Symbolic Dynamics and Sarkovskii's Theorem

1. Dynamical systems.

All spaces are metric spaces.

A *dynamical system* is a pair (X, f) consisting of a space X and a continuous map $f : X \to X$. For $x \in X$ define the *orbit* of x as $\mathcal{O}(x) = \{ f^n(x) : n \in \mathbb{N} \}$. We say $x \in X$ is *periodic* with period $n > 0$ iff $f^n(x) = x$. Such n is called a *period* of x, and the least period is sometimes called the prime period (although it need not be prime). Also, x is called *eventually periodic* iff some $f^k(x)$ is periodic; this is equivalent to requiring that the orbit of x is finite.

If (X, f) and (Y, g) are dynamical systems then a continuous surjection $h : X \to Y$ satisfying $g \circ h = h \circ f$ is called a *semi-conjugacy* from (X, f) to (Y, g), and (Y, g) is called a *quotient* of (X, f). If h is also a homeomorphism then it is called a *conjucacy* from (X, f) to (Y, g), and one says that (X, f) is conjugate to (Y, g) in this case. Of course this defines an equivalence relation on dynamical systems.

Suppose (X, f) is a dynamical system and $Y \subset X$. Then Y is *invariant* iff $fY \subset Y$, in which case the subsystem $(Y, f|Y)$ is defined.

2. Symbols and shifts.

Let A be a finite non-empty set, called the *alphabet* or set of *symbols*. We denote by A^* the set of finite words in the alphabet A. That is, A^* consists of finite sequences $\alpha = \alpha_0 \dots \alpha_n$ of elements of A. Here $|\alpha| = n + 1$ is called the *length* of α. The word of length 0 will be written ϵ; it is called the empty word. Then A^* is a monoid, with the operation of concatenation: $\alpha\beta$ is the word $\alpha_0 \dots \alpha_n \beta_0 \dots \beta_m$. Note that ϵ is the identity element. We also define $A^+ = A^* \setminus \{\epsilon\}$, the set of words of positive length.

We denote by A^∞ the set of infinite sequences $\alpha = \langle \alpha_0 \alpha_1 \dots \rangle$ of elements of A. Note that A^* acts on A^∞ on the left: If $\alpha \in A^*$ has length $n+1$ and $\beta \in A^\infty$ then $\alpha\beta \in A^\infty$ is defined as $\langle \alpha_0 \dots \alpha_n \beta_0 \beta_1 \dots \rangle$. We say $\alpha \in A^*$ is a *subword* of $\beta \in A^*$ (or A^∞) iff $\beta = \gamma\alpha\delta$ for some γ, δ. If γ is empty then α is called a *prefix* of β, and δ is called a *suffix*.

We also define the concatenation of an infinite sequence α_j of words in A^* in the obvious way. It lies in A^* if all but finitely many α_j's have length 0; otherwise it is in A^∞. In particular, if $\alpha \in A^+$ then the concatenation of infinitely many α's is written $\alpha^\infty \in A^\infty$.

Next we specify a topology on A^∞. This is just the product topology obtained by putting the discrete topology on A, so A^∞ is compact. It can be metrized as follows. For $a, b \in A$ let $\delta(a, b)$ be 0 if $a = b$, and 1 if $a \neq b$. Then for $\alpha, \beta \in A^\infty$ define

$$d(\alpha, \beta) = \sum_{k=0}^{\infty} \frac{\delta(\alpha_k, \beta_k)}{3^k}.$$

This is a metric defining the product topology on A^∞. The following is a convenient description of certain balls:

$$B(\alpha, \tfrac{1}{3^n}) = \overline{B}(\alpha, \tfrac{1}{2 \cdot 3^n}) = \{ \beta \in A^\infty : \forall j \leq n,\ \beta_j = \alpha_j \}.$$

That is, elements of A^∞ are close iff they share a sufficiently large prefix.

For a dynamical system we need a map. We define the *shift* σ from A^∞ to itself and also from A^+ to A^* by $(\sigma\alpha)_k = \alpha_{k+1}$. This is continuous on A^∞, and, if A has m elements, the dynamical system (A^∞, σ) is called the *full shift on m symbols*. If $\Sigma \subset A^\infty$ is invariant under σ then the subsystem $(\Sigma, \sigma|\Sigma)$ is called a *subshift*. Since the map in such a case is always the restriction of the shift one usually refers to Σ itself as the subshift.

We note the following dynamical properties of the full shift defined by an alphabet A with m symbols. First, a point is periodic with period n iff it has the form α^∞ for some $\alpha \in A^+$ of length n. Thus there are m^n points of period n. Also, a point is eventually periodic iff it has the form $\beta\alpha^\infty$ with α as above. The periodic points are dense in A^∞: if $\alpha \in A^\infty$ and β is the prefix of α of length $n+1$ then $d(\alpha, \beta^\infty) < 3^{-n}$. Also, there is a point with a dense orbit: for example, arrange all the words of A^+ in a sequence and form their infinite concatenation.

3. Subshifts of finite type.

Suppose F is a set of words of length 2. Then we define $\Sigma_F \subset A^\infty$ as the set of sequences α such that every subword of α of length 2 lies in F. Then Σ_F is compact and invariant under the shift, and is called a *subshift of finite type*. If the elements of A are the integers from 1 to m then such an F can be conveniently encoded by an $m \times m$, 0–1 matrix M, with the understanding that $M_{ab} = 1$ if $ab \in F$ and $M_{ab} = 0$ otherwise. In this case we write Σ_M instead of Σ_F. We will use the matrix formalism hereafter.

We can also define a *language* (that is, a subset of A^*) L_M by the same requirement on subwords of length 2 as for Σ_M.

The full shift is a subshift of finite type: just set $M_{ab} = 1$ for all a, b. The following give analogues of the dynamical properties of the full shift.

Lemma.

a) Splicing: If $\alpha, \beta \in L_M$ have positive lengths $n+1, m+1$ respectively, and $\alpha_n = \beta_0$ then

$$\alpha_0 \ldots \alpha_n \beta_1 \ldots \beta_m \in L_M$$

. *b) Limits: $\alpha \in A^\infty$ lies in Σ_M iff every prefix of α lies in L_M.*

c) Periodic points: There is a one-one correspondence between points of Σ_M of period n and words $\alpha \in L_M$ of length $n+1$ satisfying $\alpha_0 = \alpha_n$.

d) Counting words: $(M^n)_{ab}$ is the number of words $\alpha \in L_M$ of length $n+1$ satisfying $\alpha_0 = a$, $\alpha_n = b$.

e) Counting periodic points: The number of periodic points in Σ_M is $\operatorname{tr}(M^n)$.

Lemma. *Suppose*

a) For all $a, b \in A$ there is $k > 0$ such that $(M^k)_{ab} > 0$.

Then:

b) Gluing: If α, $\beta \in L_M$ then there is $\gamma \in L_M$ such that $\alpha\gamma\beta \in L_M$.

c) In Σ_M the periodic points are dense and there is a point with a dense orbit.

Moreover, these three conditions are equivalent if M contains no all-zero row or column.

4. The basic construction

Suppose (X, f) is a dynamical system and $I \subset X$ is compact. We set

$$\Lambda = \{\, x : \mathcal{O}(x) \subset I \,\} = \bigcap_{n=0}^{\infty} f^{-n} I.$$

This is the maximal invariant set contained in I; from the second form above it is clear that it is compact. Suppose A is a finite set and $\{\, I_a : a \in A \,\}$ is a collection of compact sets whose union is I. We say $\alpha \in A^{\infty}$ (or A^{+}) is an *itinerary* (or *partial itinerary*) of $x \in I$ iff $f^k x \in I_{\alpha_k}$ for all $k \geq 0$ (and $k \leq |\alpha|$). From the definitions it is clear that $x \in I$ has an itinerary iff $x \in \Lambda$.

Notice that $\alpha_0 \ldots \alpha_n$ is a partial itinerary of x iff

$$x \in I_\alpha = \{\, x : f^k \in I_{\alpha_k} \text{ for } 0 \leq x \leq n \,\} = \bigcap_{k=0}^{n} f^{-k} I_{\alpha_k}.$$

We have $I_\beta \supset I_\alpha$ if β is a prefix of α, so, for $\alpha \in A^{\infty}$,

$$I_\alpha = \bigcap_{n=0}^{\infty} I_{\alpha_0 \ldots \alpha_n}.$$

is a nested intersection. Clearly, $x \in I_\alpha$ iff α is an itinerary of x. We define Σ to be the set of itineraries of points of Λ; alternatively, $\Sigma = \{\, \alpha \in A^{\infty} : I_\alpha \neq \emptyset \,\}$.

Lemma.
a) For $\alpha \in A^{\infty}$, $I_\alpha \neq \emptyset$ iff $I_\beta \neq \emptyset$ for all prefixes β of α.
b) α is an itinerary of $x \in \Lambda$ iff each prefix of α is a partial itinerary of x.
c) For α in A^{∞} or A^{+}, $f I_\alpha \subset I_{\sigma\alpha}$.
d) Σ is a compact subshift.
e) $\Lambda = \bigcup_{\alpha \in \Sigma} I_\alpha$.

We say $\{I_a\}$ is an *expansive cover* of Λ iff I_α is a singleton for each $\alpha \in \Sigma$. In this case we define $\pi : \Sigma \to \Lambda$ by $I_\alpha = \{\pi(\alpha)\}$.

Lemma. *If $\{I_a\}$ is an expansive cover then π is a semi-conjugacy from $(\Sigma, \sigma|\Sigma)$ to $(\Lambda, f|\Lambda)$.*

We need a simple condition to ensure an expansive cover.

We say f is *expanding* on $Y \subset X$ (not necessarily invariant) iff there is $\lambda > 1$ such that $d(fx, fy) \geq \lambda d(x, y)$ for all $x, y \in Y$.

Lemma. *If f is expanding on I_a for each $a \in A$ then $\{I_a\}$ is an expansive cover.*

Next, let $B = \bigcup \{\, I_a \cap I_b : a, b \in A,\ a \neq b \,\}$. If $\alpha, \beta \in \Sigma$ are distinct and $I_\alpha \cap I_\beta \neq \emptyset$ then let k be the first index for which $\alpha_k \neq \beta_k$. Then $f^k(I_\alpha \cap I_\beta) \subset I_{\alpha_k} \cap I_{\beta_k} \subset B$.

Lemma. *For an expansive cover, π is a conjugacy iff $B \cap \Lambda = \emptyset$.*

Moreover, the failure of π to be injective can be analysed in terms of B: $\pi^{-1} x$ contains more than one point iff $f^k x \in B$ for some $k \geq 0$, and $\pi\alpha = \pi\beta$ iff $\alpha = \delta\gamma$ and $\beta = \delta\gamma'$ where γ, γ' are distinct itineraries of the same point in $B \cap \Lambda$.

If the inverse images of π are understood then we can recreate (X, Λ). Namely, introduce an equivalence relation on Σ by $\alpha \sim \beta$ iff $\pi\alpha = \pi\beta$ and let Q be the set of equivalence classes $[\alpha]$. The topology on Q is defined by declaring $U \subset Q$ to be open iff $\bigcup U$ is open in Σ. Then we define $\phi : \Sigma \to Q$, $\psi : Q \to \Lambda$, and $q : Q \to Q$ by $\phi(\alpha) = [\alpha]$, $\psi([\alpha]) = \pi\alpha$, and $q([\alpha]) = \sigma\alpha$. These are well-defined and continuous; ϕ is a semi-conjugacy from $(\Sigma, \sigma|\sigma)$ to (Q, q); and ψ is a conjugacy from (Q, q) to $(\Lambda, f|\Lambda)$.

Lemma. *Suppose (X, f) and (X', f') are dynamical systems. Suppose there are expansive covers of invariant sets $\Lambda \subset X$ and $\Lambda' \subset X'$, defining semi-conjugacies $\pi : \Sigma \to \Lambda$ and $\pi' : \Sigma' \to \Lambda'$. If $\Sigma = \Sigma'$ and $\pi^{-1}(\pi\alpha) = \pi'^{-1}(\pi'\alpha)$ for all $\alpha \in \Sigma$ then $(\Lambda, f|\Lambda)$ and $(\Lambda', f'|\Lambda')$ are conjugate.*

In general, if one can find a simpler shift than Σ which still covers Λ under π then it will be easier to analyze the dynamics of Λ. That is the point of the next condition on the I_a's. We say that $\{I_a\}$ satisfies the *Markov condition* iff each fI_a is the union of various of the I_b's. In this case, assuming $A = \{1, \ldots, m\}$, we define a matrix M by $M_{ab} = 1$ if $fI_a \supset I_b$, and $M_{ab} = 0$ otherwise.

Lemma. *If $\{I_a\}$ is an expansive Markov cover then $\Sigma_M \subset \Sigma$ and $\pi\Sigma_M = \Lambda$.*

5. Sarkovskii's Theorem.

Define an ordering $\rhd$ on the natural numbers as follows:

$$2^j(2r+1) \rhd 2^k(2s+1) \iff \begin{cases} j < k, r > 0, s > 0 & \text{or} \\ j = k, r > 0, r < s & \text{or} \\ r > 0, s = 0 & \text{or} \\ j > k, r = s = 0. \end{cases}$$

That is,

$$3 \rhd 5 \rhd 7 \rhd \ldots \rhd 2\cdot 3 \rhd 2\cdot 5 \rhd \ldots \rhd 4\cdot 3 \rhd 4\cdot 5 \rhd \ldots \rhd 2^3 \rhd 2^2 \rhd 2 \rhd 1.$$

Sarkovskii's Theorem. *If a dynamical system $([0,1], f)$ has a periodic point of least period p and $p \rhd q$ then it has a periodic point of least period q.*

The proof uses a modification of the basic symbolic dynamics construction.

So suppose $\{x_1, \ldots, x_p\}$ is a periodic orbit with $p > 1$, arranged so that $x_j < x_{j+1}$ for $1 \le j < p$. Note that f permutes these. Let $m = p - 1$.

We let $\mathcal{V}$ be the collection of intervals of the form $[x_i, x_j]$ with $i < j$, and we define $f^* : \mathcal{V} \to \mathcal{V}$ by $f^*(J) = [\min fx_k, \max fx_k]$, where the x_k's are those in J. Clearly, if $I, J \in \mathcal{V}$ and $I \cap J \ne \emptyset$ then $f^*(I \cup J) = f^*(I) \cup f^*(J)$. We say $I \in \mathcal{V}$ *covers* $J \in \mathcal{V}$, and write $I \to J$, iff $f^*(I) \supset J$. We define the *length* of an interval $I = [x_j, x_k] \in \mathcal{V}$ as $k - j$. Note that the length of I is one less than the number of points in $\mathcal{O}(x_1) \cap I$. Hence f^* does not decrease lengths.

Now let $J_a = [x_a, x_{a+1}]$ for $a \in A = \{1, \ldots, m\}$ and define the 0–1 matrix M by $M_{ab} = 1$ iff $J_a \to J_b$.

We define J_α for $\alpha \in L_M$ recursively to have the following properties:

- *a) J_α is a closed nonempty interval;*
- *b) $fJ_\alpha = J_{\sigma\alpha}$ if $|\alpha| > 1$;*
- *c) $J_{\alpha_0 \ldots \alpha_{n+1}} \subset J_{\alpha_0 \ldots \alpha_n}$ for $n \ge 0$.*

This goes as follows: Let $J_\epsilon = [x_1, x_p]$. For $|\alpha| = 1$ we have already defined

$fJ_{ab} = J_b$. For $n > 1$ we have $fJ_{\alpha_0 \ldots \alpha_{n-1}} = J_{\alpha_1 \ldots \alpha_{n-1}} \supset J_{\alpha_1 \ldots \alpha_n}$, so we can choose $J_{\alpha_0 \ldots \alpha_n}$ to continue the definition.

We now define J_α for $\alpha \in \Sigma_M$ by

$$J_\alpha = \bigcap_{n=0}^{\infty} J_{\alpha_0 \ldots \alpha_n}.$$

This is a nested intersection of non-empty compact intervals, so it is a non-empty compact interval (possibly a singleton). It follows from this and $fJ_\alpha = J_{\sigma\alpha}$ for finite α's that $fJ_\alpha = J_{\sigma\alpha}$ for $\alpha \in \Sigma_M$. If $\alpha \in \Sigma_M$ is periodic with period n then $f^n J_\alpha = J_\alpha$, so there is a point of period n in J_α.

Consider x_p. If this lies in J_α for some $\alpha \in \Sigma_M$ then $\alpha_0 = m$ and x_p is an endpoint of J_{α_0}. Moreover, fx_p is one of the endpoints of $f^*(J_m)$, and J_{α_1} is a subinterval of $f^*(J_m)$, so α_1 is uniquely determined and fx_p is an endpoint of J_{α_1}. Continuing inductively, we conclude that α is uniquely determined.

Now suppose $\alpha, \beta \in \Sigma_M$ are periodic and $J_\alpha \cap J_\beta \neq \emptyset$. Then, as in the last section, for some $n \geq 0$ we have $J_{\sigma^n\alpha} \cap J_{\sigma^n\beta} \subset \bigcup_{a \neq b} J_a \cap J_b \subset \mathcal{O}(x_1)$. Thus for some $r \geq 0$, $x_p \in J_{\sigma^r\alpha} \cap J_{\sigma^r\beta}$, so $\sigma^r\alpha = \sigma^r\beta$. By periodicity, this implies $\alpha = \beta$. Hence

Lemma. *If $\alpha, \beta \in \Sigma_M$ are periodic then $J_\alpha \cap J_\beta = \emptyset$ if $\alpha \neq \beta$. Moreover, if $\alpha \in \Sigma_M$ has least period n then J_α contains a point of least period n.*

We now turn to the proof of the theorem. Since $\rhd$ is a well-ordering on the positive integers which are not powers of 2 we can make the following minimality assumption on p:

(1) If $q < p$ and f has a periodic point of least period q then q is a power of 2.

For inductive reasons we need to prove a somewhat stronger version of the theorem:

(2) $\quad p \rhd q \implies \Sigma_M$ contains a periodic point of least period q.

In fact, this is also true with $q = p$ unless p is a power of 2.

If $p = 1$ there is nothing to prove, and if $p = 2$ we have $m = 1$ and M is the 1×1 matrix $[1]$, so there is a fixed point 1^∞, proving (2). So we may assume $p > 2$.

Now note that $fx_1 > x_1$ and $fx_p < x_p$. Thus there is a last i such that $fx_i > x_i$. Let $I_1 = [x_i, x_{i+1}]$. Then $fx_{i+1} < x_{i+1}$, so $I_1 \to I_1$. If $f^*(I_1) = I_1$ then $\{x_i, x_{i+1}\}$ would be f-invariant, contradicting $p > 2$, so there is $J \neq I_1$ such that $I_1 \to J$.

Consider chains $I_1 \to I_2 \to \cdots \to I_k$ of intervals of length 1 with $k \geq 2$ and $I_2 \neq I_1$. Let C be the union of all I_k obtained this way. Then C is a non-empty disjoint union of intervals in $\mathcal{V}$. Let l be the maximum length of a component of C and let C_1 be a component of C of length l.

Now $f^*(C_1)$ is the union of intervals $f^*(J_a)$ where the J_a's lie in C_1, so $f^*(C_1) \subset C$. Then $f^*(C_1)$ lies in a component C_2 of C. This C_2 has length $\leq l$; but it contains $f^*(C_1)$ which has length at least l. So C_2 has length l and $f^*(C_1) = C_2$. Continuing, we find components C_j of C, all of length l, satisfying $f^*(C_j) = C_{j+1}$ for all $j \geq 1$. Eventually we must have $C_k = C_j$ for some $k > j$.

Renumbering, we may assume $C_1, \ldots, C_q$ are distinct and $f^*(C_q) = C_1$. Clearly, $\mathcal{O}(x_1) \cap \bigcup C_j$ is f-invariant, so $\bigcup C_j \supset \mathcal{O}(x_1)$. We conclude that

$$q(l+1) = p. \tag{3}$$

There are now two cases:

CASE 1: $q = 1$. Hence $I_1 \subset C_1 = C = [x_1, x_p]$, so there is a chain $I_1 \to I_2 \to \cdots \to I_k \to I_1$ with each I_j of length 1 and $I_2 \neq I_1$. If we take k minimal then the I_j's are distinct; otherwise the chain would contain a loop and we could shorten it. Hence $k \leq m$. Now cycles $I_1 \to I_1 \to \ldots I_1 \to I_2 \to \cdots \to I_k \to I_1$ (with a block of n I_1's) define periodic sequences in Σ_M of least period $n+k$. If $k < m$ then either k or $k+1$ would be odd, greater than 1, and less than p, contrary to (1). Hence $k = m$ and we have produced points of least period n for all $n \geq m = p-1$. The same argument now shows that p is odd and that Σ_M contains periodic points of all least periods $\geq m$.

To finish the proof of (2) in the first case we need periodic points of least period n for all even $n < m$. We first analyse the geometry of the I_j's. (The rest of this paragraph is best understood by drawing the appropriate pictures.) We note the following

(a) $I_1 \not\to I_j,\ j > 2$;
(b) $I_j \not\to I_1,\ 1 < j < m$;
(c) $I_r \not\to I_s,\ r+1 < s$

because if any of these failed we could get a shorter cycle through I_1. From these it follows by induction that $J_j = (f^*)^j(I_1)$ is an interval of length exactly j for $0 \leq j \leq m$. Furthermore, f takes one endpoint of J_j to the other (since $f^*(J_j)$ differs from J_j by just one interval). It follows from this and (b) that the intervals I_j are added on alternate sides of I_1, so that $J_m = I_m I_{m-2} \ldots I_2 I_1 I_3 \ldots I_{m-3} I_{m-1}$ or its mirror image. Since $I_m \to I_1$ and $I_{m-2}) \to I_{m-1}$, taking the left endpoint of I_{m-2} to the right endpoint of I_{m-1}, we conclude $I_m \to I_j$ for all odd $j < m$.

The points of even period less than m required for (2) come from the cycles $I_m \to I_j \to I_{j+1} \to \cdots \to I_m$ for odd j.

CASE 2: $q > 1$. We analyse this case in two stages: first we consider f^q on C_1 and then we consider f on the intervals between the C_j's.

For $K \in \mathcal{V}$, if $x_k \in K$ then $f x_k \in f^*(K)$. So, inductively, $f^n x_k \in (f^*)^n(K)$ for all $n > 0$, and it follows that

$$(f^n)^*(K) \subset (f^*)^n(K). \tag{4}$$

We renumber so that $x_1 \in C_1$, and so $C_1 = [x_1, x_{l+1}] = J_1 \cup \ldots J_l$. Then f^q permutes the points $x_1, \ldots, x_{l+1}$, so x_1 is periodic under f^q. Since x_1 has least period p for f we conclude that it has least period $l+1$ for f^q, with orbit $\{x_1, \ldots, x_{l+1}\}$. Based on the intervals $J_1, \ldots, J_l$ we construct an $l \times l$ matrix L to represent f^q, defining Σ_L. Suppose $L_{ab} = 1$. Then $(f^q)^*(J_a) \supset J_b$. Then, by (4) with $n = q$, we can find $J_{c_j} \subset C_j$ for $1 < j \leq q$ such that, under f, $J_a \to J_{c_1} \to \cdots \to J_{c_q} \to J_b$. Applying this to consecutive letters in $\alpha \in \Sigma_L$ gives an expanded $\beta \in \Sigma_M$, which can be taken periodic if α is periodic. We conclude

(5)
If Σ_L has a periodic point with least period n then Σ_M has a periodic point of least period qn. ■

In particular, by the original argument producing I_1, Σ_L contains a fixed point, so Σ_M contains a periodic point of least period q. From (1) we conclude that q is a power of 2, say $q = 2^n$ with $n > 0$. Since $q|p$ we have $p = 2^r(2s+1)$ with $r \geq n$ and $s > 1$ or $r > n$ and $s = 0$. In either case we argue by induction on p (in the usual ordering) that Σ_L contains a periodic point of least period k for each k for which $l + 1 = 2^{r-n}(2s+1) \rhd k$. By (5), Σ_M has a periodic point of least period $t = 2^n k$ for all such k. This covers all t such that $p \rhd t$ except $t = 2^j$ for $0 \leq j < n$.

Now number the intervals of length one in $\mathcal{V}$ which do not lie in any C_j as $D_1, \ldots, D_{q-1}$. These are disjoint, since every x_j lies in some C_j. Since $(f^p)^*(D_j) = D_j$ we have $(f^*)^p(D_j) \supset D_j$ by (4), so there is a cycle of intervals of length one through D_j. However, $C_i \not\to D_k$ for all i, k, so this cycle consists only of D_k's. Such a cycle of minimal length t defines a periodic point in Σ_M of least period t. Since $t \leq q - 1$ we have $t = 2^u$ from (1). We shall show that there is exactly one such cycle for each u, $0 \leq u < n$, finishing the theorem.

First, note that these cycles must be disjoint. For if D_j was on different minimal length cycles, of lengths 2^u and 2^v with $u \leq v$, then we could form a point of Σ_M of least period $2^u + 2 \cdot 2^v$ by following the first cycle once, then the second twice, and continuing periodically. But this would have least period $k = 2^u(1 + 2^{u-v+1})$. Since $0 < u < r$ we have $k \rhd p$, contradicting (1).

So there must be a point with $u = 0$, since $q - 1$ is the sum of the lengths of the minimal cycles and $q - 1$ is odd. (Alternatively, just note that I_1 is one of the D_j's). So suppose D_j is such an interval; renumber so that $j = 1$. Since $f^*(D_1) \supset D_1$ we conclude that $D = \bigcup (f^*)^k(D_1)$ is an interval; since it contains $\mathcal{O}(x_1)$ it must be $[x_1, x_p]$. Hence there is a sequence $D_1 \to J_{a_1} \to \cdots \to J_{a_n} \to J_a$ for any a. Now if D_i, for $i \neq 1$, also satisfied $D_i \to D_i$ then we would also have such sequences starting at D_i. Hence there would be a cycle containing both D_1 and D_i so, by repeating D_1 once if necessary, and then continuing on the cycle from D_1 through D_i and back to D_1, we could construct points in Σ_M with odd least period greater than 1, contradicting (1).

Other intervals with cycles of least period 2^u for $1 \leq u < n$ now appear by induction. Namely, removing the interior of D_1 from $[x_1, x_p]$ leaves two intervals which are exchanged by f^*, and the above argument applies to f^2 to find a unique cycle of length 1 under f^2 in each component; these then form the unique cycle of length 2 under f. The formal induction is left to the reader.