actual rectangles; all we really needed was that the boundaries of the elements of the partition lie in appropriate stable and unstable sets. The use of known stable and unstable sets as above to construct Markov partitions is a completely general operation. All that is necessary is that the map preserve these sets. For example, the vertical rectangles used in the construction of the horseshoe is a Markov partition for the associated Cantor set Λ. Note that no identifications in the sequence space are necessary in this case since there are no overlapping rectangles.

Exercises

1. Let L_A be a hyperbolic toral automorphism. Prove that:
   a. transverse homoclinic points are dense in T;
   b. all points in T are nonwandering (in the sense of Exercise 1.7.2);
   c. homoclinic points are not recurrent points (in the sense of Exercise 1.7.3).

2. One may define an n-dimensional torus T^n in exact analogy with our construction of the two-dimensional torus in this section. That is, let $[x_1, \ldots, x_n]$ denote the set of all equivalence classes of points in $\mathbf{R}^n$ under the equivalence relation

$$(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$$

if and only if $x_j - y_j$ is an integer for each j. The n-torus is then simply the set of all such equivalence classes of points in $\mathbf{R}^n$. Similarly, one may define a hyperbolic toral automorphism on T^n by starting with a matrix A which satisfies the conditions in Definition 4.1. Note that the stable and unstable sets need no longer be curves in T^n.
   a. Prove that the induced hyperbolic toral automorphism on T^n has dense periodic points.
   b. Prove that if $[p] \in T^n$, then $W^s[p]$ and $W^u[p]$ are dense in T^n.
   c. Prove that a hyperbolic toral automorphism is chaotic on T^n.

3. Prove that T^n is homeomorphic to the n-fold cross product

$$S^1 \underbrace{\times \ldots \times}_{n \text{ factors}} S^1.$$

4. Consider the map $A: \mathbf{R}^n \to \mathbf{R}^n$ given by $A(\mathbf{x}) = 2\mathbf{x}$. A induces a map on T^n exactly as in the case of a hyperbolic toral automorphism, but the induced map is no longer a diffeomorphism.
   a. Prove that periodic points are dense for this map.

   b. Prove that eventually fixed points are dense.
   c. Prove that this map is chaotic on T^n.

5. Let

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

Construct a Markov partition for L_A.

6. Let L_A be a hyperbolic toral automorphism on T. Let $[p] \in W^s[0] \cap W^u[0]$ be a homoclinic point. Let ℓ_s be the segment in $W^s[0]$ connecting $[0]$ to $[p]$ and let ℓ_u be a similar segment in $W^u[0]$. Construct a rectangle R containing ℓ_s with sides in stable and unstable sets.
   a. Show that there is an integer n such that $L_A^n(R) \supset \ell_u$.
   b. Prove that we may choose $[p]$ so that $L_A^n: R \to L_A^n(R)$ is topologically conjugate to the linear map which produced the horseshoe in §2.3.

§2.5 ATTRACTORS

In this section, we introduce a third type of dynamical phenomenon which is higher dimensional in nature, the attractor. Roughly speaking, an attractor is an invariant set to which all nearby orbits converge. Hence attractors are the sets that one "sees" when a dynamical system is iterated on a computer. Thus far, all of the attractors we have encountered have been fixed or periodic points. Here we introduce two new and much more complicated attractors, the solenoid and the Plykin attractor. These are examples of a special type of attractor known as a transitive or hyperbolic attractor. We will see that these attractors are similar in many respects to the horseshoe map and the hyperbolic toral automorphisms. For example, there is a set on which the map is chaotic and, through each point in this set, there passes a stable and an unstable set. Since these are familiar phenomena, we will leave many of the details in the verification to the reader.

The solenoid is an attractor which is contained in a "solid" torus. This space is defined as follows. Let S^1 be the unit circle and let B^2 be the unit disk in the plane; that is

$$B^2 = \{(x, y) \in \mathbf{R}^2 | x^2 + y^2 \leq 1\}.$$

The Cartesian product $D = S^1 \times B^2$ is a solid torus in $\mathbf{R}^3$. Its boundary is a torus as described in the previous section. To define the solenoid, we consider the map F which maps D strictly inside itself by the formula:

$$F(\theta, p) = (2\theta, \frac{1}{10}p + \frac{1}{2}e^{2\pi i\theta})$$

where $p \in B^2$ and $e^{2\pi i\theta} = (\cos(2\pi\theta),\ \sin(2\pi\theta)) \in S^1$.

Geometrically, F may be described as follows. Let $\theta^* \in S^1$. The disk $B(\theta^*)$ which is given by $\theta = \theta^*$ and p arbitrary is mapped by F into another disk given by $B(2\theta^*)$. The image of this disk is a disk of radius 1/10 with center at the point $\frac{1}{2}(\cos(2\theta^*),\ \sin(2\theta^*))$ in $B(2\theta^*)$. See Fig. 5.1. The disk located at $\theta = \theta^* + \pi$ is also mapped into the disk given by $\theta = 2\theta^*$, but its image is a small disk of radius 1/10 diametrically opposite the image of $B(\theta^*)$ in $B(2\theta^*)$.

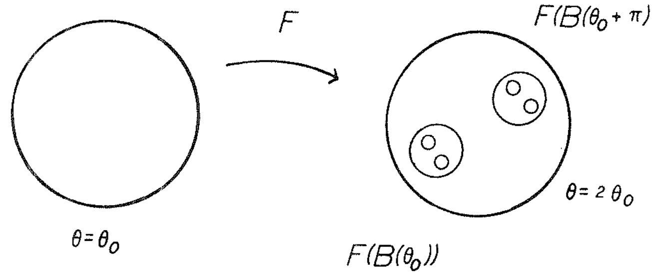

Fig. 5.1. Construction of the solenoid.

Globally, F may be interpreted as follows. In the θ coordinate, F is simply the doubling map of the circle discussed in Example 3.4 of Chapter One. In the B^2-direction, F is a strong contraction, with image a disk whose center depends on θ. The image of this disk is one-tenth the size of the original disk. Thus the image of D is another solid torus inside D which wraps twice around D. See Fig. 5.2.

The fact that F stretches in one direction and contracts in the others is, by now, a familiar phenomenon, reminiscent of both the horseshoe and the hyperbolic toral automorphisms.

Strictly speaking, F is not a diffeomorphism, since it is not onto. We think of D as a piece of a larger space and the action of F on D as just a portion of the dynamics. Since $F(D) \subset D$, it follows that all forward orbits of points in D lie in D. Regions like D have a special name.

Definition 5.1. A closed region $N \subset \mathbf{R}^n$ is a trapping region for F if $F(N)$ is contained in the interior of N.

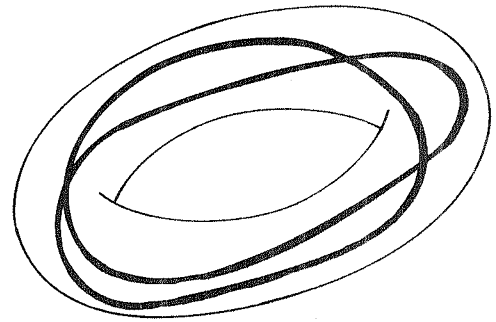

Fig. 5.2. The image of the solid torus under F is a solid torus which wraps twice around itself.

Since $F(N)$ is closed and $F(N) \subset N$, it follows that the sets $F^n(N)$ are all closed and nested for $n \geq 0$. Therefore

$$\Lambda = \bigcap_{n\geq 0} F^n(N)$$

is a closed, nonempty set. Λ is the set of points whose full orbits, both forward *and* backward, remain in N for all time. Λ will be our attractor.

Proposition 5.2. *Λ is an invariant set.*

Proof. We have

$$F(\Lambda) = F(\bigcap_{n\geq 0}^{\infty} F^n(N)) = \bigcap_{n\geq 1}^{\infty} F^n(N) \subset N.$$

But

$$\bigcap_{n\geq 0}^{\infty} F^n(N) = \bigcap_{n\geq 1}^{\infty} F^n(N)$$

since the intersections are nested. Hence $F(\Lambda) = \Lambda$ and Λ is invariant. Invariance under F^{-1} follows as well.

q.e.d.

Definition 5.3. A set Λ is called an attractor for F if there is a neighborhood N of Λ for which the closure of N is a trapping region and

$$\Lambda = \bigcap_{n\geq 0} F^n(N).$$

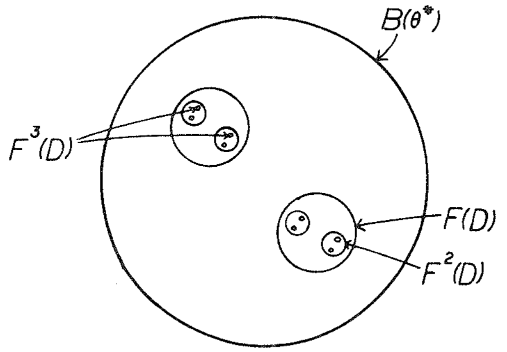

Fig. 5.3. The intersection of the $F^n(D)$ yields a Cantor set.

There are other definitions of attractors in common use. Ours is by no means standard, although it is perhaps the simplest. This definition suffers the defect that it does not produce a single, indecomposable attractor. For example, the "stadium" D for the horseshoe map of §2.3 is a trapping region. The attractor is easily seen to consist of two pieces, the fixed point in D_1 and the invariant Cantor set together with all of its unstable sets. On the other hand, the region $D_1 \subset D$ is also a trapping region, but this time the attractor is quite different; it is simply the fixed point in D_1.

To remedy this, we introduce the following terminology:

Definition 5.4. Λ is a transitive attractor for F if F is topologically transitive on Λ.

Our goal is to show that the attractor $\Lambda = \cap_{n\geq 0} F^n(D)$ for the above map is a transitive attractor and that, moreover, the dynamics of F on Λ are chaotic.

Let us investigate the nature of the set Λ. Since F stretches D in the S^1-direction and contracts it by a factor of $1/10$ in the B^2-direction, it follows that $F(D)$ is a torus of radius $1/10$ which wraps around D twice. Applying F to $F(D)$, we see that $F^2(D)$ is a torus of radius $1/100$ in the B^2-direction which wraps around D four times and which is properly contained in $F(D)$. Inductively, $F^n(D)$ is a torus of radius $1/10^n$ which wraps around D exactly 2^n times and which is contained in $F^{n-1}(D)$.

In each $B(\theta^*)$, we therefore see that $F^n(D)$ is a nested collection of 2^n disks, as in Fig. 5.3. We have seen this process before: the nesting of the $F^n(D)$ yields a Cantor set in each disk $B(\theta^*)$.

If we perform the above construction in a cylindrical piece of D of the form

$$C = \{(\theta, p) | \theta_1 \leq \theta \leq \theta_2\},$$

we see that $C \cap \Lambda$ is locally the Cartesian product of a Cantor set and an arc in the S^1-direction. The arcs are given by the nested intersection of the 2^n tubes in $F^n(D) \cap C$. Since each iteration of F contracts the radius of these tubes by $1/10$, it is intuitively clear that these arcs are continuous. Nevertheless, we will prove this later by completely different methods. In fact, it may be shown that these curves are smooth. The set Λ is called a *solenoid*.

We now turn to the dynamics of F on and near Λ. Let $x \in \Lambda$. Suppose $x = (\theta_0, p_0)$ where $\theta_0 \in S^1$ and $p_0 \in B^2$. Let $F^n(x) = (\theta_n, p_n)$. Consider the disk $B(\theta_0)$. Since F maps $B(\theta_0)$ inside $B(2\theta_0)$, it follows that $F^n(B(\theta_0)) \subset B(\theta_n)$. Moreover, each application of F contracts $B(\theta_0)$ by a factor of $1/10$. Therefore, if $y \in B(\theta_0)$, it follows that $F^n(y) \in B(\theta_n)$ and $|F^n(x) - F^n(y)| < 1/10^n$, where the absolute value is the usual one in $\mathbf{R}^2$. Consequently, $B(\theta_0)$ is part of the stable set $W^s(x)$ associated to x.

Similarly, the arc constructed above as the nested intersection of tubes about x is part of the unstable set for x which we denote by $W^u(x)$. This follows since F^{-1} contracts distances along the arc by a factor of $1/2$. We thus see that all of the points in Λ come equipped with stable and unstable sets, just as in the cases of the horseshoe and the hyperbolic toral automorphisms.

Proposition 5.5.

1. *F has sensitive dependence on initial conditions on Λ.*
2. *$Per(F)$ is dense in Λ.*
3. *F is topologically transitive on Λ.*

Proof. For sensitive dependence on initial conditions, we simply note that any point on the unstable arc associated to $x \in \Lambda$ separates from x by a factor of 2 in the θ-direction when F is iterated. To prove density of periodic points, let U be any neighborhood of $x = (\theta_0, p_0)$. There exists $\delta > 0$ and $n \in \mathbf{Z}$ such that the tube C in $F^n(D)$ defined by

$$C = \left\{(\theta, z) | |\theta - \theta_0| < \delta, |z - p_0| < \frac{1}{10^n}\right\}$$

is completely contained in U. We will produce a periodic point in C. To accomplish this, recall that $F^n(D)$ wraps around D exactly 2^n times. We may choose m so that $2^m\delta > 2^{n+1} \cdot 4\pi$. Hence $F^m(C)$ is a tube lying in

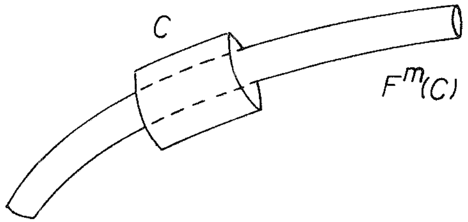

Fig. 5.4. The image $F^m(C)$ cuts through C.

$F^n(D)$ and wrapping around D at least $2 \cdot 2^n$ times. It follows that $F^m(C)$ cuts completely across C at least once as shown in Fig. 5.4. Hence there exists θ^* with $|\theta^* - \theta_0| < \delta$ such that $F^m(B(\theta^*) \cap C) \subset B(\theta^*) \cap C$. It follows that F^m has a fixed point in $B(\theta^*) \cap C$.

Similar arguments also prove topological transitivity. For if $x, y, \in \Lambda$ and U and V are neighborhoods of x and y, we may then produce tubes as above in $F^n(D)$ about x and y which are completely contained in U and V respectively. Sufficiently many iterations of these tubes produce a θ^* such that $B(\theta^*) \cap U$ is a disk which is mapped into V. It is easy to check that there is a point in Λ inside $B(\theta^*) \cap U$.

q.e.d.

As in our previous examples, we may use symbolic dynamics to model the dynamics of F on Λ. This time we use a different construction first introduced by R.F. Williams. Let $g: S^1 \to S^1$ be the doubling map $g(\theta) = 2\theta$. Our model for Λ will be the *inverse limit space*

$$\Sigma = (S^1 \overset{g}{\leftarrow} S^1 \overset{g}{\leftarrow} S^1 \ldots).$$

More precisely

$$\Sigma = \{\theta = (\theta_0\theta_1\theta_2\ldots) | \theta_j \in S^1 \text{ and } g(\theta_{j+1}) = \theta_j\}.$$

Thus Σ consists of all infinite sequences of points of S^1 subject to the restriction that θ_{j+1} is one of the two preimages of θ_j for each j. Unlike our previous sequence spaces, elements of Σ are not sequences whose entries are integers. Rather, the entries in this case are points in the circle. For example, the sequences

$$(0\,0\,0\ldots)$$
$$(0\,\pi\,\frac{\pi}{2}\,\frac{\pi}{4}\,\frac{\pi}{8}\ldots)$$

$$(\frac{\pi}{3}\,\frac{2\pi}{3}\,\frac{\pi}{3}\,\frac{2\pi}{3}\,\frac{\pi}{3}\ldots)$$

all belong to Σ. Using the doubling map g, it is helpful to think of these sequences as backward orbits:

$$0 \overset{g}{\leftarrow} 0 \overset{g}{\leftarrow} 0 \overset{g}{\leftarrow} \ldots$$

$$0 \overset{g}{\leftarrow} \pi \overset{g}{\leftarrow} \frac{\pi}{2} \overset{g}{\leftarrow} \frac{\pi}{4} \overset{g}{\leftarrow} \ldots$$

$$\frac{\pi}{3} \overset{g}{\leftarrow} \frac{2\pi}{3} \overset{g}{\leftarrow} \frac{\pi}{3} \overset{g}{\leftarrow} \frac{2\pi}{3} \overset{g}{\leftarrow} \ldots.$$

We define a metric on Σ much as we did on Σ_n. If $\Theta = (\theta_0\theta_1\theta_2\ldots)$ and $\Psi = (\psi_0\psi_1\psi_2\ldots)$ are points in Σ, we define the distance between them to be

$$d[\Theta, \Psi] = \sum_{j=0}^{\infty} \frac{|e^{2\pi i\theta_j} - e^{2\pi i\psi_j}|}{2^j}$$

where $|\alpha - \beta|$ denotes the usual Euclidean distance in the plane. It is easy to check that d is a metric on Σ. Moreover, two points are "close" if each of their first few entries are close together.

On Σ, we have a natural map, a version of the shift given by

$$\sigma(\theta_0\theta_1\theta_2\ldots) = (g(\theta_0)\theta_0\theta_1\theta_2\ldots).$$

As in previous sections, σ is easily seen to be a homeomorphism. The inverse of σ is given by a map that resembles our previous shift (but which is a homeomorphism)

$$\sigma^{-1}(\theta_0\theta_1\theta_2\ldots) = (\theta_1\theta_2\theta_3\ldots).$$

As with our previous models, this map is also easy to understand dynamically. If θ is a periodic point for g, with period n, then the repeating sequence $(\theta, g^{n-1}(\theta), g^{n-2}(\theta), \ldots, g(\theta), \theta, \ldots)$ is clearly periodic for σ with period n as well. As with our other examples, it is easy to check that σ has periodic points which are dense in Σ and that σ has a dense orbit. See Exercises 3-4.

How are σ and F related? Let $\pi: D \to S^1$ be the natural projection, i.e., $\pi(\theta, p) = \theta$. For any point $x \in \Lambda$, the map $S: \Lambda \to \Sigma$ given by

$$S(x) = (\pi(x),\ \pi F^{-1}(x),\ \pi F^{-2}(x), \ldots)$$

is well defined. This follows since we can invert F on Λ even though F^{-1} is not defined on all of D. Clearly, $S \circ F = \sigma \circ S$, since F is the doubling map in the S^1-direction.

We leave it as an exercise for the reader to prove that:

Theorem 5.6. *S gives a topological conjugacy between F on Λ and σ on Σ.*

Let us use this conjugacy to fill in the gap above where we failed to prove that the unstable sets in Λ were curves. For simplicity, let us prove this only for the fixed point which corresponds to the sequence $\mathbf{0} = (000\ldots)$. One checks easily that this is the point $\theta = 0$ and $p = (\frac{5}{9}, 0) \in B^2$.

Proposition 5.7. *The unstable set of $\mathbf{0}$ consists of precisely those sequences of the form*

$$(x, \frac{x}{2}, \frac{x}{2^2}, \frac{x}{2^3}, \ldots)$$

for any $x \in \mathbf{R}$.

Proof. By definition, we have $\sigma^{-1}(x, \frac{x}{2}, \frac{x}{4}, \ldots) = (\frac{x}{2}, \frac{x}{4}, \frac{x}{8}, \ldots)$. It therefore follows that $\sigma^{-n}(x, \frac{x}{2}, \frac{x}{4}, \ldots) \to \mathbf{0}$ as $n \to \infty$. For the converse, we first recall that if $\theta \in S^1$, then $g^{-1}(\theta)$ is one of $\frac{\theta}{2}$ or $\frac{\theta}{2} + \pi$. Now let $\Theta = (\theta_0 \theta_1 \theta_2, \ldots) \in W^u(\mathbf{0})$. There exists N such that if $n \geq N$, $|\theta_n| < 1$. Hence $\theta_N, \theta_{N+1}, \theta_{N+2}, \ldots$ all lie in the right hand semicircle in S^1. It follows that $\theta_{N+1} = \theta_N/2$, for the other preimage $(\theta_N/2) + \pi$ lies in the left semicircle. Continuing, we find $\theta_{N+k} = \frac{\theta_N}{2^k}$ and $\theta_{N-k} = 2^k \theta_N$ so that Θ assumes the desired form.

q.e.d.

Consequently, the unstable set of $\mathbf{0}$ in Σ is parametrized by $\mathbf{R}$. Under the conjugacy given by S, the unstable set of the fixed point is the continuous curve which is the image of $W^u(\mathbf{0})$.

The inverse limit construction works well for a class of attractors known as expanding attractors. These attractors are characterized by uniform expansion within the attractor itself. As in the case of the solenoid, such attractors can be suitably modeled by an inverse limit of a lower dimensional expanding map like $\theta \to 2\theta$ on S^1. The main difference in the general case is that the model space is more complicated than S^1; usually it is a "branched manifold." This concept was introduced by R.F. Williams. We will illustrate it via an example of an attractor due to Plykin. Rather than give a formula for this map, we will define it geometrically, exactly as we did for the horseshoe.

Consider the region R in the plane depicted in Fig. 5.5. R is a region with three open half-disks removed. We equip R with a foliation whose leaves are intervals as shown in Fig. 5.5. Recall that this means that there is a line segment through each point of R (the leaf) and that the leaves are mutually disjoint.

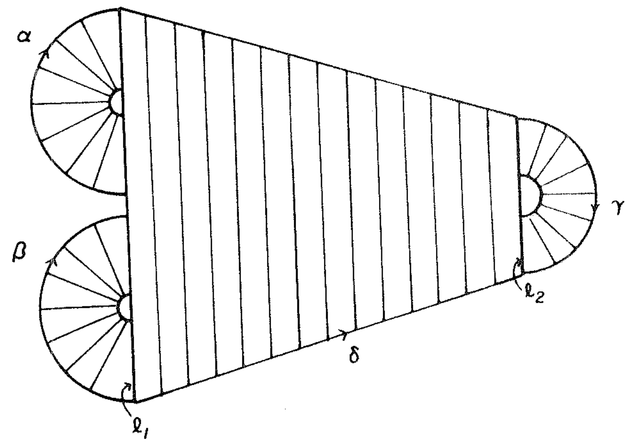

Fig. 5.5. The region R for the Plykin attractor.

Define a map $P: R \to R$ as shown in Fig. 5.6. We require that P preserve and contract the leaves of the foliation. Note that $P(R)$ is contained in the interior of R so that R is a trapping region. The set $\Lambda = \cap_{n \geq 0} P^n(R)$ is the Plykin attractor.

To understand the dynamics of P, we first note that any two points on the same leaf behave identically under iteration of P. Since the leaves are contracted, any two such points tend to the attractor in the same asymptotic manner. Thus, to understand the action of P globally, it suffices to understand the action of P on the leaves. We thus collapse each leaf to a point as in Fig. 5.7, and examine the induced map on this space. Observe that the collapsed space Γ has "branch" points along the singular leaves ℓ_1 and ℓ_2. It is called the branched "manifold" for P. We may describe the dynamics on Γ by describing how each of the four intervals α, β, γ, and δ are mapped. From Fig. 5.7 we see that the induced map g on Γ preserves the two vertices and maps the other intervals this way:

$$\alpha \to \beta$$

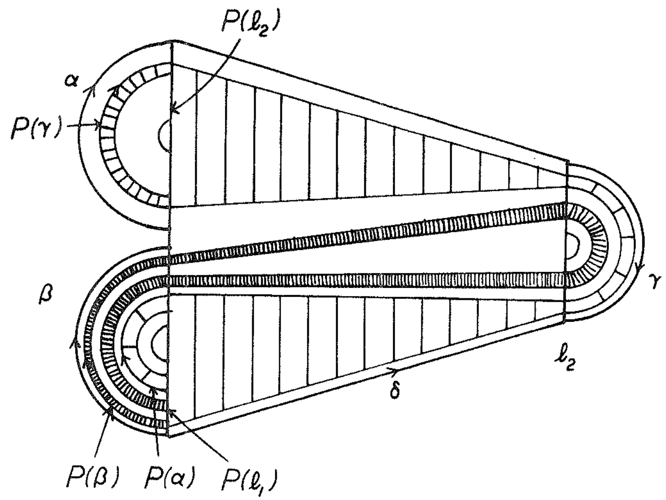

Fig. 5.6. The Plykin attractor.

$$\beta \to \beta + \delta + \gamma - \delta - \beta$$

$$\gamma \to \alpha$$

$$\delta \to \delta - \gamma - \delta$$

where the signs indicate orientations or directions in which the image crosses the given interval. We may construct such a map so that g expands all distances in the branched manifold Γ.

In the solenoid, a similar construction would have collapsed the B^2-directions (the leaves of the foliation of D) onto a circle (an unbranched manifold) on which the map g is simply $\theta \to 2\theta$. Since we understand the dynamics of $\theta \to 2\theta$ completely, we were able to use the inverse limit construction to analyze the solenoid as well. The same process works for the Plykin attractor.

For example, we may prove that $g: \Gamma \to \Gamma$ has dense periodic points as follows. Let I be any "subinterval" in Γ. Since g is expanding, it follows that there exists n such that $g^n(I)$ covers one of the four intervals $\alpha, \beta, \gamma, \delta$. Now one may check easily that there is an integer m such that $g^m(\xi) \supset \Gamma$ where ξ is any of the α, β, γ, or δ. Indeed, $g(\alpha) = \beta$, $g(\beta) \supset \gamma$, and $g(\gamma) = \alpha$ so that $g^3(\alpha) \supset \alpha$. Thus we conclude that $g^{m+n}(I) \supset \Gamma$ and so it follows that there is a periodic point in I. Using the inverse limit construction, one may then

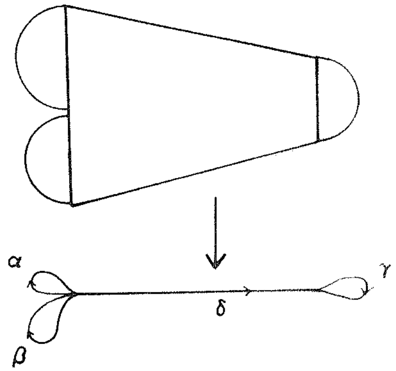

Fig. 5.7. The branched manifold for the Plykin attractor.

equate the action of P on Λ with that of the shift on $\Sigma = \Gamma \overset{g}{\leftarrow} \Gamma \overset{g}{\leftarrow} \Gamma \ldots$. We leave the details to the reader.

Remarks.

1. Much recent research has been devoted to the topic of "strange attractors." These are loosely defined as attractors which are topologically distinct from either a periodic orbit or a "limit cycle" (i.e., an invariant, attracting simple closed curve which arises often in ordinary differential equations). We prefer the term "hyperbolic" attractor for attractors like the solenoid and the Plykin example. Indeed, since we have succeeded in analyzing these maps completely, there is nothing whatsoever "strange" about them.

2. There are, however, some attractors which have thus far defied analysis. One of these is the Hénon attractor as described in Exercise 10. Numerical evidence indicates that this simple quadratic map of the plane possesses a transitive attractor, although this has never been proved rigorously. We urge the reader with access to computer graphics to plot successive iterates of a point under this map. The result is always qualitatively the same (disregarding the first few iterates) and always fascinating! We will return to this map in §2.9, where we will approach it from a different point of view.

Exercises

1. Construct a Markov partition for the solenoid.