Two Topics in Dynamics

James Denvir

A thesis submitted to the University of Warwick
for the degree of Ph.D.

Mathematics Institute, University of Warwick

September 1994

To the memories of

Rolph Schwarzenberger (1936–92)

and Justin Fielder (1968–94).

Table of Contents.

Acknowledgements

I feel very fortunate to have been supervised by Professor Robert MacKay. Robert is a respected mathematician, who has the ability to explain and teach the subject in an understandable way. I would like to thank him for always being available to discuss mathematics with me and for spending many hours helping me to produce the results given in this thesis.

I have also benefitted from mathematical discussions with other students, staff and visitors at the Mathematics Institute in Warwick University. In particular I would like to thank John Parker and Phil Boyland for the time they have given.

There are many other people who in many ways have made the writing of this thesis easier. I hope that my friends are aware of the value I place on the help and support they have given me and that they understand that there is not room to mention them all here.

I would like to thank my parents, who first instilled in me an interest in mathematics, my sister Eleanor and my partner Deborah Sherman for all their love and support.

This thesis is dedicated to two friends who died during the time I was writing it. Both of them, in different ways, have played a part in my mathematical career.

This thesis was written with partial support from the Science and Engineering Research Council and partial support from Warwick University.

Declaration

The material in this thesis is original except where otherwise stated.

Summary

This thesis consists of two independent chapters. Both present results in the field of dynamical systems.

In the first chapter we study abstract adding machines and their occurrence in unimodal maps of the interval. For unimodal maps with no aperiodic homtervals we characterize completely when adding machines occur. We also discuss their importance in relation to the boundary of positive topological entropy for two-parameter families of diffeomorphisms of the disc.

In the second chapter we prove existence of a distinguished set of geodesics on orientable Riemannian surfaces with geodesic boundary and negative Euler characteristic. This result allows us to construct a semi-equivalence between a subset of the unit tangent bundle of the surface with the arbitrary Riemannian metric and the unit tangent bundle given by the standard hyperbolic metric on this surface. The result is analogous to one of Morse [M1] for surfaces without boundary. We give a new proof of Morse's result using a method similar to the proof of our new result.

1. Adding machines in unimodal maps and maps of the disc

1·1. Introduction

In this chapter we study adding machines and their occurrence in unimodal maps of the interval $I = [-1, 1] \subset \mathbf{R}$ and discuss briefly their occurrence in diffeomorphisms of the open unit disc $D \subset \mathbf{R}^2$. An adding machine is a minimal dynamical system which is topologically conjugate to the action given by a map on a space of sequences of integers. The map acts like a counter (in some of the literature adding machines are called odometers) and adds one to the first element of a sequence, modulo a given base, with carry to the next element of the sequence.

Adding machines are interesting to study for the following reason. In parametrized families of smooth maps the first occurrence of an adding machine seems to be linked strongly with the boundary in parameter space of maps with positive topological entropy. For example, it is known for the case of one-parameter families of unimodal maps on I that the boundary of positive topological entropy is given by the Feigenbaum map, where the closure of the orbit of the critical point is a 'base 2' adding machine. This is the first occurrence of an adding machine in the family of maps. Much study has been made of the family of diffeomorphisms of the disc D known as the Hénon maps given by

$$H_{a,b}(x, y) = (1 - ax^2 + y, -bx)$$

for $a \geq -1$ and $0 \leq b \leq 1$. For $b = 0$ this reduces to the one-parameter one-dimensional case already discussed. Based on numerical experiments (see, e.g. [E-HM] and [GT]) it is conjectured that maps on the boundary of topological entropy contain an 'eventually-base-2' adding machine, and that the boundary is a piecewise smooth one-manifold in parameter space. It should be emphasized that the adding machine itself has zero topological entropy.

In this chapter we will classify adding machines of different bases up to topological conjugacy. We will then go on to examine closely the occurrence of adding machines in

one-parameter families of unimodal maps of I and identify a sub-system we call a 'blown-up adding machine' that occurs in maps of positive topological entropy. We conclude the chapter by examining the occurrence of adding machines and blown-up adding machines in the full two-sided shift on two symbols and relate this, via a theorem of Smale and Katok [S], [K], to diffeomorphisms of the disc. We conjecture that the occurrence of blown-up adding machines in such maps characterizes positive topological entropy.

1·2. Abstract Adding Machines

In this section, we will define abstract adding machines and classify them up to topological conjugacy.

Definition 1·1. *A* base sequence $\mathbf{b}$ *is a sequence of integers* $(b_n)_{n\geq 0}$, *with* $b_n \geq 2\ \forall n$.

We will use bold face letters to denote sequences which are to be regarded as base sequences.

Definition 1·2. *Let* $\mathbf{b}$ *be a base sequence. Denote by* $\mathcal{A}_{\mathbf{b}}$ *the space of sequences of integers*

$$\mathcal{A}_{\mathbf{b}} = \{a = (a_n)_{n\geq 0},\ 0 \leq a_n < b_n\ \forall n\}$$

with the metric $d(a,a') = 2^{-\min\{n : a_n \neq a'_n\}}$. *We define a map* $\alpha_{\mathbf{b}} : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$ *as follows: Let* $a \in \mathcal{A}_{\mathbf{b}}$, *suppose* $a \neq (b_0 - 1, b_1 - 1, b_2 - 1, \ldots)$. *Let* $k = \min\{n : a_n \neq b_n - 1\}$. *Then*

$$(\alpha_{\mathbf{b}}(a))_n = \begin{cases} 0, & \text{if } n < k, \\ a_n + 1, & \text{if } n = k, \\ a_n & \text{if } n > k, \end{cases}$$

and define $\alpha_{\mathbf{b}}(b_0 - 1, b_1 - 1, b_2 - 1, \ldots) = (0, 0, \ldots)$. *(We will denote* $\mathbf{0_b} = (0, 0, \ldots)$.*)*

We note that $(\mathcal{A}_{\mathbf{b}}, d)$ is a complete compact metric space.

Lemma 1·1. $\alpha_{\mathbf{b}} : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$ *is an isometry with respect to the metric* d. *It is bijective and hence a homeomorphism.*

Proof. Let $a \neq a' \in \mathcal{A}_{\mathbf{b}}$. Suppose $a_n = a'_n$ for $n < k$, $a_k \neq a'_k$, so $d(a,a') = 2^{-k}$. Then clearly $(\alpha_{\mathbf{b}}(a))_n = (\alpha_{\mathbf{b}}(a'))_n$ for $n < k$ and $(\alpha_{\mathbf{b}}(a))_k \neq (\alpha_{\mathbf{b}}(a))_k$, so $d(\alpha_{\mathbf{b}}(a), \alpha_{\mathbf{b}}(a')) = 2^{-k}$.

Given $a \in \mathcal{A}_{\mathbf{b}}$, $\alpha_{\mathbf{b}} \neq \mathbf{0}_{\mathbf{b}}$, define a' as

$$a'_n = \begin{cases} b_n - 1 & \text{if } a_k = 0 \text{ for } k \leq n \\ a_n - 1 & \text{if } a_n \neq 0 \text{ and } a_k = 0 \text{ for } k < n \\ a_n & \text{otherwise.} \end{cases}$$

Then $\alpha_{\mathbf{b}}(a') = a$, so $\alpha_{\mathbf{b}}$ is surjective. Injectivity and continuity follow from the fact that $\alpha_{\mathbf{b}}$ is an isometry. Since $\mathcal{A}_{\mathbf{b}}$ is Hausdorff and compact $\alpha_{\mathbf{b}}$ is a homeomorphism. □

Lemma 1·2. *$\alpha_{\mathbf{b}} : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$ is minimal.*

Proof. Let $a, a' \in \mathcal{A}_{\mathbf{b}}$. We have to show the orbit of a under $\alpha_{\mathbf{b}}$ passes arbitrarily close to a'.

Fix $n \geq 0$. We introduce some notation we shall use throughout our discussion of adding machines. Denote

$$B_n = \prod_{i=0}^{n} b_i$$

$$A_n = a_0 + \sum_{i=1}^{n} a_i B_{i-1} \text{ (for } n > 0;\ A_0 = a_0).$$

Similarly, write $A'_n = a'_0 + \sum_{i=1}^{n} a'_i B_{i-1}$. Then

$$\left(\alpha_{\mathbf{b}}^{B_n - A_n + A'_n}(a)\right)_k = a'_k \text{ if } k \leq n.$$

⌈To see this, first consider

$$\begin{aligned}
\alpha_{\mathbf{b}}^{A_n}(\underbrace{0,0,\ldots,0}_{n \text{ terms}}, a_{n+1}, a_{n+2}, \ldots) &= \alpha_{\mathbf{b}}^{\left(\sum_{i=1}^{n} a_i B_{i-1}\right)}(\alpha_{\mathbf{b}}^{a_0}(0,0,\ldots,0,a_{n+1},a_{n+2},\ldots)) \\
&= \alpha_{\mathbf{b}}^{\left(\sum_{i=1}^{n} a_i B_{i-1}\right)}(a_0,0,\ldots,0,a_{n+1},\ldots) \\
&= \alpha_{\mathbf{b}}^{\left(\sum_{i=2}^{n} a_i B_{i-1}\right)}\left(\alpha_{\mathbf{b}}^{a_1 B_0}(a_0,0,\ldots,0,a_{n+1},\ldots)\right) \\
&= \alpha_{\mathbf{b}}^{\left(\sum_{i=2}^{n} a_i B_{i-1}\right)}(a_0,a_1,0,\ldots,0,a_{n+1},\ldots) \\
&\ \ \vdots \\
&= a.
\end{aligned}$$

So

$$\left(\alpha_{\mathbf{b}}^{-A_n}(a)\right)_k = 0 \text{ if } k \leq n.$$

Thus

$$\left(\alpha_{\mathbf{b}}^{B_n - A_n}(a)\right)_k = 0 \text{ if } k \leq n.$$

Applying the same argument again, we have

$$\left(\alpha_{\mathbf{b}}^{A'_n}\left(\alpha_{\mathbf{b}}^{B_n - A_n}(a)\right)\right)_k = a'_k \text{ if } k \leq n.$$

Thus

$$d\left(\alpha_{\mathbf{b}}^{B_n - A_n + A'_n}(a), a'\right) \leq 2^{-n}.$$

□

Corollary 1·3. *$\mathcal{A}_{\mathbf{b}}$ has no periodic points of $\alpha_{\mathbf{b}}$.*

Proof. $\mathcal{A}_{\mathbf{b}}$ is infinite and $\alpha_{\mathbf{b}} : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$ is minimal by Lemma 1·2. A periodic orbit is a finite closed invariant set. □

Lemma 1·4. *Let $a \in \mathcal{A}_{\mathbf{b}}$ and define $T_a : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$ by*

$$T_a(a') = \lim_{n \to \infty} \alpha_{\mathbf{b}}^{A_n}(a')$$

where $A_n = a_0 + \sum_{i=1}^{n} a_i B_{i-1}$ as usual. (We call T_a a translation by a.) T_a is an isometric homeomorphism of $\mathcal{A}_{\mathbf{b}}$, and $\alpha_{\mathbf{b}} \circ T_a = T_a \circ \alpha_{\mathbf{b}}$.

Proof. First we show T_a is well defined. If $n > N$, then

$$\begin{aligned}\alpha_{\mathbf{b}}^{A_n}(a') &= \alpha_{\mathbf{b}}^{\left(\sum_{i=N+1}^{n} a_i B_{i-1}\right)}\left(\alpha_{\mathbf{b}}^{A_N}(a')\right) \\ &= \alpha_{\mathbf{b}}^{B_N\left(\sum_{i=N+1}^{n} a_i B_{i-1}/B_N\right)}\left(\alpha_{\mathbf{b}}^{A_N}(a')\right)\end{aligned}$$

and so $\alpha_{\mathbf{b}}^{A_n}(a')$ and $\alpha_{\mathbf{b}}^{A_N}(a')$ agree on the first N terms. So the sequence $\alpha_{\mathbf{b}}^{A_n}(a')$ is Cauchy and hence convergent.

T_a is clearly an isometry, hence it is injective and continuous. T_a is surjective, since, denoting

$$-a = \lim_{n \to \infty} \alpha_{\mathbf{b}}^{-A_n}(\mathbf{0}_{\mathbf{b}})$$

we have $T_a \circ T_{-a} =$ identity, as is easily verified.

Since $\mathcal{A}_{\mathbf{b}}$ is compact and Hausdorff, T_a is a homeomorphism. Finally, we note that $\alpha_{\mathbf{b}} = T_{(1,0,0,\ldots)}$ and that

$$\begin{aligned} T_a \circ T_{a'}(x) &= T_{a'} \circ T_a(x) \\ &= \lim_{n\to\infty} \alpha_{\mathbf{b}}^{A_n + A'_n}(x) \end{aligned}$$

to see that $T_a \circ \alpha_{\mathbf{b}} = \alpha_{\mathbf{b}} \circ T_a$. □

Remark 1·5. *Lemma 1·4 shows that, if $f : \Lambda \to \Lambda$ is topologically conjugate by a homeomorphism h to $\alpha_{\mathbf{b}} : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$, then we may choose arbitrarily $x \in \Lambda$ and assume $h(x) = \mathbf{0}_{\mathbf{b}}$.*

We now classify all adding machines up to topological conjugacy. It turns out that two adding machines are topologically conjugate if and only if their base sequences have the same 'prime decomposition,' up to ordering, in the sense described below. We prove this directly from the definition—another proof using spectral methods has been supplied by Buescu and Stewart [BS].

Definition 1·3. *Let* $\mathbf{b}$ *be a base sequence. For each n, write $b_n = p_{n,1} \cdot p_{n,2} \cdot \ldots p_{n,k_n}$ where each $p_{n,i}$ is prime. Then the sequence*

$$(p_{0,1}, p_{0,2}, \ldots, p_{0,k_0}, p_{1,1}, \ldots, p_{1,k_1}, \ldots, p_{2,1}, \ldots, p_{2,k_2}, \ldots)$$

is a prime decomposition of $\mathbf{b}$.

Proposition 1·6. *Let* $\mathbf{p}$ *be a prime decomposition of* $\mathbf{b}$. *Then there is a homeomorphism $h : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{p}}$ such that $h \circ \alpha_{\mathbf{b}} = \alpha_{\mathbf{p}} \circ h$. (We write this as $\mathbf{b} \sim \mathbf{p}$.)*

Proof. Denote $\mathcal{A}_{\mathbf{b}}^{\circ} = \{\alpha_{\mathbf{b}}^{m}(\mathbf{0}_{\mathbf{b}}) : m \geq 0\}$. $\mathcal{A}_{\mathbf{b}}^{\circ}$ is the set of sequences $a \in \mathcal{A}_{\mathbf{b}}$ satisfying $a_n = 0$ for $n \geq r$, for some r, and is dense in $\mathcal{A}_{\mathbf{b}}$ by Lemma 1·2. Define $h : \mathcal{A}_{\mathbf{b}}^{\circ} \to \mathcal{A}_{\mathbf{p}}$ by $h(\alpha_{\mathbf{b}}^{m}(\mathbf{0}_{\mathbf{b}})) = \alpha_{\mathbf{p}}^{m}(\mathbf{0}_{\mathbf{p}})$. h is continuous and injective on $\mathcal{A}_{\mathbf{b}}^{\circ}$ and has image $\mathcal{A}_{\mathbf{p}}^{\circ}$. We show h is uniformly continuous on $\mathcal{A}_{\mathbf{b}}^{\circ}$.

Since $\mathbf{p}$ is a prime decomposition of $\mathbf{b}$, there are integers n_k such that $P_{n_k} = B_k$ for each k. Given $n \geq 0$, there is an $n_k \geq n$. Suppose $a, a' \in \mathcal{A}_{\mathbf{b}}^{\circ}$, and $d_{\mathbf{b}}(a, a') < 2^{-k}$. Then

$a_i = a'_i$ for $0 \le i \le k$. Writing $a = \alpha_{\mathbf{b}}^m(\mathbf{0_b})$ and $a' = \alpha_{\mathbf{b}}^{m'}(\mathbf{0_b})$, it follows that $m \equiv m'$ (mod B_k). But now $h(a) = \alpha_{\mathbf{p}}^m(\mathbf{0_p})$, $h(a') = \alpha_{\mathbf{p}}^{m'}(\mathbf{0_p})$ with $m \equiv m' \pmod{P_{n_k}}$, so

$$(h(a))_i = (h(a'))_i \text{ for } 0 \le i \le n_k.$$

So $d_{\mathbf{p}}(h(a), h(a')) < 2^{-n_k} \le 2^{-n}$. Hence h is a uniformly continuous map from $\mathcal{A}_{\mathbf{b}}^\circ$ to $\mathcal{A}_{\mathbf{p}}$. So h extends continuously to all of $\mathcal{A}_{\mathbf{b}}$.

We will now show that h is surjective. Let $y \in \mathcal{A}_{\mathbf{p}}$. Pick $y^{(n)} \in \mathcal{A}_{\mathbf{p}}^\circ$ with $y^{(n)} \to y$. There exist unique $x^{(n)} \in \mathcal{A}_{\mathbf{b}}^\circ$ with $h(x^{(n)}) = y^{(n)}$. Suppose $y^{(n)} = \alpha_{\mathbf{p}}^{m_n}(\mathbf{0_p})$, so $x^{(n)} = \alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b})$. If n is suitably large, then $y_i^{(n)} = y_i$ for $0 \le i \le n_k$, i.e., for n suitably large, the sequence (m_n mod P_{n_k}) is constant. But $P_{n_k} = B_k$, so (m_n mod B_k) is eventually constant. Hence $x^{(n)}$ is convergent in $\mathcal{A}_{\mathbf{b}}$. and $h\left(\lim_{n\to\infty} x^{(n)}\right) = y$.

Now we show h is injective on all of $\mathcal{A}_{\mathbf{b}}$. Suppose $h(a) = h(a')$. Pick sequences $a^{(n)}, a^{(n)\prime} \in \mathcal{A}_{\mathbf{b}}^\circ$ with $a^{(n)} \to a, a^{(n)\prime} \to a'$. Then $h\left(a^{(n)}\right) \to h(a)$ and $h\left(a^{(n)\prime}\right) \to h(a') = h(a)$. Let $a^{(n)} = \alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b}), a^{(n)\prime} = \alpha_{\mathbf{b}}^{m'_n}(\mathbf{0_b})$. Then $h\left(a^{(n)}\right) = \alpha_{\mathbf{p}}^{m_n}(\mathbf{0_p})$, $h\left(a^{(n)\prime}\right) = \alpha_{\mathbf{p}}^{m'_n}(\mathbf{0_p})$ and for n suitably large,

$$h\left(a^{(n)}\right)_i = h\left(a^{(n)\prime}\right)_i \text{ for } 0 \le i \le P_{n_k},$$

so $m_n \equiv m'_n \pmod{B_k}$. Hence for n suitably large, $a_i^{(n)} = a_i^{(n)\prime}$ for $0 \le i \le k$, and so $a = a'$.

Hence h is a continuous bijection on $\mathcal{A}_{\mathbf{b}}$: since $\mathcal{A}_{\mathbf{b}}$ is compact and $\mathcal{A}_{\mathbf{p}}$ is Hausdorff, h is a homeomorphism. By construction we have that $h \circ \alpha_{\mathbf{b}} = \alpha_{\mathbf{p}} \circ h$, and so $\mathbf{b} \sim \mathbf{p}$. □

Proposition 1·7. *Let* $\mathbf{p}, \mathbf{q}$ *be prime base sequences (i.e., base sequences with* p_i, q_i *prime for each* i*). Then* $\mathbf{p} \sim \mathbf{q}$ *if and only if there is a bijection* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that* $p_{\sigma(i)} = q_i$ *for all* i*. (We say* $\mathbf{p}$ *is a reordering of* $\mathbf{q}$*.)*

Proof. Suppose $\mathbf{p}$ is a reordering of $\mathbf{q}$. Define $\mathcal{A}_{\mathbf{p}}^\circ, \mathcal{A}_{\mathbf{q}}^\circ$ as before, and let $h : \mathcal{A}_{\mathbf{p}}^\circ \to \mathcal{A}_{\mathbf{q}}^\circ$ be a bijection as in the proof of Proposition 1·6. Now h is uniformly continuous, since, setting $n_k = \max\{\sigma(i) : 0 \le i \le k\}$, if $a, a' \in \mathcal{A}_{\mathbf{p}}^\circ$ have $a_i = a'_i$ for $0 \le i \le n_k$, then $(h(a))_i = (h(a'))_i$ for $0 \le i \le k$. So h extends continuously to a map $h : \mathcal{A}_{\mathbf{p}} \to \mathcal{A}_{\mathbf{q}}$.

The proof that h is bijective, and hence a homeomorphism, is essentially the same as in Proposition 1·6.

Now suppose $\mathbf{p} \sim \mathbf{q}$. Let $h : \mathcal{A}_{\mathbf{p}} \to \mathcal{A}_{\mathbf{q}}$ be the appropriate homeomorphism; by Remark 1·5, we may assume $h(\mathbf{0_p}) = \mathbf{0_q}$. Given $k \geq 0$, there is an $n_k \geq 0$ such that if $a \in \mathcal{A}_{\mathbf{p}}$ with $a_i = 0$ for $0 \leq i \leq n_k$ then $(h(a))_i = 0$ for $0 \leq i \leq k$. In particular,

$$\left(h\left(a_{\mathbf{p}}^{P_{n_k}}(\mathbf{0_p})\right)\right)_i = 0 \text{ for } 0 \leq i \leq k.$$

But then

$$\left(a_{\mathbf{q}}^{P_{n_k}}\left(h(\mathbf{0_p})\right)\right)_i = 0 \text{ for } 0 \leq i \leq k,$$

i.e.,

$$\left(a_{\mathbf{q}}^{P_{n_k}}(\mathbf{0_q})\right)_i = 0 \text{ for } 0 \leq i \leq k.$$

So P_{n_k} is a multiple of Q_k. Hence the primes $q_0, \ldots, q_k$ occur in $p_0, \ldots, p_{n_k}$. Similarly, there is an m_k such that $\{p_0, \ldots, p_k\} \subset \{q_0, \ldots, q_{m_k}\}$. Hence $\mathbf{p}$ is a reordering of $\mathbf{q}$. $\square$

1·3. Kneading Theory for Unimodal Maps of the Interval

This section contains a brief summary of known material which we shall require for our section on adding machines in unimodal maps of the interval. For a full treatment of this material, see [CE] or [MT].

Definition 1·4. *Let $f : I \to I$ where I is the interval $[-1, 1] \subset \mathbf{R}$. We say that f is* unimodal *if*

(i) f is continuous,

(ii) $f(0) = 1$, and

(iii) f is strictly decreasing on $[0, 1]$ and strictly increasing on $[-1, 0]$.

Given a unimodal map f, we are interested in what happens to points under iteration of f. The method we use is to consider the itinerary of each point, as follows.

Definition 1·5. *Given a unimodal map $f : I \to I$, for each $x \in I$ we define the* itinerary *of x, $I(x)$, to be a string (either finite or infinite) of the symbols R, L, C as follows:*

(i) Suppose $f^j(x) \neq 0$ for all $j \geq 0$. Then $I(x)$ is an infinite string of R's and L's given by

$$I_j(x) = \begin{cases} L & \text{if } f^m(j) < 0 \\ R & \text{if } f^m(j) > 0 \end{cases}.$$

(ii) Suppose $f^k(x) = 0$ and $f^j(x) \neq 0$ for $0 \leq j < k$. Then $I(x)$ is a finite string $I_0(x) \ldots I_k(x)$ given by

$$I_j(x) = \begin{cases} L & \text{if } f^j(x) < 0 \\ R & \text{if } f^j(x) > 0 \\ C & \text{if } f^j(x) = 0 \end{cases}$$

for $j \leq k$.

Thus $I(x)$ is either an infinite string of R's and L's, or a finite string of R's and L's, followed by C.

If $I = I_0 \ldots I_n$ is a finite string of R's, L's and C's, and J any string of such symbols, we will write IJ for the concatenation of I and J, $IJ = I_0 \ldots I_n J_0 J_1 \ldots$. We write I^n for the n-fold concatenation of I with itself, and I^∞ for the infinite string $II \ldots$. We define the *shift map* σ on itineraries by $\sigma(I_0 I_1 \ldots) = (I_1 \ldots)$ and $\sigma(C) = I(1)$. Note that $\sigma(I(x)) = I(f(x))$. If A is a string of R's, L's and C's we will say A is an *admissible string* if either A is an infinite string of R's and L's, or if A is a finite string of R's and L's, followed by C. For any x, the itinerary $I(x)$ is an admissible string. We will now define an ordering on admissible strings. This will be done in such a way as to ensure that the map $x \mapsto I(x)$ is order preserving. First we will need to distinguish between odd and even strings of symbols.

Definition 1·6. *Let A be a finite string of the symbols R, L and C. We say that A is even (odd) if the number of R's in A is even (odd).*

Definition 1·7. *We now define an ordering $\prec$ on admissible strings. Let A, B be admissible strings, $A \neq B$. We order the individual symbols $L \prec C \prec R$. Suppose $A_i = B_i$ for $i < n$ and $A_n \neq B_n$. Then $A \prec B$ if either*

(i) $A_0 \dots A_{n-1}$ *is even and* $A_n \prec B_n$, *or*

(ii) $A_0 \dots A_{n-1}$ *is odd and* $B_n \prec A_n$.

The relation $\prec$ is a complete ordering, and we have the following:

Proposition 1·8. *If f is unimodal then the following holds:*

$$I(x) \prec I(y) \Rightarrow x < y \Rightarrow I(x) \preceq I(y).$$

Proof. See [CE], Lemma II.1.2 and II.1.3. □

We will now define maximal sequences and $*$-products, which will play a very important role in our discussion of adding machines.

Definition 1·8. *An admissible string A is* maximal *if $\sigma^k(A) \preceq A$ for $k = 1, 2, \dots, n$ if $A = A_0 \dots A_n$ and for $k = 1, 2, \dots$ if A is infinite.*

Note that every maximal sequence, with the exceptions of L^∞ and R^∞ start with the symbols $RL \dots$.

Definition 1·9. *Let A be a finite string of R's and L's, and let B be an admissible sequence. Denote $\overline{R} = L, \overline{L} = R$. We define*

$$A * B = \begin{cases} AB_0AB_1 \dots & \text{if } A \text{ is even and } B \text{ is infinite,} \\ AB_0AB_1 \dots AB_{n-1}AC & \text{if } A \text{ is even and } B = B_0 \dots B_{n-1}C \text{ is finite} \\ A\overline{B_0}A\overline{B_1} \dots & \text{if } A \text{ is odd and } B \text{ is infinite} \\ A\overline{B_0}A\overline{B_1} \dots A\overline{B_{n-1}}AC & \text{if } A \text{ is odd and } B \text{ is finite} \end{cases}$$

We also define $A * B$ for B a finite string of R's and L's, by $(A * B)C = A * (BC)$.

Suppose A,B and D are finite strings of R's and L's. Then if the length of the string A, $|A| = a - 1$ and $|B| = b - 1$ we have $|A * B| = ab - 1$. Also we have $A * (B * D) = (A * B) * D$. Now if $A^{(i)}$ is a finite string of R's and L's for each $i \in \mathbb{N}$, with $|A^{(i)}| = b_i - 1$ then for $n > N$ the product $A^{(0)} * A^{(1)} * \dots * A^{(n)} * (R^\infty)$ agrees with the product $A^{(0)} * A^{(1)} * \dots * A^{(N)} * (R^\infty)$ for the first $\prod_{i=0}^{N} b_i - 1$ terms. Hence we can define infinite $*$-products, which we write

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)}.$$

Theorem 1·9. *(See [CE] Corollary II.2.4.) If AC and B are maximal admissible strings, then $A * B$ is maximal.*

Theorem 1·10. *(See [CE] Theorem II.2.5.) If AC is a maximal admissible string, then the map $A* : B \mapsto A * B$ from the set of maximal admissible sequences to itself preserves the ordering $\prec$.*

We now relate the discussion of itineraries to discussion of orbits of points. Denote by K_f the kneading invariant of f, $K_f = I(1)$, and by $I_E(x)$ the extended itinerary of x,

$$I_E(x) = \begin{cases} I(x) & \text{if } I(x) \text{ infinite,} \\ I(x)K_f & \text{if } I(x) \text{ finite and } K_f \text{ infinite,} \\ I(x)(K_f)^\infty & \text{if } I(x) \text{ and } K_f \text{ are both finite.} \end{cases}$$

Proposition 1·11. *(See [CE] Lemma II.3.1.) $I_E(x)$ is eventually periodic (i.e., $\sigma^k(I_E(x))$ is periodic for large enough k) if and only if $f^j(x)$ converges towards a periodic orbit.*

The next theorem is very strong: it says that every 'reasonable' admissible string appears as the itinerary of some point x. By 'reasonable' we mean the string is dominated by the kneading sequence, as follows.

Definition 1·10. *An admissible string is* dominated by K_f *if, for all $k \geq 0$,*

(i) $\sigma^k(A) \prec K_f$ if K_f is infinite,

(ii) $\sigma^k(A) \prec (BL)^\infty$ if $K_f = BC$ with B even,

(iii) $\sigma^k(A) \prec (BR)^\infty$ if $K_f = BC$ with B odd.

We write this as $A \ll K_f$.

Theorem 1·12. *(See [CE] Theorem II.3.8.) Suppose f is unimodal and that A is an admissible string with*

$$I(-1) \preceq A \ll K_f.$$

Then there is $x \in [-1, 1]$ with $I(x) = A$.

There is an analogous theorem for one-parameter families of unimodal maps.

Theorem 1·13. *(See [CE] Theorem III.1.1.) Suppose f_μ, $0 \le \mu \le 1$ is a family of C^1 unimodal maps, continuous with respect to the C^1 topology. Write $K(\mu) = K_{f_\mu}$. Suppose A is a maximal string with $K(0) \prec A \prec K(1)$. Then there is a $\mu \in [0,1]$ with $K(\mu) = A$.*

We conclude this section with two new results, both lemmas about infinite $*$-products which we shall need in our discussion about adding machines.

Lemma 1·14. *For each $i \in \mathbb{N}$, let $A^{(i)}$ be a finite string of R's and L's, with $A^{(i)}C$ maximal, and let*

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)}.$$

Then X is aperiodic.

Proof. First we will show that each 'tail' $A^{(N)} * A^{(N+1)} * \cdots$ begins $RL\ldots$. Note that $A^{(N)} * A^{(N+1)} * \cdots$ is infinite and maximal. Since $A_0^{(N)} = R$ for each N,

$$\mathop{*}_{i=N}^{\infty} A^{(i)} \neq L^\infty$$

for each N. Now if $|A^{(i)}| > 1$ we have $A_1^{(i)} = L$. Suppose $A^{(N)} * A^{(N+1)} * \cdots = R^\infty$. Then we must have $|A^{(N)}| = 1$. But then

$$R^\infty = \mathop{*}_{i=N}^{\infty} A^{(i)} = R * \left(\mathop{*}_{i=N+1}^{\infty} A^{(i)} \right).$$

But this gives $A_0^{(N+1)} = L$, and then $A^{(N+1)}C$ is not maximal. Hence

$$\mathop{*}_{i=N}^{\infty} A^{(i)} \neq R^\infty.$$

Hence $A^{(N)} * A^{(N+1)} * \cdots = RL\ldots$.

Now suppose X is periodic with period n. Write $n = a_0 + \sum_{i=1}^{\infty} a_i B_{i-1}$ where $B_i = \prod_{j=0}^{i}(|A^{(j)}| + 1)$ and $0 \le a_i < B_i/B_{i-1}$. We will show that each $a_i = 0$. Now

$$X = A^{(0)} P_0 A^{(0)} P_1 \ldots$$

with $P_0 \neq P_1$. Suppose $a_0 \neq 0$. Then

$$\sigma^n X = A_{a_0}^{(0)} A_{a_0+1}^{(0)} \ldots A_{b_0-2}^{(0)} P_k A^{(0)} P_{k+1} \ldots$$

for some k. But $\sigma^n X = X$, in particular, $X_{b_0-1} = (\sigma^n X)_{b_0-1}$ gives $P_0 = A^{(0)}_{a_0-1}$ and $X_{2b_0-1} = (\sigma^n X)_{2b_0-1}$ gives $P_1 = A^{(0)}_{a_0-1}$. Hence $P_0 = P_1$, a contradiction. Thus $a_0 = 0$. Now we proceed by induction. Suppose $a_i = 0$ for $i \leq k$. Then $n = \sum_{i=k+1}^{\infty} a_i B_{i-1}$. Suppose $a_{k+1} \neq 0$. Then

$$X = (A^{(0)} * (A^{(1)} * \ldots A^{(k+1)}) P_0 (A^{(0)} * \ldots A^{(k+1)}) P_1 \ldots)$$

for some P with $P_0 \neq P_1$ and

$$\begin{aligned}\sigma^n X =& ((A^{(0)} * \cdots * A^{(k)}) A^{(k+1)}_{a_{k+1}} (A^{(0)} * \cdots * A^{(k)}) A^{(k+1)}_{a_{k+1}+1} \cdots \\ & (A^{(0)} * \cdots * A^{(k)}) A^{(k+1)}_{b_{k+1}-2} (A^{(0)} * \cdots * A^{(k)}) P_l \\ & (A^{(0)} * \cdots * A^{(k+1)}) P_{l+1} \cdots\end{aligned}$$

for some l. Now $X_{B_{k+1}-1} = (\sigma^n X)_{B_{k+1}-1}$ gives $P_0 = (A^{(0)} * \ldots A^{(k+1)})_{(b_{k+1}-a_{k+1}-1)B_k-1}$ and $X_{2B_{k+1}-1} = (\sigma^n X)_{2B_{k+1}-1}$ gives $P_1 = (A^{(0)} * \ldots A^{(k+1)})_{(b_{k+1}-a_{k+1}-1)B_k-1}$ contradicting $P_0 \neq P_1$. Hence $a_{k+1} = 0$. Thus by induction $n = 0$ and X is aperiodic. □

Lemma 1·15. *For each $i \in \mathbb{N}$, let $A^{(i)}$ be a finite string of R's and L's with $A^{(i)}C$ maximal for each i. Let*

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)}$$

and suppose that $\sigma^m(X)$ and $\sigma^n(X)$ agree on terms $0, \ldots, B_k - 1$. Then $n \equiv m \pmod{B_{k-1}}$.

Proof. Suppose $X_{m+i} = X_{n+i}$ for $0 \leq i \leq B_k - 1$. Write $m = aB_k + pB_{k-1} + r$, $n = a'B_k + qB_{k-1} + s$ where $0 \leq p, q \leq b_k - 1$ and $0 \leq r, s \leq B_{k-1} - 1$. We will show $r = s$, hence that $m \equiv n \pmod{B_{k-1}}$. Write

$$X = P * Q,$$

where

$$P = A^{(0)} * \ldots * A^{(k-1)}$$

and

$$Q = \mathop{*}_{i=k}^{\infty} A^{(i)}.$$

Note that PC and Q are maximal and that $|P| = B_{k-1} - 1$. Denote $\sigma'(Y_0 \dots Y_n) = Y_1 \dots Y_n Y_0$. Suppose $s \neq r$, without loss of generality $s > r$, and consider

$$\begin{aligned}
&\sigma'^{B_k - pB_{k-1} - r}(X_m \dots X_{m+B_k-1}) \\
&= \sigma'^{B_k - pB_{k-1} - r}\left(X_{aB_k+pB_{k-1}+r} X_{aB_k+pB_{k-1}+r+1} \dots X_{(a+1)B_k+pB_{k-1}+r-1}\right) \\
&= X_{(a+1)B_k} \dots X_{(a+1)B_k+pB_{k-1}+r-1} X_{aB_k+pB_{k-1}+r} \dots X_{aB_k+(b_k-1)B_{k-1}+B_{k-1}-1} \\
&= PQ_0PQ_1 \dots PQ_{b_k-2}PQ_{a+1}.
\end{aligned}$$

Now, similarly,

$$\begin{aligned}
&\sigma'^{B_k - pB_{k-1} - r}(X_n \dots X_{n+B_k-1}) \\
&= P_{s-r} \dots P_{B_{k-1}-2} Q_{q-p} PQ_{q-p+1} \dots PQ_{b_k-2} PQ_{a'+1} PQ_0 \dots PQ_{q-p-1} P_0 \dots P_{s-r-1}
\end{aligned}$$

with the conventions that $Q_{-1} = Q_{a'+1}$ and that any string written $R_i \dots R_{i-1}$ is empty. Now, comparing the two strings, we have $Q_0 = P_{s-r-1}$ and $Q_1 = P_{s-r-1}$, and so $Q_0 = Q_1$. But Q is maximal, infinite and, by lemma 1·14, aperiodic, so we must have $Q_0 = R$ and $Q_1 = L$. Hence we have $s = r$. □

1·4. Adding Machines in Unimodal Maps of the Interval

Throughout this section, f will denote a unimodal map from the closed interval $I \subset \mathbf{R}$ to itself, with critical point c. We say $\Lambda \subset I$ is an *embedded adding machine* if there exists a base sequence $\mathbf{b}$ and a homeomorphism $h : \Lambda \to \mathcal{A}_{\mathbf{b}}$ such that $h \circ f|_\Lambda = \alpha_{\mathbf{b}} \circ h$.

We will show that Λ is an embedded adding machine only if $c \in \Lambda$ and the kneading invariant of f (that is, the itinerary of $f(c)$) is a non-periodic infinite $*$-product. For maps f which have at most one point with any given itinerary, 'only if' becomes 'if and only if.'

Proposition 1·16. *Let $\Lambda \in I$ be an embedded adding machine. Then $c \in \Lambda$.*

Proof. Suppose not. Since Λ is compact, Λ is bounded away from c and h^{-1} is uniformly continuous. Let

$$0 < \epsilon < \inf_{x \in \Lambda} |x - c|.$$

Then there exists n such that, $\forall\, x, y \in \Lambda$, if $h(x)$ and $h(y)$ agree on the first n terms, then $|x - y| < \epsilon$, i.e., x and y lie on the same side of c. Let $x_0 = \sup \Lambda$, X be its itinerary which is an infinite string of R's and L's. Fix $r \geq 0$. Let $x = f^r(x_0)$, $y = f^{r+B_n}(x_0)$. Then $h(y) = \alpha_{\mathbf{b}}^{B_n}(h(x))$ so $h(x)$ and $h(y)$ agree on the first n terms. So x and y lie on the same side of c. So, for $r \geq 0$, $X_r = X_{r+B_n}$. So X is periodic, and hence by Proposition 1·11 the orbit of x_0 converges on a periodic orbit. Since Λ is closed and invariant, Λ contains a periodic orbit, which contradicts Corollary 1·3. □

Theorem 1·17. *Let Λ be an embedded adding machine, let $x_0 = \sup \Lambda = f(c)$, and let X be the itinerary of x_0. Then there exists a base sequence $\mathbf{c}$ for Λ and, for each $i \geq 0$, a finite string of R's and L's, $A^{(i)}$, such that*

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)},$$

$|A^{(i)} + 1| = c_i$ and $A^{(i)}C$ is maximal.

Proof. Let $h : \Lambda \to \mathcal{A}_{\mathbf{b}}$ be a homeomorphism. By Remark 1·5, we may assume that $h(x_0) = \mathbf{0}_{\mathbf{b}}$, so $h(c) = (b_0 - 1, b_1 - 1, \ldots)$. Denote

$$[a_0, a_1, \ldots, a_k] = \{a' \in \mathcal{A}_{\mathbf{b}} : a'_i = a_i \text{ for } 0 \leq i \leq k\}.$$

Let $\Lambda_0 = \Lambda \setminus h^{-1}[b_0 - 1]$, and, for $k > 0$, let

$$\Lambda_k = h^{-1}\left([b_0 - 1, b_1 - 1, \ldots, b_{k-1} - 1] \setminus [b_0 - 1, \ldots, b_k - 1]\right).$$

Now Λ_k is closed in Λ, hence Λ_k is compact; also note that $c \notin \Lambda_k$. Let

$$0 < \epsilon_k < \inf_{x \in \Lambda_k} |x - c|$$

and note that $h^{-1}|_{h(\Lambda_k)}$ is uniformly continuous. So there exist n_k such that if $x, y \in \Lambda_k$ and $h(x)$ and $h(y)$ agree on the first n_k terms, then $|x - y| < \epsilon_k$, i.e., x and y lie the same side of c. We may assume $n_{k+1} > n_k$ for $k \geq 0$. Let $C_k = B_{n_k}$, $c_0 = C_0$, $c_k = C_k/C_{k-1}$ for $k \geq 1$. Now define the $A^{(i)}$ as follows. For $0 \leq r < C_0 - 1$ and $k \geq 0$, we have that

$f^r(x_0) = x$ and $f^{r+kC_0}(x_0)$ both lie in Λ_0, and $h(x)$ and $h(y)$ agree on the first n_0 terms. Thus for $0 \leq r < b_{n_0} - 1$, and $k \geq 0$ we have

$$X_r = X_{r+kC_0}.$$

Let $A_r^{(0)} = X_r$ for $0 \leq r < C_0 - 1 = b_{n_0} - 1$. Then $X = A^{(0)} * X^{(1)}$ where

$$X^{(1)} = \begin{cases} (X_{C_0-1} X_{2C_0-1} \cdots) & \text{if } A^{(0)} \text{ even} \\ \left(\overline{X_{C_0-1}}\, \overline{X_{2C_0-1}} \cdots\right) & \text{if } A^{(0)} \text{ odd.} \end{cases}$$

Now for $0 \leq r < c_1 - 1$, and $k \geq 0$, $f^{c_0-1+rC_0}(x_0) = x$ and $f^{c_0-1+rC_0+kC_1}(x_0) = y$ both lie in Λ_1 and $h(x)$ and $h(y)$ both agree on the first n_1 terms. Thus for $0 \leq r < c_1 - 1$, $k \geq 0$, we have

$$X_{C_0-1+rC_0} = X_{C_0-1+rC_0+kC_1}.$$

Thus let $A_r^{(1)} = X_r^{(1)}$ for $0 \leq r < c_1 - 1$. Then

$$X = A^{(0)} * A^{(1)} * X^{(2)}$$

for suitable $X^{(2)}$. Continuing inductively we see that

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)}.$$

Since X is maximal, each $A^{(i)}C$ is maximal. □

Thus any embedded adding machine in a unimodal map of the interval contains the critical point, and has kneading invariant given by an infinite $*$–product. We will now complete the characterization by showing that under some extra conditions on f, any adding machine can occur on the closure of the orbit of the critical point.

A homterval is an open interval $J \subset I$ such that all points in J have the same itinerary. The condition that we require is that f has no aperiodic homtervals. This will be assured, for example, if f has negative Schwarzian derivative, see for example [CE] section II.5.

Theorem 1·18. *Let f be a map with no aperiodic homtervals and suppose $x_0 = f(c)$ has itinerary*

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)},$$

where each $A^{(i)}C$ is maximal. (Such maps exist by theorem 1·13.) Let $b_i = |A^{(i)}| + 1$ and $\Lambda = \overline{\mathcal{O}(x_0)}$. Then $f|_\Lambda$ is topologically conjugate to an adding machine with base sequence b.

Proof. We will construct a homeomorphism $h : \Lambda \to \mathcal{A}_{\mathbf{b}}$, such that $h \circ f|_\Lambda = \alpha_{\mathbf{b}} \circ h$. Let $\Lambda^\circ = \{f^m(x_0) : m \geq 0\}$, $\mathcal{A}_{\mathbf{b}}^\circ = \{\alpha_{\mathbf{b}}^m(\mathbf{0_b}) : m \geq 0\}$, which are dense subsets of Λ and $\mathcal{A}_{\mathbf{b}}$ respectively. We define $h : \Lambda^\circ \to \mathcal{A}_{\mathbf{b}}$ by $h(f^m(x_0)) = \alpha_{\mathbf{b}}^m(\mathbf{0_b})$ for $m \geq 0$. Now h is continuous by Lemma 1·15 (since, for $k \geq 0$, if $f^m(x_0)$ and $f^n(x_0)$ are suitably close, their itineraries agree on the first B_{k+1} terms, thus by Lemma 1·15 we have $n \equiv m \pmod{B_k}$ and so $d(\alpha_{\mathbf{b}}^m(\mathbf{0_b}), \alpha_{\mathbf{b}}^n(\mathbf{0_b})) \leq 2^{-k}$). h is injective since, if $\alpha_{\mathbf{b}}^m(\mathbf{0_b}) = \alpha_{\mathbf{b}}^n(\mathbf{0_b})$, $m = n$ by Corollary 1·3, finally $h(\Lambda^\circ) = \mathcal{A}_{\mathbf{b}}^\circ$. Let $x \in \Lambda$. Take a sequence m_n such that $f^{m_n}(x_0) \to x$. We will show that $\alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b})$ is convergent in $\mathcal{A}_{\mathbf{b}}$.

Suppose firstly that the itinerary of x is infinite. Then, given $k \geq 0$, there is $N \in \mathbf{N}$ such that for $n \geq N$ the itinerary of $f^{m_n}(x_0)$ agrees with the itinerary of x for the first B_{k+1} terms. Then by Lemma 1·15, if $n \geq N$, $m_n \equiv m_N \pmod{B_k}$, so $\alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b})$ agrees with $\alpha_{\mathbf{b}}^{m_N}(\mathbf{0_b})$ for the first k terms. Thus $\alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b})$ is Cauchy and hence convergent in $\mathcal{A}_{\mathbf{b}}$.

Now suppose the itinerary of x is finite. Then x is a pre-image of the critical point c, and is thus a pre-image of x_0; write $f^j(x) = x_0$. Pick r with $B_{r-1} \leq j < B_r$, and write

$$B_r - j = a_0 + \sum_{i=1}^{r} a_i B_{i-1}.$$

Now $f^{m_n+j}(x_0) \to f^j(x) = x_0$. So if $k > r$, there is N_k such that for $n \geq N_k$ the itinerary of $f^{m_n+j}(x_0)$ agrees with the itinerary of x_0 for the first B_{k+1} terms, i.e., if $n \geq N_k$ we

have $m_n + j \equiv 0 \pmod{B_k}$, or equivalently,

$$\begin{aligned} m_n &\equiv B_k - j \pmod{B_k} \\ &\equiv B_r - j + B_k - B_r \pmod{B_k} \\ &\equiv a_0 + \sum_{i=1}^{r} a_i B_{i-1} + \sum_{i=r+1}^{k} (b_i - 1) B_{i-1}. \end{aligned}$$

Hence for $n \geq N_k$, $h(f^{m_n}(x_0)) = \alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b}) = (a_0, a_1, \ldots, a_r, b_{r+1} - 1, \ldots, b_k - 1, \ldots)$. Thus $\alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b}) \to (a_0, a_1, \ldots, a_r, b_{r+1} - 1, b_{r+2} - 1, \ldots)$ as $n \to \infty$. Hence h extends continuously to Λ, and by construction we have $h \circ f|_\Lambda = \alpha_{\mathbf{b}} \circ h$.

To show h is bijective, let $a \in \mathcal{A}_{\mathbf{b}}$. Denote $A_n = a_0 + \sum_{i=1}^{n} a_i B_{i-1}$. Then $\alpha_{\mathbf{b}}^{A_n}(\mathbf{0_b}) \to a$. Let $X^{(n)}$ be the itinerary of $f^{A_n}(x_0)$. Let $n > N$. Then $A_n = A_N + \sum_{i=N+1}^{n} a_i B_{i-1}$, and $X^{(n)} = \sigma^{kB_N}(X^{(N)})$. Hence $X^{(n)}$ and $X^{(N)}$ agree on terms $0, \ldots, B_N - A_N - 2$. Thus, for any N, the terms $X_0^{(n)}, \ldots, X_{B_N - A_N - 2}^{(n)}$ are eventually constant (†).

Suppose x_n is not convergent. Then there is a k such that $X_k^{(n)}$ is not eventually constant. Suppose this is so. Then by (†), $k \geq B_N - A_N - 1$ for all N. Now $(B_{N+1} - A_{N+1} - 1) - (B_N - A_N - 1) = (b_{N+1} - 1 - a_{N+1}) \geq 0$. So the sequence of integers $(B_N - A_N - 1)$ is non-decreasing and bounded above by k, hence it is eventually constant. So there is r such that if $N \geq r$, $a_N = b_N - 1$. Pick the least such r. In the following we assume $r > 0$, the argument for the case $r = 0$ is similar.

Let $n > r$. Then

$$X^{(n)} = \sigma^{A_n}(X) = \sigma^{A_{r-1} + \sum_{i=r}^{n} a_i B_{i-1}}(X).$$

But

$$\sum_{i=r}^{n} a_i B_{i-1} = \sum_{i=r}^{n} (b_i - 1) B_{i-1} = B_n - B_{r-1}.$$

So

$$X^{(n)} = \sigma^{A_{r-1} + B_n - B_{r-1}}(X)$$

and

$$X_k^{(n)} = X_{A_{r-1} + B_n - B_{r-1} + k}.$$

If $k \geq B_{r-1} - A_{r-1}$, then

$$X_k^{(n)} = X_{k-(B_{r-1}-A_{r-1})}$$

for large enough n. If $0 \leq k < B_{r-1} - A_{r-1} - 1$ then $X_k^{(n)}$ is eventually constant by (†), say $X_k^{(n)} = Y_k$ for large enough n. Thus, for large n,

$$X^{(n)} = \begin{cases} Y_0 Y_1 \ldots Y_{B_{r-1}-A_{r-1}-2} L X_0 X_1 X_2 \ldots \\ Y_0 Y_1 \ldots Y_{B_{r-1}-A_{r-1}-2} R X_0 X_1 X_2 \ldots \end{cases}$$

and so x_n converges to a point whose itinerary is $Y_0 Y_1 \ldots Y_{B_{r-1}-A_{r-1}-2} C$. There is only one such point. Hence a has a unique pre-image in Λ, so h is a bijection. Since Λ is compact and $\mathcal{A}_{\mathbf{b}}$ is Hausdorff, h is a homeomorphism. □

We can use kneading theory to obtain another result:

Corollary 1·19. *Let $f : I \to I$ be unimodal and let $\Lambda \subset I$ be an embedded adding machine with base sequence* $\mathbf{b}$. *Then f has periodic points corresponding to arbitrarily high truncations of* $\mathbf{b}$, *that is, there is a base sequence* $\mathbf{c} \sim \mathbf{b}$ *and periodic points for f of period C_n, $\forall n \in \mathbb{N}$.*

Proof. Let X be the itinerary of the critical point.

$$X = \mathop{*}_{i=0}^{\infty} A^{(i)}$$

with $c_i = |A^{(i)}| + 1$. Then by theorem 1·12 there is $x_n \in I$ with itinerary

$$A^{(0)} * A^{(1)} * \cdots * A^{(n)} * (R^{\infty})$$

which is a periodic itinerary of period C_n. A standard argument in calculus of itineraries (see for example [CE]) shows there is a point $y_n \in I$ of period C_n. □

Thus we see that a set $\Lambda \subset I$ under a unimodal map without homtervals is topologically conjugate to an adding machine if and only if $c \in \Lambda$ and $f(c)$ has a non-periodic infinite $*$-product. If we consider a one-parameter family of C^1 unimodal maps f_μ, then an adding machine appears at a certain value of the parameter μ, when the kneading invariant

$K(\mu)$ is given by the appropriate itinerary. It is natural to ask what happens to the adding machine as the parameter moves away from this value. Specifically, suppose the kneading invariant $K(\mu)$ satisfies

$$K(\mu) \succ X = \mathop{*}_{i=0}^{\infty} A^{(i)}$$

where each $A^{(i)}$ is a finite string of R's and L's, $A^{(i)}C$ is maximal and X is non-periodic. Then there is a point $x \in I$ which has itinerary X. What are the dynamics of f on the closure of the orbit of x? The following theorem shows that there is a semi-conjugacy between this set and an adding machine, where only a countable number of points have more than one preimage.

Theorem 1·20. *Suppose $f : I \to I$ is unimodal with critical point c, contains no homtervals and that the itinerary*

$$I(f(c)) \succ X = \mathop{*}_{i=0}^{\infty} A^{(i)}$$

for some $A^{(i)}$, $i \in \mathbb{N}$ with $A^{(i)}C$ maximal for each i. Then there is a point $x_0 \in I$ with itinerary X. Denoting $b_i = |A^{(i)}| + 1$ and $\Lambda = \overline{\mathcal{O}(x_0)}$ we have the following:

(i) $f|_\Lambda$ is semi-conjugate to an adding machine of base sequence $\mathbf{b}$, i.e., there is a continuous surjection $h : \Lambda \to \mathcal{A}_\mathbf{b}$ with $h(x_0) = \mathbf{0_b}$ such that the following diagram commutes:

$$\begin{array}{ccc} \Lambda & \xrightarrow{f} & \Lambda \\ \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle h} \\ \mathcal{A}_\mathbf{b} & \xrightarrow{\alpha_\mathbf{b}} & \mathcal{A}_\mathbf{b} \end{array}$$

(ii) Given $a \in \mathcal{A}_\mathbf{b}$ then $h^{-1}\{a\}$ is a single point if a is not a pre-image of $\mathbf{0_b}$, and $h^{-1}\{a\}$ is a pair of distinct points if $\alpha_\mathbf{b}^m(a) = \mathbf{0_b}$ for some $m \geq 0$.

Proof. This proof is similar to the proof of Theorem 1·18. We construct a map $h : \Lambda \to \mathcal{A}_\mathbf{b}$. Let $\Lambda^\circ = \{f^m(x_0) : m \geq 0\}$ as before. We define, as in the proof of Theorem 1·18, a homeomorphism $h : \Lambda^\circ \to \mathcal{A}_\mathbf{b}^\circ$ by $h(f^m(x_0)) = \alpha_\mathbf{b}^m(\mathbf{0_b})$, for $m \geq 0$.

Let $x \in \Lambda$, and choose $x_n \in \Lambda^\circ$, say $x_n = f^{m_n}(x_0)$, with $x_n \to x$. The itinerary of x is infinite (since otherwise $f^k(x) = c$ for some k, then $f^{k+1}(x) = f(c) > x_0$. Then the

itinerary of $f^{k+1}(x)$ is strictly larger than X, contradicting the maximality of X). Then, given $k \geq 0$, there is N_k such that if $n \geq N_k$, the itinerary of x_n, $X^{(n)}$ say, agrees with the itinerary of x for the first B_{k+1} terms. Thus for $n \geq N_k$, we have $m_n \equiv m_{N_k} \pmod{B_k}$. Thus $\alpha_{\mathbf{b}}^{m_n}(\mathbf{0_b})$ is convergent and h extends continuously to Λ. Now we will show h is surjective. Let $a \in \mathcal{A}_{\mathbf{b}}$. As in Theorem 1·18, if a is not a preimage of $\mathbf{0_b}$, there is a unique point $x \in \Lambda$ such that $h(x) = a$. Suppose a is a preimage of $\mathbf{0_b}$, so $a_n = b_n - 1$ for all $n > r$, $a_r < b_r - 1$. We will show $h^{-1}\{a\}$ is given by two distinct points, x^L and x^R.

Suppose $n > r$. Let

$$\begin{aligned} A_n &= a_0 + \sum_{i=1}^{n} a_i B_{i-1} \\ &= A_r + \sum_{i=r+1}^{n} a_i B_{i-1} \\ &= A_r + \sum_{i=r+1}^{n} (b_i - 1) B_{i-1} \\ &= A_r + \sum_{i=r+1}^{n} (B_i - B_{i-1}) \\ &= B_n - (B_r - A_r). \end{aligned}$$

Let

$$A'_n = B_n + A_n = 2B_n - (B_r - A_r).$$

Let

$$m_n^L = \begin{cases} A'_n & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ even,} \\ A_n & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ odd} \end{cases}$$

and

$$m_n^R = \begin{cases} A_n & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ even,} \\ A'_n & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ odd.} \end{cases}$$

By (†) in Theorem 1·18, the terms $0, \ldots, B_r - A_r - 2$ of the itineraries of $f^{m_n^L}(x_0)$ and $f^{m_n^R}(x_0)$ are eventually constant, equal to $Y_0 Y_1 \ldots Y_{B_r - A_r - 2}$. Now, $X = A^{(0)} * A^{(1)} * \cdots * A^{(n-1)} * (RL \ldots)$. So, the $(B_r - A_r - 1)^{\text{th}}$ term of the itinerary of $f^{A_n}(x_0)$ is

$$X_{B_r - A_r - 1 + A_n} = X_{B_n - 1} = \begin{cases} R & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ even,} \\ L & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ odd.} \end{cases}$$

Similarly, the $(B_r - A_r - 1)^{\text{th}}$ term of the itinerary of $f^{A'_n}(x_0)$ is

$$X_{B_r - A_r - 1 + A'_n} = X_{2B_n - 1} = \begin{cases} L & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ even,} \\ R & \text{if } A^{(0)} * \cdots * A^{(n-1)} \text{ odd.} \end{cases}$$

Finally, if $k > B_r - A_r - 1$ and n is suitably large, the kth term of the itineraries of both $f^{A_n}(x_0)$ and $f^{A'_n}(x_0)$ is given by $X_{k-(B_r - A_r)}$ as in Theorem 1·18. Let x^L be the point whose itinerary is $Y_0 \ldots Y_{B_r - A_r - 2} L X$ and x^R be the point whose itinerary is $Y_0 \ldots Y_{B_r - A_r - 2} R X$. Then we have shown that $f^{m_n^L}(x_0) \to x^L$ and $f^{m_n^R}(x_0) \to x^R$. Hence $x^L, x^R \in \Lambda$ and $h(x^L) = h(x^R) = a$. These are the only candidates for preimages of a under h, so the theorem is proved. □

We will call a set Λ with the dynamics described above a 'blown-up adding machine.' We remark that if a unimodal map $f : I \to I$ contains a blown-up adding machine, then by kneading theory it contains a periodic point whose period is not a power of two. (To see this, note that the kneading invariant is strictly bigger than $R * R * \cdots$; any such kneading invariant is bigger than some periodic itinerary whose period is not a power of two. For details see [CE].) It follows (see [CE] II.8.14) that f has positive topological entropy. If f has no blown-up adding machines then the kneading invariant is at most $R * R * \cdots$ and f is known to have zero topological entropy. Thus a unimodal map f has positive topological entropy if and only if it contains a blown-up adding machine. (Note that, as we shall see in the next section, the dynamics of blown-up adding machines themselves carry zero entropy, but force positive topological entropy in the action of f on I. We will conjecture that a similar result holds for maps of the 2-disc.)

1·5. The two–sided shift and maps of the disc

In this section we will examine the occurrence of adding machines and an analogue to the blown–up adding machine in the two-sided shift on two symbols, that is, let

$$\Sigma = \{(a_n)_{n\in\mathbf{Z}} : a_n = 0 \text{ or } 1 \text{ for all } n\}$$

and let $s : \Sigma \to \Sigma$ by $(s(a))_n = a_{n+1}$ be the shift map. We will show that $s : \Sigma \to \Sigma$ contains no adding machines but every 'two–sided blown up adding machine.' The two–sided shift map is an interesting object of study due to the following theorem of Smale and Katok:

Theorem 1·21. *(See [S],[K].) Let $f : D \to D$ be a $C^{1+\epsilon}$ diffeomorphism of the 2-disc D. Then f has positive topological entropy if and only if there is some subset $\Lambda \subset D$ on which some iterate $f^n|_\Lambda : \Lambda \to \Lambda$ is topologically conjugate to $s : \Sigma \to \Sigma$.*

Thus we will see that every $C^{1+\epsilon}$ diffeomorphism of the disc D has invariant subsets $\Lambda_{\mathbf{b}}$ on which some iterate of f acts as a 2-sided blown-up adding machine with base sequence $\mathbf{b}$.

First we will show there are no embedded adding machines in the shift map.

Proposition 1·22. *There is no embedded adding machine $\Lambda \subset \Sigma$.*

Proof. (See proposition 1·16.) Suppose $h : \Lambda \to \mathcal{A}_{\mathbf{b}}$ is a homeomorphism with $h \circ s = \alpha_{\mathbf{b}} \circ h$. Then since $\mathcal{A}_{\mathbf{b}}$ is compact, h^{-1} is uniformly continuous and so there exists $n \in \mathbf{N}$ such that given any $x, y \in \mathcal{A}_{\mathbf{b}}$ with $x_i = y_i$ for $0 \le i \le n$ it follows that $(h^{-1}(x))_0 = (h^{-1}(y))_0$. In particular, fix any $x \in \mathcal{A}_{\mathbf{b}}$, let $y = \alpha_{\mathbf{b}}^{B_n}(x)$. Then $\alpha_{\mathbf{b}}^r(x)_i = \alpha_{\mathbf{b}}^r(y)_i$ for $0 \le i \le n$ and all $r \in \mathbf{Z}$, so

$$h^{-1}(\alpha_{\mathbf{b}}^r(x))_0 = h^{-1}(\alpha_{\mathbf{b}}^r(y))_0,$$

that is,

$$\begin{aligned} s^r h^{-1}(x)_0 &= s^r h^{-1}(y)_0 \\ &= s^{r+B_n} h^{-1}(x)_0, \end{aligned}$$

hence

$$h^{-1}(x)_r = h^{-1}(x)_{r+B_n}$$

for all r, so $h(x)$ is periodic under s. But then x is a periodic point of $\alpha_{\mathbf{b}} : \mathcal{A}_{\mathbf{b}} \to \mathcal{A}_{\mathbf{b}}$, contradicting Corollary 1·3. □

We will now construct, for each base sequence $\mathbf{b}$, a sequence $x^{(\mathbf{b})} \in \Sigma$ such that the action of s on the closure of the orbit of x under s, $\overline{\mathcal{O}(x^{(\mathbf{b})})}$, has the following properties:

(i) there is a semi-conjugacy $h : \overline{\mathcal{O}(x^{(\mathbf{b})})} \to \mathcal{A}_{\mathbf{b}}$, with $h(x) = 0_{\mathbf{b}}$,

(ii) given any $m \in \mathbf{Z}$, the point $\alpha_{\mathbf{b}}^m(0_{\mathbf{b}})$ has precisely two preimages under h,

(iii) given any $a \in \mathcal{A}_{\mathbf{b}}$, $a \neq \alpha_{\mathbf{b}}^m(0_{\mathbf{b}})$ for any $m \in \mathbf{Z}$, then a has precisely one preimage under h.

We call the subset $\overline{\mathcal{O}(x^{(\mathbf{b})})}$ a two-sided blown-up adding machine with base sequence $\mathbf{b}$.

We will again use the idea of infinite $*$-products. We define $*$-products on finite strings of 0's and 1's by identifying 0 with L and 1 with R in section 3. We will also use the following notation:

$$(A_n)_{n\in\mathbf{Z}} = (\ldots A_{-2}A_{-1} \cdot A_0A_1A_2 \ldots).$$

If A is a finite or semi-infinite string of 0's and 1's,

$$A = (A_0A_1 \ldots A_n) \text{ or } (A_0A_1 \ldots),$$

then

$$\overleftarrow{A} = (A_nA_{n-1} \ldots A_0) \text{ or } (\ldots A_2A_1A_0)$$

respectively.

Finally, for $k \geq 2$ we introduce some standard strings $P^{(k)}$ of length $k-1$:

(i) if k is odd, let $P^{(k)} = 101^{k-3}$;

(ii) if $k = 2^n$ for some n then let $P^{(k)} = \underbrace{1 * 1 * \cdots * 1}_{n\text{-fold}}$;

(iii) if $k = 2^n i$ for $i \geq 3$ odd, then let $P^{(k)} = P^{(2^n)} * P^{(i)}$.

(These are the min-max strings used in the proof of Šarkovskii's theorem in [CE].)

Theorem 1·23. *Let* $\mathbf{b}$ *be a base sequence, and define the semi-infinite string of 0's and 1's*

$$A^{(\mathbf{b})} = \mathop{*}_{i=0}^{\infty} P^{(b_i)}.$$

Now define

$$\widehat{A^{(\mathbf{b})}} = \overline{\mathop{*}_{i=0}^{\infty} \overline{P^{(b_i)}}}.$$

Let $x^{(\mathbf{b})} \in \Sigma$ *be given by* $x^{(\mathbf{b})} = \widehat{A^{(\mathbf{b})}}1 \cdot A^{(\mathbf{b})}$. *Then* $\overline{\mathcal{O}(x^{(\mathbf{b})})}$ *is a two-sided blown-up adding machine.*

Proof. First notice that

$$\overline{\overline{P^{(b_0)}} * \overline{P^{(b_1)}} * \cdots * \overline{P^{(b_n)}}} = P^{(b_0)} * P^{(b_1)} * \cdots * P^{(b_n)}.$$

As usual, we define $h(s^n(x^{(\mathbf{b})})) = \alpha^n(0_{\mathbf{b}})$ for $n \in \mathbf{Z}$. By Lemma 1·15 and the above remark, h is continuous on $\{s^n(x^{(\mathbf{b})}) : n \in \mathbf{Z}\}$. We extend h continuously to all of $\overline{\mathcal{O}(x^{(\mathbf{b})})}$ as in the proof of theorem 1·18. By (†) in the proof of theorem 1·18, and the above remark, we see that

$$\left(s^{kB_n}(x^{(\mathbf{b})})\right)_i = (x^{(\mathbf{b})})_i$$

for $-B_n \leq i < B_n - 1, i \neq -1$, for all $k \in \mathbf{Z}$. Following the argument of theorem 1·18 we achieve the desired result. □

Since we are dealing with two–sided sequences, the 'ambiguity' of a blown–up adding machine (that is the two–to–oneness of the map h on the $\mathbf{Z}$-orbit of $x^{(\mathbf{b})}$) persists under both forward and backward iterations of s. As we remarked at the beginning of this section we now have the following:

Corollary 1·24. *Let* $f : D \to D$ *be a* $\mathcal{C}^{1+\epsilon}$ *diffeomorphism with positive topological entropy* $h(f)$. *Then for every base sequence* $\mathbf{b}$ *there is a set* $\Lambda_{\mathbf{b}} \subset D$ *such that some iterate of* f *acts as a two–sided blown–up adding machine with base sequence* $\mathbf{b}$ *on* $\Lambda_{\mathbf{b}}$.

Proof. Theorem 1·21 and theorem 1·23. □

Given the comments we made about topological entropy and blown-up adding machines for unimodal maps, and the description of the boundary of positive topological entropy for the Hénon family conjectured in [GT], it seems reasonable to conjecture that a $C^{1+\epsilon}$ diffeomorphism of the disc D has positive topological entropy if and only if it has two-sided blown-up adding machines. Firstly we will show that two-sided blown-up adding machines themselves carry zero topological entropy.

Proposition 1·25. *Let $\Lambda \subset \Sigma$ be a two-sided blown-up adding machine. Then $s|_\Lambda$ has zero topological entropy.*

Proof. By [W] theorem 7.13 (i),

$$h(s|_\Lambda) = \lim_{n\to\infty} \frac{1}{n} \log \theta_n$$

where θ_n is the number of n-tuples $[a_0 a_1 \ldots a_{n-1}]$ of length n such that

$$\left\{ m \in \mathbb{Z} : (s^m(x^{(\mathbf{b})}))_i = a_i \text{ for } 0 \le i \le n-1 \right\}$$

is non-empty. Now, by construction of Λ we see that $\theta_{B_n} \le 2B_n$, so

$$h(s|_\Lambda) \le \lim_{n\to\infty} \frac{1}{B_n} \log 2B_n = 0.$$

□

We can prove similarly that the blown-up adding machines in section 4 and adding machines themselves carry zero entropy (or we can deduce this by constructing a semi-conjugacy from the two-sided blown-up adding machine).

Conjecture. *Let $f : D \to D$ be a $C^{1+\epsilon}$ diffeomorphism. Then f has positive topological entropy if and only if there is a set $\Lambda \subset D$ on which f acts as a two-sided blown-up adding machine.*

To prove the conjecture it would be necessary to consider the dynamics around the blown-up adding machine, for example, are there periodic points corresponding to truncations of the base sequence as in the unimodal case (Corollary 1·19)? Theorems about

co-existence of periodic points and about periodic points and topological entropy for maps in two dimensions generally require the theory of braid types, see for example [BF] and [LM]; this theory may be regarded as the two-dimensional analogue of kneading theory. However, there are technical difficulties about extending the definition of braid types from periodic orbits to adding machines which cause problems for this approach.

References

[BF] Boyland, P.L. and Franks, J. [1989] Notes on Dynamics of Surface Homeomorphisms. Preprint, University of Warwick.

[BS] Buescu, J. and Stewart, I. [1993]. Liapunov Stability and Adding Machines. Preprint, University of Warwick.

[CE] Collet, P. and Eckmann, J-P. [1980]. *Iterated Maps of the Interval as Dynamical Systems*. Progress on Physics, Vol. I. Birkhäuser-Boston.

[E-HM] El Hamouly, H. and Mira, C. [1981]. Lien Entre les Propriéties d'un endomorphisme de dimension un et celles d'un diffeomorphisme de dimension deux. *C.R. Acad. Sc. Paris*. **293**, 525-528.

[GT] Gambaudo, J-M. and Tresser, C. [1991]. How Horseshoes are Created. *Instabilities and Nonequilibrium Structures III*. 13-25, Kluwer Ac. Pub., Netherlands.

[K] Katok, A.B. [1980]. Lyapunov Exponents, Entropy and Periodic Points for Diffeomorphisms. *Publ. Math. IHES.*, **51**. 137-173.

[LM] Llibre, J. and MacKay, R.S. [1990]. A Classification of Braid Types for Diffeomorphisms of Surfaces of Genus Zero with Topological Entropy Zero. *J. London Math. Soc. (2)*, **42**. 562-576.

[MT] Milnor, J. and Thurston, W. [1977]. On Iterated Maps of the Interval I. Preprint, Princeton University.

[S] Smale, S. [1965]. Diffeomorphisms with Many Periodic Points. In *Differential and Combinatorial Topology*, S.S. Cains (ed), pp 63-80. Princeton University Press.

[W] Walters, P. [1982]. *An Introduction to Ergodic Theory*, Springer-Verlag, New York.

2. Consequences of contractible geodesics on surfaces

2·1. Introduction

In this chapter, we prove existence of an interesting set of geodesics on all Riemannian surfaces with boundary and with negative Euler characteristic, and draw conclusions about the dynamics of the geodesic flow on the unit tangent bundle generated by the Riemannian metric.

Our result is analogous to one of Morse [M1] for closed surfaces, which we now recall. Let M be a compact orientable C^2 surface of genus $g \geq 2$ without boundary, and with a Riemannian metric ρ. Then M has universal cover given by the Poincaré disc Δ and a fundamental group G of Möbius transformations on Δ. There is a standard metric h on Δ of constant negative curvature given by

$$ds^2 = \frac{dr^2 + r^2 d\theta^2}{(1 - r^2)^2},$$

which projects to a second metric on M; also the metric ρ lifts to a metric $\tilde{\rho}$ on Δ. For any metric σ denote by d_σ the distance function generated by σ. Geodesics of the metric h ('h-geodesics') are arcs of circles perpendicular to the boundary $\partial\Delta = S^1$. Two curves $\alpha, \beta \subset \Delta$ are said to be *of the same type* if there exists a constant R such that for each $x \in \alpha$ there exists $y \in \beta$ with $d_h(x, y) \leq R$ and for each $y \in \beta$ there exists $x \in \alpha$ with $d_h(y, x) \leq R$. A *class-A* $\tilde{\rho}$-geodesic is a $\tilde{\rho}$-geodesic which minimises the distance d_ρ between any two of its points. Morse's theorem then states:

Theorem 2·1. *([M1]) Given any h-geodesic l in Δ there is a class-A $\tilde{\rho}$-geodesic α of the same type as l. Conversely, given any class-A $\tilde{\rho}$-geodesic there is a unique h-geodesic of the same type.*

Hedlund ([H]) proved an analogous result for surfaces of genus 1. Klingenberg [K] has also developed further theory from [M1], with a somewhat different viewpoint. Theorem

$5 \cdot 3 \cdot 6$ of [K] is analogous to the main theorem of [M1] for manifolds of arbitrary dimension, but has weaker conclusion than our main result. In a later paper [M2], Morse showed that with certain conditions on the metric ρ ('uniform instability'), there is only one class A $\tilde{\rho}$-geodesic of any given type.

We will prove a similar result for C^2 Riemannian surfaces M with metric ρ, of genus $g \geq 0$ which have $s \geq 1$ discs removed, satisfying $2g + s \geq 3$ and having each boundary component a ρ-geodesic.

The main lemma used in proving Morse's theorem states that there is a uniform constant R such that, given an h-geodesic segment L with endpoints p, q and a $\tilde{\rho}$-shortest curve α with the same endpoints, then α does not wander an h-distance more than R from L. The proof we give of the analogous lemma for the case of a surface with geodesic boundary is very simple and follows from a careful choice of the covering space. Once this lemma is established we proceed as in [M1] to prove the main theorem.

The theorem has the following consequences. For each metric σ denote by $S^\sigma M$ the unit tangent bundle over M. The metric σ generates a C^1 flow $\phi^\sigma = \{\phi_t^\sigma : t \in \mathbf{R}\}$ on $S^\sigma M$. Denote by $\mathcal{A} \subset S^\rho M$ the subset of the unit tangent bundle determined by projections of class-A $\tilde{\rho}$-geodesics, and by $S^h M$ the unit tangent bundle of M with respect to the projection of the metric h onto M. Then the theorem we prove determines a semi-equivalence between $\mathcal{A}$ and $S^h M$, that is, a continuous surjection $\Theta : \mathcal{A} \to S^h M$ and continuous orientation-preserving reparametrization $\tau : \mathcal{A} \times \mathbf{R} \to \mathbf{R}$ with $\phi_t^h \circ \Theta = \Theta \circ \phi_\tau^\rho|_{\mathcal{A}}$. The dynamics of ϕ^h are known to be very complicated, for example there is positive topological entropy, many periodic orbits and homoclinic orbits and a Markov partition, see for example [S]. The semi-equivalence shows that the dynamics of ϕ^ρ on $\mathcal{A}$ are at least as complicated. For metrics with uniform instability ([M2]), the semi-equivalence becomes a full equivalence (that is, θ is a homeomorphism).

We indicate two interesting examples to which the theorem applies. Firstly we may apply the theorem to a torus with a simple contractible geodesic, γ, by deleting the open disc bounded by γ. This is similar to the 'torus with a big bump' results of [Ba] and [P],

Figure 1.

and was the motivating example for this work. Secondly, let M be a two-sphere with three contractible geodesics bounding disjoint discs D_1, D_2, D_3. Then we may apply the theorem to $M \setminus (D_1 \cup D_2 \cup D_3)$. See figure 1.

We proceed as follows. Firstly we will describe the elementary theory of geodesics and the construction of the covering space of M. We then prove the main theorem and discuss the dynamical consequences. We conclude this chapter with an appendix in which we adapt our proof to give a new proof of Morse's theorem; that is, we adapt the proof to cover the case $s = 0$.

2·2. Elementary theory of geodesics

We give a brief summary of the theory of geodesics we shall require. For more details see, for example, [Ba].

Let Σ be a complete compact manifold with $\mathcal{C}^2$ Riemannian metric σ. Let $\gamma : [a,b] \to \Sigma$ be a continuous curve, and for $a \leq p \leq q \leq b$ define $L_\sigma(\gamma; p, q)$ to be the length of the curve γ from $\gamma(p)$ to $\gamma(q)$ with respect to the metric σ. A *minimal geodesic segment* is a curve $\gamma : [a,b] \to \Sigma$ such that, for each $a \leq p \leq q \leq b$ there is no curve $\alpha : [0,1] \to \Sigma$ with $\alpha(0) = \gamma(p)$, $\alpha(1) = \gamma(q)$ and $L_\sigma(\alpha; 0, 1) < L_\sigma(\gamma; p, q)$. A *geodesic* is a curve $\gamma : I \to \Sigma$, where $I \subset \mathbf{R}$ is a closed interval (we allow $I = \mathbf{R}$) such that, given any $s \in I$ there is $\epsilon > 0$ such that $\gamma|_{[s-\epsilon, s+\epsilon] \cap I}$ is a minimal geodesic segment. A *minimal geodesic* is a curve $\gamma : \mathbf{R} \to \Sigma$ such that $\gamma|_{[a,b]}$ is a minimal geodesic segment for any $a \leq b \in \mathbf{R}$.

For the surfaces we are considering, we have the three following classical results, see for example [Bi], [C], [K].

Lemma 2·2. *Given $a, b \in \Sigma$, there is a shortest path $\alpha : [p,q] \to \Sigma$ with $\alpha(p) = a$ and $\alpha(q) = b$.*

Lemma 2·3. *In each non-trivial free homotopy class, there is a closed geodesic of minimum length.*

Lemma 2·4. *Let $\alpha, \beta : [a,b] \to \Sigma$ be geodesics and suppose there exists $\epsilon > 0$ such that $\alpha|_{[a,a+\epsilon]} = \beta|_{[a,a+\epsilon]}$. Then $\alpha = \beta$.*

It is also clear that

Lemma 2·5. *A geodesic $\alpha : [a,b] \to \Sigma$ is differentiable on $[a,b]$.*

In particular, a geodesic 'contains no corners.'

We will associate geodesics with points in the unit tangent bundle $S^\sigma\Sigma$ of Σ, equipped with its canonical Riemannian metric. The metric σ generates a flow $\phi^\sigma = \{\phi_t^\sigma : t \in \mathbf{R}\}$ on $S^\sigma\Sigma$, that is, a one-parameter group of diffeomorphisms of $S^\sigma\Sigma$. Each unit tangent vector determines a geodesic by lemma 2·4; the flow is determined by the motion along this geodesic with unit speed. For σ a $\mathcal{C}^2$ metric, ϕ^σ is a $\mathcal{C}^1$ flow.

We define a topology on the set of geodesics as follows. Let $(\gamma_n)_{n\in\mathbb{N}}$ be a sequence of geodesics. We say $\gamma_n \to \gamma$ if, for each $x \in S^\sigma\Sigma$ determined by the geodesic γ, there are $x_n \in S^\sigma\Sigma$ determined by the γ_n with $x_n \to x$ in the canonical metric induced by σ.

2·3. Construction of the universal cover of M

From now on, M will be a complete compact orientable surface with C^2 Riemannian metric ρ of genus $g \geq 0$ with $s \geq 1$ open discs removed. We assume $2g + s \geq 3$ and that each boundary component is a ρ-geodesic. In this section we will describe a standard construction of the covering space and fundamental group of M. We start by constructing the covering space for a surface of genus g with $s = 1$ if $g > 0$ and $s = 3$ if $g = 0$. We then show how to adapt this cover to obtain a covering space for the surface M.

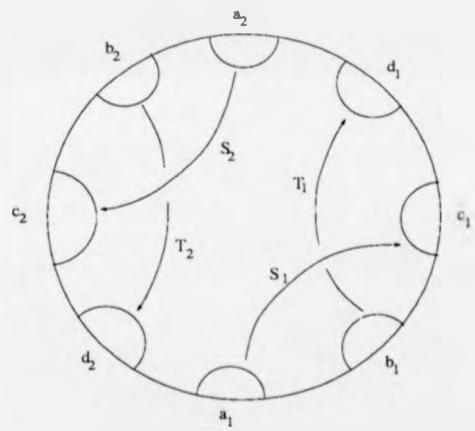

Figure 2 (for the case g=2).

Let $\Delta \subset \mathbb{C}$ be the unit disc with its standard hyperbolic metric h. The h-geodesics are arcs of circles which are perpendicular to the unit circle $\partial\Delta$ [Be]. Let M have genus $g \geq 0$ and $s \geq 1$ boundary components. Suppose firstly that $g > 0$. Choose $4g$ disjoint h-geodesics in Δ, labelled $a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, \ldots, a_g, b_g, c_g, d_g$ in anticlockwise order.

Now choose analytic isometries S_i and T_i of Δ for $i = 1, \ldots, g$, with $S_i(a_i) = c_i$ and $T_i(b_i) = d_i$, see figure 2. (Analytic isometries f of Δ are of the form $f(z) = \frac{az+c}{\bar{c}z+\bar{a}}$ with $|a|^2 -$

Figure 3.

$|c|^2 = 1$.) Let G be the group generated by the transformations $S_1, S_2, \ldots, S_g, T_1, \ldots, T_g$. Then G is free and Δ/G is homeomorphic to a surface of genus g with an 'infinite funnel,' see figure 3.

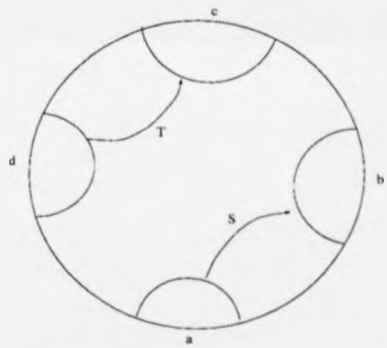

Figure 4.

Suppose alternatively that $g = 0$. We will describe the construction of a surface of genus zero with three infinite funnels. Choose four disjoint h-geodesics in Δ, labelled a, b, c, d in anticlockwise order. Choose analytic isometries $S : a \to b$ and $T : d \to c$ as in figure 4. Let G be the group generated by S and T. Then G is free and Δ/G is homeomorphic to a surface of genus zero with three infinite funnels.

We now modify the appropriate model, if necessary, to create a surface of genus g with s infinite funnels, as follows.

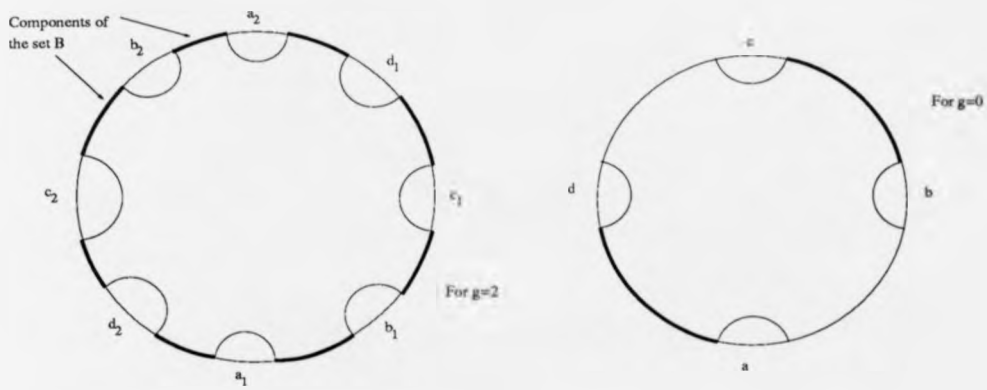

Figure 5.

To increase the number of infinite funnels by one, we will introduce two new geodesics to our model and add to the set of generators of the group an analytic isometry mapping one to the other. This will be done in such a way that there will be a boundary component of the surface containing at least one endpoint of the projection of each geodesic. We do this as follows. For the case $g > 0$, let $B \subset S^1$ be the closure of the union of the segments of S^1 lying between adjacent geodesics; that is, the segment from the right-hand endpoint of a_1 to the left-hand endpoint of b_1 containing no other endpoints of geodesics, union the segment from the right-hand endpoint of b_1 to the left-hand endpoint of c_1, and so on up to the segment from the right-hand endpoint of d_g to the left-hand endpoint of a_1. For the case $g = 0$ let $B \subset S^1$ be the union of the closed segment from the right-hand endpoint of d to the left-hand endpoint of a and the closed segment from the right-hand endpoint of b to the left-hand endpoint of c, see figure 5.

In both cases, B projects under the natural projection $\pi : \Delta \to \Delta/G$ to a single boundary component, and all h-geodesics in Δ which are sides of fundamental regions have an endpoint in B.

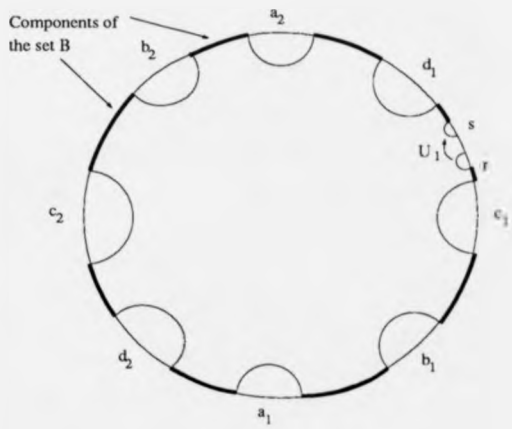

Figure 6.

Now, to increase the number of funnels by one we choose a connected component B' of B and disjoint h-geodesics r, s, both having both endpoints in B'. Let U_1 be an analytic isometry with $U_1(r) = s$. Now let G be the group generated by $S_1, T_1, \ldots, S_g, T_g, U_1$ (G is again free), and remove from the set B the segment of S^1 containing the four endpoints of r and s, see figure 6. Then Δ/G is a surface of genus g with two infinite funnels, and again the set B is the lift of a single boundary component containing at least one endpoint each of the sides of the fundamental domain.

We may now repeat this procedure to obtain the cover of a surface M_∞ of genus g with s infinite funnels. The cover has the properties that the sides of each fundamental domain are pairwise non-intersecting geodesics, and that they project to curves on the surface all of which have at least one endpoint on some fixed boundary component.

Finally, we restrict attention to a compact subset $M' \subset M_\infty$ defined as follows. Given two adjacent sides of any fundamental domain F, r and s say, there is a unique h-geodesic l perpendicular to both r and s. Let $\bar{D}$ be the collection (over all fundamental domains F and all pairs of adjacent sides r, s) of open half planes bounded by such geodesics l which intersect no sides of the fundamental domain F other than r and s. (See figure 7.) Let

$$U = \Delta \setminus \bar{D}.$$

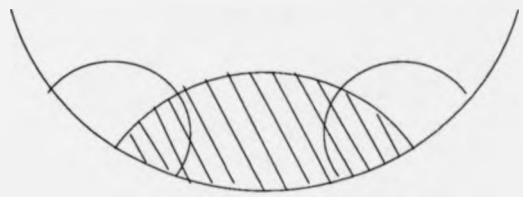

Figure 7.

Then U/G is a surface M' of genus g with s discs removed, that is, it is homeomorphic to the surface M. Let $\pi : U \to M'$ be the natural projection.

The following propostion contains the first new idea in this section and is instrumental in proving the important proposition 2·9.

Proposition 2·6. *There is a diffeomorphism $h : M' \to M$ such that $h \circ \pi$ takes sides of fundamental regions described above into minimal ρ-geodesic segments on M.*

Proof. Denote by $\alpha_1 \ldots \alpha_{2g+s-1} \subset M'$ the projections under π of the sides of the fundamental domains in U. Since the α_i are pairwise disjoint, it is enough to show that the minimal ρ-geodesics β_i in the same homotopy class as the α_i are also pairwise disjoint.

Suppose $x \in \beta_i \cap \beta_j$ for some $i \neq j$. Let $\tilde{\beta}_i$, $\tilde{\beta}_j$ be segments of β_i and β_j respectively with one endpoint on a common boundary component and the other endpoint at x. Let $\beta_i' = \beta_i \setminus \tilde{\beta}_i$, $\beta_j' = \beta_i \setminus \tilde{\beta}_j$. It then follows that $\tilde{\beta}_i \cup \beta_j'$ is a minimal geodesic in the same homotopy class as α_j, with a corner at x, contradicting Lemma 2·5. □

Definition 2·1. *A $\tilde{\rho}$-geodesic or -geodesic segment is of class A if it minimises the distance between any two of its points in U.*

Lemma 2·7. *Let $\alpha \neq \beta$ be class A $\tilde{\rho}$-geodesic segments. Then the interiors of α and β have at most one point of intersection. In particular, two distinct class A $\tilde{\rho}$-geodesics intersect at most once.*

Proof. (See [M1] theorem 3.) Suppose $x \neq y$ are points of intersection of the interiors of

α and β. Let the endpoints of α be p and q, (labelled so that the points p, x, y, q occur in order along α). Let γ_1 be the segment of α from p to x, γ_2 the segment of β from x to y and γ_3 the segment of α from y to q. Let $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3$. Now, since α and β are class A, we have $L_{\tilde{\rho}}(\alpha; x, y) = L_{\tilde{\rho}}(\beta; x, y)$ and so $L_{\tilde{\rho}}(\alpha; p, q) = L_{\tilde{\rho}}(\gamma; p, q)$. Hence γ is also a class A $\tilde{\rho}$-geodesic. But since $\alpha \neq \beta$, γ has corners at x and y, contradicting lemma 2·5. □

The following lemma will be used to establish the uniform bound of the distance between h-geodesic segments and class A $\tilde{\rho}$-geodesic segments with the same endpoints.

Lemma 2·8. *Let $L, \alpha : [0,1] \to U$ have the same endpoints, where L is an h-geodesic segment and α is a class A $\tilde{\rho}$-geodesic segment. Let $F_1, \ldots, F_m$ be the fundamental domains through which L passes, in order. Then the fundamental domains through which α passes are also, in order, $F_1, \ldots, F_m$.*

Proof. First note that neither L nor α crosses the same side of a fundamental domain more than once. This is true for L as distinct h-geodesics intersect at most once. It is true for α by lemma 2·7. Since bounding geodesics of the fundamental domains are disjoint, it is clear that if α crosses a side a of a fundamental domain that L does not cross, it enters the half plane bounded by A not containing points of L and must thus cross a again. Hence the lemma is proved. □

2·4. Properties of Class A $\widetilde{\rho}$-geodesics

We will show in this section that there is a uniform constant bounding the distance a class A $\widetilde{\rho}$-geodesic segment can wander from the h-geodesic joining its endpoints. This is analogous to Lemma 8 of Morse [M1] and Lemma 7.1 of Hedlund [H]. We will use this fact in the same way as Morse and Hedlund to prove the main theorem.

We will also establish some other properties of class A $\widetilde{\rho}$-geodesic segments similar to those in [M1] and [H]. Finally we show that any class A geodesic that is not a boundary component of $\widetilde{D}$ is disjoint from $\partial\widetilde{D}$.

Since we have removed finitely many open discs in the construction of our surface M, M is compact. Thus we may define

$$d = \sup_{x,y\in F} d_h(x,y)$$

for any fundamental domain $F \subset U$. Now we have

Proposition 2·9. *Let γ be a class A $\widetilde{\rho}$-geodesic segment in U and let L be the h-geodesic segment joining its endpoints. Then*

$$\sup_{y\in L}\inf_{x\in\gamma} d_h(x,y) \le d$$

and

$$\sup_{x\in\gamma}\inf_{y\in L} d_h(x,y) \le d.$$

Proof. Let $x \in \gamma$, and suppose $x \in F$ for some fundamental domain F. Then by lemma 2·8, $L \cap F \neq \emptyset$ so $\inf_{y\in L} d_h(x,y) \le d$, hence

$$\sup_{x\in\gamma}\inf_{y\in L} d_h(x,y) \le d.$$

Similarly,

$$\sup_{y\in L}\inf_{x\in\gamma} d_h(x,y) \le d.$$

□

Lemma 2·10. *The set of class A $\tilde{\rho}$-geodesics is closed in the set of all $\tilde{\rho}$-geodesics.*

Proof. (See [M1] lemma 9.) We need to show that if γ_n are class A $\tilde{\rho}$-geodesics and $\gamma_n \to \gamma$ then γ is also class A. Suppose not. Then there are points $p, q \in \gamma$ and a geodesic segment δ with endpoints p, q such that $L_{\tilde{\rho}}(\delta) < L_{\tilde{\rho}}(\widehat{\gamma})$ where $\widehat{\gamma}$ is the segment of γ with endpoints p, q. Let $\epsilon = L_{\tilde{\rho}}(\widehat{\gamma}) - L_{\tilde{\rho}}(\delta)$. Now, since $\gamma = \lim \gamma_n$, there is a γ_n containing a segment $\widehat{\gamma}_n$ whose endpoints r, s satisfy $d_{\tilde{\rho}}(r, p) < \epsilon/4$, $d_{\tilde{\rho}}(s, q) < \epsilon/4$ and such that $|L_{\tilde{\rho}}(\widehat{\gamma}_n) - L_{\tilde{\rho}}(\widehat{\gamma})| < \epsilon/2$. (See figure 10.) But then there is a curve α with endpoints r, s such that

$$\begin{aligned} L_{\tilde{\rho}}(\alpha) &< \epsilon/4 + \epsilon/4 + L_{\tilde{\rho}}(\delta) \\ &< L_{\tilde{\rho}}(\widehat{\gamma}) - \epsilon/2 \\ &< L_{\tilde{\rho}}(\widehat{\gamma}_n) \end{aligned}$$

contradicting the fact that γ_n is class A. □

Lemma 2·11. *Let γ be a $\tilde{\rho}$-geodesic segment or $\tilde{\rho}$-geodesic which is not a boundary component of $\tilde{D}$. Then γ does not intersect the boundary of $\tilde{D}$, except possibly at an endpoint.*

Proof. Suppose otherwise. Then by lemma 2·5, γ would meet a boundary component of $\tilde{D}$ tangentially. But this contradicts lemma 2·4. □

2·5. Proof of the main theorem

Now we have established lemma 2·9, we may proceed with the proof of the main theorem in essentially the same way as Morse. We remark that by lemma 2·11, we may assume all geodesics we are interested in lie in the interior of U.

Theorem 2·12. *Given any h-geodesic in U there is a class A $\tilde{\rho}$-geodesic of the same type. Conversely, given any class A $\tilde{\rho}$-geodesic which extends infinitely in both directions, there is a unique h-geodesic of the same type.*

Proof. Let L be an h-geodesic in U. Let $\ldots, P_{-2}, P_{-1}, P_0, P_1, P_2, \ldots$ be successive points on L with $d_h(P_n, P_0) \to \infty$ as $|n| \to \infty$. Choose a class A $\tilde{\rho}$-geodesic segment γ_n with endpoints P_{-n} and P_n. By lemma 2·9 there is a point $Q_n \in \gamma_n$ with $d_h(Q_n, P_0) \leq d$. Let $x_n \in S^{\tilde{\rho}}(U)$ be determined by Q_n and the geodesic γ_n. All the x_n lie in $S^{\tilde{\rho}}\{z \in U : d_h(z, P_0) \leq d\}$ which is compact; hence they have a limit point x. Let γ be the geodesic determined by x. Now γ is class A by lemma 2·10, and

$$\sup_{z \in \gamma} \inf_{y \in L} d_h(z, y) \leq d$$

by lemma 2·9, so γ is the same type as L.

To prove the converse, let γ be a class A geodesic extending infinitely in both directions. Let $\ldots P_{-1}, P_0, P_1, \ldots$ be successive points on γ with $d_h(P_n, P_0) \to \infty$ as $|n| \to \infty$. Let L_n be the h-geodesic segment with endpoints P_{-n}, P_n. By lemma 2·9 there is a point $Q_n \in L_n$ with $d_h(Q_n, P_0) \leq d$. As before, define $x_n \in S^h\Delta$ by Q_n and the h-geodesic L_n and let x be a limit point of the x_n. Let L be the h-geodesic defined by x. Again L is of the same type as γ since all the L_n are. Since L is of the same type as γ and $\gamma \in U$, we must have $L \in U$. To see uniqueness, note that any two h-geodesics of the same type are equal. □

We now show how to construct the semiequivalence (Θ, τ) between the set $\mathcal{A}$ and the set S^hM. Recall that $\mathcal{A} \subset S^{\tilde{\rho}}M$ is the set of unit tangent vectors which correspond to projections of class A $\tilde{\rho}$-geodesics. Let $\bar{\mathcal{A}} \subset S^{\tilde{\rho}}U$ be the set of unit tangent vectors to U

which correspond to class A $\tilde{\rho}$-geodesics. Denote by λ the natural projection $\lambda : S^{\tilde{\rho}}U \to U$ and by μ the natural projection $\mu : S^h U \to U$. Fix some fundamental domain $F \in U$ whose boundary is formed from a finite number of simultaneously class A $\tilde{\rho}$-geodesic segments and h-geodesic segments. Let $x \in \tilde{\mathcal{A}}$ with $\lambda(x) \in \partial F$ and suppose that $\lambda(\phi_t^{\tilde{\rho}}(x)) \in F$ for t small and positive (that is, x represents a vector on the boundary of F pointing into F, and corresponding to a class A $\tilde{\rho}$-geodesic). Let $y \in \tilde{\mathcal{A}}$ have $\lambda(y) \in \partial F$, $\phi_t^{\tilde{\rho}}(y) \notin F$ for $t > 0$ and y in the forward orbit of x under $\phi^{\tilde{\rho}}$ (that is, y represents the vector above the boundary of F pointing out of F, and along the same geodesic as x). Let γ be the class A $\tilde{\rho}$-geodesic defined by x and y, and let L be the h-geodesic of the same type as γ. Let $p, q \in S^h U$ satisfy $\mu(p), \mu(q) \in \partial F \cap L$, $\mu(\phi_t^h(p)), \mu(\phi_{-t}^h(q)) \in F$ for t small and positive (that is, p and q are the entry and exit vectors of L in the fundamental domain F). Define $\theta(x) = p$, $\theta(y) = q$ and define θ on points of the flow from x to y by linear interpolation along the geodesic segment, that is, suppose $z \in \tilde{\mathcal{A}}$ with $\lambda(z) \in F \cap \gamma$, and z in the orbit of x and y. There is $r \in S^h U$ with $\mu(r) \in L \cap F$, r in the orbit of p and q, and

$$\frac{d_h(\mu(p), \mu(r))}{d_h(\mu(p), \mu(q))} = \frac{d_{\tilde{\rho}}(\lambda(x), \lambda(z))}{d_{\tilde{\rho}}(\lambda(x), \lambda(y))}.$$

Then let $\theta(z) = r$. We extend the definition of θ to the whole of $\tilde{\mathcal{A}}$ by translation by the group elements $g \in G$.

The linear interpolation induces a reparametrization $\tau : \tilde{\mathcal{A}} \times \mathbf{R} \to \mathbf{R}$, by

$$\tau(z, 0) = 0 \text{ for all } z \in \tilde{\mathcal{A}},$$

$$\frac{d\tau}{dt} = \frac{d_{\tilde{\rho}}(\lambda(x), \lambda(y))}{d_h(\mu(p), \mu(q))}, \text{ and}$$

we extend the definition of τ to the whole of $\tilde{\mathcal{A}} \times \mathbf{R}$ by translation by the group elements $g \in G$ and addition of the changes in τ across each copy gF of the fundamental domain. Now we have:

Proposition 2·13. *$\theta : \tilde{\mathcal{A}} \to S^h U$ is a continuous surjection and induces a map $\Theta : \mathcal{A} \to S^h M$. τ is continuous and*

$$\phi_t^h \circ \Theta = \Theta \circ \phi_\tau^\rho | \mathcal{A}.$$

Proof. Surjectivity of θ follows from theorem 2·12. We will show continuity. Let $x_n \in \widetilde{\mathcal{A}}$ and suppose $x_n \to x$. Then $x \in \widetilde{\mathcal{A}}$ by lemma 2·10. Let γ_n be the class A $\widetilde{\rho}$-geodesic corresponding to x_n, γ the class A $\widetilde{\rho}$-geodesic corresponding to x, L_n the h-geodesic of the same type as γ_n and L the h-geodesic of the same type as γ. For each curve α denote by $\widehat{\alpha}$ the segment $\alpha \cap F$. It is clear that the $\widehat{L}_n$ accumulate on $\widehat{L}$, the $\widehat{\gamma}_n$ on $\widehat{\gamma}$ and that the lengths of the segments converge: $L_{\widetilde{\rho}}(\widehat{\gamma}_n) \to L_{\widetilde{\rho}}(\widehat{\gamma})$ and $L_h(\widehat{L}_n) \to L_h(\widehat{L})$. Thus we have

$$\lim_{n \to \infty} \theta(x_n) = \theta(x)$$

for x in the interior of any fundamental domain. To see continuity of θ across the boundaries of the fundamental domains, it is enough to observe that for $g \in G$, γ is of the same type as L if and only if $g\gamma$ is of the same type as gL. The same argument gives continuity of $\tau(z,t)$ with respect to z. Continuity of τ with respect to t is immediate.

That θ projects to a well defined map $\Theta : \mathcal{A} \to S^h M$ follows from the facts that, for any $g \in G$, the metrics $\widetilde{\rho}$ and h are invariant under g and the class A $\widetilde{\rho}$-geodesic γ is of the same type as the h-geodesic L is and only if $g\gamma$ is of the same type as gL. The final statement of the theorem is immediate. □

Given any surface Σ of genus g and with $s \geq 1$ geodesic boundary components ($2g+s \geq 3$) and Riemannian metric σ, there is a continuous dynamical system generated on $S^\sigma\Sigma$ defined by flow along the geodesics. We have shown that amongst all Riemannian metrics σ on Σ, the metrics of constant negative curvature generate the simplest dynamics. However, the dynamics generated by a metric of constant negative curvature are already complicated; there is a section of the flow on which the dynamics are essentially the same as that of a subshift of finite type, [S]. Hence we see that the dynamics generated by any Riemannian metric on such a surface are complicated.

Figure 8.

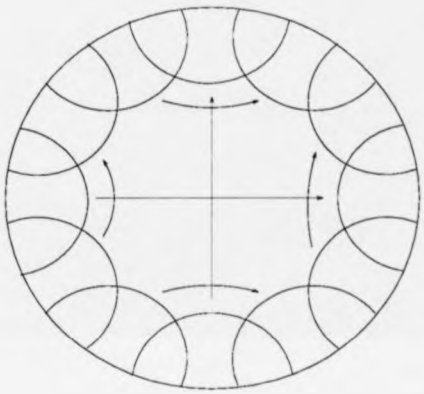

Figure 9 (for the case g=2).

Appendix: A New Proof of Morse's Theorem

In this appendix we will adapt our method of proof to obtain a new proof of the main lemma used in proving Morse's theorem ([M1] lemma 8).

Let M be a compact surface of genus $g \geq 2$ without boundary. We use Koebe's construction [BS] which defines a universal cover which has the desired properties. The free homotopy classes in which lie the curves along which the surface is cut (that is, projections of the boundaries of the fundamental regions in the universal cover Δ) are shown in figure 8. We choose the curves to be minimal ρ-geodesics in their free homotopy class by lemma 2·3. This implies, by lemma 2·7, that each curve intersects each other curve at most once. Hence, by the choice of free homotopy classes, the numbers of intersections are precisely as indicated in figure 8. The covering space is shown in figure 9 (for the case $g = 2$); each fundamental domain is a $(8g-4)$-gon. Let G be the group of transformations

with $\Delta/G \cong M$.

Since each cutting curve on M is a periodic curve intersecting at most two other cutting curves, each h-geodesic containing a side of a fundamental domain is a $\widetilde{\rho}$-geodesic intersecting only its two immediate neighbours. These curves are periodic, that is, for each such geodesic a, there is $g \in G$ with $ga = a$. We firstly show that these are class A $\widetilde{\rho}$-geodesics.

Lemma 2·14. *Let γ be a $\widetilde{\rho}$-geodesic of period g^k which projects onto the shortest closed ρ-geodesic in its homotopy class. Then γ has period g.*

Proof. First observe that γ minimises the distance between any point $p \in \gamma$ and $g^k p$; also that γ has endpoints on S^1 which are fixed points of g.

Suppose $gp \notin \gamma$. We will call the half plane bounded by γ which contains points on the right of γ when proceeding from some point q to $g^k q$ the *positive side* of γ, the other half plane bounded by γ will be the *negative side* of γ. Suppose gp lies in the positive side of γ (the case for the negative side is identical). Consider now the $\widetilde{\rho}$-geodesics $g^i\gamma$ for $i = 0, \ldots, k-1$. For each i, the point $g^{i+1}p$ lies on the positive side of $g^i\gamma$. It follows that some geodesic $g^j\gamma$ intersects γ twice between two points $q, g^k q \in \gamma$. This contradicts the minimality of γ and $g^j\gamma$. Hence $g^i p \in \gamma$ for each i. Now we have that $p, gp \in \gamma \cap g\gamma$, so by the same argument $\gamma = g\gamma$. □

Lemma 2·15. *Let γ be a $\widetilde{\rho}$-geodesic of period $g \in G$ which projects onto the shortest closed ρ-geodesic in the appropriate homotopy class. Then γ is a class A $\widetilde{\rho}$-geodesic.*

Proof. Suppose not. Then there are points $p, q \in \gamma$ and a curve α such that $L_{\widetilde{\rho}}(\alpha; p, q) < L_{\widetilde{\rho}}(\gamma; p, q)$. Let k be the smallest positive integer such that there is $\overline{p} \in \gamma$ with p, q lying in the interior of the segment of γ with enpoints $\overline{p}$ and $g^k\overline{p}$. By minimality of the curve, $k > 1$. Now let $L = L_{\widetilde{\rho}}(\gamma; \overline{p}, g\overline{p})$. Then $L_{\widetilde{\rho}}(\gamma; \overline{p}, g^k\overline{p}) = kL$. Now the curve consisting of the segment of γ from $\overline{p}$ to p, the segment of α from p to q, and the segment of γ from q to $g^k\overline{p}$, has length less than kL. Now let β be the shortest curve in Δ joining points r and $g^k r$ for some r. Then β has length less than kL. Clearly the curve consisting of β and

all its images under g^n, $n \in \mathbb{Z}$ is periodic and is a geodesic (if it contains corners at $g^n r$, then there is a shorter curve with the given properties). But by lemma 2·14, the curve so generated by β has period g, and so there is a curve of period g which projects onto a curve on M of length less than that of the projection of γ, contradicting the minimality of γ. □

Hence we see that each class A $\tilde{\rho}$-geodesic different from a boundary geodesic of the fundamental domains intersects any given boundary of the fundamental domains at most once.

The next lemma is analogous to lemma 2·8, but slightly weaker.

Lemma 2·16. *Let γ be a class A $\tilde{\rho}$-geodesic segment and let L be the h-geodesic with the same endpoints. Let L pass through a fundamental domain F. Then γ either passes through F or passes through a fundamental domain touching F at a vertex. Conversely, let γ pass through a fundamental domain F'. Then L either passes through F' or passes through a fundamental domain touching F' at a vertex.*

Proof. Suppose L passes through F. If either endpoint of L is in F, then we are done. Suppose then that no endpoints of L are in F.

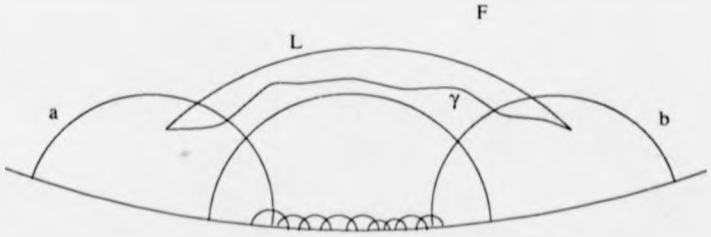

Figure 10.

Suppose firstly that L intersects two sides a, b of F which are not adjacent. Then γ must also intersect the h-geodesics containing each of these sides. Since these h-geodesics are disjoint, if $\gamma \cap F = \emptyset$ then γ must intersect a geodesic containing a side of F lying

between a and b at least twice. See figure 10. We have just shown this is impossible, so $\gamma \cap F \neq \emptyset$.

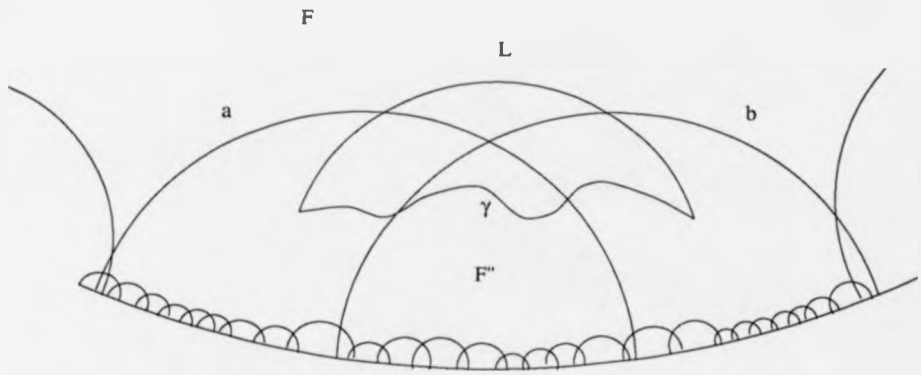

Figure 11.

Now suppose L intersects two adjacent sides a, b of F. Denote by F'' the fundamental domain which meets F only at the vertex of a and b. Again, γ must intersect the h-geodesics containing a and b. Suppose γ does so in the same order as L, then clearly $\gamma \cap F \neq \emptyset$. Suppose γ intersects a and b in the opposite order to L, then we have $\gamma \cap F'' \neq \emptyset$. See figure 11.

The proof of the converse is similar. (The only property we have used of the geodesic γ is that it does not intersect any boundary of a fundamental domain more than once. Since boundaries of fundamental domains are h-geodesics the same is true for L, so the rôles of γ and L in this lemma are interchangeable.) □

Lemma 2·17. *(See [M1] lemma 8.) Let*

$$d = \sup_{x,y \in F} h_h(x, y)$$

for any fundamental domain F. Let γ, L be as in lemma 2·16. Then

$$\sup_{x \in \gamma} \inf_{y \in L} d_h(x, y) \leq 2d$$

and

$$\sup_{y \in L} \inf_{x \in \gamma} d_h(x, y) \leq 2d.$$

Proof. This is immediate from lemma 2·16. □

Having established lemma 2·17, we may proceed in the same manner as Morse [M1] to prove theorem 2·1, as in theorem 2·12.

To establish the semiequivalence θ similar to that in theorem 2·13, we need to exercise a little more care. We proceed as follows. As before, let $\mathcal{A} \subset S^{\bar{\rho}}\Delta$ be the set of unit tangent vectors corresponding to class A $\bar{\rho}$-geodesics. Fix some fundamental domain F. Suppose $x \in \mathcal{A}$ satisfies:

(i) $\lambda(x) \in \partial F$,

(ii) $\lambda(\phi_t^{\bar{\rho}}(x)) \in F$ for t small and positive, and

(iii) if x corresponds to a class A $\bar{\rho}$-geodesic γ and L is the h-geodesic of the same type as γ, then $L \cap F \neq \emptyset$.

Then let $r \in S^h\Delta$ be the unit tangent vector corresponding to the h-geodesic L satisfying $\mu(r) \in \partial F$ and $\mu(\phi_t^h(r)) \in F$ for t small and positive, and define $\theta(x) = r$. Now suppose $y \in \mathcal{A}$ is in the forward orbit of x and is the first such vector satisfying $\lambda(y) \in \partial F'$ for some fundamental domain $F' \neq F$ with $F' \cap L \neq \emptyset$. Let s be the first vector in the forward orbit of r with $\mu(s) \in \partial F'$. Then define $\theta(y) = s$, and use linear interpolation along γ from x to y to define θ on vectors in $\mathcal{A}$ representing the segment of γ from x to y.

This defines a continuous surjection from $\mathcal{A}$ to $S^h\Delta$ which, after the corresponding time reparametrization, commutes with the flow, hence a semiequivalence between the flow generated by class A ρ-geodesics and that generated by h-geodesics on the surface M.

References

[Ba] Bangert, V. [1988]. Mather Sets for Twist Maps and Geodesics on Tori. In *Dynamics Reported*, **1**, U. Kirchgraber and H.O. Walther (eds), pp 1–56. Wiley.

[Be] Beardon, A.F. [1982]. *The Geometry of Discrete Groups*, Springer-Verlag.

[Bi] Birkhoff, G.D. [1917]. Dynamical Systems with Two Degrees of Freedom. *Trans. AMS*, **18**, 219.

[BS] Birman, J. and Series, C. [1984]. Dehn's Algorithm Revisited, with Applications to Simple Closed Curves on Surfaces. *Proc. Conf. on Combinatorial Group Theory, Utah*, Gerston and Stallings (eds), 451–478.

[C] Carathéodory, C. [1967]. *Calculus of Variations and Partial Differential Equations of the First Order. Part II: Calculus of Variations*. Holden-Day.

[H] Hedlund, G.A. [1932] Geodesics on a Two-Dimensional Riemannian Manifold with Periodic Coefficients. *Ann. Math.*, **33**, 719–739.

[K] Klingenberg, W. [1978] *Lectures on Closed Geodesics*. Springer-Verlag.

[M1] Morse, H.M. [1924] A Fundamental Class of Geodesics on any Closed Surface of Genus Greater than One. *Trans. AMS*, **26**, 25–60.

[M2] Morse, H.M. [1935] Instability and Transitivity. *Journ. de Math.* **XIV**, 49–71.

[P] Polterovich, L.V. [1989]. Geodesics on the Two-Dimensional Torus with Two Rotation Numbers. *Math. USSR Izvestiya*, **33**, 1, 101–114.

[S] Series, C. [1986]. Geometrical Markov Coding of Geodesics on Surfaces of Constant Negative Curvature. *Ergodic Theory and Dynamical Systems*, **6**, 601–625.

TITLE TWO TOPICS IN DYNAMICS.

AUTHOR James DENVIR

DEGREE Ph.D

AWARDING BODY Warwick University

DATE 1994

THESIS NUMBER DX187296

DX

187296