CHAOTIC UNIMODAL AND BIMODAL MAPS

FRED SHULTZ

ABSTRACT. We describe up to conjugacy all unimodal and bimodal maps that are chaotic, by giving necessary and sufficient conditions for unimodal and bimodal maps of slopes $\pm s$ to be transitive.

1. INTRODUCTION

In discussing chaotic dynamical systems, interval maps provide examples that can be easily visualized. Two standard examples are the full tent map $x \mapsto 1 - |1 - 2x|$, and the conjugate map $x \mapsto 4x(1-x)$. A logistic map $f_k(x) = kx(1-x)$ for $k > 4$ is also chaotic, if restricted to the set X of points in $[0, 1]$ whose orbits stay in $[0, 1]$, cf. [6]. However, here X is a Cantor set. Are there other simple examples of functions chaotic on all of $[0, 1]$? Our purpose here is to describe (up to conjugacy) all chaotic interval maps that are unimodal or bimodal, i.e., have two or three intervals of monotonicity, where by an interval map we mean any continuous map of $[0, 1]$ into $[0, 1]$.

In [3], Devaney defines a continuous map to be chaotic if it is transitive, has a dense set of periodic points, and is sensitive to initial conditions. If $\tau : [0, 1] \to [0, 1]$ is continuous, then the results of Banks [1] and of Vellekoop and Berglund [13] show that transitivity implies the other two conditions. (See [4, Thm. 3.8] for an exposition.) Thus an interval map is chaotic iff it is transitive.

If an interval map τ is piecewise monotonic and transitive, by a result of Parry [8], τ is conjugate to a *uniformly piecewise linear map*, i.e., an interval map which is piecewise linear with slopes $\pm s$ for some constant s. Thus to describe all chaotic piecewise monotonic interval maps up to conjugacy, it suffices to determine which uniformly piecewise linear maps are transitive.

If a uniformly piecewise linear map with slopes $\pm s$ with $s > 1$ has a fixed point p such that each point in a neighborhood of p has a unique preimage, then the map is not transitive, cf. Lemma 5. Our main result (Theorem 21) says roughly that the converse of this is true for maps that are unimodal or bimodal. From this result, it is easy to tell if a particular uniformly piecewise linear unimodal or bimodal map is transitive or not.

We prove this first for unimodal maps, then for bimodal maps with slope $\pm s$ with $s > 2$, then for general "down-up-down" maps, and finally for "up-down-up" maps. The arguments given here characterizing transitivity of uniformly piecewise linear unimodal and bimodal maps are easily accessible to any undergraduate.

Date: April 12, 2007.

2000 *Mathematics Subject Classification.* Primary 37E05.

Key words and phrases. interval map, piecewise monotonic, uniformly piecewise linear, conjugacy, transitive, exact, unimodal, bimodal.

Except for the "down-up-down" case, these individual results on transitivity of uniformly piecewise linear maps are already either implicitly or explicitly in the literature. See Jonker and Rand [5] or Bassein [2, Prop. 4] for the unimodal case, and Veitch and Glendinning [12, Lemmas 3.3, 3.4] for the up-down-up case. (Warning: Veitch and Glendinning call "transitive" what we call "topologically exact", cf. Definition 3). Our goal here is to pull together the necessary results and techniques and fill in the missing pieces to give a simple answer to the problem of describing transitive unimodal and bimodal maps up to conjugacy. We would like to thank Robert Devaney for a helpful reference.

2. Preliminaries

Definition 1. An interval map τ is *piecewise monotonic* if there is a partition of $[0,1]$ into finitely many subintervals on each of which τ is (strictly) monotonic. If two (respectively, three) is the minimal number of such intervals, we say τ is *unimodal* (respectively, *bimodal*).

Definition 2. A map $\tau : [0,1] \to [0,1]$ is *transitive* if for every pair U, V of nonempty open sets, there exists $n \geq 0$ such that $\tau^n(U) \cap V \neq \emptyset$.

For interval maps, transitivity implies the existence of a dense orbit, cf. [14, Thm. 5.9], and the converse holds as well. One way to show that a map τ is not transitive is to find an invariant subset J (i.e., $\tau(J) \subset J$) such that both J and its complement have non-empty interiors U, V. (Then no iterate of U meets V).

To establish transitivity, it is often easier to prove the following stronger property, which is discussed in more detail in [9].

Definition 3. A map $\tau : [0,1] \to [0,1]$ is *(topologically) exact* if for every nonempty open set V, there exists $n \geq 0$ such that $T^n(V) = [0,1]$.

We will make frequent use of the following lemmas.

Lemma 4. *Let $T : [0,1] \to [0,1]$ be uniformly piecewise linear with slopes $\pm s$ with $s > 1$. Let $[c,d]$ be a subinterval of $[0,1]$ on which T is linear and which contains a fixed point p. If J is any subinterval of $[0,1]$ containing p in its interior, then $T^k(J) = [0,1]$ for some $k \geq 0$.*

Proof. Let $e = p - c$ and $f = d - p$. Then for all $n \geq 0$, $T^{2n}([p - \frac{e}{s^{2n}}, p + \frac{f}{s^{2n}}]) = [p - e, p + f] = [c, d]$. If we choose n large enough such that $[p - \frac{e}{s^{2n}}, p + \frac{f}{s^{2n}}] \subset J$, then $T^{2n}(J) \supset [c,d]$, so $T^{2n+1}(J) = [0,1]$. □

Lemma 5. *Let $T : [0,1] \to [0,1]$ be uniformly piecewise linear with slopes $\pm s$ with $s > 1$. If for some $n \geq 1$, there is a fixed point p of T^n such that each point in a neighborhood V of p has a unique preimage under T^n, then T is not transitive.*

Proof. Assume p, n and V are as in the statement of the lemma. We first establish that for e sufficiently small, $J_e = (p - e, p + e) \cap [0,1]$ satisfies $T^{-n}(J_e) \subset J_e$.

If $p = 0$, choose $e > 0$ such that each point in $J_e = [0, e]$ has a unique preimage under T^n, and such that T^n is linear on J_e with slope $s^n > 1$. Then $T^n(J_e) = [0, s^n e] \supset J_e$, and so by uniqueness of preimages, $T^{-n}(J_e) \subset J_e$.

The case $p = 1$ is similar, so assume $0 < p < 1$. Let $V \subset (0,1)$ be an open neighborhood of p such that each point in V has a unique preimage under T^n. Choose $e > 0$ so that $J_e = (p - e, p + e) \subset V \cap T^{-n}(V)$. Since each point in V has a unique preimage, and $T^n(J_e) \subset V$, then T^n has no turning point in J_e. Thus T^n is linear on J_e, with slope $\pm s^n$. Then

$$T^n(J_e) = (p - s^n e, p + s^n e) \supset J_e.$$

Since $J_e \subset V$, each point in J_e has a unique preimage under T^n, so $T^{-n}(J_e) \subset J_e$.

Now define

$$K_e = J_e \cup T^{-1}(J_e) \cup \cdots \cup T^{-(n-1)}(J_e).$$

Since $T^{-n}(J_e) \subset J_e$, we have

$$T^{-1}(K_e) \subset T^{-1}(J_e) \cup T^{-2}(J_e) \cup \cdots \cup T^{-(n-1)}(J_e) \cup J_e = K_e.$$

Since T is bimodal and has slopes $\pm s$, for any interval J we have

$$|T^{-1}(J)| \leq 3s^{-1}|J| \leq 3|J|.$$

It follows that we can choose e so that K_e has total length strictly less than 1. Then the complement of K_e has nonempty interior W. Since $T^{-1}(K_e) \subset K_e$, forward images of W never meet K_e, and so T is not transitive. □

The following result, due to Parry [8], allows us to reduce the analysis of transitive piecewise monotonic maps to the case of uniformly piecewise linear maps.

Theorem 6. *If a piecewise monotonic map $\tau : [0,1] \to [0,1]$ is transitive, then τ is conjugate to a uniformly piecewise linear map with slopes $\pm s$ with $s > 1$.*

Proof. The map τ is said to be *strongly transitive* if for every open subinterval J a finite number of iterates of J cover $[0,1]$. Every transitive interval map is strongly transitive, cf. [9, Theorem 2.5]. Parry showed that every strongly transitive piecewise monotonic interval map is conjugate to a uniformly piecewise linear map T. Since T must be surjective, then $s \geq 1$, and $s = 1$ would force $T(x) = \pm x$, which is not transitive.

Alternatively, since transitive interval maps have positive entropy, we could apply Milnor and Thurston's result in [7] that piecewise monotonic maps of positive entropy are semiconjugate to uniformly piecewise linear maps, and then observe that transitivity forces the semiconjugacy to be a conjugacy.

We sketch a third proof. By a fixed point argument, one can show there is a non-atomic probability measure μ with support all of $[0,1]$, and a constant $s \geq 1$, such that $\mu(\tau(E)) = s\mu(E)$ for each Borel set E on which τ is injective. Define $h : [0,1] \to [0,1]$ by $h(x) = \mu([0,x])$. Then h is a homeomorphism of $[0,1]$, and is a conjugacy from τ onto a uniformly piecewise linear map with slopes $\pm s$. For details see [10, Cor. 4.4].

□

Observation. If $T : [0,1] \to [0,1]$ is uniformly piecewise linear, with slopes $\pm s$, and J is an interval on which T has k intervals of monotonicity, then $|T(J)| \geq \frac{s}{k}|J|$, where $|J|$ denotes the length of J. We will make use of this observation throughout this paper.

3. Unimodal maps

Definition 7. If $1 < s \leq 2$, the *tent map* $f_s : [0,1] \to [0,1]$ is defined by

$$f_s(x) = \begin{cases} sx & \text{if } x < 1/2 \\ -sx + s & \text{if } x \geq 1/2 \end{cases}$$

It is easy to check that f_2 is topologically exact, and thus is transitive. However, for $1 < s < 2$, the map f_s is not surjective. Even if we restrict f_s to $f_s([0,1]) = [0, s/2]$, the subinterval $[f_s(1/2), f_s^2(1/2)]$ will be invariant, so f_s will not be transitive. It is more interesting to restrict f_s to the interval $[f_s(1/2), f_s^2(1/2)]$, and rescale so that the domain is $[0,1]$.

Definition 8. If $1 < s \leq 2$, the *restricted tent map* $T_s : [0,1] \to [0,1]$ is defined by

$$T_s(x) = \begin{cases} sx - s + 2 & \text{if } x < 1 - \frac{1}{s} \\ s - sx & \text{if } x \geq 1 - \frac{1}{s} \end{cases}$$

(See Figure 1.)

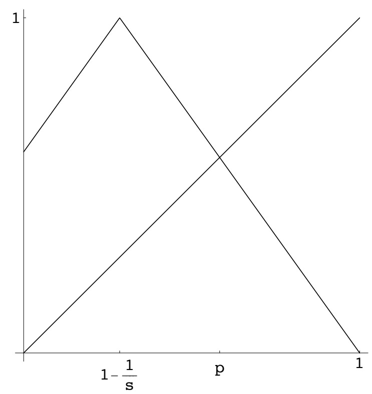

FIGURE 1

We say $f : [0,1] \to [0,1]$ *exchanges* intervals J_1 and J_2 if $f(J_1) = J_2$ and $f(J_2) = J_1$. (By "interval" we will always mean a non-degenerate interval, i.e., one consisting of more than a single point.) The following theorem is implicit in [5] and can also be found in [10, Lemma 8.1].

Theorem 9. *Let $1 < s \leq 2$, and let p be the unique fixed point of T_s.*

(i) *The restricted tent map T_s is topologically exact iff $s > \sqrt{2}$, which occurs iff $T_s(0) < p$.*
(ii) *If $s = \sqrt{2}$, then T_s is transitive but not exact. It exchanges the subintervals $[0,p]$ and $[p,1]$, and T_s^2 is exact on each of these subintervals. The condition $s = \sqrt{2}$ is equivalent to $T_s(0) = p$.*
(iii) *If $s < \sqrt{2}$, then T_s is not transitive; this occurs iff $T_s(0) > p$.*

Proof. It is straightforward to check that $T_s(0) = p$ is equivalent to $s = \sqrt{2}$, and similarly $T_s(0) > p$ iff $s < \sqrt{2}$, and $T_s(0) < p$ iff $s > \sqrt{2}$.

(i) If $s > \sqrt{2}$, let J be any subinterval of $[0,1]$. Suppose $T_s^2(J) \neq [0,1]$. Let c be the critical point of T_s. If c were in both J and $T_s(J)$, then c and $1 = T_s(c)$ both would be in $T_s(J)$, so $T_s^2(J) \supset T_s([c,1]) = [0,1]$, contrary to our assumption. Thus on at least one of the intervals J and $T_s(J)$, T_s is linear, and on the other interval T_s is either linear or unimodal. Therefore, $|T_s^2(J)| \geq \frac{s^2}{2}|J|$ and $\frac{s^2}{2} > 1$. We can repeat the argument starting with $T_s^2(J)$ in place of J. Thus $(T_s^2)^n(J) = [0,1]$ for some n.

(ii) If $T_s(0) = p$, then T_s maps $[0, p]$ onto $[p, 1]$ and maps $[p, 1]$ onto $[0, p]$, so T^2 is not exact on $[0, 1]$. However, T_s^2 maps $[0, p]$ onto itself, and maps $[p, 1]$ onto itself, and restricted to each of these intervals is conjugate to the full tent map, so is topologically exact on each of these intervals. Since T_s exchanges $[0, p]$ and $[p, 1]$, it follows that T_s is transitive on $[0, 1]$. Indeed, suppose V and W are open subsets of $[0, 1]$. If V and W both meet the interior of $[0, p]$, then exactness of T_s^2 on $[0, p]$ implies that $T_s^{2n}(V) \cap W$ is nonempty for some n. If V meets the interior of $[0, p]$ and W meets the interior of $[p, 1]$, then $T_s(W)$ meets the interior of $[0, p]$, so the previous argument applies. The remaining cases can be treated similarly.

(iii) If $T_s(0) > p$, then each point in a neighborhood of p has a unique preimage, so T_s is not transitive, cf. Lemma 5. □

Definition 10. Two interval maps τ_1 and τ_2 are *reflections* of each other if $x \mapsto 1 - x$ is a conjugacy from τ_1 to τ_2.

This terminology is motivated by the observation that if τ_2 is the reflection of τ_1, then the graph of τ_2 is the graph of τ_1 reflected in the point $(1/2, 1/2)$.

Theorem 11. *A unimodal map $\tau : [0, 1] \to [0, 1]$ is transitive iff τ is conjugate to a restricted tent map T_s with $s \geq \sqrt{2}$.*

Proof. If τ is transitive, then τ is conjugate to a uniformly piecewise linear map $T : [0, 1] \to [0, 1]$, cf. Theorem 6. We may assume that T increases and then decreases. (If not, replace T by its reflection.) By surjectivity, T takes on the value 0 at either 0 or 1. If $T(1) > 0$, then $T(0) = 0$, so 0 is a fixed point and there is a neighborhood of 0 such that each point in the neighborhood has just one preimage. By Lemma 5, this contradicts transitivity of T. Thus $T(1) = 0$, so $T = T_s$ for some $s > 1$. Now the theorem follows from Theorem 9.

□

4. Bimodal maps

If a bimodal map τ increases, then decreases, then increases we say τ is *up-down-up*, and otherwise that τ is *down-up-down*.

In this section $T : [0, 1] \to [0, 1]$ will denote a bimodal piecewise linear map with slopes $\pm s$, with $s > 1$. Recall (Theorem 6) that any transitive bimodal map $\tau : [0, 1] \to [0, 1]$ is conjugate to such a map T.

Lemma 12. *If T is transitive, then there exists a subinterval $[a, b]$ mapped linearly onto $[0, 1]$, with neither a nor b being fixed points.*

Proof. If $T([c_1, c_2]) = [0, 1]$, we're done, so suppose this does not occur. Since T is transitive, then T is surjective. We consider the possibilities for the points where T takes on the values 0 and 1.

Suppose that T takes on the values 0 and 1 only at the endpoints of $[0, 1]$. Then either $T(0) = 0$ and $T(1) = 1$, or else $T(0) = 1$ and $T(1) = 0$. In either case, we can apply Lemma 5 with $p = 0$ and $n = 1$ or $n = 2$ to conclude that T is not transitive, contrary to our hypothesis.

The remaining possibility is that T takes on the value 0 at an endpoint, and 1 at a critical point, or vice versa. For example, suppose $T(0) = 0$ and $T(c_1) = 1$, with $T(c_2) > 0$. Then there is an interval containing 0 in which each point has a unique preimage, so by Lemma 5, T is not transitive. The case $T(1) = 1$ and

$T(c_2) = 0$, with $T(c_1) < 1$ is a reflection of the previous case, so again T is not transitive. Thus we are left with the cases $T(0) = 1$ with $T(c_1) = 0$, or $T(1) = 0$ with $T(c_2) = 1$. In the former case we can take $[a, b] = [0, c_1]$, and in the latter case $[a, b] = [c_2, 1]$.

□

4.1. **The case $s > 2$.** We will consider up-down-up and down-up-down bimodal maps separately. However, we first treat these cases together for the case where $s > 2$.

We start with a lemma concerning maps with $T(0) = 1$, as illustrated in Figure 2.

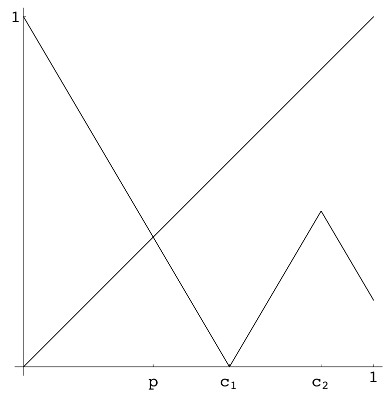

FIGURE 2

Lemma 13. *Let $c_1 < c_2$ be the critical points of T. Assume $T(0) = 1$ and $T(c_1) = 0$, and let p be the fixed point of T in $[0, c_1]$.*

(i) *If $s \leq \sqrt{2}$, then $T(c_2) < p$.*
(ii) *If $s > \sqrt{3}$, then $T(c_2) > p$.*

Proof. (i) We have

$$T(c_2) = s(c_2 - c_1) < s(1 - c_1) = s(1 - 1/s) = s - 1.$$

If $s \leq \sqrt{2}$, then $s - 1 \leq \frac{1}{s+1} = p$, so $T(c_2) < p$.

(ii) In order for T to be bimodal, $c_2 - c_1 \geq \frac{1}{2}(1 - c_1)$, so

$$T(c_2) = s(c_2 - c_1) \geq \frac{s}{2}(1 - 1/s) = \frac{1}{2}(s - 1).$$

If $s > \sqrt{3}$, then $\frac{1}{2}(s - 1) > \frac{1}{s+1} = p$, so $T(c_2) > p$.

□

The following shows that the converse to Lemma 12 holds if $s > 2$, as illustrated in Figure 3.

Theorem 14. *Assume $s > 2$. Then T is topologically exact iff there is a proper subinterval $[a, b]$ mapped linearly onto $[0, 1]$, with neither a nor b being fixed points. Furthermore, T is transitive iff it is exact.*

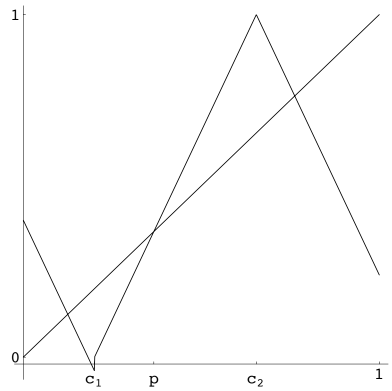

FIGURE 3. down-up-down, $s > 2$

Proof. Let $c_1 < c_2$ be the critical points of T. Suppose first that such an interval $[a, b]$ exists. Consider the case where $[a, b] = [c_1, c_2]$, so that $T([c_1, c_2]) = [0, 1]$. (See Figure 3). If J is any interval that does not contain $[c_1, c_2]$, then T is either linear or unimodal on J, so $|T(J)| \geq \frac{s}{2}|J|$. If no iterate of J contains $[c_1, c_2]$, then $|T^n(J)| \geq \left(\frac{s}{2}\right)^n |J|$ for all n. Since $\frac{s}{2} > 1$, this is impossible. Thus for some n we have $T^n(J) \supset [c_1, c_2]$, so $T^{n+1}(J) = [0, 1]$. Thus T is exact.

Consider next the case where $[a, b] = [0, c_1]$, so that $T(0) = 1$ and $T(c_1) = 0$. See Figure 2. By Lemma 13, since $s > 2 > \sqrt{3}$, then $T(c_2) > p$, where p is the fixed point of T in $[0, c_1]$. Now if J is any interval, since $s > 2$, iterates of J will expand in length until the iterate contains $[c_1, c_2]$. Since $T(c_2) > p$ and $T(c_1) = 0$, the next iterate will contain p in its interior. By Lemma 4, iterates of J will eventually equal $[0, 1]$, so T is exact.

The case where $T(1) = 0$ and $T(c_2) = 1$ is a reflection of the case just considered, and so T is again exact. Since T is bimodal, this exhausts the possibilities for an interval $[a, b]$ with the properties in the statement of the theorem, and thus the existence of such an interval implies that T is exact.

If T is transitive, the existence of the desired interval $[a, b]$ follows from Lemma 12, and exactness then follows from the first part of this proof. □

4.2. **Down-up-down maps.** For surjective down-up-down maps, with the exception of maps with $T(0) = 1$ or $T(1) = 0$, we will now show that the condition $s > 2$ is both necessary and sufficient for the map to be topologically exact or transitive.

Theorem 15. *Assume T is down-up-down, with critical points $c_1 < c_2$, and that $T(c_1) = 0$ and $T(c_2) = 1$. Let p be the fixed point of T in $[c_1, c_2]$. Then the following are equivalent.*

- (i) $s > 2$.
- (ii) $T(0) > p$.
- (iii) $T(1) < p$.
- (iv) *T is topologically exact.*
- (v) *T is transitive.*

Proof. See Figure 4.

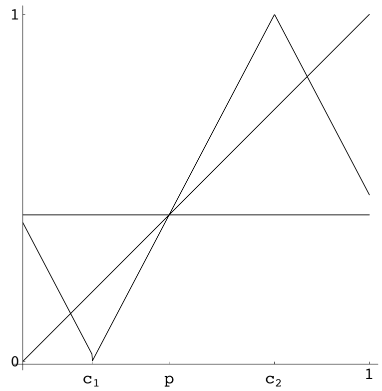

FIGURE 4. down-up-down, $s \leq 2$

(i) $\iff$ (ii) follows from $T(0) = c_1 s$ and $p = c_1 s/(s-1)$.
(i) $\iff$ (iii) follows from $T(1) = 1 - (1-c_2)s$ and $p = (c_2 s - 1)/(s-1)$.
(i) $\implies$ (iv) If $s > 2$, then T is topologically exact by Theorem 14.
(iv) $\implies$ (v) Trivial.
(v) $\implies$ (ii) If (ii) fails, then $T(0) \leq p$, so

$$T([0,p]) = T([0,c_1]) \cup T([c_1,p]) = [0,T(0)] \cup [0,p] = [0,p].$$

Thus $[0,p]$ is invariant under T, so T is not transitive.

□

By Lemma 12, there is one remaining kind of up down-up-down map that can be transitive: maps such that $T(0) = 1$ or $T(1) = 0$, cf. Figure 2. We begin with a lemma on critical points, relevant to showing that T stretches intervals until they include both critical points.

Lemma 16. *Assume T is down-up-down and $s \leq 2$. Let $c_1 < c_2$ be the critical points of T, and assume that $T(0) = 1$ and $T(c_1) = 0$. Suppose K is an interval such that no iterate of K contains both critical points.*

(i) *At least one of K, $T(K)$, $T^2(K)$ contains no critical point.*
(ii) *If $T(c_2) < c_1$, at least one of K or $T(K)$ contains no critical point.*

Proof. We first make no assumption about $T(c_2)$, and prove

$$\text{if } c_1 \in K, \text{ then } T(K) \text{ contains no critical point.} \tag{1}$$

Indeed, if $c_1 \in K$, then $0 = T(c_1) \in T(K)$. If c_1 or c_2 were in $T(K)$, then $[0, c_1] \subset K$, so $T^2(K) = [0,1]$, contrary to our assumption that no iterate of K contains both critical points. Thus we've proven (1).

We now prove (i) and (ii). By (1), we may assume $c_2 \in K$. Then $K \subset (c_1, 1]$, so

$$T(K) \subset [0, T(c_2)]. \tag{2}$$

Furthermore,

$$T(c_2) < c_2 \tag{3}$$

follows from

$$T(c_2) = s(c_2 - c_1) = sc_2 - 1 \leq 2c_2 - 1 = c_2 + (c_2 - 1) < c_2.$$

By (2) and (3), $T(K) \subset [0, c_2)$. If $c_1 \in T(K)$, by (1), $T^2(K)$ contains no critical points. If $c_1 \notin T(K)$, then $T(K) \subset [0, c_2)$ implies that $T(K)$ contains neither c_1 nor c_2, which proves (i).

If $T(c_2) < c_1$, suppose both K and $T(K)$ contain critical points. If $c_1 \in K$, we apply (1). Otherwise, as above we conclude $T(K) \subset [0, T(c_2))$. Since $T(c_2) < c_1$, then $T(K)$ contains no critical point, a contradiction. Hence (ii) follows.

□

Theorem 17. *Let $c_1 < c_2$ be the critical points of T, and assume that $T(0) = 1$ and $T(c_1) = 0$. Let p be the fixed point of T in $[0, c_1]$.*

(i) *If $T(c_2) > p$, then T is topologically exact.*
(ii) *If $T(c_2) = p$, then T is transitive, T exchanges $[0, p]$ and $[p, 1]$, and T^2 is topologically exact on each of these intervals.*
(iii) *If $T(c_2) < p$, then T is not transitive.*

Proof. See Figure 2.

(i) Assume $T(c_2) > p$. By Lemma 13, $s > \sqrt{2}$. If $s > 2$, then as shown in Theorem 14, T is exact. Thus we may assume $\sqrt{2} < s \leq 2$.

Let K be a subinterval of $[0, 1]$; we will show some iterate of K equals $[0, 1]$. If K contains both critical points, since $T(c_2) > p$ and $T(c_1) = 0$, then $T(K)$ contains the fixed point p in its interior, so by Lemma 4 some iterate of K equals $[0, 1]$. Thus it suffices to show that some iterate of K contains both critical points.

Suppose (to reach a contradiction) that K is an interval such that no iterate contains both critical points. If $T(c_2) < c_1$, by Lemma 16(ii) we have $|T^2(K)| \geq \frac{s^2}{2}|K|$. Since the same argument can be applied to $T^2(K)$, we conclude that

$$|T^{2n}(K)| \geq \left(\frac{s^2}{2}\right)^n |K| \quad \text{for all } n \geq 0.$$

Since $s > \sqrt{2}$, then $s^2/2 > 1$, so this is impossible. Thus some iterate of K contains both critical points, so is eventually mapped onto $[0, 1]$.

If $T(c_2) \geq c_1$, then a similar argument based on Lemma 16(i) shows

$$|T^{3n}(K)| \geq \left(\frac{s^3}{4}\right)^n |K| \quad \text{for all } n \geq 0. \tag{4}$$

Since $c_1 \leq T(c_2) = s(c_2 - c_1)$, then $c_1 = 1/s$ gives

$$\frac{1}{s} \leq sc_2 - 1,$$

so

$$\frac{1}{s^2} + \frac{1}{s} \leq c_2 < 1.$$

Hence $s^2 - s - 1 > 0$. Since $s > 1$, this implies $s > \frac{1+\sqrt{5}}{2} \approx 1.618$, so $s^3 > 4$. If this is combined with (4), we again reach a contradiction, and this completes the proof of (i).

(ii) If $T(c_2) = p$, then T exchanges $[0, p]$ and $[p, 1]$, and T^2 is a bimodal map from $[0, p]$ onto $[0, p]$ with slopes $\pm s^2$. By Lemma 13, $s > \sqrt{2}$, so $s^2 > 2$. If d_1, d_2 are the T-preimages in $[0, p]$ of c_1, c_2, then T^2 maps $[d_2, d_1]$ linearly onto $[0, p]$. By Theorem 14, T^2 restricted to $[0, p]$ is topologically exact. Since T exchanges $[0, p]$ and $[p, 1]$, then T is transitive on $[0, 1]$.

(iii) Assume $T(c_2) < p$. Then there is an open interval around p in which each point has a unique preimage. By Lemma 5, T is not transitive.

□

Remark. If $T(1) = 0$ and $T(c_2) = 1$, then transitivity and exactness of T can be determined by applying Theorem 17 to the reflection of T.

4.3. **Up-down-up maps.** Recall by Lemma 12 that for T to be transitive, there must be a subinterval $J = [a, b]$ mapped linearly onto $[0, 1]$, with neither a nor b being fixed points. Thus if T is up-down-up, with critical points $c_1 < c_2$, for T to have a chance of being transitive we must have $T(c_1) = 1$ and $T(c_2) = 0$. See Figure 5.

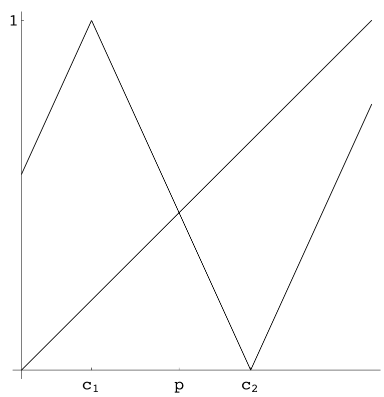

FIGURE 5. up-down-up

Lemma 18. *Let T be up-down-up, with critical points $c_1 < c_2$, with $T(c_1) = 1$ and $T(c_2) = 0$, and let p be the fixed point of T in $[c_1, c_2]$. Assume $T(1) > p$, and let K be any interval none of whose iterates contains the fixed point p in its interior.*

(i) *If K and $T(K)$ both contain critical points, then $c_2 \in K$ and $c_1 \in T(K)$*
(ii) *At most two of K, $T(K)$, $T^2(K)$ contain critical points.*
(iii) *If also $T(1) < c_2$, and both K and $T(K)$ contain critical points, then neither $T^2(K)$ nor $T^3(K)$ contains a critical point.*

Proof. Denote the interior of any interval J by J°. Hereafter K denotes an interval such that no iterate of K contains p in its interior. Since $T(c_1) = 1$ and $T(c_2) = 0$, then K can't contain two critical points (or else $T(K) = [0, 1]$ would contain p).

(i) Suppose both K and $T(K)$ contain critical points. Assume (to reach a contradiction) that $c_1 \in K$. Then $1 = T(c_1) \in T(K)$, so $T(K)$ is an interval with 1 as one endpoint. Since $T(K)$ contains just one critical point, then $c_2 \in T(K)$. Then $T^2(K)$ contains $T(c_2) = 0$ and contains $T^2(c_1) = T(1) > p$. Hence $p \in T^2(K)^\circ$, a contradiction.

Thus we've shown $c_1 \notin K$, so we must have $c_2 \in K$. Then $0 = T(c_2) \in T(K)$. Since $T(K)$ contains exactly one critical point, and contains 0, it can't contain c_2, so $c_1 \in T(K)$.

(ii) Suppose K, $T(K)$, and $T^2(K)$ all contain critical points. By (i), $c_2 \in K$ and $c_1 \in T(K)$. But by (i) applied to $T(K)$, $c_2 \in T(K)$, so $T(K)$ contains two critical points, which forces $p \in T(K)^\circ$, contrary to assumption.

(iii) Suppose $T(1) < c_2$, so that $p < T(1) < c_2$, and suppose that K and $T(K)$ contain critical points. By (ii), $T^2(K)$ contains no critical point, and by (i), $1 = T(c_1) \in T^2(K)$. Thus we can write $T^2(K) = [e, 1]$ with $e > c_2$. Then $T^3(K) = [T(e), T(1)]$. Since $p < T(1) < c_2$, and $p \notin T^3(K)^\circ$, we must have $p \le T(e)$, so $T^3(K) \subset [p, c_2)$; in particular $T^3(K)$ contains no critical point.

□

Lemma 19. *Assume T is up-down-up, with critical points $c_1 < c_2$, with $T(c_1) = 1$ and $T(c_2) = 0$. Let p be the fixed point in $[c_1, c_2]$.*

(i) *If $T(1) \ge p$, then $s > \sqrt{2}$.*
(ii) *If $T(1) \ge c_2$, then $s > 4^{1/3}$.*
(iii) *$T(0) < T(1)$ iff $s > 2$.*

Proof. (i) Note that $c_2 = c_1 + \frac{1}{s}$, so $T(1) = s(1 - c_2) = s - c_1 s - 1$. The inequality $T(1) \ge p$ is the same as

$$s - c_1 s - 1 \ge \frac{1 + c_1 s}{1 + s}.$$

Rearranging gives $s^2 - 2 \ge c_1 s(s + 2)$. Since $c_1 s(s + 2) > 0$, then $s > \sqrt{2}$.

(ii) Observe that $T(1) \ge c_2$ is equivalent to $s - c_1 s - 1 \ge c_1 + 1/s$. Rearranging gives $s^2 - s - 1 \ge c_1 s(1 + s)$. The positive root of $s^2 - s - 1$ is $\frac{1+\sqrt{5}}{2}$, so $T(1) \ge c_2$ implies $s > \frac{1+\sqrt{5}}{2} \approx 1.618$. This is larger than $4^{1/3} \approx 1.587$.

(iii) Since $T(0) = 1 - c_1 s$ and $T(1) = s - c_1 s - 1$, then (iii) follows.

□

Theorem 20. *Let T be up-down-up, with critical points $c_1 < c_2$ and $T(c_1) = 1$, $T(c_2) = 0$. Let p be the fixed point of T in $[c_1, c_2]$. Then $s > 2$ iff $T(0) < T(1)$, and in that case T is exact. If $s \le 2$, then $T(1) \le T(0)$, and we have the following cases for the location of p with respect to the interval $[T(1), T(0)]$.*

(i) *If $p < T(1)$ or $p > T(0)$, then T is topologically exact.*
(ii) *If $T(0) = p$ or $T(1) = p$, then T is transitive but not topologically exact, T exchanges $[0, p]$ and $[p, 1]$, and T^2 restricted to each of these intervals is topologically exact.*
(iii) *If $T(1) < p < T(0)$, then T is not transitive.*

Proof. By Lemma 19(iii) we have $T(0) < T(1)$ iff $s > 2$, and in that case T is exact by Theorem 14. Hereafter, we assume $s \le 2$.

(i) Assume $p < T(1)$ or $T(0) < p$. Without loss of generality, we may assume $T(1) > p$. (Otherwise, if $T(0) < p$, replace T by its reflection $\widetilde{T}$, which will satisfy $\widetilde{T}(1) > \widetilde{p}$, where $\widetilde{p} = 1 - p$ is the fixed point of $\widetilde{T}$.)

Let K be any subinterval of $[0, 1]$. Suppose (to reach a contradiction) that no iterate of K contains p in its interior (and thus also that no iterate of K contains both critical points). We consider two possibilities.

First suppose $T(1) \ge c_2$. By Lemma 19, $s > 4^{1/3}$. By Lemma 18, at most two of K, $T(K)$, $T^2(K)$ contain a critical point, so $|T^3(K)| \ge \frac{s^3}{4}|K|$. Thus each application of T^3 to any iterate of K increases its length by a factor at least $s^3/4 > 1$, which is impossible.

Now suppose $T(1) < c_2$. If either K or $T(K)$ contains no critical point, then applying T^2 increases the length of K by a factor of at least $s^2/2$, which is greater than 1 by Lemma 19(i). On the other hand, suppose both K and $T(K)$ contain critical points. By Lemma 18, $T^2(K)$ and $T^3(K)$ contain no critical points. Thus applying T^4 to K increases its length by at least a factor of $s^4/4 = (s^2/2)^2 > 1$, cf. Lemma 19(i). Thus we may repeatedly apply one of T^2 or T^4 to steadily increase the length of iterates of K by a factor strictly greater than one, a contradiction.

We conclude that some iterate of K must contain p in its interior, and thus by Lemma 4, some iterate of K equals $[0,1]$. Thus T is exact.

(ii) Assume $T(1) = p$ or $T(0) = p$. Without loss of generality, we may assume $T(1) = p$. (Otherwise, if $T(0) = p$, replace T by its reflection $\widetilde{T}$.) See Figure 6 for possible graphs of T and T^2 in the case $T(1) = p$.

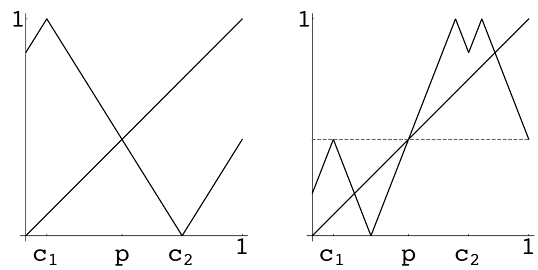

FIGURE 6. T and T^2 when $T(1) = p$

Since $s \leq 2$, by Lemma 19(iii) we have $T(0) \geq T(1) = p$. Thus T maps $[0,p]$ onto $[p,1]$, and maps $[p,1]$ onto $[0,p]$. The turning points of T^2 on $[0,p]$ are c_1 and the T-preimages of c_2, so T^2 on $[0,p]$ is either bimodal (as illustrated in Figure 6) or trimodal. On $[0,p]$ the values of T^2 at turning points are either 0 or p, and the slopes of T^2 will be $\pm s^2$, with $s^2 > 2$ by Lemma 19. Thus when T^2 is repeatedly applied to an interval in $[0,p]$, the length gets increased by at least the factor $s^2/2 > 1$ until some iterate contains two critical points of T^2 in $[0,p]$. The range of T^2 on the interval between successive critical points is $[0,p]$, so T^2 on $[0,p]$ is topologically exact.

Since T exchanges $[0,p]$ and $[p,1]$, it follows that T^2 is exact on $[p,1]$ as well, and that T is transitive on $[0,1]$. (For transitivity, see the argument in the proof of Theorem 9(ii).)

(iii) Assume $T(1) < p < T(0)$. Then there is an open neighborhood of p in which each point has a unique preimage, so by Lemma 5, T is not transitive.

□

5. SUMMARY

Theorem 21. *A unimodal or bimodal uniformly piecewise linear map $T : [0,1] \to [0,1]$ is transitive iff there is a subinterval J such that the following three conditions are satisfied.*

(i) *T is linear on J with $T(J) = [0,1]$, with neither endpoint of J fixed by T.*
(ii) *The fixed point p in J is contained in the closure of $T([0,1] \setminus J)$.*
(iii) *If p is in the boundary of $T([0,1] \setminus J)$, then T exchanges $[0,p]$ and $[p,1]$.*

Furthermore, T will be topologically exact iff an interval J can be chosen such that (i) holds and such that p is contained in the interior of $T([0,1] \setminus J)$.

Proof. If T is unimodal, the theorem follows from Theorem 9. Suppose now that T is bimodal. If T is transitive, then there exists an interval satisfying (i) by Lemma 12, so suppose that J is such an interval. (Note that there could be more than one such interval.)

We first consider the case where T is down-up-down. Let the critical points of T be $c_1 < c_2$. In the case where $J = [c_1, c_2]$, the theorem follows from Theorem 15. If $J = [0, c_1]$ or $J = [c_2, 1]$, the theorem follows from Theorem 17 (and the remark following that theorem).

Now suppose T is up-down-up. If $s > 2$ then T is exact by Theorem 20, and $T(0) < T(1)$, so either $p < T(1)$ or $p > T(0)$. In both cases p is in the interior of $T([0,1] \setminus J)$. If $s \leq 2$, the theorem follows from (i), (ii), (iii) in Theorem 20.

□

Remarks Combining Theorem 21 with Parry's theorem (cf. Theorem 6), every transitive unimodal or bimodal map τ is conjugate to a uniformly piecewise linear map T satisfying the conditions of Theorem 21. It can be shown that the map T is unique up to reflection, cf. [11].

Observe that T satisfies condition (ii) in Theorem 21 precisely if there is no open neighborhood of the fixed point p in which each point has unique preimage. Thus Theorem 21 is almost the converse of Lemma 5 (and in fact, is precisely the converse in the unimodal case). However, for a down-up-down map with $s = 2$, we have $T(1) = T(0) = p$, so there is no such neighborhood, and yet the map is not transitive, cf. Theorem 15. In this case the intervals $[0, p]$ and $[p, 1]$ are invariant, and T is transitive (in fact exact) on each.

In Theorem 21, since T is uniformly piecewise linear, we could rephrase (i) by requiring T to be monotone on J instead of linear; call that rephrased condition (i)$'$. If $\tau : [0,1] \to [0,1]$ is unimodal or bimodal, and is transitive, let $h : [0,1] \to [0,1]$ be a conjugacy from τ onto a uniformly piecewise linear map T, cf. Theorem 6. Then T is transitive, so satisfies the conditions of Theorem 21. Since h is a homeomorphism and is either strictly increasing or strictly decreasing, it follows that τ also satisfies the same conditions (i)$'$, (ii), (iii). Thus these conditions are necessary for τ to be transitive. They are not sufficient, as can be seen by considering the logistic map $\tau(x) = kx(1-x)$ where $k = 3.839$ (restricted to the interval $[\tau^2(1/2), \tau(1/2)]$, and rescaled to live on $[0,1]$). For the fixed point p, we have $\tau(0) < p$, so τ satisfies (i)$'$, (ii), (iii). However, for this value of k, τ has an attracting 3-cycle, so is not transitive, cf. [3, §1.13].

References

[1] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey, On Devaney's Definition of Chaos, *Amer. Math. Monthly* **99** (1992), no. 4, 332–334.
[2] S. Bassein, The dynamics of a family of one-dimensional maps, *Amer. Math. Monthly* **105** (1998), no. 2, 118-130.
[3] R. Devaney, *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley, 1987.
[4] S. Elaydi, *Discrete chaos*, Chapman & Hall/CRC, Boca Raton, FL, 2000.
[5] L. Jonker and D. Rand, Bifurcations in one dimension. I. The nonwandering set. *Invent. Math.* **62** (1981), no. 3, 347–365.
[6] R. Kraft, Chaos, Cantor Sets, and Hyperbolicity for the Logistic Maps, *Amer. Math. Monthly* **106** (1999), no. 5, 400–408.
[7] J. Milnor and Thurston, On iterated maps of the interval I, II. Dynamical Systems: Proc. Univ. of Maryland 1986-87 (1988), Lect. Notes In Math 1342, Springer, Berlin, NY, 465–563.

[8] W. Parry, Symbolic dynamics and transformations of the unit interval. *Trans. Amer. Math. Soc.* **122** (1966), 368–378.

[9] C. Preston, *Iterates of piecewise monotone mappings on an interval.* Lecture Notes in Mathematics, 1347. Springer-Verlag, Berlin, 1988.

[10] F. Shultz, Dimension groups for interval maps II: The transitive case. *Ergodic Theory Dyamical Systems*, to appear.

[11] F. Shultz, Uniqueness of Parry maps, and invariants for transitive piecewise monotonic maps, preprint.

[12] D. Veitch and P. Glendinning, Expliicit renormalization in piecewise linear bimodal maps, Physica D **44** (1990), 149-167.

[13] M. Vellekoop and R. Berglund, On intervals, transitivity = chaos. *Amer. Math. Monthly 101* (1994), no. 4, 353–355.

[14] P. Walters, *An introduction to ergodic theory.* Graduate Texts in Mathematics, **79**. Springer-Verlag, New York-Berlin, 1982.

Wellesley College, Wellesley, Massachusetts 02481
E-mail address: fshultz@wellesley.edu