CHAPTER 7

Stable Systems

We have, with the generalized stable manifold theorem, reached the border between elementary theory and recent research. In this final chapter we describe one direction in which the subject has developed. We have already aired the general philosophy behind this development in earlier chapters. The guiding idea is to find a family of dynamical systems which, on the one hand, contains "almost all" of them and, on the other hand, can be classified in some qualitatively significant fashion. For some years it was conjectured that structurally stable systems would fulfil both requirements. Although this turned out not to be the case except in very low dimensions, structural stability is such a natural property, both mathematically and physically, that it still holds a central place in the theory.

The two ingredients of structural stability are the topology given to the set of all dynamical systems and the equivalence relation placed on the resulting topological space. The latter is topological equivalence for flows, and topological conjugacy for diffeomorphisms. The former is the C^r topology (where $1 \leqslant r \leqslant \infty$). This topology is explained in detail in Appendix B, but the idea is clear enough. For example, two diffeomorphisms are C^r-close when their values, and also the values of corresponding derivatives up to order r, are close at every point. One may think of the derivatives as being defined via charts on the manifolds. Since the derivatives depend on the charts there is some ambiguity, but, for compact manifolds at any rate, this is unimportant provided we work with suitable finite atlases. Once we have defined the C^r topology we may be more specific about what we mean by "almost all" systems. Certainly an open dense subject of this topological space would be admirable for our purpose. Even the intersection of countably many such sets (called a *Baire* or *residual* subset†; think of the irrationals as a subset of the reals) would satisfy us.

† Such a subset would itself be dense, since the C^r-topology makes the space of dynamical systems a Baire space.

Let $\text{Diff}^r X$ $(1 \leqslant r \leqslant \infty)$ denote the set of all C^r diffeomorphisms of the smooth manifold X, provided with the C^r topology. We say that $f \in \text{Diff}^r X$ is *structurally stable* if it is in the interior of its topological conjugacy class†. That is to say, f is structurally stable if and only if any C^r-small perturbation takes it into a diffeomorphism that is topologically conjugate to f. Similarly, let $\Gamma^r X$ denote the set of all C^r vector fields on X topologized with the C^r topology. The vector field $v \in \Gamma^r X$ (or the corresponding integral flow ϕ) is *structurally stable* if it is in the interior of its topological equivalence class. The definition of structural stability is due to Andronov and Pontrjagin [1].

I. LOW DIMENSIONAL SYSTEMS

As the simplest possible example, consider a vector field v on S^1, with no zeros. Then there is precisely one orbit (exercise: why?) and this is periodic. We may identify TS^1 with $S^1 \times \mathbf{R}$. Let $\tilde{v} : S^1 \to \mathbf{R}$ be the principal part of v. Since $\tilde{v}$ is continuous and S^1 is compact $|\tilde{v}(x)|$ is bounded below by some constant $a > 0$. Any perturbation of v with C^0-size less than a does not introduce any zeros. Thus the perturbed vector field still has only one orbit, and is topologically equivalent to v. So v is structurally stable (C^0 structurally stable, in fact).

Now suppose that v has finitely many zeros, all of which are hyperbolic. We call such a vector field on S^1 *Morse–Smale* (a definition for higher dimensions follows later). Then there is an even number $2n$ of zeros, with sources and sinks alternating around S^1, as in Figure 7.1. The hyperbolic zeros are individually structurally stable (see Exercise 5.22), so a sufficiently

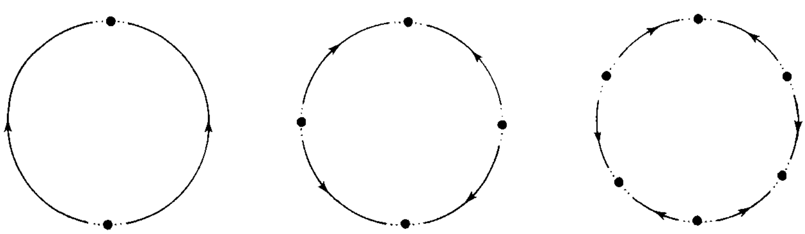


FIGURE 7.1

† Notice that this means that we have (apparently) more than one notion of structural stability for f, since we may also regard it as in $\text{Diff}^s X$ for any $s < r$. One ought, really, to talk of C^r structural stability, and we occasionally will. At the time of writing it is not known which, if any, of these notions are equivalent.

C^1-small perturbation of v leaves the orbit configuration unaltered on some neighbourhood of the zeros. But a sufficiently C^0-small perturbation introduces no further zeros outside this neighbourhood. Thus the orbits flow between the components of the neighbourhood as before, and the perturbed vector field is topologically equivalent to the original one. Thus v is C^1 structurally stable.

It is important to note that there are vector fields v which are not Morse–Smale but which nevertheless have the above orbit structures. In this case the principal part $\tilde{v}$ has zero derivative at one or more of the zeros, and the inward or outward motion near the zero is given by "higher order terms". One may also have vector fields on S^1 with "one-way" zeros, as in Figure 7.2. However, provided that v has only finitely many zeros and is not

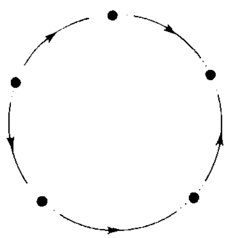


FIGURE 7.2

Morse–Smale, one can easily alter its orbit configuration. Choose a non-hyperbolic zero p. By a C^∞-small perturbation of the vector field near p, alter p to a source if it was not originally one, or to sink if it was originally a source. Since we leave all the other original zeros of v unaltered, we have changed the number of sources in either case. We have shown, then, that structurally stable vector fields on S^1 with finitely many zeros are Morse–Smale. However, a stronger result is true:

(7.3) Theorem. *A vector field on S^1 is C^1 structurally stable if and only if it is Morse–Smale. Morse–Smale vector fields form an open, dense subset of $\Gamma^r S^1$* $(1 \leqslant r \leqslant \infty)$.

Proof. We have proved sufficiency above. Let v be a vector field on S^1. Suppose that v is not the zero vector field. The tangent bundle TS^1 is isomorphic to the trivial bundle $S^1 \times \mathbf{R}$, so the image of v may be pictured as a curve on the cylinder (see Figure 7.3). The zeros of v are the points where the curve crosses the zero section, identifying the latter, as usual, with S^1. The zero is hyperbolic provided the curve is not tangent to the horizontal there.

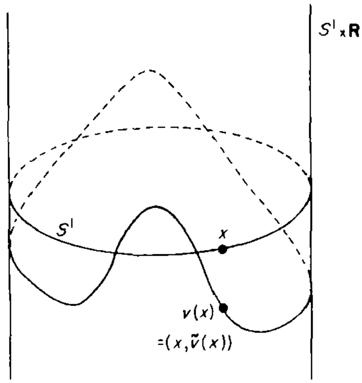


FIGURE 7.3

Now the image of the principal part $\tilde{v}$ of v is a closed subinterval $[a, b]$ with $a < b$. The critical points of $\tilde{v}$ (over which the curve is tangent to the horizontal) may form an infinite subset of S^1 but, by Sard's theorem†, their image under $\tilde{v}$ (the set of critical values) has measure zero in $\mathbf{R}$. Thus arbitrarily near $0 \in \mathbf{R}$ is a number c which is not a critical value of $\tilde{v}$. We perturb $\tilde{v}(x)$ to $\tilde{v}(x) - c$ and get a new vector field with only hyperbolic zeros. These, being isolated, are finite in number. Thus the perturbed vector field is Morse–Smale. We have proved the density part of the theorem, as the perturbation is arbitrarily C^∞-small. Note that if v were the zero vector field, we would first make a C^∞ small perturbation to get something else.

Now suppose that v is structurally stable. Then v is certainly not the zero vector field. Since v can be made Morse–Smale by an arbitrarily small perturbation, it is topologically equivalent to a Morse–Smale diffeomorphism. But we showed in the run-up to the theorem that this implies that v *is* Morse–Smale. □

The above theorem is a simple illustration of a fruitful approach to the classification problem. One attempts to prove that certain properties are *open and dense* (i.e. hold for systems in an open, dense subset of $\mathrm{Diff}^r X$ or $\Gamma^r X$) or *generic* (the same thing with "Baire" replacing "open, dense"), using *transversality* theory. This is a theory that investigates the way submanifolds of a manifold cross each other, and how a map of one manifold to another throws the first across a submanifold of the second. Sard's theorem is the opening shot in this theory; for further details see Hirsch's

† See Hirsch [1]. The proof in the present case is by dividing S^1 into n equal subintervals and letting $n \to \infty$. By Taylor's Theorem, if a subinterval contains a critical point of $\tilde{v}$ then its image under $\tilde{v}$ has a length which is bounded by a fixed multiple of the square of the length of the subinterval. Thus the total length of all such images tends to zero as $n \to \infty$.

book [1]. If a property is dense, then automatically a stable system is equivalent to a system with the property. One then anticipates, and attempts to prove, that the stable system itself has the property (since the property will often be one that is not shared by all equivalent systems). For example, the property of having only hyperbolic zeros is open and dense for flows on a compact manifold of any dimension, and any structurally stable flow possesses this property. The proof is essentially that for S^1 given above. Similarly for diffeomorphisms of a compact manifold X the following properties are generic, and satisfied by structurally stable diffeomorphisms:

(i) all periodic points are hyperbolic,

(ii) for any two periodic points x and y, the stable manifold of x and the unstable manifold of y intersect *transversally*.

(This means that the tangent spaces to the two submanifolds at any point p of intersection generate the whole tangent space of the manifold X at p: $T_pW_s(x)+T_pW_u(y)=T_pX$. See Figure 7.4. We denote this by the standard notation $W_s(x) \pitchfork W_u(y)$.) The above result is known as the Kupka–Smale theorem (Kupka [1], Smale [3], Peixoto [1]).

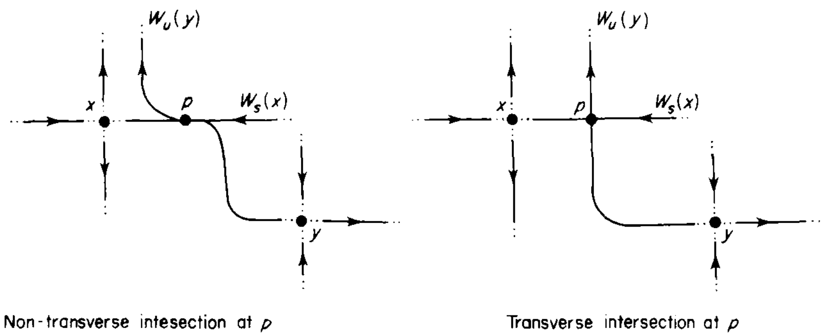


FIGURE 7.4

Theorem 7.3 also holds for vector fields on compact orientable 2-manifolds and for diffeomorphisms of S^1, a result due to Peixoto [2]. Of course we must extend our definition of the Morse–Smale property to cover these cases. In fact, we give a definition for dynamical systems on manifolds of any dimension. We say that a dynamical system is *Morse–Smale* if

(i) its non-wandering set is the union of a finite set of fixed points and periodic orbits, all of which are hyperbolic, and

(ii) if $G.x$ and $G.y$ are any two orbits of the non-wandering set then $W_s(G.x) \pitchfork W_u(G.y)$.

Thus in the case of a 2-dimensional flow the non-wandering set consists of at most five types of orbit, viz. hyperbolic sources, sinks and saddle points, and hyperbolic expanding and contracting closed orbits. The second condition is vacuous unless both x and y are saddle points, and, when they are, it reduces to the condition that $W_s(x)$ and $W_u(y)$ do not have a common orbit (apart from $\{x\}$ in the case $x = y$). One says informally "no orbit joins x to y". So the configurations in Figure 7.5 are not allowed in Morse–Smale flows. It is

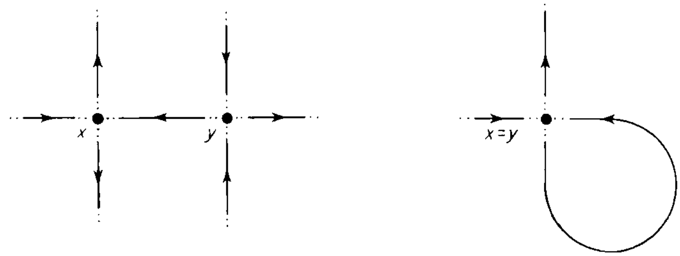


FIGURE 7.5

easy to get some idea of why condition (ii) should be necessary for structural stability. If it does not hold, we can take some point p on the offending orbit, and by making a slight (downwards in Figure 7.6) deflection of the vector

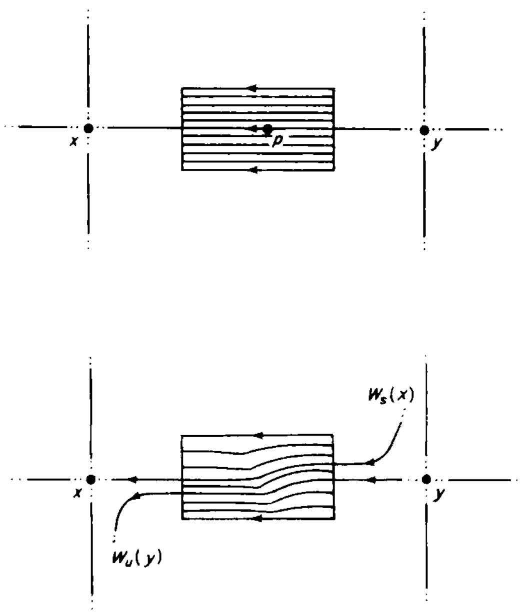


FIGURE 7.6

field in a thin flow tube containing p, we can divert $W_s(x)$ and $W_u(y)$ into different paths. Thus we have altered the orbit configuration. Or have we? It certainly looks as though we have in our picture, but it is dangerous to lean too heavily on pictures. For example we might, in our alteration of the vector field, join up two perfectly harmless orbits which happended to pass through the tube, and thus create a new orbit joining some other pair of saddle points, x' and y' say. In this case the new system might possibly be equivalent to the old one by a homeomorphism taking x' to x and y' to y. Or again, suppose that condition (i) were not in force and that there were infinitely many orbits joining pairs of saddle points. It is far from obvious that obliterating one of them would alter the topology of the picture as a whole. In fact we have to be a bit careful, and we have to assume that condition (i) holds. We take a neighbourhood U of the non-wandering set sufficiently small not to contain p. Since there are only finitely many *separatrices* (as we call orbits that begin or end at a saddle point) and since each is in U for all but a bounded period of time, each will pass through a flow tube about p in the complement of U only finitely many times. By making the tube thinner, we can ensure that $\mathbf{R}.p$ passes through the tube only once and the other separatrices miss it altogether. If we perform the perturbation in this tube we do indeed decrease the number of saddle point connections by one, and hence alter the orbit configuration topologically.

To show that a 2-dimensional Morse–Smale vector field v is structurally stable, one builds up a topological equivalence between it and any small perturbation of it first on a neighbourhood U of the non-wandering set of v (on which v is structurally stable by hyperbolicity), then on the components outside U of the separatrices (these have not been moved far by the perturbation and so still join corresponding components of U) and finally on the homogeneous looking blocks of orbits that remain. We shall not embark on the details of this programme, nor deal with the density problem. This is rather technical and special to two dimensions, and we refer the reader to Peixoto's paper [2]. We just quote the result:

(7.7) Theorem. *A vector field on a compact orientable 2-manifold X (resp. a diffeomorphism of S^1) is C^1 structurally stable if and only if it is Morse–Smale. Morse–Smale systems are open and dense in $\Gamma^r(X)$ (resp. $\mathrm{Diff}^r S^1$) for $1 \leqslant r \leqslant \infty$.*

The sufficiency of the conditions in the case when X is a 2-dimensional disc, was first suggested by Andronov and Pontrajagin [1] and proved for an analytic vector field by De Baggis [1]. In a recent paper [3] Peixoto has developed his theorem into a classification of structurally stable systems. The theorem for non-orientable 2-manifolds is still an open problem, except for some cases which have been proved by Gutierrez [1].

II. ANOSOV SYSTEMS

For some time it was hoped that the above results might generalize to higher dimensions. Unfortunately neither the characterization nor the density of structurally stable systems goes over. It is, then, a relief to find that, for all dimensions:

(7.8) Theorem. *A Morse–Smale system on a compact manifold is C^1 structurally stable.*

In fact, it is doubly welcome, because, since every compact manifold admits a Morse–Smale system†, every compact manifold admits a structurally stable system. This was an early conjecture which turned out to be hard to prove, and survived until Palis and Smale's paper [1].

By the time this paper appeared it was known that there were other structurally stable systems besides Morse–Smale systems. The first of these to emerge were the toral automorphisms (see Example 1.30). These are structurally stable, but their non-wandering sets are the whole of the tori on which they are defined, so they certainly fail (i) of Morse–Smale (in fact, they are the opposite extreme to Morse–Smale systems in this sense!). Toral automorphisms were first put forward by Thom as a counterexample to the density of Morse–Smale diffeomorphisms. The justification of this is by proving them structurally stable, and this was first done by Anosov. His proof was for a wider class of systems which we now define.

A diffeomorphism $f: X \to X$ of a manifold X is *Anosov* if X has a hyperbolic structure with respect to f. Recall that this means that the tangent bundle TX splits continuously into a Tf-invariant direct-sum decomposition $TX = E_s \oplus E_u$ such that Tf contracts E and expands E_u (with respect to some Riemannian metric on X). Trivially hyperbolic linear maps f possess this property, since one has the identification $T\mathbf{R}^n = \mathbf{R}^n \times \mathbf{R}^n$ and $Tf(x, v) = (f(x), f(v))$, so that the splitting of $\mathbf{R}^n$ into the stable and unstable manifolds of f, W_s and W_u say, gives a splitting of $T\mathbf{R}^n$ as $(\mathbf{R}^n \times W_s) \oplus (\mathbf{R}^n \times W_u)$. In the case of toral automorphisms this splitting is carried over to the torus when we make the identification, and so toral automorphisms are Anosov. Similarly a vector field on X is *Anosov* if X has a hyperbolic structure with respect to it, and as examples of such we have all suspensions of Anosov diffeomorphisms. The main result due originally to Anosov [1, 2] is:

(7.9) Theorem. *Anosov systems on compact manifolds are C^1 structurally stable.*

† Take the unit time map of the gradient flow of almost any smooth real function on the manifold (see Smale [1]).

The diffeomorphism case of the theorem has a nice functional analytic proof due to Moser [1] (see also Franks [1] or Nitecki [1]). Recall, from the preamble to Theorem 6.21, that the set $C^0(X, X)$ of all C^0 maps of the compact manifold X to itself has the structure of a Banach manifold modelled on $\Gamma^0(X)$. Let $f: X \to X$ be an Anosov diffeomorphism and let $g: X \to X$ be C^1-near f. As in Chapter 6, one considers the map $f^{\#}: C^0(X, X) \to C^0(X, X)$ sending h to fhf^{-1}. Let $G: C^0(X, X) \to C^0(X, X)$ be the C^1-nearby map sending h to ghf^{-1}. Since, as in Chapter 6, $f^{\#}$ has a hyperbolic fixed point at the identity id_X, G has a fixed point, h say, near id_X. So $h = ghf^{-1}$, or, equivalently, $hf = gh$. Now the map f is *expansive*, meaning that the orbits of different points of X are never close together. More precisely there is a number $a > 0$ such that $d(f^n(x), f^n(y)) < a$ for all $n \in \mathbf{Z}$ implies $x = y$. This is clear from the fact that the map $\phi: C_b(P) \to C_b(TX)$ defined as in the proof of Theorem 6.21 has a hyperbolic fixed point at 0, so that only iterates of 0 under ϕ stay near 0. Suppose that $h(x) = h(y)$. Then $hf^n(x) = g^n h(x) = g^n h(y) = hf^n(y)$ for all $n \in \mathbf{Z}$. Since h is near id_X, $f^n(x)$ is near $f_n(y)$ for all $n \in \mathbf{Z}$, and thus since f is expansive, $x = y$. This shows that h is injective. But it is a standard result of algebraic topology that a continuous injection $h: X \to X$ of a compact manifold is a homeomorphism. Thus f is structurally stable.

It is worth while getting as much insight as possible into the effects of hyperbolic structure, and so we give a more down-to-earth description of the existence of h and the expansiveness of f (due to Robinson [4]). The existence of a hyperbolic structure on X implies that, for any $x \in X$, $T_{f^{-1}(x)}f$ throws a product ε-disc neighbourhood $D_s \times D_u$ of 0 in $T_{f^{-1}(x)}X$ across a similar ε-disc neighbourhood of 0 in T_xX as shown in Figure 7.10. This

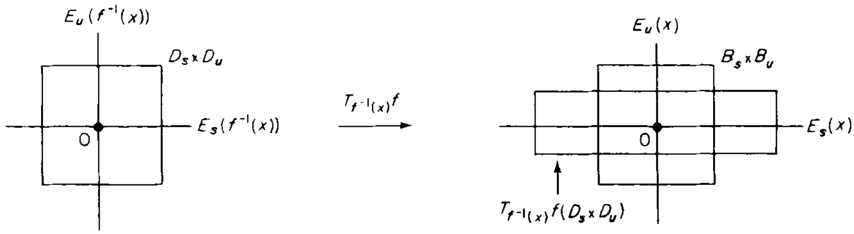


FIGURE 7.10

behaviour is echoed by f in the manifold X, provided that ε is small; it throws $D_{f^{-1}(x)} = \exp_{f^{-1}(x)}(D_s \times D_u)$ across $D_x = \exp_x(B_s \times B_u)$, as shown in Figure 7.11. We have similar product neighbourhoods D_y for all $y \in X$. We observe that f throws $D_{f^{-2}(x)}$ across $D_{f^{-1}(x)}$ as shown in Figure 7.11, and hence that $\bigcap\{f^n(D_{f^{-n}(x)}): 0 \leq n \leq 2\}$ is as shown (shaded). It is, of course, possible that

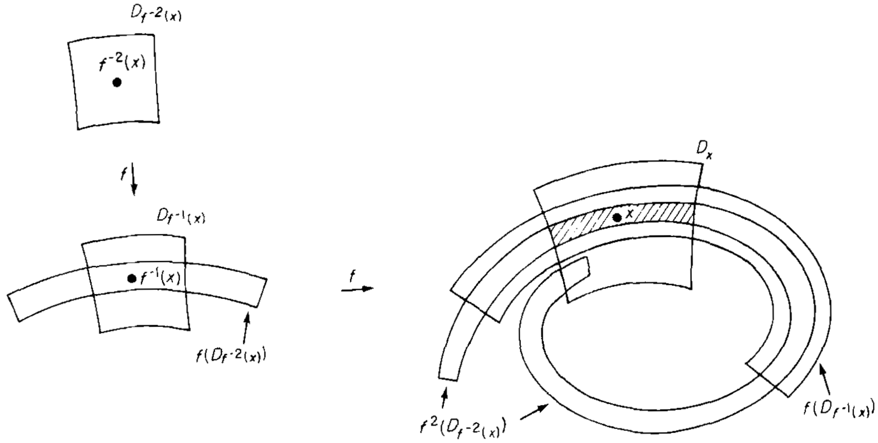


FIGURE 7.11

$f^2(D_{f^{-2}(x)})$ has further intersections with D_x, and we have attempted to indicate this in the diagram. If g is C^1-close to f then we have a similar picture replaced by $g^n(D_{f^{-n}(x)})$. It is fairly clear from the way that the (s-) height of the sets $D_x^m(g) = \bigcap\{g^n(D_{f^{-n}(x)}): 0 \leqslant n \leqslant m\}$ decreases, that $D_x^\infty(g)$ is a disc stretching across D_x in the u-direction, as shown in Figure 7.12.

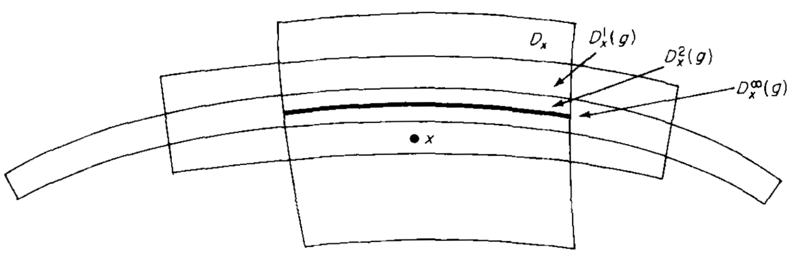


FIGURE 7.12

Since $D_x^\infty(g)$ in the set of points of D_x whose negative g-iterates stay near their corresponding f-iterates, g^{-1} maps $D_{f(x)}^\infty(g)$ into $D_x^\infty(g)$. Moreover it contracts it, since the discs are in the u-direction. Thus $\bigcap\{g^{-n}(D_{f^n(x)}^\infty(g)): n \geqslant 0\}$ is a single point, which we denote $h(x)$. We observe that $h(x)$ is the unique point of D_x such that every g-iterate of it stays near the corresponding f-iterate of x. Expansivity of f follows by putting $g = f$. In general, since $g^n(gh(x)) \in D_{f^{n+1}(x)} = D_{f^n f(x)}$ for all $n \in \mathbf{Z}$, we deduce that $hf(x) = gh(x)$.

There are, at the time of writing, several unsolved problems about Anosov diffeomorphisms. For example, is their non-wandering set always the whole manifold? Do they always have a fixed point? Not all manifolds admit Anosov diffeomorphism (in contrast to Morse–Smale diffeomorphisms). Do all n-dimensional manifolds which do admit them have $\mathbf{R}^n$ as universal covering space? One positive thing that we can say is that Anosov diffeomorphisms of tori are well understood; they are all topologically conjugate to the toral automorphisms (see Manning [1]).

III. CHARACTERIZATION OF STRUCTURAL STABILITY

The main problem now facing us is how to characterize structural stability, bearing in mind that we have to reconcile such apparently dissimilar systems as Morse–Smale and Anosov systems. The link comes when one recognizes, as Smale did, that if we generalize the Morse–Smale definition by replacing the term "closed orbits" by "basic sets" (definition shortly) then Anosov systems, and others as well, are allowed in. For example, toral automorphisms have only one basic set, the whole torus, so they trivially satisfy the new conditions. We say that a dynamical system has an *Ω-decomposition* if its non-wandering set Ω is the disjoint union of closed invariant sets Ω_1, $\Omega_2, \ldots, \Omega_k$. If the system is *topologically transitive* on Ω_i (that is, Ω_i is the closure of the orbit of one of its points) for all i, we say that $\Omega_i \cup \ldots \cup \Omega_k$ is a *spectral decomposition*, and that the Ω_i are basic sets. One could define the concept of a basic set in isolation by saying that a closed invariant set, $\Lambda \subset \Omega$ is *basic* if the system is topologically transitive on Λ but Λ does not meet the closure of the orbit of $\Omega\backslash\Lambda$. Note that a basic set is *indecomposable*, in that it is not the disjoint union of two non-empty closed invariant sets. In general there is no reason why the non-wandering set of a system should have a spectral decomposition. For example in the flow illustrated in Figure 7.13, the figure eight consisting of two orbits joining a saddle point is an

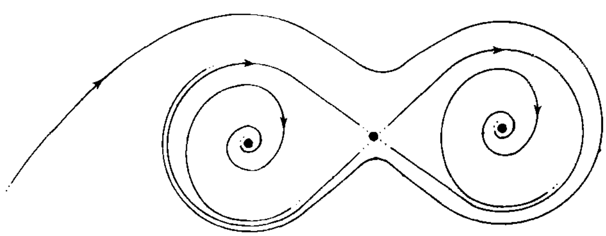


FIGURE 7.13

isolated part of the non-wandering set, but there is no way of splitting it into basic sets. However, Smale [5, 6] made the following definition:

A dynamical system is (or satisfies) *Axiom A* if its non-wandering set

(a) has a hyperbolic structure, and

(b) is the closure of the set of closed orbits of the system†,

and proved a fundamental theorem:

(7.14) Theorem. (*The Spectral Decomposition Theorem*) *The non-wandering set of an Axiom A dynamical system on a compact manifold is the union of finitely many basic sets.*

Proof. To get some idea of how Axiom A brings about this decomposition, let us examine the local structure of the non-wandering set Ω of an Axiom A diffeomorphism f of a compact manifold X. We know that Ω has a generalized stable manifold system, and the first point to realize is that Ω is given locally as the set of vertices of a sort of grid formed by the family of stable manifolds crossing the family of unstable manifolds (this is the so-called *local product structure* of Ω). To see this, note that if x and y are nearby points of Ω then, locally, by continuity of the unstable and stable manifold systems, $W_s(x)$ intersects $W_u(y)$ transversally in a single point z, and similarly $W_u(x)$ intersects $W_s(y)$ transversally in a single point t. We assert that z and t are in Ω (a slightly more general result is known as the *cloud lemma*). By Axiom A(b), continuity of the stable manifolds system and the fact that Ω is closed, we may assume that x and y are periodic, or even fixed (since periodic points are fixed by some power of f). Under positive iterates of f, any neighbourhood U of z presses up against and is spread out along $W_u(x)$, until it intersects $W_s(y)$ (see Figure 7.14). It is then pressed against $W_u(y)$ and spreads out along it until it eventually intersects U.

Also note that if N is a small open neighbourhood of x in Ω and if U is a non-empty open subset of N then the orbit of U is dense in N. For, given

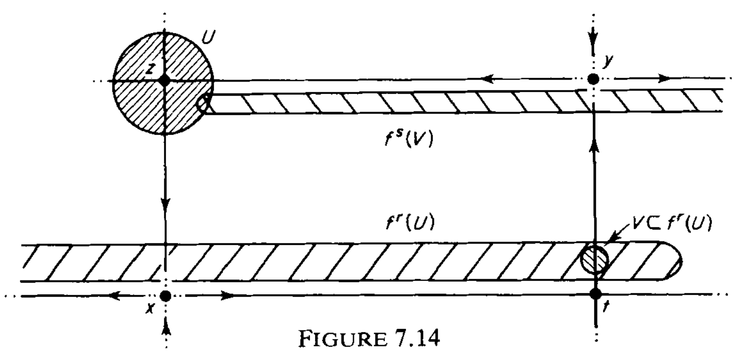


FIGURE 7.14

† It was for several years conjectured that A(a) implied A(b). This conjecture is now known to be true for diffeomorphisms of two-dimensional manifolds (Newhouse and Palis [1]) but false for diffeomorphisms of all higher dimensional manifolds (Dankner [1]).

$p \in N$, we may, as above, assume p fixed, and also assume that there is a fixed point q in U. Then $W_u(q)$ intersects $W_s(p)$ in a point z which is, by local product structure, in Ω. Some negative power of f takes z into U, and some positive power takes it arbitrarily near p. Thus p is in the closure of the orbit of U.

We now define a relation $\sim$ on Ω by $x \sim y$ *if and only if* y *is in the closure of the orbit of every neighbourhood of* x *in* Ω. We assert that $\sim$ is an equivalence relation. Reflexivity is trivial. Suppose $x \sim y$. Let P be any open neighbourhood of y in Ω, and let N be a small open neighbourhood of x in Ω (i.e. small enough for the density property mentioned in the previous paragraph to hold). Since $x \sim y$, P intersects some $f^r(N)$, and hence $N \cap f^{-r}(P)$ is a non-empty set, U say. Since the orbit of U is dense in N, so is the orbit of P, and thus $y \sim x$. See Figure 7.15. Similarly, suppose that $x \sim y$ and $y \sim z$. Let

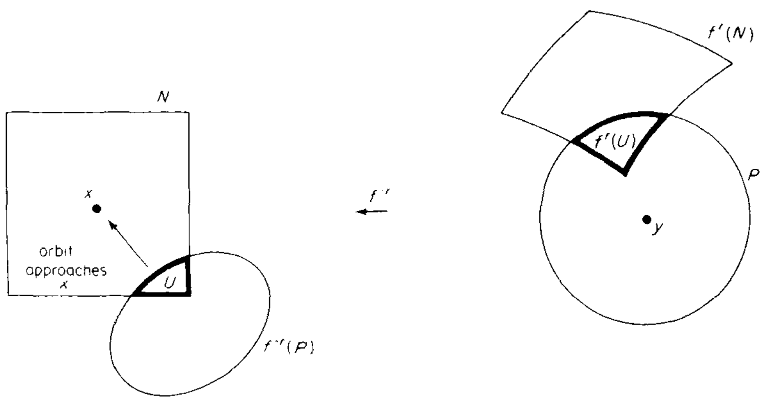


FIGURE 7.15

P be any neighbourhood of x in Ω, and let N_1 and N_2 be small neighbourhoods in Ω of y and z respectively. Then since $y \sim z$, some $f^r(N_1)$ intersects N_2, and hence $N_1 \cap f^{-r}(N_2)$ is a non-empty open subset, V say. Also, since $x \sim y$, some $f^s(P)$ intersects N_1 in a non-empty open set U_1, say. Since the orbit of U_1 is dense in N_1, some $f^t(U_1)$ intersects V. Thus $f^{r+t}(U_1)$ intersects N_2 in a non-empty open subset U_2. Since the orbit of U_2 is dense in N_2, so is the orbit of P. Thus $x \sim z$. See Figure 7.16.

Let Ω_i be the equivalence classes of Ω under $\sim$. By the proof of symmetry, we get a typical Ω_i by taking any open set N_i on Ω small enough to have the density property and taking the closure of its orbit. Thus there are only finitely many Ω_i, because we can take a finite covering of the compact set Ω by small enough open sets N_i, and, of course, the associated Ω_i, some of which may coincide, cover Ω. The Ω_i are, by the above description, closed

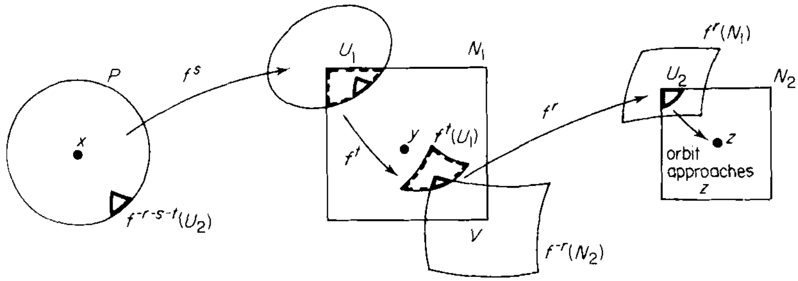


FIGURE 7.16

invariant subsets. We need finally to show that each Ω_i is topologically transitive, and this is a standard deduction from the fact that it contains an open set U_i with dense orbit. For, let V_j be a countable basis for the topology of Ω_i, and let O_j be the orbit of V_j. Then O_j is open and dense in Ω_i (because Ω_i is an $\sim$-equivalence class), and hence $\bigcap_j O_j = D$ (say) is residual, and hence dense in Ω_i. Let $x \in D$. Then the orbit of x is dense in Ω_i, because any open subset V of Ω_i contains some V_j, and $x \in$ some $f^n(V_i)$, so $f^{-n}(x) \in V_i \subset V$. $\square$

To visualize a general Axiom A system, then, one thinks of a system with finitely many fixed points and periodic orbits, all hyperbolic (such as the gradient of the height function on the torus, or a Morse–Smale system, for example Figure 7.17 for S^2), but in higher dimensions and with the fixed

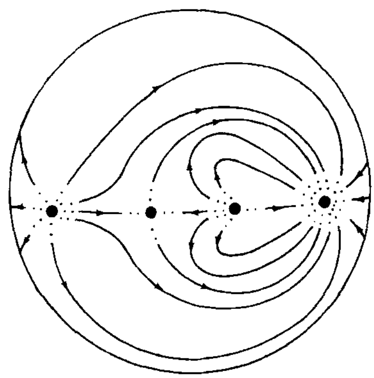


FIGURE 7.17

points and periodic orbits replaced by more general basic sets. The basic set could, for example, be a torus on which the system (a diffeomorphism) restricts to a toral automorphism. For example, if one takes the product of a toral automorphism on $S^1 \times S^1$ with the (North pole)–(South pole) diffeomorphism of S^1, one gets an Axiom A diffeomorphism of $S^1 \times S^1 \times S^1$ with two basic sets, $S^1 \times S^1 \times \{\text{North pole}\}$ and $S^1 \times S^1 \times \{\text{South pole}\}$. There is, however, no need for a basic set to be a submanifold. We shall describe two examples for which it is not. Such basic sets are usually termed *strange* or *exotic*, and a good deal of work has been done on their structure (see, for example, Williams [1, 3], and Sullivan and Williams [1]).

The first of these is the *expanding attractor* described by Smale [5]. An *attractor* of a diffeomorphism is a subset Λ that has a compact neighbourhood N such that $f(N) \subset \operatorname{int} N$ and $\bigcap_{r \geq 0} f^r(N) = \Lambda$. For a flow ϕ the corresponding conditions are $\phi^t(N) \subset N$ for all $t \supset 0$ and $\bigcap_{t \geq 0} \phi^t(N) = \Lambda$. We visualize a map f of the solid torus $S^1 \times B^2$ into itself, stretching the S^1 factor to approximately twice its length, contracting the B^2 factor into something under half its width, and wrapping it twice around the target copy, as shown in Figure 7.18. We insist that each cross-sectional 2-disc $\{[\theta]\} \times B^2$

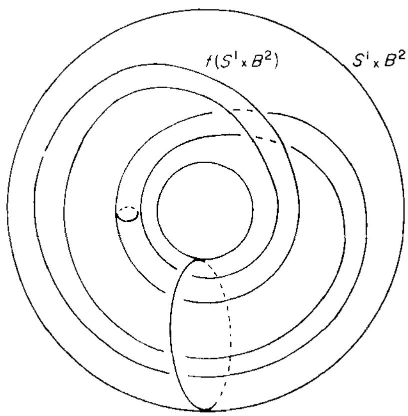


FIGURE 7.18

is mapped into $\{[2\theta]\} \times B^2$. Imagine the domain $S^1 \times B^2$ as embedded in a 3-ball B^3 in some 3-manifold X. Certainly f extends to a diffeomorphism of X. One could even produce a deformation of X, fixed outside B^3, which would slide $S^1 \times B^2$ into $f(S^1 \times B^2)$ so that x finishes up at $f(x)$. Thus the phenomenon that we are about to describe is by no means pathological. It may be met with locally in diffeomorphisms of any manifold of dimension greater than two.

Consider $f^2(S^1 \times B^2)$. This is a thinner longer tube which winds twice around $f(S^1 \times B^2)$. Similarly $f^3(S^1 \times B^2)$ is thinner and longer still, and winds twice $f^2(S^1 \times B^2)$ (so eight times around $S^1 \times B^2$). Any point x of $S^1 \times B^2$ that is outside $f^r(S^1 \times B^2)$ for some $r > 0$ has a neighbourhood U (in $S^1 \times B^2$) that is also outside $f^r(S^1 \times B^2)$. Thus $f^s(U)$ does not intersect U for $s \geq r$, and hence x is wandering. Thus, if we are interested in the non-wandering set of f in $S^1 \times B^2$, we need only look at $\Lambda = \bigcap_{r \geq 0} f^r(S^1 \times B^2)$. If we look at the intersections of $f^r(S^1 \times B^2)$ with a single cross section $\{[\theta]\} \times B^2$ (Figure 7.19), we notice a striking resemblance to the well-known

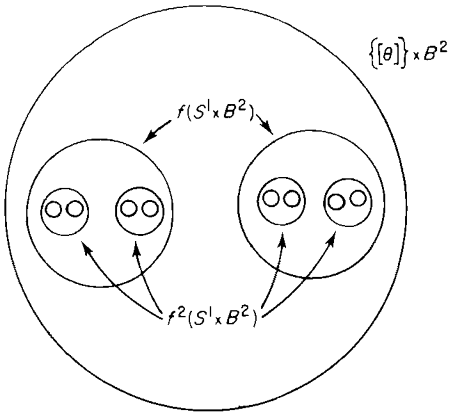


FIGURE 7.19

construction of the Cantor set from an interval by repeatedly deleting middle thirds. Thus it is no surprise to find that $\Lambda \cap (\{[\theta]\} \times B^2)$ is homeomorphic to the Cantor set. Technically speaking, Λ is a fibre bundle over S^1 with projection the product projection of $S^1 \times B^2$ and fibre the Cantor set.

It is not hard to see that Λ is, in fact, the non-wandering set of f on $S^1 \times B^2$. Let $x \in \Lambda$, and let U be any neighbourhood of x. It is clear that any segment $[a, b] \times B^2$ of $S^1 \times B^2$, however short, will, under a sufficiently large iterate f^s, wrap once round $S^1 \times B^2$ (i.e. its image will intersect every disc $\{[\theta]\} \times B^2$). A similar remark holds for any of the image tubes $f^r(S^1 \times B^2)$. We may choose a segment V of some such tube with $x \in V \subset U$. By the remark, for some $s > 0$, $f^s(V)$ intersects every cross-sectional disc of V. Thus $f^s(U) \cap U$ is non-empty, which proves that x is non-wandering. A similar argument shows that any open subset of Λ has its orbit dense in Λ, which implies that Λ is topologically transitive. Since Λ is locally like the product of a real interval and the Cantor set, we can talk of the 1-manifold part of Λ at x, meaning the fibre of the projection onto the Cantor set. The hyperbolic structure of Λ is clear: the stable subspace at $x = ([\theta], p)$ is the tangent to the cross-sectional 2-disc at x, and the unstable manifold is the tangent to the 1-manifold part at x. This expanding behaviour on the 1-manifold part is the reason we call the attractor *expanding*.

We now describe our second exotic basic set, the *Smale Horseshoe* (Smale [5]). Again the example is for a diffeomorphism, but would yield a basic set of a flow by suspension. It is fundamentally a 2-dimensional phenomenon, but is, perhaps, rather harder to grasp than the expanding attractor. This time we take a solid 2-dimensional square S, and four subrectangles, two vertical, R_0 and R_1, and two horizontal, ${}_0R$ and ${}_1R$ as shown (Figure 7.20).

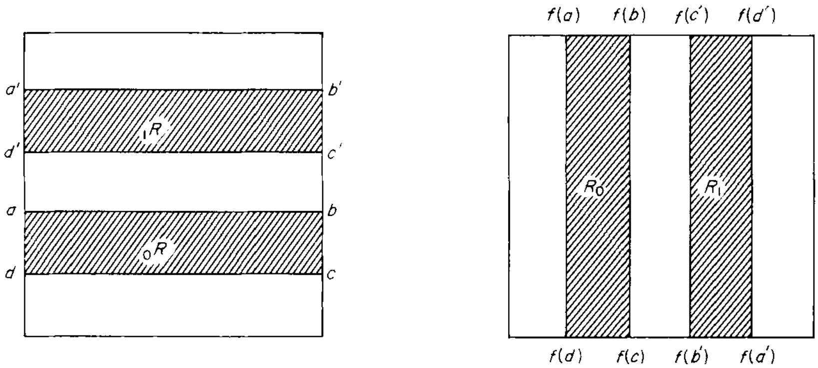


FIGURE 7.20

We map the vertices of ${}_0R$ to those of R_0 as shown, and extend linearly for a map f of ${}_0R$ to R_0. Similarly for ${}_1R$. *Note the half turn* in this second case. The map f so far defined is clearly hyperbolic. We now extend f to the whole square S, mapping it onto the shaded horseshoe shape in Figure 7.21 (whence the name of the example). We make sure that $f(S)$ is contained in the interior of $C = D_0 \cup S \cup D_1$, where D_0 and D_1 are half discs on the horizontal edges of S as shown. We then extend f over D_0 and D_1, mapping them into int D_0. We may actually arrange for $f|D_0$ to be a contraction, so that it has a unique attracting fixed point p. At this stage, f has a basic set of the type that we wish to consider. Notice again that, since we could produce f by deforming some neighbourhood of C in the plane, we can consider it as embedded locally in any 2-manifold X, as part of a diffeomorphism $f: X \to X$. We suppose this done.

We are interested in locating $\Omega(f) \cap S$. We observe that if a point $s \in S$ is not in ${}_0R$ or ${}_1R$ then it is not in $\Omega(f)$. This is because some neighbourhood U of it in S gets mapped into D_0 or D_1 by f, so into D_0 by f^2, and so into D_0 by

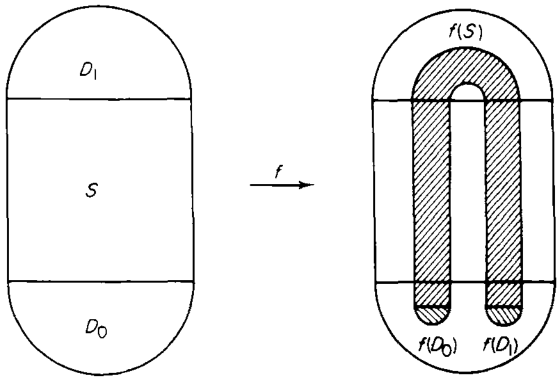


FIGURE 7.21

any higher power of f. Hence $U \cap f^r(U)$ is empty for $r > 0$. Similarly if x is not in $R_0 \cup R_1$, then some neighbourhood of it came from outside C under f, and its iterates under f^{-1} are all outside C, since $f(C) \subset C$. Summing up, if $x \in \Omega(f) \cap S$ then $x \in {}_iR_j = {}_iR \cap R_j$, for some i, j. Let $S_1 = \bigcup {}_iR_j$.

We now make a similar observation for f^2. Let $R_{ij} = f({}_iR_j)$ and ${}_{ij}R = f^{-1}({}_iR_j)$ (see Figure 7.22). Note that $R_{ij} \subset R_i$ and ${}_{ij}R \subset {}_jR$. If $x \in S$ is outside

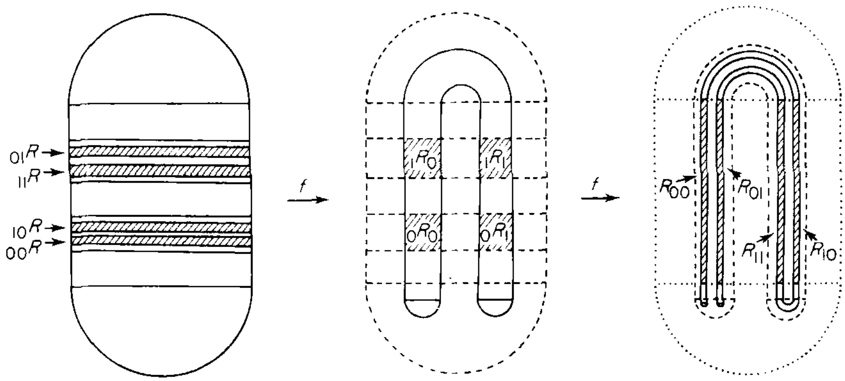
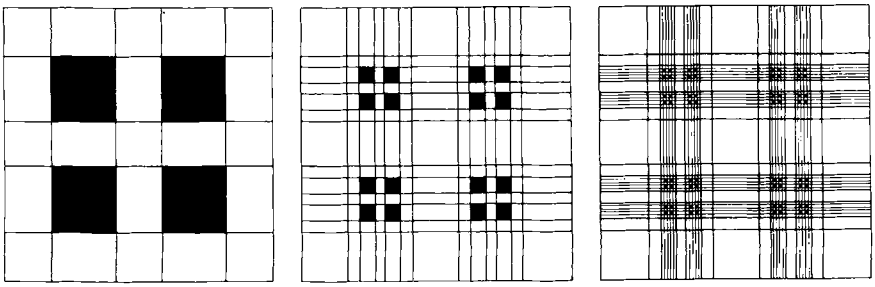


FIGURE 7.22

$\bigcup {}_{ij}R$, then so is some neighbourhood of x in S, f^2 takes the neighbourhood into $D_0 \cap D_1$ and further iterates of f keep it there. Hence $x \notin \Omega(f)$. Similarly $x \in S \backslash \bigcup R_{ij}$ is not in $\Omega(f)$. Hence $\Omega(f) \cap S \subset \bigcup({}_{ij}R \bigcap R_{kl}) = S_2$.

Now define ${}_iR_{jk} = {}_iR \cap R_{jk}$ and ${}_{ij}R_k = {}_{ij}R \cap R_k$. Notice that $f({}_{ij}R_k) = {}_iR_{jk}$. Define ${}_{ijk}R = f^{-1}({}_{ij}R_k)$ and $R_{ijk} = f({}_iR_{jk})$. Notice that $R_{ijk} \subset R_{ij}$ and ${}_{ijk}R \subset$

${}_{jk}R$. Repeat the above argument, showing that $\Omega(f) \cap S \subset \bigcup({}_{ijk}R \cap R_{lmn}) = S_3$ (see Figure 7.23). It is clear that, with a bit of care, we could define S_r



S_1 S_2 S_3

FIGURE 7.23

inductively, for all $r > 1$, and, hence, $S_\infty = \bigcap_{r \geq 1} S_r$ with the property that $\Omega(f) \subset S_\infty$. In fact S_∞ is homeomorphic to the Cantor set. There is a standard representation A of the Cantor set as infinite bisequences $a = \ldots a_{-2}a_{-1}a_0 . a_1a_2a_3 \ldots$ where each a_i is 0 or 1. If $U_n(a)$ is the set of all such bisequences that agree with a for n terms on either side of the decimal point then $\{U_n(a): n \geq 1\}$ is a basis of open neighbourhoods of a in A. We have a homeomorphism from A to S_∞ which sends the bisequence a to the (unique) point in $\bigcap_{n \geq 0} ({}_{a_{-n+1} \ldots a_0}R \cap R_{a_1 \ldots a_n})$. Identifying S_∞ with A by this homeomorphism, it is clear that $f|S_\infty$ is the well known *shift* automorphism of A, which moves the decimal point one place to the left (i.e. $f(a)_i = a_{i-1}$ for all $i \in \mathbf{Z}$). Notice that $U_n(a)$ contains many periodic points of f, for example the one whose entries are the $2n$ terms $a_{-n+1} \ldots a_n$ repeated as a block time after time. This shows that $S_\infty \subset \overline{\operatorname{Per} f}$, and hence that $S_\infty \subset \Omega(f)$, so that $S_\infty = \Omega(f) \cap S$. Also there is a bisequence c which contains every possible finite sequence as a block of consecutive entries, so that, for every $a \in A$ and $n > 0$, $f^r(c) \in U_n(a)$ for some $r \in \mathbf{Z}$. Thus the orbit of c is dense in S_∞. Finally note that the hyperbolic structure for S_∞ is induced from $f|({}_0R \cup {}_1R)$; the stable summands are horizontal and the unstable summands vertical.

Returning to the question of characterizing structural stability, and following Smale, we make another definition. We say that a system is *AS* if it satisfies both Axiom A and the *strong transversality* condition: for all x and y in the non-wandering set of the system, the stable manifold of the orbit of x intersects the stable manifold of the orbit of y transversally. This latter condition is, then, just the general version of the second condition in the definition of Morse–Smale systems. The best set of criteria of structural stability that has so far emerged is due to Robbin [1] (for C^2 diffeomorphisms) and, later, to Robinson [2, 3, 4] (for C^1 diffeomorphisms and flows).

They prove

(7.24) Theorem. *Any AS system is C^1-structurally stable.*

It seems likely that the converse of this theorem is also true, which would characterize structural stability very satisfactorily. At the time of writing the nearest approach to this is due to Franks [3] (see also Robbin [2]). A diffeomorphism $f: X \to X$ is *absolutely C^1-structurally stable* if for some C^1-neighbourhood N of f, there is a map σ (called a *selector*) associating with $g \in N$ a homeomorphism $\sigma(g)$ of X such that

(i) $\sigma(f) = id_X$,

(ii) for all $g \in N$, $g = hfh^{-1}$, where $h = \sigma(g)$,

(iii) σ is Lipschitz at f with respect to the C^0 metric d (i.e. for some $\kappa > 0$ and all $g \in N$, $d(\sigma(g), id_X) \leq \kappa d(g, f)$).

Then:

(7.25) Theorem. *Any diffeomorphism is absolutely C^1 structurally stable if and only if it is AS.*

It is known that structural stability is equivalent to AS when $\Omega(f)$ is finite (Palis and Smale [1]) and that structural stability and Axiom A imply strong transversality (Smale [5]). It is also known that a structurally stable system is *weak Axiom A* (which is Axiom A with the hyperbolic structure condition relaxed for non-periodic points of $\Omega(f)$); this uses Pugh's celebrated C^1 closing lemma (Pugh [1, 2]).

IV. DENSITY

It is in attempting to generalize the second part of Peixoto's theorem, which deals with the density of Morse–Smale systems, that things go disastrously wrong. Obviously Morse–Smale systems are not dense in higher dimensions, since there are, as we have seen, other types of structurally stable systems. One might, nevertheless, have hoped that structurally stable systems are dense. This is not the case in higher dimensions, and we are now in a position to see why. The two ingredients for a counterexample are an exotic basic set (because such a set contains arbitrarily close but topologically distinct orbits, viz. periodic and non-periodic orbits) and a failure of the strong transversality condition. The idea is to construct a system where the unstable manifold of a fixed point has a point of tangency with a member of the stable manifold system of an exotic basic set Λ. Roughly speaking, if the dimensions are right, nearby systems exhibit the same tangency, but the stable manifold in question neither consistently contains nor consistently avoids periodic points of Λ. More precisely, one ensures that the stable manifolds of periodic points of Λ are dense. Thus systems for which they are tangent to the unstable manifold of the fixed point are dense near the given system, whereas, by the Kupka–Smale theorem, the stable and unstable manifolds of periodic points intersect transversally for a structurally stable system.

The above idea was originally due to Smale [4], who constructed examples for diffeomorphisms of compact 3-manifolds and flows on compact 4-manifolds. Peixoto and Pugh [1], by similar methods, showed that structurally stable systems are not dense on any non-compact manifold of dimension $\geqslant 2$. Finally Williams [2] completed the picture by reducing the dimensions of Smale's examples by 1. Thus, for instance, structurally stable diffeomorphisms are not dense on the two-dimensional torus T^2. One starts with a toral automorphism f on T^2, and modifies it in a neighbourhood of the fixed point 0 by composing with a diffeomorphism that is the identity outside some smaller neighbourhood, preserves the y-coordinate (the unstable manifold direction) but near 0 expands strongly in the x-direction (the stable manifold direction). This has the effect of turning 0 into a source and introducing two new saddle points, one each side of it. One might almost imagine that the old unstable manifold had been split in two lengthwise and the unstable manifold of a source grafted into the hole (see Figure 7.26).

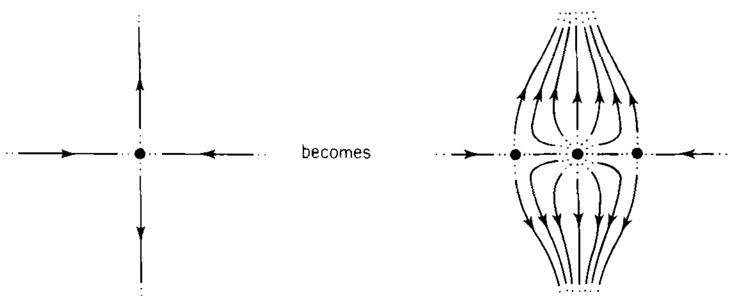


FIGURE 7.26

At this stage the system is called a *DA* (*derived from Anosov*) system. It has two basic sets, the source 0 and an exotic 1-dimensional expanding attractor Λ. The stable manifold system is still the stable manifold system of the original Anosov diffeomorphism (horizontal lines in Figure 7.26) except that one of them is broken in two at 0. We now remove a neighbourhood of the source at 0, and replace it by a plug consisting of two sources and a saddle point with the latter at 0 (see Figure 7.27). This modifies the stable manifold

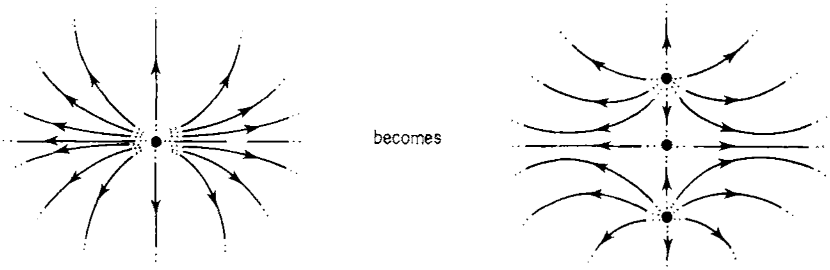


FIGURE 7.27

system of Λ to give Figure 7.28. At the moment the unstable manifold of the saddle point 0 is horizontal. We now perturb the diffeomorphism (call it g now) slightly in the neighbourhood of some point p of the unstable manifold

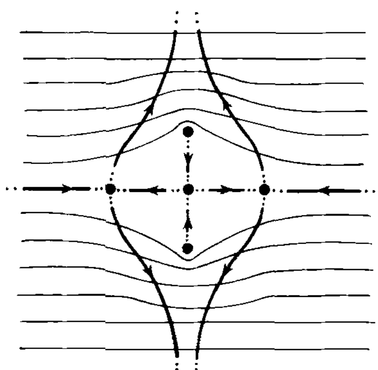


FIGURE 7.28

to produce a kink in it at $g(p)$ as shown in Figure 7.29. The diagram is oversimplified, as we do not attempt to reproduce the further kinks in $W_u(0)$ (at $g^n(p), n > 1$) and in $W_s(\Lambda)$ (at and near $g^{-n}(p), n \geqslant 0$). However we now have the tangency that we are aiming for.

There are two obvious courses of action open to us, as a response to the non-density of structural stability. We can either alter the equivalence relation on the space of all dynamical systems, in the hope that stability with respect to the new relation might be dense, or we can ask for something less than density in the given topology. We shall examine another equivalence relation in the next section. We finish this section by giving a positive result in the second direction.

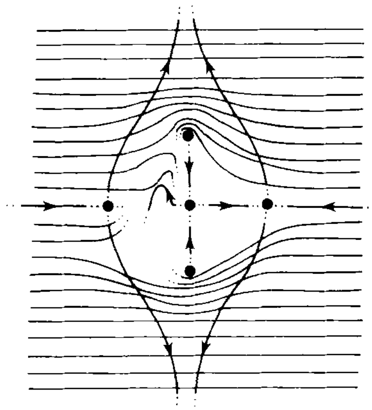


FIGURE 7.29

A natural question to ask is the following. Given an arbitrary dynamical system, can we deform it into a structurally stable system? If we can, how small a deformation do we need to make? We cannot make it arbitrarily C^1-small, since this would imply C^1-density of structural stability. But we might be able to do it by an arbitrarily C^0-small deformation. Notice that we are talking only about the *size* of the deformation needed to produce structural stability; the smoothness of the maps involved and the definition of structural stability are still as before. These questions are answered by the following theorem of Smale [9] and Shub [1] (see also Zeeman [1] and Franks [4]).

(7.30) Theorem. *Any C^r diffeomorphism* ($1 \leqslant r \leqslant \infty$) *of a compact manifold is C^r isotopic to a C^1 structurally stable system by an arbitrarily C^0-small isotopy*†.

The idea behind the proof is to triangulate the manifold X (see Munkres [1]) and hence to get a handlebody decomposition (Smale [10]; see also Mazure [1], who calls them differentiable cell decompositions). The technique then adopted is illustrated in two dimensions by Figure 7.31. The shaded picture on the left, part of the handlebody decomposition of X is originally mapped into X by f as shown. We first isotope f to a map g with the property that the image of every i-handle is contained in a union of j-handles for various $j \leqslant i$. This effect may be produced on the whole of X, using an inductive argument.

† Thus structural stability is dense in $\mathrm{Diff}^r X$ with respect to the C^0-topology. Of course it is no longer open if we use this topology. The analogous theorem for flows has been announced by D'Oliviera [1].

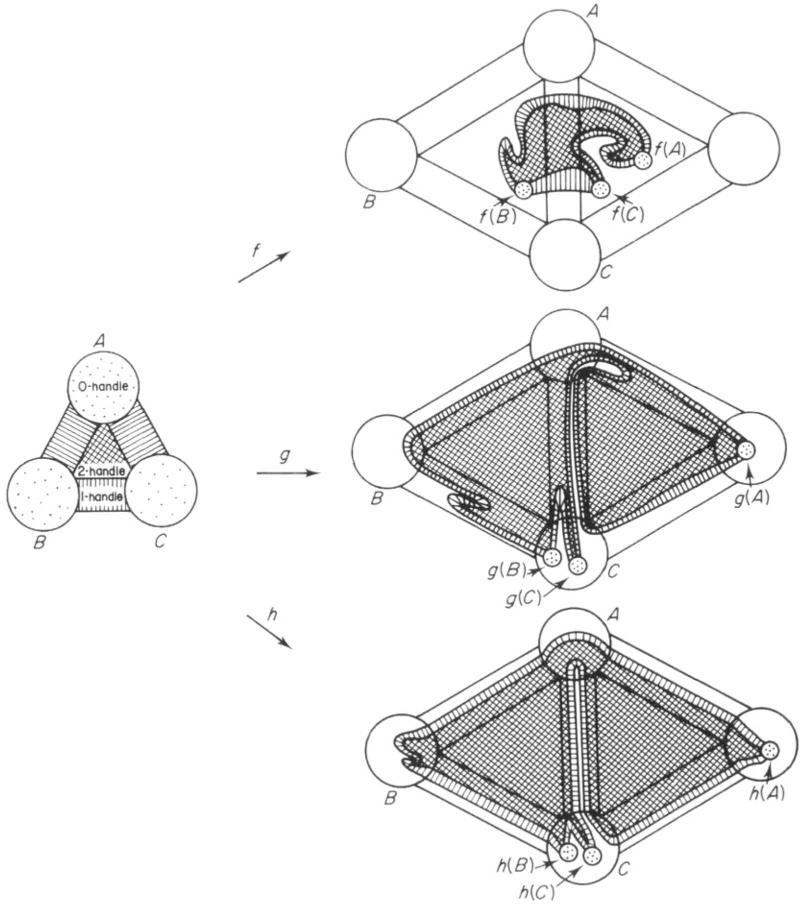


FIGURE 7.31

We then isotope g to a new map h, such that (i) the 0-handles are contracted, (ii) the 2-handles are expanded at all points of h^{-1}(2-handles), and (iii) the 1-handles are mapped in a nice linear fashion, expanding lengthwise and contracting radially, at all points of h^{-1}(1-handles). Notice the horseshoe that appears along the handle AC. With a little care, we have achieved a structurally stable map h, the Ω-set consisting of periodic sources associated with 2-handles, periodic sinks associated with 0-handles and saddle-type basic sets associated with 1-handles which may very well be exotic, for example the above horseshoe. The size of the isotopy involved depends on the fineness of the triangulation of X, and so can be made arbitrarily small, in the C^0-sense, by taking a fine enough triangulation.

V. OMEGA STABILITY

The failure of structural stability to be dense in higher dimensions led to consideration of other equivalence relations on the pace of all systems. Of these, the one that has so far aroused the most interest is *Ω-stability.* We have already introduced the notion of Ω-equivalence at the end of Chapter 2, and Ω-stability, is, of course, stability with respect to this equivalence relation. Thus, for example, a diffeomorphism $f: X \to X$ is *Ω-stable* (in the C^r sense, $1 \leq r \leq \infty$) if, for any nearby $g \in \mathrm{Diff}^r X$, there exists a topological conjugacy between $f|\Omega(f)$ and $g|\Omega(g)$. At first sight it may seem rather an extreme measure to ignore all behaviour off the Ω-set. There are two answers to this criticism. Firstly, it is only in defining the equivalence relation that the non-Ω behaviour is ignored; it is very important in determining stability with respect to the relation. Secondly, it is arguable that in applications of the theory to mathematical modelling the Ω-sets are the only parts of the system with real physical significance. Certainly if a dissipative system in dynamics gives rise to a single dynamical system then only the attractors correspond to phenomena that are physically observable.

In order to get some feeling for the way in which the non-wandering set of a system may alter when the system is perturbed, we examine a few simple examples. First consider a "one-way" zero of a flow on S^1 (Figure 7.32). By

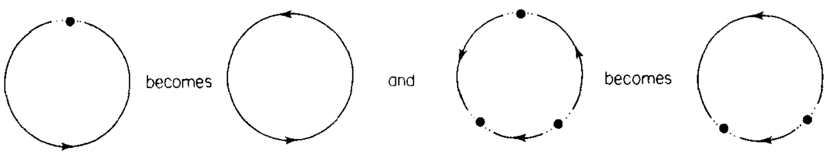


FIGURE 7.32

making a small local perturbation (adding a small velocity in the one-way direction) we eliminate the zero, and, in the first case, introduce a periodic orbit. Thus we have *exploded* the Ω-set from a single point to the whole circle. In the second case, we *implode* the Ω-set from three points to two. If we have an isolated non-hyperbolic zero which is "two-way" then we cannot entirely eliminate it by a local perturbation. However, for any non-hyperbolic zero, we can, by a C^1-small perturbation produce an interval, centred on the point, consisting entirely of zeros (the size of the interval varying with the size of the perturbation, of course), or introduce other more complicated behaviour on such an interval (Figure 7.33). Here the explosion in the Ω-set is local; its size is limited by the size of the perturbation allowed, in contrast to the previous explosion.

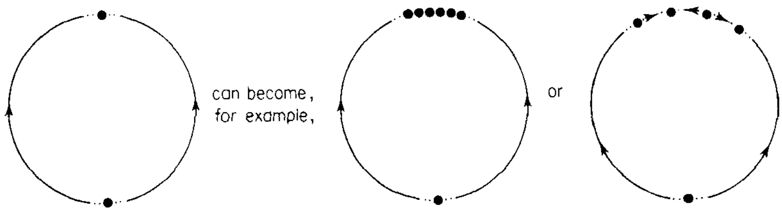


FIGURE 7.33

A more spectacular explosion occurs if we start with the first picture in Figure 7.34, which represents the Ω-set of a diffeomorphism f of S^2, though, as usual, we draw flow-like orbits for easier visualization. The Ω-set consists of two sources y and t, a sink z and a saddle point x, together with an invariant arc l joining x to itself as shown (so that $l \subset W_s(x) \cap W_u(x)$). Notice that l cannot have a hyperbolic structure (Exercise: Why not?).

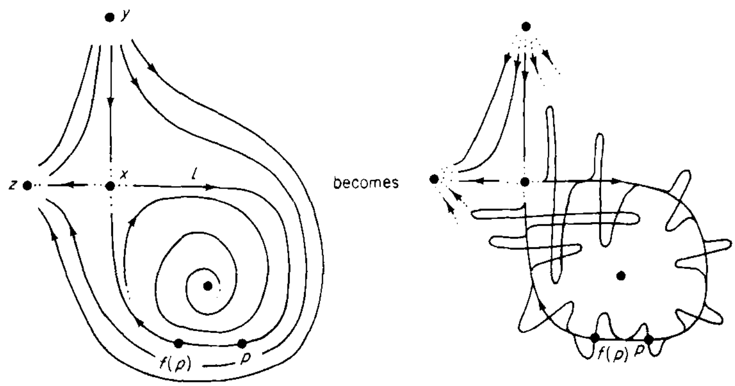


FIGURE 7.34

We perturb f on a small neighbourhood U of p, so that the resulting map g takes $l \cap U$ transversally across $l \cap g(U)$ at $f(p) = g(p)$. The unstable manifold of x with respect to g, $W_u(x)$, is the same as it was before the perturbation as one travels around it from x to p. However, after passing p moving to the left, it has a series of kinks which get larger and larger as they approach x. They press themselves up against $W_u(x)$, and eventually spread out past p into the kinks already formed, and so on. Similarly as $W_s(x)$ passes p going right, it develops a series of kinks which press themselves up against earlier parts of $W_s(x)$. Of course we have just described again the transverse homoclinic point phenomenon of Chapter 6. The point is that $W_s(x)$ now intersects $W_u(x)$ in an extremely complicated and rather random way, and, by the sort of "cloud lemma" argument used in the proof of Theorem 7.14, any such intersection is in $\Omega(g)$. We leave the reader to prove the intuitively obvious fact that $\Omega(f)$ is topologically different from $\Omega(g)$.

At first site, it appears that Ω-explosions are caused by a lack of hyperbolicity, so that we have, in Axiom A, the condition to eliminate them. This turns out to be true, in a sense, for local Ω-explosions, but Axiom A does not prevent global explosion. To see this consider the example (Smale [6]) of a diffeomorphism of S^2 with Ω-set consisting of six hyperbolic fixed points, of which two are sinks, two sources and two saddles, laid out as in Figure 7.35. To see that the Ω-set is as described note that any point outside the closed curve has a neighbourhood that either falls into the sink c under positive iterates of f, or into the source a under negative iterates. Similarly for points inside the closed curve. A small neighbourhood of a point such as p, under positive iterates of f, presses up against the underside of $W_u(x)$ and the right-hand side of the portion of $W_u(y)$ going to d. It does not come back and re-enter U, so p is wandering. Note that under negative iterates of f, U presses itself against the topside of $W_u(x) \cap W_s(y)$ and so positive and negative iterates of U come arbitrarily close together. Thus if we, by local perturbation at q on $W_u(x) \cap W_s(y)$, bring $W_u(x)$ above $W_s(y)$, we cause the positive and negative iterates $f^n(U)$ and $f^{-n}(U)$ to intersect for large n (Figure 7.35). Hence p becomes non-wandering, and similarly, any other

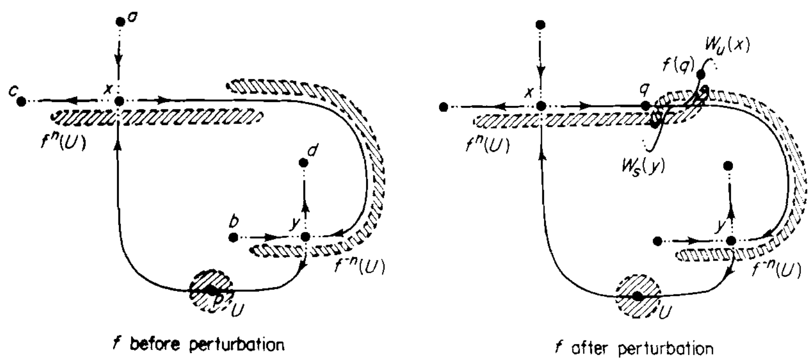


FIGURE 7.35

point of $W_s(x) \cap W_u(y)$. On the other hand the points of $W_u(x) \cap W_s(y)$ are wandering, just as they were before. Hence the new Ω-set is the six points together with the arc $W_s(x) \cap W_u(y)$. Note that, again, we cannot have a hyperbolic structure. If we make another perturbation so that $W_s(x)$ and $W_u(y)$ cross transversally, we cause another explosion. We have forced intersections $W_s(x) \cap W_u(x)$ and $W_s(y) \cap W_u(y)$ to appear and these join the new versions of $W_s(x) \cap W_u(y)$ and $W_s(y) \cap W_u(x)$ and the six fixed points to give the Ω-set of the new map. The set is similar to, but even harder to visualize than the one in the last example.

We now give an example of a system that is Ω-stable but not structurally stable. Consider the gradient flow of the height function on the torus T^2 already described in Example 3.3 and illustrated again in Figure 7.36. A

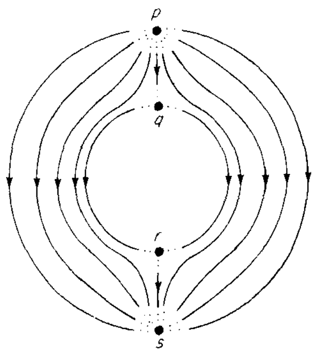


FIGURE 7.36

slight perturbation of the vector field will make the separatrices beginning at q end at s instead of r. However for all C^1-small perturbations the orbit structure is the same near the zeros, and elsewhere the orbits continue to move downwards, which precludes any global recurrence. Thus the Ω-set is still the four zeros, and so the system is Ω-stable.

Since this last example does not have strong transversality, it is clear that this property is not a relevant consideration when trying to characterize Ω-stability. We must examine the above examples to see what causes the explosions. We can distinguish two types, local explosions and global explosions. Local explosions are ones where the old and new Ω-sets are close together, the degree of closeness depending on the size of the perturbation. Global explosions are ones where arbitrarily small perturbations produce Ω-sets that are not close to the original ones†. The examples in Figures 7.33 and 7.34 are local, those in Figures 7.32 and 7.35 are global. As we have already hinted, local explosions are due to a lack of hyperbolicity. The basic sets of an Axiom A diffeomorphism f are locally stable. That is to say, given any basic set Λ, there is a neighbourhood U of Λ such that, for all g

† In some references (e.g. Shub [4]) the term "Ω-explosion" is only applied to global explosions.

sufficiently C^1-near f, $U \cap \Omega(g)$ has a subset Λ' such that

(i) Λ' contains every g-invariant subset of U, and

(ii) $g|\Lambda'$ is topologically conjugate to $f|\Lambda$.

Thus there may be other points of $\Omega(g)$ in U, but their orbits go outside U. This stability property is a close relative to the stability theorem for Anosov diffeomorphisms, and, not surprisingly its proof (see Smale [6] and Hirsch and Pugh [1]) has a lot in common with that of Theorem 7.9 sketched above.

Several times in this chapter, in proving that a point is non-wandering, we have used arguments that involve open sets spreading out from basic set to basic set until they get back near where they started. Any two consecutive basic sets in the route have to be linked together, or the open set cannot make it across from the one to the other. The link consists of an unstable manifold of the one and a stable manifold of the other with non-empty intersection. If the manifolds meet transversally the open set can always get across; if not then it may still be lucky or it may not (see Figure 7.37).

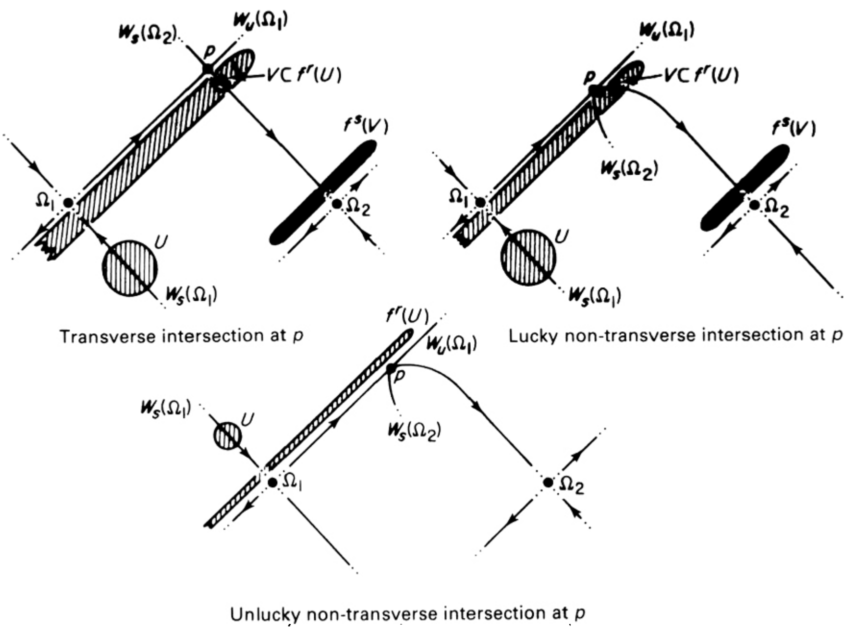


FIGURE 7.37

Another unlucky situation, illustrated in the first picture in Figure 7.35 occurs when the open set reaches the first basic set of a link but cannot get near the intersection point of the link. This can happen when, due to an earlier lack of transversality, the open set does not intersect the stable manifold of the basic set ($f^n(U)$ does not intersect $W_s(y)$ in the picture). These unlucky situations, since they are all due to non-transversality, may be eliminated by arbitrarily small perturbations producing transversality. Thus a recipe for a global Ω-explosion of an Axiom A system is to start with a collection of basic sets, linked up to form a circular route, but with one or more unlucky situations incorporated (of course—we could not have separate basic sets otherwise). Then, by a small perturbation, we can remove all the bad luck situations, simultaneously if we do not insist on the perturbation being a local one, and we have our explosion.

The above description motivates the following definition (Rosenberg [1]). An *n-cycle* of an Axiom A dynamical system is a sequence of basic sets $\Omega_0, \Omega_1, \dots, \Omega_n$, with $\Omega_0 = \Omega_n$ and $\Omega_i \neq \Omega_j$ otherwise, and such that $W_u(\Omega_{i-1}) \cap W_s(\Omega_i)$ is non-empty for all i, $1 \leq i \leq n$. An Axiom A dynamical system satisfies the *no-cyclic condition* if it has no n-cycles for all $n \geq 1$. The main theorem on Ω-stability, due to Smale [6] in the diffeomorphism case, and to Pugh and Shub [1] in the flow case is:

(7.38) Theorem. *If an Axiom A system on a compact manifold has the no-cycle property then it is Ω-stable.*

The idea of the proof is to show first, as indicated above, that Axiom A rules out local explosions. We then have a situation rather similar to that illustrated in Figure 7.36. If $W_u(\Omega_i) \cap W_s(\Omega_j)$ is non-empty, we can think of Ω_i as being situated above Ω_j (the no-cycle condition ensures that we do not find, paradoxically, that Ω_i is above itself) and the movement between basic sets as being downwards. It seems quite clear that C^1-small perturbations cannot produce an upwards motion and so global recurrence cannot occur in the perturbed system. Technically the proof uses a *filtration* (see Shub and Smale [1]). This is an increasing sequence of compact manifolds with boundary $X_1 \subset \cdots \subset X_r = X$, such that $\dim X_i = \dim X$, $f(X_i) \subset \operatorname{int} X_i$ and each $X_i \backslash X_{i-1}$ contains a single basic set. We picture a suitable filtration for the Figure 7.36 example in Figure 7.39. As with structural stability, it is conjectured that we have here a complete characterization; that Ω-stability implies Axiom A and no cycles. Palis [1] has the partial result that Axiom A and Ω-stability imply no cycles. There is, again, a notion of *absolute* stability, and this is characterized by Axiom A and no cycles (see Franks [3], Guckenheimer [1]).

Unfortunately Ω-stability is no more successful than structural stability as far as density is concerned. Abraham and Smale [1] took the product of a toral automorphism and a horseshoe diffeomorphism on S^2,, and by making the former *DA* over one of the two fixed points ($\dots 00{\cdot}00 \dots$ and $\dots 11{\cdot}11 \dots$) of the horseshoe, they constructed a

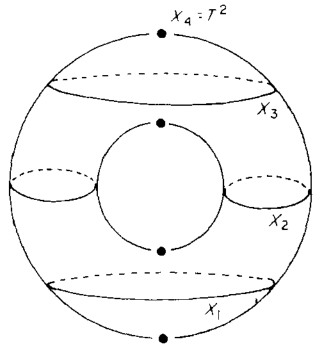


FIGURE 7.39

diffeomorphism of $T^2 \times S^2$ that has a C^1-neighbourhood all of whose members are neither Axiom A nor Ω-stable. Finally, Newhouse [1], by modifying the horseshoe diffeomorphism, constructed a counterexample to C^2-density of Axiom A and of Ω-stability for diffeomorphisms of any surface.

VI. BIFURCATION

Several other notions of stability have been put forward in the hope that they might prove to be generic, for example *future stability* (Shub and Williams [1]), *tolerance stability* (Takens [1], White [1]), *finite stability* (Robinson and Williams [1]) and *topological Ω-stability* (Hirsch, Pugh and Shub [1]). Some have survived longer than others; none has yet been wholly successful. In the course of time, a feeling has developed that perhaps it is too optimistic to expect to find a single natural equivalence relation with respect to which stability is dense (see, for example, Smale [8]). Recently, more attention has been focused on the interesting and important question of bifurcation of systems.

Up to now we have not devoted much time to systems such as the one-way zero in Figure 7.32. Our feeling has been that the most important systems are ones which can be used to model the dynamics of real life situations. But no real life situation can ever be exactly duplicated, and we should expect this to lead to slight variations in the model system. Thus a theory making use of qualitative features of a dynamical system is not convincing unless the features are shared by nearby systems. That is to say, good models should possess some form of qualitative stability. Hence our contempt for the one-way zero, which is extremely unstable. Now this argument has some force, but it is too simple-minded. Firstly, there may, in a physical situation, be factors present that rule out certain dynamical systems as models. Conservation laws have this effect, and so has symmetry in the physical situation. In this case, the subset of those dynamical systems that *are* allowable as models may be nowhere dense in the space of all systems, and so the stable systems that we have considered in this book are totally irrelevant. One has to consider afresh which properties are generic in the space of admissible systems. This has been done for Hamiltonian systems (see Robinson [1]). Secondly, even if the usual space of systems is the relevant one, the way in which a stable system loses its stability as it is gradually perturbed may be of importance, since the model for an event may consist not of a single system but of a whole family of systems. In his theory of morphogenesis, Thom [1] envisages a situation where the development of an organism, say, is governed by a collection of dynamical systems, one for each point of space time. The dynamical systems are themselves controlled by a number of parameters, which could for example be quantities like temperature and concentrations of chemicals in the neighbourhood of the point in question. A *catastrophe* is a point where the form of the organism changes discontinuously, and it corresponds to a topological change in the orbit structure of the dynamical systems. We say that the family of dynamical systems *bifurcates* there.

We shall give some examples of bifurcations of flows. These particular ones are local changes that can take place on any manifold. By taking a suitable chart we may work in Euclidean space. First consider a vector field v_α on **R** given by

$$v_\alpha(x) = \alpha + x^2.$$

Here α is a single real parameter, so for each $\alpha \in \mathbf{R}$ we have a vector field on R. We are interested in how the orbit structure varies with α. The analysis is very easy, and we find that

(i) for $\alpha > 0$ we have no zeros, and the whole of **R** is an orbit, oriented positively,

(ii) for $\alpha = 0$ we have a single zero at $x = 0$, which is a one way zero in the positive direction,

(iii) for $\alpha < 0$, we have two zeros, a sink at $-\sqrt{-\alpha}$ and a source at $\sqrt{-\alpha}$.

Thus the bifurcation occurs at $\alpha = 0$. Figure 7.40 illustrates the bifurcation.

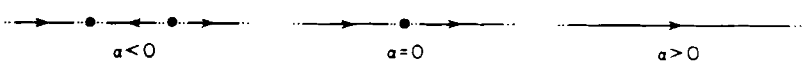


FIGURE 7.40

If one takes the product of v_α with a fixed (i.e. independent of α) vector field on $\mathbf{R}^{n-1}$ having a hyperbolic fixed point at 0, one obtains a bifurcation of the resulting vector field on $\mathbf{R}^n$. Of course we have essentially added nothing to the original bifurcation in doing this. All such bifurcations are known as *saddle-node bifurcations*, because of the picture that one gets for $n = 2$ (Figure 7.41). A saddle point and a node come together, amalgamate and cancel each other out.

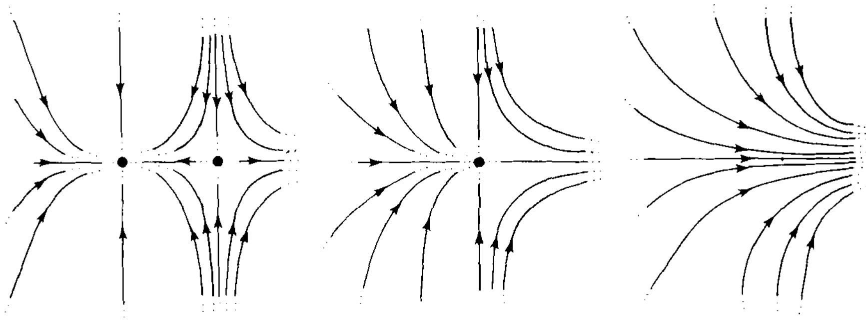


FIGURE 7.41

Saddle-node bifurcations are important because (when properly defined) they are stable as bifurcations of one-parameter families. Roughly speaking, then, one-parameter families near to a family with a saddle-node bifurcation also exhibit something that is topologically like a saddle node bifurcation near (in the positions of the amalgamating zeros, for example) the original one. One speaks of them as being *codimension one* bifurcations; one is then visualizing the set of systems exhibiting zeros of the $\alpha = 0$ type above as (in some sense) a submanifold of codimension one in $\Gamma^r(X)$, and the one-parameter family as being given by an arc in $\Gamma^r(X)$ crossing the submanifold transversally. Notice that the bifurcation illustrated in Figure 7.42 (a node bifurcating into two nodes and a saddle point) is *not* stable for one-parameter families. It can be perturbed slightly so that there is a saddle-node bifurcation near to but not at the original node.

The saddle-node bifurcation is the typical bifurcation obtained when the sign of a real eigenvalue of (the differential at) a zero is changed by varying a single parameter governing the system. There is, similarly, a typical codimension one bifurcation which comes about when the sign of the real part of a complex conjugate pair of eigenvalues is changed by varying a single parameter. This is known as the *Hopf bifurcation*, since it was first described by E. Hopf [1]. Consider the vector field v_α on $\mathbf{R}^2$ given by

$$v_\alpha(x, y) = (-y - x(\alpha + x^2 + y^2), x - y(\alpha + x^2 + y^2)).$$

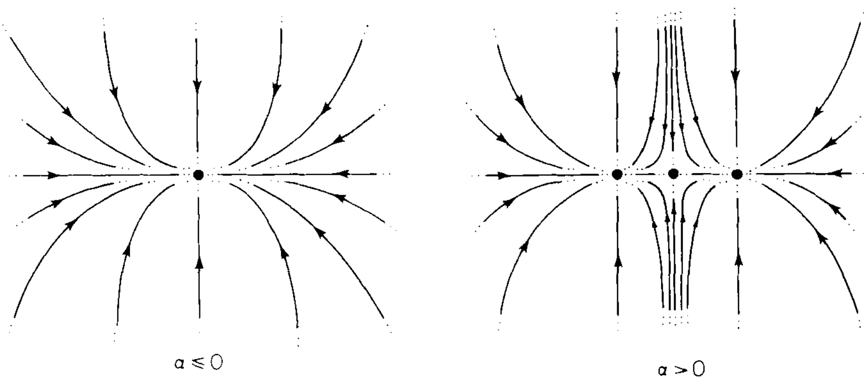


FIGURE 7.42

Again α is a single real parameter. For all $\alpha \in \mathbf{R}$, v_α has a zero at the origin, and the linear terms make this a spiral source for $\alpha < 0$ and a spiral sink for $\alpha > 0$. For $\alpha = 0$ the linear terms would give a centre (recall that this is a zero surrounded by closed orbits), but the cubic terms make the orbits spiral weakly inwards. The interesting feature of the bifurcation is not, however, the zero, but the unique closed orbit that we get at $x^2 + y^2 = -\alpha$, for each $\alpha < 0$. The bifurcation is illustrated in Figure 7.43. The periodic attractor

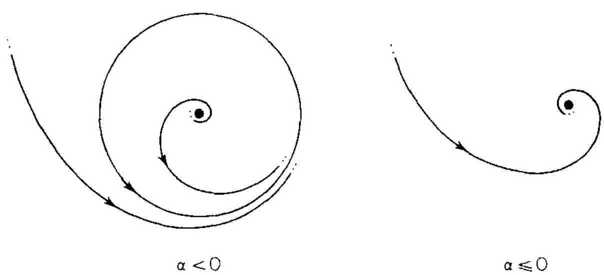


FIGURE 7.43

decreases in size until it amalgamates with the spiral source to form a spiral sink. The reverse bifurcation, in which a spiral sink splits into a spiral source and a periodic attractor, is intriguing because of the feeling one has that from something inert and dead (the source) one has created something pulsating and alive (the periodic orbit). This makes it very popular as a component in mathematical models.

It is not possible at present to give a coherent account of bifurcation theory for dynamical systems, since the subject is still in an early stage of development. Papers of Sotomayor [1, 2, 3] give some general principles for tackling the problem and prove some generic properties of bifurcations of one parameter families of vector fields (see also Guckenheimer [3]). Newhouse and Palis [2, 3] have examined the situation after one passes the first bifurcation of a Morse–Smale vector field, and found that it may be very complicated. Newhouse and Peixoto have investigated conditions under which pairs of structurally stable vector fields may be connected by one-parameter families with only finitely many stable bifurcations (see Newhouse and Peixoto [1] and Newhouse [3]). Takens [2, 3, 4] has written papers on various aspects of bifurcations of singularities of dynamical systems. Guckenheimer [2] has shown that the analysis of bifurcation of functions on manifolds is fundamentally different from and simpler than the corresponding theory for gradient vector fields.

The above discussion has mainly been for bifurcations in the context of topological equivalence of flows. The theory for topological conjugacy of diffeomorphisms is less well developed, and seems to differ in some respects. Also, of course, any other equivalence relation on dynamical systems has an associated bifurcation theory. Very little work has been done so far on such bifurcations. It seems likely, though, that for any interesting relation it will be a difficult task to establish a satisfactory general theory.