the odd parameter value for which there is a non-hyperbolic periodic point.

Example 4.7 The family of quadratic maps $Q_c(x) = x^2 + c$. There are no fixed points for $c > 1/4$, two for $c < 1/4$ and one for $c = 1/4$ (just solve $x = x^2 + c$). For $c = 1/4$, the unique fixed point is non-hyperbolic.

Take $c < 1/4$. The fixed points are at $p = (1 - (1 - 4c)^{1/2})/2$, $q = (1 + (1 - 4c)^{1/2})/2$. Now $Q_c'(x) = 2x$, so $Q_c'(q) = 1 + (1 - 4c)^{1/2} > 1$, so q is repelling, and $Q_c'(p) = 1 - (1 - 4c)^{1/2}$, so $Q_c'(p) < 1$ always and $Q_c'(p) > -1$ iff $1 - 4c < 4$, i.e. $c > -3/4$. So p is attracting for $-3/4 < c < 1/4$. For $c = -3/4, p = -1/2$ is non-hyperbolic; it can be shown to be weakly attracting. For $c < -3/4$, we have $Q_c'(p) < -1$ and so p is repelling.

5 Monotonic maps

The setting in this chapter is that, unless otherwise stated, f will be a continuous function either $\mathbb{R} \to \mathbb{R}$ or $[a, b] \to [a, b]$ for some $a, b \in \mathbb{R}$.

Definition 5.1 A function f is said to be:

(a) *strictly increasing* if $x < y \;\; \Rightarrow \;\; f(x) < f(y)$;

(b) *increasing* if $x \leqslant y \;\; \Rightarrow \;\; f(x) \leqslant f(y)$;

(c) *strictly decreasing* if $x < y \;\; \Rightarrow \;\; f(x) > f(y)$;

(d) *decreasing* if $x \leqslant y \;\; \Rightarrow \;\; f(x) \geqslant f(y)$.

In all of these cases, f is said to be *monotonic*

The dynamics of monotonic maps are uninteresting. Suppose $f : \mathbb{R} \to \mathbb{R}$ is increasing. If $x \in \mathbb{R}$ and $f(x) \geqslant x$, then $f(f(x)) \geqslant f(x)$, $f(f(f(x))) \geqslant f(f(x))$, *et cetera*; i.e. the sequence $f^n(x)$ $(n = 1, 2, 3, \dots)$ is an increasing sequence. It therefore is either unbounded, in which case $f^n(x) \to +\infty$, or it is convergent. However, by Theorem 3.3, if $(f^n(x))$ converges, its limit must be a fixed point. Likewise, if $f(x) \leqslant x$, then $f^n(x)$ either tends to $-\infty$ or converges down to a fixed point. This argument is the analytic proof corresponding to most pieces of graphical iteration.

If f is decreasing, the situation is only slightly less boring. In this case, the function f^2 is increasing, and so the sequences $(f^{2n}(x))_{n=1}^{\infty}$ and $(f^{2n+1}(x))_{n=0}^{\infty}$ are one decreasing and the other increasing. If $(f^{2n}(x))_{n=1}^{\infty}$ is bounded, then, as above, it converges to a point p which is a fixed point of f^2. Then $f^{2n+1}(x) \to f(p)$, so "$f^n(x)$ converges to the 2-cycle $\{p, f(p)\}$". Note that we could have $f(p) = p$, in which case $f^n(x) \to p$. A similar argument applies if $(f^{2n+1}(x))_{n=0}^{\infty}$ is bounded. If both sequences are unbounded, then $f^{2n}(x) \to +\infty$ and $f^{2n+1}(x) \to -\infty$, or vice versa.

To summarize:

Theorem 5.2 *Let $f : \mathbb{R} \to \mathbb{R}$ be monotonic and $x \in \mathbb{R}$.*

(a) If f is increasing and $f(x) \geqslant x$, then either $f^n(x) \uparrow p$ for some fixed point p, or $f^n(x) \uparrow +\infty$.

(b) If f is increasing and $f(x) \leqslant x$, then either $f^n(x) \downarrow p$ for some fixed point p, or $f^n(x) \downarrow -\infty$.

(c) If f is decreasing, then either the iterates $f^n(x)$ converge to a 2-cycle $\{p, f(p)\}$ (possibly with $f(p) = p$), or they converge to the "2-cycle" $\{\pm\infty\}$.

Example 5.3 In Example 3.4, we considered the function $f(x) = -x^3$ and observed that $\mathrm{Fix}(f) = \{0\}$ and $\mathrm{Per}_2(f) = \{-1, 0, +1\}$, by solving the equation $f^2(x) = x$. The fact that f is decreasing (in fact, strictly decreasing) now tells us that for any $x \in \mathbb{R}$, the sequence of iterates $f^n(x)$ either tends to the fixed point 0 or to the 2-cycle $\{+1, -1\}$ or to $\{+\infty, -\infty\}$. None of these would be possible if x were a periodic point of order greater than 2. Therefore $\mathrm{Per}(f) = \mathrm{Per}_2(f)$, and we have not even had to think about the equations $f^n(x) = x$ to prove this.

Perhaps we have overdone the ennui of monotonic maps. It is possible to find a strictly increasing function $f : [0, 1] \to [0, 1]$ whose set of fixed points is uncountable; (in fact, with $\mathrm{Fix}(f)$ equal to the "Cantor set" − see the Fractals course).

Nevertheless, monotonic maps are easily understood. We shall next turn our attention to maps which have just one maximum or minimum: *unimodal* maps. The dynamics are similar whether the map has a maximum or a minimum; (see the chapter on "topological conjugacy"); so we shall look at maps with a single maximum.

6 The quadratic family

The most obvious examples of functions with a single maximum are quadratic maps. It is found that they have many, but not all, of the features of general unimodal maps. Amongst the quadratic maps, we shall see later, (in the chapter on topological conjugacy), that it suffices to consider either the family Q_c ($c < 1/4$) mentioned above or, as we shall prefer, the family

$$F_\mu(x) = \mu x(1-x) \qquad (\mu > 1). \tag{1}$$

For convenience, we write F in place of F_μ where there will be no confusion. When studying the iterates of F, we are, in effect, studying solutions of the difference equation

$$x_{n+1} = \mu x_n(1-x_n);$$

the so-called *logistic equation*.

Before we embark on a study of the F_μ, it will be interesting to see how the logistic equation arises quite naturally in a biological setting.

Example 6.1 Population dynamics. Let P_n denote the population at year n. A very simple model, with fixed birth and death rates would be

$$P_{n+1} = aP_n$$

leading to exponential growth: $P_n = a^n P_0$.

A more sophisticated model has an extra term for enhanced death rate due to overpopulation — competition for food, cannibalism, etc.:

$$P_{n+1} = aP_n - bP_n^2.$$

The transformation $x_n = bP_n/a$ turns this into the canonical form

$$x_{n+1} = ax_n(1-x_n).$$

The continuous analogue of this is the differential equation

$$\frac{dx}{dt} = (a-1)x - ax^2$$

which has a perfectly respectable solution (separate the variables). It therefore comes as a surprise that the difference equation (which is more realistic, for species one would sample each year after the breeding season) exhibits more complicated, including chaotic, behaviour. The paper which brought this to prominence is

R.M.May "Simple mathematical models with very complicated dynamics", *Nature*, 261 (10 June 1976), 459-467
(included in both the Hao Bai-Lin and the Cvitanovic reprint selections).

There is an interesting account of recent research demonstrating chaotic behaviour in flour beetle populations in the chapter "Beetlemania: Chaos in Ecology" of Barry Cipra's "What's Happening in the Mathematical Sciences 1998-1999" (American Mathematical Society, 1999). In this case, cannibalism causes the nonlinearity. The equations are complicated by the fact that four stages of the beetle's life cycle — egg, larva, pupa and adult — need to be considered. However, the equations do show similar dynamics to that of the logistic equation and *chaotic fluctuations of population have been demonstrated experimentally*, albeit with artificially enhanced mortality rates.

Henceforth, we assume $\mu \geqslant 1$. The elementary facts about F_μ are as follows.

Proposition 6.2 *1.* $F(0) = F(1) = 0$*;*

2. $F(x) = x$ *iff* $x = 0$ *or* $x = p_\mu := (\mu-1)/\mu$*;*

3. $0 \leqslant p_\mu < 1$;

4. $\max F(x) = F(\frac{1}{2}) = \frac{\mu}{4}$ *so* $F : [0,1] \to [0,1]$ *iff* $\mu \leqslant 4$;

5. $p_\mu \leqslant \mu/4$ *with equality iff* $\mu = 2$;

6. $F'_\mu(0) = \mu$ *and* $F'_\mu(p_\mu) = 2 - \mu$; *hence 0 is repelling for* $\mu > 1$, *and* p_μ *is attracting for* $1 < \mu < 3$, *repelling for* $\mu > 3$ *and non-hyperbolic for* $\mu = 3$. *For* $\mu = 1$, *the single fixed point* $p_\mu = 0$ *is non-hyperbolic.*

The proof is by A-level calculus. For convenience, we write p in place of p_μ where there will be no confusion.

[PICTURE: Graph of F_4.]

The dynamics of F outside [0,1] are trivial.

Proposition 6.3 *If* $x < 0$ *or* $x > 1$, *then* $F^n(x) \to -\infty$ *as* $n \to \infty$.

Proof. If $x < 0$, then $\mu(1-x) > 1$ so $F(x) = \mu(1-x)x < x$. Since F is strictly increasing on $(-\infty, 0]$, it follows that the sequence $F^n(x)$ is strictly decreasing, so either it converges or $\to -\infty$. However, if it converges, it must converge to a fixed point, and the only fixed points are 0 and p, both non-negative. Therefore $F^n(x) \to -\infty$ for $x < 0$.

For $x > 1$, the first step gives $F(x) = \mu x(1-x) < 0$, so $F^n(x) = F^{n-1}(F(x)) \to -\infty$ as $n \to \infty$. $\Diamond$

[PICTURE OR COMPUTER DEMO: Proof of Proposition on F for large and small x]

Since 0 is a fixed point and 1 is eventually fixed ($F(1) = 0$), this proposition says that all the interesting dynamics is confined to the open interval (0,1).

[The rest of this chapter contains some detailed argument. You will be expected to understand the principles involved in this argument, but not to reproduce the details as bookwork.]

We study the family $F(x) = \mu x(1-x)$ for $1 \leqslant \mu \leqslant 3$.

Lemma 6.4 *For* $1 \leqslant \mu \leqslant 3$ *the function* F^2 *has no fixed points except 0 and* $p = (\mu-1)/\mu$.

Proof. Using MAPLE, we solve $F^2(x) = x$ and obtain the roots:

$$0, \quad p_\mu, \quad \frac{\mu + 1 \pm \sqrt{\mu^2 - 2\mu - 3}}{2\mu}.$$

The roots of the equation $\mu^2 - 2\mu - 3 = 0$ are -1 and $+3$, so for $\mu \in (-1, 3)$, the last two roots of $F^2(x) = x$ are non-real. At $\mu = 3$, the last two roots are real and coincident — and equal to $\frac{2}{3} = p_\mu$; again we have only two distinct real roots. $\Diamond$

Proposition 6.5 *If* $1 \leqslant \mu \leqslant 3, 0 < x_0 < 1$, *then* $F^n(x_0) \to p$ *as* $n \to \infty$.

[COMPUTER DEMOS throughout proof .]

Proof.

Case $\mu = 1$. If $\mu = 1$, then $F(x) \leqslant x$ for all x, so the sequence $x_0, F(x_0), F^2(x_0), \ldots$ is monotone decreasing in (0,1). It therefore converges to a fixed point of F in [0,1]. The only fixed point is $p = 0$.

Case $1 < \mu \leqslant 2$. In this case, $0 < \mu - 1 \leqslant \frac{1}{2}\mu$, so $0 < p \leqslant \frac{1}{2}$.

(a) Let $0 < x_0 \leqslant p$. For $0 < x \leqslant p$, we have

$$0 < F(x) \leqslant F(p) = p,$$

because F is strictly increasing on $[0, \frac{1}{2}]$. Moreover,

$$\begin{aligned} \mu x(1-x) - x &= (\mu - 1)x - \mu x^2 \\ &= \mu x(p - x) \\ &\geqslant 0. \end{aligned} \qquad (2)$$

Hence $F^n(x_0)$ is an increasing sequence in the interval $(0, p]$, so it converges to a point which must be a fixed point and therefore must be p.

(b) Let $p < x_0 \leqslant \frac{1}{2}$. For $p < x \leqslant \frac{1}{2}$, we have

$$\begin{aligned} p &= F(p) \\ &< F(x) \\ &\leqslant F(\frac{1}{2}) \\ &= \mu/4 \\ &\leqslant \frac{1}{2} \end{aligned}$$

because F is strictly increasing on $[0, \frac{1}{2}]$, so $F : (p, \frac{1}{2}] \to (p, \frac{1}{2}]$. Moreover,

$$\mu x(1-x) - x = \mu x(p-x) < 0.$$

Hence $F^n(x_0)$ is an strictly decreasing sequence in the interval $(p, \frac{1}{2})$, again it converges to p.

(c) Let $\frac{1}{2} < x_0 < 1$. Now

$$0 < F(x) < F(\frac{1}{2}) = \mu/4 \leqslant \frac{1}{2} \quad (\frac{1}{2} < x < 1),$$

and so, again, $F^n(x_0) \to p$.

Case $2 < \mu \leqslant 3$. In this case, $p > \frac{1}{2}$ so $p \in (\frac{1}{2}, \frac{\mu}{4})$.

(a) Let $0 < x_0 < \frac{1}{2}$. Now if $0 < x < \frac{1}{2}$, then $0 < F(x) < \mu/4$ and, by (2) above, $F(x) > x$. Thus the sequence $F^n(x_0)$ is strictly increasing. If it were contained in the interval $(0, \frac{1}{2})$, then it would tend to a fixed point in $(0, \frac{1}{2}]$, which is impossible. Therefore this sequence escapes: there exists k with $\frac{1}{2} \leqslant F^k(x_0) \leqslant \mu/4$. Its subsequent adventures are described in the following three cases.

(b) Now consider $x_0 \in [\frac{1}{2}, p)$. Since F is strictly decreasing on $[\frac{1}{2}, 1]$, if $\frac{1}{2} \leqslant x < p$ then $p < F(x) \leqslant \mu/4$, and $F(\mu/4) \leqslant F^2(x) < p$. Now $F(\mu/4) = \mu^2(4-\mu)/16 > \frac{1}{2}$, so F^2 maps the interval $[\frac{1}{2}, p)$ into itself. Since $F^2(\frac{1}{2}) > \frac{1}{2}$, we must have $F^2(x) > x$ for all $x \in [\frac{1}{2}, p)$, because if $F^2(x) \leqslant x$, then, by the Intermediate Value Theorem, there would be a point $y \in (\frac{1}{2}, x]$ with $F^2(y) = y$, contradicting the fact that the only fixed points of F^2 are 0 and p. Therefore, the sequence $F^{2n}(x_0)$ is strictly increasing, bounded above (by p), and so converges to a fixed point of F^2. By the previous lemma, the only fixed points of F^2 are 0 and p. Therefore $F^{2n}(x_0) \to p$ as $n \to \infty$. Since F is continuous, $F^{2n+1}(x_0) \to p$, and hence $F^n(x_0) \to p$.

(c) If $x_0 = p$ then, of course $F^n(x_0) = p$ for all n.

(d) Now consider $x_0 \in (p, \mu/4]$. Then, as above, we have

$$\frac{1}{2} < F(\mu/4) \leqslant F(x_0) < F(p) = p,$$

and so $F^n(x_0) = F^{n-1}(F(x_0)) \to p$ as $n \to \infty$. This completes the discussion of case (a).

(e) Finally, consider $x_0 \in (\mu/4, 1) \subseteq (p, 1)$. Since F is strictly decreasing on $[\frac{1}{2}, 1]$, $0 < F(x_0) < p$, and so $F^n(x_0) \to p$, by the preceding four cases.

This completes the proof. $\Diamond$

7 The bifurcations of F_μ

$1 \leqslant \mu < 3$. Here, F_μ has an attracting fixed point $p_\mu = (\mu - 1)/\mu$. It also has an repelling fixed point at 0. Generally, repelling fixed points are invisible to the computer and of little physical interest, so we ignore them.

$\mu = 3$. Here, p_μ changes to repelling and is ignored, and a 2-cycle is born.

$3 < \mu < 1 + \sqrt{6} = 3.449499\ldots$. The 2-cycle is stable.

$\mu = 1 + \sqrt{6}$. The 2-cycle becomes unstable and a stable 4-cycle is born.

$\mu = 3.544090$. The 4-cycle becomes unstable and a stable 8-cycle is born.

Generally, a stable 2^k-cycle is born at a_k and becomes unstable at a_{k+1}, where:

$$\begin{array}{lllllll} a_1 & = & 3 & & a_5 & = & 3.568759\ldots \\ a_2 & = & 3.449499\ldots & & a_6 & = & 3.569692\ldots \\ a_3 & = & 3.544090\ldots & & a_7 & = & 3.569891\ldots \\ a_4 & = & 3.564407\ldots & & a_8 & = & 3.569934\ldots, \end{array}$$

and, approximately,

$$a_k = a_\infty - c\delta^{-k},$$

where $a_\infty = 3.569946\ldots, c = 2.6327\ldots, \delta = 4.669202\ldots$.

Beyond this *period doubling cascade* we find chaos mixed with windows of stability. For example, a stable 3-cycle emerges at $\mu = 1 + \sqrt{8} = 3.828427\ldots$, and becomes unstable at $\mu = 3.8415$, at which point a stable 6-cycle is born. This, in turn, gives way to a 12-cycle, a 24-cycle, *et cetera* , with the values of μ at which the transitions occur converging to 3.8495. Likewise there are windows where a stable N-cycle emerges at $\mu = b$, and becomes unstable at $\mu = c$, at which point a stable $2N$-cycle is born, which gives way to a $4N$-cycle, an $8N$-cycle, *et cetera* , at points converging to $\mu = d$, as shown in the table below.

N	b	c	d
1	1.0000	3.0000	3.569946
6	3.6265	3.6304	3.6327
5	3.7382	3.7411	3.7430
3	3.8284	3.8415	3.8495
5	3.9056	3.9061	3.9065
6	3.937516	3.937596	3.937649
4	3.9601	3.9608	3.9612
6	3.977760	3.977784	3.977800
5	3.99026	3.99030	3.99032
6	3.997583	3.997585	3.997586

The paper: N.Metropolis, M.L.Stein and P.R.Stein "On finite limit sets for transformations on the unit interval" *J. Combinatorial Theory,* **15**, (1973), 25-44, lists all the stable N-cycles up to $N = 11$ in order of increasing μ.

[COMPUTER DEMOS:

(1) Building up the bifurcation diagram line-by-line.

(2) Automatic drawing of the bifurcation diagram using FRACTINT.

[HANDOUT: Picture of the bifurcation diagram]

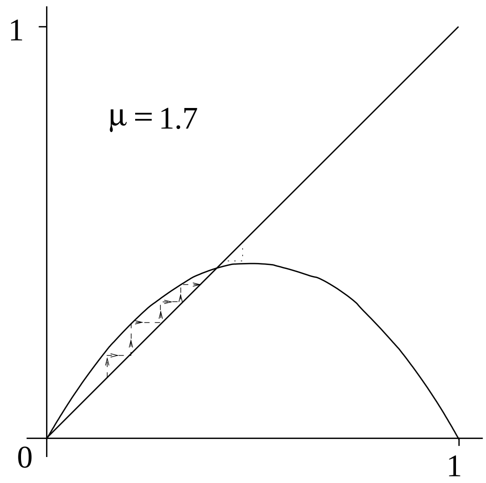
For a long time it was a famous unsolved problem whether the "windows of periodicity" are dense in the interval [1,4]. The expected result, that they are dense, was proved by J. Graczyk and G. Swiatek,'Generic hyperbolicity in the logistic family', *Ann. of Math.* (2) **146** (1997), no. 1, 1–52;

MR 99b:58079. A more complete treatment of their theory appears in their book 'The real Fatou conjecture' (Princeton University Press, Annals of Math. Studies **144**, 1998) Main Lib. 3 PER 510.5.

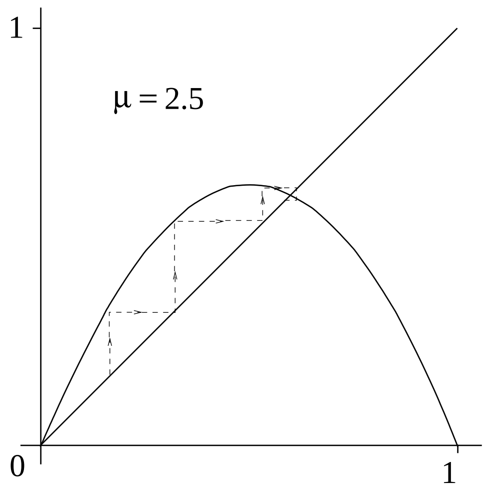
We now look at the way that qualitative changes in the dynamics of F_μ occur as the parameter μ passes through certain critical values.

7.1 Period doubling bifurcations

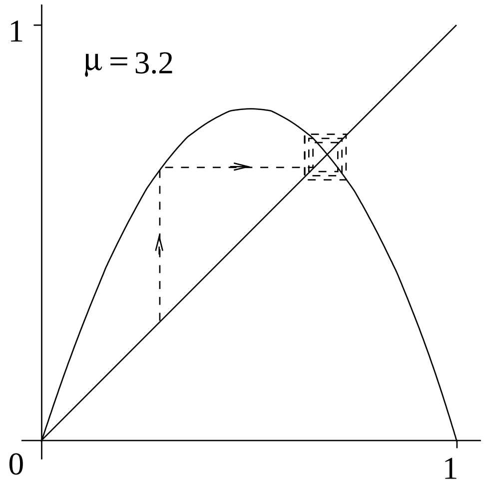
We observe the first bifurcation of this type at $\mu = 3$. The non-zero fixed point changes from attracting to repelling and, simultaneously, an attracting 2-cycle is born there.
Graphical analysis near $\mu = 3$.



The iterates converge to the non-zero fixed point p from below (and likewise from above).



When the iterates get sufficiently close to p, they alternate above and below it, converging to it.



Diagrams of F_μ^2 near $\mu = 3$ are shown on below (Figures 1 to 3).

Just below $\mu = 3$ the function F_μ^2 has fixed points only at the fixed points of F_μ, namely 0 and $p = (\mu - 1)/\mu$, and it can be seen that 0 is repelling (gradient of the graph is > 1 there) and p is attracting (gradient of the graph in $(-1, 1)$).

At $\mu = 3$ the function F_μ^2 has the same fixed points, but at p, the graph of F_μ^2 is tangent to the diagonal, indicating that this point is non-hyperbolic.

This bifurcation is easily understood by considering how the graph of F_μ^2 changes with μ. For μ just below 3, it cuts the diagonal just once with gradient just less than 1. At $\mu = 3$, it cuts the diagonal once and is tangent to it. For μ just above 3, it cuts the diagonal three times: once at a point corresponding to the (repelling) fixed point p_μ of F_μ, and, on either side of p_μ, at points which correspond to the new 2-cycle. The fact that the 2-cycle is attracting (provided μ is only just above 3) can be seen from the fact that gradient of F_μ^2 at these points has modulus less than 1: in fact, the gradient starts at $+1$ and decreases with increasing μ.

If we plot the position of the periodic points against μ, as in the 'bifurcation diagram', we can see why this type of bifurcation is sometimes called a "pitchfork bifurcation".

Notice that the fixed point remains present after the bifurcation but, as it has become repelling, it does not show up on computer pictures. It is a general feature of the family F_μ that as μ increases, new periodic orbits may be born, but none disappear (J. Milnor and W. P. Thurston, "On iterated maps of the interval" in "Dynamical systems: Proceedings Univ. Maryland 1986–7" (Springer, Lecture Notes in Mathematics **1342**, (1989) pp. 465–563). This contrasts with the situation for one-parameter families of maps on $\mathbb{R}^2$, where there are typically parameter values such that infinitely many births and deaths of periodic orbits occur densely (I. Kan and J. Yorke, "Antimonotonicity: concurrent creation and annihilation of periodic orbits", *Bull. Amer. Math. Soc.*, **23**(1990), 469–476).

The graph of F_μ^2 inside the box resembles the graph of F_ν in the box $[0, 1] \times [0, 1]$, inverted and scaled down; where ν runs from 1 to 3 as μ runs from 3 to $1 + \sqrt{6}$. This resemblance may be made precise (with some difficulty) and it may be shown that the process of repeated bifurcation continues—a *period-doubling cascade*—with the sequence $a_1, a_2, \ldots$ of values of μ at which bifurcations occur having the property

$$\frac{a_{n+1} - a_n}{a_{n+2} - a_{n+1}} \to \delta = 4.6692016091029\ldots$$

as $n \to \infty$.

Amazingly, this constant δ is *universal.* If we replace the family F_μ by another family G_μ in which μ controls the magnitude of the nonlinearity with non-zero gradient and where G_μ has a single quadratic maximum, then the same period-doubling phenomena are observed with exactly the same constant δ.

Figure 1: $\mu < 3$

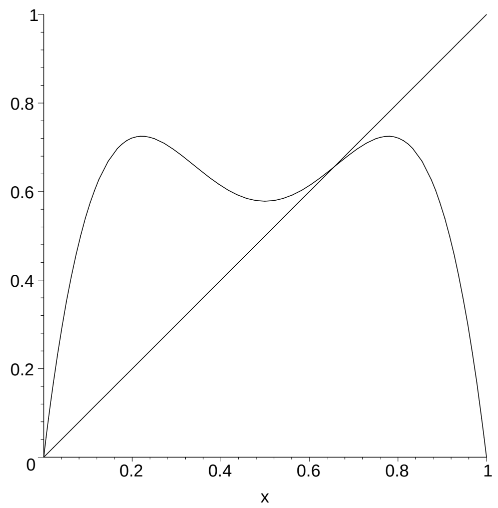
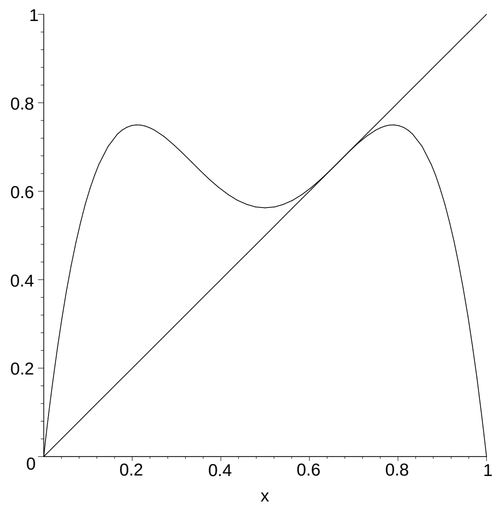


Figure 2: $\mu = 3$



An example of such a family is

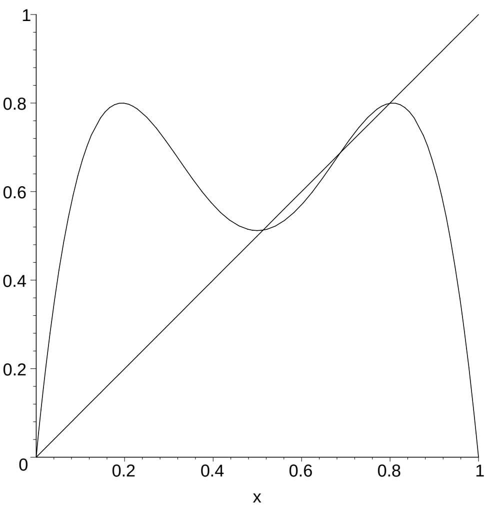
$$G_\mu(x) = \mu \sin(\pi x) \qquad (x \in [0,1],\ \mu \geqslant 0).$$

(Devotees of FRACTINT will find its bifurcation diagram there under fractal type "bif=sinpi".)

Saying that G_μ has a "quadratic maximum" means that if its maximum is attained at x_0 then, $G''_\mu(x_0) < 0$. If, instead, G_μ had a maximum of order $2n$, (i.e. $G_\mu^{(2n)}(x_0)$ is the first non-vanishing derivative at the maximum x_0), then we would again see a period doubling phenomenon, but with a different constant δ, depending (only) on n.

Another universal constant associated with the period-doubling cascade is obtained as follows. At the bifurcation point a_1, the fixed point p of F gives rise to two attracting fixed points q_1, q_2 of F^2. These points q_1, q_2 move apart as μ increases. Let s_1 denote the distance between them at the next bifurcation a_2. Choosing either of the branches, follow the resulting two points to the next bifurcation a_3, when their distance apart is s_2, and so on.

Figure 3: $\mu > 3$



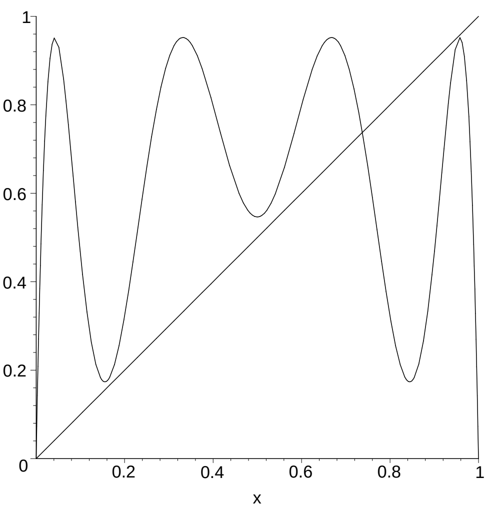
Then

$$\frac{s_n}{s_{n+1}} \rightarrow \alpha = 2.5029078750957\ldots$$

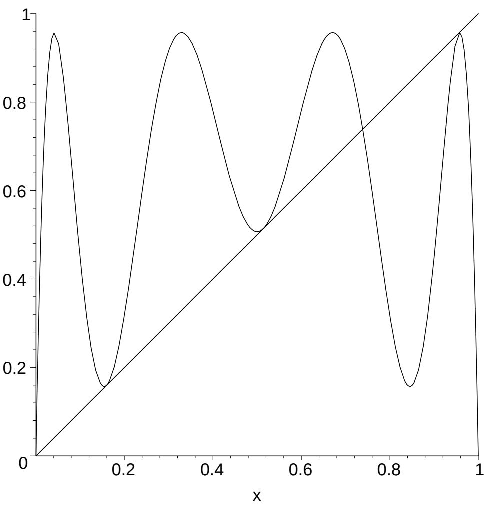
as $n \rightarrow \infty$. This constant α, like δ, is universal for all families G_μ with single quadratic maxima in which μ controls the magnitude of the nonlinearity with non-zero gradient. For families with maxima of higher order $2n$, the constant α depends on n.

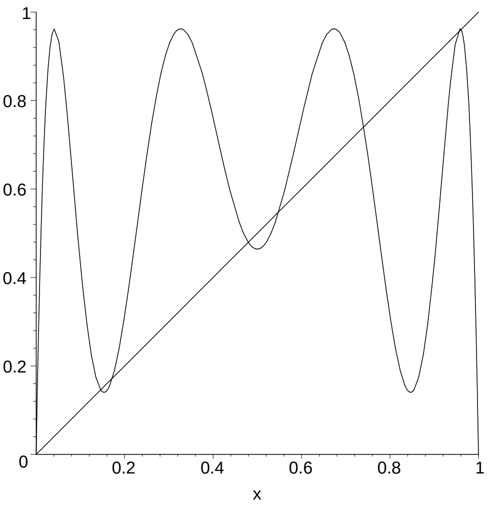
7.2 Tangent bifurcations.

The other type of bifurcation, called a *tangent* or *saddle-node bifurcation*, is most clearly observe at the point where an attracting periodic cycle of order three emerges. To see how this happens, we draw graphs of F_μ^3 for values of μ just below and just above the critical value ($\mu_c = 1 + \sqrt{8} = 3.828427\ldots$). F_μ^3 near $\mu = 1 + \sqrt{8}$.



At $\mu = \mu_c$, the graph of F_μ^3 is tangent to the diagonal at three points; the points of the emerging 3-cycle. As μ increases, these three non-hyperbolic fixed points of F_μ^3 each turn into a pair of fixed





points, one attracting and one repelling. Thus two 3-cycles of F_μ have appeared; one attracting and one repelling. The repelling cycle does not show up on the bifurcation diagram on the computer screen, of course. The attracting cycle will subsequently undergo a period-doubling cascade as μ increases.

One of the most interesting phenomena associated with the tangent bifurcation occurs at values of μ just below μ_c. Although the onset of pure period 3 behaviour is sudden, the behaviour for μ just below μ_c clearly presages it. At these values one observes long sequences of nearly periodic oscillations of period three, interspersed with short bursts of chaos. As μ approaches μ_c, the intervals of periodicity get longer. Such behaviour has also been observed experimentally in electrical circuits and in oscillatory chemical reactions (the Belousov–Zhabotinsky reaction). The phenomenon is known as *intermittency.*

To understand the reason for it, we look at graphical iteration of F_μ^3 at a value of μ just less than μ_c. We concentrate on that part of the graph which makes a close approach to the diagonal and then veers away. The iterates of a point at the left of this region remain near there for some while before emerging to the right. This correspond to the near-period-three behaviour. There follows a burst of chaos until the iterates are again trapped near this region, or one of the other two regions where the graph of F_μ^3 nears the diagonal.

In the words of Prof. Cvitanović:

> "For many iterations the system is being fooled into believing that it is converging toward a fixed point [of F_μ^3], only to discover that the fixed point is not there after all; it then wanders

away again, in the hope of finding a true fixed point."

Predrag Cvitanović, Introduction to "Universality in Chaos".

There is some universality in intermittency, too, in the form of the *Manneville–Pomeau scaling law*:

$$\text{duration of near-periodic motion} \quad \propto \quad (\mu_c - \mu)^{-1/2}.$$

The "$-1/2$" is universal for all families G_μ with single quadratic maxima, *et cetera* .

8 Period 3 implies chaos

This chapter is something of a special topic. It concerns a neat little theorem about continuous functions $f : \mathbb{R} \to \mathbb{R}$ discovered by Li and Yorke and published in the American Mathematical Monthly **82** (1975) 985-992 under the title of our chapter.

Theorem 8.1 *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. If f has a periodic point of order 3, then f has periodic points of all orders.*

Amazingly, for such a recently discovered theorem, the proof uses no more in the way of prerequisites than the Intermediate Value Theorem (which states that if $f : [a, b] \to \mathbb{R}$ is a continuous function and $y \in [f(a), f(b)]$, then there exists $c \in [a, b]$ such that $y = f(c)$).

Definition 8.2 Let I, J be compact intervals. We shall say that I *f-covers* J, and write $I \xrightarrow{f} J$, or just $I \to J$, if $f(I) \supseteq J$. If $f(I) = J$, we write $I \overset{f}{\twoheadrightarrow} J$ and say that I *precisely f-covers* J.

Lemma 8.3 *If I is a compact interval and $I \xrightarrow{f} I$, then f has a fixed point in I.*

Proof. If $I = [a, b]$ and $I \xrightarrow{f} I$, then there exist $c, d \in [a, b]$ such that $a = f(c)$ and $b = f(d)$. Then

$$f(c) - c = a - c \leqslant 0 \leqslant b - d = f(d) - d$$

so, by the Intermediate Value Theorem, there is a point x between c and d with $f(x) - x = 0$. $\diamondsuit$

The following properties are elementary.

Lemma 8.4 *Let I, J, K be compact intervals.*
(i) If $I \xrightarrow{f} J \supseteq K$ then $I \xrightarrow{f} K$.
(ii) If $I \xrightarrow{f} J \xrightarrow{g} K$ then $I \xrightarrow{g\circ f} K$.
(iii) If $I \overset{f}{\twoheadrightarrow} J \overset{g}{\twoheadrightarrow} K$ then $I \overset{g\circ f}{\twoheadrightarrow} K$.
(iv) If $I \xrightarrow{f} J$ then there is a closed subinterval $Q \subseteq I$ with $Q \overset{f}{\twoheadrightarrow} J$.

Proof. To prove (iv), let $J = [c, d]$ and let $a, b \in I$ with $f(a) = c$ and $f(b) = d$. Suppose $a < b$, the case $a > b$ being similar.

[PICTURE]

Let $x = \sup\{t \in [a, b] : f(t) \leqslant c\}$. Then for every positive integer n, since $x - \frac{1}{n}$ is not an upper bound for the set $\{t \in [a, b] : f(t) \leqslant c\}$, there exists $x_n \in [x - \frac{1}{n}, x]$ with $f(x_n) \leqslant c$. Since $x_n \to x$, by the continuity of f, we have $f(x) = \lim f(x_n) \leqslant c$. On the other hand, $x + \frac{1}{n} > x$ for all n, so $f(x + \frac{1}{n}) > c$, so $f(x) = \lim f(x + \frac{1}{n}) \geqslant c$. Thus $f(x) = c$.

Let $y = \inf\{t \in [x, b] : f(t) \geqslant d\}$. By similar reasoning, $f(y) = d$. Then $Q = [x, y]$ is the desired interval. $\diamondsuit$

Lemma 8.5 *If we have a finite sequence of compact intervals $I_0, I_1, \ldots, I_k$ with*

$$I_0 \xrightarrow{f} I_1 \xrightarrow{f} \ldots \xrightarrow{f} I_k = I_0,$$

then f^k has a fixed point p in I_0 with $f^j(p) \in I_j$ $(1 \leqslant j \leqslant k)$.

Proof. We construct, inductively, a sequence of compact intervals Q_j such that

$$I_0 = Q_0 \supseteq Q_1 \supseteq Q_2 \supseteq \ldots \supseteq Q_k$$

and

$$Q_j \xrightarrow{f^j}\!\!\!\!\rightarrow I_j \quad (0 \leqslant j \leqslant k).$$

The start, $Q_0 = I_0$, is trivial. Suppose $Q_0, Q_1, \ldots, Q_j$ have been constructed. Then

$$Q_j \xrightarrow{f^j}\!\!\!\!\rightarrow I_j \xrightarrow{f} I_{j+1}$$

so

$$Q_j \xrightarrow{f^{j+1}} I_{j+1},$$

so, by Lemma 8.4(iv), there is a closed subinterval $Q_{j+1} \subseteq Q_j$ such that $Q_{j+1} \xrightarrow{f^{j+1}}\!\!\!\!\rightarrow I_{j+1}$. This completes the induction.

Thus we have

$$Q_k \xrightarrow{f^k} I_k = I_0 \supseteq Q_k,$$

so $Q_k \xrightarrow{f^k} Q_k$. Lemma 8.3 then implies that f^k has a fixed point p in Q_k. Since $Q_k \subseteq Q_j$, we have $f^j(p) \in f^j(Q_j) = I_j$ $(1 \leqslant j \leqslant k)$. ◊

Proof of Theorem. Let a be a periodic point of f of order 3. Let $b = f(a), c = f^2(a)$. Replacing a by b or c if necessary, we may assume that either $a < b < c$ or $a > b > c$. The two cases are similar: we consider the former.

Let k be a positive integer. We show that f has a periodic point of order k. For $k > 1$, let

$$\begin{aligned} I_j &= [b, c] \quad (0 \leqslant j \leqslant k-2), \\ I_{k-1} &= [a, b], \\ I_k &= [b, c]. \end{aligned}$$

For $k = 1$, let $I_0 = I_1 = [b, c]$. Since $f(b) = c$ and $f(c) = a$ and f is continuous, $[b, c] \xrightarrow{f} [a, c]$ and so $[b, c] \xrightarrow{f} [a, b]$ and $[b, c] \xrightarrow{f} [b, c]$. Likewise, $[a, b] \xrightarrow{f} [b, c]$. Thus, for all $k \geqslant 1$,

$$I_0 \xrightarrow{f} I_1 \xrightarrow{f} \ldots \xrightarrow{f} I_k = I_0.$$

Lemma 8.5 applies to produce a point $p \in I_0$ which is a fixed point of f^k, and therefore a periodic point of f of *period* k. For $k = 1$, this is enough. For $k > 1$, we must show that k is the *order* of p. Suppose the order of p is $\ell < k$. Then $k - 1 \equiv \ell - 1 \pmod{\ell}$, so $f^{\ell-1}(p) = f^{k-1}(p)$. But $f^{k-1}(p) \in I_{k-1} = [a, b]$, whilst $f^{\ell-1}(p) \in I_{\ell-1} = [b, c]$. Therefore, $f^{k-1}(p) = b$. Hence $p = f^k(p) = f(b) = c$, which is impossible if $k = 2$ because $f^2(c) \neq c$. Hence, $a = f(p) \in I_1 = [b, c]$ if $k > 2$, giving another contradiction. Therefore, k is the order of p, and the proof is complete. ◊

Li and Yorke's paper made a substantial impact. It was probably responsible for introducing the word "chaos" into mathematics. The story of this theorem does not end there, however. We take up James Gleick's account of subsequent events:

> A few years later, attending an international conference in East Berlin, James Yorke took some time out for sightseeing and went for a boat ride on the Spree. Suddenly he was approached by a Russian trying urgently to communicate something. With the help of a Polish friend, Yorke finally understood that the Russian was claiming to have proved the same result. The Russian refused to give details, saying only that he would send his paper. Four months later it arrived. A. N. Sarkovskii had indeed been there first, in a paper titled "Coexistence of cycles of a continuous map of a line into itself".

Sarkovskii's paper was published in Ukrainian Math. J. **16** (1964) 61. It is hardly surprising that it failed to make an impression in the West. It was, however, unfortunate that it was ignored, for it is at the same time one of the most beautiful and one of the most elementary pieces of modern real analysis.

Definition 8.6 The *Sarkovskii ordering* of the positive integers is the following linear ordering. $3 \prec 5 \prec 7 \prec 9 \prec \ldots 6 \prec 10 \prec 14 \prec 18 \prec \ldots 12 \prec 20 \prec 28 \prec \ldots 24 \prec 40 \prec \ldots\ldots \quad \ldots 64 \prec 32 \prec 16 \prec 8 \prec 4 \prec 2 \prec 1$.

Theorem 8.7 (Sarkovskii) *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. If f has a periodic point of order n, then f has periodic points of all orders greater than n in the Sarkovskii ordering. Further, for each positive integer n there is a continuous function $f : \mathbb{R} \to \mathbb{R}$ which has a periodic point of order n, but no periodic points of order smaller than n in the Sarkovskii ordering.*

Corollary 8.8 *If f has a periodic point of order 3, then f has periodic points of all orders.*

The proof of Sarkovskii's Theorem is complicated, and we do not have time to go into the details, but, like the case proved by Li and Yorke, it uses nothing "higher" than the Intermediate Value Theorem. Probably the shortest proof to date is that of B. Gaweł(B. Gaweł, "On the theorems of Šarkovskii and Štefan on cycles", *Proc. Amer. Math. Soc.,* **107** (1989), 125–132).

Remark 8.9 Notice that we do need the little hypotheses of Sarkovskii's Theorem. Without the hypothesis of continuity, one could easily make up functions for which the theorem fails: just decide which cycles you want and define the function accordingly.

More surprisingly, the theorem is critically dependent on the underlying set being the real line, or an interval thereof. Let S^1 denote the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ with the usual metric. Then the function defined by $f(z) = e^{2\pi i/3}z$ is a continuous function from S^1 to itself such that every point is periodic of order 3. Likewise, on the plane, a rotation by $2\pi/3$ has every point except the centre periodic of order 3.

For the second part of Sarkovskii's Theorem – the examples which show that the first part is best possible – we shall give here the proof that period 5 does not imply period 3; various homework exercises will illustrate the other cases and, in effect, complete the proof.

Theorem 8.10 *There is a continuous function $f : \mathbb{R} \to \mathbb{R}$ with a periodic point of order 5 but no periodic points of order 3.*

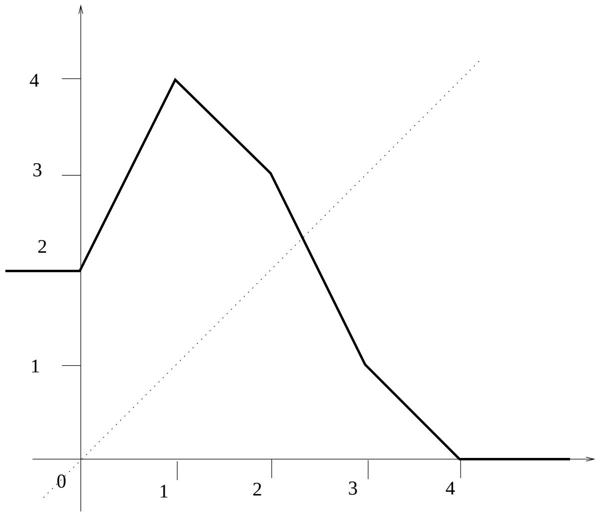
Proof. Define f by

$$\begin{aligned} f(x) &= 2 \quad (x \leqslant 0), \\ f(1) &= 4, \\ f(2) &= 3, \\ f(3) &= 1, \\ f(x) &= 0 \quad (x \geqslant 4), \end{aligned}$$

and f is linear on the intervals $(0, 1), (1, 2), (2, 3), (3, 4)$.

Clearly $(0, 2, 3, 1, 4)$ is a 5-cycle for f.

Clearly, also, there are no periodic points for f outside [0,4], since $f(\mathbb{R}) = [0, 4]$.



Now

$$[1,2] \xrightarrow{f} [3,4] \xrightarrow{f} [0,1] \xrightarrow{f} [2,4] \xrightarrow{f} [0,3] \xrightarrow{f} [1,4] \xrightarrow{f} [0,4].$$

Therefore,

$$\begin{aligned} f^3([1,2]) &= [2,4], \\ f^3([3,4]) &= [0,3], \\ f^3([0,1]) &= [1,4]. \end{aligned}$$

Thus, f^3 can have no fixed points in $[0,1] \cup [1,2] \cup [3,4]$, except, possibly the points 2,3 or 1. However, we know that the points 2,3 and 1 are periodic order 5 for f, so we can discount this possibility. There remains the possibility that f^3 has a fixed point in (2,3). Indeed it does: there is a fixed point of f in (2,3)! However, each of the maps

$$\begin{aligned} f : [2,3] &\to [1,3], \\ f : [1,3] &\to [1,4], \\ f : [1,4] &\to [0,4], \end{aligned}$$

is strictly decreasing. Therefore their composition

$$f^3 : [2,3] \to [0,4]$$

is strictly decreasing and so can have no more than one fixed point. Thus f^3 has no fixed point other than the fixed point of f. Therefore f has no periodic points of order 3. $\Diamond$

Exercise 8.11 1. Show that there is a function $f : \mathbb{R} \to \mathbb{R}$ with a periodic point of order 7 but no periodic points of order 5.

2. Given a continuous function $f : [0,1] \to [0,1]$ we define the *double* F of f by

$$F(x) = \begin{cases} \frac{2}{3} + \frac{1}{3} f(3x) & (0 \leqslant x \leqslant \frac{1}{3}) \\ \alpha x + \beta & (\frac{1}{3} < x < \frac{2}{3}) \\ x - \frac{2}{3} & (\frac{2}{3} \leqslant x \leqslant 1), \end{cases}$$

where the constants α, β are chosen to make F continuous at 1/3 and 2/3.

(a) Draw sketch graphs of a *typical* f (draw a random squiggle) and the corresponding F to illustrate this construction.

(b) Show that F has a fixed point in the interval (1/3,2/3). Call it x_0. By observing that

$$|F(x) - x_0| \geqslant 2|x - x_0| \qquad (x \in (1/3, 2/3)),$$

or otherwise, prove that F has no periodic points in (1/3,2/3) except x_0.

(c) Show that there are no periodic points of odd order in $[0, 1/3] \cup [2/3, 1]$

(d) Show that if $p \in [0, 1]$ is a periodic point of order n for f, then the points $p/3$ and $(p + 2)/3$ are periodic of order $2n$ for F.

(e) Show that all the periodic points of F in $[0, 1/3] \cup [2/3, 1]$ are of the above form.

To summarize: you have shown that the orders of the periodic points of F are twice the orders of the periodic points of f, together with 1.

3. Deduce that there is

   (a) a continuous function $F : \mathbb{R} \to \mathbb{R}$ with periodic points of order 10 but no periodic points of order 6 and no periodic points of whose orders are odd and greater than 1;

   (b) a continuous function $F : \mathbb{R} \to \mathbb{R}$ with periodic points of order 14 but no periodic points of order 10 and no periodic points of whose orders are odd and greater than 1;

4. Find a continuous function $f : [0, 1] \to [0, 1]$ with at least one fixed point, but no periodic points of order greater than 1. Using the doubling construction, deduce that, for every $n \geqslant 0$, there is a continuous function $F_n : [0, 1] \to [0, 1]$ with at least one periodic point of order 2^n, but no periodic points of orders other than $1, 2, 4, 8, \ldots, 2^n$.

5. Find a continuous function $G : \mathbb{R} \to \mathbb{R}$ with periodic points of orders all the powers of two and only those.

9 Topological conjugacy

When are two groups "essentially the same"? When they are *isomorphic*: i.e. when there is a bijective homomorphism of one onto the other.

When are two vector spaces "essentially the same"? When there is a bijective linear map of one onto the other.

When are two subsets of $\mathbb{R}^n$ "essentially the same"? When there is a bijective continuous map with continuous inverse of one onto the other. Such a map is called a *homeomorphism.* (This is the answer which is most useful for many purposes, including our present concerns. It allows the possibility of distorting the "geometry" of these subsets, while preserving their "topology". A different answer, which preserves the geometry, is the there should exist an "isometric" (distance-preserving) bijection between the sets.)

We refer to a pair (f, X), where f is a continuous mapping of a set $X \subseteq \mathbb{R}^N$ into itself, as a *dynamical system.* When are two dynamical systems "essentially the same"?

Definition 9.1 We say that two dynamical systems (f, X), (g, Y) are *topologically conjugate* iff there is a homeomorphism $h : X \to Y$ with $hf = gh$. We say that h is a *topological conjugacy.*

Topological conjugacy preserves most of the features of a dynamical system which interest us. Some examples follow.

Proposition 9.2 *Let $f : X \to X$ and $g : Y \to Y$ be topologically conjugate by $h : X \to Y$. If $p \in X$ is a periodic point of f of order m, then $h(p)$ is a periodic point of g of the same order.*

Proof. We observe that

$$hf^n = (hf)f^{n-1} = (gh)f^{n-1} = g(hf)f^{n-2} = g^2hf^{n-2} = \ldots = g^nh$$

for $n = 1, 2, 3, \ldots$. Therefore

$$f^n(p) = p \quad \Rightarrow \quad g^n(h(p)) = h(f^n(p)) = h(p).$$

Conversely,

$$g^n(h(p)) = h(p) \quad \Rightarrow \quad f^n(p) = h^{-1}(h(f^n(p))) = h^{-1}(g^n(h(p))) = h^{-1}(h(p)) = p.$$

Thus p is periodic of period n iff $h(p)$ is periodic of period n. Since the set of periods of p and $h(p)$ are the same, the least periods (i.e. orders) are the same. $\Diamond$

We immediately have an example of an interesting property preserved under topological conjugacy. We recall that a subset $D \subseteq X$ is said to be *dense* if every point in X is the limit of a sequence of (not necessarily distinct) points of D.

If D is dense in X and $h : X \to Y$ is a homeomorphism, then $h(D)$ is dense in Y. [Proof: if $y \in Y$, then $y = h(x)$ for some $x \in X$. Then $x = \lim d_n$ for some sequence of points $d_n \in D$. Since h is continuous, $y = \lim h(d_n)$.]

Proposition 9.3 *If (f, X) is a dynamical system in which the periodic points are dense (i.e. $Per(f)$ is a dense subset of X), then the same is true of any dynamical system topologically conjugate to (f, X).*

We now consider some examples of topologically conjugate systems.

Example 9.4 Consider the quadratic families

$$Q_c(x) = x^2 + c,$$

and

$$F_\mu(x) = \mu x(1 - x)$$

both acting on $\mathbb{R}$. We show that these are topologically conjugate. We find a homeomorphisms $h : \mathbb{R} \to \mathbb{R}$ of the form $h(x) = ax + b$ for suitable constants a, b depending on μ such that $hQ_c = F_\mu h$ for suitable $c = c(\mu)$. We have

$$h(Q_c(x)) = a(x^2 + c) + b = ax^2 + (ac + b)$$

and

$$F_\mu(h(x)) = \mu(ax + b)(1 - ax - b) = -\mu a^2 x^2 + \mu a(1 - 2b)x + \mu b(1 - b).$$

These two expressions agree if and only if

$$\begin{aligned} a &= -1/\mu, \\ b &= 1/2, \\ ac + b &= \mu b(1 - b). \end{aligned}$$

The last equation gives

$$c = -b(\mu b - \mu + 1)/a = \mu(2 - \mu)/4.$$

Thus the Q_c for $c \leqslant 1/4$ correspond to the F_μ for $\mu \geqslant 1$ (and to the F_μ for $\mu \leqslant 1$). We tabulate the values of c corresponding to the interesting values of μ.

μ	c
1	1/4
2	0
3	-3/4
$1 + \sqrt{5}$	-1
$1 + \sqrt{6}$	-5/4
$1 + \sqrt{7}$	-3/2
$1 + \sqrt{8}$	-7/4
4	-2

Another way of viewing this conjugacy is to think of these dynamical systems in terms of the corresponding difference equations:

$$y_{n+1} := \mu y_n(1 - y_n) \tag{3}$$
$$x_{n+1} := x_n^2 + c. \tag{4}$$

Then the substitution $y = ax + b$, (i.e. $y_n = ax_n + b$ and $y_{n+1} = ax_{n+1} + b$), where a, b, c are chosen as above, turns (3) into (4). This would certainly cause anyone studying these as difference equations to remark that they are "equivalent".

Notice that to show two systems are topologically conjugate, all we need do is to find a topological conjugacy between them. On the other hand, to show two systems are *not* topologically conjugate, we must show that there is no topological conjugacy between them, and this is best done by finding a property, possessed by one but not by the other, which would have been preserved by a topological conjugacy.

Example 9.5 Consider the following dynamical systems $(f_i, \mathbb{R})$:

$$\begin{aligned} f_1(x) &= x \qquad (x \in \mathbb{R}) \\ f_2(x) &= -x \qquad (x \in \mathbb{R}) \\ f_3(x) &= x^2 \qquad (x \in \mathbb{R}) \\ f_4(x) &= -x^2 \qquad (x \in \mathbb{R}). \end{aligned}$$

Then

1. in $(f_1, \mathbb{R})$ every point is fixed;
2. in $(f_2, \mathbb{R})$ only one point is fixed, all the others being periodic of order 2;
3. in $(f_3, \mathbb{R})$ precisely two points are fixed and there are no other periodic points;
4. in $(f_4, \mathbb{R})$ precisely two points are fixed and there are no other periodic points.

The number of fixed points is invariant under topological conjugacy, so while there *might* (or might not) be a topological conjugacy between $(f_3, \mathbb{R})$ and $(f_4, \mathbb{R})$, there cannot be any other topological conjugacies between these four systems.

In fact, $(f_3, \mathbb{R})$ and $(f_4, \mathbb{R})$ *are* topologically conjugate, by the topological conjugacy $h : x \mapsto -x$. The map h is obviously a homeomorphism and

$$hf_3(x) = h(x^2) = -x^2 = -(-x)^2 = -(h(x))^2 = f_4(h(x).$$

Of course, we could have considered the map $h : x \mapsto -x$ as a candidate for a topological conjugacy between $(f_1, \mathbb{R})$ and $(f_2, \mathbb{R})$ and shown that $hf_1(x) = -x$ but $f_2h(x) = +x$, but this would not help as it would not preclude the possibility that some other map might be a topological conjugacy between those systems.

Exercise 9.6 For $q \neq 0$, transform $x_{n+1} := px_n^2 + qx_n + r$ into the form $y_{n+1} := \mu y_n(1 - y_n)$ by a substitution $y = ax + b$. (The manipulation is messy — use MAPLE.)

Example 9.7 We define the "tent map", $T : [0, 1] \to [0, 1]$ by

$$T(x) = \begin{cases} 2x & (0 \leqslant x \leqslant \frac{1}{2}) \\ 2 - 2x & (\frac{1}{2} \leqslant x \leqslant 1) \end{cases}$$

It is easily proved by induction that

$$T^n(x) = \begin{cases} 0 & (x = 2k2^{-n},\ 0 \leqslant k \leqslant 2^{n-1}) \\ 1 & (x = (2k-1)2^{-n},\ 1 \leqslant k \leqslant 2^{n-1}) \end{cases}$$

and T^n is linear in between these points. It follows that the diagonal cuts the graph of T^n precisely once in each interval $[k2^{-n}, (k+1)2^{-n})$ for $0 \leqslant k \leqslant 2^n - 1$. Thus T^n has precisely 2^n fixed points, so T has precisely 2^n periodic points of period n. Further, there is a periodic point of T in every interval $[k2^{-n}, (k+1)2^{-n})$. But every interval $(a, b) \subseteq [0, 1]$ contains such a subinterval, for sufficiently large n and suitable k, and so contains a periodic point of T. Thus the periodic points of T are dense in $[0, 1]$.

Consider the map $F_4 : [0, 1] \to [0, 1]$, $F_4(x) = 4x(1 - x)$. The maps F_4 and T are topologically conjugate by the conjugacy

$$h(x) = \left(\sin \frac{\pi}{2} x\right)^2 \quad (x \in [0, 1]),$$

since

$$\begin{aligned}
(F_4 \circ h)(x) &= 4\left(\sin \frac{\pi}{2}x\right)^2 \left(1 - \left(\sin \frac{\pi}{2}x\right)^2\right) \\
&= 4\left(\sin \frac{\pi}{2}x\right)^2 \left(\cos \frac{\pi}{2}x\right)^2 \\
&= (\sin \pi x)^2 \\
&= (\sin(\pi - \pi x))^2 \\
&= \begin{cases} \left(\sin \frac{\pi}{2} 2x\right)^2 & (0 \leqslant x \leqslant \frac{1}{2}) \\ \left(\sin \frac{\pi}{2}(2 - 2x)\right)^2 & (\frac{1}{2} \leqslant x \leqslant 1) \end{cases} \\
&= (h \circ T)(x).
\end{aligned}$$

It follows that F_4 has 2^n periodic points of period n and that the periodic points of F_4 are dense in [0,1].

10 Holomorphic dynamics

A function $f : \mathbb{C} \to \mathbb{C}$ is said to be *holomorphic* or *analytic* if it is locally defined by a Taylor series. In particular, every polynomial is everywhere holomorphic. This chapter concerns some of the more famous aspects of the theory of the dynamics of holomorphic functions.

If $f : \mathbb{C} \to \mathbb{C}$ is holomorphic everywhere and injective, then it may be shown by complex function theory that f is of the form $f(z) = az + b$ $(z \in \mathbb{C})$ for some $a, b \in \mathbb{C}$. The dynamics of these maps are trivial. We therefore proceed to consider two-to-one maps—i.e. maps f for which $f^{-1}(\{z\})$ has at most two elements for each $z \in \mathbb{C}$. It may be shown that the two-to-one maps which are holomorphic everywhere are precisely the quadratic polynomials. A topological conjugacy of the form $h(z) = cz + d$ then allows us to reduce consideration to just those quadratic polynomials of the form $Q_c(z) = z^2 + c$ for some $c \in \mathbb{C}$. We recall from an earlier chapter that, for real c, z, the mappings Q_c with $c \leqslant 1/4$ are topologically conjugate to the F_μ with $\mu \geqslant 1$, the value of c corresponding to a given value of μ being given by

$$c = \frac{\mu(2 - \mu)}{4}.$$

We reproduce the table showing key values of μ and the corresponding values of c.

μ	c
1	1/4
2	0
3	-3/4
$1+\sqrt{6}$	-5/4
$1+\sqrt{8}$	-7/4
4	-2

We begin studying the dynamics of the maps $Q_c : \mathbb{C} \to \mathbb{C}$ by considering the case $c = 0$: i.e., we are looking at the map $z \mapsto z^2$. On the circle $\mathbb{T} = \{z : |z| = 1\}$, this is simply the angle-doubling map discussed in Chapter 1 (Example 1.7). If $|z| > 1$ then $|Q_0^n(z)| = |z^{2^n}| \to \infty$ as $n \to \infty$. If $|z| < 1$ then $Q_0^n(z) \to 0$ as $n \to \infty$.

The theory of attracting and repelling periodic points for holomorphic functions is similar to that for continuously differentiable real functions. Thus if $f : \mathbb{C} \to \mathbb{C}$ is holomorphic with a fixed point p and if $f'(p)$ denotes the (complex) derivative of f at p, then we say that p is *repelling* if $|f'(p)| > 1$, *non-hyperbolic* if $|f'(p)| = 1$, *attracting* if $|f'(p)| < 1$ and *superattracting* if $f'(p) = 0$. If p is a periodic point of order m, we say p is *attracting*, etc. if $|(f^m)'(p)| < 1$, etc.. As in the real case, 'attracting' implies 'weakly attracting' (Ernst Schröder, 1870) and 'repelling' implies 'weakly repelling'. (The terms 'weakly attracting' and 'weakly repelling' are, of course, defined for any continuous function f on any subset of $\mathbb{R}^n$ and we regard $\mathbb{C}$ as a subset of $\mathbb{R}^2$.)

For the function Q_0, we have $Q_0'(z) = 2z$, so $Q_0'(0) = 0$ and the fixed point 0 is (super)attracting, and therefore weakly attracting. For the fixed point 1, we have $Q_0'(1) = 2$, so it is repelling. The other periodic points are all on the unit circle. Let p be any one of these, or order m, with orbit $\{p_1, p_2, \ldots, p_m\}$, say. Then, applying the Chain Rule, as before,

$$|(Q_0^m)'(p)| = |Q_0'(p_1) \ldots Q_0'(p_m)| = |2p_1 \ldots 2p_m| = 2^m > 1,$$

since $|p_i| = 1$ $(1 \leqslant i \leqslant m)$. Thus the other periodic points are all repelling, and they are dense in the circle. It follows from these remarks that the set where the interesting dynamics takes place, is $\mathbb{T}$, and we can characterize it as the closure of the repelling periodic points. Recall that the *closure* $\bar{A}$ of a set A is defined by

$$\bar{A} := \{x \in X : \forall \varepsilon > 0 \quad B(x; \varepsilon) \cap A \neq \emptyset\}.$$

Alternatively, looking back to our discussion in Chapter 1, Example 1.7 of the eventually fixed points of the angle-doubling map, we can see that $\mathbb{T}$ is the closure of the set of all iterated inverse images of the repelling fixed point 1. (By the *set of all iterated inverse images* of a point p under a function f, we mean the set

$$\{p\} \cup f^{-1}(p) \cup f^{-1}(f^{-1}(p)) \cup \ldots,$$

that is, the set of all x such that, for some $n = n(x)$, we have $f^n(x) = p$.)

Let us now ask: for which values of c does Q_c have an attracting fixed point?

If p is a fixed point of Q_c, then $Q_c(p) = p$; i.e.

$$p^2 + c = p. \tag{5}$$

Since $Q_c'(p) = 2p$, the fixed point p is attracting if and only if $|p| < \frac{1}{2}$. Therefore the set M_0 of points c for which Q_c has an attracting fixed point is

$$M_0 = \{c : c = p - p^2, \quad |p| < \frac{1}{2}\} = g(D),$$

where $g : z \mapsto z - z^2$ and $D = B(0; 1/2)$.

The set $g(\partial D)$ (the image of the boundary of D) is the curve

$$c = \frac{1}{2}e^{i\theta} - \frac{1}{4}e^{2i\theta} \quad (0 \leqslant \theta \leqslant 2\pi). \tag{6}$$

This is a *cardioid*: imagine a fixed disc, centre 0, radius $\frac{1}{4}$, and a moving disc of the same radius which starts with its centre at $\frac{1}{2}$ and rolls round the fixed disc in an anticlockwise direction. Mark a point P on the circumference of the moving disc so that P is initially at the point $\frac{1}{4}$. The locus traced out by P is the cardioid (6), the variable θ representing the argument of the point of contact between the two discs.

Since g is continuous and injective on D, it follows that $g(D)$ consists of the region 'inside' the cardioid. This is intuitively 'obvious', but hard to prove. We shall give a partial justification, but not a complete proof.

First, we observe that g is injective on $\overline{D}$; for if $p - p^2 = q - q^2$, then $(p - q)(1 - p - q) = 0$, so either $p = q$ or $p + q = 1$. However, if $|p| \leqslant 1/2$ and $|q| \leqslant 1/2$, then $p + q = 1$ implies $p = q = 1/2$. Thus, in either case, $p = q$.

To show that every point of $g(D)$ lies inside the cardioid, we begin by noting that $g(0) = 0$ which is inside the cardioid. Suppose there is some z with $|z| < 1/2$ and $g(z)$ outside the cardioid. Then the straight line from 0 to z would be mapped by g to a curve joining 0 and $g(z)$, which would have to meet the cardioid at a point $g(z_1) = g(z_2)$ where z_1 is on the line from 0 to z and $z_2 \in \partial D$. Therefore $|z_1| < |z_2|$, so $z_1 \neq z_2$ contradicting the injectivitiy of g. (This is not quite a formal proof, since it relies implicitly on the (rather deep) 'Jordan Curve Theorem' which, roughly, asserts that a simple closed curve has an 'inside' and an 'outside'.)

The converse assertion, that every point inside the cardioid is in $g(D)$, we shall also dismiss as intuitively obvious: because as r increases from 0 to 1/2 the curve $g(re^{i\theta})$ $(0 \leqslant \theta \leqslant 2\pi)$ changes continuously from the point 0 to the cardioid and 'so' sweeps out all the points in between. (This argument too can be replaced by a formal proof using suitable theorems on topology.)

With these provisos, we have proved the following.

Theorem 10.1 *The function $Q_c : \mathbb{C} \to \mathbb{C}$ has an attracting fixed point if and only if c lies strictly inside the cardioid (6).*

For c strictly inside this cardioidal region, the dynamics of Q_c look essentially the same as those of Q_0. The closure of the set of repelling periodic points is a set J_c, the *Julia set*, which is homeomorphic to a circle, though it looks crinkly—for $c \neq 0$, it has fractional "Hausdorff dimension" (see PMA443 'Fractals'). It is an example of a type of set called a "quasi-circle" by virtue of its "quasi-self-similarity" properties (the precise definition of which is rather complicated). Again J_c is the closure of the set of iterated inverse images of the (unique) repelling fixed point. The latter characterization is used to produce a computer picture of J_c.

A similar, but more complicated, calculation shows that the set of c for which Q_c has an attracting periodic point of order 2 is the disc

$$\{c : |c + 1| < \frac{1}{4}\}.$$

Other regions corresponding to the existence of attracting periodic points of various orders may be found. These are called the *hyperbolic regions of the Mandelbrot set.* In each case, when c is in such a region, the closure J_c of the set of repelling periodic points is connected and J_c is the closure of the set of all iterated inverse images of one of the repelling fixed points. Note that Q_c has two fixed points, p_1, p_2, the roots of (5), which coincide only if $c = 1/4$. If c lies inside the cardioidal region, one fixed point is attracting and the other repelling, since $p_1 + p_2 = -1$ (the sum of the roots of the quadratic (5)), so $|p_1| < \frac{1}{2}$ implies $|p_2| > \frac{1}{2}$. Strictly outside the cardioid, both are repelling and, in this case, taking the closure of iterated inverse images, starting from either of the fixed points, yields the set J_c.

Theorem 10.2 (Fatou's Critical Points Theorem, 1905) *For every holomorphic function f, every attracting periodic orbit of f attracts at least one critical point of f (point z with $f'(z) = 0$).*

For $f = Q_c$, there is just one critical point, namely 0, so we have the following corollaries

Corollary 10.3 *For each c, the function Q_c has at most one attracting periodic orbit.*

Corollary 10.4 *If Q_c has an attracting periodic orbit, the the sequence $(Q_c^n(0))_{n=1}^{\infty}$ tends to this orbit and, in particular, this sequence is bounded.*

For every $c \in \mathbb{C}$, we define the *Julia set* J_c to be the closure of the set of all repelling periodic points. It may be shown that J_c equals the closure of the set of all iterated inverse images of either one of the repelling fixed points, and it is easy to see that $Q_c(J_c) \subseteq J_c$. For general c, however, particularly when c is far from 0, the set J_c need not be connected.

Let K_c denote the *filled-in Julia set*

$$K_c = \{z : |Q_c^n(z)| \not\to \infty\}.$$

For example, $K_0 = \{z : |z| \leqslant 1\}$. Generally, it can be shown that $J_c \subseteq K_c$ and $\mathbb{C} \setminus K_c$ is connected.

Proposition 10.5 *If p is an attracting periodic point of Q_c, then $0 \in B(0;\delta) \subseteq W^s(p) \subseteq K_c \setminus J_c$ for some $\delta > 0$.*

Definition 10.6 The *Mandelbrot set* M is the set of all c such that J_c is connected.

Plotting Mandelbrot and Julia sets on the computer screen
The following two theorems give some ways of plotting these sets.

Theorem 10.7 *The set J_c is the closure of the set of all iterated inverse images of either one of the repelling fixed points of Q_c.*

[The fixed points of Q_c are determined by solving the quadratic equation $z = z^2 + c$. A fixed point p is repelling iff $|2p| = |Q_c'(p)| > 1$. Inverse images are computed by $z \mapsto \pm\sqrt{z-c}$. Plotting the iterated inverse images gives, for practical purposes, the plot of the J_c.]

Theorem 10.8 *(i) $Q_c^n(0) \to \infty$ iff $c \notin M$;*

(ii) $|Q_c^n(0)| \leqslant 2$ for all n iff $c \in M$.

[To plot M, we take each value of c in turn, one corresponding to each pixel on the screen, and compute the sequence $Q_c^n(0)$ $(1 \leqslant n \leqslant N)$ for some suitably large N. If, for some n, we find $|Q_c^n(0)|^2 > 4$ then $c \notin M$. Otherwise, we say $c \in M$ (this is precise only for $N = \infty$!).]

We omit the proof of (i), but prove that (ii) follows from (i). Clearly, if $|Q_c^n(0)| \leqslant 2$ for all n, then $Q_c^n(0) \not\to \infty$ and so $c \in M$, by (i). The more substantial result is that if $|Q_c^n(0)| > 2$ for some n, then $Q_c^n(0) \to \infty$, and so $c \notin M$, by (i).

We consider two cases.

Case $|c| > 2$

Since $Q_c(0) = c$ this means $|Q_c^n(0)| > 2$ for $n = 1$ and we have to show that $Q_c^n(0) \to \infty$. Suppose $|c| = 2 + \varepsilon$. Then we prove by induction on n that $|Q_c^n(0)| \geqslant 2 + n\varepsilon$ for $n = 1, 2, 3, \ldots$; whence $|Q_c^n(0)| \to \infty$ as $n \to \infty$.

At $n = 1$, we have $|Q_c(0)| = |c| = 2 + \varepsilon$, so the result holds. Suppose $|Q_c^n(0)| \geqslant 2 + n\varepsilon$. Then

$$\begin{aligned}
|Q_c^{n+1}(0)| &= |Q_c^n(0)^2 + c| \\
&\geqslant |Q_c^n(0)|^2 - |c| \\
&\geqslant (2 + n\varepsilon)^2 - (2 + \varepsilon) \\
&= 2 + (4n - 1)\varepsilon + n^2\varepsilon^2 \\
&\geqslant 2 + (n + 1)\varepsilon,
\end{aligned}$$

since $n \geqslant 1$ implies $4n - 1 \geqslant n + 1$. This completes the induction. Since $\varepsilon > 0$ we have $2 + n\varepsilon \to \infty$ and so $|Q_c^n(0)| \to \infty$.

Case $|c| \leqslant 2$

Suppose there exists N such that $|Q_c^N(0)| > 2$; say $|Q_c^N(0)| = 2 + \varepsilon$. We show by induction on r that $|Q_c^{N-1+r}(0)| \geqslant 2 + r\varepsilon$ for $r = 1, 2, 3, \ldots$; whence $|Q_c^n(0)| \to \infty$ as $n \to \infty$.

At $r = 1$, we have $|Q_c^N(0)| = 2 + \varepsilon$, so the result holds. Suppose $|Q_c^{N-1+r}(0)| \geqslant 2 + r\varepsilon$. Then

$$\begin{aligned}
|Q_c^{N-1+r+1}(0)| &= |Q_c^{N-1+r}(0)^2 + c| \\
&\geqslant |Q_c^{N-1+r}(0)|^2 - |c| \\
&\geqslant (2 + r\varepsilon)^2 - 2 \\
&= 2 + 4r\varepsilon + r^2\varepsilon^2 \\
&\geqslant 2 + (r + 1)\varepsilon,
\end{aligned}$$

since $r \geqslant 1$ implies $4r \geqslant r + 1$. This completes the induction. Thus $|Q_c^n(0)| \geqslant 2 + (n - N + 1)\varepsilon$ for all $n \geqslant N$. Since $\varepsilon > 0$ we have $2 + (n - N + 1)\varepsilon \to \infty$ and so $|Q_c^n(0)| \to \infty$.

Corollary 10.9 $M \subseteq \{c \in \mathbb{C} : |c| \leqslant 2\}$.

This follows from the first case in the above proof.

Corollary 10.10 $M = \bar{M}$ *(i.e.* M *is 'closed').*

The proof is omitted, but is an easy exercise for students of metric spaces (since $M = \bigcap_{n=1}^{\infty} \{c : |Q_c^n(0)| \leqslant 2\}$).

Combining Theorem 10.8 (i) with Corollary 10.4, we have the following.

Corollary 10.11 *If* Q_c *has an attracting periodic orbit, then* $c \in M$.

The converse is false.

The structure of M is unusually complicated. This is, perhaps, not unexpected, since we know that the bifurcation diagram of F_μ and so of Q_c for $c \in (-\infty, \frac{1}{4})$ is very complicated. Travelling down the real axis through the Mandelbrot set, we pass first through the cardioidal region, corresponding to an attracting fixed point, then, from $-\frac{3}{4}$ to $-\frac{5}{4}$ through the region corresponding to an attracting periodic point of order 2, then through the region corresponding to an attracting periodic point of order 4, and so on. Later, we come to the period 3 window, with its corresponding bifurcation sequence, and we find ourselves passing through a small "copy" (approximately) of M corresponding to this. Moving out from the cardioidal region in other directions takes us along curved paths qualitatively similar to the real 'spike', with, likewise small 'copies' of M. At first sight, on a computer screen, these look like islands, but in fact an important theorem of Douady and Hubbard (1982) tells us that they are not.

Theorem 10.12 *The Mandelbrot set is connected.*

Definition 10.13 A point $c \in M$ is a *Misiurewicz point* if 0 is eventually periodic for Q_c.

Example 10.14 If $c = -2$ then Q_c maps

$$0 \mapsto -2 \mapsto 2 \mapsto -2 \ldots.$$

If $c = i$ then Q_c maps

$$0 \mapsto i \mapsto -1 + i \mapsto -i \mapsto -1 + i \ldots.$$

Theorem 10.15 *If* c *is a Misiurewicz point, then* $J_c = K_c$.

In this case, (i.e. J_c connected and $J_c = K_c$,) we say that J_c is a *dendrite*.

What happens when $c \notin M$? For c real and $c < -2$, Q_c on $\mathbb{R}$ is topologically conjugate to some F_μ with $\mu > 4$. Such an F_μ has its interesting dynamics confined to a set homeomorphic to the Cantor set on which it has the easily describable dynamics. Therefore, Q_c will also exhibit this behaviour, and in fact the set on which its interesting dynamics occurs is precisely the Julia set J_c. The same is true for all $c \notin M$: the interesting dynamics occurs on J_c.

A great deal more can be said about the mappings Q_c, their Julia sets and the Mandelbrot set, but lack of time (and two or three courses in complex analysis) forces us to stop here.

Appendix: a simple chaos program

We can study the effect of iterating these maps by the following QBASIC program on a PC.

```
10 X=0.4137
20 INPUT "MU=";MU
30 PRINT X
40 PITCH = 220 * 4 ^ X
50 SOUND PITCH, 4
60 X=MU*X*(1-X)
70 GOTO 30
```

Notes.

Line 10 sets X to a value chosen randomly; almost any $X \in (0, 1)$ will do.

Line 20 inputs the parameter μ.

Line 30 prints the value of X on the screen.

Line 40 interprets X as a pitch, measured in Hertz.

Line 50 creates the sound, at the chosen pitch, for a time 4, measured in units of 1/18.2 seconds.

Line 60 iterates the function to get the next value of X.

Line 70 loops back. The program continues indefinitely until the ESCAPE key is pressed.

This program may be used on any computer with a BASIC interpreter/compiler, with appropriate modifications to lines 40 and 50.

Suitable values of μ are:-

2.5 rapid convergence to a fixed point;

2.9 slow convergence to a fixed point;

3.3 period 2;

3.5 period 4;

3.56 period 8;

3.57 chaos; the limit of the 2^n cycles;

3.74 period 5;

3.83 period 3;

3.90 chaotic;

3.906 period 5;