Smale Spaces

Ian F. Putnam

February 14, 2012

Contents

Chapter 1

Definitions

1.1 Dynamical preliminaries

In this section, we will introduce some simple notions in the study of topological dynamical systems: periodic points, non-wandering, irreducibilty and mixing. Each can be regarded as a type of recurrence property.

We will always work in the topological category and, more specifically, usually with metric spaces. Let (X, d) be a metric space. For any x in X and $\epsilon > 0$, we will let $X(x, \epsilon)$ denote the open ball of radius ϵ centred at X. This notation is preferable the more common $B(x, \epsilon)$ when dealing with more than a single space.

There is a wide range of notions of topological dynamical systems, but for us here it will be a homeomorphism $f : X \to X$. We begin with the simplest of the concepts, that of a periodic point.

Definition 1.1.1. *Let (X, d, f) be as above. We say that x in X is a* fixed-point *of f if $f(x) = x$. If n is a positive integer and x is in X, we say that x is a* periodic point of period n *if $f^n(x) = x$. The least such positive integer n is called the* period *of x. For any positive integer m, we let $Per_m(X, f)$ denote the set of all periodic points of period m. We also let*

$$Per(X, f) = \cup_{m \geq 1} Per_m(X, f)$$

denote the set of all periodic points.

Next, we consider the notion of non-wandering. As we mentioned above, this is a kind of recurrence condition on the points of X. There are a number of these available, but this is the most natural for the systems which we will consider later.

Definition 1.1.2. *Let (X, d, f) be as above and let x be any point of X. We say that x is* non-wandering *if, for every non-empty open set, U, containing x, there is a positive integer n such that $f^n(U) \cap U$ is non-empty. Any point which is not non-wandering is called* wandering.

We let $NW(X, f)$ denote the set of non-wandering points of X. We say that (X, d, f) is non-wandering *if every point of X is non-wandering.*

Let us make a few simple remarks about the definition.

1. A point is non-wandering if and only if, for every non-empty set U containing the point, there is a z in U and positive integer n with $f^n(z)$ also in U. (If y is in $f^n(U) \cap U$, let $z = f^{-n}(y)$.)

2. Every periodic point is clearly non-wandering. Moreover, if the periodic points of X are dense, then every point is non-wandering.

3. The set of non-wandering points is f-invariant. That is, x is non-wandering if and only if $f(x)$ is also.

4. The set of non-wandering points is closed. To see this, we show the compliment is open. If x is wandering, then there is an open set U containing x such that $f^n(U)$ and U are disjoint for every positive integer n. But then, every point of U is also wandering, using the same set U as neighbourhood.

Next, we turn to the definition of irreducibility.

Definition 1.1.3. *We say the system (X, d, f) is* irreducible *if, for every (ordered) pair of non-empty open sets, U, V, there is a positive integer n such that $f^n(U) \cap V$ is non-empty.*

It is clear that every irreducible system is non-wandering. The converse is false. For example the identity map on a set X (having at least two points) is non-wandering but not irreducible.

We note the following standard result (for example, see Theorem 5.9 of [**?**]).

Theorem 1.1.4. *Let (X, d, f) be as above with X compact. The following conditions are equivalent.*

1. *(X, d, f) is irreducible.*

2. *(X, d, f^{-1}) is irreducible.*

3. *There is a dense G_δ-set in X such that every point x in this set has a forward orbit, $\{f^n(x) \mid n \geq 0\}$, which is dense in X.*

4. *There exists a point x whose forward orbit is dense in X.*

Finally, we turn to the definition of mixing.

Definition 1.1.5. *We say the system (X, d, f) is* mixing *if, for every (ordered) pair of non-empty open sets, U, V, there is a positive integer N such that $f^n(U) \cap V$ is non-empty for all $n \geq N$.*

We observe that every mixing system is also irreducible and hence non-wandering as well. The converse is false. Consider the case that X consists of two points which the map f exchanges. This is irreducible, but not mixing.

1.2 An introduction to Smale spaces

In this section, we will provide a heuristic discussion of Smale spaces. This is intended as motivation and will be rather short on rigour. It is important to proceed in this way because the rigourous definition - which we will see in the next section - is really quite opaque without a preliminary discussion to provide some kind of insight.

We will consider a compact metric space (X, d) and a homeomorphism $f : X \to X$. We will require some extra structure. This will take quite some time to describe.

First, for every point x in X, we will have two closed sets E_x and F_x, which we will call the local stable set of x and the local unstable set of x, respectively, having a number of special properties. We require

P1

$$E_x \cap F_x = \{x\}$$

P2 The cartesian product $E_x \times F_x$ is homeomorphic to a neighbourhood of x in X.

This second item is really too vague. The proper definition in the next section will actually specify this homeomorphism and a number of its properties. But for the moment, this will be enough. That is, locally, X is the product of E_x and F_x.

It is worth noting at this point that the sets E_x and F_x are not unique. For example, if we make both smaller, so long as x is still in the interior of the cartesian product, the result would also satisfy our conditions. They are unique in the sense that any two such choices for E_x will be equal in some neighbourhood of x.

Next, we want to require that these sets be invariant under f. This is a little too much to ask, especially in view of the comments of the last paragraph. Instead, we only require invariance in a local sense.

P3 For all x in X,

$$\begin{aligned} f(E_x) \cap V &= E_{f(x)} \cap V \\ f(F_x) \cap V &= F_{f(x)} \cap V \end{aligned}$$

for some neighbourhood, V, of $f(x)$.

Finally, we come to the crucial conditions: f is contracting on the sets E_x while it is expanding on F_x. For various technical reasons, it is much better to say that f^{-1} is contracting on F_x. Specifically, there is a constant $0 < \lambda < 1$ such that

P4 for all y, z in E_x, we have

$$d(f(y), f(z)) \leq \lambda d(y, z)$$

P5 and for all y, z in F_x, we have

$$d(f^{-1}(y), f^{-1}(z)) \leq \lambda d(y, z).$$

This condition also re-inforces our earlier statement that exact invariance of E_x is not reasonable. We expect that $f(E_x)$ will be smaller than $E_{f(x)}$.

The next section will provide the rigourous definition of Smale space. Alternately, the reader can pass on to the examples in the subsequent section. Most of these should be understandable with the vague notion of Smale space which we have now, although there will be a little new notation in the next section.

1.3 The definition of Smale space

We are now ready to give a precise definition of a Smale space in this section. The definition is rather long. As we go, we will try to provide some comparison with the heuristic version given in the last section.

We begin with a compact metric space (X, d). We let $f : X \to X$ be a homeomorphism of X.

We assume that there is a constant ϵ_X and a map defined on

$$\Delta_{\epsilon_X} = \{(x, y) \mid d(x, y) \leq \epsilon_X\}$$

taking values in X. The map should be continuous in the natural product topology. The image of (x, y) is denoted $[x, y]$. We assume that this satisfies certain axioms.

Before beginning the axioms, let us mention that if one has the heuristic description of Smale space in the last section, then $[x, y]$ should be thought of as the intersection of the sets E_x and F_y. If we had been more rigourous in the earlier discussion, we should have added hypotheses that such sets would meet in exactly one point, provided x and y are within ϵ_X.

We require $[,]$ to satisfy the following

B1 $[x, x] = x$,

B2 $[x, [y, z]] = [x, z]$, whenever both sides are defined,

B3 $[[x, y], z] = [x, z]$, whenever both sides are defined,

B4 $[f(x), f(y)] = f([x, y])$, whenever both sides are defined.

In terms of the description of the last section and the idea that $[x, y]$ is the intersection of E_x and F_y, B1 and B4 are equivalent to P1 and P3 of the last section. The axioms B2 and B3 should be regarded as implying a kind of compatibility between the local product structures of X at nearby points.

Finally, we require that there is a constant $0 < \lambda < 1$ such that, for all x in X, we have the following two conditions.

C1 For y, z such that $d(x, y), d(x, z) \leq \epsilon_X$ and $[y, x] = x = [z, x]$, we have

$$d(f(y), f(z)) \leq \lambda d(y, z).$$

C2 For y, z such that $d(x, y), d(x, z) \leq \epsilon_X$ and $[x, y] = x = [x, z]$, we have

$$d(f^{-1}(y), f^{-1}(z)) \leq \lambda d(y, z).$$

These axioms are obviously analogous to P4 and P5 of the last section.

Definition 1.3.1. *A* Smale space *is any quadruple* $(X, d, f, [,])$ *satisfying the axioms B1, B2, B3, B4, C1 and C2.*

A word of warning is in order. The most annoying thing in dealing with Smale spaces is that the bracket map is only defined on points which are close. It is very important to check this at all times, because it is quite easy to be lead to false conclusions if this is ignored. As an example, it is tempting to say that, if x, y are in X and n is a positive integer and $d(x, y), d(f^n(x), f^n(y))$ are both less than ϵ_X, then

$$[f^n(x), f^n(y)] = f^n([x, y]).$$

In fact, this may be false, unless $d(f^k(x), f^k(y))$ is less than ϵ_X for every $1 \leq k \leq n$.

For convenience, we define, for each x in X and $0 < \epsilon \leq \epsilon_X$, sets

$$X^s(x, \epsilon) = \{y \mid d(x, y) \leq \epsilon, [y, x] = x\} \quad (1.1)$$

$$X^u(x, \epsilon) = \{y \mid d(x, y) \leq \epsilon, [x, y] = x\}. \quad (1.2)$$

These will be referred to as the *local stable set* and *local unstable set* at x. We quickly observe the alternate characterization of these points.

Lemma 1.3.2. *Suppose* $d(x, y) \leq \epsilon_X$.

1. $[x, y] = x$ *if and only if* $[y, x] = y$.
2. $[x, y] = y$ *if and only if* $[y, x] = x$.

Proof. We prove only the "only if" part of the first statement. The others are similar. We calculate

$$\begin{aligned} [y, x] &= [y, [x, y]] && \text{by hypothesis,} \\ &= [y, y] && \text{by axiom B2,} \\ &= y && \text{by axiom B1.} \end{aligned}$$

□

Lemma 1.3.3. *Suppose that* x *and* y *are in* X *and* $d(x, y), d(x, [x, y])$ *and* $d(y, [x, y])$ *are all less than* ϵ_X. *Then we have*

$$\begin{aligned} [x, y] &\in X^s(x, \epsilon_X) \\ [x, y] &\in X^u(y, \epsilon_X) \end{aligned}$$

Proof. For the first, we check

$$\begin{aligned} [[x, y], x] &= [x, x] && \text{by hypothesis B3,} \\ &= x && \text{by hypothesis B1.} \end{aligned}$$

For the second, we check

$$\begin{aligned} [y, [x, y]] &= [y, y] && \text{by hypothesis B2,} \\ &= y && \text{by hypothesis B1.} \end{aligned}$$

□

The sets $X^s(x,\epsilon)$ and $X^u(x,\epsilon)$ are exactly the sets E_x and F_x of the last section, except that we have added a parameter ϵ to allow us to control their size. We have now set things up in such a way that our earlier hypothesis P2 is now a consequence of the other axioms.

Theorem 1.3.4. *There is $0 < \epsilon'_X \leq \epsilon_X/2$ such that, for every $0 < \epsilon \leq \epsilon'_X$, the map*

$$[\,,\,] : X^u(x,\epsilon) \times X^s(x,\epsilon) \to X$$

is a homeomorphism to its image, which is a neighbourhood of x. We will denote this range by $U(x,\epsilon)$.

Proof. First we note that the map is well-defined, since if both y and z are within ϵ of x and $\epsilon \leq \epsilon_X/2$, then $d(y,z) \leq \epsilon_X$ by the triangle inequality. Moreover, since $[\,,\,]$ is jointly continuous, we may find $0 < \delta \leq \epsilon_X$ such that, for all x, y with $d(x,y) \leq \delta$, we have $d(x,[x,y]) \leq \epsilon_X/2$ and $d(x,[y,x]) \leq \epsilon_X/2$. We choose $0 < \epsilon'_X \leq \epsilon_X/2$ so that for all y, z with $d(x,y) \leq \epsilon'_X$ and $d(x,z) \leq \epsilon'_X$, we have $d(x,[y,z]) \leq \delta$. Then we can define a map h on a neighbourhood of x by $h(y) = ([y,x],[x,y])$. By the choice of ϵ'_X this map is defined on the range of $[\,,\,]$. It is also clearly continuous. It is clear from the axioms B1, B2 and B3 that the composition $[\,,\,] \circ h$ is the identity. Moreover, if we begin with y in $X^u(x,\epsilon)$ and z in $X^s(x,\epsilon)$, then we have

$$\begin{aligned} h([y,z]) &= ([[y,z],x],[x,[y,z]]) \\ &= ([y,x],[x,z]) && \text{by Axioms B2 and B3,} \\ &= (y,z) && \text{by Lemma 1.3.2.} \end{aligned}$$

The conclusion follows. □

The first important fact which we want to establish is that the choice of the bracket map is unique (up to the choice of its domain). Once this is established, then we can speak of (X,d,f) being a Smale space when such a bracket exists.

We begin with the following lemma.

Lemma 1.3.5. *Suppose that $(X,d,f,[\,,\,])$ is a Smale space. Then there is a constant $0 < \epsilon_1$ satisfying the following, for all $0 < \epsilon \leq \epsilon_1$.*

ES *If x and y are in X and $d(f^n(x),f^n(y)) \leq \epsilon$, for all $n \geq 0$, then y is in $X^s(x,\epsilon)$.*

EU *If x and y are in X and $d(f^n(x), f^n(y)) \leq \epsilon$, for all $n \leq 0$, then y is in $X^u(x, \epsilon)$.*

Proof. Choose $0 < \epsilon_1 \leq \epsilon_X$ so that, for all x, y with $d(x, y) < \epsilon_1$, we have $d([y, x], x) < \epsilon_X$. It follows then that for x, y satisfying the hypothesis of ES, we have $[f^n(y), f^n(x)]$ is in $X^u(f^n(x), \epsilon_X)$, for all $n \geq 0$. Now we apply hypothesis B4 to note that

$$f^{-1}[f^n(y), f^n(x)] = [f^{n-1}(y), f^{n-1}(x)]$$

provided n is positive. Then we apply hypothesis C2 to assert

$$\begin{aligned} d(f^{n-1}(x), [f^{n-1}(y), f^{n-1}(x)]) &= d(f^{-1}f^n(x), f^{-1}[f^n(y), f^n(x)]) \\ &\leq \lambda d(f^n(x), [f^n(y), f^n(x)]). \end{aligned}$$

Then an easy induction shows that, for all $n \geq 0$, we have

$$\begin{aligned} d(x, [y, x]) &\leq \lambda^n d(f^n(x), [f^n(y), f^n(x)]) \\ &\leq \lambda^n \epsilon_X. \end{aligned}$$

Since $\lambda < 1$, we conlcude that $x = [y, x]$ and hence the conclusion. We have shown the first statement is true; the proof of the second is analogous and we omit the details. □

This lemma has an immediate consequence of some interest, that is, the systems we are considering are expansive. This means that there is a positive constant (here, ϵ_1) so that any two distinct points, no matter how close, may be separated by at least this constant, by applying the map f (or f^{-1}) a number of times to both.

Corollary 1.3.6. *The map f is expansive for the constant ϵ_1; i.e. if x and y are in X and $d(f^n(x), f^n(y)) \leq \epsilon_1$, for all integers n, then $x = y$.*

Theorem 1.3.7. *Let $(X, d, f, [,])$ be a Smale space and let ϵ_1 be as in the last lemma. If x, y are in X and $d(x, y), d(x, [x, y]), d(y, [x, y])$ are all less than $\epsilon_1/2$, then*

$$\begin{aligned} \{[x, y]\} &= \{z \mid d(f^n(x), f^n(z)) < \epsilon_1/2, \\ &\qquad d(f^{-n}(y), f^{-n}(z)) < \epsilon_1/2, \text{ for all } n \geq 0\} \\ &= X^s(x, \epsilon_1/2) \cap X^u(y, \epsilon_1/2). \end{aligned}$$

Proof. The last lemma asserts that the second set in the statement is contained in the third. It is easy to check from the definitions and the axioms C1 and C2 that the reverse containment also holds.

It follows from the definitions, the triangle inequality and expansiveness that the second set is at most a single point. Finally, we note that $[x, y]$ is contained in the second set by Lemma 1.3.2 and the conclusion follows. $\square$

We note that this theorem says, among other things, that the bracket map is uniquely determined by (X, d, f), provided that it exists.

Chapter 2

Examples

In this chapter, we will introduce a number of examples of Smale spaces.

2.1 Shifts of Finite Type

The following class of examples are, in a certain sense, the most important. These are the *shifts of finite type* (or occassionally, the *subshifts* of finite type). We will first give the abstract definition. We will not provide any proofs of the assertions we make. Most are fairly easy and all can be found in [**?**].

Let $\mathcal{A}$ denote a finite non-empty set, sometimes called the alphabet. We consider the space of doubly infinite sequences in $\mathcal{A}$

$$\mathcal{A}^{\mathbb{Z}} = \{(a_n)_{n \in \mathbb{Z}} \mid a_n \in \mathcal{A}, \text{ for all } n \in \mathbb{Z}\}.$$

For each a in $\mathcal{A}^{\mathbb{Z}}$ and pair of integers $i \leq j$, we let

$$a_{[i,j]} = (a_i, a_{i+1}, \ldots, a_j),$$

whch is in $\mathcal{A}^{j-i+1}$. We define a metric on $\mathcal{A}^{\mathbb{Z}}$ by

$$d(a, b) = \inf\{1, 2^{-|n|} \mid n \geq 1, a_{[1-n,n]} = b_{[1-n,n]}\}.$$

where $a = (a_n)_{n \in \mathbb{Z}}, b = (b_n)_{n \in \mathbb{Z}}$ are in $\mathcal{A}^{\mathbb{Z}}$. The most useful interpretation of this metric is the following easy result.

Lemma 2.1.1. *Let a, b be in $\mathcal{A}$. For any $n \geq 0$, we have $d(a, b) < 2^{-n}$ if and only if $a_i = b_i$, for $-n \leq i \leq n+1$.*

This is a compact metric space whose topology is generated by sets which are both closed and open. We call such sets clopen. For an example of such a set, fix a finite length sequence $(a_m, a_{m+1}, \ldots, a_n)$ of elements in $\mathcal{A}$, for integers $m \leq n$ and we let

$$U = \{b \in \mathcal{A}^{\mathbb{Z}} \mid b_i = a_i, m \leq i \leq n\}.$$

We define the *shift map* σ_A or just σ by

$$(\sigma_A(a))_n = a_{n+1}$$

for any a in $\mathcal{A}^{\mathbb{Z}}$ and n in $\mathbb{Z}$. If one writes out the elements of $\mathcal{A}^{\mathbb{Z}}$ as an infinite string, it is a little hard to make sense of the map; it looks like nothing is happening. The point is that one must keep track of the 0 entry of the sequence. We can do this by inserting a dot between the entries 0 and 1. Then our map looks like

$$\sigma_A(\ldots a_{-2}a_{-1}a_0.a_1a_2\ldots) = (\ldots a_{-2}a_{-1}a_0a_1.a_2\ldots)$$

so that every entry is moved one place to the left. (Sometimes for emphasis, we call σ_A the *left* shift.) It is easy to check that σ_A is a homeomorphism of $\mathcal{A}^{\mathbb{Z}}$.

We define the bracket on X as follows: for a and b in $\mathcal{A}^{\mathbb{Z}}$, we define

$$[a, b]_n = \begin{cases} b_n & n \leq 0 \\ a_n & n \geq 1 \end{cases}$$

for all n in $\mathbb{Z}$. At this point, the bracket is defined for all pairs a and b and we leave it as an easy exercise to see that conditions B1, B2 and B3 are satisfied. Condition B4 however, is only satisfied if we set $\epsilon_{\mathcal{A}^{\mathbb{Z}}} = 1$; for if $d(a, b) < 1$, then $a_0 = b_0$ and $a_1 = b_1$ and it is then an easy matter to check that $[\sigma(a), \sigma(b)] = \sigma[a, b]$. Turning to condition C1, suppose that $d(a, b), d(a, c) < 1$ and $[b, a] = a$ and $[c, a] = a$. It follows from the first condition that $a_0 = b_0 = c_0$ and from the second that $a_n = b_n = c_n$, for all $n \geq 1$. Assuming that $b \neq c$, $d(b, c) = 2^n$, where n is the greatest integer (necessarily negative) such that $b_n \neq c_n$. It easily follows that $d(\sigma(b), \sigma(c)) = 2^{n-1} = 2^{-1}d(b, c)$. The condition C2 is proved in an analogour way.

Here we extend the class of systems considered in the last section by taking subsystems. Of course, we must establish a convenient condition for the subsystem to retain the property of being a Smale space.

If $w = (w_1, \ldots, w_n)$ is a finite sequence in elements of $\mathcal{A}$, we say that w is a *word* in $\mathcal{A}$. Given a in $\mathcal{A}^{\mathbb{Z}}$, we say that w appears in a if, for some $k \in \mathbb{Z}$,

$$(a_{k+1}, \ldots, a_{k+n}) = (w_1, \ldots, w_n).$$

Let $\mathcal{F}$ be a finite (possibly empty) collection of words in $\mathcal{A}$. We define

$$X_{\mathcal{F}} = \{a \in \mathcal{A}^{\mathbb{Z}} \mid \text{ no word in } \mathcal{F} \text{ appears in } a\}.$$

It is easy to see that this is a closed subset of $\mathcal{A}^{\mathbb{Z}}$. It is possible for this to be empty at this stage. Moreover, it is invariant under σ_A; a is in $X_{\mathcal{F}}$ if and only if $\sigma_A(a)$ is.

The restriction of σ_A to $X_{\mathcal{F}}$ is denoted by $\sigma_{\mathcal{F}}$ or just σ. Any non-empty system obtained as $(X_{\mathcal{F}}, \sigma_{\mathcal{F}})$ is called a shift of finite type.

Now that we have given the abstract definition, we will not use it again. We will instead produce two classes of examples of shifts of finite type. However, our classes are exhaustive in the sense that every shift of finite type is topologically conjugate to one in each class.

First, we let N be a positive integer and let A be an $n \times N$-matrix with entries which are either 0 or 1. The (i, j) entry of A is denoted $A(i, j)$.

We define

$$\Sigma_A = \{(a_n)_{n\in\mathbb{Z}} \mid A(a_i, a_{i+1}) = 1, \text{ for all } i \in \mathbb{Z}\}.$$

and σ_A is the left shift map on this space. This is the shift of finite type associated with the collection of words

$$\mathcal{F} = \{(i, j) \mid A(i, j) = 0\}.$$

It is best to think of N states which are labelled $1, \ldots, N$. The matrix describes which transitions between states are allowed. The transition from i to j is allowed when $A(i, j) = 1$. Then the elements of Σ_A can be thought of as infinite sequences of states where each successive transition is allowed.

We construct another example of a shift of finite type as follows. Let G be a (finite) directed graph. That is, G consists of a vertex set, G^0, and an edge set, G^1, and two maps $i, t : G^1 \to G^0$. This can be viewed geometrically as follows. Each vertex is a point and each edge, e, is drawn as an arrow from $i(e)$ ("i" for "initial") to $t(e)$ ("t" for "terminal").

We then define

$$\Sigma_G = \{(e_n)_{n=-\infty}^{\infty} \mid e_n \in G^1, t(e_n) = i(e_{n+1}), \text{ for all } n\},$$

which can be viewed as the space of doubly infinite paths in the graph. We assume the graph has at least one cycle so that Σ_G is non-empty. We define a metric on Σ_G exactly as before, by

$$d(e, f) = \inf\{1, 2^{-|n|} \mid n \geq 1, e_{[1-n,n]} = f_{[1-n,n]}\}.$$

where $e = (e_n)_{-\infty}^{\infty}, f = (f_n)_{-\infty}^{\infty}$ are in Σ_G. The map σ_G is just the left shift on these sequences of edges.

To see this is a shift of finite type, we let $\mathcal{A} = G^1$ and $\mathcal{F} = \{(e, f) \mid t(e) \neq i(f)\}$. It is easy to see that $X_{\mathcal{F}} = \Sigma_G$.

We want to describe one more method of constructing a shift of finite type. If A is any $N \times N$-matrix with non-negative integer entries, we define Σ_A as follows. Let G be the directed graph with $G^0 = \{1, 2, \ldots, N\}$ and an edge set G^1 so that from vertex i to vertex j there are exactly $A(i, j)$ edges. We then define $\Sigma_A = \Sigma_G$. It is easy to see that the result, up to a relabelling of symbols, is independent of G.

However, this creates a small problem. If A is an $N \times N$-matrix with 0-1 entries, we now have two distinct definitions of Σ_A. In the first, the symbol set for our shift is $\{1, 2, \ldots, N\}$, while in the second it is the set of edges of some graph and the number of symbols is just the total number of ones in the matrix. We will never quite resolve this problem except to say that usually, in this case, we mean the latter. We will, however, describe methods that take us back and forth between these two constructions.

Let A be an $N \times N$-matrix with 0-1 entries. We have our shift Σ_A as first described above. Next, we construct a graph G with vertex set $G^0 = \{1, 2, \ldots, N\}$ and edge set $G^1 = \{(i, j) \mid A(i, j) = 1\}$. We set $i(i, j) = i$ and $t(i, j) = j$; that is, (i, j) is an edge from i to j. We claim that the two shifts Σ_A and Σ_G are topologically conjugate. If (a_n) is any sequence in Σ_A, then $((a_n, a_{n+1})_{n\in\mathbb{Z}})$ is in Σ_G. This map is clearly continuous and shift commuting. It is easy to check that it is both injective and surjective and hence a topological conjugacy.

Somewhat less obviously, there is a way to begin with a directed graph G and find a square matrix, A, with 0-1 entries so that Σ_G and Σ_A are topologically conjugate. We let N be the number of edges in G and list the edge set $G^1 = \{e_1, e_2, \ldots, e_N\}$. We define $A(i, j) = 1$, if $t(e_i) = i(e_j)$ and $A(i, j) = 0$ otherwise. If (a_n) in is Σ_A, then (e_{a_n}) is in Σ_G. This is clearly a continuous, shift commuting map. We leave it to the reader to check that it is a bijection.

So we have ways to pass back and forth between our two classes of shifts of finite type. A word of warning is in order. These are not inverses to each other. That is, if one begins with a 0-1 matrix A, constructs a graph G and then from the graph constructs a 0-1 matrix A', then A' will not be A.

Finally, we mention that any shift of finite type is conjugate to one in either class above. We will not give a proof of this fact, but we will use it frequently.

Now let us turn to the issue of why such a system is a Smale space. We will use the graph version Σ_G. For any $e = (e_n)_{n\in\mathbb{Z}}$ in Σ_G, we define the sets

$$\begin{aligned} E_e &= \{e' \in \Sigma_G \mid e'_n = e_n, \text{ for all } n \geq 0\} \\ F_e &= \{e' \in \Sigma_G \mid e'_n = e_n, \text{ for all } n \leq 0\} \end{aligned}$$

It is clear that these two sets intersect exactly at e. Moreover, given e' in E_e and e'' in F_e, we form the sequence

$$f_n = \begin{cases} e'_n & n \leq 0 \\ e''_n & n \geq 0 \end{cases}$$

It is easy to check that the construction is a homeomorphism between $E_e \times F_e$ and the set $\{f \in \Sigma_G \mid f_0 = e_0\}$ which is a neighbourhood of e.

Let us consider the contracting/expanding structure of σ on these sets. If e', e'' are both in E_e, then $e'_n = e_n = e''_n$ for all $n \geq 0$ and so $d(e', e'') = 2^{-n}$, where n is the largest positive integer such that $e_n \neq f_n$. Similarly, $d(\sigma(e'), \sigma(e'')) = 2^{-m}$, where m is the largest positive integer such that $\sigma(e')_m \neq \sigma(e'')_m$. Then from the definiton of σ, we see that $m = n + 1$ and so

$$d(\sigma(e'), \sigma(e'')) = \frac{1}{2} d(e', e'').$$

In exactly the same way, it can be shown that, if e' and e'' are in F_e, then

$$d(\sigma^{-1}(e'), \sigma^{-1}(e'')) = \frac{1}{2} d(e', e'').$$

If we want to consider the more rigourous definition of Smale space, we define the operation $[,]$ as follows. We set $\epsilon_X = 1/2$. Now if e and f are in Σ_G and $d(e, f) \leq \epsilon_X$, it follows from the definition of the metric that we must have $e_0 = f_0$. In this case, we set

$$[e, f] = \begin{cases} f_n & n \leq 0 \\ e_n & n \geq 0 \end{cases}$$

The important thing to observe at this point is that $[e, f]$ is again in Σ_G. This is a consequence of the definition of Σ_G and is not true for more general closed subsets of the space of bi-infinite sequences.

2.2 Anosov Diffeomorphisms

We will begin with a very specific example and then discuss some generalizations. Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Observe that $\det(A) = 1$. We first regard A as a linear map of $\mathbb{R}^2$. As $\det(A) = 1$, $A\mathbb{Z}^2 = \mathbb{Z}^2$ and so A induces a map, f, of the quotient $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, which is the 2-torus. Let q denote the quotient map from $\mathbb{R}^2$ to $\mathbb{T}^2$. The metric we put on $\mathbb{T}^2$ is the quotient one. This means that for points x, y in $\mathbb{R}^2$ which are sufficiently close, $d(q(x), q(y)) = |x - y|$. We claim that $(\mathbb{T}^2, f)$ is a Smale space. (Actually, $(\mathbb{R}^2, A)$ would have been a Smale space, except for the fact that $\mathbb{R}^2$ is not compact; this is a crucial axiom.)

To see the local product structure, we need a description of the eigenvalues and eigenvectors of A. Let $\gamma = (1 + \sqrt{5})/2$, which satisfies $\gamma^2 = \gamma + 1$ and $\gamma > 1$. The eigenvalues of A are λ^2 and λ^{-2}. The associated eigenvectors are $v_1 = (\gamma, 1)$ and $v_2 = (-1, \gamma)$. For any point x in $\mathbb{R}^2$, we define

$$\begin{aligned} E_{q(x)} &= \{q(x + tv_2) \mid |t| \leq \epsilon\} \\ F_{q(x)} &= \{q(x + tv_1) \mid |t| \leq \epsilon\} \end{aligned}$$

where $\epsilon > 0$ is some sufficiently small fixed parameter. If $y = x + tv_2, z = x + sv_2$, for $|s|, |t| \leq \epsilon$, then we have

$$\begin{aligned} d(f(q(y)), f(q(z))) &= d(q(Ay), q(Az)) \\ &= |Ay - Az| \\ &= |A(x + tv_2) - A(x + sv_2)| \\ &= |(t - s)Av_2| \\ &= |(t - s)\gamma^{-2}v_2| \\ &= \gamma^{-2}|(t - s)v_2| \\ &= \gamma^{-2}d(q(y), q(z)) \end{aligned}$$

This shows the contracting nature of f on $E_{q(x)}$, since $\gamma^{-2} < 1$. The contracting nature of f^{-1} on $F_{q(x)}$ is done in a similar way. The fact that the vectors v_1, v_2 form a basis for $\mathbb{R}^2$ means that the map sending the pair $(q(x + tv_2), q(x + sv_1)$ in $E_{q(x)} \times F_{q(x)}$ to $q(x + tv_2 + sv_1$ is a homeomorphism to a neighbourhood of $q(x)$ in $\mathbb{T}^2$.

To see the bracket operation in this example, we can do no better than our original discussion. The point $[q(x), q(y)]$ is the unique point in $E_{q(x)} \cap F_{q(y)}$.

There are many generalizations of this example possible. First, let A be any $N \times N$ matrix with integer entries and determinant either 1 or -1. In exactly the same fashion as above, we can construct a map f of the N-torus, $\mathbb{T}^N$. If we assume that the matrix has no eigenvalues of absolute value 1, then we construct $E_{q(x)}$ as before, using all eigenvectors whose associated (complex) eiganvalues have absolute value less than 1. Similarly, $F_{q(x)}$ is constructed from all eigenvectors whose eigenvalues are greater than 1 in absolute value. Obviously, some care must be taken in the case ob complex eigenvalues and eigenvectors, but we leave this to the reader to sort out.

These are all examples of Anosov diffeomorphisms. Let M be a compact Riemannian manifold and let f be a diffeomorphism of M. We say that (M, f) is an Anosov diffeomorphism if we may find constants $C \geq 0$ and $0 < \lambda < 1$ and a splitting of the tangent space of M

$$TM = E^s \oplus E^U$$

into Tf-invariant sub-bundles such that, for all $n \geq 1$, we have

$$\begin{aligned} \| T(f^n)\xi \| &\leq C\lambda^n \| \xi \| \quad \text{for all } \xi \in E^s, \\ \| T(f^{-n})\eta \| &\leq C\lambda^n \| \eta \| \quad \text{for all } \eta \in E^u. \end{aligned}$$

The equations above look reminiscent of the definition of Smale space, but slightly different. This can be improved. In Exercises 6.4.1 and 6.4.2 of [**?**], it is shown that the definition givenabove is equivalent to requiring the existence is a Riemannian metric in which we have

$$\begin{aligned} \| (Tf)\xi \| &\leq \lambda \| \xi \| \quad \text{for all } \xi \in E^s, \\ \| T(f^{-1})\eta \| &\leq \lambda \| \eta \| \quad \text{for all } \eta \in E^u. \end{aligned}$$

which looks considerably more like the condition we want.

2.3 Basic sets of Axiom A Systems

We will now spend a short time discussing Smale's Axiom A systems. In a certain sense this doesn't belong in a section on examples; particularly since we won't present any explicit ones. However, this class of dynamical systems has been of great interest and, in a certain sense, be regarded as the raison d'etre for Smale spaces. We refer the reader to [**?**] for more extensive discussions.

Smale's program for differential dynamics begins with a compact manifold, M, with a diffeomorphism $f : M \to M$. We consider the set of non-wandering points, $NW(f)$. The key ingredient in the definition of Axiom is to suppose that the tangent bundle of M, when restricted to $NW(f)$ has a global splitting

$$T_{NW(f)}M = E^s \oplus E^u$$

and the same conditions hold on these spaces as for Anosov diffeomorphisms: for all $n \geq 1$, we have

$$\begin{aligned} \| T(f^n)\xi \| &\leq C\lambda^n \| \xi \| \quad \text{for all } \xi \in E^s, \\ \| T(f^{-n})\eta \| &\leq C\lambda^n \| \eta \| \quad \text{for all } \eta \in E^u. \end{aligned}$$

To say this another way, an Anosov diffeomorphism is an Axiom A system in which every point is non-wandering.

The other requirement for Axiom A systems is that the periodic points are dense in the non-wandering set. From our point of view here, this is needed to prove that the non-wandering set is actually a Smale space.

Smale also proved that the non-wandering set for an Axiom A system had a canonical decomposition into a finite number of irreducible pieces. We will not make this notion of irreducible precise at the moment (although we will in section **??**). These sets, Smale called basic sets.

Theorem 2.3.1. *If (M, f) is an Axiom A system, then $(NW(f), f \mid NW(f))$ and all the basic sets are Smale spaces.*

We will not give a proof. The essential features of a proof may be found in section 6.4 of [**?**]. (See especially 6.4.9 and 6.4.13.)

Smale proposed the class of Axiom A systems for study for several reasons. First, he believed that they should be generic in a certain sense. Second, they should display structural stability: any sufficiently small perubation of such a map should actually be topologically conjugate to the original map. It seems that they may actually coincide with the class of structurally stable maps. Finally, Smale hoped that they could be classified by relatively simple combinatorial data in the same sort of fashion that Morse-Smale systems could be described. We will not concern ourselves here with all the developments of this program, but [] is an excellent reference.

One of Smale's great insights was that, even though one began with a system which was smooth, the non-wandering set itself would not usually be a sub-manifold. The first example of this was the horse-shoe. It is a diffeomorphism of the two-sphere where the non-wandering set consists of a repelling fix-point, an attracting fix-point and an invariant Cantor set where stable and unstable sub-bundles of the tangent bundle are both one-dimensional. This phenomenon has now become very well known and the non-wandering set is very typically some sort of fractal object. This is our motivation for moving from the smooth category to the topological one.

2.4 Solenoids

We begin this section with a simple class of examples. Let K be a finite directed graph. We regard this as a topological space, with a metric d. Let $f : K \to K$ be a map satisfying the following conditions.

1. f maps vertices to vertices.
2. f is continuous.
3. f is surjective.
4. the restriction of f to each edge is locally expanding; that is, there are constants $\delta > 0$, $\lambda > 1$ such that, if x, y are on the same edge and $d(x, y) \leq \delta$, then $d(f(x), f(y)) \geq \lambda d(x, y)$.
5. f is 'flattening' at the vertices; that is, there is a constant $k \geq 1$ such that each vertex, v, has a neighbourhood V such that $f^k(V)$ is homeomorphic to an open interval with $f^k(v)$ in its interior.

Let us consider an explicit example: suppose K has one vertex v, and two edges a and b. We describe f as $a \to aab$ and $b \to ab$. By this, we mean that a is divided into three equal length subintervals. The first is mapped homeomorphically onto a (and is uniformaly stretched by 3), the second is mapped to a also and the third to b. The interval b is divided into two equal subintervals. The first is mapped to a (uniformly stretched by 2) and the second to b.

For the locally expanding axiom, the constant $\delta = \frac{1}{6}$ (notice that there are two distinct points in the interior of the a edge both mapped to v) and

$\lambda = 2$. In the flattening axiom, we may use $k = 1$. Notice that the image of a small open ball around v (which looks like a point with four 'legs' is an interval containing the start of the a edge and the end of the b edge.

Now we let X be the inverse limit of the system

$$K \xleftarrow{f} K \xleftarrow{f} K \xleftarrow{f} \cdots .$$

More explicitly, we write

$$X = \{(x_0, x_1, x_2, \ldots) \mid x_n \in K, f(x_{n+1}) = x_n, \text{ for all } n \geq 0\}.$$

We define a metric d on X by

$$d((x_0, x_1, x_2, \ldots), (y_0, y_1, y_2, \ldots)) = \sum_{n \geq 0} 2^{-n} d(x_n, y_n),$$

for all $(x_0, x_1, x_2, \ldots), (y_0, y_1, y_2, \ldots)$ in X. The map on X, also denoted by f, is defined as

$$f(x_0, x_1, x_2, \ldots) = (f(x_0), f(x_1), f(x_2), \ldots),$$

for all $(x_0, x_1, x_2, \ldots)$ in X. Notice that the inverse is given by

$$f^{-1}(x_0, x_1, x_2, \ldots) = (x_1, x_2, x_3, \ldots).$$

In this example, the local contracting sets are totally disconnected, while the local expanding sets are homeomorphic to intervals in the real line.

This example (or a small variation of it) is due to R.F. Wiliams. He also gave a more general construction where the space K is a branched manifold, of arbitrary dimension. In these examples, the local contracting set is totally disconnected, while the local exanding expanding set is homeomorphic to an open ball in Euclidean space. However, it should be possible to give some kind of definition which does not use any kind of manifold structure. The resulting space should have quite general looking local expanding sets.

2.5 Substitution Tiling Systems

We work in Euclidean space $\mathbb{R}^d$, $d \geq 1$. In fact, all our examples with be with $d = 1, 2$. For x in $\mathbb{R}^d$ and $r > 0$, we let $B(x, r)$ denote the open ball of radius r centred at x.

We assume to have a finite number of subsets $p_1, \ldots, p_N$ of $\mathbb{R}^d$. These should be homeomorphic to a closed ball, but in fact, we may even assume for the moment that each is a polyhedron. We may also allow that two of them are the same subset, but carry different labels. We call these sets the proto-tiles.

We also have a constant $\lambda > 1$ and, for each $i = 1, \ldots, N$, $\omega(p_i)$ which is a collection of subsets, each of which is a translate of one of the originals, whose interiors are pairwise disjoint and whose union is the set λp_i.

We define a tile to be any translate of of one of the proto-tiles. We extend our definition of ω by setting $\omega(p_i + x) = \omega(p_i) + \lambda x$, for any i and x in $\mathbb{R}^d$.

A partial tiling T is a collection of tiles whose interiors are pairwise disjoint. A tiling is a partial tiling whose union is $\mathbb{R}^d$. We may extend our definition of ω to collections of tiles by $\omega(T) = \cup_{t \in T} \omega(t)$. Note that this is again a partial tiling. This now also allows us to iterate ω; $\omega^k(p_i)$ makes sense for any $k \geq 1$.

If T is a partial tiling, x in is $\mathbb{R}^d$ and $r > 0$, we use a slight abuse of notation by setting

$$T \cap B(x, r) = \{t \in T \mid t \subset B(x, r)\}.$$

We define Ω to be the set of tilings T such that, for any $r > 0$, there is $k \geq 1$, $1 \leq i \leq N$ and x in $\mathbb{R}^d$ such that

$$T \cap B(0, r) \subset \omega^k(p_i) + x.$$

The first basic facts are summarized as follows.

Lemma 2.5.1. *1. Ω is non-empty.*

2. $\omega(\Omega) = \Omega$.

The next step is to introduce a metric on Ω. The idea is that two elements, T, T' are close if, after a small translation, they agree on a large ball around the origin. More precisely, $d(T, T')$ is the infimum of all $\epsilon > 0$ such that, there exist x, x' in $\mathbb{R}^d$ such that

$$(T - x) \cap B(0, \epsilon^{-1}) = (T' - x') \cap B(0, \epsilon^{-1}).$$

(If no such $\epsilon > 0$ exists, we set $d(T, T') = 1$.)

We say that Ω has *finite local complexity* or FLC if, for every $r > 0$, modulo translation, there are only finitely many collections $T \cap B(x, r)$, T

in Ω, x in $\mathbb{R}^d$. In the case that our proto-tiles are polyhedra and in the substitution ω, tiles meet full face to full face, this is automatic.

The remaining important properties are summarized below.

Lemma 2.5.2. *1. ω is continuous.*

2. Ω is compact if and only if it has finite local complexity.

3. $\omega : \Omega \to \Omega$ is injective if and only if Ω contains no periodic tilings.

Theorem 2.5.3. *Suppose the substitution tiling system ω has finite local complexity and that the space Ω contains no periodic tilings. Then (Ω, ω, d) is a Smale space.*

Proof. We choose a constant $\epsilon_\omega > 0$ such that $(2\epsilon_\omega)^{-1}$ is greater than the diameter of each prototile and such that each prototile contains an open ball of radiuus $2\epsilon_\omega$.

It is then a fairly simple matter to check that, if T and T' are two tilings in Ω with $d(T, T') \leq \epsilon_\omega$, the x and x' such that $(T - x) \cap B(0, \epsilon_\omega^{-1}) = (T' - x') \cap B(0, \epsilon_\omega^{-1})$ are unique. We then define

$$[T, T'] = T' - x' + x.$$

While the precise details are slightly more complicated, it is fairly easy to see why this makes Ω a Smale space. We give the main ideas.

First, it is quite easy to prove that, for $\epsilon > 0$ sufficiently small, we have

$$\begin{aligned} \Omega^s(T, \epsilon) &= \{T' \mid T \cap B(0, \epsilon^{-1}) = T' \cap B(0, \epsilon^{-1})\} \\ \Omega^u(T, \epsilon) &= \{T + x \mid x \in B(0, \epsilon)\}. \end{aligned}$$

Next, it is a simple matter to see that if T and T' agree on $B(0, \epsilon)$, then $\omega(T)$ and $\omega(T')$ agree on $B(0, \lambda\epsilon)$. This immediately translates into the fact that, if $[T, T'] = T'$, then

$$d(\omega(T), \omega(T')) \leq \lambda^{-1} d(T, T').$$

Finally, it is clear that $\omega(T + x) = \omega(T) + \lambda x$. Using the fact that ω is injective, we also see that $\omega^{-1}(T + x) = \omega^{-1}(T) + \lambda^{-1} x$. This translates into the fact that, if $[T, T'] = T$, then $T' = T + x$ and

$$\begin{aligned} d(\omega^{-1}(T), \omega^{-1}(T + x)) &= d(\omega^{-1}(T), \omega^{-1}(T) + \lambda^{-1} x) \\ &= \lambda^{-1} |x| \\ &= \lambda^{-1} d(T, T + x). \end{aligned}$$

□

Chapter 3

Basic theory

3.1 Stable and Unstable Equivalence

In this section, we want to introduce and investigate the notions of stable and unstable equivalence for the points of a Smale space.

Before stating the definitions, we establish the following preliminary result.

Lemma 3.1.1 (flochomeo). *Let x be in X and $0 < \delta < \epsilon_X$. If y is in $X^s(x, \epsilon_X - \delta)$, then*

$$f(X(y, \delta) \cap X^s(x, \epsilon_X))$$

is an open subset of $X^s(f(x), \epsilon_X)$ in its relative topology.

Also, if $f(y)$ is in $X^u(f(x), \epsilon_X - \delta)$, then

$$f^{-1}(X(f(x), \delta) \cap X^u(f(x), \epsilon_X))$$

is an open subset of $X^u(f(x), \epsilon_X)$ in its relative topology.

Proof. We will only consider the first statement, the other being similar. We first note that the fact that the range is contained in the given set follows from the definitions and the properties of the map f. We must show that the range is open. Begin with z in $X(y, \delta) \cap X^s(x, \epsilon_X)$. Choose $0 < \delta' < \epsilon_X$ sufficiently small so that $f^{-1}(X(f(z), \delta')) \subset X(z, \delta - d(y, z))$. It follows from the triangle inequality that

$$f^{-1}(X(f(z), \delta') \cap X^s(f(x), \epsilon_X)) \subset X(y, \delta).$$

We also note for any w in $X(f(z), \delta') \cap X^s(f(x), \epsilon_X)$, we have

$$[f^{-1}(w), x] = f^{-1}([w, f(x)]) = f^{-1}(f(x)) = x$$

and so $f^{-1}(w)$ is also in $X^s(x, \epsilon_X)$. From this we see that $X(f(z), \delta') \cap X^s(f(x), \epsilon_X)$ is in the image of $X(y, \delta) \cap X^s(x, \epsilon_X)$ and so the range is open, as desired. □

Definition 3.1.2. *Let (X, d, f) be a Smale space. We say two points x, y in X are* stably equivalent *and write $x \stackrel{s}{\sim} y$ if*

$$\lim_{n \to +\infty} d(f^n(x), f^n(y)) = 0.$$

We say that x, y are unstably equivalent *and write $x \stackrel{u}{\sim} y$ if*

$$\lim_{n \to -\infty} d(f^n(x), f^n(y)) = 0.$$

It is immediate that each of these is an equivalence relation. It is also fairly clear that $x \stackrel{s}{\sim} y$ (or $x \stackrel{u}{\sim} y$) if and only if $f(x) \stackrel{s}{\sim} f(y)$ (or $f(x) \stackrel{u}{\sim} f(y)$, respectively).

It should also be fairly clear from the definitions of the last section that if y is in $X^s(x, \epsilon_X)$, then one can show inductively that, for every positive n, $f^n(y)$ is in $X^s(f^n(x), \epsilon_X)$ and that

$$d(f^n(y), f^n(x)) \leq \lambda^n d(x, y)$$

and since $\lambda < 1$, we have $x \stackrel{s}{\sim} y$. One might even have anticipated this result from the fact that we called $X^s(x, \epsilon_X)$ the local stable set of x. In a similar way, every point in $X^u(x, \epsilon_X)$ is unstably equivalent to x.

We can take these last comments a step further by observing that if $f^n(y)$ is in $X^s(f^n(x), \epsilon_X)$ for any positive integer n, then $f^n(x) \stackrel{s}{\sim} f^n(y)$ and hence, $x \stackrel{s}{\sim} y$. We will show that this is a complete description.

Proposition 3.1.3. *Let x be in X and $0 < \epsilon \leq \epsilon_X$. The equivalence class of x under $\stackrel{s}{\sim}$ is*

$$\bigcup_{n \geq 0} f^{-n}(X^s(f^n(x), \epsilon))$$

and its equivalence class under $\stackrel{u}{\sim}$ is

$$\bigcup_{n \geq 0} f^{n}(X^s(f^{-n}(x), \epsilon)).$$

Proof. We will show only the first statement. The second is obtained in the same way. We have already argued that any point in the set given is in the stable equivalence class of x. Conversely, suppose that $y \stackrel{s}{\sim} x$. Then we may choose $N \geq 0$ sufiiciently large so that

$$d(f^n(x), f^n(y)) \leq \epsilon' = \min\{\epsilon, \epsilon_1\}$$

for all $n \geq N$, where ϵ_1 is as given in Lemma **??**. By Lemma **??**, we have $f^N(y)$ is in $X^s(f^N(x), \epsilon)$. This completes the proof. $\square$

Let us consider an example. Let G be a graph and X_G be the associated shift of finite type. Recall from the definitions that for any e and f in X_G, we have f is in $X^s(e, \epsilon_X)$ if and only if $e_n = f_n$, for all $n \geq 0$. It follows that $\sigma^N(f)$ is in $X^s(\sigma^N(e), \epsilon_X)$ if and only if $e_n = f_n$, for all $n \geq N$. So we see that $e \stackrel{s}{\sim} f$ if and only if there is some $N \geq 0$ such that $e_n = f_n$, for all $n \geq N$. This is usually refered to as *right tail equivalence*. Analogous statements are available for unstable equivalence, including the notion of left tail equivalence.

Henceforth, the $\stackrel{s}{\sim}$ equivalence class of a point x in X is denoted $X^s(x)$ while its unstable class is denoted by $X^u(x)$. These are called the stable and unstable sets of x, respectively. We could have introduced the notation ealier, but its similarity with that of the local stable and unstable sets seemed premature before we had established the last result.

Theorem 3.1.4. *Let (X, d, f) be a mixing Smale space and let x be in X. Then $X^s(x)$ and $X^u(x)$ are dense in X.*

Proof. Let δ be positive and let y be in X. We will show that $X^s(x)$ meets $X(y, \delta)$. First, we choose $\epsilon_X > \delta' > 0$ such that if $d(x', y') < \delta'$, then $d([x', y'], y') < \delta/2$. Considering $f^n(x), n \geq 0$, find a subseqence $f^{n_i}(x)$ which converges to some point x_0 in X. Let $U = X(y, \delta/2)$ and $V = X(x_0, \delta'/2)$ and apply the definition of mixing to find a positive integer N. Then find $n_i \geq N$ such that $f^{n_i}(x)$ is in $X(x_0, \delta'/2)$. From the choice of N, there is z in U with $f^{n_i}(z)$ in V. This means that $f^{n_i}(z)$ and $f^{n_i}(x)$ are in $X(x_0, \delta'/2)$. It follows that $w = f^{-n_i}[f^{n_i}(x), f^{n_i}(z)]$ is well-defined. More over, $f^{n_i}(w)$ is in $X^u(f^{n_i}(z), \delta/2)$. It follows that $d(w, z) \leq \lambda^{n_i}\delta/2$ and from this that w is in $X(y, \delta)$. On the other hand, $f^{n_i}(w)$ is in $X^s(f^{n_i}(x), \delta/2)$ and from this it follows that w is in $X^s(x)$. This completes the proof. $\square$

For a given x in X, the set $X^s(x)$ is a subset of X and has a relative topology from X. However, there is another much more natural topology on it. To see this, we will use the description of $X^s(x)$ given in the last result. Notice that for a given n, we have

$$f^{-n}(X^s(f^n(x), \epsilon)) \subset f^{-n-1}(X^s(f^{n+1}(x), \epsilon)).$$

To see this, let y be in the former set. This means $f^n(y)$ is in $X^s(f^n(x), \epsilon)$. It follows from the definitions that $f^{n+1}(y)$ is in $X^s(f^{n+1}(x), \epsilon)$ and this implies the result.

Now each set, $f^{-n}(X^s(f^n(x), \epsilon))$ is given the relative topology of X and the set $X^s(x)$ is given the inductive limit topology. A subset, U, is open if and only if its intersection with $f^{-n}(X^s(f^n(x), \epsilon))$ is open, for all but finitely many n. The unstable set is given a topology in a similar fashion.

Just to get an idea of what is happening here, let us consider the first example of the Anosov diffeomorphism given in 2.2. Let $x = q(0,0)$. The reason this is a nice choice is that it is fixed and that simplifies our computation of the stable set of x. The local stable set is

$$X^s(x,\epsilon) = \{q(tv_2) \mid |t| \leq \epsilon\}$$

where q is the quotient map from $\mathbb{R}^2$ to $\mathbb{T}^2$ and $v_2 = (-1,\gamma)$ is the contracting eigenvector for A. It is easy to check then that

$$f^{-n}(X^s(x,\epsilon)) = \{q(tv_2) \mid |t| \leq \gamma^{2n}\epsilon\}$$

and that

$$X^s(x) = \{q(tv_2) \mid t \in \mathbb{R}\}.$$

Moreover, this set is dense in T^2 and rather horrid in the relative topology. In particular, it is not locally compact. However, the inductive limit topology makes the map from $t \in \mathbb{R}$ to $q(tv_2)$ a homeomorphism.

Theorem 3.1.5. *Let x be a point in the Smale space (X,d,f).*

1. *The sets $X^s(x)$ and $X^u(x)$, endowed with the inductive limit topology above, are locally compact and Hausdorff.*

2. *A sequence y_n converges to y in $X^s(x)$ if and only if it converges to y in the usual topology of X and $[y_n,y] = y$, for all n sufficiently large.*

3. *A sequence y_n converges to y in $X^u(x)$ if and only if it converges to y in the usual topology of X and $[y,y_n] = y$, for all n sufficiently large.*

4. *Sets of the form $X^s(y,\epsilon)$, where y is in $X^s(x)$ and $0 < \epsilon \leq \epsilon_X$, form a neighbourhood base for the inductive limit topology on $X^s(x)$.*

5. *Sets of the form $X^u(y,\epsilon)$, where y is in $X^u(x)$ and $0 < \epsilon \leq \epsilon_X$, form a neighbourhood base for the inductive limit topology on $X^u(x)$.*

Proof. That $X^s(x)$ is Hausdorff follows from the fact that each set in the inductive limit is Hausdorff. We consider the fact that it is locally compact. Let y be a point of $X^s(x)$, which means that $f^N(y)$ is in $X^s(f^N(x),\epsilon_X)$, for some $N \geq 0$. By replacing N by $N+1$ if necessary, we may assume that $d(f^N(y), f^N(x)) < \epsilon_X$. Choose $0 < \delta < \epsilon_X - d(f^N(y), f^N(x))$ and let $U = f^{-N}(X^s(f^N(x), \delta))$. This set is equal to

$$f^{-N}(X(f^N(x), \delta)) \cap f^{-N}(X^s(f^N(x), \epsilon_X))$$

and, hence, open in the relative topology of $f^{-N}(X^s(f^N(x), \epsilon_X))$. We want to show the same is true for each of the sets $U \subset f^{-n}(X^s(f^n(x), \epsilon_X))$, for every $n \geq N$. For each such n, $f^n(U)$ is open in the relative topology of $X^s(f^n(x), \epsilon_X)$, from Lemma **??**. The desired conclusion follows since f is a homeomorphism. So the set U is open in the inductice limit topology of $X^s(x)$. Its closure is compact in $f^{-N}(X^s(f^N(x), \epsilon_X))$ and hence in $X^s(x)$. This completes the proof of the first part.

For the second statement, suppose that y_n converges to y in the relative topology. Repeating the argument in the first part, for some $N \geq 0$, y is in $f^{-N}(X^s(f^N(x), \epsilon_X))$ and for some $0 < \delta$, the set

$$U = f^{-N}(X^s(f^N(x), \delta))$$

is a neighbourhood of y in $X^s(x)$. So for all n sufficiently large, y_n is in U. Also for all n sufficiently large, y_n is sufficiently close to y so that $d(f^{-k}(y_n), f^{-k}(y)) < \epsilon_X$, for all $0 \leq k \leq N$. Then we have

$$\begin{aligned} [y_n, y] &= f^{-N}[f^N(y_n), f^N(y)] \\ &= f^{-N}(f^N(y)) \\ &= y \end{aligned}$$

because both $f^N(y_n)$ and $f^N(y)$ are in $X^s(f^N(x), \epsilon_X)$.

Now suppose that y_n converges to y and $[y_n, y] = y$, for all n sufficiently large. Then for some N, $f^N(y)$ is in $X^s(f^N(x), \epsilon_X)$ and, again replacing N by $N + 1$ if necessary, we may assume that $d(f^N(y), f^N(x)) < \epsilon_X$. Then for n sufficiently large, we have $d(f^N(y_n), f^N(x)) \leq \epsilon_X$ and we can compute

$$\begin{aligned} [f^N(y_n), f^N(x)] &= [[f^N(y_n), f^N(y)], f^N(x)] \\ &= [f^N[y_n, y], f^N(x)] \\ &= [f^N(y), f^N(x)] \\ &= f^N(x) \end{aligned}$$

This means that $f^N(y_n)$ is converging to $f^N(y)$ in the relative topology of $X^s(f^N(x), \epsilon_X)$ and, hence, y_n is converging to y in the relative topology of $f^{-N}(X^s(f^N(x), \epsilon_X))$. Now we note that the inclusion of this set in the inductive limit is continuous and the desired conclusion follows. $\square$

There is one more equivalence relation on the points of X which is quite important. It is just the intersection of stable and unstable equivalence, but it has a number of very nice features which we will exploit.

Definition 3.1.6. *Two points x and y in X are* homoclinic *if they are both stably and unstably equivalent. That is, we have*

$$\lim_{|n|\to\infty} d(f^n(x), f^n(y)) = 0.$$

In this case, we write $x \stackrel{h}{\sim} y$. Here, we denote the equivalence class of x by $X^h(x)$.

It is worth considering an example at this point. If we once again think about our shift of finite type X_G, we see that $e \stackrel{h}{\sim} f$ if and only if $e_n = f_n$ for all but finitely many n.

3.2 Shadowing

In this section we discuss a critical property of Smale spaces called shadowing.

First of all, if a is in $\mathbb{Z} \cup \{-\infty\}$ and b is in $\mathbb{Z} \cup \{\infty\}$, then we say that $I = (a, b) = \{n \in \mathbb{Z} \mid a < n < b\}$ is an interval in $\mathbb{Z}$.

We begin with a pair of definitions.

Definition 3.2.1. *Let (X, d) be a compact metric space and let $f : X \to X$ be a homeomorphism. For any $\epsilon > 0$, an* ϵ-pseudo-orbit *over a non-empty interval, I, is a collection of points x_n, for each n in I, such that*

$$d(f(x_n), x_{n+1}) \leq \epsilon$$

provided n and $n+1$ are in I.

Observe first that if x is in X, then for any I, the points $x_n = f^n(x)$ are an ϵ-pseudo-orbit, for any positive ϵ. That is, orbits are pseudo-orbits.

Definition 3.2.2. *Let $\epsilon > 0$ and $\delta > 0$. If x_n and y_n are ϵ-pseudo-orbits over the same interval I, then we say that one* δ-shadows *the other if*

$$d(x_n, y_n) \leq \delta$$

for all n in I. If x is in X, we also say that x_n is δ-shadowed *by (the orbit of) x, if*

$$d(x_n, f^n(x)) \leq \delta$$

for all n in I.

Our objective is to prove the following result.

Theorem 3.2.3. *Let (X, d, f) be a Smale space. For any $\delta > 0$, there is an $\epsilon > 0$ such that every ϵ-pseudo-orbit in X is δ-shadowed by an orbit of X.*

We will need the following result in the proof.

Lemma 3.2.4. *Suppose that $0 < \delta_1 \leq \epsilon_X$. Then there is $\epsilon > 0$ such that, if $d(f(x), x') < \epsilon$, then for all z in $X^s(x, \delta_1)$, $d(x', f(z)) \leq \epsilon_X$ and $[x', f(z)]$ is in $X^s(x', \delta_1)$.*

Proof. First, it is clear that $[x', f(z)]$ is well-defined and in the local stable set of x'. We must find ϵ so that $d(x', [x', f(z)]) \leq \delta_1$.

Consider the set

$$A = \{(x, y, z) \mid d(x, y), d(y, z) \leq \epsilon_X/2, [y, z] = z\}$$

which is compact in $X \times X \times X$. Consider also the function h defined on A by

$$h(x, y, z) = d(x, [x, z]) - d(y, z)$$

which is clearly continuous and hence uniformly continuous. On the set,

$$B = \{(x, y, z) \in A \mid x = y\}$$

which is compact, we have

$$\begin{aligned} h(x, y, z) &= h(y, y, z) \\ &= d(y, [y, z]) - d(y, z) \\ &= d(y, z) - d(y, z) \\ &= 0. \end{aligned}$$

Therefore, there is $\epsilon > 0$ such that, if $d(x, y) < \epsilon$ and $(x, y, z) \in A$, then

$$|h(x, y, z)| < \delta_1(1 - \lambda).$$

Also choose ϵ sufficiently small so that $\epsilon < (1-\lambda)\delta_1$.

Now consider x, x', z as in the statement. First we have $d(x, z) < \epsilon$ and hence

$$\begin{aligned} d(x', f(z)) &\leq d(x', f(x)) + d(f(x), f(z)) \\ &\leq \epsilon + \lambda d(x, z) \\ &\leq (1-\lambda)\delta_1 + \lambda\delta_1 \\ &\leq \delta_1 \\ &\leq \epsilon_X \end{aligned}$$

Also, we have $(x', f(x), f(z))$ is in the set A and $d(x', f(x)) < \epsilon$ and so we conclude that

$$\begin{aligned} d(x', [f(x), f(z)]) &\leq h(x', f(x), f(z)) + d(f(x), f(z)) \\ &\leq \delta_1(1-\lambda) + \lambda d(x, z) \\ &\leq \delta_1(1-\lambda) + \lambda\delta_1 \\ &= \delta_1 \end{aligned}$$

This completes the proof. □

Proof of Theorem 3.2.3. First we choose $0 < \delta_1 \leq \epsilon_X/2$ such that

$$[X^u(x, \delta_1), X^s(x, \delta_1)] \subset X(x, \delta),$$

for all x in X. Next, we choose $\epsilon > 0$ as in Lemma **??** which holds for both (X, d, f) and (X, d, f^{-1}).

We will first show the conclusion holds in the case where $I = (a, b)$ is finite and $a < 0$ and $b > 0$. For each $a < i < b-1$, we define a map $g_i : X^s(x_i, \delta_1) \to X^s(x_{i+1}, \delta_1)$ by $g_i(z) = [x_{i+1}, f(z)]$. The fact that g_i is well-defined follows from the conclusions of the last Lemma.

In an analogous fashion, we may define a map $h_i : X^u(x_i, \delta_1) \to X^u(x_{i-1}, \delta_1)$ by $h_i(z) = [f^{-1}(z), x_{i-1}]$, for $a+1 < i < b$. We let g and h be the union of the functions g_i, $a < i < b-1$, and h_i, $a+1 < i < b$, respectively.

We define sets

$$S_i = \left[h^{b-1+i}(X^u(x_{b-1}, \delta_1), g^{1-a+i}(X^s(x_{a+1}, \delta)\right],$$

for every $a < i < b$, and we claim that any point x in S_0 will shadow the pseudo-orbit. It is clear that

$$\begin{aligned} h^{b-1+i}(X^u(x_{b-1}, \delta_1) &\subset X^u(x_i, \delta_1) \\ g^{1-a+i}(X^s(x_{a+1}, \delta_1) &\subset X^s(x_i, \delta_1) \end{aligned}$$

and then from the choice of δ_1, we have $S_i \subset X(x_i, \delta)$. Then it suffices for us to show that $f(S_i) = S_{i+1}$.

Choose y in $X^s(x_{a+1}, \delta_1)$ and z in $X^u(x_{b-1}, \delta_1)$. Let $y' = g^{i-a-1}(y)$ and $z' = h^{i-b}(z)$ so that we have $[h(z'), y'] \in S_i$ and $[z', g(y')] \in S_{i+1}$. Any element of S_i can be obtained in this way. We will show that f carries the former to the latter. In the following computation, one must verify that all bracket operations are defined. We leave this tedious aspect of the proof to the reader. We have

$$\begin{aligned} f([h(z'), y']) &= f(\left[\left[f^{-1}(z'), x_i\right], y'\right]) \\ &= \left[f(\left[f^{-1}(z'), x_i\right]), f(y')\right] \\ &= [[z', f(x_i)]), f(y')] \\ &= [z', f(y')] \\ &= [z', [x_{i+1}, f(y')]] \\ &= [z', g(y')] \end{aligned}$$

This completes the proof.

We now address the problem when the interval is infinite. In fact, we consider the case only for $I = \mathbb{Z}$ and leave the half-open cases for the reader.

We begin by considering $I_n = (-n, n)$, for any positive integer n. Notice that the choice of δ_1 is independent of n. This also means that every point of

$$S^{(n)} = \left[h^{n-1}(X^u(x_{n-1}, \delta_1)), g^{n-1}(X^s(x_{-n+1}, \delta_1))\right]$$

will δ-shadow the pseudo-orbit over the interval I_n. Each set $S^{(n)}$ is closed and it follows directly from the definitions that $S^{(n)} \supset S^{(n+1)}$ for all n. The intersection of all $S^{(n)}$ is non-empty and any point in this intersection will δ-shadow the pseudo-orbit over all of $\mathbb{Z}$. This completes the proof. □

At this point we note the following corollary.

Corollary 3.2.5. *Let (X, d, f) be a Smale space. Then the set of periodic points for f, $Per(X, f)$, is dense in $NW(X, f)$. In particular, if X is non-wandering, then $Per(X, f)$ is dense in X.*

Proof. Let x_0 be a non-wandering point of X and let ϵ_0 be positive. Let $\delta = \epsilon_1/2$, where ϵ_1 is the expansiveness constant for (X, d, f) **??**. Choose $\epsilon_2 > 0$ so that every ϵ_2-pseudo-orbit is δ-shadowed by an orbit. Let ϵ be the minimum of ϵ_0, δ and ϵ_2 and let $V = X(x_0, \epsilon)$ Since x_0 is non-wandering, there is a positive integer n and a point x in V, with $f^n(x)$ also in V. Now we define $x_{in+j} = f^j(x)$, for any i in $\mathbb{Z}$ and $0 \leq j < n$. It is easy to verify that this is a ϵ-pseudo-orbit over $\mathbb{Z}$. Then we may find a point y whose orbit δ-shadows x_n. In particular, y is in $X(x, \delta)$ and hence in $X(x_0, \epsilon_0)$. We claim that y is periodic. To see this we note that, for any integer i, $x_{n+i} = x_i$ and so we have

$$\begin{aligned} d(f^i(y), f^{i+n}(y)) &\leq d(f^i(y), x_i) + d(x_{i+n}, f^{i+n}(y)) \\ &\leq \delta + \delta \\ &\leq \epsilon_1. \end{aligned}$$

Applying the expansiveness condition **??** to the points y and $f^n(y)$, we see that are equal and hence, y is periodic. □

3.3 Decomposition of Smale spaces

Smale spaces admit very simple descriptions of how they decompose into 'irreducible' pieces (meant in the non-technical sense. These fall into two classes: the first describes how the non-wandering set is decomposes into basic pieces, while the second describes how the wandering set fits around them.

3.3.1 Decomposition of the non-wandering set

We have introduced the notions of non-wandering, irreducibility and mixing and we have noticed that mixing implies irreducibity, which implies non-wandering. These implications hold in generality and the converses do not.

We now want to restrict our consideration to the case of Smale spaces. The converse directions are still false. For example, the finite disjoint union of irreducible Smale spaces is still non-wandering but no longer irreducible. The remarkable fact which we will prove is that every non-wandering Smale space arises in exactly this way; that is, it may be decomposed into a finite number of irreducible components.

Theorem 3.3.1. *Let (X, d, f) be a non-wandering Smale space. Then there are open, closed, pairwise disjoint, f-invariant subsets $X_1, \ldots, X_n$ of X, whose union is X, and so that $(X_i, d, f|X_i)$ is irreducible, for each $1 \leq i \leq n$. Moreover, these sets are unique up to relabelling.*

Proof. We define an equivalence relation $\sim$ on the periodic points of f in X as follows. Let x and y be periodic points in X. We say $x \sim y$ if, $X^s(x) \cap X^u(y)$ and $X^u(x) \cap X^s(y)$ are non-empty. First, we must show that this is an equivalence relation. In the case $x = y$, then x is in both sets, so the relation is reflexive. It is clearly symmetric. Finally we suppose that $x \sim y$ and $y \sim z$. We will prove that $X^u(x) \cap X^s(z)$ is non-empty. The other set, $X^s(x) \cap X^u(z)$, is done in a similar way and so $x \sim z$. Let p be the product of the periods of x, y and z. Let u be any point in $X^u(x) \cap X^s(y)$ and let w be any point of $X^u(y) \cap X^s(z)$. Then as n tends to plus infinity, the sequences $f^{np}(u)$ and $f^{-np}(w)$ both tend to y Choose n suficiently large so that $d(f^{np}(u), f^{-np}(w))$ is less than ϵ_X. Let $v = [f^{-np}(w), f^{np}(u)]$. Now the point v is stably equivalent to $f^{-np}(w)$ which is stably equivalent to $f^{-np}(z)$ which is just z, by the choice of p. Similarly, v is unstably equivalent to x and this completes the proof.

Next, we want to observe that if x and y are periodic points with $d(x, y) \leq \epsilon_X$, then $x \sim y$. This is because $[x, y]$ and $[y, x]$ are in the sets $X^s(x) \cap X^u(y)$ and $X^u(x) \cap X^s(y)$, respectively.

Let x be a periodic point and consider the closure of its equivalence class in X. This is certainly closed and, by the result of the last paragraph, it is also open. Moreover, any point, x, is the limit of a sequence of periodic points. Again by the result of the last paragraph, these will eventually all lie in the same equivalence class and so x will be in its closure.

This means that these closures of equivalence classes cover all of X. They are all closed and open. For any point x, all periodic points within $\epsilon_X/2$ of x must all be equivalent and so x will lie in at most one of these sets. That is, these sets form a partition of X into clopen sets. As X is compact, this partition must be finite.

It is also easily seen that for any periodic points x, y, $x \sim y$ if and only if $f(x) \sim f(y)$. This implies that f permutes the elements of the partition. We let $X_1, \ldots, X_n$ denote the orbits of the elements of the partition under f. That is, we choose an element of our partition, Z. Since f permutes the partition, there is a positive integer m so that

$$X_1 = \cup_{i \in \mathbb{Z}} f^i(Z) = \cup_{i=0}^{m-1} f^i(Z).$$

If this is not all of X, then we choose another element of our partition and repeat to get X_2. We continue in this way until X is covered.

It remains only to show that the restriction of f to X_i is irreducible. We will do this for X_1 using Z and m as above. Let U and V be non-empty open sets in X_1. Then for some i, j, $U \cap f^i(Z)$ and $V \cap f^j(Z)$ are nonempty. Now $f^{j-i}(U) \cap f^j(Z)$ and $V \cap f^j(Z)$ are both non-empty open sets in $f^j(Z)$, which is one on the elements of our partition of X. We may find periodic points x in the former and y in the latter. By the construction of the elements of the partition, x and y must be limits of periodic points in the same equivalence class. Again using our continuity result, this means that $x \sim y$. Choose a point z in $X^u(x) \cap X^s(y)$ Let p be the product of the periods of x and y. Then as n goes to $+\infty$, $f^{np}(z)$ tends to y which is in V. So find $n > i - j$ sufficiently large so that $f^{np}(z)$ is in V. Similarly, we may find $l < 0$, so that $f^{lp}(z)$ is in $f^{j-i}(U)$. This means that $f^{np}(z)$ is in V and in $f^{j-i+(n-l)p}(U)$. This proves that $f|X_1$ is irreducible.

As for the uniqueness of this decomposition, it suffices to show that if W is any clopen f-invariant subset of X with $f|W$ irreducible, then W is equal to one of the X_i. Given such a set W, let $1 \leq i \leq N$. The sets $U = W \cap X_i$ and $V = (X - W) \cap X_i$ are clopen and f-invariant. If both were non-empty, this would contradict the ireducibilty of $f|X_i$. So X_i is either disjoint from W or a subset of it. In the latter case, we consider the sets $W \cap X_i$ and $W \cap (X - X_i)$ which are clopen and f-invariant. If second is also non-empty, this would contradict the irreducibilty of $f|W$. We conclude that $W = X_i$ if $W \cap X_i$ is non-empty. This completes the proof. □

We have now seen that every non-wandering Smale space can be decomposed into irreducible pieces in a very nice fashion. The second result concerns the difference between the conditions of irreducibilty and mixing. The latter implies the former and the converse is still false for Smale spaces. As an example, take any mixing Smale space X. Let N be a positive integer and consider $X \times \{1, \ldots, N\}$ with the homeomorphism $f \times \sigma$, where σ is a cyclic permutation of $\{1, \ldots, N\}$. This new system is still irreducible. However, if U and V are non-empty open subsets of $X \times \{1\}$, then $f^n(U) \cap V$ is non-empty only if n is a multiple of N and this means that our new system is not mixing (if $N > 1$). We will next show that this type of situation only way irreducible, non-mixing Smale spaces can arise. More precisely, we have the following result.

Theorem 3.3.2. *Let (X, d, f) be an irreducible Smale space. Then there are open, closed, pairwise disjoint sets $X_1, X_2, \ldots, X_N$ whose union is X. These sets are cyclicly permuted by f and $f^N|X_i$ is mixing for every $1 \leq i \leq N$.*

Proof. In fact, we have already done most of the work already in the proof of the last result. We proceed exactly as before arriving at the point where we have partitioned X into the closures of the $\sim$-equivalence classes. This is a finite partition of X into clopen sets. In this situation, we denote them as $X_1, X_2, \ldots, X_N$.

We must show that f permutes them cyclicly. If not, the union of some finite subcollection is f-invariant, as is its compliment. If we let U denote this union and V be its compliment, then $f^n(U) \cap V$ is empty for every n and so X is not irreducible.

Next, we must show that $f^N|X_i$ is mixing, for any fixed i. We consider the case $i = 1$ only. To see this, let U and V be non-empty open sets in X_1. Choose periodic points x and y in U and V, respectively. Let p be the product of their periods. It is clear that N divides p; let $q = p/N$. For each $0 \leq i < q$, the point $y_i = f^{iN}(y)$ is a periodic point in X_1. It follows, just as in the last proof, that x is equivalent to each of these. So we may find $z_i \in X^u(x) \cap X^s(y_i)$. Because x and y are fixed by f^p, the sequence $f^{np}(z_i)$ converges to x as n tends to minus infinity and to y_i as n tends to plus infinity. Note that $f^{iN}(V)$ is a neighbourhood of y_i, for each $0 \leq i < q$. Therefore, we may find $m_i < 0 < n_i$ such that $f^{mp}(z_i) \in U$ for $m \leq m_i$ and $f^{np}(z_i) \in f^{iN}(V)$, for all $n \geq n_i$.

We let

$$K = \max\{n_i q \mid 0 \leq i < q\} - \min\{m_i q \mid 0 \leq i < q\}.$$

and we claim that $(f^N)^k(U) \cap V$ is not empty for $k \geq K$. It follows that (X_1, f^N) is mixing as desired. We may write $k = lq - i$, where $0 \leq i < q$. If $k \geq K$, then we may write $l = n - m$, where $n \geq n_i$ and $m \leq m_i$ for all $0 \leq i < q$. This means that $z = f^{mqN}(z_i) = f^{mp}(z_i) \in U$. We also have

$$\begin{aligned} (f^N)^k(z) &= (f^N)^k(f^{nqN}(z_i)) \\ &= f^{-iN}(f^{np}(z_i)) \\ &\in f^{-iN}(f^{iN}(V)) \\ &= V. \end{aligned}$$

This completes the proof.

□

3.3.2 Decomposition of the wandering set

Before beginning, a few words of warning are in order regarding Smale spaces with wandering points.

Smale's key idea for Axiom A systems was that the non-wandering set should have a hyperbolic structure and it was this idea that Ruelle interpreted in purely topological terms to give the definition of Smale space. Somewhat curiously, Ruelle did *not* require that every point in a Smale space be non-wandering.

Let us consider a couple of examples to clarify the point. As we showed in **??**, every shift of finite type, is a Smale space and may contain wandering points: consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The shift space Σ_A is homeomorphic to the two-point compactification of the integers $\mathbb{Z}$, $\{-\infty, \ldots, -2, -1, 0, 1, 2, \ldots, +\infty\}$ with the map $\sigma(n) = n - 1$. Here the non-wandering set is $\{-\infty, +\infty\}$.

On the other hand, Smale's horseshoe (see []) is an Axiom A system, but it is *not* a Smale space. The point being that the hyperbolic structure does not exist on the whole manifold, but only on the non-wandering set. The restriction of the map to its non-wandering set, *is* a Smale space.

A simpler example than the horseshoe may be obtained by considering $X = [0, 1]$ (which is not a manifold) with the map $f(x) = x^2$. The non-wandering set is $\{0, 1\}$. Again, (X, f) is not a Smale space as, for any $x \neq 0, 1$, every point in $(0, 1)$ is stably and unstably equivalent to x. This situation is curious when compared to the shift of finite type above. They are remarkably similar dynamically, yet one is a Smale space and the other is not. The resolution of this paradox is the gap that exists between 0 and 1. One may check from the definitions given in **??** that in Σ_A, we have $[x, y] = y$ for all $x, y \leq 0$ while $[x, y] = x$ for all $x, y \geq 1$ and is not defined otherwise.

There is one final subtlety to mention. It is not too difficult to produce an example of a dynamical system (X, f) is such that $NW(f)$ has wandering points: meaning that a point x may be non-wandering in X and hence lie in $NW(f)$, but may be wandering when considered as an element of $NW(f)$. in other terms, we have $NW(f \mid NW(f)) \neq NW(f)$. This seems odd at first, but the idea is that, for a given x with neighbourhood U, the points which lie in $f^n(U) \cap U$ may themselves be wandering.

In [], Dankner gives an example of such a system where X is a compact manifold, f is a diffeomorphism and $NW(f)$ has a hyperbolic structure. This phenomenon is prevented in an Axiom A system by Smale's condition that the periodic points be dense in the non-wandering set. It is this condition (only) that Dankner's example fails to satisfy. Just like the horseshoe, Dankner's example is not a Smale space, but its restriction to its non-wandering set is a Smale space (having wandering points).

We now describe our decomposition for the wandering set of a Smale space (X, d, f). First, let $X_1, \ldots, X_n$ be the irreducible components of X given in Theorem **??**.

Lemma 3.3.3. *Let x be a wandering point in the Smale space. The set of accumulation points of $\{f^n(x) \mid n \geq 1\}$ is contained in one of the irreducible components of (X, f) as given in Theorem* **??**. *Similarly, set of accumulation points of $\{f^n(x) \mid n \leq 0\}$ is contained in one of the irreducible components of (X, f). Moreover, these two components are distinct.*

Proof. Suppose that $\{f^n(x) \mid n \geq 1\}$ has accumulation points x_1 and x_2 in X_j. Let $\epsilon_1 > 0$ any expansive constant for (X, f). Choose $\epsilon_1 > \epsilon_2 > 0$ such that every ϵ_2-pseudo-orbit in X is $\epsilon_1/2$-shadowed by an orbit. It follows from the hypotheses that we may find positive integers $k < l < m$ such that

$$d(f^k(x), x_1), d(f^l(x), x_2), d(f^m(x), x_1) < \epsilon_2/2.$$

Thus the bi-infinite sequence obtained by repeating $f^k(x), f^{k+1}(x), f^{m-1}(x)$ is an ϵ_2-pseudo-orbit and hence is $\epsilon_1/2$-shadowed by the orbit of some point x'. Then $f^{m-k}(x')$ also $\epsilon_1/2$-shadows this pseudo-orbit, we see that $d(f^{m-k+i}(x'), f^i(x')) \leq \epsilon_1$, for all integers i. By expansiveness, we conclude that x' is periodic (of period $m - k$). We also note that

$$d(x', x_1) \leq d(x', f^k(x)) + d(f^k(x), x_1) \leq \epsilon_1/2 + \epsilon_2/2 < \epsilon_1,$$

and

$$d(f^{l-k}(x'), x_2) \leq d(f^{l-k}(x'), f^l(x)) + d(f^l(x), x_2) \leq \epsilon_1/2 + \epsilon_2/2 < \epsilon_1,$$

Choosing a sequence of values of ϵ_1 approaching zero, for each, we have a periodic point whose orbit lies within ϵ_1 of x_1 and x_2. Of course, each orbit must lie in a single irreducible component of X and by passing to a subsequence we may assume all orbits lie in the same irreducble component. It follows that x_1 and x_2 are in the same component.

A similar argument deals with accumulation points of $\{f^n(x) \mid n \leq 0\}$. Now suppose that x_1 is an accumulation point of $\{f^n(x) \mid n \geq 1\}$ x_2 is an accumulation point of $\{f^n(x) \mid n \leq 0\}$ and assume that these points are in the same irreducible component of X. Let $\delta > 0$. Find $\epsilon > 0$ such that any ϵ-pseudo-orbit is δ-shadowed by an orbit. Next, find integers $k \geq 1$ and $l \leq 0$ such that

$$d(f^k(x), x_1), d(f^l(x), x_2) < \epsilon/2.$$

As x_1 and x_2 are in the same irreducible component of X, we may find x' in X and a positive integer n with

$$d(x', x_1) < \epsilon/2, d(f^n(x'), x_2) < \epsilon/2.$$

Then the sequence $x, f(x), \ldots, f^{k-1}(x), x', f(x'), \ldots, f^{n-1}(x'), f^l(x), f^{l+1}(x), \ldots, x$ is an ϵ-pseudo-orbit and is δ-shadowed by an orbit. This means we can find a point in $X(x, \delta) \cap f^{k+n+l}(X(x, \delta))$. As δ was arbitrary, we conclude that x is non-wandering, which is a contradiction.

□

Definition 3.3.4. *Let X_i and X_j be irreducible components of the Smale space (X, f).*

1. *Let $W(X_i, X_j)$ denote the set of wandering points x in X such that the set of accumulation points of $\{f^n(x) \mid n \leq 0$ is contained in X_i while the set of accumulation points of $\{f^n(x) \mid n \geq 1$ is contained in X_j*

2. *We define $X_i \prec X_j$ if $W(X_i, X_j)$ is non-empty. We also define $X_i \preceq X_j$ if either $X_i \prec X_j$ or $X_i = X_j$.*

Theorem 3.3.5. *The relation $\preceq$ is a partial order on the set of irreducible components of a Smale space (X, f). Moreover, the wandering set of X may be written as a disjoint union*

$$X \setminus NW(f) = \cap_{X_i \prec X_j} W(X_i, X_j).$$

Proof. □

Chapter 4

Maps

4.1 Introduction

In this section, we give some basic definitions of maps and factor maps between dynamical systems and establish basic properties of them, particularly in the case that both domain and range are Smale spaces.

Definition 4.1.1. *Let (X, f) and (Y, g) be dynamical systems. A* map $\pi : (Y, g) \to (X, f)$, *is a continuous function $\pi : Y \to X$ such that $\pi \circ g = f \circ \pi$. A* factor map *from (Y, g) to (X, f) is a map for which $\pi : Y \to X$ is also surjective.*

As a very simple example, suppose that G and H are two graphs. A graph homomorphism $\theta : H \to G$ consists of two maps $\theta^0 : H^0 \to G^0$ and $\theta^1 : H^1 \to G^1$ such that $t_G \circ \theta^1 = \theta^0 \circ t_H$ and $i_G \circ \theta^1 = \theta^0 \circ i_H$. In an obvious way, θ induces a map from $\theta : (\Sigma_H, \sigma) \to (\Sigma_G, \sigma)$.

Definition 4.1.2. *Let (Y, g) and (X, f) be dynamical systems and let*

$$\pi : (Y, g) \to (X, f)$$

be a map. We say that π is finite-to-one *if there exists a positive integer M such that*

$$\#\pi^{-1}\{x\} \leq M,$$

for all x in X.

We note the following easy result.

Theorem 4.1.3. *Let (Y, g) and (X, f) be dynamical systems and let*

$$\pi : (Y, g) \to (X, f)$$

be a map. If y is a periodic point of g, then $\pi(y)$ is a periodic point of f. Moreover, if π is finite-to-one, the converse also holds.

Proof. From $\pi \circ g = f \circ \pi$, it follows that $\pi \circ g^n = f^n \circ \pi$, for every integer n and that π maps orbits onto orbits. Then we observe that a point is periodic if and only if its orbit is finite. Both conclusions follow. □

We now want to state a couple of technical results which will be useful later. The first is purely topological.

Lemma 4.1.4. *Let $\pi : Y \to X$ be a continuous map and let x_0 be in X with $\pi^{-1}\{x_0\} = \{y_1, y_2, \ldots, y_N\}$ finite. For any $\epsilon > 0$, there exists $\delta > 0$ such that $\pi^{-1}(X(x_0, \delta)) \subset \cup_{n=1}^{N} Y(y_n, \epsilon)$.*

Proof. If there is no such δ, we may construct a sequence $x^k, k \geq 1$ in X converging to x_0 and a sequence $y^k, k \geq 1$ with $\pi(y^k) = x^k$ and y^k not in $\cup_{n=1}^{N} Y(y_n, \epsilon)$. Passing to a convergent subsequence of the y^k, let y be the limit point. Then we know that y is not in $\cup_{n=1}^{N} Y(y_n, \epsilon)$ since that set is open, while $\pi(y) = \lim_k \pi(y^k) = \lim_k x^k = x_0$. This is a contradiction to $\pi^{-1}\{x_0\} = \{y_1, y_2, \ldots, y_N\}$. □

Lemma 4.1.5. *Let $\pi : (Y, g) \to (X, f)$ be a finite-to-one factor map between Smale spaces and suppose that (X, f) is non-wandering. There exists a periodic point x in X such that*

$$\#\pi^{-1}\{x\} = \min\{\#\pi^{-1}\{x'\} \mid x' \in X\}.$$

Proof. Choose x_0 in X which minimizes $\#\pi^{-1}\{x_0\}$ and let $\pi^{-1}\{x_0\} = \{y_1, \ldots, y_N\}$. Choose $\epsilon_Y/2 > \epsilon > 0$ so that the sets $Y(y_n, \epsilon), 1 \leq n \leq N$ are pairwise disjoint. Apply the last Lemma to find δ satisfying the conclusion there for this ϵ.

As (X, f) is non-wandering, the periodic points are dense, so choose a periodic point x in $X(x, \delta)$. We claim that, for any $1 \leq n \leq N$, the set $\pi^{-1}\{x\} \cap Y(y_n, \epsilon)$ contains at most one point. Note that from this, it follows at once that $\#\pi^{-1}\{x\} \leq N = \#\pi^{-1}\{x_0\}$. The reverse inequality is trivial from the choice of x_0 and the proof will be complete. Suppose that both y and y' are in $\pi^{-1}\{x\} \cap Y(y_n, \epsilon)$. As they both map to x under π, they are both periodic. As $\epsilon < \epsilon_Y/2$, we may bracket y and y'. We have

$$\pi[y, y'] = [\pi(y), \pi(y')] = [x, x] = x.$$

This implies that $[y, y']$ is also periodic. The periodic points y and $[y, y']$ are stably equivalent and hence must be equal. Similarly, the periodic points y' and $[y, y']$ are unstably equivalent and hence must be equal. We conclude that $y = y'$ as desired. □

The most important basic fact about maps between Smale spaces is that they must preserve the bracket, as follows.

Theorem 4.1.6. *Let (Y, g) and (X, f) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be a map. There exists $\epsilon_\pi > 0$ such that, for all y_1, y_2 in Y with $d(y_1, y_2) \leq \epsilon_\pi$, then both $[y_1, y_2], [\pi(y_1), \pi(y_2)]$ are defined and

$$\pi([y_1, y_2]) = [\pi(y_1), \pi(y_2)].$$

Proof. Let ϵ_X, ϵ_Y be the Smale space constants for X and Y, respectively. As Y is compact and π is continuous, we may find a constant $\epsilon > 0$ such that, for all y_1, y_2 is in Y with $d(y_1, y_2) < \epsilon$, we have $d(\pi(y_1), \pi(y_2)) < \epsilon_X/2$. From the continuity of the bracket map, we may choose ϵ_π such that $0 < \epsilon_\pi < \epsilon_Y$ and for all y_1, y_2 in X with $d(y_1, y_2) \leq \epsilon_\pi$, we have

$$d(y_1, [y_1, y_2]), d(y_2, [y_1, y_2]) < \epsilon.$$

Now assume y_1, y_2 are in Y with $d(y_1, y_2) \leq \epsilon_\pi$. It follows that, $[y_1, y_2]$ is defined and we have the estimates above. Then, inductively for all $n \geq 0$, we have

$$d(g^n(y_1), g^n[y_1, y_2]) \leq \lambda^n d(y_1, [y_1, y_2]) \leq \epsilon$$

and also

$$d(g^{-n}(y_2), g^{-n}[y_1, y_2]) \leq \lambda^n d(y_2, [y_1, y_2]) \leq \epsilon$$

It follows from the choice of ϵ that, for all $n \geq 0$, we have

$$\begin{aligned} d(\pi(g^n(y_1)), \pi(g^n[y_1, y_2])) &\leq \epsilon_X/2 \\ d(f^n(\pi(y_1)), f^n(\pi[y_1, y_2])) &\leq \epsilon_X/2 \end{aligned}$$

and similarly

$$d(f^{-n}(\pi(y_2)), f^{-n}(\pi[y_1, y_2])) \leq \epsilon_X/2.$$

On the other hand, these two estimates are also satisfied replacing $f^n(\pi[y_1, y_2])$ by $f^n[\pi(y_1), \pi(y_2)]$ and so, by expansiveness of (X, f), we have the desired conclusion. □

4.2 s/u-resolving maps and s/u-bijective maps

In this section, we discuss special classes of maps called s-resolving, u-resolving, s-bijective and u-bijective maps. These maps possess many nice properties.

It is an easy consequence of the definitions that if (Y, g) and (X, f) are Smale spaces and

$$\pi : (Y, g) \to (X, f)$$

is a map, then $\pi(Y^s(y)) \subset X^s(\pi(y))$ and $\pi(Y^u(y)) \subset X^u(\pi(y))$. We recall the following definition due to David Fried [**?**].

Definition 4.2.1. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be a map. We say that π is s-resolving *(or* u-resolving*) if, for any y in Y, its restriction to $Y^s(y)$ (or $Y^u(y)$, respectively) is injective.*

The following is a useful technical preliminary result.

Proposition 4.2.2. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be an s-resolving (or u-resolving) map. With ϵ_π as in Theorem **??***, if y_1, y_2 are in Y with $\pi(y_1)$ in $X^u(\pi(y_2), \epsilon_X)$ (or $\pi(y_1)$ in $X^u(\pi(y_2), \epsilon_X)$, respectively) and $d(y_1, y_2) \leq \epsilon_\pi$, then $y_2 \in Y^u(y_1, \epsilon_\pi)$ ($y_2 \in Y^s(y_1, \epsilon_\pi)$, respectively).*

In particular, if $\pi(y_1) = \pi(y_2)$ and $d(y_1, y_2) \leq \epsilon_\pi$, then $y_2 \in Y^u(y_1, \epsilon_\pi)$ ($y_2 \in Y^s(y_1, \epsilon_\pi)$, respectively).

Proof. It follows at once the from hypotheses that

$$\pi[y_1, y_2] = [\pi(y_1), \pi(y_2)] = \pi(y_1).$$

On the other hand $[y_1, y_2]$ is stably equivalent to y_1 and, since π is s-resolving, $[y_1, y_2] = y_1$. This completes the proof. □

Resolving maps have many nice properties, the first being that they are finite-to-one. We establish this, and a slight variant of it, as follows.

Theorem 4.2.3. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be an s-resolving map. There is a constant $M \geq 1$ such that

1. *for any x in X, there exist $y_1, \ldots, y_K$ in Y with $K \leq M$ such that*

$$\pi^{-1}(X^u(x)) = \cup_{k=1}^{K} Y^u(y_k),$$

and

2. *for any x in X, we have $\#\pi^{-1}\{x\} \leq M$. In particular, π is finite-to-one.*

Proof. Cover Y with balls of radius $\epsilon_\pi/2$, then extract a finite subcover, whose elements we list as $B_m, 1 \leq m \leq M$. We claim this M satisfies the desired conclusions.

For the first statement, given x in X and y in $\pi^{-1}(X^u(x))$, it is clear that $Y^u(y) \subset \pi^{-1}(X^u(x))$. We must show that there exist at most M unstable equivalence classes in $\pi^{-1}(X^u(x))$. For this, it suffices to show that if $y_i, 1 \leq i \leq M+1$, are in Y with $\pi(y_i)$ and $\pi(y_j)$ unstably equivalent, for all i, j, then y_i and y_j are unstably equivalent for some $i \neq j$. Choose $n \leq 0$ such that $f^n(\pi(y_i))$ is in $X^u(f^n(\pi(y_j)), \epsilon_X)$, for all $1 \leq i, j \leq M+1$. From the pigeon hole principle, there exists distinct i and j such that $g^n(y_i)$ and $g^n(y_j)$ lie in the same B_m, for some $1 \leq m \leq M$. These points satisfy the hypotheses of **??** and it follows that they are unstably equivalent. Then y_i are y_j are also unstably equivalent.

For the second statement, suppose $\pi^{-1}\{x\}$ contains distinct points $y_1, \ldots, y_{M+1}$. Let δ denote the minimum distance, $d(y_i, y_j)$, over all $i \neq j$. Choose $n \geq 1$ such that $\lambda^n \epsilon_\pi < \delta$. Consider the points $g^n(y_i), 1 \leq i \leq M+1$. By the pigeon-hole principle, there exists $i \neq j$ with $g^n(y_i)$ and $g^n(y_j)$ in the same set B_m. We have

$$\pi(g^n(y_i)) = f^n(\pi(y_i)) = f^n(x) = f^n(\pi(y_j)) = \pi(g^n(y_j)).$$

From Proposition **??**, $g^n(y_i)$ is in $Y^u(g^n(y_j), \epsilon_\pi)$. This implies that y_i is in $Y^u(y_j, \lambda^n \epsilon_\pi)$. As $\lambda^n \epsilon_\pi < \delta$, this is a contradiction. □

Although the definition of s-resolving is given purely at the level of the stable sets as sets, various nice continuity properties follow.

Theorem 4.2.4. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be either an s-resolving or a u-resolving map. For each y in Y, the maps

$$\pi : Y^s(y) \to X^s(\pi(y)), \pi : Y^u(y) \to X^u(\pi(y))$$

are continuous and proper, where the sets above are given the topologies of Proposition **??**.

Proof. From the symmetry of the statement, it suffices to consider the case that π is s-resolving.

We use the characterization of limits in $Y^s(y)$ and $X^s(\pi(y))$ given in Proposition **??**. From this, and Theorem **??**, it is easy to see that π is continuous on $Y^s(y)$. The same argument covers the case of π on $Y^u(y)$.

To see the map π on $Y^s(y)$ is proper, it suffices to consider a sequence y_n in $Y^s(y)$ such that $\pi(y_n)$ is convergent in the topology of $X^s(\pi(y))$, say with limit x, and show that it has a convergent subsequence. As Y is compact in its usual topology, we may find y' which is a limit point of a convergent subsequence $y_{n_k}, k \geq 1$. It follows that

$$\pi(y') = \pi(\lim_k y_{n_k}) = \lim_k \pi(y_{n_k}) = \lim_n \pi(y_n) = x.$$

We also have, for k sufficiently large,

$$\pi[y_{n_k}, y'] = [\pi(y_{n_k}), \pi(y')] = [\pi(y_{n_k}), x] = x,$$

since $\pi(y_{n_k})$ is converging to x in the topology on $X^s(\pi(y))$ and using Proposition **??**. We know that $\pi^{-1}\{x\}$ is finite and contains y' and $[y_{n_k}, y']$, for all k sufficiently large. Moreover, y' is the limit of the sequence $[y_{n_k}, y']$. It follows that there is K such that $[y_{n_k}, y'] = y'$, for all $k \geq K$. From this, we see that y' is in $Y^s(y_{n_k}) = Y^s(y)$ and that the subsequence y_{n_k} converges to y' in $Y^s(y)$.

To see the map π on $Y^u(y)$ is proper, we begin in the same way with a sequence y_n such that $\pi(y_n)$ has limit x in the topology of $X^u(\pi(y))$. Again we obtain a subsequence y_{n_k} with limit y' in Y. Then we have, for k sufficiently large,

$$\pi[y_{n_k}, y'] = [\pi(y_{n_k}), \pi(y')] = [\pi(y_{n_k}), x] = \pi(y_{n_k}),$$

since $\pi(y_{n_k})$ is converging to x in the topology on $X^u(\pi(y))$ and using Proposition **??**. On the other hand, $[y_{n_k}, y']$ and y_{n_k} are stably equivalent and since π is s-resolving, this implies they are equal. It follows that y' is in $Y^u(y_{n_K}) = Y^u(y)$ and y_{n_k} is converging to y' in the topology of $Y^u(y)$. □

There has been extensive interest in s/u-resolving maps. We will need a slightly stronger condition, which we refer to as s/u-bijective maps.

Definition 4.2.5. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be a map. We say that π is s-bijective *(or* u-bijective*) if, for any y in Y, its restriction to $Y^s(y)$ (or $Y^u(y)$, respectively) is a bijection to $X^s(\pi(y))$ (or $X^u(\pi(y))$, respectively).*

It is relatively easy to find an example of a map which is s-resolving, but not s-bijective and we will give one in a moment. However, one important distinction between the two cases should be pointed at at once. The image of a Smale space under an s-resolving map is not necessarily a Smale space. The most prominent case is where the domain and range are both shifts of finite type and the image is a sofic shift, which is a much broader class of systems. (See [**?**].) This is not the case for s-bijective maps (or u-bijective maps).

Theorem 4.2.6. *Let (Y, g) and (X, ϕ) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be either an s-bijective map or a u-bijective map. Then $(\pi(Y), f|_{\pi(Y)})$ is a Smale space.

Proof. The only property which is not clear is the existence of the bracket: if y_1 and y_2 are in Y and $d(\pi(y_1), \pi(y_2)) < \epsilon_X$, then it is clear that $[\pi(y_1), \pi(y_2)]$ is defined, but we must see that it is in $\pi(Y)$. If π is s-bijective, then $[\pi(y_1), \pi(y_2)]$ is stably equivalent to $\pi(y_1)$ and hence in the set $\pi(Y^s(y_1))$ and hence in $\pi(Y)$. A similar argument deals with the case π is u-bijective. $\square$

If $\pi : (Y, g) \to (X, f)$ is a factor map and every point in the system (Y, g) is non-wandering (including the case that (Y, g) is irreducible), then it follows that the same is true of (X, f) and in this case, any s-resolving factor map is also s-bijective, as we will show.

Example 4.2.7. *Consider (Y, g) to be the shift of finite type associated with the following graph:*

$$\begin{tikzcd}
v_1 \arrow[loop left] \arrow[r] & v_2 \arrow[loop right] \\
v_3 \arrow[loop left] \arrow[r] & v_4 \arrow[loop right]
\end{tikzcd}$$

and (X, f) to be the shift of finite type associated with the following graph:

$$\begin{tikzcd} w_1 \arrow[loop left] \arrow[r, bend left] \arrow[r, bend right] & w_2 \arrow[loop right] \end{tikzcd}$$

It is clear that there is a factor map from (Y, g) to (X, f) obtained by mapping the loops in the first graph to those in the second 2-to-1, while mapping the other two edges injectively. The resulting factor map is s-resolving and u-resolving but not s-bijective or u-bijective.

Theorem 4.2.8. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be an s-resolving factor map. Suppose that each point of (Y, g) is non-wandering. Then π is s-bijective.

Before beginning the proof, we need a version of Lemma **??** for local stable sets.

Lemma 4.2.9. *Let $\pi : (Y, g) \to (X, f)$ be a factor map between Smale spaces and suppose x_0 in X is periodic and $\pi^{-1}\{x_0\} = \{y_1, y_2, \ldots, y_N\}$. Given $\epsilon_0 > 0$, there exist $\epsilon_0 > \epsilon > 0$ and $\delta > 0$ such that*

$$\pi^{-1}(X^s(x_0, \delta)) \subset \cup_{n=1}^{N} Y^s(y_n, \epsilon).$$

Proof. First, since x_0 is periodic, so is each y_n. Choose $p \geq 1$ such that $g^p(y_n) = y_n$, for all $1 \leq n \leq N$, and hence $f^p(x_0) = x_0$. The system (Y, g^p) is also a Smale space. Choose $\epsilon_0 > \epsilon > 0$ to be less than the Smale space constant for this. Also, choose ϵ sufficiently small so that the sets $Y(y_n, \epsilon), 1 \leq n \leq N$ are pairwise disjoint and so that $g^p(Y(y_n, \epsilon)) \cap Y(y_m, \epsilon) = \emptyset$, for $m \neq n$. Use the Lemma **??** to find δ such that $\pi^{-1}\{x\} \subset \cup_{n=1}^{N} Y(y_n, \epsilon)$, for all x in $X(x_0, \delta)$.

Now suppose that x is in $X^s(x_0, \delta)$ and $\pi(y) = x$. It follows that y is in $Y(y_m, \epsilon)$, for some m. Now consider $k \geq 1$. We have $\pi(g^{kp}(y)) = f^{kp}(x)$ which is in $X^s(x_0, \lambda^{kp}\delta) \subset X(x, \delta)$. It follows that $g^{kp}(y)$ is in $\cup_{n=1}^{N} Y^s(y_n, \epsilon)$ for all $k \geq 1$. It then follows from the choice of ϵ and induction that $g^{kp}(y)$ is in $Y^s(y_m, \epsilon)$ for all $k \geq 1$. This means that y is in $Y^s(y_m, \epsilon)$. □

We are now prepared to give a proof of Theorem **??**.

Proof. In view of the structure Theorem **??**, it suffices for us to consider the case that (X, f) is irreducible. First, choose a periodic point x_0 satisfying the conclusion of Lemma **??**. Let $\pi^{-1}\{x_0\} = \{y_1, \ldots, y_N\}$. We will first show that, for each $1 \leq n \leq N$, $\pi : Y^s(y_n) \to X^s(x_0)$ is open and onto. We choose $\epsilon_0 > 0$ so that the sets $Y(y_n, \epsilon_0), 1 \leq n \leq N$, are pairwise disjoint. We then choose $\epsilon_0 > \epsilon > 0$ and $\delta > 0$ as in Lemma **??**. Let x be any point in $B(x_0, \delta)$. We know that $\pi^{-1}(\{x\})$ is contained in $\cup_{n=1}^{N} Y^s(y_n, \epsilon_0)$. As the map π is s-resolving, it is injective when restricted to each of the sets $Y^s(y_n, \epsilon)$. This means that $\pi^{-1}\{x\}$ contains at most one point in each of these sets. On the other hand, it follows from our choice of x_0 that $\pi^{-1}\{x\}$ contains at least N points. We conclude that, for each n, $\pi^{-1}\{x\} \cap Y^s(y_n, \epsilon)$ contains exactly one point. Let $W_n = \pi^{-1}(X(x, \delta)) \cap Y^s(y_n, \epsilon)$. The argument above shows that π is a bijection from W_n to $X^s(x_0, \delta)$, for each n. It is clearly continuous and we claim that is actually a homeomorphism. To see this, it suffices to show that, for any sequence y^k in W_n such that $\pi(y^k)$ converges to some x in $X^s(x_0, \delta)$, it follows that y^k converges to some y in W_n. As $\epsilon < \epsilon_0$, the closure of W_n is a compact subset of $Y^s(y_n, \epsilon_0)$. So the sequence y^k has limit points; let y be one of them. By continuity, $\pi(y) = x$. On the other hand, there is a unique point y' in W_n such that $\pi(y') = x$. Thus, y and y' are both in $Y^s(y_n, \epsilon_0)$ and have image x under π. As π is s-resolving, $y = y'$ and so y is in W_n. So the only limit point of the sequence y^k is y' and this completes the proof that π is a homeomorphism.

Since W_n is an open subset of $Y^s(y_n, \epsilon)$, we know that $Y^s(y_n) = \cup_{l \geq 0} g^{-l}(W_n)$ and the topology is the inductive limit topology. Similarly, $X^s(x_0) = \cup_{l \geq 0} f^{-l}(X^s(x_0, \delta))$ and the topology is the inductive limit topology. It follows at once that π is a homeomorphism from the former to the latter.

Now we turn to arbitrary point y in Y and $x = \pi(y)$ in X and show that $\pi : Y^s(y) \to X^s(x)$ is onto. We choose x_0 and $\{y_1, \ldots, y_N\}$ to be periodic points as above so that $\pi : Y^s(y_n) \to X^s(x_0)$ are homeomorphisms. By replacing x_0 by another point in its orbit (which will satisfy the same condition), we may assume that x is in the closure of $X^s(x_0)$. Then, we may choose y_n such that y is in the closure of $Y^s(y_n)$. There exists a point y' is $Y^s(y_n)$ in $Y^u(y, \epsilon_Y/2)$ and so that $x' = \pi(y')$ is in $X^u(x, \epsilon_X/2)$. The map π may be written as the composition of three maps. The first from $Y^s(y, \epsilon_Y/2)$ to $Y^s(y', \epsilon_Y)$ sends z to $[y', z]$. The second from $Y^s(y', \epsilon_Y)$ to $X^s(x')$ is simply π. The third is the map from $X^s(x', \epsilon_X/2)$ to $X^s(x, \epsilon_X)$ sends z to $[x, z]$. Each is defined on an open set containing y, y' and x', respectively and is an open map. The conclusion is that there exists some $\epsilon' > 0$ such that $\pi(Y^s(y, \epsilon')) = U$ is an open set in $X^s(x)$ containing x. It then follows that

$$X^s(x) = \cup_{l \geq 0} f^{-l}(U) = \cup_{l \geq 0} \pi(g^{-l}(Y^s(y, \epsilon))) \subset \pi(Y^s(y)).$$

This completes the proof. □

Now we want to observe that although the property of a map being s-bijective is defined purely at the level of stable sets, continuity properties follow as a consequence.

Theorem 4.2.10. *Let (X, f) and (Y, g) be Smale spaces and let*

$$\pi : (Y, g) \to (X, f)$$

be an s-bijective (or u-bijective) map. Then for each y in Y, the map $\pi : Y^s(y) \to X^s(\pi(y))$ (or $\pi : Y^u(y) \to X^u(\pi(y))$, respectively) is a homeomorphism.

Proof. The proof is the general fact that if A, B are locally compact Hausdorff spaces and $f : A \to B$ is a continuous, proper bijection, then f is a homeomorphism. This can be seen as as follows. Let A^+ and B^+ denote the one-point compactifications of A and B, respectively. That the obvious extension of f to a map between these spaces is continuous follows from the fact that f is proper. Since this extension is a continuous bijection between compact Hausdorff spaces, it is a homeomorphism. The result follows from this argument and Theorem **??**. □

As we are discussing s/u-resolving maps between shifts of finite type, we describe a simple condition on the underlying graphs which is related.

Definition 4.2.11. *Let G and H be graphs. A graph homomorphism $\theta : H \to G$ is* left-covering *if it is surjective and, for every v in H^0, the map $\theta : t^{-1}\{v\} \to t^{-1}\{\theta(v)\}$ is a bijection. Similarly, π is* right-covering *if it is surjective and, for every v in H^0, the map $\theta : i^{-1}\{v\} \to i^{-1}\{\theta(v)\}$ is a bijection.*

The following result is obvious and we omit the proof.

Theorem 4.2.12. *If G and H are graphs and $\theta : H \to G$ is a left-covering (or right-covering) graph homomorphism, then the associated map $\theta : (\Sigma_H, \sigma) \to (\Sigma_G, \sigma)$ is an s-bijective (or u-bijective, respectively) factor map.*

4.3 Degree

Definition 4.3.1. *Let $\pi : (Y, g) \to (X, f)$ be a finite-to-one map between two dynamical systems. We define the* degree *of π, denoted* $\deg(\pi)$ *by*

$$\deg(\pi) = \inf\{\#\pi^{-1}\{x\} \mid x \in X\}.$$

We re-state Theorem **??**, just for emphasis.

Theorem 4.3.2. *Let $\pi : (Y, g) \to (X, f)$ be a factor map between two Smale spaces. If (X, f) is non-wandering, then there exists a periodic point x in X with*

$$\deg(\pi) = \#\pi^{-1}\{x\}$$

Consider an s-resolving map $\pi : (Y, g) \to (X, f)$ between two Smale spaces. By definition, if we have two points y, y' in Y with $\pi(y) = \pi(y')$ and

$$\lim_{n\to\infty} d(g^n(y), g^n(y')) = 0,$$

then $y = y'$. In fact, the second condition may be relaxed considerably.

Lemma 4.3.3. *Let $\pi : (Y, g) \to (X, f)$ be an s-resolving map between two Smale spaces. Let y, y' be in Y with $\pi(y) = \pi(y')$ and*

$$\liminf_{n\to+\infty} d(g^n(y), g^n(y')) < \epsilon_\pi$$

(as in Lemma **??***). Then $y = y'$.*

Proof. The hypothesis means that we have infinitely values of $n \geq 0$ such that

$$d(g^n(y), g^n(y')) < \epsilon_\pi$$

For such n, it follows from Lemma **??** that $g^n(y)$ is in $Y^u(g^n(y'), \epsilon_\pi)$. From this, we see that

$$d(g^{n-j}(y), g^{n-j}(y')) \leq \lambda^j \epsilon_{pi},$$

for all $j \geq 0$. Letting $j = n$ and using the fact that there are infinitely many positive n for which the estimate holds, we see $y = y'$. □

We know already that there exists a periodic point with minimum fibre size. We next want to show that for an s-resolving map, points in the quotient having a dense forward orbit will also have minimum size fibres.

Theorem 4.3.4. *Let $\pi : (Y, g) \to (X, f)$ be a s-resolving (or u-resolving) map between two Smale spaces. If (X, f) is irreducible, then for every point x in X with a dense forward orbit (backward orbit, respectively) we have*

$$\deg(\pi) = \#\pi^{-1}\{x\}$$

Proof. We assume that π is s-resolving. Suppose that $\pi^{-1}\{x\} = \{y_1, \ldots, y_k\}$, so that $k \geq \deg \pi = d$ by definition. We know there is a point x' in X with precisely d pre-images under π. Find a sequence of positive integers n_i such that $f^{n_i}(x), i \geq 1$, converges to x'. By passing to a subsequence, we may assume that the sequence $g^{n_i}(y_j), i \geq 1$, converges for each $j = 1, 2, \ldots k$. By Lemma **??**, the limit points of these sequences are all distinct, but clearly they all lie in $\pi^{-1}\{x'\}$. We conclude that $k \leq d$. □

We would now like to understand a little more about how stable and unstable equivalence behaves on pre-images of points. Of course, if the map is s-resolving, no two points in a pre-image of a singleton can be stably equivalent. However, we can also show the degree of the map is also a lower bound on the number of unstable classes.

Theorem 4.3.5. *Let $\pi : (Y, g) \to (X, f)$ be a s-resolving (or u-resolving) map between two Smale spaces. For any x in X, there exists at least* $\deg(\pi)$ *points in $\pi^{-1}\{x\}$, no two of which are either stably or unstably equivalent.*

Proof. As we mentioned above, the fact that π is s-resolving means no two points in the pre-image of x can be stably equivalent. It suffices to prove that there are at least $d = \deg(\pi)$ distinct unstable classes. If not, we may replace x by $f^k(x)$ (for some large negative value of k to assume that there are $d - 1$ points in $\pi^{-1}\{x\}$ such that the $\epsilon_\pi/2$-balls around them cover $\pi^{-1}\{x\}$. Now choose x' in X with a dense forward orbit and an increasing sequence n_i such that $f^{n_i}(x')$ convergse to x. By passing to a sebsequence, we may assume that $g^{n_i}(y')$ also converges, for each y' in $\pi^{-1}\{x'\}$. Each must converge to some point in $\pi^{-1}\{x\}$. This means that there are a pair of distinct elements of $\pi^{-1}\{x'\}$ satisfying the hypotheses of Lemma 4.3.3, which is a contradiction. □

Chapter 5

Products

5.1 Products of Smale spaces

First, if (X, d_X) and (Y, d_Y) are metric spaces, we let $d_X \times d_Y$ be the metric on the product space $X \times Y$ defined by

$$d_{X\times Y}((x,y),(x',y')) = \max\{d_X(x,x')d_Y(y,y')\},$$

for all x, x' in X and y, y' in Y.

Theorem 5.1.1. *If (X, d_X, f) and (Y, d_Y, g) are Smale spaces, then so is there product $(X \times Y, d_X \times d_Y, f \times g)$.*

Proof. Let $\epsilon_{X\times Y} = \min\{\epsilon_X, \epsilon_Y\}$ and define

$$[(x,y),(x',y')] = ([x,x'],[y,y']),$$

for all $(x,y),(x',y')$ in $X \times Y$ with $d_{X\times Y}((x,y),(x',y')) \leq \epsilon_{X\times Y}$. It is a routine matter to verify the axioms are satisfied. □

We will note that $(X \times Y)^s((x,y),\epsilon) = X^s(x,\epsilon) \times Y^s(y,\epsilon)$ and similarly for the local unstable sets. Similar formulae also holds for the global stable and unstable sets.

5.2 Fibred products

We first review the standard construction of the fibred product.

Definition 5.2.1. *Let X, Y_1 and Y_2 be topological spaces and $\pi_1 : Y_1 \to X$ and $\pi_2 : Y_2 \to X$ be continuous maps. The* fibred product *of π_1 and π_2 is*

$$Z = \{(y_1, y_2) \in Y_1 \times Y_2 \mid \pi_1(y_1) = \pi_2(y_2).$$

We have natural maps $\rho_i : Z \to Y_i$ defined by $\rho_i(y_1, y_2) = y_i$, for $i = 1, 2$ which satisfy $\pi_1 \circ \rho_1 = \pi_2 \circ \rho_2$.

Theorem 5.2.2. *Let X, Y_1 and Y_2 be topological spaces, $\pi_1 : Y_1 \to X$ and $\pi_2 : Y_2 \to X$ be continuous maps and let Z, ρ_1, ρ_2 be their fibred product. For each of the following properties, if π_1 has the property, then so does ρ_2 and if π_2 has the property, then so does ρ_1:*

1. *injective*

2. *surjective*

3. *finite-to-one.*

Proof. Let y_1 be in Y_1. It is easy to see that

$$\rho_1^{-1}\{y_1\} = \{y_1\} \times \pi_2^{-1}\{\pi_1(y_1)\}$$

The first statement follows at once. The second case is obviously similar. □

It is clear from the defintions that if the three spaces are each part of a dynamical system and the maps are maps of the systems, then the fibred product is also a dynamical system and the maps ρ_1 and ρ_2 are maps of systems. We are interested in analogue of the result above for Smale spaces and s/u-resolving maps and s/u-bijective maps.

Theorem 5.2.3. *Let X, Y_1 and Y_2 be Smale spaces and $\pi_1 : Y_1 \to X$ and $\pi_2 : Y_2 \to X$ be maps. Also, let $(Z, \zeta), \rho_1, \rho_2$ be their fibred product. If π_1 is s-resolving (u-resolving, s-bijective and u-bijective) then ρ_2 is s-resolving (u-resolving, s-bijective and u-bijective, respectively). A similar statement holds for π_2 and ρ_1.*

Proof. We first suppose that π_1 is s-resolving. Let $(y_1, y_2), (y_1', y_2')$ be stably equivalent in Z and suppose that the have the same image under ρ_2. The first part means that y_1 and y_1' are stably equivalent and the second simply means that $y_2 = y_2'$. It follows that $\pi_1(y_1) = \pi_2(y_2) = \pi_2(y_2') = \pi_1(y_1')$. Then $y_1 = y_1'$ since π_1 is s-resolving.

Next, we assume that π_1 is s-bijective. Suppose that (y_1, y_2) is in Z, y_2' is in Y_2 and $\rho_2(y_1, y_2)$ and y_2' are stably equivalent. It follows that y_2 and y_2' are stably equivalent. Therefore, $\pi_2(y_2) = \pi_1(y_1)$ and $\pi_2(y_2)$ are stably equivalent. As π_1 is s-bijective, we may find y_1' which is stably equivalent to y_1 and $\pi_1(y_1) = \pi_2(y_2)$. This means that (y_1', y_2') is in Z and it is clearly stably equivalent to (y_1, y_2).

□

Chapter 6

Markov Partitions

In this section, we introduce the notion of a Markov partition for a non-wandering Smale space. We prove the existence of such items. Finally, we draw important conclusions from the existence. One of these is that, for any irreducible Smale space (X, d, f), there is a positive integer N, an irreducible 0-1 $N \times N$ matrix A and a map

$$\pi : \Sigma_A \to X$$

from the associated shift of finite type to X.

This means that the shifts of finite type have a certain universal property among Smale spaces. Furthermore, we will prove that if, in the situation above, X is totally disconnected, then the map π can be chosen to be an isomorphisim.

6.1 Rectangles

Let us first introduce the notion of a rectangle in a Smale. This is a feature of the local product structure which we showed in **??**.

Definition 6.1.1. *Let (X, d, f) be a Smale space. We say that $R \subset X$ is a* rectangle *if it has diameter less than $\epsilon_X/2$ and $[R, R] = R$.*

It is worth noting that the inclusion $[R, R] \supset R$ always holds for a set of diameter less than ϵ_X.

Proposition 6.1.2. *Let A and B be non-empty subsets of a Smale space X such that $d(x, y) \leq \epsilon_X$, for all x in A and y in B, and so that the diameters of A, B and $[A, B]$ are all less than $\epsilon_X/2$. Then $R = [A, B]$ is a rectangle.*

Proof. Let x, x' be in A and y, y' be in B. From the hypotheses, we see that both $[[x, y], [x', y']]$, $[x, [x', y']]$ and $[x, y']$ are all defined and it follows from the definition that they are equal. This implies that $[R, R] \subset R$ and hence the conclusion. □

Proposition 6.1.3. *1. If R and R' are rectangles, then so is $R \cap R'$.*

2. If R is a rectangle, then so is $f(R)$ (and $f^{-1}(R)$), provided $f(R)$ ($f^{-1}(R)$, respectively) has sufficiently small diameter.

Proof. It is clear that the diameter of the intersection is smaller than the diameter of either set. Moreover, we have

$$[R \cap R', R \cap R'] \subset [R, R] = R.$$

In a similar way, $[R \cap R', R \cap R'] \subset R'$ and hence $[R \cap R', R \cap R'] \subset R \cap R'$. The second part is clear from the fact $[f(R), f(R)] = f[R, R]$, provided both sides are defined. □

The following will be useful.

Theorem 6.1.4. *Let R and R' be rectangles whose union has diameter lessthan $\epsilon_X/2$. The following are equivalent.*

1. *$[R, R'] = R \cap R'$,*
2. *for all x in $R \cap R'$, we have $[R, x] \subset [R', x]$ and $[x, R] \supset [x, R']$,*
3. *for some x in $R \cap R'$, we have $[R, x] \subset [R', x]$ and $[x, R] \supset [x, R']$.*

Proof. We assume the first condition and prove the second holds. Let x be in $R \cap R'$. Then we have

$$[R, x] = [[R, R'], x] = [R \cap R', x] \subset [R', x].$$

A similar argument shows the other condition holds.

It is clear that the second condition inplies the third. We now show the third implies the first. First, we note that from the first part of Proposition 6.1.3, we have

$$R \cap R' = [R \cap R', R \cap R'] \subset [R, R'].$$

Letting x be as in condition three, we also have

$$[R, R'] = [[R, x], R'] \subset [[R', x], R'] = [R', R'] = R'.$$

Using the other containment, we also have $[R, R'] \subset R$. This completes the proof. □

Definition 6.1.5. *Let A be a subset of a Smale space X. We define its* stable boundary*, denoted $\partial^s(A)$, as those points x which are boundary points of $A \cap X^s(x)$ in the relative topology of X^s. Similarly, the* unstable boundary*, denoted $\partial^u(A)$, as those points x which are boundary points of $A \cap X^u(x)$ in the relative topology of X^u.*

6.2 Markov partitions

Definition 6.2.1. *We say that $\mathcal{R}$ is a Markov partition for X if it is a finite collection of rectangles satisfying the following conditions.*

1. *For each R in $\mathcal{R}$, the diameters of R and $f^{-1}(R)$ are both less than $\epsilon_X/2$.*
2. *Each rectangle is a regular closed set.*
3. *The union of the rectangles is all of X.*
4. *The interiors of the rectangles are pairwise disjoint.*
5. *If R and R' are rectangles in $\mathcal{R}$ and $Int(R) \cap Int(f^{-1}(R'))$ is non-empty, then*
$$[f^{-1}(R'), R] = f^{-1}(R') \cap R.$$
(Here, $Int(A)$ denotes the interior of the set A.)

The first condition is really a technical one for convenience. We refer to the last condition as the Markov property and it is the crucial one. If R and R' are in $\mathcal{R}$ and are such that $Int(R) \cap Int(f^{-1}(R'))$ is non-empty, then we write
$$R \rightsquigarrow R'.$$
This means that if R and $f^{-1}(R')$ meet in a non-trivial fashion (not just on their boundaries), then $f^{-1}(R')$ "stretches completely across" R in the stable direction, while R "stretches completely across $f^{-1}(R')$ in the unstable direction. Notice that this kind of thing is reasonable since the map f^{-1} will expand the stable sets and contract the unstable ones. Let us make this more precise.

Let us give two examples before we continue. First, we consider A, an $N \times N$ matrix with 0-1 entries., and its associated shift of finite type, (X_A, σ_A). For each integer $1 \leq i \leq N$, we define
$$R_i = \{a \in X_A \mid a_0 = i\}.$$
There is a small problem in this example with the diameters of the sets (or rather their f-preimages) being too large. But if we ignore this, it is easy to verify that $\{R_1, \ldots, R_N\}$ is a Markov partition and we leave this to the

reader. It is a particularly nice one because the sets are actually clopen rather than just regular closed. Let us point out one very important feature since it will become a running theme in what follows later. In this example, we have $R_i \rightsquigarrow R_j$ (meaning that R_i meets $f^{-1}(R_j)$) if and only if $A(i,j) = 1$.

For our second example, we consider the substitution tiling systems as described in Section **??**. We list our prototiles $p_1, \ldots, p_N$ and define, for each $1 \leq i \leq N$, the set

$$R_i = \{T \mid 0 \in t \in T, t \text{ is a translate of } p_i\}.$$

This means that R_i consists of all those tiles where the origin in $\mathbb{R}^d$ lies in a tile which is a translate of p_i. If the origin in the tiling lies on the boundary of several tiles, then this tiling will lie in several of the R_i's. We leave it to the reader to verify that these form a Markov partition. One can also check that $R_i \rightsquigarrow R_j$ here exactly when a translate of p_j appears in the inflation of p_i, $\omega(p_i)$.

We let $\partial^s(\mathcal{R})$ ($\partial^s(\mathcal{R})$) denote the union of the stable boundaries (unstable bounaries, respectively) of all the elements of a Markov partition $\mathcal{R}$.

We also note that from the Markov property, one can show

$$\begin{aligned} f(\partial^s(\mathcal{R})) &\subset \partial^s(\mathcal{R}) \\ f^{-1}(\partial^u(\mathcal{R})) &\subset \partial^u(\mathcal{R}). \end{aligned}$$

We now state Bowen's fundamental result on the existence of Markov partitions.

Theorem 6.2.2 (MPexist)**.** *Let (X, d, f) be a non-wandering Smale space and let ϵ be any psoitive real number. Then there exists a Markov partition for X, with each element having diameter less than ϵ.*

Chapter 7

Measures

In this section, we will discuss certain measures on a given Smale space, (X, f). First of all, we will restrict our attention to Borel probability measures; that is, measures μ which are defined on the σ-algebra of Borel subsets of X, $\mathcal{B}(X)$, are positive and have $\mu(X) = 1$. We say that such a measure, μ, is *f-invariant* (or simply *invariant*), if

$$\mu(f^{-1}(E)) = \mu(E)$$

for every Borel set $E \in \mathcal{B}(X)$.

Let us first mention that there are a lot of such measures. To any periodic point, x in X, we may take counting measure on the orbit of x and normalize it so as to have measure 1. That is, if n is the least positive integer such that $f^n(x) = x$, we let

$$\mu = \frac{1}{n} \sum_{i=1}^{n} \delta_{f^i(x)}$$

where δ_y denotes the atomic measure concentrated at y. Since Smale spaces have many periodic points (**??**), there are many invariant measures.

Here, we will concentrate our attention on one specific measure which will be called the Bowen measure, characterized by a specific property. We will show that it has a number of nice features. This should not give the reader the impression that the others are not important. There is a great deal of interest in other measures.

7.1 The Parry Measure

We will consider first a shift if finite type. We will write, quite explicitly, a measure and show that it has many nice features. This is called the Parry measure. One drawback is that seems at the outset to depend on the symbolic presentation of the shift. The fact that it does not will be described in the next subsection.

Let N be a positive integer and let A be an $N \times N$ matrix with $0, 1$ entries. We lat X_A, σ_A be the shift of finite type as in section **??**.

We could immediately give the definition of the Parry measure 7.1.1, but we will we take a few moments to put the construction into a more general framework.

Recall that X_A is a closed subset of $\Pi_{n \in \mathbb{Z}} \{1, 2, \ldots, N\}$ which we will denote by $\tilde{X}$. Also σ_A is the restriction to X_A of the shift map on $\tilde{X}$, which we denote by σ. Let $p = (p(1), \ldots, p(N))$ be a *probability vector*. That is, we have

$$\begin{aligned} p(i) &\geq 0, \text{ for all i =1, \ldots,N} \\ \sum_{i=1}^{N} p(i) &= 1. \end{aligned}$$

Also, let P denote a *stochastic* $N \times N$ matrix. That is, we have

$$\begin{aligned} P(i,j) &\geq 0, \text{ for all } i,j \\ \sum_j P(i,j) &= 1, \text{ for all } i. \end{aligned}$$

Also assume that p is a left eigenvector for P; that is, $pP = \lambda p$, for some scalar λ. If one sums the entries of the two sides of the equation $pP = \lambda p$, one finds that $\lambda = 1$. So we see that

$$\sum_i p(i)P(i,j) = p(j), \text{ for all } j.$$

Given this data one can form a Borel probability measure μ on $\tilde{X}$ called a *Markov measure*. One begins by defining μ on the so-called cylinder sets. If $k \leq l$ are integers and $(\xi_k, \ldots, \xi_l)$ is a sequence with $1 \leq \xi_i \leq N$, for $k \leq i \leq l$, then we define

$$U(k,l,\xi) = U(k,l,(\xi_k,\ldots,\xi_l)) = \{x \in \tilde{X} \mid x_i = \xi_i, \text{ for all } k \leq i \leq l\}.$$

We define

$$\mu(U(k,l,\xi)) = \mu(U(k,l,(\xi_k,\ldots,\xi_l))) = p(\xi_k)P(\xi_k,\xi_{k+1})\cdots P(\xi_{l-1},\xi_l).$$

Of course, a cylinder set for a given k, l can be written as the disjoint union of cylinder sets for fixed $k' \leq k$ and $l' \geq l$. The fact that our measure is additive on cylinder sets follows from the basic properties of p and P. This measure then extends uniquely to a probability measure on $\tilde{X}$ as explained in detail in [**?**]. The invariance of the measure under σ is immediate.

For our Smale space (X_A, σ_A), we would like this measure to be concentrated on the set $X_A \subset \tilde{X}$. For this, we require $0 \leq P_{i,j} \leq A(i,j)$, which is just to say that $P(i,j) = 0$ whenever $A(i,j) = 0$.

For the Parry measure, we make the following special choices for the vector and matrix. We apply the Perron-Frobenius Theorem to the matrix A, noting again our assumption that it is irreducible. This asserts that the matrix has an eigenvalue, λ, which is a positive real number and larger in absolute value than all other eigenvalues. Moreover, its (algebraic) multiplicity is one and there is an associated eigenvector, u, whose entries are all positive real numbers. This means that it is the right eigenvector for A; we have

$$\sum_{j=1}^{N} A(i,j)u(j) = u(i),$$

for all $i = 1, \ldots, N$. We call λ the *Perron* eigenvalue of A and u the *right Perron* eigenvector. Of course, the matrix also has a left Perron eigenvctor, which we denote by v. This can be obtained by applying the theorem to the matrix A^T. Here, we have

$$\sum_{j=1}^{N} A(j,i)v(j) = v(i),$$

for all $i = 1, \ldots, N$.

We may normalize the vectors u and v so that $\sum_i u(i)v(i) = 1$. Now we choose the probability vector

$$p = (u(1)v(1), \ldots, u(N)v(N))$$

and the stochastic matrix

$$P(i,j) = \lambda^{-1}A(i,j)u(j)u(i)^{-1}$$

for all i, j. It is straight-forward to verify the correct conditions hold.

Let us note that for this choice of matrix we have the following simple formula for the Parry measure of a cylinder set.

Theorem 7.1.1. *Let A be an irreducible $N \times N$ $0,1$-matrix with associated Perron eigenvalue λ and right and left eigenvectors u and v, respectively. Let (X_A, σ_A) be the shift of finite type associated with A. If $k \leq l$ are any integers and $\xi = (\xi_k, \cdots, \xi_l)$ is any sequence with $1 \leq \xi_i \leq N$, for $k \leq i \leq l$ and $A(\xi_i, \xi_{i+1}) = 1$, for all $k \leq i < l$, then the Parry measure of the corresponding cylinder set is given by*

$$\mu(U(k,l,\xi)) = \mu(U(k,l,(\xi_k, \ldots, \xi_l))) = \lambda^{k-l}v(\xi_k)u(\xi_l).$$

We will not give a proof; it is a simple computation using the formulae above.

The basic feature of a Smale space is the local product structure. We want to see that the Parry measure respects this. More specifically, we will show that, locally, it decomposes as a product measure.

Suppose that $k \leq 0 \leq l$, let ξ be as above and let x be in $U(\xi)$. We define

$$\begin{aligned} U^s(l, \xi, x) &= \{y \in U(\xi) \mid y_i = x_i, \text{ for all } i \geq 0\} \\ U^u(k, \xi, x) &= \{y \in U(\xi) \mid y_i = x_i, \text{ for all } i \leq 0\} \end{aligned}$$

Of course, these are just subsets of the local stable and unstable sets at x and $[,]$ defines a homeomorphism from $U^u(k, \xi, x) \times U^s(l, \xi, x)$ to $U(k, l, \xi)$.

We define measures, μ^s_x and μ^u_x on $U^s(0, (x_0), x)$ and $U^u(0, (x_0), x)$, respectively, by setting

$$\begin{aligned} \mu^s_x(U^s(k, \xi, x)) &= \lambda^k v(\xi_k) \\ \mu^u_x(U^u(l, \xi, x)) &= \lambda^{-l} u(\xi_l) \end{aligned}$$

for any ξ as above with $k \leq 0 \leq l$ such that $x \in U(k, l, \xi)$.

Of course, one must repeat the type of calculations which we meantioned earlier to show that these are additive on the cylinder sets and extend to measures. It is easy to see that we obtain the desired product structure; specifically, we have

$$\mu(U(k, l, \xi)) = \mu^s_x(U^s(k, \xi, x))\mu^u_x(U^u(l, \xi, x))$$

It is important to note that the measures μ^s_x and μ^u_x are not themselves invariant, even in the case that x is a fixed point of f. The following formulae are easily checked

$$\begin{aligned} \mu^s_{\sigma(x)}(\sigma(U^s(k, \xi, x))) &= \mu^s_{\sigma(x)}(U^s(k-1, \xi x_0, \sigma(x))) \\ &= \lambda^{k-1} v(\xi_k) \\ &= \lambda^{-1}\mu^s_x(U^s(k, \xi, x)) \end{aligned}$$

and

$$\begin{aligned} \mu^u_{\sigma^{-1}(x)}(\sigma^{-1}(U^u(l, \xi, x))) &= \mu^u_{\sigma^{-1}(x)}(U^s(l+1, \xi x_0, \sigma^{-1}(x))) \\ &= \lambda^{-(l+1)} u(\xi_l) \\ &= \lambda^{-1}\mu^u_x(U^u(l, \xi, x)) \end{aligned}$$

We may summarize the properties of these measures in the following theorem. Most of these we have already established or discussed. The others are easy computations from the formulae.

Theorem 7.1.2. *Let (X_A, σ_A) be an irreducible shift of finite type given by a $N \times N$ $0,1$ matrix A. Let the measures μ, μ^s_x, μ^u_x be given as above.*

1. *μ is a σ_A-invariant probability measure on X.*
2. *For any x in $_A$ and $0 < \epsilon < \epsilon_X/2$, the map $[,]$ is a homeomporphism from $V^u(x,\epsilon) \times V^s(x,\epsilon)$ to a neighbourhood of in X and*
$$\mu \circ [,] = \mu^u_x \times \mu^s_x,$$
on this set.
3. *For any x in X and y in $V^s(x, \epsilon_X/2)$, we have*
$$\mu^u_x(E) = \mu^u_y([E, y]),$$
whenever E is a Borel subset of $V^u(x, \epsilon_X/2)$ such that $[E, y]$ is contained in $V^u(y, \epsilon_X/2)$.
4. *For any x in X and y in $V^u(x, \epsilon_X/2)$, we have*
$$\mu^s_x(E) = \mu^s_y([y, E]),$$
whenever E is a Borel subset of $V^s(x, \epsilon_X/2)$ such that $[y, E]$ is contained in $V^s(y, \epsilon_X/2)$.

7.2 The Bowen Measure

Now we want to find to extend the situation to general Smale spaces and find a Measure which has the same nice features as the Parry measure. At the same time, this will provide a description of the Parry measure whihc is independent of the symbolic presentation.

The key ingredient here is the notion of *entropy*. We will not give a definition, since this would take us rather far afield and because there are a number of excellent references [**?**]. We summarize by saying that, to any invariant measure, μ, on X, we may assign a number $0 \leq h(f, \mu) \leq +\infty$.

In ergodic theory, one tends to take the measure μ as part of the given data. Here, μ is considered as a variable.

Definition 7.2.1. *For a Smale space* (X, f)*, the* Bowen measure *is the measure which maximizes* $h(f, \mu)$*, as* μ *varies over the set of all invariant Borel probability measures on* X*.*

Of course, it is not clear such a measure exists and if it does, it may not be unique. It turns out that the situation for Smale spaces is as nice as one might hope. The following may be found as Theorem 18.3.9 and Theorem 20.1.3 of [**?**].

Theorem 7.2.2. *For any irreducible Smale space* (X, f)*, the Bowen measure exists. Moreover, there is a unique measure* μ *which maximizes the entropy and* $h(f, \mu)$ *is finite.*