2.1 Affine equivalence

An *affine transformation* of the real line is a transformation of the form

$$x \mapsto h(x) = Ax + B$$

where A and B are real constants with $A \neq 0$. So an affine transformation consists of a change of scale (and possibly direction if $A < 0$) given by the factor A, followed by a shift of the origin given by B. In the study of linear phenomena, we expect that the essentials of an object be invariant under a change of scale and a shift of the origin of our coordinate system.

For example, consider the logistic transformation, $L_\mu(x) = \mu x(1-x)$ and the affine transformation

$$h_\mu(x) = -\mu x + \frac{\mu}{2}.$$

We claim that

$$h_\mu \circ L_\mu \circ h_\mu^{-1} = Q_c \tag{2.1}$$

where

$$Q_c(x) = x^2 + c \tag{2.2}$$

and where c is related to μ by the equation

$$c = -\frac{\mu^2}{4} + \frac{\mu}{2}. \tag{2.3}$$

In other words, we are claiming that if c and μ are related by (2.3) then we have

$$h_\mu(L_\mu(x)) = Q_c(h_\mu(x)).$$

To check this, the left hand side expands out to be

$$-\mu[\mu x(1-x)] + \frac{\mu}{2} = \mu^2x^2 - \mu^2 x + \frac{\mu}{2},$$

while the right hand side expands out as

$$(-\mu x + \frac{\mu}{2})^2 - \frac{\mu^2}{4} + \frac{\mu}{2} = \mu^2x^2 - \mu^2 x + \frac{\mu}{2}$$

giving the same result as before, proving (2.1).

We say that the transformations L_μ and $Q_c, c = -\frac{\mu^2}{4} + \frac{\mu}{2}$ are *conjugate* by the affine transformation, h_μ.

More generally, let $f : X \to X$ and $g : Y \to Y$ be maps of the sets X and Y to themselves, and let $h : X \to Y$ be a one to one map of X onto Y. We say that h conjugates f into g if

$$h \circ f \circ h^{-1} = g,$$

or, what amounts to the same thing, if

$$h \circ f = g \circ h.$$

We shall frequently write this equation in the form of a *commutative diagram*

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ h \downarrow & & \downarrow h \\ Y & \xrightarrow[g]{} & Y \end{array}$$

The statement that the diagram is commutative means that going along the upper right hand path (so applying $h \circ f$) is equal to traversing the left lower path (which is $g \circ f$).

Notice that if $h \circ f \circ h^{-1} = g$, then

$$g^{circn} = h \circ f^{\circ n} \circ h^{-1}.$$

So the problem of studying the iterates of g is the same (up to the transformation h) as that of f, *providing* that the properties we are interested in studying are not destroyed by h.

Certainly affine transformations will always be allowed. Let us generalize the preceding computation by showing that *any* quadratic transformation (with non-vanishing leading term) is conjugate (by an affine transformation) to a transformation of the form Q_c for suitable c. More precisely:

Proposition 2.1.1 *Let $f = ax^2 + bx + d$ then f is conjugate to Q_c by the affine map $h(x) = Ax + B$ where*

$$A = a, \quad B = \frac{b}{2}, \quad \textit{and} \quad c = ad + \frac{b}{2} - \frac{b^2}{4}.$$

Proof. Direct verification.

Let us understand the importance of this result. The general quadratic transformation f depends on three parameters a, b and d. But if we are interested in the qualitative behavior of the iterates of f, it suffices to examine the one parameter family C_c. Any quadratic transformation (with non-vanishing leading term) has the same behavior (in terms of its iterates) as one of the Q_c. The family of possible behaviors under iteration is one dimensional, depending on a single parameter c. We may say that the family Q_c (or for that matter the family L_μ) is *universal* with respect to quadratic maps as far as iteration is concerned.

2.2 Conjugacy of T and L_4

Let $T : [0, 1] \to [0, 1]$ be the map defined by

$$T(x) = 2x, \; 0 \le x \le \frac{1}{2}, \quad T(x) = -2x + 2, \; \frac{1}{2} \le x \le 1.$$

So the graph of T looks like a tent, hence its name. It consists of the straight line segment of slope 2 joining $x = 0, y = 0$ to $x = \frac{1}{2}, y = 1$ followed by the segment of slope -2 joining $x = \frac{1}{2}, y = 1$ to $x = 1, y = 0$.

Of course, here L_4 is our old friend, $L_4(x) = 4x(1-x)$. We wish to show that

$$L_4 \circ h = h \circ T$$

where

$$h(x) = \sin^2\left(\frac{\pi x}{2}\right).$$

In other words, we claim that the diagram of section 1 commutes when $f = T$, $g = L_4$ and h is as above. The function $\sin\theta$ increases monotonically from 0 to 1 as θ increases from 0 to $\pi/2$. So, setting

$$\theta = \frac{\pi x}{2},$$

we see that $h(x)$ increases monotonically from 0 to 1 as x increases from 0 to 1. It therefore is a one to one continuous map of $[0, 1]$ onto itself, and thus has a continuous inverse. It is differentiable everywhere with $h(x) > 0$ for $0 < x < 1$. But $h'(0) = h'(1) = 0$. So h^{-1} is not differentiable at the end points, but is differentiable for $0 < x < 1$.

To verify our claim, we substitute

$$\begin{aligned} L_4(h(x)) &= 4\sin^2\theta(1-\sin^2\theta) \\ &= 4\sin^2\theta\cos^2\theta \\ &= \sin^2 2\theta \\ &= \sin^2 \pi x. \end{aligned}$$

So for $0 \le x \le \frac{1}{2}$ we have verified that

$$L_4(h(x)) = h(2x) = h(T(x)).$$

For $\frac{1}{2} < x \le 1$ we have

$$\begin{aligned} h(T(x)) &= h(2-2x) \\ &= \sin^2(\pi - \pi x) \\ &= \sin^2 \pi x \\ &= \sin^2 2\theta \\ &= 4\sin^2\theta(1-\sin^2\theta) \\ &= L_4(h(x)) \end{aligned}$$

where we have used the fact that $\sin(\pi - \alpha) = \sin\alpha$ to pass from the second line to the third. So we have verified our claim in all cases.

Many interesting properties of a transformation are preserved under conjugation by a homeomorphism. (A *homeomorphism* is a bijective continuous map

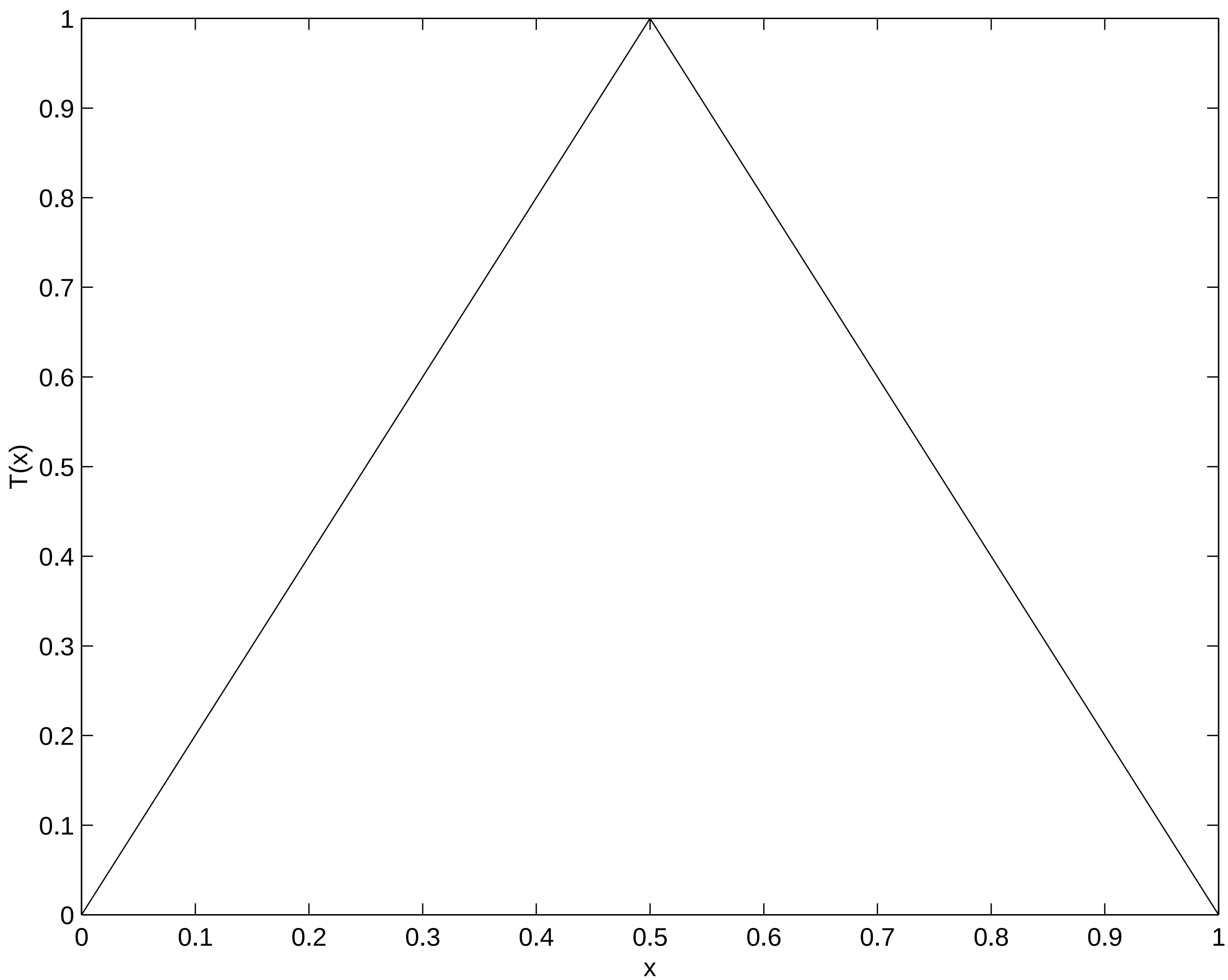

Figure 2.1: The tent map.

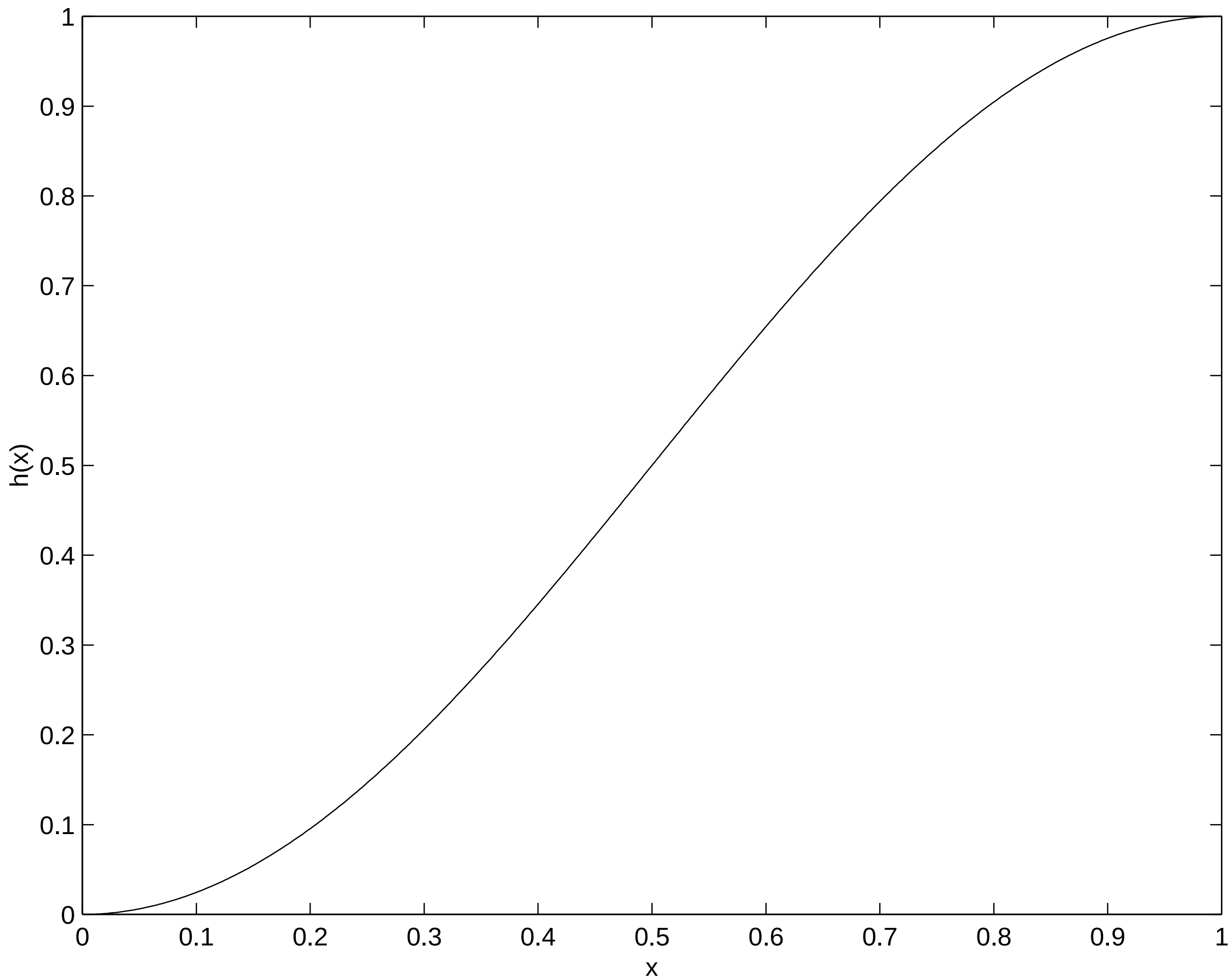

Figure 2.2: $h(x) = \sin^2\left(\frac{\pi x}{2}\right)$.

with continuous inverse.) For example, if p is a periodic point of period n of f, so that $f^{\circ n}(p) = p$, then

$$g^{\circ n}(h(p)) = h \circ f^{\circ n}(p) = h(p)$$

if $h \circ f = g \circ h$. So periodic points are carried into periodic points of the same period under a conjugacy. We will consider several other important properties of a transformation as we go along, and will prove that they are invariant under conjugacy. So what our result means is that if we prove these properties for T, we conclude that they are true for L_μ. Since we have verified that L_4 is conjugate to Q_{-2}, we conclude that they hold for Q_{-2} as well.

Here is another example of a conjugacy, this time an affine conjugacy. Consider

$$V(x) = 2|x| - 2.$$

V is a map of the interval $[-2, 2]$ into itself. Consider

$$h_2(x) = 2 - 4x.$$

So $h_2(0) = 2, h_2(1) = -2$. In other words, h_2 maps the interval $[0, 1]$ in a one to one fashion onto the interval $[-2, 2]$. We claim that

$$V \circ h_2 = h_2 \circ T.$$

Indeed,

$$V(h_2(x)) = 2|2 - 4x| - 2.$$

For $0 \leq x \leq \frac{1}{2}$ this equals $2(2 - 4x) - 2 = 2 - 8x = 2 - 4(2x) = h_2(Tx)$. For $\frac{1}{2} \leq x \leq 1$ we have $V(h_2(x)) = 8x - 6 = 2 - 4(2 - 2x) = h_2(Tx)$. So we have verified the required equation in all cases. The effect of the affine transformation, h_2 is to enlarge the graph of T, shift it, and turn it upside down. But as far as iterations are concerned, these changes do not effect the essential behavior.

2.3 Chaos

A transformation F is called (topologically) *transitive* if for any two open (non empty) intervals, I and J, one can find initial values in I which, when iterated, will eventually take values in J. In other words, we can find an $x \in I$ and an integer n so that $F^n(x) \in J$.

For example, consider the tent transformation, T. Notice that T maps the interval $[0, \frac{1}{2}]$ onto the entire interval $[0, 1]$, and also maps the interval $[\frac{1}{2}, 1]$ onto the entire interval, $[0, 1]$. So $T^{\circ 2}$ maps each of the intervals $[0, \frac{1}{4}], [\frac{1}{4}, \frac{1}{2}], [\frac{1}{2}, \frac{3}{4}]$ and $[\frac{3}{4}]$ onto the entire interval $[0, 1]$. More generally, $T^{\circ n}$ maps each of the 2^n intervals $[\frac{k}{2^n}, \frac{k+1}{2^n}]$, $0 \leq k \leq 2^n - 1$ onto the entire interval $[0, 1]$. But any open interval I contains some interval of the form $[\frac{k}{2^n}, \frac{k+1}{2^n}]$ if we choose n sufficiently large. For example it is enough to choose n so large that $\frac{3}{2^n}$ is less than the length of I. So for this value on n, $T^{\circ n}$ maps I onto the entire interval $[0, 1]$, and so, in particular, there will be points, x, in I with $F(x) \in J$.

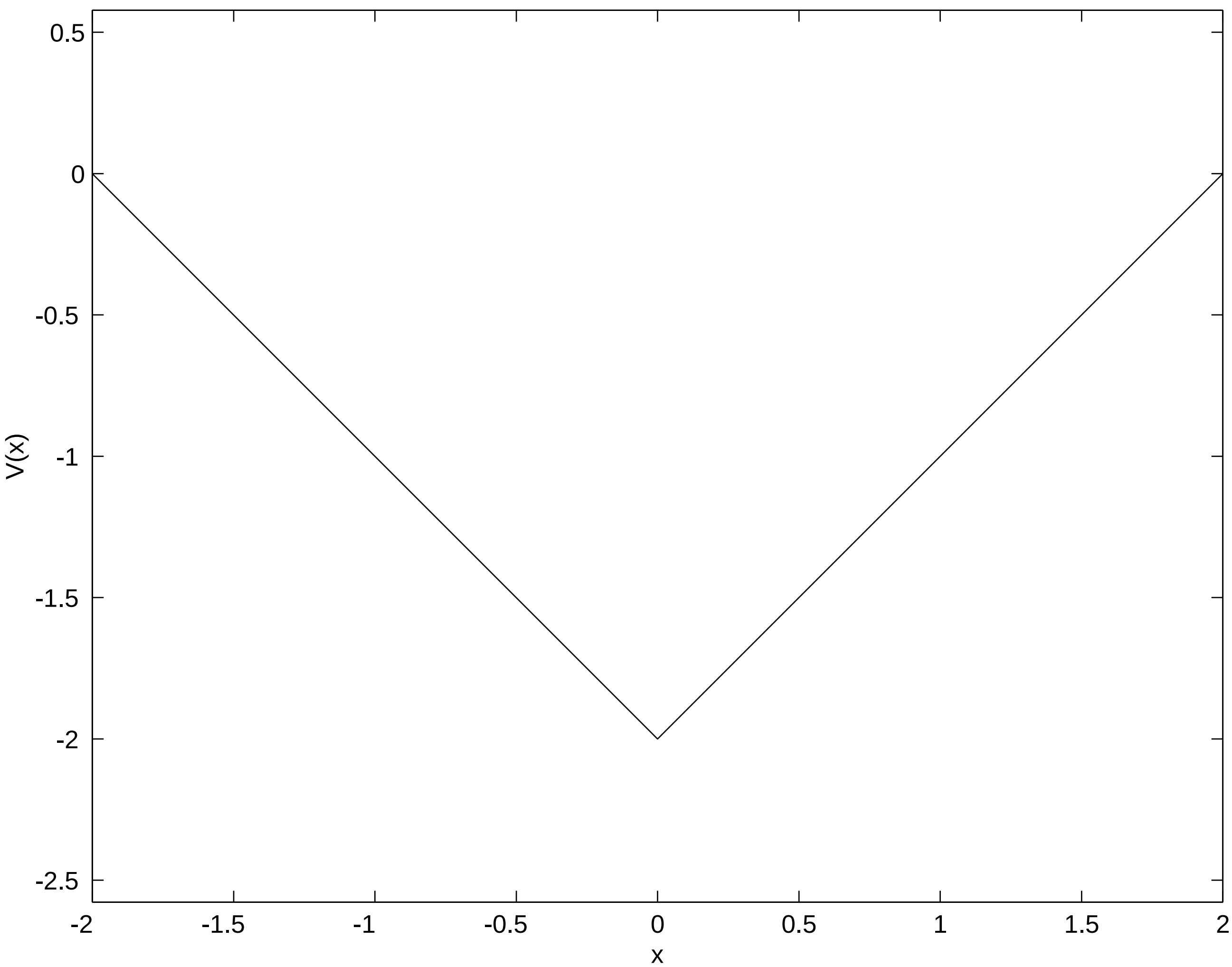

Figure 2.3: $V(x) = 2|x| - 2$.

Proposition 2.3.1 *Suppose that $g \circ h = h \circ f$ where h is continuous and surjective, and suppose that f is transitive. Then g is transitive.*

Proof. We are given non-empty open I and J and wish to find an n and an $x \in I$ so that $g^{\circ n}(x) \in J$. To say h is continuous means that $h^{-1}(J)$ is a union of open intervals. To say that h is surjective implies that $h^{-1}(J)$ is not empty. Let L be one of the intervals constituting $h^{-1}(J)$. Similarly, $h^{-1}(I)$ is a union of open intervals. Let K be one of them. By the transitivity of f we can find an n and a $y \in K$ with $f^{\circ n}(y) \in L$. Let $x = h(y)$. Then $x \in I$ and $g^{\circ n}(x) = g^{\circ n}(h(y)) = h(f^{\circ n}(y)) \in h(L) \subset J$. QED.

As a corollary we conclude that if f is conjugate to g, then f is transitive if and only if g is transitive. (Just apply the proposition twice, once with the roles of f and g interchanged.) But in the proposition we did not make the hypothesis that h was bijective or that it had a continuous inverse. We will make use of this more general assertion.

A set S of points is called *dense* if every non-empty open interval, I, contains a point of S. The behavior of density under continuous surjective maps is also very simple:

Proposition 2.3.2 *If $h : X \to Y$ is a continuous surjective map, and if D is a dense subset of X then $h(D)$ is a dense subset of Y.*

Proof. Let $I \subset Y$ be a non-empty open interval. Then $h^{-1}(I)$ is a union of open intervals. Pick one of them, K and then a point $y \in D \cap K$ which exists since D is dense. But then $f(y) \in f(D) \cap I$. QED

We define PER(f) to be the set of periodic points of the map f. If $h \circ f = g \circ h$, then $f^{\circ n}(p) = p$ implies that $g^{\circ n}(h(p)) = h(f^{\circ n}(p)) = h(p)$ so

$$h[\mathrm{PER}(f)] \subset \mathrm{PER}(g).$$

In particular, if h is continuous and surjective, and if PER(f) is dense, then so is PER(g).

Following Devaney and recent work (1992) by J. Banks et.al. *Amer. Math. Monthly* **99** (1992) 332-334, let us call f **chaotic** if f is transitive and PER(f) is dense. It follows from the above discussion that

Proposition 2.3.3 *If $h : X \to Y$ is surjective and continuous, if $f : X \to X$ is chaotic, and if $h \circ f = g \circ h$, then g is chaotic.*

We have already verified that the tent transformation, T, is transitive. We claim that PER(T) is dense on $[0, 1]$ and hence that T is chaotic. To see this, observe that T^n maps the interval $[\frac{k}{2^n}, \frac{k+1}{2^n}]$ onto $[0, 1]$. In particular, there is a point $x \in [\frac{k}{2^n}, \frac{k+1}{2^n}]$ which is mapped into itself. In other words, every interval $[\frac{k}{2^n}, \frac{k+1}{2^n}]$ contains a point of period n for T. But any non-empty open interval I contains an interval of the type $[\frac{k}{2^n}, \frac{k+1}{2^n}]$ for sufficiently large n. Hence T is chaotic.

From the above propositions it follows that L_4, Q_{-2}, and V are all chaotic.

2.4 The saw-tooth transformation and the shift

Define the function S by

$$S(x) = 2x, \quad 0 \le x < \frac{1}{2}, \qquad S(x) = 2x - 1, \quad \frac{1}{2} \le x \le 1. \tag{2.4}$$

The map S is discontinuous at $x = .5$. However, we can find a continuous, surjective map, h, such that $h \circ S = T \circ h$. In fact, we can take h to be T itself! In other words we claim that

$$\begin{array}{ccc} I & \xrightarrow{S} & I \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ I & \xrightarrow[T]{} & I \end{array}$$

commutes where $I = [0, 1]$. To verify this, we successively compute both $T \circ T$ and $T \circ S$ on each of the quarter intervals:

$$\begin{array}{lclcll}
T(T(x)) & = & T(2x) & = & 4x & \text{for } 0 \le x \le 0.25 \\
T(S(x)) & = & T(2x) & = & 4x & \text{for } 0 \le x \le 0.25 \\
T(T(x)) & = & T(2x) & = & -4x+2 & \text{for } 0.25 < x < 0.5 \\
T(S(x)) & = & T(2x) & = & -4x+2 & \text{for } 0.25 \le x < 0.5 \\
T(T(x)) & = & T(-2x+2) & = & 4x-2 & \text{for } 0.5 \le x \le 0.75 \\
T(S(x)) & = & T(2x-1) & = & 4x-2 & \text{for } 0.5 \le x \le 0.75 \\
T(T(x)) & = & T(-2x+2) & = & -4x+4 & \text{for } 0.75 < x \le 1 \\
T(S(x)) & = & T(2x-1) & = & -4x+4 & \text{for } 0.75 < x \le 1
\end{array}$$

The h that we are using (namely $h = T$) is not one to one. That is why our diagram can commute even though T is continuous and S is not.

We now give an alternative description of the saw-tooth function which makes it clear that it is chaotic. Let X be the set of infinite (one sided) sequences of zeros and ones. So a point of X is a sequence $\{a_1 a_2 a_3 \ldots\}$ where each each a_i is either 0 or 1. However we exclude all points with a tail consisting of infinite repeating $1'$s. So a sequence such as $\{00111111111\ldots\}$ is excluded. We will identify X with the half open interval $[0, 1)$ by assigning to each point $x \in [0, 1)$ its binary expansion, and by assigning to each sequence $a = \{a_1 a_2 a_3 \ldots\}$ the number

$$h(a) = \sum \frac{a_i}{2^i}.$$

The map $h : X \to [0, 1)$ just defined is clear. The inverse map, assigning to each real number between 0 and 1 its binary expansion deserves a little more discussion: Take a point $x \in [0, 1)$. If $x < \frac{1}{2}$ the first entry in its binary expansion is 0. If $\frac{1}{2} \le x$ then the first entry in the binary expansion of x is 1. Now apply S. If $S(x) < \frac{1}{2}$ (which means that either $0 \le x < \frac{1}{4}$ or $\frac{1}{2} \le x < \frac{3}{4}$) then the second entry of the binary expansion of x is 0, while if $\frac{1}{2} \le S(x) < 1$

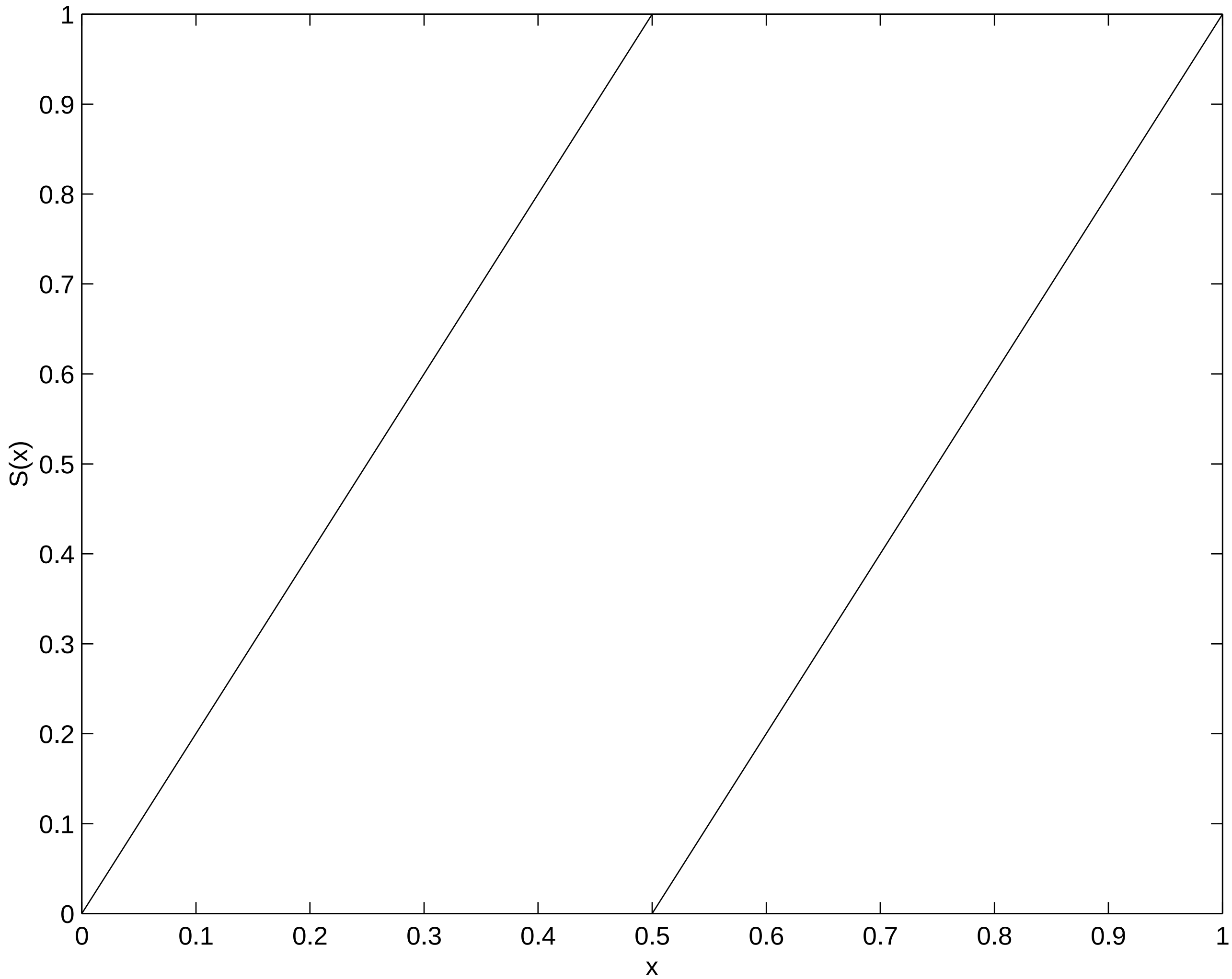

Figure 2.4: The discontinuous function S.

then the second entry in the binary expansion of x is 1. Thus the operator S provides the algorithm for the computation of the binary expansion of x. Let us consider, for example, $x = \frac{7}{16}$. Then the sequence $\{S^k(x)\}, k = 0, 1, 2, 3, \ldots$ is

$$\frac{7}{16}, \frac{7}{8}, \frac{3}{4}, \frac{1}{2}, 0, 0, 0, \ldots.$$

In general it is clear that for any number of the form $\frac{k}{2^n}$, after $n-1$ iterations of the operator S the result will be either 0 or $\frac{1}{2}$. So all $S^k(x) = 0, k \geq n$. In particular, no infinite sequence with a tail of repeating 1's can arise. We see that the binary expansion of $h(a)$ gives us a back, so we may (and shall) identify X with $[0, 1)$. Notice that we did not start with any independent notion of topology or metric on X. But now that we have identified X with $[0, 1)$, we can use use standard notions of distance on the unit interval but expressed in terms of properties of the sequences. For example, if the binary expansions of x and y agree up to the kth position, then

$$|x - y| < 2^{-k}.$$

So we define the distance between to sequences a and b to be 2^{-k} where k is the first place they do not agree. (Of course we define the distance from an a to itself to be zero.)

The expression of S in terms of the binary representation is very simple:

$$S : .a_1a_2a_3a_4 \cdots \mapsto .a_2a_3a_4a_5 \ldots.$$

It consists of throwing away the first digit and then shifting the entire sequence one unit to the left.

From this description it is clear that PER(S) consists of points with eventually repeating binary expansions, these are the rational numbers. They are dense. We can see that S is transitive as follows: We are given intervals I and J. Let $y = .b_1b_2b_3 \ldots$ be a point of J, and let $z = .a_1a_2a_3 \ldots$ be a point of I which is at a distance greater than 2^{-n} from the boundary of I. We can always find such a point if n is sufficiently large. Indeed, if we choose n so that the length of I is greater than $\frac{1}{2^{(n-1)}}$, the midpoint of I has this property. In particular, any point whose binary expansion agrees with z up to the $n-$th position lies in I. Take x to be the point whose first n terms in the binary expansion are those of z, followed by the binary expansion of y, so

$$x = 0.a_1a_2a_3 \ldots a_nb_1b_2b_3b_4 \ldots.$$

The point x lies in I and $S^n(x) = y$. Not only is S transitive, we can hit *any* point of J by applying S^n (with n fixed, depending only on I) to a suitable point of I. This is much more than is demanded by transitivity. Thus S is chaotic on $[0, 1)$.

Of course, once we know that S is chaotic on the open interval $[0, 1)$, we know that it is chaotic on the closed interval $[0, 1]$ since the addition on one extra point (which gets mapped to 0 by S) does not change the definitions.

Now consider the map $t \mapsto e^{2\pi it}$ of $[0,1]$ onto the unit circle, S^1. Another way of writing this map is to describe a point on the unit circle by $e^{i\theta}$ where θ is an angular variable, that is θ and $\theta + 2\pi$ are identified. Then the map is $t \mapsto 2\pi t$. This map, h, is surjective and continuous and is one to one except at the end points: 0 and 1 are mapped into the same point on S^1. Clearly

$$h \circ S = D \circ h$$

where

$$D(\theta) = 2\theta.$$

Or, if we write $z = e^{i\theta}$, then in terms of z, the map D sends

$$z \mapsto z^2.$$

So D is called the doubling map or the squaring map. We have proved that it is chaotic. We can use the fact that D is chaotic to give an alternative proof of the fact that Q_{-2} is chaotic. Indeed, consider the map $h : S^1 \to [-2, 2]$

$$h(\theta) = 2\cos\theta.$$

It is clearly surjective and continuous. We claim that

$$h \circ D = Q_{-2} \circ h.$$

Indeed,

$$h(D(\theta)) = 2\cos 2\theta = 2(2\cos^2\theta - 1) = (2\cos\theta)^2 - 2 = Q_{-2}(h(\theta)).$$

This gives an alternative proof that Q_{-2} (and hence L_4 and T) are chaotic.

2.5 Sensitivity to initial conditions

In this section we prove that if f is chaotic, then f is sensitive to initial conditions in the sense of the following

Proposition 2.5.1 (Sensitivity.) *Let $f : X \to X$ be a chaotic transformation. Then there is a $d > 0$ such that for any $x \in X$ and any open set J containing x there is a point $y \in J$ and an integer, n with*

$$|f^{\circ n}(x) - f^{\circ n}(y)| > d. \tag{2.5}$$

In other words, we can find points arbitrarily close to x which move a distance at least d away. This for any $x \in X$. We begin with a lemma.

Lemma 2.5.1 *There is a $c > 0$ with the property that for any $x \in X$ there is a periodic point p such that*

$$|x - f^{\circ k}(p)| > c, \quad \forall k.$$

Proof of lemma. Choose two periodic points, r and s with distinct orbits, so that $|f^{\circ k}(r) - f^{\circ l}(s)| > 0$ for all k and l. Choose c so that $2c < \min |f^{\circ k}(r) - f^{\circ l}(s)|$. Then for all k and l we have

$$\begin{aligned} 2c &< |f^{\circ k}(r) - f^{\circ l}(s)| \\ &= |f^{\circ k}(r) - x + x - f^{\circ l}(s)| \\ &\leq |f^{\circ k}(r) - x| + |f^{\circ l}(s) - x|. \end{aligned}$$

If x is within distance c to *any* of the points $f^{\circ}l(s)$ then it must be at a greater distance than c from *all* of the points $f^{\circ k}(r)$ and vice versa. So one of the two, r or s will work as the p for x.

Proof of proposition with $d = c/4$. Let x be any point of X and J any open set containing x. Since the periodic points of f are dense, we can find a periodic point q of f in

$$U = J \cap B_d(x),$$

where $B_d(x)$ denotes the open interval of length d centered at x,

$$B_d(x) = (x - d, x + d).$$

Let n be the period of q. Let p be a periodic point whose orbit is of distance greater than $4d$ from x, and set

$$W_i = B_d(f^{\circ i}(p)) \cap X.$$

Since $f^{\circ i}(p) \in W_i$, i.e. $p \in f^{-i}(W_i) = (f^{\circ i})^{-1}(W_i)$ for all i, we see that the open set

$$V = f^{-1}(W_1) \cap f^{-2}(W_2) \cap \cdots \cap f^{-n}(W_n)$$

is not empty.

Now we use the transitivity property of f applied to the open sets U and V. By assumption, we can find a $z \in U$ and a positive integer k such that $f^k(z) \in V$. Let j be the smallest integer so that $k < nj$. In other words,

$$1 \leq nj - k \leq n.$$

So

$$f^{nj}(z) = f^{nj-k}(f^k(z)) \in f^{nj-k}(V).$$

But

$$\begin{aligned} f^{nj-k}(V) &= f^{nj-k}\left(f^{-1}(W_1) \cap f^{-2}(W_2) \cap \cdots \cap f^{-n}(W_n)\right) \\ &\subset f^{nj-k}(f^{-(nj-k)}W_{nj-k}) \\ &= W_{nj-k}. \end{aligned}$$

In other words,

$$|f^{nj}(z) - f^{nj-k}(p)| < d.$$

On the other hand, $f^{nj}(q) = q$, since n is the period of q. Thus

$$\begin{aligned} |f^{nj}(q) - f^{nj}(z)| &= |q - f^{nj}(z)| \\ &= |x - f^{nj-k}(p) + f^{nj-k}(p) - f^{nj}(z) + q - x| \\ &\geq |x - f^{nj-k}(p)| - |f^{nj-k}(p) - f^{nj}(z)| - |q - x| \\ &\geq 4d - d - d = 2d. \end{aligned}$$

But this last inequality implies that either

$$|f^{nj}(x) - f^{nj}(z)| > d$$

or

$$|f^{nj}(x) - f^{nj}(q)| > d$$

for if x were within distance d from both of these points, they would have to be within distance $2d$ from each other, contradicting the preceding inequality. So one of the two, z or q will serve as the y in the proposition with $m = nj$.

2.6 Conjugacy for monotone maps

We begin this section by showing that if f and g are continuous strictly monotone maps of the unit interval $I = [0, 1]$ onto itself, and if their graphs are both strictly below (or both strictly above) the line $y = x$ in the interior of I, then they are conjugate by a homeomorphism. Here is the precise statement:

Proposition 2.6.1 *Let f and g be two continuous monotone strictly increasing functions defined on $[0, 1]$ and satisfying*

$$\begin{aligned} f(0) &= 0 \\ g(0) &= 0 \\ f(1) &= 1 \\ g(1) &= 1 \\ f(x) &< x \quad \forall x \neq 0, 1 \\ g(x) &< x \quad \forall x \neq 0, 1. \end{aligned}$$

Then there exists a continuous, monotone increasing function h defined on $[0, 1]$ with

$$h(0) = 0, \qquad h(1) = 1,$$

and

$$h \circ f = g \circ h.$$

Proof. Choose any point (x_0, y_0) in the open square

$$0 < x < 1, \quad 0 < y < 1.$$

If (x_0, y_0) is to be a point on the curve $y = h(x)$, then the equation $h \circ f = g \circ h$ implies that the point (x_1, y_1) also lies on this curve, where

$$x_1 = f(x_0), \quad y_1 = g(y_0).$$

By induction so will the points (x_n, y_n) where

$$x_n = f^n(x_0), \quad y_n = g^n(y_0).$$

By hypothesis

$$x_0 > x_1 > x_2 > ...,$$

and since there is no solution to $f(x) = x$ for $0 < x < 1$ the limit of the x_n, $n \to \infty$ must be zero. Also for the y_n. So the sequence of points (x_n, y_n) approaches $(0,0)$ as $n \to +\infty$. Similarly, as $n \to -\infty$ the points (x_n, y_n) approach $(1,1)$. Now choose any continuous, strictly monotone function

$$y = h(x),$$

defined on

$$x_1 \leq x \leq x_0$$

with

$$h(x_1) = y_1, \quad h(x_0) = y_0.$$

Extend its definition to the interval $x_2 \leq x \leq x_1$ by setting

$$h(x) = g(h(f^{-1}(x))), \qquad x_2 \leq x \leq x_1.$$

Notice that at x_1 we have

$$g(h(f^{-1}(x_1))) = g(h(x_0)) = g(y_0) = y_1,$$

so the definitions of h at the point x_1 are consistent. Since f and g are monotone and continuous, and since h was chosen to be monotone on $x_1 \leq x \leq x_0$, we conclude that h is monotone on $x_2 \leq x \leq x_1$ and hence continuous and monotone on all of $x_2 \leq x \leq x_0$. Continuing in this way, we define h on the interval $x_{n+1} \leq x \leq x_n$, $n \geq 0$ by

$$h = g^n \circ h \circ f^{-n}.$$

Setting $h(0) = 0$, we get a continuous and monotone increasing function defined on $0 \leq x \leq x_0$. Similarly, we extend the definition of h to the right of x_0 up to $x = 1$. By its very construction, the map h conjugates f into g, proving the proposition.

Notice that as a corollary of the method of proof, we can conclude

Proposition 2.6.2 *Let f and g be two monotone increasing functions defined in some neighborhood of the origin and satisfying*

$$f(0) = g(0) = 0, \quad |f(x)| < |x|, \ |g(x)| < |x|, \ \forall x \neq 0.$$

Then there exists a homeomorphism, h defined in some neighborhood of the origin with $h(0) = 0$ and

$$h \circ f = g \circ h.$$

Indeed, just apply the method (for $n \geq 0$) to construct h to the right of the origin, and do an analogous procedure to construct h to the left of the origin. As a special case we obtain

Proposition 2.6.3 *Let f and g be differentiable functions with $f(0) = g(0) = 0$ and*

$$0 < f'(0) < 1, \quad 0 < g'(0) < 1. \tag{2.6}$$

Then there exists a homeomorphism h defined in some neighborhood of the origin with $h(0) = 0$ and which conjugates f into g.

The mean value theorem guarantees that the hypotheses of the preceding proposition are satisfied.

Also, it is clear that we can replace (2.6) by any of the conditions

$$\begin{array}{lr} 1 < f'(0), & 1 < g'(0) \\ 0 > f'(0) > -1, & 0 > g'(0) > -1 \\ -1 > f'(0), & -1 > g'(0), \end{array}$$

and the conclusion of the proposition still holds.

It is important to observe that if $f'(0) \neq g'(0)$, then the homeomorphism, h, can not be a diffeomorphism. That is, h can not be differentiable with a differentiable inverse. In fact, h can not have a non-zero derivative at the origin. Indeed, differentiating the equation $g \circ h = h \circ f$ at the origin gives

$$g'(0)h'(0) = h'(0)f'(0),$$

and if $h'(0) \neq 0$ we can cancel it form both sides of the equation so as to obtain

$$f'(0) = g'(0). \tag{2.7}$$

What is true is that if (2.7) holds, and if

$$|f'(0)| \neq 1, \tag{2.8}$$

then we can find a differentiable h with a differentiable inverse which conjugates f into g.

We postpone the proof of this result until we have developed enough machinery to deal with the n-dimensional result. These theorems are among my earliest mathematical theorems. A complete characterization of transformations of $\mathbf{R}$ near a fixed point together with the conjugacy by smooth maps if (2.7) and (2.8) hold, were obtained and submitted for publication in 1955 and published in the Duke Mathematical Journal. The discussion of equivalence under homeomorphism or diffeomorphism in n-dimensions was treated for the case of contractions in 1957 and in the general case in 1958, both papers appearing in the American Journal of Mathematics. We will return to these matters in Chapter ??.

2.7 Sequence space and symbolic dynamics.

In this section we will illustrate a powerful method for studying dynamical systems by examining the quadratic transformation

$$Q_c : x \mapsto x^2 + c$$

for values of $c < -2$.

For any value of c, the two possible fixed points of Q_c are

$$p_-(c) = \frac{1}{2}(1 - \sqrt{1-4c}), \quad p_+(c) = \frac{1}{2}(1 + \sqrt{1-4c})$$

by the quadratic formula. These roots are real with $p_-(c) < p_+(c)$ for $c < 1/4$. The graph of Q_c lies above the diagonal for $x > p_+(c)$, hence the iterates of any $x > p_+(c)$ tend to $+\infty$. If $x_0 < -p_+(c)$, then $x_1 = Q_c(x_0) > p_+(c)$, and so the further iterates also tend to $+\infty$. Hence all the interesting action takes place in the interval $[-p_+, p_+]$. The function Q_c takes its minimum value, c, at $x = 0$, and

$$c = -p_+(c) = -\frac{1}{2}(1 + \sqrt{1-4c})$$

when $c = -2$. For $-2 \leq c \leq 1/4$, the iterate of any point in $[-p_+, p_+]$ remains in the interval $[-p_+, p_+]$. But for $c < -2$ some points will escape, and it is this latter case that we want to study.

To visualize the what is going on, draw the square whose vertices are at $(\pm p_+, \pm p_+)$ and the graph of Q_c over the interval $[-p_+, p_+]$. The bottom of the graph will protrude below the bottom of the square. Let A_1 denote the open interval on the x-axis (centered about the origin) which corresponds to this protrusion. So

$$A_1 = \{x | Q_c(x) < -p_+(c)\}.$$

Every point of A_1 escapes from the interval $[-p_+, p_+]$ after one iteration.

Let

$$A_2 = Q_c^{-1}(A_1).$$

Since every point of $[-p_+, p_+]$ has exactly two pre-images under Q_c, we see that A_2 is the union of two open intervals. To fix notation, let

$$I = [-p_+, p_+]$$

and write

$$I \backslash A_1 = I_0 \cup I_1$$

where I_0 is the closed interval to the left of A_1 and I_1 is the closed interval to the right of A_1.Thus A_2 is the union of two open intervals, one contained in I_0 and the other contained in I_1. Notice that a point of A_2 escapes from $[-p_+, p_+]$ in exactly two iterations: one application of Q_c moves it into A_1 and another application moves it out of $[-p_+, p_+]$.

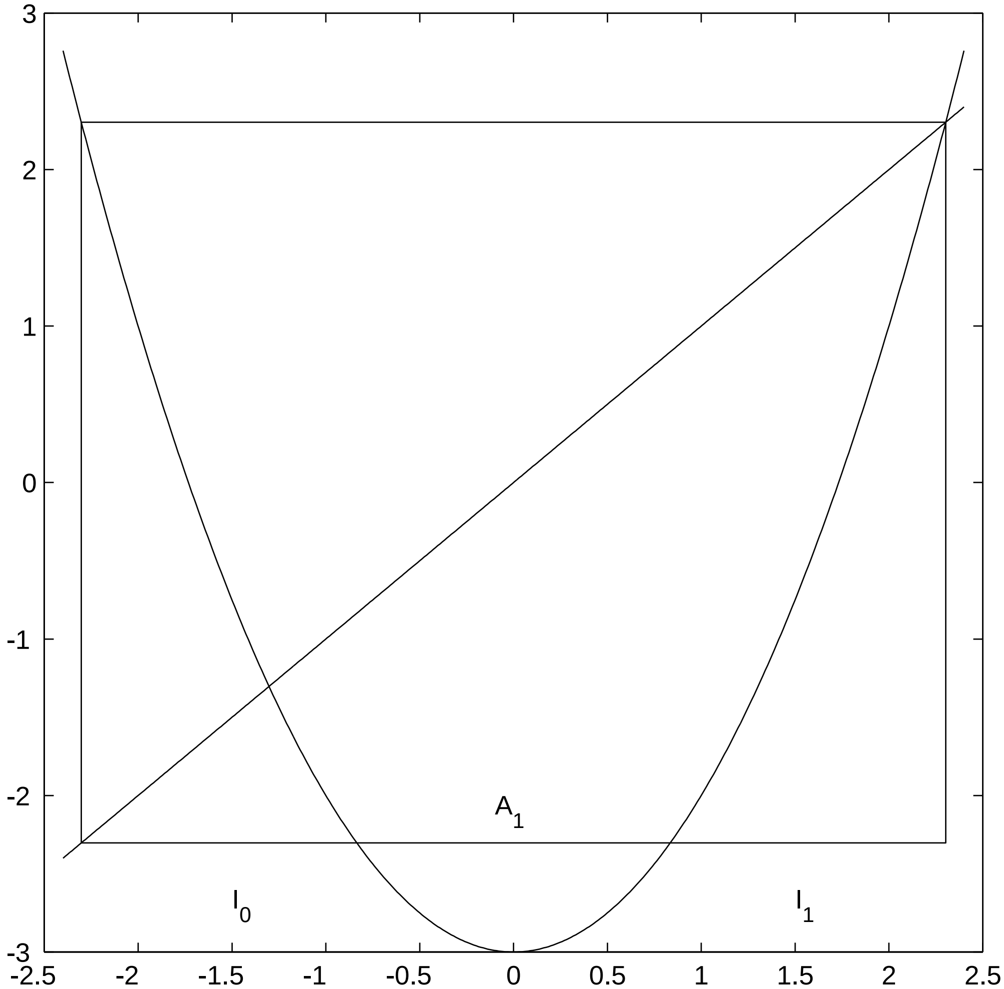

Figure 2.5: Q_3.

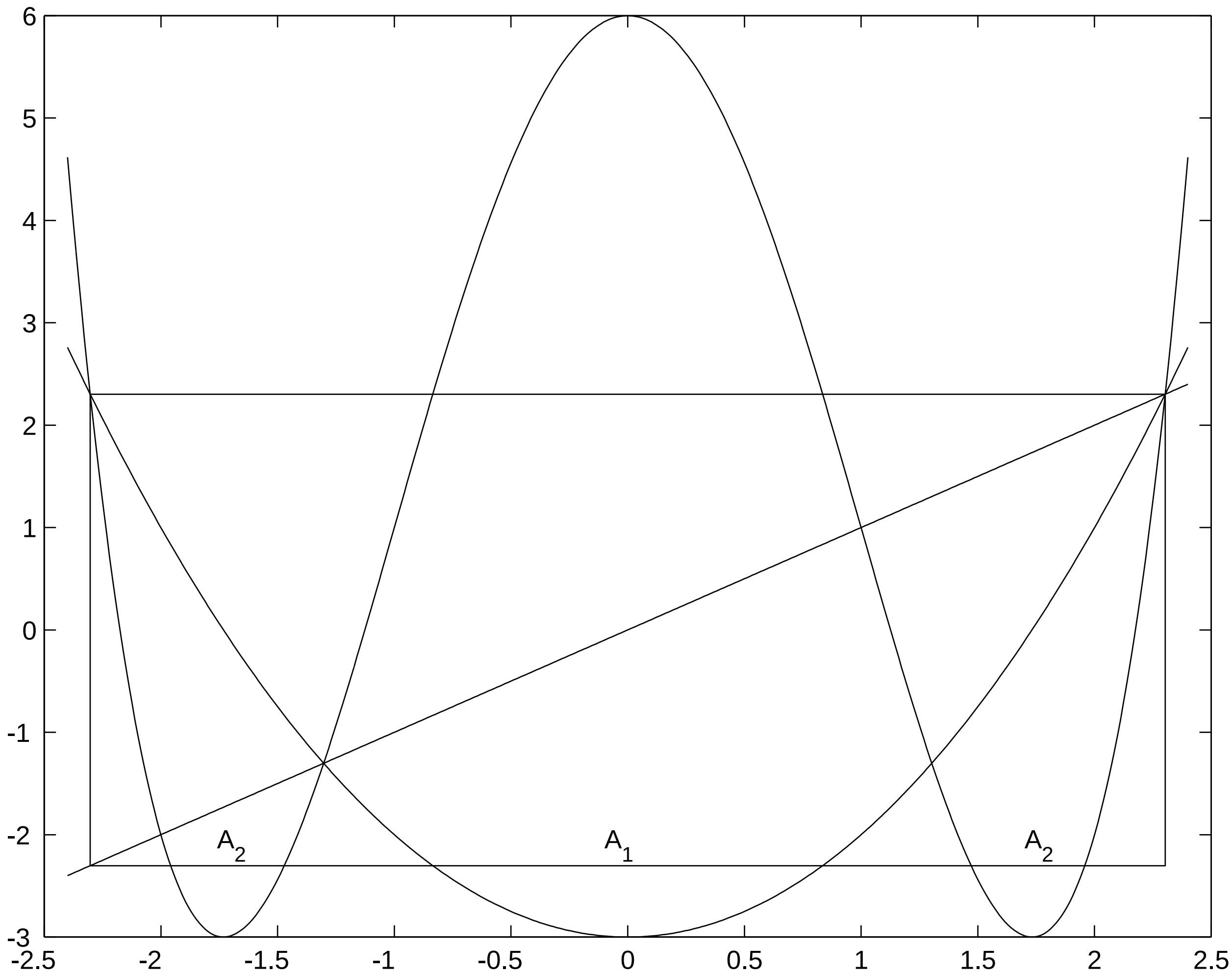

Figure 2.6: Q_3 and $Q_3^{\circ 2}$.

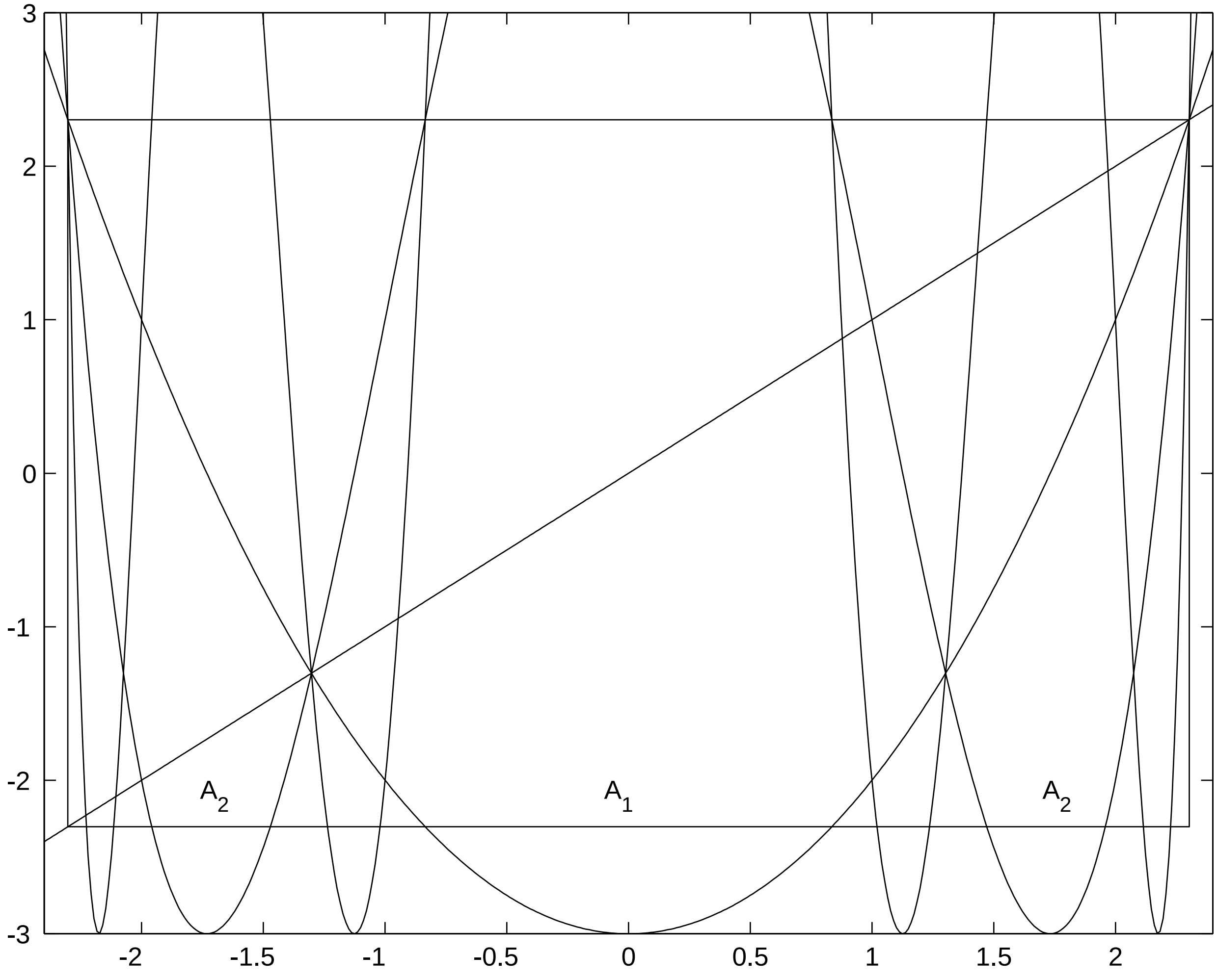

Figure 2.7: Q_3, $Q_3^{\circ 2}$, and $Q_3^{\circ 3}$.

Conversely, suppose that a point x escapes from $[-p_+, p_+]$ in exactly two iterations. After one iteration it must lie in A_1, since these are exactly the points that escape in one iteration. Hence it must lie in A_2.

In general, let
$$A_{n+1} = Q_c^{-\circ n}(A_1).$$
Then A_{n+1} is the union of 2^n open intervals and consists of those points which escape from $[-p_+, p_+]$ in exactly $n+1$ iterations. If the iterates of a point x eventually escape from $[-p_+, p_+]$, there must be some $n \geq 1$ so that $x \in A_n$. In other words,
$$\bigcup_{n\geq 1} A_n$$
is the set of points which eventually escape. The remaining points, lying in the set
$$\Lambda := I \backslash \bigcup_{n\geq 1} A_n m,$$
are the points whose iterates remain in $[-p_+, p_+]$ forever. The thrust of this section is to study Λ and the action of Q_c on it.

Since Λ is defined as the complement of an open set, we see that Λ *is closed.* Let us show that Λ *is not empty.* Indeed, the fixed points, $p_\pm$ certainly belong to Λ and hence so do all of their inverse images, $Q_c^{-n}(p_\pm)$. Next we will prove

Proposition 2.7.1 *If*
$$c < -\frac{5+2\sqrt{5}}{4} \doteq -2.368\ldots \tag{2.9}$$
then Λ is totally disconnected, that is, it contains no interval.

In fact, the proposition is true for all $c < -2$ but, following Devaney [?] we will only present the simpler proof when we assume (2.9). For this we use

Lemma 2.7.1 *If (2.9) holds then there is a constant $\lambda > 1$ such that*
$$|Q_c'(x)| > \lambda > 1, \quad \forall x \in I \backslash A_1. \tag{2.10}$$

Proof of Lemma. We have $|Q_c'(x)| = |2x| > \lambda > 1$ if $|x| > \frac{1}{2}\lambda$ for all $x \in I\backslash A_1$. So we need to arrange that A_1 contains the interval $[-\frac{1}{2}, \frac{1}{2}]$ in its interior. In other words, we need to be sure that
$$Q_c(\frac{1}{2}) < -p_+.$$
The equality
$$Q_c(\frac{1}{2}) = -p_+$$
translates to
$$\frac{1}{4} + c = -\frac{1+\sqrt{1-4c}}{2}.$$

Solving the quadratic equation gives

$$c = -\frac{5 + 2\sqrt{5}}{4}$$

as the lower root. Hence if (2.9) holds, $Q_c(\frac{1}{2}) < -p_+$.
Proof of Prop. 2.7.1. Suppose that there is an interval, J, contained in Λ. Then J is contained either in I_0 or I_1. In either event the map Q_c is one to one on J and maps it onto an interval. For any pair of points, x and y in J, the mean value theorem implies that

$$|Q_c(x) - Q_c(y)| > \lambda|x - y|.$$

Hence if d denotes the length of J, then $Q_c(J)$ is an interval of length at least λd contained in Λ. By induction we conclude that Λ contains an interval of length $\lambda^n d$ which is ridiculous, since eventually $\lambda^n d > 2p_+$ which is the length of I. QED.

Now consider a point $x \in \Lambda$. Either it lies in I_0 or it lies in I_1. Let us define

$$s_0(x) = 0 \quad \forall x \in I_0$$

and

$$s_0(x) = 1 \quad \forall x \in I_1.$$

Since all points $Q_c^{\circ n}(x)$ are in Λ, we can define $s_n(x)$ to be 0 or 1 according to whether $Q_c^{\circ n}(x)$ belongs to I_0 or I_1. In other words, we define

$$s_n(x) := \begin{cases} 0 & \text{if} \quad Q_c^{\circ n}(x) \in I_o \\ 1 & \text{if} \quad Q_c^{\circ n}(x) \in I_1 \end{cases}. \tag{2.11}$$

higher iterates of Q_c.

So let us introduce the **sequence space**, Σ, defined as

$$\Sigma = \{(s_0 s_1 s_2 \dots) \mid s_j = 0 \text{ or } 1\}.$$

Notice that in contrast to the space X we introduced in Section 3.4, we are not excluding any sequences. Define the notion of *distance* or *metric* on Σ by defining the distance between two points

$$\mathbf{s} = (s_0 s_1 s_2 \dots)$$

and

$$\mathbf{t} = (t_0 t_1 t_2 \dots)$$

to be

$$d(\mathbf{s}, \mathbf{t}) \stackrel{def}{=} \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i}.$$

It is immediate to check that d satisfies all the requirements for a metric: It is clear that $d(\mathbf{s},\mathbf{t}) \geq 0$ and $d(\mathbf{s},\mathbf{t}) = 0$ implies that $|s_i - t_i| = 0$ for all i, and hence that $\mathbf{s} = \mathbf{t}$. The definition is clearly symmetric in $\mathbf{s}$ and $\mathbf{t}$. And the usual triangle inequality

$$|s_i - u_i| \leq |s_i - t_i| + |t_i - u_i|$$

for each i implies the triangle inequality

$$d(\mathbf{s},\mathbf{u}) \leq d(\mathbf{s},\mathbf{t}) + d(\mathbf{t},\mathbf{u}).$$

Notice that if $s_i = t_i$ for $i = 0, 1, \ldots, n$ then

$$d(\mathbf{s},\mathbf{t}) = \sum_{j=n+1}^{\infty} \frac{|s_j - t_j|}{2^j} \leq \sum_{j=n+1}^{\infty} \frac{1}{2^j} = \frac{1}{2^n}.$$

Conversely, if $s_i \neq t_i$ for some $i \leq n$ then

$$d(\mathbf{s},\mathbf{t}) \geq \frac{1}{2^j} \geq \frac{1}{2^n}.$$

So if

$$d(\mathrm{s},\mathrm{t}) < \frac{1}{2^n}$$

then $s_i = t_i$ for all $i \leq n$.

Getting back to Λ, define the map

$$\iota : \Lambda \to \Sigma$$

by

$$\iota(x) = (s_0(x)s_1(x)s_2(x)s_3(x)\ldots) \tag{2.12}$$

where the $s_i(x)$ are defined by (2.11).

The point $\iota(x)$ is called the *itinerary* of the point x. For example, the fixed point, p_+ lies in I_1 and hence do all of its images under Q_c^n since they all coincide with p_+. Hence its itinerary is

$$\iota(p_+) = (111111\ldots).$$

The point $-p_+$ is carried into p_+ under one application of Q_c and then stays there forever. Hence its itinerary is

$$\iota(-p_+) = (01111111\ldots).$$

It follows from the very definition that

$$\iota(Q_c(x)) = S(\iota(x))$$

where S is our old friend, the shift map,

$$S : (s_0s_1s_2s_3\ldots) \mapsto (s_1s_2s_3s_4\ldots)$$

applied to the space Σ. In other words,

$$\iota \circ Q_c = S \circ \iota.$$

The map ι conjugates Q_c, acting on Λ into the shift map, acting on Σ. To show that this is a legitimate conjugacy, we must prove that ι is a homeomorphism. That is, we must show that ι is one-to one, that it is onto, that it is continuous, and that its inverse is continuous:

One-to one: Suppose that $\iota(x) = \iota(y)$ for $x, y \in \Lambda$. This means that $Q_c^n(x)$ and $Q^n(y)$ always lie in the same interval, I_0 or I_1. Thus the interval $[x, y]$ lies entirely in either I_0 or I_1 and hence Q_c maps it in one to one fashion onto an interval contained in either I_0 or I_1. Applying Q_c once more, we conclude that Q_c^2 is one-to-one on $[x, y]$. Continuing, we conclude that Q_c^n is one-to-one on the interval $[x, y]$, and we also know that (2.9) implies that the length of $[x, y]$ is increased by a factor of λ^n. This is impossible unless the length of $[x, y]$ is zero, i.e. $x = y$.

Onto. We start with a point $\mathbf{s} = (s_0 s_1 s_2 \dots) \in \Sigma$. We are looking for a point x with $\iota() = \mathbf{s}$. Consider the set of $y \in \Lambda$ such that

$$d(\mathrm{s}, \iota(y)) \leq \frac{1}{2^n}.$$

This is the same as requiring that y belong to

$$\Lambda \cap I_{s_0 s_1 \dots s_n}$$

where $I_{s_0 s_1 \dots s_n}$ is the interval

$$I_{s_0 s_1 \dots s_n} = \{y \in I |\ y \in I_{s_0}, Q_c(y) \in I_{s_1}, \dots Q_c^{\circ n}(y) \in I_{s_n}\}.$$

So

$$\begin{aligned} I_{s_0 s_1 \dots s_n} &= I_{s_0} \cap Q_c^{-1}(I_{s_1}) \cap \dots \cap Q_c^{-n}(I_{s_n}) \\ &= I_{s_0} \cap Q_c^{-1}(I_{s_1} \cap \dots \cap Q_c^{-(n-1)}(I_{s_n})) \\ &= I_{s_0} \cap Q_c^{-1}(I_{s_1 \dots s_n}) \qquad (2.13) \\ &= I_{s_0 s_1 \dots s_{n-1}} \cap Q_c^{-n}(I_{s_n}) \subset I_{s_0 \dots s_{n-1}}. \qquad (2.14) \end{aligned}$$

The inverse image of any interval, J under Q_c consists of two intervals, one lying in I_0 and the other lying in I_1. For $n = 0$, I_{s_0} is either I_0 or I_1 and hence is an interval. By induction, it follows from (2.13) that $I_{s_0 s_1 \dots s_n}$ is an interval. By (2.14), these intervals are nested. By construction these nested intervals are closed. Since every sequence of closed nested intervals on the real line has a non-empty intersection, there is a point x which belongs to all of these intervals. Hence all the iterates of x lie in I, so $x \in \Lambda$ and $\iota(x) = \mathbf{s}$.

Continuity. The above argument shows that the interiors of the intervals $I_{s_0 s_1 \dots s_n}$ (intersected with Λ) form neighborhoods of x that map into small neighborhoods of $\iota(x)$.

Continuity of ι^{-1}. Conversely, any small neighborhood of x in Λ will contain one of the intervals $I_{s_0 \dots s_n}$ and hence all of the points t whose first n coordinates agree with s $= \iota(x)$ will be mapped by ι^{-1} into the given neighborhood of x.

To summarize: we have proved

Theorem 2.7.1 *Suppose that c satisfies (2.9). Let $\Lambda \subset [-p_+, p_+]$ consist of those points whose images under Q_c^n lie in $[-p, p_+]$ for all $n \geq 0$. Then Λ is a closed, non-empty, disconnected set. The itinerary map ι is a homeomorphism of Λ onto the sequence space, Σ, and conjugates Q_c to the shift map, S.*

Just as in the case of the space X in section 3.4, the periodic points for S are precisely the periodic or "repeating" sequences. Thus we can conclude from the theorem that there are exactly 2^n points of period (at most) n for Q_c. Also, the same argument as in section 3.4 shows that the periodic points for S are dense in Σ, and hence the periodic points for Q_c are dense in Λ. Finally, the same argument as in section 3.4 shows that S is transitive on Σ. Hence, the restriction of Q_c to Λ is chaotic.