

LECTURES ON DYNAMICAL SYSTEMS

Chapters I-III

Oscar E. Lanford III

ETH-Zürich

Disclaimer: This is a preliminary text for a course of lectures given in the academic year 1987-88; it is intended for the use of the students in the course and not for wider distribution. While I will be happy if anyone else finds it useful, the reader should be aware that this manuscript is not in final form and should therefore be used with care. One particularly glaring defect is the lack of any attempt to attribute credit for the results discussed.

O. E. Lanford

Mathematik
ETH-Zentrum
8092 Zürich (Switzerland)

November 1987

CHAPTER I. INTRODUCTION

The fundamental motivation for the study of dynamical systems is the desire to understand, qualitatively, the long-time behavior of solutions of differential equations, even if these differential equations cannot be solved explicitly. This subject can be studied on a number of different levels of generality and abstraction. We can situate, roughly, the subject of these lectures by saying that, in contrast to ergodic theory, we will emphasize the *differentiable* character of the objects being studied but that, on the other hand, we will not investigate the properties of Hamiltonian motion which derive specifically from its conservative character. It is customary to refer to the general subject to be addressed in these lectures as the theory of *dissipative differentiable dynamical systems*, where the word “dissipative” is to be taken to mean merely “non-Hamiltonian”; it is rare that one can ever exploit the genuinely dissipative features of systems with friction.

Let us consider, then, a differential equation

$$(DE) \quad \frac{dx}{dt} = X(x).$$

Note that we are supposing that our differential equation is *autonomous*, i.e., that t does not appear explicitly on the right-hand side. We will refer to the set of possible x 's as the *state space*, and generally denote it by M ; it might be Euclidean space, or an open set in Euclidean space, or an open set in a Banach space (infinite system of coupled equations, or partial differential equation) or it might be a manifold, or a Banach manifold. The natural framework for most of our investigations will be manifolds, possibly infinite dimensional ones. Nevertheless, we will write all formulas for the simple case of differential equations on Euclidean or Banach spaces. The reason for doing this is that, on the one hand, non-trivial global properties of manifolds play almost no role in the investigations to be described here—most of these analyses, once carried out in Euclidean spaces, are easy to extend to manifolds—and, on the other hand, the systematic use of the formalism of differentiable manifold theory, with co-ordinate patches, tangent bundles, etc., is a serious distraction from the real issues. This expository strategy doesn't quite work, since we will want sometimes to assume that M is compact, but I don't intend to worry about this; the audience will be assumed to be translating everything into the language of differentiable manifold theory. (Mike Shub has suggested that most theorems about differentiable dynamical systems are proved assuming the state space to be a compact Euclidean space.)

This may be a good place to comment on the philosophy to be adopted in the lectures about whether or not to assume that the state space is finite dimensional. The first point to be made is that many of the most penetrating analyses of finite dimensional dynamical systems rest on the use of non-linear functional analysis on infinite spaces, so, even if we are really only interested in finite dimensional systems, we need to develop a fair amount of infinite-dimensional theory. Once that theory is at hand, it turns out that a great deal of dynamical systems theory can be done, at little or no extra cost, without assuming finite dimensionality. Indeed, as is often the case, the discipline

of avoiding unnecessary assumptions of finite dimensionality frequently leads to better proofs. I have therefore adopted the point of view that, whenever finite-dimensionality assumptions can be avoided, without making an argument *essentially* more difficult, avoid them.

One should not, however, overstate what is accomplished by avoiding unnecessary finite-dimensionality assumptions. It does not, in itself, make the theory applicable to interesting infinite dimensional systems like the Navier-Stokes equation. The problem with the Navier-Stokes equation, and most other partial-differential evolution equations is that it is definitely not a smooth dynamical system. If one writes it in the schematic form (DE), the $X(x)$ appearing on the right-hand side is not a smooth function of x but involves, e.g., unbounded operators. To a certain extent, the lack of smoothness of $X(x)$ can be circumvented by exploiting the fact that the equation of motion, like the heat equation, produces smoothing when run forward in time, so time evolution operators for positive times may be smooth even if the infinitesimal time evolution operators are not. Analyzing the dynamical properties of such evolution equations as the Navier-Stokes equation will certainly need to make use of the methods of infinite dimensional dynamical systems theory, but will require in addition a detailed study of the particular evolution equation. That kind of analysis of particular partial differential equations will not be given here.

If $X(x)$ in (DE) is Lipschitz continuous, then solutions exist, at least locally, for arbitrary initial data, and they are unique. As simple and apparently well-behaved an equation as

$$\frac{dx}{dt} = x^2$$

shows that solutions need not exist for all times. In any event, we introduce the associated solution mapping f^t , defined as follows: $f^t(x_0)$ is defined if and only if the differential equation admits a solution $x(t)$ defined for all t between 0 and t_0 , and its value in this case is $x(t_0)$. By the uniqueness of solutions of the differential equation,

$$f^{t_1}(f^{t_2}(x)) = f^{t_1+t_2}(x),$$

whenever the left-hand side is defined. This condition is expressed by saying that f^t is a semi-flow. If $f^t(x)$ is defined for all (x, t) , then each f^t is a diffeomorphism (i.e., a smooth mapping of M onto itself with a smooth inverse) and $f^{t_1}f^{t_2} = f^{t_1+t_2}$. Furthermore, $f^t(x)$ is jointly smooth in (t, x) . We thus say that f^t is a smooth flow. Conversely, any smooth flow f^t is the solution flow for the differential equation with

$$X(x) = \left. \frac{df^t}{dt}(x) \right|_{t=0}.$$

A flow is a one-parameter group of diffeomorphisms with parameter $t \in \mathbb{R}$. Often, one considers instead discrete one-parameter groups, i.e., the set of all iterates f^n of a single diffeomorphism. Sometimes, moreover, one drops the condition that f be invertible, i.e., one considers the positive iterates of an arbitrary smooth map.

There are a number of reasons for considering maps rather than flows. One is the fact that most theorems about maps have analogues for flows, and vice versa, but the versions for maps is often technically simpler and exhibits the main ideas in a less encumbered form. Secondly, if f^t is a flow, then the large- t behavior of f^t is essentially determined by the behavior of iterates of any single f^{t_0} with $t_0 \neq 0$, e.g., $t_0 = 1$. Thus any theorem about maps gives information about flows. A third reason is the "surface of section" idea of Poincaré which, in some cases, reduces the study of a flow in an m -dimensional space to that of a mapping in an $m - 1$ -dimensional space. The idea is as follows: Let f^t be a flow in an m -dimensional state space M and let Σ be a hypersurface (codimension-one submanifold) of M which is, at each of its points, transverse to the generator $X(x)$ of f^t . If the forward orbit of $x \in \Sigma$ crosses Σ again, define $\Phi(x)$ to be the first such crossing point (and leave $\Phi(x)$ undefined otherwise.) The mapping Φ so defined is called the first return map or Poincaré map for Σ . It follows easily from the Implicit Function Theorem that the domain of Φ is an open subset of Σ and that Φ is smooth on its domain. It is often the case that, if Σ is judiciously chosen, quite a lot of the dynamics of f^t is captured by the dynamics of the iterates of Φ . For example, a fixed or periodic point x_0 for Φ corresponds to a closed orbit (periodic solution) for f^t , and the long-time behavior of orbits with initial point near x_0 can be determined from the behavior of iterates of Φ near x_0 . Fourthly, iteration of maps provides a natural way to analyze the behavior of solutions of differential equations which are not autonomous but periodic. Consider a differential equation

$$\frac{dx}{dt} = X(t, x),$$

where X is periodic in t : $X(t + \tau, x) = X(t, x)$. If solutions exist for all time, then we can construct the corresponding solution mapping f^t . This will not be a flow—the group property will fail—but it will remain true that $f^{n\tau} = (f^\tau)^n$. Thus, the iterates of the time τ mapping essentially determine the long-term behavior of solutions of the differential equation. Finally, interesting phenomena occur in state spaces of smaller dimension for mappings than for flows and may therefore be more comprehensible. Roughly speaking, flows on spaces of dimension less than three cannot exhibit any interesting asymptotic behavior, while invertible maps can already do interesting things in two dimensions and non-invertible maps already in one dimension.

CHAPTER II. RUDIMENTS OF SPECTRAL THEORY FOR GENERAL BOUNDED OPERATORS ON BANACH SPACES

§1. Orientation. Any normal operator A on a finite dimensional (complex) Hilbert space $\mathcal{X}$ has a very simple structure: If the eigenvalues of A are $\lambda_1, \dots, \lambda_m$, and if $\mathcal{X}_i$ is the set of vectors ξ such that $A\xi = \lambda_i\xi$ (i.e., the eigenvectors with eigenvalue λ_i together with the zero vector), then $\mathcal{X}$ is the direct sum of the $\mathcal{X}_i$.

This simple structure theory can be extended satisfactorily in two directions. The first extension is to *general* linear operators on *finite-dimensional* vector spaces. The principal complication here is the necessity to take into account the possible existence of *generalized eigenvectors*, i.e., vectors annihilated by some power of $A - \lambda 1$, λ a complex number, but not necessarily by $A - \lambda 1$ itself. If, again, we let the eigenvalues of A be $\lambda_1, \dots, \lambda_m$, and if we now let $\mathcal{X}_i$ denote the set of all vectors ξ such that $(A - \lambda_i 1)^j \xi = 0$ for some integer $j \geq 1$ then, again, each $\mathcal{X}_i$ is a linear subspace mapped into itself by A , and $\mathcal{X}$ is the direct sum of the $\mathcal{X}_i$. To complete the structure theory it is only necessary to work out what A restricted to any one of the $\mathcal{X}_i$ can look like, and this is a relatively straightforward task.

The second extension is to *normal* operators on infinite dimensional Hilbert spaces. Leaving aside problems concerning unbounded operators, the main complication here is the fact that such an operator need not have any eigenvectors at all. The direct-sum decomposition into eigenspaces in the finite dimensional case is replaced in this more general situation by an integral representation $A = \int \lambda P(d\lambda)$ where $P(d\lambda)$ is a projection valued measure on the complex plane. The intuitive interpretation of this projection valued measure is that the range of projection $P(X)$ corresponding to a Borel set X is the proper infinite-dimensional analogue of the linear span of the eigenvectors with eigenvalues in X in the finite-dimensional case.

For a general bounded operator A on a Banach space $\mathcal{X}$, our knowledge is much less complete. There is, however, a good definition of the spectrum in this general situation, and one can construct useful spectral projections $P(X)$ associated with subsets X of the spectrum which are at a non-zero distance from the rest of the spectrum.

§2. The spectrum. From now on in this chapter, A will denote a bounded linear operator on a Banach space $\mathcal{X}$. By the resolvent set of A we mean the set of complex numbers λ such that $\lambda 1 - A$ has a bounded (everywhere defined) inverse. By the Closed Graph Theorem, this is the same as saying that $\lambda 1 - A$ is injective—i.e., that λ is *not* an eigenvalue of A —and that the range of $\lambda 1 - A$ is all of $\mathcal{X}$. The inverse $(\lambda 1 - A)^{-1}$ is called the resolvent (of A , or of λ , or of (A, λ) , depending on the context.) The spectrum of A means the complement of the resolvent set; we will denote it by $\sigma(A)$.

One of the principal tools of elementary spectral theory is the systematic exploitation of the Neumann series. This means the following: Let U be an operator with a bounded inverse, and V some other operator. Then, formally, the infinite series

$$(2.1) \quad U^{-1} - U^{-1}VU^{-1} + \dots + (-1)^n (U^{-1}V)^n U^{-1} + \dots$$

should give the inverse of $U + V$. It is easy to check that this is true in the sense that, whenever the series (2.1) converges in norm, $U + V$ is invertible and the sum of the series

$$\sum_{n=0}^{\infty} (-1)^n (U^{-1}V)^n U^{-1} = \frac{1}{1+U^{-1}V} U^{-1} = \frac{1}{U+V}$$

is $(U + V)^{-1}$. Furthermore, the series is easily seen to converge if $\|V\| < (\|U^{-1}\|)^{-1}$ or, more generally, if $\|(U^{-1}V)^m\| < 1$ for some positive integer m . This remark shows, for example that the resolvent set for any A is open, and that the resolvent $(\lambda 1 - A)^{-1}$ varies analytically with λ on the resolvent set, i.e., that it is given locally in the neighborhood of each point by a norm-convergent power series. Since the spectrum is defined to be the complement of the resolvent set, it is closed.

It is quite a non-trivial fact that the spectrum of an operator A cannot be empty. To prove it, assume the contrary, i.e., that the resolvent set of some operator A is the whole complex plane. Then $(\lambda 1 - A)^{-1}$ is an entire operator valued function, which furthermore is easily seen—using the Neumann series—to go to zero at infinity. Thus, for any ξ in $\mathcal{X}$ and any ϕ in the dual of $\mathcal{X}$, $\phi((\lambda 1 - A)^{-1}\xi)$ is an entire complex valued analytic function vanishing at infinity which therefore, by Liouville's Theorem, vanishes identically. Since this is true for all ξ, ϕ , it follows that $(\lambda 1 - A)^{-1}$ vanishes for all λ , and this is manifestly impossible.

The argument in the preceding paragraph illustrates a general method of extending results from complex analysis to vector- and operator-valued functions. Many important results are easy to extend in this way, among them the Cauchy Integral Theorem and Cauchy Integral Formula; the integrals involved can be interpreted simply as Riemann integrals.

By elementary properties of the norm, $\|A^{n+m}\| \leq \|A^n\| \|A^m\|$. From this it follows easily that $\|A^n\|^{1/n}$ converges as n goes to infinity, and that the limit is actually the infimum of the $\|A^n\|^{1/n}$. The limit is called the spectral radius of A ; we will denote it by $\rho(A)$. The terminology is justified by the fact that the spectral radius turns out to be equal to the supremum of $|\lambda|$ over the spectrum of A . To see this, we first invoke the Neumann series to write

$$(2.2) \quad (\lambda 1 - A)^{-1} = \sum_{n=0}^{\infty} A^n / (\lambda)^{n+1}$$

wherever the right-hand side converges. Since the right-hand side certainly converges for $|\lambda| > \rho(A)$, this shows that the supremum of $|\lambda|$ over the spectrum of A is at most equal to $\rho(A)$. On the other hand, integrating (2.2) term-by-term gives

$$A^n = \frac{1}{2\pi i} \oint_{\gamma_R} \lambda^n (\lambda 1 - A)^{-1} d\lambda,$$

where γ_R is the circle $\{|\lambda| = R\}$, traversed in the counterclockwise sense, and R is any positive number larger than $\rho(A)$. By the Cauchy Integral Theorem we can reduce the radius R to any number strictly larger than the supremum of $|\lambda|$ over the spectrum. Estimating the integral in a straightforward way gives a bound $\|A^n\| \leq B_R R^n$ for any such R and any positive integer n , and from this bound it follows that $R \geq \rho(A)$; Hence, the supremum of $|\lambda|$ over the spectrum cannot be strictly smaller than $\rho(A)$, so equality is established.

Since $\rho(A) = \inf_n \|A^n\|^{1/n}$, we always have $\rho(A) \leq \|A\|$. The 2×2 matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

shows that this inequality cannot always be made an equality, even with the freedom to replace the norm by an equivalent one. However, we can do almost as well:

PROPOSITION 2.1. *If $R > \rho(A)$, there exists a norm $||| \cdot |||$, equivalent to the original norm $\| \cdot \|$, such that $|||A||| \leq R$*

PROOF: Since $R > \rho(A)$, $\|(A/R)^n\|$ goes to zero as n goes to infinity. Hence, if we put

$$|||\xi||| = \sup_{n \geq 0} \|(A/R)^n \xi\|,$$

we get

$$\|\xi\| \leq |||\xi||| \leq \left(\sup_{n \geq 0} \|(A/R)^n\| \right) \|\xi\|,$$

i.e., the norm $||| \cdot |||$ is equivalent to the norm $\| \cdot \|$. On the other hand,

$$|||A\xi||| = \sup_{n \geq 0} \|(A/R)^n A\xi\| = R \sup_{n \geq 0} \|(A/R)^{n+1} \xi\| \leq R |||\xi|||,$$

so $|||A||| \leq R$, as desired.

§3. Spectral projections. A projection will mean a bounded operator P such that $P^2 = P$, i.e., such that $P\xi = \xi$ for all ξ in the range of P . Note that P is a projection if and only if $1 - P$ is. To specify a projection is essentially the same thing as to specify a direct-sum decomposition of $\mathcal{X}$ into two closed complementary subspaces. In one direction: Given a projection P , we let $\mathcal{X}_1$ denote the range of P (or, equivalently, the null space of $1 - P$) and $\mathcal{X}_2$ the null space of P . Only the zero vector is in both $\mathcal{X}_1$ and $\mathcal{X}_2$. If ξ is any vector, then $P\xi \in \mathcal{X}_1$ and $\xi - P\xi \in \mathcal{X}_2$; hence, $\mathcal{X}_1$ and $\mathcal{X}_2$ span $\mathcal{X}$. Thus $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. In the other direction: Given a direct sum decomposition, the mapping $\xi_1 \oplus \xi_2 \mapsto \xi_1$ is a projection and the Closed Graph Theorem implies that it is continuous if $\mathcal{X}_1$ and $\mathcal{X}_2$ are closed. In this case, we speak of the projection onto $\mathcal{X}_1$ along $\mathcal{X}_2$.

Note that it is necessary to specify both the range and the null space in order to specify a projection. This is contrary to experience with spectral theory for normal operators, where the projections encountered are all self-adjoint or *orthogonal* projections, i.e., ones for which the null space is the orthogonal complement of the range; evidently, in this case, it suffices to specify one of them. It is easy to see that a projection P commutes with a bounded operator A if and only if both the range and null space of P are mapped into themselves by A (i.e., are "invariant under A "). Again, the invariance of the null space follows from that of the range and vice versa if A is a normal operator and P an orthogonal projection, but the two conditions are independent in the general case.

* $\left. \begin{matrix} AA^* = A^*A \\ A^2 = A \end{matrix} \right\} \Rightarrow \left. \begin{matrix} A = A^* \\ \ker A \oplus \text{ran } A = \mathcal{X} \end{matrix} \right\}$

6

$$PA = AP$$

$$A \ker P = A\{0\} = \{0\}$$

$$PA \ker P = AP \ker P = \{0\}$$

$$P A \xi \Rightarrow A \xi \in \ker P$$

$$A \ker P \subseteq \ker P$$

$$A \text{ran } P \subseteq \text{ran } P$$

$$PA\xi = PA(\xi_1 + \xi_2)$$

$$= P(A\xi_1) + P(A\xi_2) = \underbrace{PA\xi_1}_{\in \ker P} + \underbrace{PA\xi_2}_{\in \text{ran } P} = \underbrace{PA\xi_1}_{\in \ker P} + \underbrace{PA\xi_2}_{\in \text{ran } P} = AP\xi$$

$$A \text{ran } P = AP\eta = P A \eta \in \text{ran } P$$

The main piece of spectral theory which persists in the general case we are considering is the following: To any splitting $\sigma(A) = S_1 \cup S_2$ into two *disjoint closed* subsets (which, by the compactness of the spectrum, must be at a non-zero distance from each other), there corresponds a direct sum splitting $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$ of the space into two closed invariant subspaces such that the spectrum of the restriction of A to $\mathcal{X}_1$ is S_1 and that of the restriction of A to $\mathcal{X}_2$ is S_2 . Furthermore, the splitting is invariant not only under A itself, but under any operator which commutes with A , and this condition, together with what was said above, determines it uniquely.

The construction of this splitting is trivial in the case of a diagonalizable operator on a finite-dimensional space— $\mathcal{X}_1$ is the linear span of the eigenvectors with eigenvalues belonging to S_1 and $\mathcal{X}_2$ the linear span of the eigenvectors with eigenvalues belonging to S_2 . Described in this way, the construction does not generalize; the key to generalizing it is the observation that the projection onto $\mathcal{X}_1$ along $\mathcal{X}_2$ can be written as

$$\frac{1}{2\pi i} \oint_{\gamma} (\lambda 1 - A)^{-1} d\lambda,$$

where γ is a contour enclosing all the points of S_1 and none of the points of S_2 . To see that the above integral formula does give the indicated projection, it is only necessary to verify that the integral is the identity on eigenvectors with eigenvalues in S_1 and zero on eigenvectors with eigenvalues in S_2 , and this follows easily, for diagonalizable operators, from the Cauchy Integral Formula. We should note here that the description of the conditions on the “contour” γ were a little imprecise. What we need, precisely, is that γ is a *cycle*, i.e., a formal sum of closed paths, in the complement of the spectrum, which winds once around each point of S_1 and zero times around each point of S_2 , i.e., $(1/2\pi i) \oint_{\gamma} (\lambda - \lambda')^{-1} d\lambda$ is one for λ' in S_1 and zero for λ' in S_2 .

With this simple example as motivation, we formulate the following more general result:

PROPOSITION 3.1. *Let A be a bounded operator on the Banach space $\mathcal{X}$, $\sigma(A) = S_1 \cup S_2$ a decomposition of the spectrum of A into two closed disjoint subsets, and γ a cycle in $\sigma(A)$ winding once around each point of S_1 and zero times around each point of S_2 . Then*

$$P = \frac{1}{2\pi i} \oint_{\gamma} (\lambda 1 - A)^{-1} d\lambda$$

is a projection which commutes with every bounded operator commuting with A (so in particular commutes with A itself). The integral depends only on S_1 , S_2 , not on γ . Furthermore, the spectrum of the restriction of A to the range of P is S_1 and the spectrum of the restriction of A to the null space of P is S_2 . The properties of P listed determine it uniquely.

We will refer to P as the spectral projection of A corresponding to the subset S_1 of its spectrum.

PROOF: We first remark that it follows easily from a sufficiently general form of the Cauchy Integral Theorem (see, e.g., L. Ahlfors, *Complex Analysis*, 2nd. ed., Theorem

18 of Chapter 4) that the integral does not depend on γ . Also, in the case where S_2 is empty and γ is a sufficiently large circle centered at the origin, we have already noted that integrating the Neumann series term-by-term gives $P = 1$. From this and the fact that P does not depend on γ , it follows that interchanging the roles of S_1 and S_2 replaces P by $1 - P$.

We now want to show that P , as defined by the contour integral, is a projection. For this purpose, we will need the following topological remark: Let γ be as in the statement of the proposition. Then there exists another cycle γ_1 , which also winds once around each point of S_1 and zero times around each point of S_2 , and which is "inside" γ in the sense that γ winds once around each point of γ_1 but γ_1 winds zero times around each point of γ . One way to guarantee this is to take γ_1 to lie in a small neighborhood of S_1 . Then

$$\begin{aligned} P^2 &= \frac{1}{2\pi i} \oint_{\gamma} (\lambda 1 - A)^{-1} d\lambda \frac{1}{2\pi i} \oint_{\gamma_1} (\lambda_1 1 - A)^{-1} d\lambda_1 \\ &= \left(\frac{1}{2\pi i} \right)^2 \oint_{\gamma} \oint_{\gamma_1} (\lambda 1 - A)^{-1} (\lambda_1 1 - A)^{-1} d\lambda d\lambda_1 \\ &= \left(\frac{1}{2\pi i} \right)^2 \oint_{\gamma_1} (\lambda_1 1 - A)^{-1} \oint_{\gamma} \frac{1}{\lambda - \lambda_1} d\lambda d\lambda_1 \\ &\quad - \left(\frac{1}{2\pi i} \right)^2 \oint_{\gamma} (\lambda 1 - A)^{-1} \oint_{\gamma_1} \frac{1}{\lambda_1 - \lambda} d\lambda_1 d\lambda, \end{aligned}$$

where we have used the identity

$$U^{-1} - V^{-1} = U^{-1} (V - U) V^{-1}$$

with $U = \lambda 1 - A$ and $V = \lambda_1 1 - A$ to write

$$(\lambda 1 - A)^{-1} (\lambda_1 1 - A)^{-1} = (\lambda - \lambda_1)^{-1} \left((\lambda_1 1 - A)^{-1} - (\lambda 1 - A)^{-1} \right).$$

Finally, since γ_1 is inside γ , we have

$$\oint_{\gamma} \frac{1}{\lambda - \lambda_1} d\lambda = 2\pi i; \quad \oint_{\gamma_1} \frac{1}{\lambda_1 - \lambda} d\lambda_1 = 0,$$

and hence $P^2 = P$, as desired.

It is clear from the definition that P commutes with any bounded operator which commutes with A . In particular, it commutes with A , so the range and null space of P are invariant subspaces for A . We want next to show that the spectrum of A restricted to the range of P is contained in S_1 , i.e., that, if $\lambda \in S_2$, then $(\lambda 1 - A)$ is invertible on the range of P . We prove this by constructing its inverse explicitly as a contour integral. Let γ be as in Proposition 1; let $\lambda \in S_2$; and let

$$B = \frac{1}{2\pi i} \oint_{\gamma} \frac{1}{\lambda - \lambda_1} (\lambda_1 1 - A)^{-1} d\lambda_1.$$

Then B commutes with A ; hence also with P , and

$$\begin{aligned} (\lambda 1 - A)B &= \frac{1}{2\pi i} \oint_{\gamma} \frac{\lambda - \lambda_1 + (\lambda_1 1 - A)}{\lambda - \lambda_1} (\lambda_1 1 - A)^{-1} d\lambda_1 \\ &= \frac{1}{2\pi i} \oint_{\gamma} \left[(\lambda_1 1 - A)^{-1} + \frac{1}{\lambda - \lambda_1} \right] d\lambda_1 \\ &= P + \frac{1}{2\pi i} \oint_{\gamma} \frac{1}{\lambda - \lambda_1} d\lambda_1 \\ &= P, \end{aligned}$$

since $\lambda \in S_2$ and γ winds zero times around S_2 . Thus, if $P\xi = \xi$, $(\lambda 1 - A)B\xi = \xi$, and hence, since A and B commute, the restriction of B to the range of P is a two-sided inverse to the restriction of $(\lambda 1 - A)$ to the same subspace.

Interchanging the roles of S_1 and S_2 shows that the spectrum of the restriction of A to the null space of P is contained in S_2 . But since the union of the spectral of the restrictions of A to S_1 and S_2 must fill up the whole spectrum of A , it follows that the spectrum of the restriction of A to the range of P (respectively the null space of P) must be all of S_1 (respectively, all of S_2).

It remains to prove the uniqueness of P . Thus, let $\hat{P}$ be another projection which commutes with every operator commuting with A and such that the spectrum of A restricted to the range (null space) of $\hat{P}$ is S_1 (S_2). Then P and $\hat{P}$ commute, so $P(1 - \hat{P})$ is again a projection, and it commutes with A since both P and $\hat{P}$ do. Since its range is contained in both the range of P and the null space of $1 - \hat{P}$, the spectrum of A restricted to its range is contained both in S_1 and in S_2 . Since S_1 and S_2 are disjoint, this range must be zero-dimensional—Recall that a bounded operator on a Banach space of non-zero dimension has non-empty spectrum—so $P - P\hat{P} = 0$, i.e., $P = P\hat{P}$. Interchanging the roles of P and $\hat{P}$ gives $\hat{P} = \hat{P}P$, so $P = \hat{P}$, as desired.

General spectral theory gives a bit more information if the operator A is *compact*, i.e., if the closure of the image of the unit ball under A is compact in the norm topology. In this situation it can be shown that the spectrum of A is a finite or countable set with no accumulation point except, perhaps, zero. Thus, there is a spectral projection for each non-zero point of the spectrum. Furthermore, each of these spectral projections is finite-dimensional, i.e., if $\lambda \neq 0$ is in the spectrum of the compact operator A , and if P is the associated spectral projection, then the range of P is a finite-dimensional space invariant for A . The restriction of A to this invariant space is thus a linear operator on a finite-dimensional space whose spectrum consists of the single point λ ; it thus has a Jordan Canonical Form with all diagonal elements equal to λ .

It is illuminating to see what the structure of one-dimensional spectral projections is. Let P be a projection with one-dimensional range, and let ξ_0 be a vector spanning its range. The null space of P is a codimension-one subspace of $\mathcal{X}$, i.e., a hyperplane. Any hyperplane is the null space of some linear functional which is uniquely determined up to multiplication by a non-zero constant. Let ϕ_0 be a linear functional whose null

space is the null space of P . Since ξ_0 is not in the null space of P , $\phi_0(\xi_0) \neq 0$; by multiplying ϕ_0 by an appropriate constant, we can arrange to have $\phi_0(\xi_0) = 1$. We now claim that, for a general vector ξ , $P\xi = \phi_0(\xi)\xi_0$. The proof of this is almost immediate; $\xi \mapsto \phi_0(\xi)\xi_0$ defines a linear operator which is easily checked to be a projection whose range is the set of multiples of ξ_0 , i.e., the range of P , and whose null space is the null space of ϕ_0 , i.e., the null space of P .

Now suppose that P is a spectral projection for some operator A . Then, since the range of P is invariant for A , ξ_0 must be mapped by A to a multiple of itself, i.e., must be an eigenvector of A with some eigenvalue λ . We define the *adjoint* A^* of A to be the mapping of the dual $\mathcal{X}^*$ of $\mathcal{X}$ to itself defined by $(A^*\phi)(\xi) = \phi(A\xi)$ for all $\xi \in \mathcal{X}$ and all $\phi \in \mathcal{X}^*$. Since the null space of ϕ_0 is the same as the null space of P and hence is mapped into itself by A , it follows that $A^*\phi_0$ vanishes on the null space of ϕ_0 and hence that $A^*\phi_0$ is a multiple of ϕ_0 , i.e., that ϕ_0 is an eigenvector of A^* . Furthermore,

$$\lambda = \lambda\phi_0(\xi_0) = \phi_0(A\xi_0) = A^*\phi_0(\xi_0),$$

which implies that the corresponding eigenvalue is again λ . Summing up the results of the above arguments: If λ is an isolated point of the spectrum of A , and if the spectral projection corresponding to $\{\lambda\}$ is one dimensional, then that spectral projection may be written as

$$\xi \mapsto \phi_0(\xi)\xi_0,$$

where ξ_0 is an eigenvector of A with eigenvalue λ , ϕ_0 is an eigenvector of A^* with the same eigenvalue, and these eigenvectors are normalized to that $\phi_0(\xi_0) = 1$. One of the lessons to be drawn from this is that the eigenvectors of A^* are as important as those of A in determining the "spectral representation" of A .

§4. Real vector spaces. There is one more small twist needed for most applications to dynamical systems theory. In the above discussion, we assumed tacitly that $\mathcal{X}$ is a Banach space over the complex numbers; even in the finite-dimensional case, an operator on a real vector space need not have any spectrum on the real axis. In dynamical systems theory, on the other hand, we usually want to work in *real* vector spaces, but would still like to be able to use simple spectral theory.

One satisfactory way to develop spectral theory for operators on a real vector space $\mathcal{X}$ is to work systematically with an associated complex vector space, called the *complexification* $\mathcal{X}_C$ of $\mathcal{X}$. As an additive group, $\mathcal{X}_C$ is simply $\mathcal{X} \times \mathcal{X}$, and multiplication by complex numbers is defined by

$$(\alpha + i\beta)(\xi_1, \xi_2) = (\alpha\xi_1 - \beta\xi_2, \alpha\xi_2 + \beta\xi_1).$$

It is easy to check that this does make $\mathcal{X} \times \mathcal{X}$ into a complex vector space. A norm on $\mathcal{X}$ can be extended to a norm on $\mathcal{X}_C$ by

$$\|(\xi_1, \xi_2)\| = \sup_{\theta} \|\cos \theta \xi_1 + \sin \theta \xi_2\|;$$

again it is easy to check that this does define a norm on the complex vector space $\mathcal{X}_C$ and that the complexified space is complete if the initial real space is.

If A is any real-linear operator on $\mathcal{X}$, we construct a corresponding complex linear operator A_C , called the complexification of A , on $\mathcal{X}_C$, by $A_C(\xi_1, \xi_2) = (A\xi_1, A\xi_2)$; again, it is necessary—but trivial—to check that this does define a complex-linear operator and that its norm is the same as the norm of A . With the construction available, we can then interpret any statement about spectral properties of A as referring instead to A_C .

Although this device permits us to apply spectral theory to operators on real-linear vector spaces, it is not entirely satisfactory because it entails working in the auxiliary space $\mathcal{X}_C$, and this is often inconvenient. There are some statements in spectral theory, however, which, although most conveniently proved by passing through the complexification, can be formulated directly in the original space $\mathcal{X}$. The following proposition gives one such statement:

PROPOSITION 4.1. *Let A be a bounded linear operator on the real Banach space $\mathcal{X}$. Then the spectrum of A (i.e., the spectrum of A_C) is invariant under complex conjugation. Let $\sigma(A) = S_1 \cup S_2$ be a decomposition of the spectrum of A into two disjoint closed subsets invariant under complex conjugation. Then there is a unique projection P on $\mathcal{X}$, commuting with every operator which commutes with A , such that the spectrum of the restriction of A to the range of P is S_1 and the spectrum of the restriction of A to the null space of P is S_2 .*

PROOF: Not every complex-linear operator on $\mathcal{X}_C$ is the complexification of a real-linear operator on $\mathcal{X}$. It is not difficult to see, in fact, that an operator is a complexification if and only if it commutes with the complex conjugation operator $K : (\xi_1, \xi_2) \mapsto (\xi_1, -\xi_2)$. Furthermore, $A_C - \lambda 1$ is invertible if and only if $K^{-1}(A_C - \lambda 1)K = A_C - \bar{\lambda} 1$ is, so the spectrum of A_C is invariant under complex conjugation.

Now let S_1, S_2 be as in the statement of the proposition, and let P_C denote the spectral projection belonging to the subset S_1 of the spectrum of A_C . The uniqueness statements in Proposition 3.1 imply that $K^{-1}P_C K = P_C$, i.e., that P_C is, as its name suggests, the complexification of a real-linear operator P on $\mathcal{X}$. It is nearly immediate that this operator has the properties asserted in the proposition. To prove uniqueness, we note that the complexification of any operator with the asserted properties has the properties which characterize the spectral projection P_C .

A linear operator is A said to be hyperbolic if its spectrum does not intersect the unit circle.

COROLLARY 4.2. *Let A be a hyperbolic linear operator on the real Banach space $\mathcal{X}$. Then there is a splitting $\mathcal{X} = \mathcal{X}_s \oplus \mathcal{X}_u$ into subspaces $\mathcal{X}_s, \mathcal{X}_u$ each of which is invariant under all operators commuting with A , so that the spectrum of the restriction of A to $\mathcal{X}_s$ is inside the unit circle and the spectrum of the restriction of A to $\mathcal{X}_u$ is outside the unit circle.*

We will refer to $\mathcal{X}_s, \mathcal{X}_u$ respectively as the stable and unstable subspaces of A .

EXERCISES:

$$A_C K(\xi_1, \xi_2) = A_C(\xi_1, -\xi_2) = (A\xi_1, -A\xi_2) = K A_C(\xi_1, \xi_2)$$

$$2) \quad \begin{cases} AK = KA & A: \mathcal{X}_C \rightarrow \mathcal{X}_C \\ A(\xi_1, \xi_2) = (A_1 \xi_1, A_2 \xi_2) & \text{linear} \Rightarrow A_1 = A_2 \end{cases}$$

$$1) \quad \begin{cases} A(\xi_1, \xi_2) = (A_1(\xi_1, \xi_2), A_2(\xi_1, \xi_2)) \\ Ak = kA \Rightarrow A_2(\xi_1, \xi_2) = A_2(\xi_2) \\ \text{linear} \Rightarrow A_1(\xi_1, \xi_2) = A_1(\xi_1) \end{cases}$$

1. Evidently, a complex-linear vector space $\mathcal{X}$ is also real-linear and so can re-complexified. This doubles its dimension if it is finite dimensional. Similarly for complex-linear operators. What is the relation between the spectrum and spectral projections of a complex linear operator A and of its re-complexification A_G ?

2. Again let $\mathcal{X}$ be a complex-linear space, A a complex-linear operator on it. Suppose $\mathcal{X}$ is equipped with a complex conjugation, i.e., a real-linear operator K such that $K^2 = 1$ and $K\lambda\xi = \bar{\lambda}K\xi$. What is the relation between the spectra of the complex-linear operator A and the real-linear operator KA ?

3. In the notation of Corollary 4.2, show that $\mathcal{X}_s$ can be characterized as the set of vectors ξ such that $A^n\xi$ is uniformly bounded in n . Give an analogous characterization of $\mathcal{X}_u$ which does not assume that A is invertible.

1. $\mathcal{X} \subset \mathbb{C}\text{-VR}$, $A: \mathcal{X} \rightarrow \mathcal{X}$ bounded linear operator with spectrum $\sigma(A)$

A_G Komplettifizierung of A , $\lambda = \alpha + i\beta$

$$A_G(\xi_1, \xi_2) = A(\xi_1, \xi_2) = (A\xi_1 - \alpha\xi_1 + \beta\xi_2, A\xi_2 - \alpha\xi_2 - \beta\xi_1) = \begin{pmatrix} A - \alpha & \beta \\ -\beta & A - \alpha \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$$

$$\begin{aligned} A_G - \lambda \text{ non invertible} &\iff (A - \alpha)^2 + \beta^2 \text{ non invertible} \\ &\iff (A - \alpha - i\beta)(A - \alpha + i\beta) \text{ non invertible} \\ &\iff (\alpha + i\beta) \in \sigma(A) \text{ oder } (\alpha - i\beta) \in \sigma(A) \end{aligned}$$

that means: $\sigma(A_G) = \sigma(A) \cup \overline{\sigma(A)}$

2. $(KA - \lambda) = K(A - \lambda) = (KA - \bar{\lambda})$

$$\implies \sigma(KA) = \overline{\sigma(A)}$$

$$\sigma(A_G) = \sigma(A) \cup \overline{\sigma(A)}$$

$$\sigma(KA_G) = \sigma((KA)_G) = \sigma(A) \cup \overline{\sigma(A)}$$

3. Choose R so that $\rho(A|_{\mathcal{X}_s}) < R < 1$
 Prop. 2.1: $\exists \|\cdot\|$ $\|A|_{\mathcal{X}_s}\| \leq R$ $\|\cdot\|$ equiv. $\|\cdot\|$. For $\xi \in \mathcal{X}_s$

$$\|A^n \xi\| \rightarrow 0 \quad (n \rightarrow \infty), \text{ because } \|A^n \xi\| \leq R^n$$

$$\implies \|A^n \xi\| \rightarrow 0 \quad (n \rightarrow \infty) \implies A^n \xi \text{ is uniformly bounded}$$

$$\rho(A^n|_{\mathcal{X}_s}) = \rho(A|_{\mathcal{X}_s})^n < R^n \xrightarrow{n \rightarrow \infty} 0$$

$\tilde{A}_n := A^n + 2 \cdot R^n$ is invertible for n sufficiently big

$$\|\tilde{A}_n - A^n\| = 2R^n \rightarrow 0 \quad n \rightarrow \infty$$

Define $\tilde{\mathcal{X}}_u := \{ \xi \mid \tilde{A}_n^{-1} \xi \xrightarrow{n \rightarrow \infty} 0 \}$ Claim: $\tilde{\mathcal{X}}_u = \mathcal{X}_u$

$$\begin{aligned} \xi \in \mathcal{X}_u & \quad (A|_{\mathcal{X}_u} \text{ is invertible}) \cdot (A|_{\mathcal{X}_u})^{-n} \xi \rightarrow 0 \quad n \rightarrow \infty \\ (\|\tilde{A}_n^{-1} \xi - A^{-n} \xi\| \rightarrow 0) & \implies \|\tilde{A}_n^{-1} \xi\| \rightarrow 0, \tilde{A}_n^{-1} \xi \rightarrow 0 \iff \xi \in \tilde{\mathcal{X}}_u \end{aligned}$$

~~$\xi \in \tilde{\mathcal{X}}_u$~~



are defined by an obvious extension of this definition, and we can think of $D^n f$ as taking values in the space of bounded n -linear mappings from $\mathcal{X} \times \cdots \times \mathcal{X}$ (n times) to $\mathcal{Y}$. A function defined on an open set $\mathcal{U}$ whose n -th derivative exists at each point of $\mathcal{U}$, and such that, furthermore, $x \mapsto D^n f(x)$ is continuous, is said to be of class C^n , and a function of class C^n for all n is said to be of class C^∞ .

It will be useful to define also non-integral degrees of differentiability. A function f defined on an open subset $\mathcal{U}$ of a Banach space $\mathcal{X}$, with values in a Banach space $\mathcal{Y}$, is said to be Hölder continuous of order α if there is a constant C such that $\|f(x_1) - f(x_2)\| \leq C \|x_1 - x_2\|^\alpha$ for all x_1, x_2 in $\mathcal{U}$. Here, in principle, α can be any strictly positive real number, but the notion is not very useful for $\alpha > 1$: A Hölder continuous function of order greater than one is everywhere differentiable, with derivative zero, and hence is locally constant. The term Lipschitz continuous is used as a synonym for Hölder continuous of order one.

Now let r be a strictly positive real number, and write $r = n + \alpha$, with n an integer and $0 \leq \alpha < 1$. We then say that a function f defined on $\mathcal{U}$ is of class C^r if it is of class C^n and if $D^n f$ is Hölder continuous of order α on a neighborhood of each point of $\mathcal{U}$.

As already noted, a great deal of standard multidimensional differential calculus extends to the infinite-dimensional situation in a completely straightforward way. Among the most important such results are the chain rule, Taylor's Theorem with remainder in integral form, and the fact that existence and continuity of partial derivatives implies differentiability. Another important result of this kind, to be discussed in the following subsections, is the Implicit Function Theorem. There are, nevertheless, differences between the finite and infinite dimensional situations, and care is needed to avoid pitfalls. Perhaps the most obvious pitfall concerns results which depend on the local compactness of the domain space. For example, a continuous function on an infinite dimensional Banach space need not be bounded on bounded sets; in fact, even a function which is C^∞ on the whole space need not be bounded on bounded sets (See Exercise 1.1). Similarly, continuity apparently does not imply local uniform continuity, i.e., a continuous function apparently does not need to be uniformly continuous on a neighborhood of each point in its domain of definition.

A more subtle difference concerns the existence of smooth partitions of unity. On a given Banach space, there need not exist non-zero smooth functions vanishing outside bounded sets. Such functions of course exist on Hilbert spaces and other spaces whose norms are smooth away from zero. But many spaces arising naturally do not have smooth norms. The most complete results of which I am aware concern the sequence spaces ℓ^p . It is easy to see that $\|\cdot\|_p^p$ is of class C^n on ℓ^p for any integer $n < p$. If p is an even integer, then $\|\cdot\|_p^p$ is everywhere p times differentiable, and that the p -th derivative is constant, so $\|\cdot\|_p^p$ is in fact of class C^∞ . (If, on the other hand, p is an odd integer, the p -th derivative fails to exist on the dense set of points (x_n) for which $x_n = 0$ for some n .) If we exclude the trivial case of even p , it is not hard to see that $\|\cdot\|_p^p$ is nowhere n -times differentiable for any $n > p$. The fact that the standard norm has only a limited number of derivatives does not in any obvious way imply that there are no smoother equivalent norms, to say nothing of nonconstant smoother function of bounded support.

d. bedeutet. f ist zwar Gateaux diffbar aber nicht Fréchet diffbar

CHAPTER III. DIFFERENTIAL CALCULUS ON BANACH SPACES

Much of elementary multidimensional differential calculus extends in a straightforward way to Banach spaces. In this section, we sketch this extension, concentrating mostly on pointing out the few places where the standard arguments do not work.

§1. Differentiable mappings. Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, f a mapping defined on some subset $\mathcal{D}$ of $\mathcal{X}$ and taking values in $\mathcal{Y}$. We say that f is differentiable at a point x of $\mathcal{D}$ provided that x is an interior point of $\mathcal{D}$ and that there is a bounded linear operator $L : \mathcal{X} \rightarrow \mathcal{Y}$ such that

$$f(x + \xi) - f(x) - L\xi = o(\|\xi\|) \quad \text{as } \xi \rightarrow 0$$

i.e., such that for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|f(x + \xi) - f(x) - L\xi\| \leq \epsilon \|\xi\| \quad \text{when } \|\xi\| \leq \delta.$$

If this is the case, the linear operator L is unique, and is called the derivative of f at x ; we will denote it by $Df(x)$.

The kind of differentiability we have defined above is sometimes called Fréchet differentiability, to distinguish it from the weaker notion of Gateaux differentiability, which means that $t \mapsto f(x + t\xi)$ is differentiable at 0 (or perhaps just right-differentiable at 0) for each ξ in $\mathcal{X}$, but without any uniformity in ξ . There is also a notion of weak Gateaux differentiability, which means that $t \mapsto \phi(f(x + t\xi))$ is differentiable at 0 for each ϕ in the dual of $\mathcal{Y}$ and each ξ in $\mathcal{X}$. The difference between the various notions is not very great. For example, if f is Gateaux differentiable at each point of some non-empty open set; if, for each x in that set,

$$\xi \mapsto L_x(\xi) \equiv \left. \frac{df(x + t\xi)}{dt} \right|_{t=0}$$

is linear and bounded and if $x \mapsto L_x \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is continuous (in the norm topology), then f is also Fréchet differentiable and $L_x = Df(x)$. We will not take the time here to investigate the extent to which the results we discuss can be generalized to Gateaux differentiable functions; Fréchet differentiability generally suffices for the purposes of dynamical systems theory.

We have explained what is meant by the derivative of a function at a point. The function Df is defined at exactly those points of $\mathcal{X}$ where f is differentiable, and its value at such a point x is $Df(x)$, which is in $\mathcal{L}(\mathcal{X}, \mathcal{Y})$, the space of bounded linear mappings from $\mathcal{X}$ to $\mathcal{Y}$. Since $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is itself a Banach space, it makes sense to speak of the derivative of the function Df , and this derivative is what is called the second derivative D^2f of f . As defined, this second derivative takes values in $\mathcal{L}(\mathcal{X}, \mathcal{L}(\mathcal{X}, \mathcal{Y}))$, but it is often useful to use the observation that this space is naturally identified with the space of all continuous bilinear mappings from $\mathcal{X} \times \mathcal{X}$ to $\mathcal{Y}$ and to think of the second derivative as taking values in this space of bilinear mappings. Higher derivatives

O.E. Lanford: Dynamical Systems (Preliminary Version)

Printed November 11, 1987

$\Gamma \quad \eta := f(x + r(s)) - f(x) - L_x r(s) \quad \text{mit } r(s) \xrightarrow{s \rightarrow 0} 0$

$$\begin{aligned} \|f(x + \eta) - f(x) - L_x \eta\| &= \|f(x + r(s) + \xi) - f(x) - L_{x+r(s)} \xi - (L_{x+r(s)} - L_x) \xi - L_x r(s)\| \\ &\leq \|f(x + r(s) + \xi) - f(x) - L_{x+r(s)} \xi\| + \|(L_{x+r(s)} - L_x) \xi\| + \|L_x r(s)\| \\ &\leq \underbrace{\epsilon + \|r(s)\| + \frac{\delta}{2}}_{\substack{\leq \frac{\delta}{2} + \|r(s)\| + \frac{\delta}{2} \\ \text{da } f \text{ Gateaux diffbar}}} + \underbrace{\|L_{x+r(s)} - L_x\| \|\xi\|}_{\substack{\leq \frac{\delta}{2} + \|r(s)\| + \frac{\delta}{2} \\ \text{da } L_x \text{ stetig}}} + \underbrace{\|L_x\| \|r(s)\|}_{\substack{\leq C \cdot \frac{\delta}{2} \\ \text{da } L_x \text{ beschränkt}}} \\ &\leq C_\epsilon \cdot \|\eta\| \quad \text{falls } \|\eta\| \leq \delta \end{aligned}$$

$G \subseteq X$ offen, $x \in G$
 f Gateaux diffbar in G
 L_x ist linear, beschränkt
 $x \mapsto L_x$ stetig
 $\Rightarrow f$ ist Fréchet diffbar in G , $Df(x) = L_x$

Nevertheless, it has been shown by R. Bonic and J. Frampton ("Differentiable functions on certain Banach spaces", *Bull. Amer. Math. Soc.*, **71**, (1965) 393-395.) that, on ℓ^p , p not an even integer, there exist no non-constant functions of class C^n , $n > p$, which vanish outside the unit ball. They prove this as a corollary of the following remarkable "Maximum Modulus Principle":

THEOREM (BONIC AND FRAMPTON). *Let $p \geq 1$ be a real number which is not an even integer; let D be a connected bounded open set in ℓ^p , and let f be a real-valued function defined and continuous on the closure of D and n times differentiable on D itself for some integer $n \geq p$. Then the image under f of the boundary of D is dense in the image under f of D itself.*

EXERCISES:

1.1 Construct a real-valued function f , defined and of class C^∞ everywhere on a Banach space $\mathcal{X}$, which vanishes outside the unit ball but which is nevertheless unbounded.

1.2 (Easy) If $\mathcal{X}, \mathcal{Y}$ are Banach spaces show that $A \mapsto A^{-1}$ is differentiable on the (open) set of invertible linear mappings $\mathcal{X} \rightarrow \mathcal{Y}$ with bounded inverses, and compute its derivative. The main point of this exercise is to get the "functorial" aspects right, i.e. the derivative is to be exhibited as a linear mapping from $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ into $\mathcal{L}(\mathcal{Y}, \mathcal{X})$.

1.3 We first have to establish some notation, and we do it only in a simple case: For n a non-negative integer, let $C^n(-1, 1)$ denote the space of all real valued functions of class C^n on $(-1, 1)$ whose n -th derivatives are bounded on that interval. Equipped with the norm

$$\|f\|_{C^n} \equiv \max \left(\sup_{-1 < x < 1} \{|f(x)|\}, \sup_{-1 < x < 1} \{|f^{(n)}(x)|\} \right)$$

$C^n(-1, 1)$ becomes a Banach space. (This construction can evidently be generalized to define Banach spaces of C^r mappings, r not necessarily an integer, from an open set in one Banach space into another.)

Now consider the mapping $f \mapsto f \circ f$ which sends the open unit ball of $C^n(-1, 1)$ into itself. Show that this mapping is nowhere differentiable from $C^n(-1, 1)$ to itself, but that it is differentiable—in fact, of class C^1 —from the open unit ball of $C^n(-1, 1)$ into $C^{n-1}(-1, 1)$.

The reason for the non-differentiability will probably strike you as an inessential pathology, but the non-differentiability of composition operators turns out to have serious consequences.

§2. Fixed point theorems and Newton's method.

The starting point for this subsection is the Contraction Mapping Principle:

THEOREM 2.1 (CONTRACTION MAPPING THEOREM). *Let (X, ρ) be a complete metric space; f a mapping of X to itself such that, for some $\kappa < 1$,*

$$\rho(f(x_1), f(x_2)) \leq \kappa \rho(x_1, x_2) \quad \text{for all } x_1, x_2 \text{ in } X$$

Then f has exactly one fixed point $\bar{x} \in X$, and, for any $x_0 \in X$, $f^n(x_0) \xrightarrow{n \rightarrow \infty} \bar{x}$.

The essential observation is that

$$\rho(f^{n+1}(x_0), f^n(x_0)) \leq \kappa^n \rho(f(x_0), x_0).$$

From this it also follows that

$$\rho(x_0, \bar{x}) \leq (1 - \kappa)^{-1} \rho(x_0, f(x_0)).$$

This estimate implies immediately the following corollary, which shows that the fixed point varies continuously with the map:

COROLLARY 2.2. *Let (X, ρ) be as in Theorem 2.1; f_1 and f_2 two mappings of X to itself, each contractive with contraction constant κ ; $\bar{x}_1$ and $\bar{x}_2$ their respective fixed points. Then*

$$\rho(\bar{x}_1, \bar{x}_2) \leq (1 - \kappa)^{-1} \rho(\bar{x}_1, f_2(\bar{x}_1)) \leq (1 - \kappa)^{-1} \sup_x \rho(f_1(x), f_2(x)).$$

For a differentiable mapping (on a convex set), the obvious way to check contractivity is to estimate the norm of the derivative:

REMARK: Let x_1, x_2 be two points of a Banach space $\mathcal{X}$, and write $[x_1, x_2]$ for the straight line segment joining them. Let f be defined and differentiable on $[x_1, x_2]$. Then

$$\|f(x_1) - f(x_2)\| \leq \sup_{x \in [x_1, x_2]} \{\|Df(x)\|\} \|x_1 - x_2\|.$$

$$\sup_{x \in [x_1, x_2]} \|Df(x)\| \geq \sup_{x \in [x_1, x_2]} \frac{\|f(x) - f(x_1)\|}{\|x - x_1\|} \geq \frac{\|f(x_2) - f(x_1)\|}{\|x_2 - x_1\|} = \frac{\|f(x_2) - f(x_1)\|}{\|x_2 - x_1\|}$$

To apply the Contraction Mapping Theorem to a differentiable mapping on a ball in a Banach space, it is not enough to know that the derivative of the mapping has norm less than one on that ball; one also needs to know that the ball is mapped into itself. The following extremely simple proposition provides a way to check this which is often useful in practice:

PROPOSITION 2.3. *Let f be differentiable on the (closed) ball of radius ρ about x_0 in the Banach space $\mathcal{X}$, and suppose there is a constant $\kappa < 1$ such that $\|Df\| \leq \kappa$ on that ball. Then if*

$$\|f(x_0) - x_0\| \leq (1 - \kappa)\rho$$

f has exactly one fixed point in that ball.

PROOF: We have only to check that f maps the ball into itself. But $\|x - x_0\| \leq \rho$ implies

$$\|f(x) - x_0\| \leq \|f(x) - f(x_0)\| + \|f(x_0) - x_0\| \leq \kappa\rho + (1 - \kappa)\rho = \rho,$$

as desired.

The Contraction Mapping Theorem is a simple and powerful tool, but it does have the disadvantage that it only works for contractions. By using the circle of ideas around

Newton's method, its range of applicability can be extended considerably. Let us review the idea of Newton's method. Let f be a mapping defined on an open subset of a Banach space $\mathcal{X}$, with values in a Banach space $\mathcal{Y}$. To fix ideas, suppose that f is of class $\mathcal{C}^2$ and that $Df(x)$ is invertible for each x . We want to find a zero of f , i.e., a $\bar{x} \in \mathcal{X}$ such that $f(\bar{x}) = 0$. The idea is as follows: If we have an approximate zero x_0 , then, near x_0 , $f(x)$ may be approximated by the affine mapping $f(x_0) + Df(x_0)(x - x_0)$, which has a zero at $x_0 - Df(x_0)^{-1}f(x_0)$. We can then hope that this zero of the affine approximation to f is a better approximation to a zero of f itself than is the starting point x_0 and repeat the process starting there. In other words: We note that a zero of f is the same thing as a fixed point of

$$\Phi(x) = x - Df(x)^{-1}f(x),$$

and we seek a fixed point for Φ by iteration. If we suppose that $\bar{x}$ is a zero of f , we can see that this iteration must converge if we start near enough to $\bar{x}$ by computing the derivative of Φ :

$$\begin{aligned} D\Phi(x) &= 1 - D(Df(x)^{-1})f(x) + Df(x)^{-1}Df(x) \\ &= -D(Df(x)^{-1})f(x), \end{aligned}$$

which vanishes at $\bar{x}$. Thus, simply by continuity, Φ is a contraction on a sufficiently small ball about the fixed point $\bar{x}$, so iterating Φ starting at any point in such a ball gives a sequence which converges to $\bar{x}$. Furthermore, the convergence is asymptotically extremely fast, provided that f is of class $\mathcal{C}^3$. In fact, if f is of class $\mathcal{C}^3$, then Φ is of class $\mathcal{C}^2$, and if $\|D^2\Phi\| \leq K$ on a ball of radius ρ about $\bar{x}$, then, if $\|x - \bar{x}\| \leq \rho$,

$$\|\Phi(x) - \bar{x}\| = \|\Phi(x) - \Phi(\bar{x})\| \leq (K/2)\|x - \bar{x}\|^2,$$

i.e., the error after $n + 1$ iterations is of the order of the *square* of the error after n iterations, i.e., *the number of decimal places of accuracy doubles with each iteration*.

In the above discussion we have, for simplicity, started from the assumption of the existence of a zero of f . It is not difficult, however, to use the preceding proposition to give a concrete sufficient condition for the existence of a zero which also shows that it can be found by Newton's method. The formulation, admittedly a little awkward, is as follows:

PROPOSITION 2.4. *Let f be of class $\mathcal{C}^2$, with bounded second derivative, on the ball of radius ρ about x_0 . Assume that $Df(x_0)$ is invertible with bounded inverse. We will use the following notation: $M = \|Df(x_0)^{-1}\|$, $\epsilon = \|f(x_0)\|$, $K_1 = \sup_{\|x-x_0\| \leq \rho} \|Df(x)\|$, $K_2 = \sup_{\|x-x_0\| \leq \rho} \|D^2f(x)\|$. We now assume*

$$(1) \quad K_1 \rho M < 1$$

$$(2) \quad \kappa \equiv K_2 \left(\frac{M}{1 - K_1 \rho M} \right)^2 (\epsilon + K_1 \rho) < 1$$

$$(3) \quad M\epsilon \leq \rho(1 - \kappa).$$

Then f has a unique zero in the ball of radius ρ about x_0 .

PROOF: The proof consists simply in checking that the various hypotheses suffice to guarantee that

$$\Phi(x) \equiv x - (Df(x))^{-1} f(x)$$

is defined on the ball of radius ρ about x_0 and satisfies the hypotheses of Proposition 2.3. We note, first of all, that for x in that ball, $\|Df(x) - Df(x_0)\| \leq K_1\rho$, and hence the Neumann series implies that $Df(x)$ is invertible and that

$$\|(Df(x))^{-1}\| \leq \frac{M}{1 - K_1\rho M}.$$

Second, as we have already noted, $D\Phi(x) = -D(Df(x)^{-1})f(x)$. This actually calls for a little interpretation. By definition, $D\Phi(x)$ is a linear operator from $\mathcal{X}$ to $\mathcal{Y}$, whereas $D(Df(x)^{-1})$ is a linear mapping from $\mathcal{X}$ to $\mathcal{L}(\mathcal{Y}, \mathcal{X})$; what the above equation means is that, for any $\xi \in \mathcal{X}$,

$$D\Phi(x)\xi = - \underbrace{(D(Df(x)^{-1})(\xi))}_{\in \mathcal{L}(\mathcal{Y}, \mathcal{X})} f(x).$$

Again using the Neumann series, one checks easily that

$$D(Df(x)^{-1})(\xi)(\eta) = -Df(x)^{-1}D^2f(x)(Df(x)^{-1}\xi, Df(x)^{-1}\eta).$$

Also, $\|f(x)\| \leq \epsilon + K_1\rho$. Combining all these remarks we see that $\|D\Phi(x)\| \leq \kappa$ for all x with $\|x - x_0\| \leq \rho$. Furthermore, $\|\Phi(x_0) - x_0\| = \|Df(x_0)^{-1}f(x_0)\| \leq M\epsilon$, so our final assumption is just what is needed to guarantee that the final assumption of the Proposition 2.3 holds.

The preceding proposition is inconvenient to apply because it requires estimating the second derivative. For this reason, it is sometimes advantageous to replace Newton's method by another approximation scheme which is easier to control. The idea is to redefine Φ by replacing $Df(x)$ by some fixed constant invertible operator Γ ; if Γ is a good enough approximation to $Df(x)$ for all x in the ball of radius ρ about x_0 , the redefined Φ is still a contraction on this ball. The precise conditions which have to be satisfied are quite simple:

PROPOSITION 2.5. Let $B_\rho(x_0)$ denote the closed ball with radius ρ centered at x_0 ; let f be a mapping defined on a neighborhood of $B_\rho(x_0)$ taking values in $\mathcal{Y}$ and differentiable at each point of $B_\rho(x_0)$; and let Γ be an invertible linear mapping $\mathcal{X} \rightarrow \mathcal{Y}$. If there exists $\kappa < 1$ such that

- (1) $\|\Gamma^{-1}(Df(x) - \Gamma)\| \leq \kappa$ for all $x \in B_\rho(x_0)$.
- (2) $\|\Gamma^{-1}f(x_0)\| \leq (1 - \kappa)\rho$

PROOF: To prove (1), we simply apply Proposition 2.5 to the mapping $f^{(z)} : x \mapsto f(x, z)$, with 0 as an approximate zero, $\Gamma = D_1 f(0, 0)$, and $z \in \mathcal{Z}$ as a "parameter." We have to choose ρ and κ . At $z = 0$, we have $\Gamma^{-1}(Df^{(z)}(0) - \Gamma) = 0$ and $\Gamma^{-1}f^{(z)}(0) = 0$. By continuity, there exist $\delta_x, \delta_z > 0$ such that $\|\Gamma^{-1}(Df^{(z)}(x) - \Gamma)\| \leq 1/2$ for $\|x\| \leq \delta_x$ and $\|z\| \leq \delta_z$. By making δ_z smaller if necessary, we can also ensure that $\|\Gamma^{-1}f^{(z)}(0)\| \leq \delta_x/2$ if $\|z\| \leq \delta_z$. Proposition 2.5 now applies, with $\kappa = 1/2$ and $\rho = \delta_x$, for each z with $\|z\| \leq \delta_z$ and proves (1).

We now turn to the proof of (2). Note that the formula for $DX(z)$ would follow immediately by differentiating the identity $f(X(z), z) = 0$, if we knew that X was differentiable. Proving that X is differentiable, assuming that f is, is an instructive exercise in the definition of the derivative. To get started, we need to know that X is continuous. This follows from the proof of Proposition 2.5 together with Corollary 2.2. To prove differentiability, fix z , and, for sufficiently small δz , define $\delta X = X(z + \delta z) - X(z)$. From $f(X, z) = 0$, $f(X + \delta X, z + \delta z) = 0$, and the definition of derivative, we get

$$D_2 f(X, z) \delta z + D_1 f(X, z) \delta X = o(\|\delta z\| + \|\delta X\|).$$

Since we have set things up so that we know that $D_1 f(X, z)$ is invertible, we get

$$\delta X = -(D_1 f(X, z))^{-1} D_2 f(X, z) \delta z + o(\|\delta z\| + \|\delta X\|).$$

A first application of this estimate together with the continuity of X shows that $\|\delta X\| = O(\|\delta z\|)$, and hence $o(\|\delta z\| + \|\delta X\|)$ can be replaced simply by $o(\|\delta z\|)$. A second application of the estimate shows, directly from the definition of derivative, that X is differentiable at z and that its derivative is as indicated in (2).

Once the formula for $DX(z)$ is established, it is straightforward to prove—by differentiating that formula repeatedly—that X is n times differentiable if f is. This process also leads to an explicit (if unwieldy) formula for the n -th derivative of X in terms of the partial derivatives (of total order not greater than n) of f . From this formula it follows at once that if the partial derivatives of f of order n are Hölder continuous of order α , then so is the n -th derivative of X .

COROLLARY 3.2 (INVERSE FUNCTION THEOREM). *Let f be a continuously differentiable mapping defined on a neighborhood of $x_0 \in \mathcal{X}$, taking values in $\mathcal{Y}$. Assume that $Df(x_0)$ is invertible. Then f maps any sufficiently small ball about x_0 homeomorphically onto a neighborhood of $f(x_0)$. The inverse homeomorphism is continuously differentiable, and of class C^r if f is. We have*

$$Df^{-1}(f(x)) = (Df(x))^{-1}.$$

PROOF: To prove the local invertibility of f , we need to show that $f(x) = y$, regarded as an equation to be solved for x with y fixed, has a unique solution near x_0 for each y sufficiently near to $f(x_0)$. This follows from the Implicit Function Theorem applied to $(y, x) \mapsto f(x) - y$. Differentiability of the inverse function also follows from the Implicit

then f has a unique zero in $B_\rho(x_0)$.

In fact, the exact conditions of the preceding proposition are not often very important, but it is simpler to state them explicitly than to formulate qualitatively what it is about them which is actually used.

PROOF: The proof is almost immediate. Put $\Phi(x) = x - \Gamma^{-1}f(x)$. Evidently, a fixed point of Φ is the same thing as a zero of f . By a very simple computation,

$$D\Phi(x) = \Gamma^{-1}(\Gamma - Df(x)),$$

so the two conditions of Proposition 2.5 are simply reformulations of those of Proposition 2.3.

It should be noted that, although this modified Newton's method is easier to control, and requires less regularity, than Newton's method itself, it does not give the extremely rapid convergence that Newton's method does (unless $\Gamma = Df(\bar{x})$, which is hard to arrange in practice).

§3. The Implicit Function Theorem.

The ideas of the preceding section can be used to prove the Implicit Function Theorem:

THEOREM 3.1 (IMPLICIT FUNCTION THEOREM). *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be Banach spaces, f a function defined on a neighborhood $\mathcal{U}$ of 0 in $\mathcal{X} \times \mathcal{Z}$, with values in $\mathcal{Y}$. Assume:*

- (1) *For each $(x_0, z_0) \in \mathcal{U}$, $f(x, z_0)$ is differentiable in x on a neighborhood of x_0 , and the partial derivative $D_1f(x, z)$ varies continuously with $(x, z) \in \mathcal{U}$.*
- (2) *$f(0, 0) = 0$.*
- (3) *$D_1f(0, 0)$ is invertible.*

Then

- (1) *If $\delta_x > 0$ is sufficiently small, there exists $\delta_z > 0$ such that, for each z with $\|z\| < \delta_z$, there is exactly one x with $\|x\| < \delta_x$ such that $f(x, z) = 0$. (Intuitively: For each small z , the equation $f(x, z) = 0$ has a unique small solution.)*

We will from now on suppose that δ_z, δ_x have been fixed once and for all so that the condition of (1) holds, and are also small enough so that $D_1f(x, z)$ is invertible for all (x, z) in

$$\Delta \equiv \{(x, z) : \|z\| < \delta_z \text{ and } \|x\| < \delta_x\}.$$

We will denote the x corresponding to z by $X(z)$

- (2) *If, in addition, f is differentiable on Δ , X is differentiable on*

$$\Delta_z \equiv \{z : \|z\| < \delta_z\},$$

and

$$DX(z) = -(D_2f(X(z), z))^{-1} D_1f(X(z), z).$$

- (3) *If f is of class C^r , $r \geq 1$ on Δ , then X is of class C^r on Δ_z .*

Function Theorem; the formula for the derivative can be derived from the formula given above for the derivative of an implicitly defined function, but it is easier simply to differentiate the identity $f^{-1}(f(x)) = x$ and apply the chain rule.

~~The~~ Theorem 3.1 is frequently referred to as the “soft” Implicit Function Theorem, to distinguish it from the much deeper “hard” Implicit Function Theorems of Nash and Moser.

LECTURES ON DYNAMICAL SYSTEMS

Chapter IV

Oscar E. Lanford III

ETH-Zürich

Disclaimer: This is a preliminary text for a course of lectures given in the academic year 1987-88; it is intended for the use of the students in the course and not for wider distribution. While I will be happy if anyone else finds it useful, the reader should be aware that this manuscript is not in final form and should therefore be used with care. One particularly glaring defect is the lack of any attempt to attribute credit for the results discussed.

O. E. Lanford

Mathematik
ETH-Zentrum
8092 Zürich (Switzerland)

November 1987

§1. Orientation. Let f be a map. A periodic point of f means a point x_0 such that $f^p(x_0) = x_0$ for some strictly positive integer p . The smallest $p > 0$ for which this is true is called the period of f . A periodic cycle of f means a sequence $x_0, f(x_0), \dots, f^{p-1}(x_0)$, where x_0 is a periodic point of period p . Study of the behavior of a map near a periodic point can be reduced to study of the behavior near a fixed point by replacing f by f^p . Somewhat more precisely: The motion induced by f near a periodic cycle of period p is, in a natural sense, the semi-direct product of a cyclic permutation of order p with the motion induced by f^p near any point of the cycle. The cyclic permutation is “trivial”; we therefore focus our attention on the case of a fixed point.

Thus, let f be a differentiable mapping defined on an open set in a Banach space $\mathcal{E}$, with a fixed point x_0 . Near this fixed point, f is well approximated by the linear mapping $x \mapsto f(x_0) + Df(x_0)(x - x_0)$. The question we want to address in this section is: To what extent is the dynamics of f near the fixed point like the dynamics of $Df(x_0)$?

One nice situation, which happens reasonably frequently, is that, near the fixed point, f differs from $Df(x_0)$ only by a change of coordinates. More precisely, what we mean is that there is an invertible mapping ϕ , defined on a neighborhood of 0 and mapping 0 to x_0 , such that $f = \phi Df(x_0) \phi^{-1}$ on a neighborhood of 0. We speak of such a ϕ as a linearization of f (near x_0), or we say that ϕ conjugates $Df(x_0)$ to f (near x_0). It matters a lot how much regularity ϕ is required to have. In order to be of any interest at all, a linearization must be at least a homeomorphism, in which case we speak of a topological linearization. As we shall see later on, topological linearizations exist quite often—not always—but the information they give about the dynamics near the fixed point is too coarse to be very useful.

Linearizations with some degree of differentiability are more useful than ones which are merely topological, but conditions for their existence are more complicated. For example, among maps of class C^∞ , “most” (in a sense which we will not define here) admit C^∞ linearizations, but a dense set of them do not. Because the existence or non-existence of a smooth linearization depends on delicate properties of the map, these linearizations are not very effective tools for analyzing general smooth maps. Nevertheless, it can be a helpful *heuristic* device to imagine that linearizations exist. Judiciously chosen “obvious” facts about linear maps often extend to give useful and less trivial properties of general maps. For example: Suppose f admits a C^1 linearization, with C^1 inverse, at the fixed point x_0 . Without loss of generality, we can assume that $D\phi(0) = 1$.* Now if $\mathcal{E}_1$ is a linear subspace of $\mathcal{E}$ invariant under $Df(x_0)$, $\phi\mathcal{E}_1$ is an invariant manifold for f . (Actually, since we are only supposing that ϕ is locally defined, this manifold is in general only locally invariant in an appropriate sense. This manifold passes through x_0 , and its tangent space there is exactly the subspace $\mathcal{E}_1$. Thus, if a smooth linearization exists, invariant subspaces for $Df(x_0)$ —about which spectral theory provides detailed information—give rise to locally invariant manifolds for f . It is then natural to ask

*In general, we get from the chain rule that $Df(x_0) = D\phi(0)Df(x_0)D\phi(0)^{-1}$, i.e., $D\phi(0)$ commutes with $Df(x_0)$. Hence, replacing ϕ by $\phi \circ [D\phi(0)^{-1}]$ gives a new linearization whose derivative at 0 is the identity.

to what extent this correspondence can be established *without* assuming the existence of a smooth linearization. It turns out, in fact, that a strict correspondence between invariant subspaces and invariant manifolds breaks down in general, for roughly the same reasons as does the existence of smooth linearizations, but that there are some “dynamically defined” invariant subspaces for $Df(x_0)$ to which there do correspond invariant manifolds. The study of these questions will be the main topic of this chapter, but, before beginning on it, we will look at a simpler relation between the dynamics of f near a fixed point x_0 and the dynamics of $Df(x_0)$, manifested in what is called “linear stability analysis”.

§2. Linear stability analysis. A fixed point x_0 for a mapping f will be said to be attracting if there exists a neighborhood $\mathcal{U}$ of x_0 whose images under f shrink down uniformly to $\{x_0\}$ in the sense that, for any neighborhood $\mathcal{V}$ of x_0 , $f^n\mathcal{U} \subset \mathcal{V}$ for all sufficiently large n . (Although we have not explicitly assumed that $\mathcal{U}$ is mapped into itself by f , we can add this assumption without making the definition more restrictive; simply replace $\mathcal{U}$ by $\bigcup_{n=0}^{\infty} f^n\mathcal{U}$.)

This definition calls for two comments on terminology. The first comment is that what we call an attracting fixed point is commonly called a stable or (in the older literature) asymptotically stable fixed point. The reason for this terminology is that, if we imagine perturbing an attracting fixed point slightly, we get another initial point whose orbit converges back to that of the unperturbed fixed point. The term “stable” is, however, overused in dynamical systems theory, and I have preferred to reserve it primarily to refer to properties of a system which are preserved under small changes in that dynamical system (as opposed to properties of an orbit which are preserved under small changes in initial condition). The second comment is that the above definition of attracting fixed point is never satisfied by a Hamiltonian dynamical system (at least on a finite-dimensional state space); such systems, because they preserve volumes in state space, cannot map a set $\mathcal{U}$ of nonzero volume into an arbitrarily small subset of itself. There is another definition of stability more appropriate for Hamiltonian systems: A fixed point x_0 is said to be Lyapunov stable if, for any neighborhood $\mathcal{U}$ of x_0 , there exists a (smaller) neighborhood $\mathcal{V}$ such that $f^n\mathcal{V} \subset \mathcal{U}$ for all $n \geq 0$, i.e., any orbit starting sufficiently near to x_0 stays near forever.

Whether or not a fixed point is attracting can often be determined from the spectrum of the derivative at that fixed point.

PROPOSITION 2.1. *Let f be a mapping of class C^1 ; x_0 a fixed point of f .*

- (1) *If the spectral radius of $Df(x_0)$ is strictly less than one, x_0 is attracting.*
- (2) *If $Df(x_0)$ is a compact operator, and if the spectral radius of $Df(x_0)$ is strictly greater than one, x_0 is not attracting.*

Thus, at least in the case where $Df(x_0)$ is compact—and in particular in the finite-dimensional situation which is our principal interest—the spectral radius of $Df(x_0)$ determines whether or not x_0 is attracting unless it is exactly one. I do not know whether (2) is true without some assumption like compactness to permit the splitting off of the part of the spectrum with largest modulus.

PROOF: We will prove only (1). We choose a norm such that $\|Df(x_0)\| < 1$; then, by continuity of the derivative, there exist $\kappa < 1$ and $\rho > 0$ such that $\|Df(x)\| < \kappa$ for all x with $\|x - x_0\| < \rho$. We take $\mathcal{U}$ to be the ball of radius ρ about x_0 ; then for $n = 1, 2, \dots$, $f^n\mathcal{U}$ is contained in the ball of radius $\kappa^n\rho$ about x_0 . Any neighborhood $\mathcal{V}$ of x_0 contains the ball of radius $\kappa^n\rho$ for all sufficiently large n . ■

The analogue of the above considerations for flows is as follows: A stationary solution of a differential equation

$$\frac{dx}{dt} = X(x)$$

means an x_0 such that $X(x_0) = 0$. Then x_0 is a fixed point for the solution flow f^t for all t . (Recall: It is one of our standing hypotheses for flows that $X(x)$ possesses enough regularity to guarantee uniqueness of solutions.) A stationary solution x_0 is attracting if there is a neighborhood $\mathcal{U}$ of x_0 such that, for every neighborhood $\mathcal{V}$ of x_0 , $f^t(\mathcal{U}) \subset \mathcal{V}$ for all sufficiently large t . By the uniform continuity of $\{f^t : 0 \leq t < 1\}$, it is enough to guarantee that for every neighborhood $\mathcal{V}$, $f^n\mathcal{U} \subset \mathcal{V}$ for all sufficiently large integers n , i.e., a stationary solution of a differential equation is attracting if and only if it is an attracting fixed point for the corresponding time-one solution map.

The spectrum of the derivative of the time-one solution mapping at a stationary solution is also easy to compute—given the spectrum of the linearization $DX(x_0)$ of the differential equation—using the first variational equation. The idea is as follows: Let $x(t)$, $0 \leq t \leq t_0$ be a solution of the differential equation. Write $V(t)$ for $Df^t(x(0))$. Then $V(t)$ satisfies the linear differential equation

$$\frac{dV}{dt} = V(t)DX(x(t)) \quad \text{with initial condition } V(0) = 1.$$

$$\begin{aligned} \frac{dx}{dt} &= X(x) \\ \frac{d(f^t(x_0))}{dt} &= X(f^t(x_0)) \\ \frac{d}{dt} Df^t(x_0) &= Df^t(X(x_0)) \\ &= Df^t(X(x_0)) \end{aligned}$$

This equation, furthermore, determines $V(t)$. In the simple case where $x(t) \equiv x_0$, we get

$$V(t) = e^{tDX(x_0)},$$

and hence we have

COROLLARY 2.2. A stationary solution x_0 of the differential equation

$$\frac{dx}{dt} = X(x(t))$$

is attracting if the spectrum of $DX(x_0)$ is contained in the open left half-plane; in finite dimensions, such a stationary solution is not attracting if the spectrum of $DX(x_0)$ is not contained in the closed left half-plane.

§3. The stable manifold: Setting up the problem. We will say that a fixed point x_0 for a differentiable mapping f defined on an open set in a Banach space $\mathcal{E}$ is hyperbolic if $Df(x_0)$ is a hyperbolic linear operator. (Recall that in §II.4 we defined a linear operator to be hyperbolic if its spectrum is disjoint from the unit circle.) By Corollary

set of points near x_0 whose orbits remain near x_0 and in fact converge to it with distance decreasing faster than $(\rho_- + \epsilon)^n$ for all $\epsilon > 0$.

(3) $\rho_- = 1$, in which case we speak of a center-stable manifold.

We will use the following notation: We write the space on which f acts as $\mathcal{E}_- \oplus \mathcal{E}_+$, and we put the fixed point x_0 at the origin. When referring to points of $\mathcal{E}_- \oplus \mathcal{E}_+$, we will generally use x (with subscripts, accents, etc.) to refer to the $\mathcal{E}_-$ component and y to refer to the $\mathcal{E}_+$ component. We denote by Λ_- and Λ_+ the restrictions of $Df(0)$ to $\mathcal{E}_-$ and $\mathcal{E}_+$ respectively. The spectral radius of Λ_- is ρ_- ; Λ_+ is invertible and the spectral radius of Λ_+^{-1} is ρ_+^{-1} . We can therefore equip $\mathcal{E}_-$ and $\mathcal{E}_+$ with norms such that

$$\|\Lambda_- \| \|\Lambda_+^{-1}\| < 1.$$

If $\rho_- < 1$, we also arrange that

$$\|\Lambda_- \| < 1.$$

We equip $\mathcal{E}_- \oplus \mathcal{E}_+$ with the norm

$$\|(x, y)\| = \max\{\|x\|, \|y\|\}.$$

We then split off the linear part of f near the origin by writing:

$$(3.1) \quad f(x, y) = (\Lambda_- x + f_-(x, y), \Lambda_+ y + f_+(x, y)),$$

where the $f_{\pm}$ vanish, together with their first derivatives, at the origin.

We will refer to the $f_{\pm}$ as the nonlinear part of f . A first very important observation is that, because we want to prove a result which is local at the fixed point, we can without loss of generality assume that the $f_{\pm}$ are “small”. This goes as follows: By assumption, the $f_{\pm}$ are continuously differentiable and vanish, together with their first derivatives, at $(0, 0)$. Hence, the norms of their first derivatives can be made as small as we like on the ball of radius ϵ about $(0, 0)$ by making ϵ small. We can then replace f by $z \mapsto \frac{1}{\epsilon} f(\epsilon z)$ to get a new mapping with the same linear terms as f but whose non-linear terms have first derivatives which are uniformly as small as we like on the unit ball. That is: Working on the unit ball with the $f_{\pm}$ small in the C^1 sense is equivalent to working on a small ball for a general C^1 mapping f . We will refer to this reduction as “magnification”.

The next step is to determine when the graph of a mapping w from $\mathcal{E}_-$ to $\mathcal{E}_+$ is mapped to itself by f . To answer this question we write out, in our detailed notation (3.1) for f , the image under f of a general point of the graph of w :

$$f(x, w(x)) = (\Lambda_- x + f_-(x, w(x)), \Lambda_+ w(x) + f_+(x, w(x))).$$

This will again be in the graph of w if and only if

$$\Lambda_+ w(x) + f_+(x, w(x)) = w(\Lambda_- x + f_-(x, w(x))),$$

II.4.2, there is then a splitting of $\mathcal{E}$ as a direct sum of two closed linear subspaces $\mathcal{E}_s$ and $\mathcal{E}_u$, each invariant under $Df(x_0)$, such that the spectrum of the restriction of $Df(x_0)$ to $\mathcal{E}_s$ is entirely inside the unit circle and that of the restriction to $\mathcal{E}_u$ is entirely outside the unit circle. Our principal objective in this chapter is to prove the existence, uniqueness, and smoothness of an invariant manifold for f , passing through x_0 and tangent there to $\mathcal{E}_s$. This invariant manifold is known as the stable manifold for f at x_0 . To be slightly more precise: We will concentrate, for the first part of the analysis, on the behavior in the immediate vicinity of the fixed point, and will construct only a little piece of invariant manifold. In nice cases, it is possible to extend the small invariant manifold we construct here in a natural way to a global invariant manifold. Thus, what we are concerned with here is more properly called a local stable manifold. Invariant manifolds tangent to $\mathcal{E}_u$ are also important; they are called unstable manifolds. If f is invertible, an unstable manifold for f can be constructed as the stable manifold for f^{-1} . As we will see later on, however, unstable manifolds can also be constructed without the assumption that f is invertible.

It turns out that the argument we are going to give for the existence and smoothness of the stable manifold works, with almost no extra effort, to prove the existence—and to some extent the smoothness—of more general invariant manifolds, including what are known as center manifolds. It is worthwhile, therefore, to give this argument in a slightly more general context. For this purpose, we introduce the following notation: By a gap in the modulus of the spectrum of an operator A we mean an open subinterval (ρ_-, ρ_+) of the positive real axis such that the spectrum of A does not intersect the annulus $\{\lambda : \rho_- < |\lambda| < \rho_+\}$ but does intersect both of the circles $\{\lambda : |\lambda| = \rho_{\pm}\}$. With this terminology, the condition that A is hyperbolic can be expressed by saying that 1 belongs to a gap in the modulus of the spectrum of A . Given any gap in the spectrum of an operator A on a Banach space $\mathcal{E}$, there is a corresponding direct sum decomposition

$$\mathcal{E} = \mathcal{E}_- \oplus \mathcal{E}_+,$$

where $\mathcal{E}_-$ is the spectral subspace corresponding to the part of the spectrum of A inside the annulus $\{\lambda : \rho_- < |\lambda| < \rho_+\}$ and $\mathcal{E}_+$ is the spectral subspace corresponding to the part of the spectrum outside this annulus. We are going to study invariant manifolds for a continuously differentiable mapping f with fixed point x_0 which are tangent to a subspace $\mathcal{E}_-$ corresponding to a gap in the modulus of the spectrum of $Df(x_0)$.

We will call such manifolds generalized stable manifolds. A few particularly important cases are:

- (1) The gap in the modulus of the spectrum contains one, in which case the fixed point is hyperbolic and the invariant manifold is the ordinary stable manifold. The term stable manifold is also sometimes used in the case $\rho_+ = 1$, i.e., a stable manifold means an invariant manifold tangent to the spectral subspace corresponding to the part of the spectrum strictly inside the unit circle.
- (2) $\rho_+ < 1$ (which implies $\rho_- < 1$). In this case, we speak of a strong stable manifold. As we shall see, a strong stable manifold can be characterized dynamically as the

i.e., if and only if

$$(3.2) \quad w(x) = \Lambda_+^{-1} (w(\Lambda_- x + f_-(x, w(x))) - f_+(x, w(x))).$$

Now the right-hand side of (3.2) can be taken as defining a non-linear operator on function w , i.e., (3.2) can be interpreted as the fixed-point equation $w = \mathcal{F}(w)$ where

$$(3.3) \quad \mathcal{F}(w)(x) = \Lambda_+^{-1} (w(\Lambda_- x + f_-(x, w(x))) - f_+(x, w(x))).$$

The final result, then, is that the graph of w is mapped into itself by $\mathcal{F}$ if and only if w is a fixed point for $\mathcal{F}$ as defined in (3.3). Note that the fixed point condition includes the requirement that the right-hand side of (3.3) is defined on all of the domain of w , i.e., that $\Lambda_- x + f_-(x, w(x))$ is in the domain of w for all x in the domain of w .

The idea now is to choose a function space on which the operator $\mathcal{F}$ is a contraction. At first glance, it appears that it could be made to be contractive on the space of continuous functions with the supremum norm because of the overall contractive factor of Λ_+^{-1} . For a variety of reasons this turns out not to be a good approach; a better idea is to observe that the operator should contract the first derivative of w by a factor of about $\|\Lambda_+^{-1}\| \|\Lambda_-\|$. Showing that the operator is contractive in a supremum norm on first derivatives is not quite satisfactory either; it leads to complicated estimates and does not seem to give the strongest possible results. The space which turns out to work well is a space of Lipschitz continuous functions, equipped with a judicious modification of the supremum norm. The great advantage of working with Lipschitz functions is that a Lipschitz condition is almost as good as a bound on the first derivative for making estimates but Lipschitz continuity, unlike differentiability, is well-behaved under uniform—or even pointwise—limits.

We need to introduce a little formalism. If (X_1, ρ_1) , (X_2, ρ_2) are metric spaces, we say that a mapping $f : X_1 \rightarrow X_2$ is Lipschitz continuous if there exists a constant k such that

$$\rho_2(f(x), f(x')) \leq k \rho_1(x, x') \quad \text{for all } x, x' \in X_1.$$

The set of k 's for which this inequality holds is closed; its infimum will be denoted by $\text{Lip}(f)$. If the range space is a normed vector space, then the set of Lipschitz continuous functions is a vector space and $\text{Lip}(f)$ is almost a norm on this vector space. Although it fails to be a norm because it vanishes on constant functions, we will nevertheless refer to $\text{Lip}(f)$ as the Lipschitz norm of f .

The Lipschitz norm has very nice properties with respect to composition of functions:

PROPOSITION 3.2. *Let (X_1, ρ_1) , (X_2, ρ_2) , (X_3, ρ_3) be metric spaces, $f : X_1 \rightarrow X_2$ and $g : X_2 \rightarrow X_3$ Lipschitz continuous mappings. Then $g \circ f$ is Lipschitz continuous and*

$$\text{Lip}(g \circ f) \leq \text{Lip}(g) \text{Lip}(f).$$

PROOF: For x, x' in X_1 ,

$$\rho_3(g(f(x)), g(f(x'))) \leq \text{Lip}(g) \rho_2(f(x), f(x')) \leq \text{Lip}(g) \text{Lip}(f) \rho_1(x, x').$$

Once we have decided to look for the invariant set as the graph of a Lipschitz continuous function, it is no longer necessary to assume that f is continuously differentiable. Instead of requiring that $f_{\pm}$ be continuously differentiable on the unit ball with uniformly small first derivatives, it suffices to assume instead that they are Lipschitz continuous with small Lipschitz norm.

We have thus reworked the problem we want to solve into the following: We have two Banach spaces $\mathcal{E}_-$ and $\mathcal{E}_+$, equipped with operators Λ_- and Λ_+ respectively. We assume that Λ_+ is invertible and that

$$\|\Lambda_-\| \|\Lambda_+^{-1}\| < 1.$$

(This condition depend on what norms we are using on the Banach spaces $\mathcal{E}_-$ and $\mathcal{E}_+$. We can always choose the norms so that the condition holds provided that the spectral radius ρ_- of Λ_- is strictly smaller than ρ_+ , the infimum of $|\lambda|$ over the spectrum of Λ_+ .) It will be convenient to equip the direct sum space $\mathcal{E} = \mathcal{E}_- \oplus \mathcal{E}_+$ with the norm

$$\|(x, y)\| = \sup \{\|x\|, \|y\|\}.$$

We consider maps f defined on the unit ball of $\mathcal{E}$ of the form

$$(3.1) \quad f(x, y) = (\Lambda_-x + f_-(x, y), \Lambda_+y + f_+(x, y)),$$

where the $f_{\pm}$ are Lipschitz continuous and vanish at the origin. We want to show that, provided the Lipschitz norms of the $f_{\pm}$ are small enough, there is a unique mapping w from the unit ball of $\mathcal{E}_-$ into $\mathcal{E}_+$, with $w(0) = 0$ and $\text{Lip}(w) \leq 1$, which satisfies the functional equation (3.2) on the unit ball of $\mathcal{E}_-$.

To simplify the formulas arising in the course of the proof, it is convenient to introduce a streamlined notation. We write (3.2) and (3.3) as

$$(3.9.a) \quad w(x) = F(x, w(x), w(v(x, w(x)))) \equiv \mathcal{F}(w)(x),$$

where

$$(3.9.b) \quad F(x, y, y') = \Lambda_+^{-1} [y' - f_+(x, y)]$$

and

$$(3.9.c) \quad v(x, y) = \Lambda_-x + f_-(x, y).$$

Note that

$$v(0, 0) = 0 \quad \text{and} \quad F(0, 0, 0) = 0,$$

(and, if the $f_{\pm}$ are differentiable with first derivatives vanishing at the origin, then $D_1F(0, 0, 0) = 0$.) When we consider functions of several variables, like F and v , we will write Lip_i for the Lipschitz norm with respect to the i -th variable with the other

variables held fixed (i.e., in the differentiable case, the supremum norm of the i -th partial derivative.) From the definitions, we have

$$\text{Lip}_1(F) \leq \|\Lambda_+^{-1}\| \text{Lip}_1(f_-),$$

$$\text{Lip}_2(F) \leq \|\Lambda_+^{-1}\| \text{Lip}_1(f_+),$$

$$\text{Lip}_3(F) = \|\Lambda_+^{-1}\|,$$

$$\text{Lip}_1(v) \leq \|\Lambda_-\| + \text{Lip}_1(f_-),$$

and

$$\text{Lip}_2(v) \leq \text{Lip}_2(f_-).$$

Thus, all these Lipschitz norms except $\text{Lip}_3(F)$ and $\text{Lip}_1(v)$ can be made as small as we like by making $f_{\pm}$ small in Lipschitz norm, and the remaining two can be made almost as small as $\|\Lambda_+^{-1}\|$ and $\|\Lambda_-\|$ respectively. Finally, we will write $\mathcal{X}$ and $\mathcal{Y}$ for the two Banach spaces rather than $\mathcal{E}_-$ and $\mathcal{E}_+$.

§4. Generalized stable manifolds: the functional equation.

To construct a solution to the invariant manifold problem as set up in the end of the preceding section, and to prove its uniqueness, we will construct a complete metric space of candidates for w on which $\mathcal{F}$ is defined and contractive. The space we use will be the set $\mathcal{X}_{\mathcal{F}}$ of Lipschitz continuous mappings w from the unit ball in $\mathcal{X}$ to $\mathcal{Y}$ with $w(0) = 0$ and $\text{Lip}(w) \leq 1$. We equip $\mathcal{X}_{\mathcal{F}}$ with the metric

$$\|w_1 - w_2\| = \sup_{x \neq 0} \frac{\|w_1(x) - w_2(x)\|}{\|x\|}$$

Since the w_i , $i = 1, 2$, vanish at zero and have Lipschitz norm majorized by one, the supremum is no larger than two. Note that we are *not* using the Lipschitz norm as a metric. Instead, we are using essentially a weighted supremum norm, which is much weaker, but which nevertheless makes $\mathcal{X}_{\mathcal{F}}$ into a complete metric space.

Rather than stating a formal result at this point, we are going to proceed to develop an argument, deriving along the way a set of conditions on F and v needed to make the argument work; the conclusions will be collected at a proposition formulated at the end.

The first point we have to deal with is that of making sure that, for all w in $\mathcal{X}_{\mathcal{F}}$, the right-hand side of (3.9.a) is defined on the whole unit ball of $\mathcal{X}$, i.e., that for any $w \in \mathcal{X}_{\mathcal{F}}$, the mapping $x \mapsto v(x, w(x))$ sends the unit ball into itself. Since it maps 0 to 0, it suffices to show its Lipschitz norm is not greater than one. But the Lipschitz norm of this mapping is certainly not greater than

$$\text{Lip}_1(v) + \text{Lip}_2(v)\text{Lip}(w) \leq \text{Lip}_1(v) + \text{Lip}_2(v).$$

We will for the remainder of this section use the (not strictly consistent) notation of writing $\text{Lip}(v)$ for $\text{Lip}_1(v) + \text{Lip}_2(v)$, so our first condition on v and F can be written as

$$(4.1) \quad \text{Lip}(v) \leq 1.$$

Looking back at the original problem, we see that it will be possible to satisfy this condition if $\|\Lambda_-\| < 1$, i.e., if $\rho_- < 1$. We will introduce shortly a device to get around this limitation, but for the moment we impose (4.1).

Next, to show that $\mathcal{F}$ maps $\mathcal{X}_{\mathcal{F}}$ to itself, we have to verify that $\mathcal{F}w(0) = 0$ (which is immediate from $w(0) = 0$, $v(0, 0) = 0$, and $F(0, 0, 0) = 0$) and that $\text{Lip}(\mathcal{F}w) \leq 1$. But

$$\begin{aligned} \text{Lip}(\mathcal{F}w) &\leq \text{Lip}_1(F) + \text{Lip}_2(F)\text{Lip}(w) + \text{Lip}_3(F)\text{Lip}(w) [\text{Lip}_1v + \text{Lip}_2v\text{Lip}(w)] \\ &\leq \text{Lip}_1(F) + \text{Lip}_2(F) + \text{Lip}_3(F) [\text{Lip}_1v + \text{Lip}_2v] \end{aligned}$$

where we have used $\text{Lip}(w) \leq 1$. Thus, to guarantee that $\mathcal{F}$ maps $\mathcal{X}_{\mathcal{F}}$ into itself, we require

$$(4.2) \quad \text{Lip}_1(F) + \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) \leq 1$$

To find conditions under which $\mathcal{F}$ is contractive on $\mathcal{X}_{\mathcal{F}}$, we take an arbitrary pair of element w_1 and w_2 in $\mathcal{X}_{\mathcal{F}}$, and an arbitrary $x \neq 0$, and estimate

$$\begin{aligned} \mathcal{F}w_1(x) - \mathcal{F}w_2(x) &= F(x, w_1, w_1(v_1)) - F(x, w_2, w_1(v_1)) \\ &\quad + F(x, w_2, w_1(v_1)) - F(x, w_2, w_2(v_2)) \\ &\quad + F(x, w_2, w_2(v_1)) - F(x, w_2, w_2(v_2)) \end{aligned}$$

where we have suppressed many arguments, using the convention that w_i without argument means $w_i(x)$ and that v_i means $v(x, w_i(x))$. We then use the estimates:

$$\|v_1 - v_2\| \leq \text{Lip}_2(v)\|w_1 - w_2\|\|x\|,$$

$$\|w_2(v_1) - w_2(v_2)\| \leq \|w_1 - w_2\|\text{Lip}(v)\|x\|,$$

to get

$$\|\mathcal{F}w_1(x) - \mathcal{F}w_2(x)\| \leq [\text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) + \text{Lip}_3(F)\text{Lip}_2(v)]\|w_1 - w_2\|\|x\|.$$

Dividing by $\|x\|$ and taking the supremum over x , we see that $\mathcal{F}$ is a contraction on $\mathcal{X}_{\mathcal{F}}$ provided

$$(4.3) \quad \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) + \text{Lip}_3(F)\text{Lip}_2(v) < 1.$$

Thus: If (4.1), (4.2) and (4.3) all hold, then (3.9) has a unique solution, as desired. We will, from now on, denote this unique fixed point simply by w .

We next want to introduce a new idea which makes a version of the above argument work without the assumption $\text{Lip}(v) \leq 1$. The idea is to cut off the non-linear terms in f away from the origin to get a mapping which is *globally* close, in a Lipschitz sense, to a linear mapping and to apply the analysis developed above to prove the existence of a global invariant Lipschitz manifold W for the cut-off mapping. The cutting-off is done in such a way as to make the cut-off mapping agree exactly with f on a neighborhood of the fixed point; then the manifold W which is globally invariant for the cut-off mapping is *locally invariant* for f in the sense that, for $\mathcal{U}$ a small enough neighborhood of the fixed point, if $x \in \mathcal{U} \cap W$, and $f(x) \in \mathcal{U}$, then $f(x) \in W$.

More explicitly: Consider a mapping f of the standard form

$$f(x, y) = (\Lambda_-x + f_-(x, y), \Lambda_+y + f_+(x, y)).$$

with $f_{\pm}$ small in Lipschitz norm. Choose a Lipschitz-continuous real-valued function ψ on $\mathcal{X}$ which is identically equal to one on a neighborhood of 0 and identically zero for $\|x\| \geq 1 - \epsilon$ for some strictly positive ϵ . Such a function can easily be constructed by composing the norm with a smooth function on the positive real axis which is one on a neighborhood of 0 and 0 to the right of $1 - \epsilon$. (Note, however, that the remarks in §III.1, it may not be possible to take ψ to be differentiable if $\mathcal{X}$ is infinite dimensional.) Define

$$\begin{aligned}\widetilde{f}_-(x, y) &= \psi(x)f_-(x, y) \\ \widetilde{f}_+(x, y) &= \psi(x)f_+(x, y) \\ \widetilde{f}(x, y) &= (\Lambda_-x + \widetilde{f}_-(x, y), \Lambda_+y + \widetilde{f}_+(x, y))\end{aligned}$$

Although f might not have been defined on all of $\mathcal{E}$, we can evidently extend $\widetilde{f}_-$, $\widetilde{f}_+$ to vanish when $\|x\| \geq 1$ and extend $\widetilde{f}$ correspondingly. By making $\text{Lip}(f_-)$ and $\text{Lip}(f_+)$ small on the unit ball, we can guarantee that the Lipschitz norms of $\widetilde{f}_-$ and $\widetilde{f}_+$ taken over the set of all (x, y) with $\|y\| \leq \|x\|$ with no restriction at all on x are as small as we like.* Dropping the tildes, we are led to consider a *global* version of the invariant manifold problem, i.e., to consider a mapping f which is defined and uniformly near to linear in the sense of Lipschitz norm on—essentially—all of $\mathcal{X} \oplus \mathcal{Y}$ and to look for a function w defined and satisfying the condition $\text{Lip}(w) \leq 1$ on all of $\mathcal{X}$ whose graph is mapped into itself by f . Reexamination of the argument in the first few paragraphs of this section shows that it works perfectly in this situation, without requiring $\text{Lip}(v) \leq 1$, provided that all the Lipschitz norms are understood to be taken over the appropriate domains, e.g., the $\text{Lip}_i(F)$ is to be understood as taken over the domain

$$\{(x, y, y') : \|y\| \leq \|x\| \quad \text{and} \quad \|y'\| \leq \|x\| \quad \text{with no restriction on } x.\}$$

*We could have taken a cutoff function ψ which cuts off in both x and y , in which case the condition $\|y\| \leq \|x\|$ would be unnecessary. Cutting off only in x has the advantage that it may be easier to find smooth cutoff functions on $\mathcal{E}_-$ than on $\mathcal{E}$ (particularly if $\mathcal{E}_-$ is finite-dimensional and $\mathcal{E}_+$ infinite-dimensional.)

From now on, we will carry simultaneously two versions of the argument, a local version where $F(x, y, y')$ and $v(x, y)$ are defined for x, y, y' in the unit balls of their respective spaces and Lipschitz norms taken over these unit balls, and a global version where x runs over all of $\mathcal{X}$ and y, y' are constrained only to have norms not greater than that of x . In the local case, we always assume in addition that

$$\text{Lip}(v) \leq 1;$$

this guarantees that there are no domain problems involved in making sense of the functional equation.

Although the main arguments are exactly the same in the local and global cases, the results obtained in the global case, i.e., in the invariant manifold problem from which we started, for $\rho_- \geq 1$, turn out to be considerably less satisfactory than those for $\rho_- < 1$. First of all, in the infinite-dimensional situation, there is the problem of whether the cut-off function ψ can be taken to be smooth, or, more generally, whether we can find a way of cutting off f to produce a smooth $\tilde{f}$ with globally small non-linear terms. Unless this can be done, there would seem to be no hope of showing that w is smooth. We will see shortly that, if we add to our hypotheses the assumption that $f_{\pm}$ are continuously differentiable with derivatives vanishing at zero, then the function w is also continuously differentiable and has derivative which vanishes at zero, i.e., its graph is a C^1 submanifold tangent to $\mathcal{X}$ at the zero. The next difficulty arises if one wants to prove more smoothness for w . The assumption that $\tilde{f}$ is of class C^n —which can certainly be arranged if f itself is of class C^n and if there exists a cut-off function ψ of class C^n —does not suffice to guarantee that w is of class C^n ; it is also necessary to assume that $\|\Lambda_-\|^n \|\Lambda_+^{-1}\| < 1$. If the spectral radius of Λ_- is greater than 1, this condition can only be satisfied for n not too large. Finally, there are problems with the uniqueness of the invariant manifold. The Contraction Mapping Theorem assures us that the global function w is uniquely determined by the condition of global invariance of its graph under the global cut-off mapping. This does not imply—and we will later see from examples that it is not true—that the restriction of w to a small neighborhood of the fixed point is independent of the cut-off. That is: For $\rho_- \geq 1$ there may be many locally invariant manifolds tangent to $\mathcal{E}_-$. This is a reflection of the fact that, for $\rho_- \geq 1$, local invariance is not really a local property: The invariant manifold can contain orbits which start out arbitrarily near to the fixed point but which get well away from it. In spite of these drawbacks, the center-stable manifold (the invariant manifold given by this argument when $\rho_- = 1$) is still quite useful.

Next, we will assume that F and v are continuously differentiable, with $D_1 F(0, 0, 0) = 0$ and seek conditions which guarantee that the fixed point w of $\mathcal{F}$ is also continuously differentiable. In outline, the argument is going to go as follows: If we knew that w was differentiable, we could differentiate the functional equation w to get the functional equation

$$Dw(x) = D_1 F + D_2 F Dw + D_3 F Dw(v)[D_1 v + D_2 v Dw]$$

for its derivative. Here again, we have suppressed arguments, and, from here on, we will normally leave them out when it is obvious from the context what they should be. We

first argue that this functional equation, with w fixed, uniquely determines Dw . That is: We define an operator $\mathcal{K}$ acting on a space of mappings from $\mathcal{X}$ to $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ by

$$(4.4) \quad \mathcal{K}\sigma(x) = D_1F + D_2F\sigma + D_3F\sigma(v)[D_1v + D_2v\sigma],$$

and show that this mapping has a unique fixed point in an appropriate function space. Again, we denote this unique fixed point simply by σ . Thus, although we do not yet know that w is differentiable, we know that if it is differentiable its derivative must be σ . We then use the functional equations for w and σ to verify that σ has the defining property of the derivative of w —i.e., that $w(x + \delta x) - w(x) - \sigma\delta x$ is $o(\delta x)$ —and hence that w is indeed differentiable.

We will work in the space $\mathcal{X}_{\mathcal{K}}$ of continuous mappings σ from $\mathcal{X}$ to $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ with $\sigma(0) = 0$ and $\|\sigma(x)\| \leq 1$ for all x . (This last bound is chosen to correspond to the condition $\text{Lip}(w) \leq 1$.) We equip $\mathcal{X}_{\mathcal{K}}$ simply with the supremum norm. From the definition (4.4) of $\mathcal{K}$, and from $v(0, 0) = 0$, $w(0) = 0$, and $D_1F(0, 0, 0) = 0$, it follows that $\mathcal{K}\sigma(0) = 0$ if $\sigma(0) = 0$. Also, simply by taking the norm of the defining equations, and using estimates like $\|D_1F\| \leq \text{Lip}_1(F)$, we see that $\|\mathcal{K}\sigma(x)\| \leq 1$ for all x provided $\|\sigma(x)\| \leq 1$ for all x and provided

$$\text{Lip}_1(F) + \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) \leq 1.$$

This latter condition is exactly (4.2) (and this is not an accident; (4.2) was what was necessary to guarantee that $\mathcal{F}$ preserves the condition $\text{Lip}(w) \leq 1$.)

To find conditions which make $\mathcal{K}$ a contraction, we use the definition to write

$$\begin{aligned} \mathcal{K}\sigma_1(x) - \mathcal{K}\sigma_2(x) = & D_2F(\sigma_1 - \sigma_2) + D_3F(\sigma_1(v) - \sigma_2(v))[D_1v + D_2\sigma_1] \\ & + D_3F\sigma_2D_2v(\sigma_1 - \sigma_2). \end{aligned}$$

Taking norms and estimating in a completely straightforward way shows that $\mathcal{K}$ is a contraction provided

$$\text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) + \text{Lip}_3(F)\text{Lip}_2(v) < 1,$$

and this is exactly (4.3), i.e., the condition needed to make $\mathcal{F}$ contractive.

We thus now let σ denote the unique fixed point for $\mathcal{K}$ in $\mathcal{X}_{\mathcal{K}}$. We want to show that $\sigma(x)$ is the derivative of w at x for each point x . For this purpose, define

$$M \equiv \sup_x \limsup_{\delta x \rightarrow 0} \frac{\|w(x + \delta x) - w(x) - \sigma(x)\delta x\|}{\|\delta x\|}.$$

The idea will be to use the functional equations to show that $M \leq \kappa M$ for some $\kappa < 1$. Since M is finite, this implies $M = 0$, which is exactly the statement that $\sigma(x)$ is the derivative of w at x for all x . (Finiteness of M follows at once from the finiteness of $\text{Lip}(w)$ and uniform boundedness of $\sigma(x)$.)

To keep the formulas of reasonable size, we introduce the abbreviations:

$$\delta w = w(x + \delta x) - w(x)$$

$$\delta v = v(x + \delta x, w + \delta w) - v(x, w)$$

From the functional equation for w , and the assumed differentiability of F ,

$$\delta w = D_1 F \delta x + D_2 F \delta w + D_3 F [w(v + \delta v) - w(v)] + o(\delta x).$$

We then multiply δx by the expression given by the functional equation for $\sigma(x)$ and subtract from the above equation; this gives

$$\begin{aligned} \delta w - \sigma \delta x &= D_2 F [\delta w - \sigma \delta x] + D_3 F [w(v + \delta v) - w(v) - \sigma(v) \delta v] \\ &\quad + D_3 F \sigma(v) [\delta v - D_1 v \delta x - D_2 v \delta w] \\ (4.5) \quad &\quad + D_3 F \sigma(v) D_2 v [\delta w - \sigma \delta x] + o(\delta x). \end{aligned}$$

By the differentiability of v ,

$$\delta v - D_1 v \delta x - D_2 v \delta w = o(\delta x).$$

Also,

$$\|\delta v\| \leq \text{Lip}(v) \|\delta x\|,$$

and hence

$$\limsup_{\delta x \rightarrow 0} \frac{\|w(v + \delta v) - w(v) - \sigma(v) \delta v\|}{\|\delta x\|} \leq \text{Lip}(v) M.$$

Taking norms of both sides of (4.5), dividing by the norm of δx , letting δx tend to 0, and inserting the preceding two estimates gives

$$\limsup_{\delta x \rightarrow 0} \frac{\|\delta w - \sigma(x) \delta x\|}{\|\delta x\|} \leq [\text{Lip}_2(F) + \text{Lip}_3(F) \text{Lip}(v) + \text{Lip}_3(F) \text{Lip}_2 v] M,$$

so w is differentiable, with derivative σ , provided

$$\text{Lip}_2(F) + \text{Lip}_3(F) \text{Lip}(v) + \text{Lip}_3(F) \text{Lip}_2 v < 1,$$

and this is precisely (4.3).

Next, we assume that the first derivatives of F and v are α -Hölder continuous, with $0 < \alpha \leq 1$, and try to prove that σ is also α -Hölder continuous. It will be convenient to define a "Hölder norm" generalizing the Lipschitz norm already introduced: If g is an α -Hölder continuous function from (X_1, ρ_1) to (X_2, ρ_2) , then $\text{Hol}_\alpha(g)$ will denote the smallest constant k such that

$$\rho_2(g(x), g(x')) \leq k \rho_1(x, x')^\alpha \quad \text{for all } x, x' \text{ in } X_1.$$

The idea will be to try to find a bound on $\text{Hol}_\alpha(\sigma)$ which is propagated by $\mathcal{K}$. We can then restrict $\mathcal{X}_\mathcal{K}$ to the subspace of σ 's satisfying this bound; the restricted $\mathcal{X}_\mathcal{K}$ is still complete in the original metric, so $\mathcal{K}$ has a fixed point in it, and this fixed point must be Dw by uniqueness of the fixed point in the unrestricted $\mathcal{X}_\mathcal{K}$.

Thus, let $\sigma \in \mathcal{X}_\mathcal{K}$ be α -Hölder continuous. For any $x_1 \neq x_2$, we have

$$\begin{aligned}
 \mathcal{K}\sigma(x_1) - \mathcal{K}\sigma(x_2) &= D_1F_1 - D_1F_2 + (D_2F_1 - D_2F_2)\sigma_1 \\
 &\quad + D_2F_2(\sigma_1 - \sigma_2) + (D_3F_1 - D_3F_2)\sigma(v_1)(D_1v_1 + D_2v_1\sigma_1) \\
 (4.6) \quad &\quad + D_3F_2(\sigma(v_1) - \sigma(v_2))(D_1v_1 + D_2v_1\sigma_1) \\
 &\quad + D_3F_2\sigma(v_2)[(D_1v_1 - D_1v_2) + (D_2v_1 - D_2v_2)\sigma_1 + D_2v_2(\sigma_1 - \sigma_2)]
 \end{aligned}$$

Here we have written subscripts 1 and 2 to indicate objects to be evaluated at $x = x_1$ and $x = x_2$ respectively. This notation is superimposed on our usual suppression of arguments. Thus, for example, D_1F_1 means $D_1F(x_1, w(x_1), w(v(x_1, w(x_1))))$.

In equation (4.6), each of the terms $D_1F_1 - D_2F_2$, $D_2F_1 - D_2F_2$, $D_3F_1 - D_3F_2$, $D_1v_1 - D_1v_2$, and $D_2v_1 - D_2v_2$ is bounded by a constant multiple of $\|x_1 - x_2\|^\alpha$, with a constant which doesn't depend on the x 's nor on σ . Also,

$$\|\sigma(v_1) - \sigma(v_2)\| \leq \text{Hol}_\alpha(\sigma)\|v_1 - v_2\|^\alpha \leq \text{Hol}_\alpha(\sigma)\text{Lip}(v)^\alpha\|x_1 - x_2\|^\alpha.$$

Taking norms of both sides of (4.6) and using these remarks, we get a bound of the form

$$\begin{aligned}
 \|\mathcal{K}\sigma(x_1) - \mathcal{K}\sigma(x_2)\| &\leq A\|x_1 - x_2\|^\alpha \\
 &\quad + [\text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v)^{1+\alpha} + \text{Lip}_3(F)\text{Lip}_2(v)] \text{Hol}_\alpha(\sigma)\|x_1 - x_2\|^\alpha,
 \end{aligned}$$

with A independent of σ , x_1 , and x_2 . Thus, if we require

$$(4.7) \quad \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v)^{1+\alpha} + \text{Lip}_3(F)\text{Lip}_2(v) < 1,$$

and if we write κ for the left-hand side of (4.7) and B for $A/(1 - \kappa)$, we get

$$\text{Hol}_\alpha(\sigma) \leq B \quad \text{implies} \quad \text{Hol}_\alpha(\mathcal{K}\sigma) \leq B$$

and so

$$\text{Hol}_\alpha(Dw) \leq B.$$

We next take up the question of higher differentiability. To start with, we want to assume that F is of class C^2 and show that w is also. We can write the functional equation satisfied by Dw in approximately the same form as the functional equation satisfied by w :

$$(4.8a) \quad Dw(x) = \bar{F}(x, Dw(x), Dw(\bar{v}(x)))$$

where

$$(4.8b) \quad \bar{F}(x, \sigma, \sigma') = D_1 F + D_2 F \sigma + D_3 F \sigma' [D_1 v + D_2 v \sigma]$$

and

$$(4.8c) \quad \bar{v}(x) = v(x, w(x)).$$

We already know that, provided the appropriate bounds hold, the solution of (4.8) exists and is unique in the space of continuous functions with norm bounded by one, and that it is Lipschitz continuous if the first derivative of F is. We would like to use the argument given above to deduce the differentiability of w from that of F and v to deduce the differentiability of Dw from that of $\bar{F}$ and $\bar{v}$, which in turn follows if we know that F is of class C^2 .

The functional equations (3.9) and (4.8) are, however, sufficiently different so that it is necessary to repeat the argument with appropriate modifications; the result we want does not follow formally from the argument already given. On the one hand, equation (4.8) is considerably simpler to treat than (3.9) because $\bar{v}$ depends only on x , and not on $\sigma(x)$. (We are now supposed to regard w as "known" and fixed.) This simplification already permitted us to treat the existence and uniqueness problem for (4.8) directly in a space of continuous functions equipped with the supremum norm, without the refinements needed to treat these questions for (3.9). On the other hand, the specific bound $\text{Lip}(\sigma) \leq 1$, analogous to the bound $\text{Lip}(w) \leq 1$ for (3.9) may *not* be satisfied.

Differentiating (4.8) formally shows that the derivative of $D^2 w$, if it exists, must satisfy

$$D^2 w(x) = D_1 \bar{F} + D_2 \bar{F} D^2 w + D_3 \bar{F} D^2 w(\bar{v}) D \bar{v}.$$

Accordingly, define an operator $\mathcal{F}^{(2)}$ by

$$(4.10) \quad \mathcal{F}^{(2)} \sigma^{(2)}(x) = D_1 \bar{F} + D_2 \bar{F} \sigma^{(2)} + D_3 \bar{F} \sigma^{(2)}(\bar{v}) D \bar{v}.$$

Note that $\mathcal{F}^{(2)}$ is *linear* (or, more precisely, affine.) We are going to let this operator act on the space of all bounded continuous functions equipped with the supremum norm. It is nearly immediate that

$$(4.11) \quad \text{Lip}_2(\bar{F}) + \text{Lip}_3(\bar{F}) \text{Lip}(\bar{v}) < 1,$$

implies that $\mathcal{F}^{(2)}$ is a contraction in the supremum norm and hence that it has a unique bounded continuous fixed point, which we will now denote simply by $w^{(2)}$. (A further, very minor, difference between this analysis and that given for first derivatives is that $w^{(2)}(0)$ need not vanish. We simply do not put the condition $\sigma^{(2)}(0) = 0$ into the definition of the space, and so do not need to verify that the operator $\mathcal{F}^{(2)}$ preserves it.)

It is useful to re-express the condition (4.11) in terms of the original functions F and v . From the definitions (4.8bc) it follows almost immediately that

$$\text{Lip}_2(\bar{F}) \leq \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}_2(v)$$

$$\text{Lip}_3(\bar{F}) \leq \text{Lip}_3(F)\text{Lip}(v)$$

$$\text{Lip}(\bar{v}) \leq \text{Lip}(v).$$

Hence (4.11) follows from

$$(4.12) \quad \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v)^2 + \text{Lip}_3(F)\text{Lip}_2(v) < 1,$$

and this is the same as condition (4.7) with $\alpha = 1$.

If we now look back at the analysis showing that the unique fixed point of $\mathcal{K}$ is indeed the derivative of the unique solution of (3.9), we note that *the only terms for which we had to use the estimate $\|\sigma(x)\| \leq 1$ were multiplied by D_2v and hence do not appear at all in the present analysis.* That proof therefore applies to the present case as well and shows that Dw is continuously differentiable provided that (4.11) holds, and hence provided that (4.12) holds. Similarly, the argument showing that Dw is α -Hölder continuous applies to show that, if D^2F is α -Hölder continuous, and if

$$\text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v)^{2+\alpha} + \text{Lip}_3(F)\text{Lip}_2(v) < 1,$$

then D^2w is α -Hölder continuous.

If we then proceed to analyze the *third* derivative of w , assuming that F is of class C^3 , we find that the functional equation satisfied by D^2w has all the properties that we needed to assume about the equation for Dw in order to make the preceding analysis. Hence, we use induction to deal immediately with arbitrary degrees of differentiability. We can sum up the results obtained in the following proposition:

PROPOSITION 4.1. *Let F, v be as at the end of the preceding section. Assume further that F and v are of class C^r , with $1 \leq r \leq \infty$, and that*

$$\text{Lip}_1(F) + \text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) \leq 1$$

$$\text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v) + \text{Lip}_3(F)\text{Lip}_2(v) < 1$$

$$\text{Lip}_2(F) + \text{Lip}_3(F)\text{Lip}(v)^r + \text{Lip}_3(F)\text{Lip}_2(v) < 1$$

(this last condition should be replaced by $\text{Lip}(v) \leq 1$ for $r = \infty$.) Then (3.9) has a unique solution w with $w(0) = 0$ and $\text{Lip}(w) \leq 1$; this solution is of class C^r .

A discussion of other approaches to proving smoothness should go here.

§5. **Expanding and contracting wedges.** We consider a mapping f of the form described at the end of §3, and we also suppose that

$$\|\Lambda_-\| + \text{Lip}(f_-) < 1,$$

and that there is a mapping w with $w(0) = 0$ and $\text{Lip}(w) \leq 1$ whose graph is mapped into itself by f . If $z = (x, w(x))$ is in this invariant graph, then the x component of $f(z)$ is given by $\Lambda_-x + f_-(x, w(x))$, and the Lipschitz norm of the mapping

$$x \mapsto \Lambda_-x + f_-(x, w(x))$$

is bounded by $\|\Lambda_-\| + \text{Lip}(f_-)$. Furthermore, because of the norm we have chosen on $\mathcal{X} \oplus \mathcal{Y}$, $\|(x, w(x))\| = \|x\|$ for any x . Putting these remarks together, and writing r_- for $\|\Lambda_-\| + \text{Lip}(f_-)$, we see that

$$\|f(z)\| \leq r_- \|z\| \quad \text{for all } z \text{ in the graph of } w.$$

Since the graph is invariant, we can immediately iterate this remark to show that, for any z in the graph of w , $f^n(z)$ goes to zero at least as fast as r_-^n .

We want to prove here a complementary result showing that, if z is *not* in the invariant manifold, then $\|f^n(z)\|$ cannot go to zero too fast. To prove this, we introduce a new idea which has many applications in dynamical systems theory. To illustrate the idea, we consider a particularly simple case: a continuously differentiable mapping f with a hyperbolic fixed point. The key observation in this situation is that, if z_1 and z_2 are two points near the fixed point whose separation lies approximately in the expanding direction, then

- (a) $\|f(z_1) - f(z_2)\| \geq k\|z_1 - z_2\|$ for some constant $k > 1$ independent of z_1, z_2 .
- (b) the separation between $f(z_1)$ and $f(z_2)$ still lies approximately in the expanding direction.

Because of (b), this observation can be applied repeatedly to show that the orbits of z_1 and z_2 diverge exponentially as long as they both remain near the hyperbolic fixed point, and hence that they cannot both remain near the fixed point forever.

A precise version of this argument, adequate for our present purposes, is as follows: We let, as usual, Λ_- , Λ_+ be linear operators on Banach spaces $\mathcal{E}_-$, $\mathcal{E}_+$ respectively, with

$$\|\Lambda_-\| < \|\Lambda_+^{-1}\|^{-1},$$

and let

$$f(x, y) = (\Lambda_-x + f_-(x, y), \Lambda_+y + f_+(x, y))$$

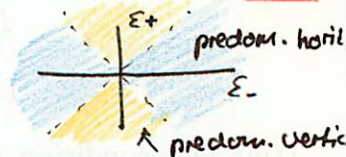
with $f_{\pm}(0, 0) = 0$. We use the norm $\|(x, y)\| = \max(\|x\|, \|y\|)$ on the direct sum space. If $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ are two points of $\mathcal{E}_- \oplus \mathcal{E}_+$, we will say that the separation between z_1 and z_2 is predominantly vertical if

$$\|x_2 - x_1\| < \|y_2 - y_1\|,$$

and predominantly horizontal if

$$\|y_2 - y_1\| < \|x_2 - x_1\|.$$

Also, write r_+ for $\|\Lambda_+^{-1}\|^{-1} - \text{Lip}(f_+)$ and r_- for $\|\Lambda_-\| + \text{Lip}(f_-)$, where the Lipschitz norms are taken over the unit ball.



PROPOSITION 5.1. *Assume that the $\text{Lip}(f_{\pm})$ are small enough so that*

$$r_- < r_+.$$

and let z_1, z_2 be two points of the unit ball. Then

- (a) *If the separation between z_1 and z_2 is predominantly vertical, then so is the separation between $f(z_1)$ and $f(z_2)$, and*

$$\|f(z_1) - f(z_2)\| \geq r_+ \|z_1 - z_2\|.$$

- (b) *If the separation between $f(z_1)$ and $f(z_2)$ is predominantly horizontal, then so is the separation between z_1 and z_2 , and*

$$\|f(z_1) - f(z_2)\| \leq r_- \|z_1 - z_2\|.$$

PROOF: Write (x_i, y_i) for z_i and $(\bar{x}_i, \bar{y}_i)$ for $f(z_i)$, $i = 1, 2$. We assume first that the separation between z_1 and z_2 is predominantly vertical, i.e., that $\|y_1 - y_2\| > \|x_1 - x_2\|$. Note that in this case, because of our choice of norm for the direct sum space, $\|z_1 - z_2\| = \|y_1 - y_2\|$. Thus

$$\begin{aligned} \|\bar{y}_2 - \bar{y}_1\| &\geq \|\Lambda_+(y_2 - y_1)\| - \text{Lip}(f_+) \|z_2 - z_1\| \\ &\geq (\|\Lambda_+^{-1}\|^{-1} - \text{Lip}(f_+)) \|z_1 - z_2\| = r_+ \|z_1 - z_2\|. \end{aligned}$$

Similarly,

$$\begin{aligned} \|\bar{x}_2 - \bar{x}_1\| &< (\|\Lambda_-\| + \text{Lip}(f_-)) \|z_1 - z_2\| \\ &= r_- \|z_1 - z_2\| \leq \|\bar{y}_2 - \bar{y}_1\|, \end{aligned}$$

(where, to get the strict inequality on the first line, we used the fact that $\|x_1 - x_2\|$ is strictly less than $\|z_1 - z_2\|$.) These two inequalities together prove (a).

It then follows from (a) and continuity that

$$\|y_1 - y_2\| \geq \|x_1 - x_2\| \quad \text{implies} \quad \|\bar{y}_1 - \bar{y}_2\| \geq \|\bar{x}_1 - \bar{x}_2\|.$$

Hence, if the separation between $f(z_1)$ and $f(z_2)$ is predominantly horizontal, the separation between z_1 and z_2 must also be predominantly horizontal. Once we know this, it is nearly immediate that

$$\|f(z_1) - f(z_2)\| = \|\bar{x}_1 - \bar{x}_2\| \leq [\|\Lambda_- + \text{Lip}(f_-)\|] \|z_1 - z_2\| = r_- \|z_1 - z_2\|,$$

as desired. ■

COROLLARY 5.2. *Let the hypotheses and notation be as in Proposition 5.1 but assume in addition*

- (a) $\|\Lambda_-\| + \text{Lip}(f_-) < 1$, (with Lipschitz norm taken over the unit ball)
- (b) *There is a function w from the unit ball in $\mathcal{E}_-$ to $\mathcal{E}_+$, with $w(0) = 0$ and $\text{Lip}(w) \leq 1$, whose graph is mapped into itself by f .*
- (c) $f_{\pm}$ are continuously differentiable with derivatives vanishing at $(0,0)$.

Then, for any z in the graph of w ,

$$\limsup_n \frac{1}{n} \log(\|f^n(z)\|) \leq \log \rho_-,$$

while, if z is in the unit ball but not in the graph of w , either some $f^n(z)$ is not in the unit ball or

$$\liminf_n \frac{1}{n} \log(\|f^n(z)\|) \geq \log r_+.$$

If z is not in the graph of w , but if the orbit of z nevertheless remains in the unit ball and converges to zero, then

$$\liminf_n \frac{1}{n} \log(\|f^n(z)\|) \geq \log \rho_+.$$

If $r_+ > 1$, and if z is in the unit ball but not in the graph of w , then $f^n(z)$ is outside the unit ball for some n .

PROOF: We have already seen, in effect, that if z is in the graph of w ,

$$\limsup_n \frac{1}{n} \log(\|f^n(z)\|) \leq \log(r_-).$$

Thus, for an arbitrarily small ball $\mathcal{U}$ centered at the origin, $f^n(z)$ is inside $\mathcal{U}$ for all n greater than some n_0 , and the term $\text{Lip}(f_-)$ in the definition of r_- can be replaced by the Lipschitz norm of f_- over $\mathcal{U}$. But because f_- is continuously differentiable with derivative vanishing at zero, this Lipschitz norm can be made as small as we like by making $\mathcal{U}$ small, so we get

$$\limsup_n \frac{1}{n} \log(\|f^n(z)\|) \leq \log(\|\Lambda_-\|).$$

Since the left-hand side is unchanged if the norm is replaced by an equivalent one, we can minimize the right-hand side over norms and thus get

$$\limsup_n \frac{1}{n} \log(\|f^n(z)\|) \leq \log(\rho_-),$$

as desired.

Now suppose that $z = (x, y)$ is not in the graph of w , and let $z' = (x, w(x))$. Then the separation between z and z' is predominantly vertical, so we can apply Proposition 5.1 to conclude that, as long as $f^n(z)$ and $f^n(z')$ remain in the unit ball,

$$\|f^n(z) - f^n(z')\| \geq r_+^n \|z - z'\|,$$

completing the proof.

If z is not in the graph of w but $f^n(z)$ nevertheless converges to zero—which, by the preceding estimate, is impossible unless $r_+ < 1$ —the argument like that given for the asymptotic rate of convergence of orbits in the invariant manifold shows that

$$\liminf_n \frac{1}{n} \log(\|f^n(z)\|) \geq \log \rho_+.$$

Note also that, by restricting to a small enough neighborhood of zero, and adjusting the norm, we can bring r_+ as close as we like to ρ_+ . ■

§6. Reformulation of results. We now want to re-express the main results obtained so far in a way which is less closely tied to specific choices of coordinates, norms, and so on. We use the following terminology: A set X is said to be locally invariant for f at z_0 if there is a neighborhood $\mathcal{U}$ of z_0 such that

$$z \in X \cap \mathcal{U} \text{ implies } f(z) \in X.$$

i.e., such that $f(X \cap \mathcal{U}) \subset X$, and we say that two sets X_1 and X_2 coincide near z_0 if there is a neighborhood $\mathcal{U}$ of z_0 such that

$$\mathcal{U} \cap X_1 = \mathcal{U} \cap X_2.$$

THEOREM 6.1 (GENERALIZED STABLE MANIFOLD THEOREM). Let f be a mapping of class C^r , z_0 a fixed point for f , (ρ_-, ρ_+) a gap in the modulus of the spectrum of $Df(z_0)$, $\mathcal{E} = \mathcal{E}_- \oplus \mathcal{E}_+$ the corresponding splitting of the state space. If $\rho_- = 1$, assume r is finite; if $\rho_- > 1$, assume $\rho_-^r < \rho_+$.

1. There exists a C^r manifold $W_{\text{loc}}^{(-)}$, passing through z_0 and tangent there to $\mathcal{E}_-$ and locally invariant for f at z_0 .
2. If $\rho_- < 1$, $W_{\text{loc}}^{(-)}$ can be taken to be invariant in the literal sense, i.e., mapped into itself by f . In this case, the invariant manifold is locally unique in the sense that any two locally invariant C^1 manifolds passing through z_0 and tangent there to $\mathcal{E}_-$ coincide near z_0 . If z is in such an locally invariant manifold and near enough to z_0 , then $f^n(z)$ converges to z_0 . More precisely, if $\mathcal{U}$ is any sufficiently small neighborhood of z_0 , then any such locally invariant manifold coincides near z_0 with the set of z such the $f^n(z)$ is in $\mathcal{U}$ for all $n = 0, 1, \dots$ and such that

$$(6.1) \quad \limsup_n \frac{1}{n} \log(\|f^n(z) - z_0\|) \leq \log(\rho_-)$$

3. If $\rho_+ > 1$, we have the partial uniqueness result that, for any locally invariant manifold tangent to $\mathcal{E}_-$, and for any sufficiently small neighborhood $\mathcal{U}$ of z_0 , any z such that $f^n(z)$ is in $\mathcal{U}$ for all $n = 0, 1, \dots$ is in $W_{\text{loc}}^{(-)}$.

4. If $\rho_- < 1 < \rho_+$ (hyperbolic case), then for any sufficiently small neighborhood $\mathcal{U}$ of z_0 , $W_{\text{loc}}^{(-)}$ coincides near z_0 with the set of z such that $f^n(z)$ is in $\mathcal{U}$ for all $n = 0, 1, \dots$.

§7. Generalized unstable manifolds. Given a splitting $\mathcal{E}_- \oplus \mathcal{E}_+$ as in the preceding sections, we want to investigate invariant manifolds tangent to $\mathcal{E}_+$. We will call these generalized unstable manifolds. The most important case is that where $\rho_- \leq 1 < \rho_+$; a corresponding invariant manifolds are called an unstable manifold. Similarly, if $\rho_- > 1$, we speak of a strong unstable manifold, and, if $\rho_+ = 1$, of a center-unstable manifold. If f is invertible, we can deduce the existence and properties of generalized unstable manifolds by applying our results about stable manifolds to f^{-1} . It turns out, however, that a theory of unstable manifolds completely parallel to that for stable manifolds can be developed without assuming the invertibility of f . This is a useful thing to do, since in applications to partial differential equations one expects that derivatives of the mappings encountered will often be compact operators and hence not invertible but with finite-dimensional $\mathcal{E}_+$'s, so restricting to a generalized unstable manifold may give a means of extracting the essential finite-dimensional part from such infinite-dimensional dynamical systems.

Although the generalized unstable manifolds parallels that of generalized stable manifolds, I have not been able to invent one formalism which encompasses both cases. Furthermore, unfortunately, the functional equations arising in the theory of unstable manifolds seem to be significantly more complicated than those for stable manifolds.

The initial reductions are just the same: We want to study mappings of the form

$$(3.1) \quad f(x, y) = (\Lambda_- x + f_-(x, y), \Lambda_+ y + f_+(x, y)),$$

with $f_{\pm}$ small in Lipschitz norm and vanishing at the origin, and we seek invariant sets which are graphs of mappings w from $\mathcal{E}_+$ to $\mathcal{E}_-$ with Lipschitz norm no greater than one. The appropriate invariance condition is, however, different if the mapping is non-invertible. To see what is going on, consider the case where f is exactly linear, i.e., where $f_{\pm} = 0$, but where Λ_- has non-trivial null space. In this case, it is easy to see that $\mathcal{E}_-$ satisfies

$$f^{-1}\mathcal{E}_- = \mathcal{E}_- \quad \text{i.e.,} \quad z \in \mathcal{E}_- \quad \text{if and only if} \quad f(z) \in \mathcal{E}_-,$$

and, in the class of graphs of bounded linear mappings from $\mathcal{E}_-$ to $\mathcal{E}_+$, is characterized by the (logically weaker) condition

$$f^{-1}\mathcal{E}_- \supset \mathcal{E}_- \quad \text{i.e.,} \quad z \in \mathcal{E}_- \quad \text{implies} \quad f(z) \in \mathcal{E}_-.$$

On the other hand, it need not satisfy

$$f\mathcal{E}_- \supset \mathcal{E}_-.$$

Appropriately localized versions of these statements are true in the non-linear case: For sufficiently small z (and writing W for the graph of w), we have $f(z) \in W$ if and only if $z \in W$ (The “only if” part of the preceding assertion has not been proved, but a proof is not difficult to supply using the “expanding wedges” idea.) Furthermore, we constructed W by writing a functional equation expressing the requirement that $z \in W$ implies $f(z) \in W$ and showing that this equation has a unique solution in the class of graphs of Lipschitz mappings from $\mathcal{E}_-$ to $\mathcal{E}_+$.

The situation for the unstable manifold is different. Returning to the linear case, it is easy to see that $\mathcal{E}_+$ satisfies

$$f\mathcal{E}_+ = \mathcal{E}_+,$$

and is characterized in the class of graphs of linear mappings from $\mathcal{E}_+$ to $\mathcal{E}_-$ by

$$f\mathcal{E}_+ \supset \mathcal{E}_+.$$

It does not in general satisfy

$$f^{-1}\mathcal{E}_+ = \mathcal{E}_+,$$

since the left-hand side contains in particular the null space of f .

The lesson we draw from all this is that we should try to construct generalized unstable manifolds by writing a functional equation for w corresponding to the requirement that its graph W satisfies

$$fW \supset W.$$

Thus, if $z = (w(y), y)$, we want to find $\bar{z} = (w(\bar{y}), \bar{y})$ such that $f(\bar{z}) = z$. Using the detailed form for f we find the pair of conditions

$$(7.1.a) \quad y = \Lambda_+ \bar{y} + f_+(w(\bar{y}), \bar{y})$$

$$(7.1.b) \quad w(y) = \Lambda_- w(\bar{y}) + f_-(w(\bar{y}), \bar{y}).$$

We treat these conditions as follows: We first argue, under appropriate smallness conditions on f_+ and the assumption $\text{Lip}(w) \leq 1$, that equation (7.1.a) can be solved uniquely for $\bar{y}$ in terms of y . That is, the y -component of the equation $f(\bar{z}) = z$ already determines $\bar{z}$ if $\bar{z}$ is required to lie on the graph of w . We stress: Solvability here does not depend on any special property of w ; it works for all w with Lipschitz norm not greater than one. We will denote the solution by $v(y, w)$; in contrast to the corresponding notation in the case of generalized stable manifolds, v depends in principle on the whole function w .

Solving the y -component of the equation $f(\bar{z}) = z$ does not, however, suffice to solve the equation completely; the x -component of the equation imposes a constraint on w of the form

$$(7.2) \quad w(y) = F(w(v(y, w)), v(y, w)) \quad \text{with} \quad F(x, \bar{y}) = \Lambda_- x + f_-(x, \bar{y}).$$

Thus: Assuming solvability of (7.1.a) for all relevant y , the graph W of w satisfies $fW \supset W$ if and only if w satisfies the functional equation (7.2).

The analysis of (7.2) goes along the same general lines as that already given for (3.9), but with significant differences of detail. The first step is to show that (7.1.a) is solvable. We have two cases to consider: Either y is allowed to run over all of $\mathcal{E}_+$ or is restricted to the unit ball. Let us start with the first case. The equation to be solved can be rewritten as a fixed point problem:

$$\bar{y} = \Lambda_+^{-1}y + \Lambda_+^{-1}f_+(w(\bar{y}), \bar{y}),$$

and the right-hand side is a Lipschitz continuous function of $\bar{y}$ with Lipschitz constant not larger than $\|\Lambda_+^{-1}\|\text{Lip}f_+$. Thus, if

$$(7.3) \quad \|\Lambda_+^{-1}\|\text{Lip}f_+ < 1,$$

(7.1.a) is uniquely solvable for $\bar{y}$ for all y . Now consider the case where y is restricted to the unit ball. The same estimate holds for the Lipschitz norm of the right-hand side of (7.3), with y restricted to the unit ball and $\text{Lip}(f_+)$ interpreted as meaning the Lipschitz norm over the unit ball of $\mathcal{E}_- \oplus \mathcal{E}_+$. To apply the contraction mapping principle, we need to check that the right-hand side of (7.3) maps the unit ball into itself. This will be the case provided that

$$(7.4) \quad \|\Lambda_+^{-1}\| [1 + \text{Lip}(f_+)] \leq 1$$

Thus, if (7.4) holds, we can study the functional equation (7.2) on the unit ball without the necessity of cutting off the nonlinear terms. If we start from a gap (ρ_-, ρ_+) in the modulus of the spectrum of the derivative of a differentiable mapping at a fixed point, it is clear that, by choosing the norm appropriately and magnifying sufficiently, we can make (7.4) hold provided that $\rho_+ > 1$.

As indicated above, we denote $\bar{y}$ solving (7.1.a) by $v(y, w)$. We next want to estimate the Lipschitz norm of v with respect to y . It in fact turns out to be useful to make a slightly more general estimate. Let y_1, y_2 be two points of $\mathcal{E}_+$ or of the unit ball of $\mathcal{E}_+$, and let x_1, x_2 two points of $\mathcal{E}_-$ such that $\|x_i\| \leq \|y_i\|$ and such that $\|x_1 - x_2\| \leq \|y_1 - y_2\|$. Then

$$\begin{aligned} \|\{\Lambda_+ y_1 + f_+(x_1, y_1)\} - \{\Lambda_+ y_2 + f_+(x_2, y_2)\}\| &\geq \|\Lambda_+ (y_1 - y_2)\| - \text{Lip}(f_+) \|y_1 - y_2\| \\ &\geq [\|\Lambda_+^{-1}\| - \text{Lip}(f_+)] \|y_1 - y_2\|. \end{aligned}$$

We now *define* $\text{Lip}(v)$ to be the smallest constant such that

$$\|\{\Lambda_+ y_1 + f_+(x_1, y_1)\} - \{\Lambda_+ y_2 + f_+(x_2, y_2)\}\| \geq \frac{\|y_1 - y_2\|}{\text{Lip}(v)}$$

whenever the x_i and y_i are as above; the computation we have just done shows that

$$\text{Lip}(v) \leq \frac{\|\Lambda_+^{-1}\|}{1 - \|\Lambda_+^{-1}\|\text{Lip}(f_+)} = \|\Lambda_+^{-1}\| + O(\text{Lip}(f_+)).$$

It is evident from the definitions that, for any w with $\text{Lip}(w) \leq 1$, the mapping $y \mapsto v(y, w)$ has Lipschitz norm no larger than the constant we have called $\text{Lip}(v)$, and this is the justification for the terminology although it does not appear to be necessarily true that, for arbitrary $f_{\pm}$, $\text{Lip}(v)$ is the smallest bound on the Lipschitz norms of all the mappings $y \mapsto v(y, w)$. We note for later use that, if the $f_{\pm}$ are continuously differentiable and if A is a linear operator from $\mathcal{E}_+$ to $\mathcal{E}_-$ with $\|A\| \leq 1$,

$$(7.5) \quad \left\| [\Lambda_+ + D_1 f_+ A + D_2 f_+]^{-1} \right\| \leq \text{Lip}(v);$$

this can be seen by applying the definition of $\text{Lip}(v)$ with y_1 and y_2 very close together.

We now want to find conditions which guarantee that the operator $\mathcal{F}$ defined by the right-hand side of (7.2) is a contraction on the space of mapping w from $\mathcal{E}_+$ (or, as appropriate, the unit ball in $\mathcal{E}_+$) to $\mathcal{E}_-$, with $w(0) = 0$ and $\text{Lip}(w) \leq 1$. From the equation (7.1.a) which determines v , together with $f_+(0, 0) = 0$, it follows that $v(0, w) = 0$, and this, together with $F(0, 0) = 0$, implies that $\mathcal{F}$ preserves the condition $w(0) = 0$. Also, the Lipschitz norm of the right-hand side of (7.2) is bounded by $\text{Lip}(F)\text{Lip}(v)$, so

$$(7.6) \quad \text{Lip}(F)\text{Lip}(v) \leq 1$$

suffices to guarantee that $\mathcal{F}$ preserves the condition $\text{Lip}(w) \leq 1$.

To check contractivity we take two mappings w_1, w_2 satisfying all the conditions and write

$$\begin{aligned} (\mathcal{F}(w_1))(y) - (\mathcal{F}(w_2))(y) &= F(w_1(v_1), v_1) - F(w_2(v_1), v_1) \\ &\quad + F(w_2(v_1), v_1) - F(w_2(v_2), v_2), \end{aligned}$$

where we have written v_i for $v(y, w_i)$, $i = 1, 2$. Now

$$\|F(w_1(v_1), v_1) - F(w_2(v_1), v_1)\| \leq \text{Lip}_1(F)\|w_1 - w_2\|\|v_1\| \leq \text{Lip}_2(F)\|w_1 - w_2\|\text{Lip}(v)\|y\|,$$

and

$$\|F(w_2(v_1), v_1) - F(w_2(v_2), v_2)\| \leq \text{Lip}(F)\|v_1 - v_2\|.$$

To make effective use of this last bound, we need to estimate $\|v_1 - v_2\|$ carefully. Writing the two equations for v_1 and v_2 gives

$$\Lambda_+ v_1 + f_+(w_1(v_1), v_1) = \Lambda_+ v_2 + f_+(w_2(v_2), v_2)$$

or

$$\{\Lambda_+ v_1 + f_+(w_1(v_1), v_1)\} - \{\Lambda_+ v_2 + f_+(w_1(v_2), v_2)\} = f_+(w_1(v_2), v_2) - f_+(w_2(v_2), v_2).$$

Taking norms, and using the definition of $\text{Lip}(v)$, gives

$$\begin{aligned} \|v_1 - v_2\| &\leq \text{Lip}(v)\text{Lip}_1(f_+)\|w_1(v_2) - w_2(v_2)\| \\ &\leq \text{Lip}(v)\text{Lip}_1(f_+)\|w_1 - w_2\|\|v_2\| \leq \text{Lip}(v)^2\text{Lip}_1(f_+)\|w_1 - w_2\|\|y\| \end{aligned}$$

By analogy with the corresponding definitions in the stable manifold case, we define

$$\text{Lip}_2(v) \equiv \text{Lip}(v)^2 \text{Lip}_1(f_+),$$

and we can thus write the preceding estimate at

$$\|v_1 - v_2\| \leq \text{Lip}_2(v) \|w_1 - w_2\| \|y\|.$$

Returning to our analysis of the contractivity of $\mathcal{F}$, we find

$$\|(\mathcal{F}(w_1))(y) - (\mathcal{F}(w_2))(y)\| \leq [\text{Lip}_1(F)\text{Lip}(v) + \text{Lip}(F)\text{Lip}_2(v)] \|w_1 - w_2\| \|y\|,$$

and hence we see that $\mathcal{F}$ is contractive provided that

$$(7.7.) \quad [\text{Lip}_1(F)\text{Lip}(v) + \text{Lip}(F)\text{Lip}_2(v)] < 1$$

Thus, if (7.6) and (7.7) hold, $\mathcal{F}$ has a unique fixed point in the space considered. We will denote this fixed point simply by w .

The next step is to investigate the differentiability of w . Again this goes roughly as before: We first differentiate the functional equation for w to get a functional equation which must be satisfied by Dw if w is differentiable; we then prove that this functional equation has a unique solution; then finally we show that the solution is indeed the derivative of w . If w is differentiable, then, differentiating the equation for v gives

$$Dv(y) = [\Lambda_+ + D_1 f_+(w(v), v) Dw(v) + D_2 f_+(w(v), v)]^{-1};$$

then differentiating the equation for w gives

$$Dw(y) = [D_1 F + D_2 F Dw] Dv(y),$$

where suppressed arguments are now to be understood to be $(w(v), v)$. Hence, if w is differentiable, then, writing σ for Dw , we have

$$\sigma(y) = [D_1 F + D_2 F \sigma] [\Lambda_+ + D_1 f_+ \sigma(v) + D_2 f_+]^{-1}.$$

We will denote the right-hand side of this equation as $\mathcal{K}\sigma(x)$, and we want to show that $\mathcal{K}$ is a contraction on the space of continuous mappings $y \mapsto \sigma(y)$ with $\sigma(0) = 0$ and $\|\sigma(y)\| \leq 1$ for all y , equipped with the supremum norm. From $D_1 F(0, 0) = 0$ it follows that $\mathcal{K}$ preserves the condition $\sigma(0) = 0$. It is easy to see that

$$\|[\Lambda_+ + D_1 f_+ \sigma(v) + D_2 f_+]^{-1}\| \leq \text{Lip}(v),$$

and hence $\mathcal{K}$ preserves the condition $\|\sigma(y)\| \leq 1$ for all y provided

$$[\text{Lip}_1(F) + \text{Lip}_2(F)] \text{Lip}(v) < 1,$$

and this is (7.6) again.

To find conditions which make $\mathcal{K}$ contractive, we choose σ_1 and σ_2 and write

$$\begin{aligned} \mathcal{K}\sigma_1(y) - \mathcal{K}\sigma_2(y) &= D_2F[\sigma_1(v) - \sigma_2(v)] [\Lambda_+ + D_1f_+\sigma_1(v) + D_2f_+]^{-1} \\ &\quad + [D_1F + D_2F\sigma_2(v)] [\Lambda_+ + D_1f_+\sigma_1(v) + D_2f_+]^{-1} D_1f_+ \times \\ &\quad [\sigma_2(v) - \sigma_1(v)] [\Lambda_+ + D_1f_+\sigma_1(v) + D_2f_+]^{-1}, \end{aligned}$$

where we have used the standard operator identity

$$U^{-1} - V^{-1} = U^{-1}(V - U)V^{-1}.$$

Using the bound (7.5) repeatedly, and then the definition to $\text{Lip}_2(v)$ to simplify the expression obtained, we get

$$\|\mathcal{K}\sigma_1(y) - \mathcal{K}\sigma_2(y)\| \leq [\text{Lip}_2(F)\text{Lip}(v) + \text{Lip}(F)\text{Lip}_2(v)] \|\sigma_1(v) - \sigma_2(v)\|,$$

which shows that $\mathcal{K}$ is contractive provided that (7.7) holds.

Once again we let σ denote the unique fixed point for $\mathcal{K}$; we write

$$M = \sup_y \limsup_{\delta y \rightarrow 0} \frac{\|w(y + \delta y) - w(y) - \sigma(y)\delta y\|}{\|\delta y\|};$$

and we want to use the functional equation to show that $M \leq \kappa M$ for some $\kappa < 1$. We write

$$\delta v = v(y + \delta y, w) - v(y, w);$$

$$\delta w = w(v(y, w) + \delta v) - w(v(y, w)).$$

This last is a little unnatural, but is convenient because $w(v(y, w))$ occurs frequently on the right-hand sides of the functional equations whereas $w(y)$ does not. Note also that suppressed arguments in $\mathcal{E}_+$ are to be taken to be $v(y, w)$ and in $\mathcal{E}_-$ are to be taken to be $w(v(y, w))$.

Writing the equations satisfied by $v(y, w)$ and by $v(y + \delta y, w)$, subtracting, and using the differentiability of f_+ , gives

$$\begin{aligned} \delta y &= \Lambda_+ \delta v + D_1f_+ \delta w + D_2f_+ \delta v + o(\delta y) \\ &= [\Lambda_+ + D_1f_+\sigma + D_2f_+] \delta v + D_1f_+ [\delta w - \sigma \delta v] + o(\delta y). \end{aligned}$$

Similarly, writing the functional equation for w and using the differentiability of the $f_{\pm}$, we get

$$\begin{aligned} w(y + \delta y) - w(y) &= D_1F\delta w + D_2F\delta v + o(\delta y) \\ &= [D_1F\sigma + D_2F] \delta v + D_1F[\delta w - \sigma \delta v] + o(\delta y) \\ &= [D_1F\sigma + D_2F] [\Lambda_+ + D_1f_+\sigma + D_2f_+]^{-1} \delta y + D_1F[\delta w - \sigma \delta v] \\ &\quad - [D_1F\sigma + D_2F] [\Lambda_+ + D_1f_+\sigma + D_2f_+]^{-1} D_1f_+ [\delta w - \sigma \delta v] + o(\delta y). \end{aligned}$$

The first term on the right of this last equation can be recognized as $\sigma(y)\delta y$, so we get

$$w(y + \delta y) - w(y) - \sigma(y)\delta y = \left\{ D_1 F - [D_1 F \sigma + D_2 F] [\Lambda_+ + D_1 f_+ \sigma + D_2 f_+]^{-1} D_1 f_+ \right\} [\delta w - \sigma \delta v] + o(\delta y).$$

Taking norms, dividing by $\|\delta y\|$, letting δy tend to zero, and using

$$\limsup_{\delta y \rightarrow 0} \frac{\|\delta w - \sigma \delta v\|}{\|\delta y\|} \leq M \text{Lip}(v),$$

we get

$$\limsup_{\delta y \rightarrow 0} \frac{\|w(y + \delta y) - w(y) - \sigma(y)\delta y\|}{\|\delta y\|} \leq [\text{Lip}_1 F \text{Lip}(v) + \text{Lip}(F) \text{Lip}_2(v)] M.$$

Hence, (7.7) implies $M = 0$, i.e., implies that w is differentiable with derivative σ .

The analysis of Hölder continuity of the first derivative and of higher differentiability goes in almost exactly the same way as for generalized stable manifolds. The outcome of the analysis is summed up in the following result:

PROPOSITION 7.1. *Let F, v be as above. Assume that F and f_+ are of class C^r , with $1 \leq r \leq \infty$, and such that*

$$\text{Lip}(F) \text{Lip}(v) \leq 1$$

$$[\text{Lip}_1(F) + \text{Lip}(F) \text{Lip}_1(f_+) \text{Lip}(v)] \text{Lip}(v) < 1$$

$$[\text{Lip}_1(F) + \text{Lip}(F) \text{Lip}_1(f_+) \text{Lip}(v)] \text{Lip}(v)^r < 1$$

(with the last condition replaced by $\text{Lip}(v) \leq 1$ for $r = \infty$.) Then the functional equation (7.2) has a unique solution with $w(0) = 0$ and $\text{Lip}(w) \leq 1$, and this solution is of class C^r .

We now want to apply the expanding and contracting wedges idea of §5. A backward orbit for a point z will mean a sequence $z_{-1}, z_{-2}, \dots, z_{-n}$ with $f(z_{-1}) = z$ and $f(z_{-j}) = z_{-j+1}$ for $j = 2, 3, \dots, n$.

PROPOSITION 7.2. *Let the hypotheses and notation be as in Proposition 5.1, but assume in addition that*

(a) $r_+ > 1$

(b) *There is a function w from the unit ball in $\mathcal{E}_+$ into $\mathcal{E}_-$ with $w(0) = 0$ and $\text{Lip}(w) \leq 1$ whose graph is mapped onto itself by f .*

(c) $f_{\pm}$ are continuously differentiable with derivatives vanishing at $(0, 0)$.

Then, for any z in the graph of w admits a unique infinite backward orbit (z_{-n}) in the graph of w . This backward orbit converges to $(0, 0)$ fast enough so that

$$\limsup_n \frac{1}{n} \log \|z_n\| \leq -\log(\rho_+)$$

Conversely, a z in the unit ball but not in the graph of w does not admit an infinite backward orbit in the unit ball such that

$$\limsup_n \frac{1}{n} \log \|z_n\| \leq -\log(r_-).$$

We now drop the assumption that $r_+ > 1$ and assume instead that

$$r_- < 1;$$

we also let the various Lipschitz norms be taken over the domain $\|y\| \leq \|x\|$ with x unrestricted, and we assume $\text{Lip}(w) \leq 1/2$ rather than $\text{Lip}(w) \leq 1$. Then, if z is a point of the unit ball such that the $f^j(z)$ up to $j = n$ are all in the unit ball, the distance from $f^n(z)$ to the graph of w is not greater than $3r_-^n$. In particular, any orbit which stays in the unit ball for all n converges to the graph of w , and any orbit which starts very near to fixed point—and so stays in the unit ball for a long time—but which nevertheless eventually escapes from the unit ball must escape “along” the graph of w .

PROOF: Under our hypotheses, f is injective on the graph of w and maps a subset of this graph onto the whole graph. Hence, there is a uniquely determined sequence of successive preimages z_{-n} , and these are easily seen to converge to $(0,0)$ at least as fast as r_+^{-n} . Improving this rate of convergence to the one indicated is done by the same arguments as those used to prove Corollary 5.2.

Next suppose that z is not in the graph of w but that it admits successive preimages $z_{-1}, z_{-2}, \dots, z_{-n}$, all in the unit ball. Write $z = (x, y)$; put $z' = (w(y), y)$; and write z'_j for the successive preimages of z' in the graph. By construction, the separation between z and z' is predominantly horizontal, so the separations between the z_j and z'_j are all predominantly horizontal, so

$$\|z - z'\| \leq (r_-)^n \|z_n - z'_n\|.$$

Since the left-hand side of this inequality is non-zero and independent of n , $\|z_n - z'_n\|$ cannot go to zero faster than $(r_-)^{-n}$. But we already know that $\|z'_n\|$ goes to zero at least as fast as $(r_+)^{-n}$, so $\|z_n\|$ cannot go to zero faster than $(r_-)^{-n}$.

We now turn to the last assertion. If $f^n(z)$ is in the graph of w , there is nothing to prove. Otherwise, the argument of the preceding paragraph with z of that paragraph replaced by $f^n(z)$ shows that there is a z' on the graph of w whose separation from z is predominantly vertical and such that the distance from $f^n(z)$ to the graph of w is not greater than $r_-^n \|z - z'\|$. Since we are no longer assuming $r_+ > 1$, we cannot immediately conclude that z' is in the unit ball, but from the facts that z' is on the graph of w and has predominantly vertical separation from a point of the unit ball imply $\|z - z'\| \leq 3$.

To apply this last result, we need to be able to prove the existence of mappings w with $\text{Lip}(w) \leq 1/2$ whose graphs are invariant. This can be deduced by applying the result with $\text{Lip}(w) \leq 1$ but with the norm on $\mathcal{E}_-$ doubled (and the norm on $\mathcal{E}_+$ unchanged.)

This change of norms multiplies some of the $\text{Lip}_i(f_{\pm})$ by 2 and others by 1/2 but the norms of Λ_- and Λ_+^{-1} are unchanged; it thus only changes the smallness conditions on $\text{Lip}(f_{\pm})$ necessary to make the argument work.

We can now reformulate the results of this section in the following theorem:

THEOREM 7.3 (GENERALIZED UNSTABLE MANIFOLD THEOREM). *Let f be a mapping of class C^r , z_0 a fixed point for f , (ρ_-, ρ_+) a gap in the modulus of the spectrum of $Df(z_0)$, $\mathcal{E} = \mathcal{E}_- \oplus \mathcal{E}_+$ the corresponding splitting of the state space. If $\rho_+ = 1$, assume r is finite; if $\rho_+ < 1$, assume $\rho_+^r > \rho_-$.*

1. *There exists a C^r manifold W_{loc}^+ , passing through z_0 and tangent there to $\mathcal{E}_+$ and locally invariant for f at z_0 .*
2. *If $\rho_+ > 1$, W_{loc}^+ can be taken to be "backwards invariant", i.e., mapped onto itself by f . In this case, the invariant manifold is locally unique in the sense that any two locally invariant C^1 manifolds passing through z_0 and tangent there to $\mathcal{E}_+$ coincide near z_0 . Any z is in such an locally invariant manifold and near enough to z_0 admits an infinite sequence of successive preimages converging to z_0 . More precisely, if $\mathcal{U}$ is any sufficiently small neighborhood of z_0 , then any such locally invariant manifold coincides near z_0 with the set of z admitting an infinite sequence z_{-n} of successive preimages, all in $\mathcal{U}$, with*

$$\limsup_n \frac{1}{n} \log(\|z_{-n} - z_0\|) \leq -\log(\rho_+)$$

3. *If $\rho_- < 1$, any orbit which remains near z_0 for all time must converge to W_{loc}^+*
4. *If $\rho_- < 1 < \rho_+$ (hyperbolic case), then for any sufficiently small neighborhood $\mathcal{U}$ of z_0 , W_{loc}^+ coincides near z_0 with the set of z admitting an infinite sequence of successive preimages in $\mathcal{U}$.*

§8. Complements.

8.1 Real analytic mappings. If f is real-analytic, and if $\rho_- < 1$, it is easy to show that the invariant manifold tangent to $\mathcal{E}_-$ is real-analytic: Instead of giving the fundamental proof of existence in a space of Lipschitz continuous functions on a real neighborhood of zero, work instead in a space of complex-analytic functions on a complex neighborhood of zero. The proof of contractivity is just the same, and gives immediately an analytic w , thus bypassing completely the complicated analysis of smoothness given above. Similarly for the invariant manifold tangent to $\mathcal{E}_+$ when $\rho_+ > 1$. In the cases $\rho_- \geq 1$ and $\rho_+ \leq 1$, however, the procedure we have adopted of cutting off the non-linear terms spoils analyticity so this argument does not work. In fact, we will give shortly an example showing that analytic mappings do not necessarily admit analytic invariant manifolds.

8.2 Nonexistence of smooth invariant manifolds. We will give here an example to show that the condition $\rho_-^r < \rho_+$ appearing in Theorem 6.2 for the existence of a C^r generalized stable manifold (in the case where $\rho_- > 1$) is, in at least one instance, sharp. Let $\lambda > 1$, and let

$$f(x, y) = (\lambda x, \lambda^2 y + x^2).$$

The origin is a fixed point, and $Df(0,0)$ is the diagonal matrix with diagonal entries λ, λ^2 . Thus, (λ, λ^2) is a gap in the modulus of the spectrum of $Df(0,0)$, and the corresponding spaces $\mathcal{E}_-$ and $\mathcal{E}_+$ are the x and y axes respectively. Since

$$\rho_-^r = \lambda^r < \rho_+ = \lambda^2 \quad \text{if and only if } r < 2,$$

our general theory asserts that there is, for each $r < 2$, a locally invariant manifold of class C^r tangent to the x axis. We will show that there is no such manifold of class C^2 .

A manifold tangent to the x axis at $(0,0)$ is, locally near $(0,0)$, the graph of a function $y = w(x)$. The condition for invariance is easy to work out directly: The image under f of a point $(x, w(x))$ of the graph is $(\lambda x, \lambda^2 w(x) + x^2)$; this is again in the graph if and only if

$$w(\lambda x) = \lambda^2 w(x) + x^2;$$

hence, the graph is locally invariant if and only if this equation holds for all sufficiently small x . It is nearly trivial to see that this equation cannot have a C^2 solution: Differentiate twice to get

$$\lambda^2 w''(\lambda x) = \lambda^2 w''(x) + 2;$$

then put x equal to 0 to get the contradiction

$$\lambda^2 w''(0) = \lambda^2 w''(0) + 2.$$

8.3 Flows. Consider a differential equation

$$\frac{dz}{dt} = X(z)$$

with solution flow f^t , and let 0 be a stationary solution of this differential equation. The analogue of a gap in the modulus of the spectrum at a fixed point of a map is an interval (σ_-, σ_+) in the real axis such that the spectrum of $DX(0)$ does not intersect the vertical band

$$\{\lambda : \sigma_- < \Re(\lambda) < \sigma_+\},$$

but does intersect the vertical lines $\{\Re(\lambda) = \sigma_\pm\}$. Corresponding to such a band is a splitting of the state space into the direct sum of the spectral subspaces $\mathcal{E}_-$ associated with the part of the spectrum of $DX(0)$ in the left half-plane $\{\Re(\lambda) \leq \sigma_-\}$ and $\mathcal{E}_+$ associated with the part of the spectrum in the right half-plane $\{\Re(\lambda) \geq \sigma_+\}$. As we saw earlier,

$$Df^t(0) = \exp(tDX(0)),$$

so, for each t , $\mathcal{E}_- \oplus \mathcal{E}_+$ is the direct sum decomposition corresponding to a gap

$$(\exp(t\sigma_-), \exp(t\sigma_+))$$

in the modulus of the spectrum of $Df^t(0)$, and there are therefore invariant manifolds for f^t tangent to $\mathcal{E}_-$ and $\mathcal{E}_+$. These invariant manifolds might, in principle, depend

on t . What we want to show here is that there is a single manifold tangent to $\mathcal{E}_-$ locally invariant for all the f^t simultaneously, and similarly for $\mathcal{E}_+$. We will restrict our discussion to invariant manifolds tangent to $\mathcal{E}_-$, i.e., generalized stable manifolds.

In the case $\sigma_- < 0$ (stable and strong stable manifolds), this is quite easy: Let W be a manifold locally invariant for f^1 tangent to $\mathcal{E}_-$. From the Stable Manifold Theorem for maps, we know that such a manifold exists and that it is "locally unique". Since the f^t commute, $f^t W$ is another such locally invariant manifold, and so must coincide near 0 with W . Thus, for any t , W is locally invariant for f^t at 0.

The case $\sigma_- \geq 0$ requires a little more care. In this case, for a map f , we constructed a locally invariant manifold for f by constructing a globally invariant manifold for a cut-off version of f , and this globally invariant manifold is unique. In the case at hand, if we cut off the non-linear terms in the solution mappings f^t , it is hard to see how we will preserve the commutativity needed to apply the argument of the preceding paragraph. What we do instead is to cut off the non-linear terms in the differential equation. That is, we write

$$X(x, y) = (A_-x + X_-(x, y), A_+y + X_+(x, y)),$$

where the $X_{\pm}$ vanish, together with their first derivatives, at the origin. As in the case of mappings, we can make the $X_{\pm}$ small on the unit ball by magnification, and we can then multiply them by a smooth function which is identically one on a neighborhood of the origin and identically zero outside the unit ball to get a globally defined differential equation with globally small non-linear terms. For any t , the time- t solution mappings for the cut-off differential equation agrees with that for the original equation on a neighborhood of the origin (but this neighborhood may become very small as $|t|$ becomes large.)

We now consider only the cut-off equation, which we denote by the same symbols as before. It is not difficult to see that the solution mappings have globally small non-linear terms. Hence, the time-one mapping, for example, admits a unique global invariant manifold W which is the graph of a mapping with Lipschitz norm not greater than one. We would like to argue that W is mapped into itself by all the f^t . By the group property of the solution flow, it is enough to prove this for small t . By commutativity, $f^t W$ is mapped into itself by f^1 . Hence, if we can show that $f^t W$ is the graph of a mapping with Lipschitz norm not greater than one, we are done, by the uniqueness of such a globally invariant graph. One way to proceed here is as follows: By imposing further smallness condition on the non-linear terms in f^1 , we can ensure that W is in fact the graph of a mapping with Lipschitz norm not greater than $1/2$, say. Uniqueness still holds, of course, in the class of maps with Lipschitz norm not greater than one. By making t small enough, we can make f^t globally as near as we like to the identity in the C^1 sense. By making it near enough, we can guarantee that that image under f^t of the graph of any mapping with Lipschitz norm not greater than $1/2$ is again the graph of a mapping, this time with Lipschitz norm not greater than one. Thus, the desired invariance follows by uniqueness.

There are a number of respect in which the arguments we have given are less than optimal:

1. In the case $\sigma_- < 0$, we would like to show how to construct an invariant manifold which is mapped into itself by all f^t with $t > 0$. Such a manifold would be realized as the graph of a mapping of the unit ball in $\mathcal{E}_-$ into $\mathcal{E}_+$, and to make it invariant for all positive t it is necessary to choose the norm on $\mathcal{E}_-$ with more care than we have used. The remark we need is that, for any $s > \sigma_-$, it is possible to choose a norm on $\mathcal{E}_-$ which is equivalent to the original norm and for which

$$\|\exp(tA_-)\| \leq \exp(ts) \quad \text{for all } t \geq 0.$$

If we use such a norm, with $s < 0$, it is not difficult to see that, by making the non-linear terms in X sufficiently small on the unit ball, we can arrange that the invariant manifold for f^1 defined as the graph of a mapping defined on the unit ball in $\mathcal{E}_-$ is indeed mapped into itself by all f^t with $t > 0$.

B. The argument we have given for the case $\sigma_- \geq 0$, involving cutting off the non-linear terms in the differential equation, does not seem to be very well adapted to the case of partial-differential evolution equations. In this latter case, the f^t for small t need not be C^1 -near the identity. It is more likely, however, that the cutting off can be arranged to make them C^1 -near to $\exp(tDX(0))$. If this can be done it is not difficult to see that the *preimage* under f^t , t small and positive, of the graph of a mapping with Lipschitz norm not greater than $1/2$ is the graph of a mapping with Lipschitz norm not greater than one, so uniqueness will imply that $(f^t)^{-1}W = W$ for small positive t and hence for all positive t . A different approach which may work better in this case is suggested on p. 48 of Marsden-McCracken.

8.4 Center manifolds. Suppose we have a fixed point z_0 for a smooth mapping f such that spectrum of $Df(z_0)$ does intersect the unit circle non-trivially but also such that the part of the spectrum on the unit circle is separated from the rest. There is then a splitting of the state space into a direct sum of three subspaces:

$$\mathcal{E} = \mathcal{E}_s \oplus \mathcal{E}_c \oplus \mathcal{E}_u,$$

where $\mathcal{E}_c$, the *center subspace*, is the spectral subspace associated with the part of the spectrum on the unit circle and $\mathcal{E}_s$, $\mathcal{E}_u$ are, as before, the spectral subspaces associated with the parts of the spectrum inside and outside the unit circle respectively. We want to argue that there exists a locally invariant manifold passing through z_0 and tangent there to $\mathcal{E}_c$. Such an invariant manifold is called a *center manifold*. We do this in two steps. We first apply the Generalized Unstable Manifold Theorem to see that there is a locally invariant manifold W_{cu} (a *center-unstable manifold*) tangent to $\mathcal{E}_c \oplus \mathcal{E}_u$. We know that such an invariant manifold exists; we don't know that it is unique, but we know that it can be chosen to be of class C^r , r finite, if f is of class C^r . The idea now is to apply the Generalized Stable Manifold Theorem to the restriction of f to the center-unstable manifold. The spectrum of the derivative of the restriction of f is just the spectrum of the restriction of $Df(z_0)$ to $\mathcal{E}_c \oplus \mathcal{E}_u$, i.e., the part of the spectrum of $Df(z_0)$ on and outside the unit circle. The Generalized Stable Manifold Theorem applies to show that there is a locally invariant manifold for f , contained in the center-unstable

manifold and tangent to $\mathcal{E}_c$ at z_0 . Again, this center manifold can be taken to be of class C^r for any finite r if f is. Evidently, a similar construction can be done for a separated part of the spectrum contained in an annulus centered at the origin, except that smoothness need not be preserved in this more general situation.

The center manifold, although useful for many things, turns out to be a somewhat unnatural object which only barely manages to exist and which is subject to a number of pathologies. We will give here two simple examples showing some of the sorts of things which can go wrong.

1. Non-uniqueness. Consider the pair of differential equations

$$\frac{dx}{dt} = -x; \quad \frac{dy}{dt} = y^2.$$

We can regard them as forming a single differential equation in the plane (of a degenerate form). The origin is a stationary solution; the spectrum of the linearization at the origin has eigenvalues 0 and -1 , with corresponding eigenspaces the y and x axes respectively. A center manifold is a locally invariant manifold tangent to the y axis. In fact, the y axis itself is a center manifold.

There are, however, many more. The equations are easy to solve explicitly:

$$f^t(x, y) = \left(e^{-t}x, \frac{y}{1 - ty} \right)$$

with $t < 1/y$ for $y > 0$ and $t > 1/y$ for $y < 0$. Any solution curve starting in the lower half-plane converges at t goes to infinity to the origin, and it has an infinite-order contact with the y axis there. Thus, any such solution curve can be put together with the origin itself and the positive y axis to form a C^∞ invariant manifold tangent to the y axis at the origin, i.e., a C^∞ center manifold. There are thus continuously many C^∞ center manifolds. Only one of them, the y axis, is analytic. In general, in fact, an analytic mapping or flow cannot admit more than one analytic center manifold because the Taylor series of the center manifold is uniquely determined; see subsection 8.5 below.

2. An analytic mapping with no C^∞ center manifold. We start by generalizing the example given above showing non-existence of a C^2 invariant manifold. Let $\lambda = 2^{1/n}$ and consider the mapping

$$(x, y) \mapsto (\lambda x, 2y + h(x))$$

with h analytic and vanishing to second order at 0. We claim that, unless $h^{(n)}(0) = 0$, there can exist no locally invariant manifold through $(0, 0)$ of the form $y = w(x)$ with w of class C^n . The argument is essentially as before: In order for $y = w(x)$ to be locally invariant, w must satisfy

$$w(\lambda x) = 2w(x) + h(x) \quad \text{for small enough } x.$$

If w is of class C^n near zero, we can differentiate this equation n times, put $x = 0$, and use $\lambda^n = 2$ to get the relation

$$2w^{(n)}(0) = 2w^{(n)}(0) + h^{(n)}(0),$$

which is impossible unless $h^{(n)}(0) = 0$. Thus, if h is not a polynomial, there are infinitely many parameter values λ accumulating at 1 (namely, those of the form $2^{1/n}$ with n such that $h^{(n)}(0) \neq 0$) for which the mapping

$$(x, y) \mapsto (\lambda x, 2y + h(x))$$

does not admit a locally invariant set of the form $y = w(x)$ with w infinitely differentiable.

We now apply an idea due to Ruelle and Takens which is extremely useful in bifurcation theory: We consider the one-parameter family of two-dimensional mappings $(x, y) \mapsto (\lambda x, 2y + h(x))$ as a single three-dimensional mapping by treating the parameter as a third variable with trivial evolution, and we look at center manifolds of this augmented system. That is, we consider

$$(x, y, \lambda) \mapsto (\lambda x, 2y + h(x), \lambda).$$

For each λ , $(0, 0, \lambda)$ is a fixed point, but we want to look particularly at the fixed point with $\lambda = 1$. The derivative of the mapping at that fixed point is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

There are two eigenvalues on the unit circle; the x - λ plane is the center subspace; and so a center manifold will be the graph of a mapping $y = w(x, \lambda)$. We want to show that no such graph, with w infinitely differentiable, can be locally invariant. The key observation is that, if the graph of such a w is locally invariant, then the graph with λ fixed (near enough to one) is locally invariant for

$$(x, y) \mapsto (\lambda x, 2y + h(x)).$$

We have just seen that there is a sequence of λ 's converging to one for which this mapping does not admit a locally invariant set which is the graph of an infinitely differentiable w , so there cannot exist an infinitely differentiable $w(x, \lambda)$ whose graph is locally invariant for the augmented mapping. (This example is modelled on one due to S. J. van Strien, "Center manifolds are not C^∞ ", *Math. Z.* 166 143-145 (1979).)

8.5 Derivatives of invariant manifolds. We have shown how to reduce the problem of finding generalized stable and unstable manifolds to solving functional equations, and have proved existence theorems for those functional equations. We have not, however, been able to give "formulas" for the invariant manifolds. We want to note here that at least the derivatives at the fixed point of the function whose graph is the invariant manifold can be computed algebraically. Consider for definiteness a generalized stable manifold which is the graph of a function w satisfying

$$w(x) = \Lambda_+^{-1} \{w(\Lambda_- x + f_-(x, w(x))) - f_+(x, w(x))\}.$$

We have already seen what the equation given by differentiating this one once looks like. Differentiating a second time gives a discouragingly complicated equation which reduces to

$$D^2w(0) = \Lambda_+^{-1} \{D^2w(0)\Lambda_-^2 - D_1^2f_+(0,0)\}$$

at $x = 0$. This is a simple inhomogeneous linear equation for $D^2w(0)$ which can be solved provided

$$\|\Lambda_+^{-1}\| \|\Lambda_-\|^2 < 1.$$

(It may also be solvable otherwise, but more delicate considerations are involved.)

It is not difficult to see that this procedure can be repeated: If f is of class C^n and if

$$\|\Lambda_+^{-1}\| \|\Lambda_-\|^n < 1,$$

(which guarantees that w is of class C^n), then

$$D^n w(0) = \Lambda_+^{-1} D^n w(0) \Lambda_-^n + \Psi_n,$$

where Ψ_n is an expression involving derivatives of order up to n of the $f_{\pm}$ and derivatives up to order $n-1$ of w , all at the origin. Thus, the $D^n w(0)$ can in principle be computed explicitly.

One consequence of the above analysis, which we will have occasion to use later, is that, if the $f_{\pm}$ vanish to order n at zero, so does w .

8.6 Global stable and unstable manifolds. If f is invertible, it is possible to define global stable and unstable manifolds. We need only consider the stable manifold, since the unstable manifold is simply the stable manifold for f^{-1} . Let z_0 be a hyperbolic fixed point for f , the (global) stable manifold for f means the set W^s of all z such that

$$f^n(z) \rightarrow z_0 \quad \text{as } n \rightarrow \infty.$$

This definition of course makes sense whether or not f is invertible; the reason for restricting to invertible f is to ensure that this set has a nice structure. To see this, let $W_{(loc)}^s$ be a "nice" local stable manifold, i.e., the graph of a smooth mapping from a small ball in $\mathcal{E}_s$ to $\mathcal{E}_u$, where the norm on $\mathcal{E}_s$ is chosen in such a way as to make $W_{(loc)}^s$ be mapped into itself by f . Then $W_{(loc)}^s$ is contained in W^s , and, furthermore, any orbit which stays in a small enough neighborhood of z_0 for all time must be in $W_{(loc)}^s$. Thus, for any $z \in W^s$, $f^n(z) \in W_{(loc)}^s$ for all sufficiently large n . Thus, W^s is a increasing union of a sequence of pieces $f^{-n}W_{(loc)}^s$, each diffeomorphic to a ball in $\mathcal{E}_s$. The global manifold W^s can nevertheless have a complicated structure; loosely speaking, it can double back on itself (but cannot cross itself). Technically, it is a *connected injectively immersed submanifold* of the state space, locally diffeomorphic to $\mathcal{E}_s$.

Strong stable and unstable manifolds can similarly be globalized. If, for example, $r < 1$ belongs to a gap in the modulus of the spectrum of $Df(z_0)$, with z_0 a fixed point of the differentiably invertible mapping f , then the set of z such that

$$\limsup_n \frac{1}{n} \log(\|f^n(z) - z_0\|) < \log(r)$$

is again an injectively immersed submanifold of the state space. Here, the state space can be allowed to be an arbitrary manifold with $\|f^n(z) - z_0\|$ understood to mean the distance between $f^n(z)$ and z_0 with respect to any Riemannian manifold.

We mention, without going fully into the details, that there are, in fact, reasonable definitions of global stable and unstable manifolds even without the assumption that f is invertible. The idea is roughly that the stable manifold is the maximal invariant submanifold of the stable set W^s as defined above. The precise definition is that the global stable manifold is the set of all z such that, for some n ,

- (a) $f^n(z) \in W_{loc}^s$
- (b) The range of $Df^n(z)$ together with the tangent space to W_{loc}^s at $f^n(z)$ spans $\mathcal{E}$ (i.e., f^n is transverse to W_{loc}^s at z .)

At least if $\mathcal{E}$ is finite-dimensional, it is then not difficult to show that this set is a submanifold in the same sense as is the global stable manifold for an invertible mapping, and it is almost obvious that it is mapped into itself by f and contained in the stable set of z_0 .

8.7 Example. Hyperbolic linear torus automorphisms. Let $\mathbf{T}^2$ denote the quotient of $\mathbf{R}^2$ by the additive subgroup $\mathbf{Z}^2$ of points with integer coordinates, i.e., the standard realization of the two-dimensional torus. Let $\tilde{f}$ denote the linear mapping of $\mathbf{R}^2$ to itself generated by the 2×2 matrix

$$F = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

i. e., $\tilde{f}(x, y) = (2x + y, x + y)$. Since $\tilde{f}$ is linear and maps $\mathbf{Z}^2$ to itself, it induces a mapping which we will call f from $\mathbf{T}^2$ to itself. Because the matrix F has unit determinant, its inverse also has integer entries and thus maps $\mathbf{Z}^2$ to itself, and this implies that f is an invertible mapping of $\mathbf{T}^2$ to itself. The image of $(0, 0)$ under the quotient mapping

$$\pi : \mathbf{R}^2 \longrightarrow \mathbf{T}^2$$

is evidently a fixed point for f , and is in fact easily seen to be its only fixed point. In the obvious coordinates, the derivative of f at this fixed point (and everywhere else) is the matrix F . The eigenvalues of F are easy to compute by solving the characteristic equation; their values are

$$\lambda_{\pm} = \frac{3 \pm \sqrt{5}}{2}.$$

Since neither value is on the unit circle, F is hyperbolic, so the fixed point of f is hyperbolic. If we let $\phi_{\pm}$ denote (real) eigenvectors with eigenvalues $\lambda_{\pm}$, then it follows from the "linearity" of f that the global stable manifold of the fixed point is the image of the line

$$\{t\phi_- : -\infty < t < \infty\}$$

under the quotient mapping of $\mathbf{R}^2$ onto $\mathbf{T}^2$. Since this line has irrational slope, its image wraps densely around the torus without ever crossing itself. The global unstable

manifold is similarly the image under the quotient mapping of the line generated by ϕ_+ which again wraps densely around the torus.

This example gives some feeling for how complicated the global structure of stable and unstable manifolds can be even in the simplest cases. It is worth noting, however, that the situation is much more complicated than the preceding analysis shows. For any $n = 1, 2, 3, \dots$, the image under the quotient mapping of $\frac{1}{n}\mathbf{Z}^2$ is a finite subset of $\mathbf{T}^2$ which is mapped into itself by f . Since f is invertible, this set is in fact mapped onto itself by f , and this in turn implies that every point of the set is periodic for f . Since this is true for all n , we have shown that any point of $\mathbf{T}^2$ which admits a representative in $\mathbf{R}^2$ with rational coordinates is periodic for f . In particular, the periodic points are dense in $\mathbf{T}^2$. (It is not difficult to show, using the fact that 1 is not in the spectrum of F^n for any n that, conversely, any periodic point for f admits a representative with rational coordinates.)

If $z_0 = \pi(\bar{z}_0)$ is a periodic point of period p for f , then the derivative of f^p at z_0 , again in the obvious co-ordinates, is just F^p . The eigenvectors of F^p are the same as those of F , and from this it easily follows that the unstable and stable manifolds for f^p at z_0 are the images under the quotient mapping of the lines

$$\{\bar{z}_0 + t\phi_{\pm} : -\infty < t < \infty\},$$

i.e., the stable manifolds (and the unstable manifolds) of periodic points are all "parallel" to each other and each of them is a line wrapping densely around the torus.

This example can be generalized. Everything remains essentially unchanged if we replace $\mathbf{T}^2$ by $\mathbf{T}^m$ and F by any $m \times m$ matrix with integer entries with determinant ± 1 all of whose eigenvalues have modulus different from 1.

8.8 Example. The Hénon mapping. The Hénon mapping is a mapping from the plane $\mathbf{R}^2$ to itself given by the formula

$$(8.3) \quad f(x, y) = (1 - ax^2 + by, x),$$

with a and b constants (parameters). (The right-hand side is often taken to be $(1 - ax^2 + y, bx)$, and it was in this form that Hénon originally wrote the mapping. The two forms differ only by a rescaling of y by a factor of b , and the one we have chosen is better adapted to looking at what happens for small b .) We are going to be interested primarily in small b and, roughly, a between 1 and 2.

It is illuminating to split the Hénon mapping into a sequence of simple steps:

$$(x, y) \mapsto (x, by) \mapsto (x, 1 - ax^2 + by) \mapsto (1 - ax^2 + by, x).$$

The first step is a vertical compression. The second is a *vertical shear*: Each vertical line $x = x_0$ is shifted up or down along itself by an amount $1 - ax_0^2$ depending on x_0 . The last step is simply a reflection through the line $y = x$. The first step multiplies areas by $|b|$, while the second and third steps preserve areas, so the composite mapping (8.3) multiplies areas by a factor of $|b|$. This can also be seen algebraically by checking

that the Jacobian of (8.3) is $-b$. Thus, small b corresponds to strong contraction of areas. (At the other extreme, for $b = \pm 1$ the mapping is area preserving. This is a very interesting case, but not one we will investigate here.) Note also that f is *orientation preserving* for b negative, and orientation reversing for b positive. This gives the cases of positive and negative b slightly different flavors. For example: The eigenvalues of Df^p at a periodic point of period p are either real or complex conjugates of each other. Since their product is $(-b)^p$, the second alternative is not possible for $b > 0$ and p odd. Thus: In the vicinity of a periodic point of odd period p , f^p can look like a rotation followed by a constant contraction (by $|b|^{p/2}$) if b is negative but not if b is positive. We are going to concentrate, for definiteness, on the case $b > 0$.

Somewhat surprisingly, in view of the x^2 term it contains, f is *invertible* as a mapping of the whole plane to itself (provided that $b \neq 0$.) This can be seen either by noting that each of the three elementary steps into which f is decomposed is separately invertible, or by computing its inverse in a straightforward way:

$$f^{-1}(x, y) = \left(y, \frac{1}{b}[x - (1 - ay^2)] \right).$$

Numerical studies first performed by Hénon, and since repeated and extended by many others, strongly suggest that the above mapping f , with $a = 1.4$ and $b = .3$, admits a “fractal strange attractor”, i.e., that almost all orbits starting in a reasonably large region of the plane converge to a set with a self-similar structure which is Cantor-set like in some directions. This is best seen by looking at pictures showing the results of numerical experiments. This situation, simple as it seems, remains outside the grasp of mathematical investigation. As we shall see, the situation for $b = 0$ is fairly well understood, and the most promising approach to understanding the Hénon attractor seems to be regard it as a small modification of the $b = 0$ situation. Even for very small b , our degree of control is very poor. To put our ignorance in a sharp form: No one has exhibited a parameter pair a, b , with b small but non-zero, for which it can be shown that there is a set of orbits with non-zero planar Lebesgue measure none of which converges to an attracting periodic orbit. To put things more concretely if less precisely: It is not excluded that the “Hénon attractor” does not exist at all and that what one sees in the numerical experiments referred to above is a numerical artifact—the true orbit is converging to an attracting cycle with a large period, but the convergence is masked by the effects of round-off error in the computation. I stress that these are only conceptual possibilities; it seems overwhelmingly unlikely that they are true. Part of the difficulty in ruling them out lies in the fact that it seems practically certain that the set of parameter pairs for which f admits an attracting periodic orbit is dense, and fairly likely even that the set of parameter pairs for which f admits infinitely many attracting periodic orbits is dense. Delicate estimates would be needed to show either that there are some parameter pairs with no attracting periodic orbits or that, when there are attracting periodic orbits, they don't always attract essentially all orbits.

Let me return, then, to describing some of what we *do* understand about the Hénon mapping. Suppose we fix a at some value between 1 and 2. If we put $b = 0$, the mapping

is of course no longer invertible; the whole plane is squeezed down onto the parabola $x = 1 - ay^2$. The square

$$\Delta_0 = \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}$$

is mapped onto the arc

$$\{(1 - ay^2, y) : -1 \leq y \leq 1\}.$$

This arc is contained in Δ_0 but touches its boundary at three points. If we open up Δ_0 slightly, say to the rectangle

$$\Delta_\epsilon = \{(x, y) : -(1 + \epsilon) \leq x \leq 1 + \epsilon, -(1 + 2\epsilon) \leq y \leq 1 + 2\epsilon\},$$

with ϵ small, then Δ_ϵ is mapped into its own interior by f (with $b = 0$.) By continuity, then, Δ_ϵ will be mapped into itself by all f with sufficiently small b .

Suppose we pick and fix such a b which is not exactly zero. From the fact that f maps Δ_ϵ into itself it follows that

$$\Delta_\epsilon \supset f\Delta_\epsilon \supset f^2\Delta_\epsilon \supset \cdots \supset f^n\Delta_\epsilon \cdots$$

On the other hand, the area of $f^n\Delta_\epsilon$ is $|b|^n$. Hence

$$X = \bigcap_{n=0}^{\infty} f^n\Delta_\epsilon$$

is a compact set of zero Lebesgue measure, mapped onto itself by f , to which all orbits starting in Δ_ϵ converge. Although the above simple construction of an invariant rectangle doesn't work for Hénon's parameter values $a = 1.4$, $b = .3$, Hénon was nevertheless able, by looking at the mapping in slightly more detail, to construct an invariant quadrilateral in his case as well, so there is also a set X in that case.

It might be thought that the set X , since it attracts all orbits which start in the much larger set Δ_ϵ , would qualify to be called an *attractor*. On closer examination, this does not seem to be a good idea. To see why, suppose for example that f has an attracting periodic orbit in Δ_ϵ . (This certainly happens for some parameter pairs to which our analysis applies.) This orbit—indeed, any periodic orbit, attracting or not—must be in X . Around each point of the periodic cycle there will be a neighborhood of points whose forward orbits are asymptotic to the cycle and which are therefore, in a natural sense, *non-recurrent*. It is easy to see that X , as the decreasing intersection of a sequence of sets each homeomorphic to a rectangle, is connected, so some of these transient points have to belong to X . It seems natural to make attractors as small as possible by leaving out all transient points, and X does not meet this requirement. Accordingly, it is customary to distinguish between *attractors* and *attracting sets*. An *attracting set* simply means a compact set X mapped onto itself by f which has an open neighborhood $\mathcal{U} \supset X$ whose forward images $f^n\mathcal{U}$ shrink down to X in the sense that, for any open set $\mathcal{V} \supset X$, $f^n\mathcal{U} \subset \mathcal{V}$ for all sufficiently large n . What we earlier defined as an attracting fixed point is simply an attracting set with only one element.

Just as in the fixed point case, one can add the requirement that $f\mathcal{U} \subset \mathcal{U}$ without making the definition more restrictive. An attractor then means an attracting set with some further properties. It is still controversial exactly what these properties should be, but two things one certainly wants are that every point of an attractor should be (in some sense) recurrent and that the motion on the attractor should be in some sense indecomposable. The set X constructed above is certainly an attracting set, but need not be an attractor.

It is immediate from the formula (8.3) for f that any fixed point must lie on the line $y = x$. The point (x_0, x_0) is a fixed point if and only if

$$1 - ax_0^2 + bx_0 = x_0.$$

For $a > 0$, whatever the value of b , this equation has one positive solution and one negative solution, so f has exactly two fixed points. As b goes to zero, these fixed points converge to the two points of intersection of the parabola $x = 1 - ay^2$ with the line $y = x$. We are going to study the solution with $x_0 > 0$, for small b . It is easy to check from the equation that this fixed point lies above and to the right of the parabola $x = 1 - ay^2$ for small positive b . Since the fixed point and the derivative of f at it can be computed explicitly, it is straightforward to determine the spectrum of the derivative of f at the fixed point. In this way it can be shown that the fixed point is hyperbolic, with one-dimensional unstable manifold, if and only if $a > 3(1 - b)^2/4$, and that the expanding eigenvalue is negative. For $b = 0$, the contracting eigenvalue is zero, and it is easy to guess what the local stable and unstable manifolds are: The local stable manifold is a vertical line segment through the fixed point, and the local stable manifold is a small segment of the parabola $x = 1 - ay^2$.

(Incidentally, the Hénon mapping with $b = 0$ provides a tame example of the problems with global stable and unstable manifolds for non-invertible mappings. It is nearly obvious that the parabolic arc

$$\{(1 - ay^2, y) : 1 - a \leq y \leq 1\}$$

is mapped into itself by f and it can be proved that the successive images of an arbitrarily small segment of this arc containing the fixed point expand to fill up the whole arc. Thus, if W_{loc}^u is a local unstable manifold, i.e., a small segment of the parabola, then

$$\bigcup_{n=0}^{\infty} f^n W_{loc}^u = \{(1 - ay^2, y) : 1 - a \leq y \leq 1\},$$

and the segment on the right-hand side of this equation is not a submanifold because it contains its end points.)

As soon as b is different from 0, the pathologies of the preceding paragraph have to disappear. What happens for a large enough and b small is that the global unstable manifold winds back and forth infinitely often, always staying close to the above parabolic arc, never crossing itself, with sharper and sharper bends.

We will next give an argument indicating that the Hénon attractor, if it exists, cannot be too different from the closure of the global unstable manifold of the fixed point with $x_0 > 0$. For small b , the unstable manifold has a "first branch" lying roughly along

$$\{(1 - ay^2, y) : 1 - a \leq y \leq 1\},$$

and there is a large piece of the stable manifold lying roughly along the vertical line through the fixed point. The stable and unstable manifolds therefore cross at a point near $(x_0, -x_0)$. (A point of intersection of stable and unstable manifolds for a single hyperbolic fixed point is called a *homoclinic point*; the fixed point itself, of course, doesn't count. If the crossing takes place at non-zero angle, as will be the case here, one speaks of a *transverse homoclinic point*. We will see later on that the existence of a transverse homoclinic point implies that the mapping has quite a lot of complicated structure, and in particular has infinitely many periodic orbits.) The above argument for the existence of a homoclinic point only works for b small, but for the rest of our argument all we need is the existence of a homoclinic point and $|b| < 1$.

Let Δ denote the D-shaped region bounded by the segments of the stable and unstable manifolds joining the fixed point to the homoclinic point. We want to argue that all orbits starting in Δ converge to the global unstable manifold. The proof is very simple: Consider $\Delta_n = f^n \Delta$ with n large. Like Δ , Δ_n is homeomorphic to a disk and its boundary is made up of a piece of stable manifold and a piece of unstable manifold. The piece of stable manifold is however the image under f^n of the piece of stable arc joining the fixed point to the homoclinic point and so is small; most of the boundary is unstable manifold. On the other hand, the area of Δ_n is $|b|^n$ times the area of Δ , and so is small. This means that, for any fixed $\epsilon > 0$, Δ_n for n sufficiently large does not contain any disk of radius ϵ , i.e., every point of Δ_n is within a distance ϵ of the boundary. If, furthermore, n is big enough so that the part of the boundary consisting of stable manifold has length no larger than ϵ , it follows that every point of Δ_n is at distance no greater than 2ϵ from the unstable manifold, and thus every orbit starting in Δ converges to the unstable manifold.

Finally, we present a clever and convenient procedure, due to Franceschini and Russo, for numerical computation of the stable and unstable manifolds of the fixed points of the Hénon mapping. Franceschini and Russo pose the problem as follows: They look for analytic functions $x(t)$, $y(t)$, and a number λ , such that

$$(8.4) \quad f(x(t), y(t)) = (x(\lambda t), y(\lambda t)).$$

That is: They look for a representation of the invariant manifolds as parametrized curves, with parametrization chosen so that the action of f on the curve itself, using the parameter as a coordinate, is simply multiplication by λ . It is clear that if such a representation exists, it is non-unique in a trivial sense; the parameter t can be multiplied by an arbitrary non-zero constant. We can thus expect to have to impose a normalization condition to remove this arbitrariness.

The first observation to be made about (8.4) is that any solution must have $(x(0), y(0))$ a fixed point for f , and we should expect to be able to choose here either of the fixed

points. From the explicit formula for f , the equation (8.4) is equivalent to the pair of equations

$$x(t) = y(\lambda t) \quad \text{and} \quad x(\lambda t) = 1 - ax(t)^2 + by(t).$$

It is now easy to eliminate $y(t)$ from this equation, so what we really have to solve is

$$(8.5) \quad x(\lambda t) = 1 - ax(t)^2 + bx(t/\lambda).$$

Now suppose $x(t)$ has the Taylor series expansion

$$x(t) = \sum_{n=0}^{\infty} x_n t^n.$$

Inserting this expansion into (8.5) gives

$$\sum_{n=0}^{\infty} \left(x_n \lambda^n + a \sum_{j=0}^n x_j x_{n-j} - b \lambda^{-n} \right) t^n = 1.$$

Identifying the constant terms gives

$$x_0 = 1 - ax_0^2 + bx_0,$$

which is exactly the fixed point condition. Identifying the linear terms gives

$$x_1 \lambda + 2ax_0 x_1 - bx_1/\lambda = 0$$

or

$$(\lambda^2 + 2ax_0 \lambda - b) x_1 = 0.$$

The equation

$$\lambda^2 + 2ax_0 \lambda - b = 0$$

is exactly the condition for λ to be an eigenvalue of the derivative of f at the fixed point (x_0, x_0) . If we require this, then x_1 is undetermined and we can fix the normalization for t by putting $x_1 = 1$. (If λ is not an eigenvalue, then $x_1 = 0$ and it is not hard to see that all higher x_n 's are zero. This does give us a solution of (8.4) but not a very interesting one.) For $n \geq 2$ we get

$$x_n \lambda^n + 2ax_0 x_n + a \sum_{j=1}^{n-1} x_j x_{n-j} - b \lambda^{-n} x_n = 0,$$

or

$$(\lambda^n + 2ax_0 - b \lambda^{-n}) x_n = -a \sum_{j=1}^{n-1} x_j x_{n-j}.$$

$$\begin{aligned} x_n &= \frac{-a \sum_{j=1}^{n-1} x_j x_{n-j}}{(\lambda^n - b \lambda^{-n} + 2ax_0)} \\ x_0 &= \frac{b \cdot 1 + \sqrt{(1-b)^2 + 4a}}{2a^2} \\ \lambda &= \frac{-2ax_0 \pm \sqrt{4a^2 x_0^2 + 4b}}{2} \end{aligned}$$

From this last equation it is straightforward to compute as many x_n 's as desired numerically. It is not difficult to see, directly from (8.4) that, if an analytic solution exists at all, it is entire, and the existence of an analytic solution follows from general theory. Thus, the series computed in this way converges for arbitrary t to give as much as desired of the global stable and unstable manifolds (subject, of course, to the limitations imposed by round-off error.)

cut off function
in the theorem of .
unstable unit.

LECTURES ON DYNAMICAL SYSTEMS

Chapter V. Linearizations

Chapter VI. Bifurcations

Oscar E. Lanford III

ETH-Zürich

Disclaimer: This is a preliminary text for a course of lectures given in the academic year 1987-88; it is intended for the use of the students in the course and not for wider distribution. While I will be happy if anyone else finds it useful, the reader should be aware that this manuscript is not in final form and should therefore be used with care. One particularly glaring defect is the lack of any attempt to attribute credit for the results discussed.

O. E. Lanford

Mathematik
ETH-Zentrum
8092 Zürich (Switzerland)

November 1987

§1. The one-dimensional case. We begin our discussion of linearizations by looking at the one-dimensional case, which is much simpler than the general multidimensional situation but nevertheless instructive.

The question we want to address is: Given a mapping f with a fixed point, which we may as well take to be at zero, when does there exist an invertible (i.e., strictly monotone) mapping Ψ defined on a neighborhood of zero, with $\Psi(0) = 0$, such that $\Psi f \Psi^{-1}$ is linear, i.e., such that

$$(1.1) \quad \Psi f \Psi^{-1}(x) = \lambda x,$$

λ a real number, for all x near zero. Such a Ψ is called a linearization of f , or we say that Ψ conjugates f to the linear mapping λx . If Ψ is just a homeomorphism, we will speak of a topological linearization; if Ψ and Ψ^{-1} are of class C^r , we speak of a C^r linearization. Note that the fundamental conjugation equation (1.1) can be rewritten (replacing x by $\Psi(x)$) as

$$\Psi \circ f = \lambda \Psi,$$

which shows that it is a *linear* condition on Ψ . Hence, multiplying any linearization by an arbitrary nonzero constant gives another linearization, so, multiplying Ψ by -1 if necessary, we can arrange to make it strictly increasing. We will assume this from now on.

We will be mostly interested in differentiable mappings f , but we will consider both differentiable and non-differentiable Ψ 's. One of the lessons to be learned from studying the one-dimensional case is that linearizations which are not required to be differentiable can be constructed by crude methods and are not at all unique whereas differentiable linearizations are harder to construct but unique (up to multiplication by constants) when they exist. Note, to begin with, that if f admits a linearization Ψ differentiable at 0 and $\Psi'(0) \neq 0$, then λ must equal $f'(0)$. This need not be the case if Ψ is merely required to be a homeomorphism. For example, if f is already linear, say $f(x) = \lambda x$, and if we take $\Psi(x) = \operatorname{sgn}(x)|x|^\alpha$, with α any strictly positive number, then

$$\Psi f \Psi^{-1}(x) = \operatorname{sgn}(\lambda)|\lambda|^\alpha x.$$

Thus, for example, if f admits a topological linearization with $0 < \lambda < 1$, it also admits a topological linearization with $\lambda = 1/2$. On the other hand, the sign of λ is determined since conjugations preserve the property of being increasing or decreasing.

The next thing to observe is that linearizations with $\lambda = 1$ are *very* special—if (1.1) holds with $\lambda = 1$ then $f(x) = x$ near 0. Those with $\lambda = -1$ are almost as special, since the existence of such a linearization implies that $f^2(x) = x$ near 0. Thus, among mappings $f'(0) = \pm 1$, those which admit differentiable linearizations are exceptional. If, on the other hand, f admits a topological linearization with $\lambda \neq \pm 1$, then the fixed point at 0 of f must be either attracting ($|\lambda| < 1$) or repelling ($|\lambda| > 1$), since these

properties are evident for λx and preserved under conjugation. Thus, for example, the mapping $f(x) = x + x^4$ admits no topological linearization because it attracts from the left and repels to the right (or alternatively because $f(x) > x$ both for small positive and small negative x and no linear mapping has this property.) On the other hand, as we shall see, $f(x) = x + x^3$ does admit a linearization (with, say, $\lambda = 2$.)

We have thus accumulated some necessary conditions for f to admit a topological linearization. These necessary conditions are also sufficient, as the following proposition shows.

PROPOSITION 1.1. *Let f be a continuous strictly monotone function defined on a neighborhood of 0, with $f(0) = 0$. Then f admits a topological linearization at 0 if and only if one of the four following (mutually exclusive) conditions holds:*

- (a) f is increasing and $f(x) = x$ for sufficiently small x .
- (b) f is decreasing and $f^2(x) = x$ for sufficiently small x .
- (c) 0 is an attracting fixed point for f .
- (d) 0 is an attracting fixed point for f^{-1} .

In case (c), we can take $\lambda = 1/2$ if f is increasing and $\lambda = -1/2$ if f is decreasing. In case (d), we can take $\lambda = 2$ if f is increasing and $\lambda = -2$ if f is decreasing.

PROOF: We have already proved the necessity of these conditions; it remains to prove sufficiency. Case (a) is immediate; in case (b), it is easy to verify that

$$\Psi(x) = \begin{cases} x & \text{for } x \geq 0 \\ -f(x) & \text{for } x < 0 \end{cases}$$

is a linearization; and case (d) follows from case (c) applied to f^{-1} . Thus, we only need to consider case (c). We will give the argument only for increasing f ; the proof for decreasing f is similar. We can assume without loss of generality that f is a strictly increasing mapping of $[-1, 1]$ to itself and that $f^n(x) \rightarrow 0$ for all $x \in [-1, 1]$. From this it follows that $f(x) < x$ for all $x \in (0, 1]$ and $f(x) > x$ for all $x \in [-1, 0)$.

Consider $x \in (0, 1]$. Since every orbit in this interval converges to zero, we must have $f(x) < x$ for at least some $x \in (0, 1]$. If this were not true for all such x , the continuity of f would imply that $f(x_0) = x_0$ for some $x_0 \in (0, 1]$, and it would then not be true that $f^n(x_0) \rightarrow 0$. It is worth noting that, conversely, $f(x) < x$ for all $x \in (0, 1]$ implies that $f^n(x) \rightarrow 0$ for all such x . To see this, consider any $x \in (0, 1]$. Then $f^n(x) > 0$ for all n , and $f^n(x)$ is decreasing in n . Hence, $f^n(x)$ converges to something, and this something must—by continuity—be a fixed point of f . But since we are assuming $f(x) < x$ on $(0, 1]$, the only possibility is that the limit is 0.

We are going to argue that any continuous strictly increasing Ψ mapping $[f(1), 1]$ onto $[1/2, 1]$ can be extended (uniquely) to a continuous strictly increasing mapping of $[0, 1]$ to itself satisfying

$$(1.2) \quad \Psi(f(x)) = \frac{1}{2}\Psi(x);$$

the interval $[-1, 0]$ can be treated similarly. We make the extension in a sequence of steps. For the first step, we note that the right-hand side of (1.2) is defined on $[f(1), 1]$, so there is only one way to extend Ψ to $[f^2(1), f(1)]$ consistent with this equation. The extension made in this way is strictly increasing on $[f^2(1), f(1)]$, and the two determinations of $\Psi(f(1))$ agree since we chose Ψ on its initial interval of definition to satisfy $\Psi(f(1)) = 1/2$. Thus, the extended Ψ is continuous and strictly increasing on $[f^2(1), 1]$, and satisfies (1.2) where that equation makes sense, i.e., on $[f(1), 1]$. Now repeat this construction to extend Ψ to $[f^3(1), f^2(1)]$, then to $[f^4(1), f^3(1)]$, etc. Doing this infinitely often, and using $f^n(1) \rightarrow 0$, we get an extension of Ψ to a continuous strictly increasing function defined on $(0, 1]$ and satisfying (1.2) everywhere. From (1.2), it also follows that

$$\Psi(f^n(1)) = \frac{1}{2^n},$$

so we can extend Ψ to a continuous strictly increasing mapping of $[0, 1]$ to itself by putting $\Psi(0) = 0$, and this extension obviously does not spoil (1.2).

It is tempting to try to extend this argument to attracting fixed points for mappings in more than one dimensions. The hard step is to get started: Suppose we consider a mapping f sending the closed unit ball $\bar{B}$ of some Banach space homeomorphically into the open unit ball B , with $f(0) = 0$ and with $\sup_{z \in B} \|f^n(z)\| \rightarrow 0$. The problem is to find a contractive linear mapping Λ and a homeomorphism Ψ of $\bar{B} \setminus fB$ onto $\bar{B} \setminus \Lambda B$ such that $\Psi \circ f = \Lambda \circ \Psi$ on the unit sphere. This is easy to do in one dimension, but not so obvious in more than one. If such a Ψ can be found, the argument given in Proposition 1.1 can be used to extend it to a topological conjugacy between f and Λ . These topological difficulties can easily be circumvented if f is continuously differentiable with differentiable inverse and if the norm of $Df(0)$ is strictly smaller than one. In this case, by cutting off the non-linear terms in f , we can make a diffeomorphism $\tilde{f}$ of B which agrees with f on a neighborhood of 0 and with $Df(0)$ on $\bar{B} - \tilde{f}B$; we can then simply take $\Lambda = Df(0)$ and $\Psi(x) = x$ on $B - \tilde{f}B$.

The degree of arbitrariness in Ψ shows that a topological linearization is only loosely connected with f . Our next observation is that differentiable linearizations, when they exist, are essentially unique. To see this, assume that f has 0 as an attracting fixed point and that some one of the many topological linearizations Ψ of f is differentiable at zero with non-vanishing derivative. (We don't need to require that it be differentiable anywhere else.) By multiplying this linearization by an appropriate constant, we can arrange that $\Psi'(0) = 1$. We now claim that Ψ is uniquely determined. In fact, we are going to show that, for each sufficiently small x ,

$$\lim_{n \rightarrow \infty} \frac{f^n(x)}{\lambda^n}$$

exists and is equal to $\Psi(x)$. To see this, we first note that the linearization identity can be iterated to give

$$\Psi(f^n(x)) = \lambda^n \Psi(x),$$

or, equivalently,

$$\Psi(x) = \frac{\Psi(f^n(x))}{\lambda^n}$$

But since $\Psi(0) = 0$ and $\Psi'(0) = 1$,

$$\frac{\Psi(f^n(x))}{f^n(x)} \longrightarrow 1 \quad \text{provided} \quad f^n(x) \longrightarrow 0. \quad \left. \begin{array}{l} \psi(\kappa) = \frac{\psi(f^n(x))}{\lambda^n} \\ \psi(\kappa) = \lim_{n \rightarrow \infty} \frac{\psi(\kappa)}{\lambda^n} \end{array} \right\}$$

Thus, the desired assertion holds for all x small enough so that $f^n(x) \rightarrow 0$. ■

Next note that, if we write

$$\underline{\Psi_n(x) \equiv \frac{f^n(x)}{\lambda^n}},$$

then

$$\Psi_n \circ f = \lambda \Psi_{n+1}$$

and hence, if $\lim_{n \rightarrow \infty} \Psi_n$ exists and is strictly increasing, this limit is a linearization. To check whether the limit exists, we compute

$$\frac{d}{dx} \Psi_n(x) = \prod_{i=0}^{n-1} \left(\frac{f'(f^i(x))}{\lambda} \right) = \prod_{i=0}^{n-1} (1 + \eta(f^i(x))).$$

Here, we have introduced $\eta(x)$ defined by

$$\underline{f'(x) = \lambda(1 + \eta(x))}.$$

If f is continuously differentiable at zero; if $\lambda = f'(0)$ has modulus less than one; and if x is small enough, then $f^n(x)$ converges to zero exponentially fast, so $\eta(f^n(x))$ converges to zero. Simple continuity of the derivative of f at zero is not enough to imply the convergence of the sequence of products

$$\prod_{i=0}^{n-1} (1 + \eta(f^i(x))),$$

and it is not hard to make an example of a continuously differentiable f for which $\Psi'_n(x)$ goes to infinity for all small non-zero x , with enough uniformity to guarantee that $\Psi_n(x)$ goes to $+\infty$ for all small positive x and to $-\infty$ for all small negative x . Thus, not every C^1 function f with $|f'(0)| < 1$ admits a differentiable linearization. If, however, f is of class C^r for some $r > 1$, so $f'(x)$ is Hölder continuous, then $|\eta(f^n(x))|$ goes to zero *exponentially* with n , so the sequence of products converges uniformly on a neighborhood of zero. Thus, f admits a C^1 linearization Ψ on a neighborhood of zero, and the derivative of Ψ is given by the infinite product

$$\Psi'(x) = \prod_{i=0}^{\infty} (1 + \eta(f^i(x))).$$

It is easy to see from this formula that Ψ is also of class C^r for $1 < r \leq \infty$ if f is, and also that Ψ is analytic near zero if f is.

C-linear in $\mathbb{R}$

PROPOSITION 1.2. *Let f be a mapping of class C^r , $1 < r \leq \infty$, defined on a neighborhood of zero in $\mathbb{R}$, with $f(0) = 0$. Write λ for $f'(0)$. If $\lambda \neq 0, \pm 1$, and if $\mathcal{U}$ is a sufficiently small neighborhood of zero, then f admits exactly one topological linearization Ψ defined on $\mathcal{U}$ which is differentiable at zero and has $\Psi'(0) = 1$. This linearization is of class C^r , and is real-analytic if f is.*

We have already proved the proposition for $|\lambda| < 1$; the case $|\lambda| > 1$ can be reduced to this one by replacing f by f^{-1} .

The result in the analytic case is a classical theorem of Schröder (1871). It can be proved in a number of ways other than the one we have chosen, e.g., by computing a formal power series which formally linearizes f and proving that it has a non-zero radius of convergence. It is not necessary that λ be real. It should also be mentioned here that Siegel has shown that analytic linearizations also exist for λ 's with modulus one whose arguments are sufficiently irrational relative to 2π ; the proof is much more difficult than in the case considered here.

§2. Formal theory. We turn now to the multidimensional case. Let f be a mapping of class at least C^n with $f(0) = 0$, and denote $Df(0)$ by Λ . We want to investigate whether there exists a change of coordinates which eliminates all non-linear terms in f of degree up to n , i.e., a smooth mapping Φ defined on a neighborhood of zero, with $\Phi(0) = 0$ and $D\Phi(0) = 1$, such that

$$\Phi f \Phi^{-1}(z) = \Lambda z + o(\|z\|^n).$$

We might call such a Φ an n -th order linearization of f .

The question of whether such a Φ exists is a purely algebraic one involving only the Taylor coefficients for f and Φ at zero up to order n . In fact, the problem is even linear in the Taylor coefficients of Φ since it can be rewritten as

$$\Phi(f(z)) = \Lambda \Phi(z) + o(\|z\|^n).$$

Thus, in principle, the solvability of this problem reduces to a question of the non-vanishing of certain determinants constructed out of the Taylor coefficients of f . It turns out, however, that the non-vanishing of these determinants is equivalent to the absence of certain simple algebraic relations among the eigenvalues of Λ , i.e., in a vague sense, to the absence of certain kinds of "resonances". The equivalence between the solvability of the n -th order linearization problem and the absence of resonances is the main idea in the theory of smooth linearizations; getting from the existence of n -th order linearizations to the existence of exact linearizations involves some hard technical work but is much less enlightening.

We rewrite the approximate linearization equation as

$$\Phi(z) = \Lambda^{-1} \Phi(f(z)) + o(\|z\|^n),$$

or, equivalently,

$$D^j \Phi(0) = \Lambda^{-1} D^j [\Phi \circ f](0) \quad \text{for } j = 2, \dots, n.$$

It is easy to see by induction that (for $j \geq 2$)

$$D^j [\Phi \circ f] = [(D^j \Phi) \circ f] Df^j + [(D\Phi) \circ f] D^j f + \text{lower order terms},$$

where the lower order terms mean ones involving only derivatives of order $j - 1$ and smaller. Writing f_j for $D^j f(0)$ and Φ_j for $D^j \Phi(0)$, we can thus write

$$\Lambda^{-1} D^j [\Phi \circ f](0) = \Lambda^{-1} \Phi_j \Lambda^j + \Lambda^{-1} f_j + W_j,$$

where the W_j are expressions involving only the f_k and Φ_k for $k < j$. The equations to be solved are thus

$$(2.1) \quad \Phi_j - \Lambda^{-1} \Phi_j \Lambda^j = \Lambda^{-1} f_j + W_j \quad \text{for } 2 \leq j \leq n,$$

and the idea is to solve them successively starting with $j = 2$.

Recall that the Φ_j and f_j are symmetric j -linear mappings from $\mathcal{E} \times \dots \times \mathcal{E}$ to $\mathcal{E}$. We will denote the space of all such mappings by $\mathcal{L}_j$; it is again a finite-dimensional vector space. Each of the equations (2.1) has the schematic form

$$\Phi_j - K_j \Phi_j = a_j,$$

where K_j is the linear operator on the vector space $\mathcal{L}_j$ given by

$$K_j \Phi_j(z_1, \dots, z_j) = \Lambda^{-1} \Phi_j(\Lambda z_1, \dots, \Lambda z_j).$$

Equation (2.1) will be solvable for arbitrary f_j if and only if one is not in the spectrum of K_j . We therefore want to determine this spectrum.

If Λ is a diagonalizable operator then it is easy to see that K_j is also diagonalizable and that the spectrum of K_j is the set of products

$$\lambda_0^{-1} \lambda_1 \lambda_2 \dots \lambda_j \quad \text{with the } \lambda_i \text{ in the spectrum of } \Lambda \text{ (not necessarily distinct).}$$

The result is exactly the same in the non-diagonalizable case. We can see this by approximation, i.e., we use

- (1) every linear operator can be approximated arbitrarily well by diagonalizable ones;
- (2) the mapping $\Lambda \mapsto K_j$ is continuous;
- (3) the spectrum of a linear operator varies continuously with the operator.*

Thus: Equations (2.1) can be solved for all $f_2, \dots, f_n$ if and only if no relation of the form

$$\lambda_0 = \lambda_1 \dots \lambda_j$$

holds among the eigenvalues of Λ for $2 \leq j \leq n$. We have thus proved:

*Remark and question: The argument given only works on finite-dimensional spaces. Is the result nevertheless true in the infinite-dimensional situation?

PROPOSITION 2.1. Let f be a mapping of class C^n , $n \geq 2$, which leaves 0 fixed, and let Λ denote $Df(0)$. If no relation

$$(*) \quad \lambda_0 = \lambda_1 \lambda_2 \cdots \lambda_j \quad \text{with } \lambda_0, \dots, \lambda_j \text{ eigenvalues of } \Lambda$$

holds, for $2 \leq j \leq n$, then there exists a C^∞ mapping Φ from $\mathcal{E}$ to itself, with

$$\Phi(z) = z + \text{higher order terms},$$

such that

$$\Phi \circ f \circ \Phi^{-1}(z) = \Lambda z + o(\|z\|^n),$$

and the derivatives of Φ at zero up to order n are uniquely determined. Conversely, if one or more relations of the form $(*)$ do hold, then there is a set of non-trivial algebraic conditions which the derivatives of f of order $2, \dots, j$ at 0 must satisfy in order that there exist such a Φ .

Note that, if Λ has an eigenvalue λ on the unit circle, then, since the complex conjugate $\bar{\lambda}$ of λ is also an eigenvalue, and since

$$\lambda = |\lambda|^2 \bar{\lambda} = \lambda^2 \bar{\lambda},$$

we cannot in general eliminate terms of degree 3 in f . Thus, we cannot expect smooth linearizations to exist with any generality at non-hyperbolic fixed points. Furthermore, for hyperbolic fixed points, there is a qualitative difference between the situation where all eigenvalues are on one side of the unit circle (i.e., they are either all inside or all outside) and the non-trivially hyperbolic situation. In the first case, assuming for definiteness that all eigenvalues are inside the unit circle, no relation of the form $(*)$ can hold for j big enough so that

$$|\lambda_{\max}|^j < |\lambda_{\min}|,$$

where $\lambda_{\max}$ and $\lambda_{\min}$ denote eigenvalues with maximal and minimal moduli. Thus, in the neighborhood of any particular such Λ there are only *finitely many* conditions to be avoided in order to be sure that terms of arbitrarily high degree can be eliminated. This implies that the set of Λ 's for which no relation $(*)$ holds for any j is open (and dense) in the set of Λ 's with $\rho(\Lambda) < 1$. On the other hand, it is easy to see that the set of Λ for which some relation $(*)$ holds is dense in the set of all Λ 's with eigenvalues both inside and outside the unit circle. The complementary set where no such relation holds is also dense, and in fact *residual*, i.e., a countable intersection of open dense sets. Thus, loosely, among C^∞ f 's admitting 0 as a nontrivially hyperbolic fixed point, those which do not admit a C^∞ linearization are dense. On the other hand, we see that in some sense "most" such f 's admit at least approximate linearizations of arbitrarily high order.

If relations of the form $(*)$ hold, one can try to choose a Φ which eliminates as many terms in the Taylor series for f as possible, i.e., to make a coordinate change reducing f to some *normal form*, where the normal form depends on exactly which relations $(*)$ hold. This does indeed work in many useful cases, but is a little awkward to formulate in general.

C[∞] linearization in the neighborhood of a fixed point.
 Abb. his zur Ordnung n

§3. Attracting fixed points. In this section we find conditions which guarantee the existence of smooth linearizations for the case where the spectral radius of $Df(0)$ is strictly less than one. It is striking that the argument we are going to give may require the existence of very high order derivatives of f in order to construct a linearization which is even once differentiable, but that, once it works at all, it produces a linearization which is as smooth as f is.

We denote the nonlinear part of f by Λg , i.e., we write

$$f(z) = \Lambda\{z + g(z)\}.$$

The idea is going to be to argue that if $g(z)$ is smooth enough and vanishes to high enough order at 0 then f admits a differentiable linearization. This part of the argument does not involve any non-resonance conditions. To get from a general f to one with a g which vanishes to high order, however, we use the proposition of the preceding section which does require non-resonance conditions. We write the linearization Ψ we are looking for as

$$\Psi(z) = z + \psi(z),$$

with $\psi(z) = O(z^2)$ (at least). We can thus rewrite the linearization equation as

$$\Psi(f(z)) = \Lambda z + \Lambda g(z) + \psi(f(z)) = \Lambda \Psi(z) = \Lambda z + \Lambda \psi(z)$$

or

$$\Lambda \psi(z) - \psi(f(z)) = \Lambda g(z)$$

or

$$(3.1) \quad \psi(z) - \Lambda^{-1}\psi(f(z)) = g(z).$$

This equation has the schematic form

$$\psi - K\psi = g,$$

and we can try to solve it by constructing a function space which contains g and on which K is a contraction, or, alternatively, by proving that the formal Neumann series solution

$$\sum_{j=0}^{\infty} K^j g(z) = \sum_{j=0}^{\infty} \Lambda^{-j} g(f^j(z))$$

converges in some sufficiently strong sense. To start with, let us just investigate the pointwise convergence of this series. The factor Λ^{-n} grows with n , but it multiplies $g(f^n(z))$ which goes to zero because $f^n(z)$ converges to zero and g vanishes there. In fact, by making g vanish to high enough order we can make $g(f^n(z))$ go to zero fast enough to beat the growth of Λ^{-n} .

The details go as follows: We choose a norm with respect to which $\|\Lambda\| < 1$; then an $\epsilon > 0$ and a $\kappa < 1$ such that $\|Df(z)\| \leq \kappa$ on the ball of radius ϵ about zero. By rescaling we can assume that $\epsilon = 1$. Then, for any z in the unit ball,

$$\|f^n(z)\| \leq \kappa^n \|z\| \quad \text{for } n \geq 0,$$

so, if g vanishes to order N at zero,

$$\|g(f^n(z))\| \leq \text{const } \kappa^{Nn} \|z\|^N.$$

Thus, we have only to take N big enough so that

$$(3.2) \quad \kappa^N \|\Lambda^{-1}\| < 1$$

to ensure that the series converges uniformly on the unit ball.

Next we see whether the differentiated series series converges uniformly. The derivative of the n -th term is

$$\Lambda^{-n} Dg(f^n(z)) Df(f^{n-1}(z)) Df(f^{n-2}(z)) \dots Df(z).$$

If g is of class C^N and vanishes to order N at zero, then Dg vanishes to order $N - 1$, so the factor $Dg(f^n(z))$ only gives us a decay like $\kappa^{(N-1)n}$. However, we also have n factors of Df , each of which contributes a factor of κ to the estimates, so the final result is just the same as for the undifferentiated series—condition (3.2) suffices to guarantee that the series converges in C^1 and hence that f admits a C^1 linearization.

This argument can be extended in a completely straightforward way to higher derivatives; what comes out is that Ψ is of class C^N and in fact of class C^n , $n \geq N$, if f is. The details of the extension, done in this way, are a little messy, and it is easier to carry out the proof by constructing a norm on C^n with respect to which K is a contraction. The proof of the following proposition shows in detail how to do this. In formulating this proposition, we have taken the opportunity to prove something a bit more general than what we need here; the more general version will be useful when we discuss linearizations for general hyperbolic fixed points. Also for convenience later on, we include in the proposition a generalized version of the formal theory of linearization to finite order developed in the preceding section.

Since the precise formulation of the proposition is a little intricate, it may be helpful to start with an informal survey of what it asserts. It gives sufficient conditions for equations of the form

$$\psi(x) - L(x)\psi(f(x)) = a(x)$$

to be uniquely solvable for ψ , which is to be a smooth mapping from the unit ball in some vector space $\mathcal{E}$ into some (other) vector space $\mathcal{F}$. Here, L , f , and a are given;

- (a) L is a smooth mapping taking values in the space of linear mappings from $\mathcal{F}$ to itself;
- (b) f is a contractive smooth mapping from the unit ball of $\mathcal{E}$ into itself with the origin as a fixed point; and
- (c) a is a smooth mapping from the unit ball in $\mathcal{E}$ into $\mathcal{F}$.

They are all assumed to be of class C^n for some sufficiently large n ; how large n has to be is related to how strongly f contracts and how large $L(x)$ gets. There is a non-resonance condition saying that eigenvalues of $L(0)$ cannot be written as products of fewer than n reciprocals of eigenvalues of $Df(0)$. However, resonance of *low* order—of order less than some $n_0 < n$ —are not excluded; instead, we require that a vanish to order n_0 at the origin and look only for ψ 's which also vanish to this order. For the purposes of the statement and proof of this proposition, "the unit ball" will mean the *open* unit ball, and functions of class C^n on the unit ball are assumed to have all derivatives up to order n bounded on the unit ball.

PROPOSITION 3.1. *Let $\mathcal{E}, \mathcal{F}$ be finite-dimensional normed vector spaces, and let $0 \leq n_0 < n$ be integers. Also, let f a mapping of class C^n from the unit ball of $\mathcal{E}$ into itself with the origin as a fixed point and $x \mapsto L(x)$ a mapping of class C^n from the unit ball in $\mathcal{E}$ into the space $\mathcal{L}(\mathcal{F}, \mathcal{F})$ of linear operators on $\mathcal{F}$. Assume that*

$$\kappa \equiv \sup_{\|x\| < 1} \|Df(x)\| < 1,$$

and define M as

$$M \equiv \sup_{\|x\| < 1} \|L(x)\|.$$

Assume:

- (1) $\kappa^n M < 1$.
- (2) (Non-resonance) No non-zero eigenvalue of $L(0)$ can be written as the reciprocal of a product of j eigenvalues of $Df(0)$, with $n_0 \leq j < n$.

Let a be a mapping of class C^n from the unit ball in $\mathcal{E}$ into $\mathcal{F}$ with $D^j a(0) = 0$ for $0 \leq j < n_0$. (This latter condition is empty if $n_0 = 0$.) Then there is exactly one mapping ψ of class C^n from the unit ball of $\mathcal{E}$ into $\mathcal{F}$, such that

$$(3.3) \quad \psi(x) - L(x)\psi(f(x)) = a(x) \quad \text{for all } x \text{ in the unit ball of } \mathcal{E}.$$

and such that

$$D^j \psi(0) = 0 \quad \text{for } 0 \leq j < n_0.$$

PROOF: The proof splits into an algebraic part and an analytic part. The algebraic part—the only place where the non-resonance condition (2) is used—is to show that there is a unique polynomial

$$\psi^{(0)}(x) = \sum_{j=n_0}^{n-1} \psi_j^{(0)}(x) \quad \text{with } \psi_j^{(0)} \text{ a homogeneous polynomial of degree } j$$

such that

$$\psi^{(0)}(x) - L(x)\psi^{(0)}(f(x)) - a(x) \text{ vanishes to order } n \text{ at } 0.$$

The analytic part is to show that, if $a(x)$ vanishes to order n at zero, then (3.3) has a unique C^n solution with $D^j\psi(0) = 0$ for $j < n$.

The proof of the algebraic part is an easy generalization of the proof of Proposition 2.1. Writing Λ for $Df(0)$, we first note that

$$[D^j L(x)\psi(f(x))]_{x=0} = L(0)D^j\psi(0)\Lambda^j + \text{lower order terms.}$$

where the "lower order terms" involve ψ only through its derivatives at 0 of order strictly less than j . We can thus determine successively the derivatives of order $n_0, \dots, n-1$ at 0 of the desired $\psi^{(0)}$ provided that, for each $j = n_0, \dots, n-1$, the operator $\mathcal{L}_j$ on the space of symmetric j -linear mappings ϕ from $\mathcal{E} \times \dots \times \mathcal{E}$ to $\mathcal{F}$ given by

$$\mathcal{L}_j\phi(x_1, \dots, x_j) = L(0)\phi(\Lambda x_1, \dots, \Lambda x_j)$$

does not have 1 in its spectrum. But the spectrum of this operator is easily verified to be exactly $\sigma(L(0))\sigma(\Lambda)^j$, i.e., the set of all products of one eigenvalue of $L(0)$ and j not-necessarily-distinct eigenvalues of Λ , and the non-resonance condition (2) of the proposition says precisely that no such product is equal to one.

To prove the analytic part, we let K denote the operator defined by

$$(K\psi)(x) = L(x)\psi(f(x))$$

on the space of functions ψ of class C^n from the unit ball of $\mathcal{E}$ into $\mathcal{F}$ with $D^j\psi(0) = 0$ for $j < n$. We will construct a norm which makes this space of functions into a Banach space and with respect to which K is a contraction. For such a function ψ , and for $j \leq n$, define

$$\|\psi\|_j = \sup_{0 < \|x\| < 1} \frac{\|D^j\psi(x)\|}{\|x\|^{n-j}}.$$

These suprema are all finite, and the space in question is complete with respect to, say, the supremum over $0 \leq j \leq n$ of the $\|\cdot\|_j$. It is not very likely that the operator K will be contractive in this particular norm, but we are going to show how to construct an equivalent one in which K is contractive. It is easy to see that

$$D^j K\psi(x) = L(x)D^j\psi(f(x))Df(x)^j + \text{lower order terms,}$$

where the lower order terms are ones involving derivatives of ψ only up to $j-1$ -st order. Because

$$\|f(x)\| \leq \kappa\|x\|,$$

we get

$$\frac{\|D^j\psi(f(x))\|}{\|x\|^{n-j}} \leq \kappa^{n-j}\|\psi\|_j,$$

and the norm of $Df(z)^j$ is majorized by κ^j . Hence,

$$\sup_x \frac{\|L(x)D^j\psi(f(x))Df^j(x)\|}{\|x\|^{n-j}} \leq M\kappa^n \|\psi\|_j,$$

and thus we get an estimate

$$\|K\psi\|_j \leq M\kappa^n \|\psi\|_j + B_j \sup_{k < j} \{\|\psi\|_k\},$$

where B_j is a constant independent of ψ .

The estimates we get in this way have the general structure

$$\|K\psi\|_j \leq a_j \|\psi\|_j + B_j \sup_{k < j} \{\|\psi\|_k\} \quad \text{with } a_j < 1.$$

The idea now is that, no matter how big B_j is, it can be made negligible by adjusting the relative weights of the various norms. Concretely, we let C_j be a decreasing sequence of positive numbers with $C_0 = 1$ and

$$B_j \frac{C_j}{C_{j-1}} \leq \frac{1 - a_j}{2},$$

and put

$$\|\psi\| = \sup_{j \leq n} \{C_j \|\psi\|_j\}.$$

Then

$$\begin{aligned} C_j \|K\psi\|_j &\leq a_j \|\psi\| + B_j \sup_{k < j} \left\{ \left(\frac{C_j}{C_k} \right) C_k \|\psi\|_k \right\} \\ &\leq \left(a_j + B_j \frac{C_j}{C_{j-1}} \right) \|\psi\| \leq \left(\frac{1 + a_j}{2} \right) \|\psi\| \end{aligned}$$

Thus,

$$\|K\psi\| \leq \sup_{j \leq n} \{(1 + a_j)/2\} \|\psi\|,$$

and hence K is contractive relative to $\|\cdot\|$, as desired, so the proof of Proposition 3.1 is complete.

As an almost immediate consequence, we have:

THEOREM 3.2. *Let f be a differentiable mapping; z_0 a fixed point for f ; and assume that $Df(z_0)$ has spectral radius strictly smaller than one. Write Λ for $Df(z_0)$, and let the eigenvalues of Λ be $\lambda_1, \dots, \lambda_m$, labelled in such a way that*

$$|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_m| (< 1).$$

Let N be the smallest integer such that

$$|\lambda_m|^N < |\lambda_1|.$$

Assume

- (a) f is of class C^n , $N \leq n \leq \infty$.
- (b) $\lambda_{i_0} \neq \lambda_{i_1} \dots \lambda_{i_j}$ for all $i_0, i_1, \dots, i_j$ and all j with $2 \leq j < N$

Then there exists a C^r linearization for f , and the linearization is unique (up to normalization) in C^N .

PROOF: Let $\kappa > |\lambda_1|$ and $M > |\lambda_m|^{-1}$ be such that $\kappa^N M < 1$. Choose a norm so that $\|\Lambda\| < \kappa$ and $\|\Lambda^{-1}\| < M$. By magnifying sufficiently, we can arrange that $\|Df(z)\| \leq \kappa$ for all z in the unit ball, so Proposition 3.1 applies. Note that, as we saw in the preceding section, the inequality

$$\lambda_{i_0} \neq \lambda_{i_1} \dots \lambda_{i_j}$$

always holds for $j \geq N$.

§4. General hyperbolic fixed points. Now suppose that zero is a non-trivially hyperbolic fixed point for f , i.e., that $Df(0)$ has spectrum both inside and outside of the unit circle. We are going to show that, if f is smooth enough, and if no relations

$$\lambda_0 = \lambda_1 \lambda_2 \dots \lambda_j$$

hold between the eigenvalues of $\Lambda = Df(0)$ for j up to some finite N , then f admits a C^r linearization. The number of derivatives f has to have, and the N up to which resonances have to be excluded, depend on r and on "coarse" information about Λ . It seems to be necessary, to be sure that f has a C^r linearization, to assume that f has many more than r derivatives.

We begin by making a number of preliminary reductions. We start by splitting the space into the expanding and contracting subspaces for Λ and writing, as usual,

$$f(x, y) = (\Lambda_s x + f_s(x, y), \Lambda_u y + f_u(x, y)).$$

It is very convenient to arrange things so that the stable and unstable manifolds are exactly the linear subspaces $\{y = 0\}$ and $\{x = 0\}$ respectively. This is done as follows: Suppose that the stable manifold is $\{y = w_s(x)\}$ and the unstable manifold $\{x = w_u(y)\}$. Put $x' = x - w_u(y)$, $y' = y - w_s(x)$. Because the derivatives of w_s and w_u vanish at zero, the mapping $(x, y) \mapsto (x', y')$ has derivative 1 at zero, so (x', y') are good local coordinates at zero, and they are chosen precisely so that the unstable manifold is the set where y' vanishes and the stable manifold the set where x' vanishes. We will from now on assume that we have made this coordinate change, and we drop the primes. Note that $\{x = 0\}$ is an invariant set if and only if $f_s(0, y) = 0$ for all y , so the vanishing of f_s for $x = 0$ is the same as the statement that the unstable manifold is $\{x = 0\}$.

We will also want to assume that the non-linear terms in f are *globally* small. We can arrange this by constructing a new f in which the non-linear terms are multiplied by a smooth cut-off function which is one on a neighborhood of zero. Cutting off in this way does not spoil, e.g. the fact that $f_s(0, y) = 0$ for all y , i.e., does not spoil the way we have arranged the stable and unstable manifolds in the preceding paragraph. A linearization for the cut-off f will also be a linearization for f itself on a small enough neighborhood of zero.

We are looking for a linearization Ψ with $D\Psi(0) = 1$, i.e., one which can be written $\Psi(z) = z + \psi(z)$ with ψ vanishing to second order at zero. We can split ψ into its $\mathcal{E}_s$ and $\mathcal{E}_u$ components, i.e., we write

$$\psi(x, y) = (\psi_s(x, y), \psi_u(x, y)).$$

The linearization condition $\Psi(f(z)) = \Lambda\Psi(z)$ then splits into two conditions

$$\Lambda_s x + f_s(x, y) + \psi_s(f(x, y)) = \Lambda_s x + \Lambda_s \psi_s(x, y),$$

$$\Lambda_u y + f_u(x, y) + \psi_u(f(x, y)) = \Lambda_u y + \Lambda_u \psi_u(x, y).$$

These are two independent equations for ψ_s and ψ_u which can be attacked separately. Furthermore, the ψ_u equation is just the ψ_s equation for f^{-1} , so we have only to show that the ψ_s equation has a solution.

Hence, cleaning up the equation a bit, what we have to show is that there exists a smooth function ψ_s from $\mathcal{E}$ into $\mathcal{E}_s$ which vanishes to second order at zero and which satisfies

$$(4.1) \quad \psi_s(z) - \Lambda_s^{-1} \psi_s(f(z)) = \Lambda_s^{-1} f_s(z).$$

We will denote the right-hand side of this equation by $a(z)$. Our approach will be to try to make the formal series solution

$$(4.2) \quad \sum_{n=0}^{\infty} \Lambda_s^{-n} a(f^n(z))$$

converge in C^r , as we did in the case of an attracting fixed point.

Note one curious feature of this series. If z is not on the stable manifold, then $f^n(z)$ will eventually be driven away from zero. Thus, since we have arranged to make the non-linear terms in f vanish outside a small neighborhood of the zero, there will be only finitely many non-zero terms in the series, so it certainly converges and the sum is as smooth as f itself. *The whole problem is to control its behavior on and near the stable manifold.*

Next we note that, using the expanding/contracting wedges idea and the fact that f is globally near to Λ , it is easy to arrange the choice of norms and cutoffs so that the norm of the x component of $f(x, y)$ is smaller than the norm of x by a factor of $\kappa < 1$, where κ is independent of (x, y) . Thus, writing a general orbit

$$f^n(x, y) = (x_n, y_n),$$

we get an estimate

$$\|x_n\| \leq \kappa^n \|x\|,$$

i.e., every orbit converges exponentially fast to the unstable manifold, and we have uniform control over the rate of convergence. Thus, if

$$\|a(x, y)\| \leq k\|x\|^N \quad \text{uniformly in } y,$$

where N is large enough so that

$$\kappa^N \|\Lambda_s^{-1}\| < 1,$$

then, just as in the proof for the case of an attracting fixed point, the series (4.2) converges uniformly on bounded sets, and so defines at least a continuous solution of (4.1).

Controlling the differentiability of the sum is also straightforward, as in the attracting case, but, in contrast to the attracting case, it makes new demands on the rate at which $a(x, y)$ vanishes with x . If $a(x, y)$ vanishes to order N in x uniformly in y , then $Da(x, y)$ vanishes to order $N - 1$. Using

$$D\Lambda_s^{-n}a(f^n(z)) = \Lambda_s^{-n}Da(f^n(z))Df(f^{n-1}(z)) \dots Df(z),$$

we get an estimate

$$\|D\Lambda_s^{-n}a(f^n(z))\| \leq \text{const} \|\Lambda_s^{-1}\|^n \kappa^{n(N-1)} M^n,$$

where M is a global bound on $\|Df(z)\|$. Thus, we can conclude that the sum (4.2) is continuously differentiable provided that $a(x, y)$ vanishes to order N in x , for all y , where N is large enough so that

$$\kappa^{N-1} \|\Lambda_s^{-1}\| M < 1.$$

Continuing in this way, we see that there is a C^n solution provided that N is big enough so that

$$\kappa^{N-n} \|\Lambda_s^{-1}\| M^n < 1.$$

What we want to do, then, is to find a change of coordinates which eliminates all terms up to some given order in x for all y simultaneously. In fact, we will do this only locally, i.e. we will find a coordinate change which eliminates all terms up to a given order in x for all y in the unit ball. After making this coordinate change, we again multiply the non-linear terms by a cut-off function to make them vanish exactly for y outside the unit ball.

The procedure for finding the desired coordinate change is very much like the procedure by which we constructed finite-order linearizations in §2, except that y is carried along as a parameter. We look for a coordinate change of the form

$$\Phi(x, y) = (\phi(x, y), y) \quad \text{where } \phi(0, y) = 0 \text{ for all } y.$$

A straightforward computation shows that, if Φ has this form and if we write

$$\Phi f \Phi^{-1}(x, y) = (\Lambda_s x + \hat{f}_s(x, y), \Lambda_u y + \hat{f}_u(x, y)),$$

then

$$D_x^j \hat{f}_s(0, y) = 0 \quad \text{for } 1 \leq j < N$$

if and only if

$$(4.3) \quad D_x^j (\phi \circ f)(0, y) = \Lambda_s D_x^j \phi(0, y) \quad \text{for } 1 \leq j < N.$$

We are going to use Proposition 3.1 of the preceding section to show that this set of equations has a solution.

As usual, we will solve the equation (4.3) successively for $j = 1, \dots, N - 1$, and, also as usual, we start by noting that

$$D_x^j \phi \circ f(x, y) = (D_x^j \phi)(f(x, y)) (\Lambda_s + D_x f_s(x, y))^j + \text{lower order terms},$$

where the lower order terms involve x derivatives of ϕ only of order $j - 1$ or smaller. We also note that the mapping

$$y \mapsto \Lambda_u y + f_u(0, y) = \mathcal{E}_u\text{-component of } f(0, y)$$

is invertible, and we write F for its inverse. If the non-linear terms in f are small enough, F is globally contractive. We also write $\phi_j(y)$ for $D_x^j \phi(0, y)$ and $\Lambda_s(y)$ for $\Lambda_s + D_x f_s(0, y)$. With this notation, we can write the j -th equation of (4.3) as

$$\phi_j(F^{-1}(y))(\Lambda_s(y))^j - \Lambda_s \phi_j(y) = a_j(y),$$

where $a_j(y)$ is "known" once the earlier equations have been solved. Replacing y by $F(y)$ and multiplying on the right by $(\Lambda_s(F(y)))^{-j}$, we can rewrite this equation as

$$\phi_j(y) - \Lambda_s \phi_j(F(y))(\Lambda_s(F(y)))^{-j} = b_j(y),$$

with b_j also known.

This equation has the form considered in Proposition 3.1 with

- (a) $\mathcal{E} = \mathcal{E}_u$, $f = F$,
- (b) $\mathcal{F}$ the space of symmetric j -linear mappings from $\mathcal{E}_s \times \dots \times \mathcal{E}_s$ to $\mathcal{E}_s$,
- (c) $L(y)$ the operator defined by

$$L(y)\phi(x_1, \dots, x_j) = \Lambda_s \phi(\Lambda_s(F(y))^{-1}x_1, \dots, \Lambda_s(F(y))^{-1}x_j).$$

In particular, the spectrum of $L(0)$ is

$$\sigma(\Lambda_s)^{-1} \sigma(\Lambda_s)^j.$$

Since

$$DF(0) = \Lambda_u^{-1},$$

the non-resonance condition (2) of Proposition 3.1, which we can write as

$$1 \notin \sigma(L(0)) \sigma(DF(0))^k$$

is a consequence of the ordinary non-resonance condition (of higher order)

$$\sigma(\Lambda) \cap \sigma(\Lambda)^{j+k} = \emptyset,$$

provided $j+k \geq 2$. For $j \geq 2$, this holds for all k , and in this case we apply Proposition 3.1 with $n_0 = 0$. For $j = 1$, we need to do a little more: We want ϕ_1 to satisfy $\phi_1(0) = 1$ so we apply Proposition 3.1 to the equation satisfied by $\phi_1(y) - 1$, with $n_0 = 1$. The smoothness conditions on a required in Proposition 3.1 can be guaranteed inductively by requiring enough smoothness of f .

Thus, given any N , there exists $\overline{N}$ depending only on N , $\rho(\Lambda_u)$, and $\rho(\Lambda_s^{-1})$ such that, if f_s, f_u are of class $C^{\overline{N}}$, and if

$$\sigma(\Lambda) \cap \sigma(\Lambda)^j = \emptyset \quad \text{for } 2 \leq j < \overline{N},$$

then there exist $\phi_1, \dots, \phi_N$ of class C^N such that (4.3) holds. Once this has been done, we construct the coordinate change Φ as described above, cutting off in x if necessary to make it globally invertible. Conjugating with Φ gives a mapping $\hat{f}$ such that

$$D_x^j(\hat{f}_s)(0, y) = 0 \quad \text{for all } y, \text{ for } 1 \leq j < N.$$

We do all this only for $\|y\| < 1$, but, after making the coordinate change, we multiply the non-linear terms in $\hat{f}$ by a smooth cutoff function which is identically one on a neighborhood of zero; the argument at the beginning of this section can then be used to conclude that this re-cutoff $\hat{f}$ admits a smooth linearization and hence that f itself admits a smooth linearization. Thus we have proved:

THEOREM 4.1 (STERNBERG LINEARIZATION THEOREM). *Let $0 < \alpha < \beta < 1$, and let $n \geq 1$ be a positive integer. There is an integer N depending on α, β , and n such that, if f is a mapping of class C^N ; if z_0 is a fixed point of f ; if the spectrum of $Df(z_0)$ is contained in the union of the two annuli*

$$\{\alpha \leq |\lambda| \leq \beta\} \cup \{\alpha^{-1} \leq |\lambda| \leq \beta^{-1}\}$$

and if the non-resonance condition

$$\sigma(Df(z_0)) \cap \sigma(Df(z_0))^j = \emptyset \quad \text{for } 2 \leq j < N$$

*Cⁿ linearisierung
für f bei dyperiod. Fixpunkt*

holds, then f admits a C^n linearization at z_0 .

It would not be difficult to make an explicit estimate for N . Since N does not depend on the dimension of $\mathcal{E}$, it seems likely that the theorem remains true for Banach spaces, although the proof we have given does not work in that generality. (A Banach space version would, however, be much less satisfactory than the finite-dimensional version since the non-resonance condition is no longer "generic" if $Df(z_0)$ has continuous spectrum.)

It should also be noted that Theorem 4.1 has no uniqueness statement. In fact, smooth linearizations are not unique in the general hyperbolic case, even in the absence of resonances, or, more precisely, smooth *local* linearizations are not *locally* unique. To see why this is, we first note that f has a locally unique normalized smooth linearization at z_0 if and only if the only smooth Φ with $\Phi(0) = 0$ and $D\Phi(0) = 1$ which commutes near zero with $\Lambda = Df(z_0)$ is the identity.

If Ψ is one normalized smooth linearization, and if Φ commutes with Λ , then $\Phi \circ \Psi$ is another smooth linearization; conversely, if Ψ_1 and Ψ_2 are normalized smooth linearizations, then $\Phi = \Psi_1 \Psi_2^{-1}$ commutes with Λ .

We want to show therefore that there exist non-trivial Φ 's which commute with Λ on a neighborhood of 0. We can re-express the commutation condition by writing $\Phi(z) = z + \phi(z)$ with ϕ vanishing to second order at 0. Commutativity of Φ with Λ on a neighborhood of 0 is equivalent to

$$\phi(z) = \Lambda^{-1} \phi(\Lambda z)$$

on that neighborhood. What we want to do, therefore, is to show that this equation has non-trivial smooth solutions.

We can exhibit such solutions using the techniques developed above. Let, for example, a be any C^∞ mapping of $\mathcal{E}_s \oplus \mathcal{E}_u$ into itself, with compact support, vanishing identically on a neighborhood of 0 and vanishing together with all its derivatives everywhere on $\mathcal{E}_u$. By what we have shown above,

$$\phi(z) \equiv \sum_{j=0}^{\infty} \Lambda^{-j} a(\Lambda^j z)$$

converges in C^n for all n ; hence, is of class C^∞ , and satisfies

$$\phi(z) - \Lambda^{-1} \phi(\Lambda z) = a(z).$$

Since, by assumption, a vanishes on a neighborhood of 0, this equation means that ϕ commutes with Λ on a neighborhood of 0.

We still have to show that a can be chosen in such a way as to make ϕ not vanish identically on any neighborhood of 0. This is easy to do if, for example, Λ is diagonalizable with positive eigenvalues: In this case, we work in a basis of eigenvectors for Λ , and we choose a so that each component is *strictly positive* everywhere in

$$\{1 \leq \|z\| \leq \|\Lambda\|\}$$

except on $\mathcal{E}_u$. This is certainly possible to do. Then for every z which is in the unit ball but not in either $\mathcal{E}_s$ or in $\mathcal{E}_u$ there is at least one j such that $a(\Lambda^j z) \neq 0$. But since there are no cancellations in the sum defining ϕ , this means that $\phi(z)$ is non-zero everywhere on the unit ball except on $\mathcal{E}_s$ and $\mathcal{E}_u$. It does, nevertheless, vanish to infinite order at zero.

§5. The C^∞ case. Let f be a C^∞ mapping with fixed point z_0 , and suppose that there are no resonances at all, i.e., that

$$\sigma(Df(z_0)) \cap \sigma(Df(z_0))^j = \emptyset \quad \text{for all } j \geq 2.$$

Then the theorem of the preceding section says that f admits a C^n linearization for each n , but, because of the lack of uniqueness, does not directly imply that there is a C^∞ linearization. We will prove here that a C^∞ linearization does nevertheless exist.

The argument is essentially that of the preceding section, with one additional twist. The essential new step is to find a change of coordinates which makes $f_s(x, y)$ vanish to infinite order as x goes to zero, uniformly in y . Once this has been done, the argument of the preceding section gives an infinitely differentiable ψ_s . It is also clear from the preceding section how to construct the change of coordinates as a formal power series in x with coefficients depending on y . It looks extremely unlikely that this power series converges, especially if f is not analytic. The new twist is the remark that it is not necessary to show that the series converges, but only to find a C^∞ function of (x, y) whose derivatives with respect to x at 0 are as specified. Such a function, it turns out, can be shown to exist without knowing anything at all about the rate of growth of these derivatives with order. What we need to prove is an easy generalization of the standard but surprising fact the every formal power series is the Taylor series of a C^∞ function. We will start by proving this fact and then discuss the generalization we really need.

It will be convenient to use here the multi-index notation. We fix a positive integer n ; a *multi-index* α will mean an order n -tuple $(\alpha_1, \alpha_2, \dots, \alpha_n)$ of non-negative integers. If $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (or $\mathbb{C}^n$), x^α means $\prod_{i=1}^n x_i^{\alpha_i}$, and D^α means the differential operator $D_1^{\alpha_1} \dots D_n^{\alpha_n}$. We write $|\alpha|$ for $\sum_{i=1}^n \alpha_i$ and $\alpha!$ for $\prod_{i=1}^n \alpha_i!$.

PROPOSITION 5.1. Let c_α be an arbitrary family of real numbers indexed by the multi-indices. Then there is a C^∞ real-valued function ψ on $\mathbb{R}^n$ such that

$$(D^\alpha \psi)(0) = \alpha! c_\alpha \quad \text{for all } \alpha,$$

(i.e., the formal power series $\sum_\alpha c_\alpha x^\alpha$ is the Taylor series for ψ .)

PROOF: The idea is to cut off the Taylor series vigorously enough to make it converge, without spoiling the derivatives at zero. Let χ be a C^∞ real-valued function of compact support on $\mathbb{R}^n$ which is identically one on some neighborhood of zero. We are going to show that we can take ψ to be a function of the form

$$\psi(x) = \sum_\alpha c_\alpha x^\alpha \chi(\eta_\alpha x),$$

where the η_α are positive real numbers increasing rapidly enough as $|\alpha|$ goes to infinity. Just from $\eta_\alpha \rightarrow \infty$ it follows that for each non-zero x there are only finitely many non-zero terms in the series defining $\psi(x)$, so the series certainly converges everywhere. For similar reasons the sum is C^∞ everywhere except perhaps at zero. Some care will be required, however, to guarantee that it is also C^∞ at zero and that its derivatives there are what they appear to be. For this, it will suffice to arrange things so that

$$\sum_{\alpha} c_{\alpha} D^{\beta} (x^{\alpha} \chi(\eta_{\alpha} x))$$

converges uniformly in x for any multi-index β .

Once β has been fixed, we can limit the sum to those terms with $|\alpha| > |\beta|$. Also,

$$D^{\beta} (x^{\alpha} \chi(\eta_{\alpha} x)) = \eta_{\alpha}^{|\beta| - |\alpha|} D^{\beta} (y^{\alpha} \chi(y)) \Big|_{y = \eta_{\alpha} x}.$$

Now let

$$M_{\alpha} \equiv \sup_{|\beta| < |\alpha|} \sup_{y \in \mathbb{R}^n} |D^{\beta} y^{\alpha} \chi(y)|,$$

and, for j a positive integer, write N_j for the number of multi-indices α with $|\alpha| = j$. Now choose $\eta_{\alpha} \geq |\alpha|$ so that

$$\frac{|c_{\alpha}| M_{\alpha}}{\eta_{\alpha}} \leq \frac{1}{2^{|\alpha|} N_{|\alpha|}},$$

i.e., so that

$$\eta_{\alpha} \geq 2^{|\alpha|} |c_{\alpha}| M_{\alpha} N_{|\alpha|}.$$

Then, for $|\alpha| \geq |\beta|$,

$$\begin{aligned} |c_{\alpha} D^{\beta} (x^{\alpha} \chi(\eta_{\alpha} x))| &\leq \frac{|c_{\alpha}| M_{\alpha}}{\eta_{\alpha}^{|\alpha| - |\beta|}} \\ &\leq \frac{|c_{\alpha}| M_{\alpha}}{\eta_{\alpha}} \leq \frac{1}{2^{|\alpha|} N_{|\alpha|}}. \end{aligned}$$

Since

$$\sum_{|\alpha| > |\beta|} \frac{1}{2^{|\alpha|} N_{|\alpha|}} = \sum_{j > |\beta|} 2^{-j} < \infty,$$

the desired convergence follows and the proposition is proved.

The result we really need is the following generalization:

PROPOSITION 5.2. *Let $c_{\alpha}(y)$ be an arbitrary family of real-valued C^∞ functions on $\mathbb{R}^m$ indexed by the multi-indices. Then there is a real-valued C^∞ function ψ on $\mathbb{R}^{n+m}$ such that*

$$(D_x^{\alpha} \psi)(0, y) = \alpha! c_{\alpha}(y) \quad \text{for all } \alpha, y.$$

PROOF: It is tempting to try to repeat the proof of the preceding proposition but making the η_α depend on y . This has the disadvantage of making it necessary to worry about derivatives of the η_α with respect to y . It works better to take the η_α constant, as follows: It will suffice to show that the series

$$\sum_{\alpha} D_y^{\beta_y} c_{\alpha}(y) D_x^{\beta_x} (x^{\alpha} \chi(\eta_{\alpha} x))$$

converges uniformly for $x \in \mathbb{R}^n$ and for $y \in \mathbb{R}^m$ with $\|y\| \leq k$ for all pairs β_x, β_y of multi-indices and all positive integers k . Without affecting the convergence, we can leave out of the sum the finite set of terms with $|\alpha| \leq |\beta_x|$ or with $|\alpha| \leq |\beta_y|$ or with $|\alpha| \leq k$. We then choose η_{α} so that

$$\frac{|D_y^{\beta_y} c_{\alpha}(y)| M_{\alpha}}{\eta_{\alpha}} \leq \frac{1}{2^{|\alpha|} N_{|\alpha|}},$$

for all β_y with $|\beta_y| < |\alpha|$ and all y with $\|y\| < |\alpha|$. The proof of uniform convergence then goes just as before.

This proposition, and the argument sketched at the beginning of this section, prove:

THEOREM 5.3. *Let z_0 be a hyperbolic fixed point for the C^{∞} mapping f such that*

$$\sigma(Df(z_0)) \cap \sigma(Df(z_0))^j = \emptyset \quad \text{for all } j \geq 2.$$

Then f admits a C^{∞} linearization at z_0 .

§1. Persistence of fixed points. We are going to develop here an argument, based on the Implicit Function Theorem, which says that fixed points of mappings depending on a parameter cannot easily disappear as the parameter changes. Let f_μ be a mapping depending on a parameter; for definiteness, let us assume that $(x, \mu) \mapsto f_\mu(x)$ is jointly smooth in x and μ for $\mu_0 < \mu < \mu_1$ and for $x \in \mathcal{E}$. Here $\mathcal{E}$ can be any manifold, but we will write the formulas assuming it is an open set in a Banach space. Suppose that, for some $\bar{\mu} \in (\mu_0, \mu_1)$, $f_{\bar{\mu}}$ has a fixed point $\bar{x}$ such that 1 is not in the spectrum of $Df_{\bar{\mu}}(\bar{x})$. Then zero is not in the spectrum of the derivative (with respect to x) of $f_{\bar{\mu}}(x) - x$ at $\bar{x}$ and so, by the Implicit Function Theorem, there is a smooth function x_μ defined for μ near $\bar{\mu}$ such that $f_\mu(x_\mu) - x_\mu = 0$ and $x_{\bar{\mu}} = \bar{x}$, i.e., the fixed point $\bar{x}$ lies on a smooth family of fixed points depending on the parameter μ . Furthermore, the uniqueness part of the Implicit Function Theorem says that, if $\mathcal{U}$ is a sufficiently small neighborhood of $\bar{x}$ in $\mathcal{E}$, and if μ is near enough to $\bar{\mu}$, then f_μ has no fixed point in $\mathcal{U}$ other than x_μ .

We next make this argument a little more global. By Zorn's Lemma, there exists an open parameter interval (μ_-, μ_+) and continuous parameter-dependent fixed point x_μ defined on (μ_-, μ_+) which cannot be extended to any strictly larger open interval. We want to see what happens at the ends of this maximal parameter interval, and for definiteness we will discuss the behavior as μ approaches μ_+ . There are a certain number of trivial possibilities: It may be that $\mu_+ = \mu_1$, i.e., that the interval of definition of x_μ is limited by the limits on the interval of definition of f_μ . It may also be that, as μ approaches μ_+ , x_μ approaches infinity, or the boundary of the domain of definition (in x) of the f_μ , or, in an infinite-dimensional space, that it oscillates in such a complicated way that no subsequence converges. If, however, there is some sequence μ_n increasing to μ_+ such that x_{μ_n} converges to a limit x_+ in the interior of the domain of f_{μ_+} , then necessarily 1 is in the spectrum of $Df_{\mu_+}(x_+)$; if not, then the Implicit Function Theorem would imply that there is a smooth curve of fixed points parametrized by μ and passing through x_+ ; the uniqueness statement in the Implicit Function Theorem would imply that the x_μ for μ near μ_+ coincides with that part of that curve with $\mu < \mu_+$; and the part of the curve lying on the other side of x_+ would provide an extension of x_μ which we have assumed not to exist. We can sum up the result of this argument by saying that the only non-trivial way in which a fixed point of a mapping depending on a parameter can "disappear" as the parameter is changed is by passing first through a situation where the spectrum of the derivative of the mapping at the fixed point contains
1.

By applying the Implicit Function Theorem to the Poincaré mapping associated with a periodic orbit, one gets a persistency result for periodic orbits: If a flow f_μ^t depending on a parameter μ admits a periodic orbit for $\mu = \bar{\mu}$, and if one is not in the spectrum of the derivative of the associated Poincaré mapping, then the given periodic orbit lies in a smooth μ -dependent family of periodic orbits defined for μ near $\bar{\mu}$. Globalization is not so simple as for fixed points—among the "trivial" ways in which a parameter-

dependent periodic orbit can disappear are that the orbit becomes unbounded, or that the size of the orbit shrinks to zero (so that the orbit becomes a stationary solution) or that some point on the orbit approaches a stationary solution without the size of the orbit going to zero (in which case the period goes to infinity.)

§2. The saddle-node bifurcation in one dimension. The considerations of the preceding section motivate a closer examination of what happens when a mapping belonging to a smooth one-parameter family has a fixed point at which its derivative has spectrum containing one. In general, the behavior can be very complicated, but it has been discovered that what “usually” happens is quite simple and nearly always the same. That is: There is a single *generic* behavior, which happens in all but exceptional situations. The analysis of this generic behavior goes in two steps: There is an argument which shows what typically happens for mappings in one dimension and a separate argument showing that the behavior of higher-dimensional mappings, properly understood, is the same as that of one-dimensional mappings. Most of the insight comes from the analysis of the one-dimensional case which will be given in this section. The extension to higher dimensions will be done using the Center Manifold Theorem; that argument will be given in the following section.

We consider a mapping of $\mathbb{R}$ to itself depending on a real parameter μ ; we will write this mapping as $f(\mu, x)$ rather than as $f_\mu(x)$. We assume that this mapping is jointly C^2 in (μ, x) ; that it has 0 as a fixed point for $\mu = 0$; and that $D_x f(0, 0) = 1$. We also assume

$$(2.1) \quad D_\mu f(0, 0) \neq 0.$$

As we will soon see, this assumption leads to a very simple picture of the behavior of f near 0 for μ near 0. It is clear that this condition does not always hold, but it seems very plausible that it will *normally* hold: As we follow the parameter-dependent fixed point, we would expect to pass through at most a discrete set of parameter values at which $D_\mu f$ vanishes, and there is no reason to expect that one of them will coincide exactly with a place where $D_x f$ is equal to one. We thus refer to (2.1) as a *genericity* or *non-degeneracy* assumption.

Given the assumption (2.1) and given that

$$f(\mu, x) - x \text{ vanishes for } x = 0 \text{ and } \mu = 0$$

(i.e., that 0 is a fixed point for f at $\mu = 0$), we can use the Implicit Function Theorem to solve the fixed point equation

$$f(\mu, x) - x = 0$$

for μ as a function of x . That is, there is a continuously differentiable function $\mu(x)$, defined on a neighborhood of zero and vanishing at zero, such that x is a fixed point for f at $\mu = \mu(x)$. The usual formula for the derivative of an implicitly defined function gives

$$\frac{d\mu(x)}{dx} = \frac{D_x f - 1}{D_\mu f}.$$

Since we are assuming that $D_x f(0, 0) = 1$, it follows that $d\mu(x)/dx$ vanishes at $x = 0$. Thus, if $d^2\mu(x)/dx^2$ does not vanish at $x = 0$, $\mu(x)$ has a local minimum or maximum there. By a straightforward computation

$$\frac{d^2\mu(x)}{dx^2} = \frac{\left(D_x^2 f + D_\mu D_x f \frac{d\mu}{dx}\right) D_\mu f - (D_x f - 1) D_\mu^2 f \frac{d\mu}{dx}}{(D_\mu f)^2},$$

which reduces at $x = 0$ to $D_x^2 f(0, 0)/D_\mu f(0, 0)$. We thus make a second genericity assumption, that

$$(2.2) \quad D_x^2 f(0, 0) \neq 0.$$

It then follows that either $\mu(x) \geq 0$ for all small x or $\mu(x) \leq 0$ for all small x . These two situations are essentially the same; one can be converted into the other by using $-\mu$ rather than μ as the parameter. We can therefore assume without losing any essential generality that $\mu(x) > 0$ for small non-zero x .

The above description, with μ expressed in terms of x , is not what we really want; rather, we want to see how the behavior of f near the origin changes as we vary μ . This, however, is not difficult to sort out from what we now know. For small positive μ , f has two fixed points $x_\pm(\mu)$ with

$$x_\pm \approx \pm C\sqrt{\mu}, \quad \text{where } C \text{ is a non-zero constant.}$$

Furthermore, one of these fixed points is attracting and the other repelling. To see this, write

$$D_x f(\mu, x_\pm) = D_x f(\mu, 0) + \int_0^{x_\pm} D_x^2 f(\mu, x) dx.$$

If μ is near enough to zero, the $D_x^2 f(\mu, x)$ is approximately equal to $D_x^2 f(0, 0)$, which we have assumed to be non-zero, and the length of the interval of integration is $O(\sqrt{\mu})$. On the other hand, $D_x f(\mu, 0) = 1 + O(\mu)$. Hence,

$$D_x f(\mu, x_\pm(\mu)) = 1 \pm C_1 f(0, 0)\sqrt{\mu} + o(\sqrt{\mu}),$$

where we have written C_1 for $CD_x^2 f(0, 0)$. If, for example, C_1 is positive, then the derivative at x_+ is slightly larger than 1 so x_+ is repelling and the derivative at x_- is slightly smaller than 1 so x_- is attracting; if C_1 is negative, it is the other way around.

For small negative μ , on the other hand, f has no fixed point at all near the origin. Intuitively, the two parameter-dependent fixed points collide at the origin when $\mu = 0$, and "annihilate each other." A very simple example of this bifurcation is provided by the family

$$f(\mu, x) = x - \mu + x^2.$$

We need to go back and look again at our genericity assumptions. The heuristic argument by which we justified them can be criticized because it could equally well

be applied to justify adding as a further genericity assumption the requirement that $d\mu(x)/dx$ is non-zero at $x = 0$, and this, we know, does not *ever* hold when the initial assumption (2.1) does. The justification can be improved by noting that a family in which (2.1) or (2.2) (or both) do not hold can be perturbed arbitrarily little to produce one in which both nondegeneracy conditions do hold. If, for example, (2.1) fails for $f(\mu, x)$, it holds for $f(\mu, x) + \epsilon\mu$ for arbitrarily small nonzero ϵ , and the perturbed family still has the origin as a fixed point for $\mu = 0$ and x -derivative equal to one there. Similarly, if (2.2) does not hold, it can be made to hold by adding ϵx^2 for arbitrarily small nonzero ϵ , again without spoiling the other conditions.

§3. The Ruelle-Takens center manifold argument. Let f_μ be a mapping depending on a parameter; assume that $f_\mu(z)$ is of class C^r , $1 \leq r < \infty$ jointly in μ and z . Suppose that $f_0(0) = 0$ and that $\Lambda \equiv Df_0(0)$ has non-trivial center subspace. (To be precise, in the infinite dimensional case, we mean that the intersection of the spectrum of Λ with the unit circle is non-empty and separated from the rest of the spectrum.) Split the space as usual into

$$\mathcal{E}_s \oplus \mathcal{E}_c \oplus \mathcal{E}_u.$$

Assume that there exists a real-valued function of class C^r on $\mathcal{E}_c$ which vanishes outside the unit ball but is not identically zero. By applying the Center Manifold Theorem to the augmented mapping

$$(\mu, z) \mapsto (\mu, f_\mu(z)) \quad \text{on } \mathbf{R} \oplus \mathcal{E},$$

one proves that there exist open neighborhoods $(-\epsilon, \epsilon)$ of 0 in $\mathbf{R}$, $\mathcal{U}_c$ of 0 in $\mathcal{E}_c$, and $\mathcal{U}_{su}$ of 0 in $\mathcal{E}_s \oplus \mathcal{E}_u$, and a function of class C^r

$$(\mu, x) \mapsto w_\mu(x) \quad \text{from } (-\epsilon, \epsilon) \times \mathcal{U}_c \text{ into } \mathcal{U}_{su},$$

whose graph is locally invariant in the sense that, if $|\mu| < \epsilon$, $z \in W_\mu$, and $f_\mu(z) \in \mathcal{U}$, then $f_\mu(z) \in W_\mu$. (We have used here two new pieces of notation: W_μ for the graph of $x \mapsto w_\mu(x)$ and $\mathcal{U}$ for $\mathcal{U}_c \times \mathcal{U}_{su}$.) Furthermore

- (a) If, for some z in $\mathcal{U}$, the forward orbit $f_\mu^n(z)$, remains in $\mathcal{U}$ for all $n > 0$ then it converges to W_μ . Similarly, for f_μ invertible, if the backward orbit $f_\mu^{-n}(z)$ remains in $\mathcal{U}$ for all $n > 0$ then it converges to W_μ . (For f_μ not necessarily invertible, the statement is that *any* backward orbit in $\mathcal{U}$ defined for all n must converge to W_μ .)
- (b) Any z in $\mathcal{U}$ whose full orbit—forward and backward—remains in $\mathcal{U}$ must be in W_μ . (For f_μ non-invertible: Any z whose forward orbit remains in $\mathcal{U}$ and which admits an infinite backward orbit remaining in $\mathcal{U}$ must be in W_μ .) In particular, any fixed point for f_μ which is in $\mathcal{U}$ must be in W_μ ; similarly for any periodic cycle all points of which are in $\mathcal{U}$.
- (c) If z_0 is a fixed point for f_μ , then the spectrum of $Df_\mu(z_0)$ (in the full ambient space $\mathcal{E}$) is the union of the spectrum of $Df_\mu(z_0)$ in W_μ and the spectrum of

an operator on $\mathcal{E}_s \oplus \mathcal{E}_u$ which is near (in the norm topology) to the restriction of $Df_0(0)$ to $\mathcal{E}_s \oplus \mathcal{E}_u$. If $\mathcal{E}_u$ is zero-dimensional and if the fixed point z_0 is attracting in the submanifold W_μ , it is attracting in $\mathcal{E}$. (This last statement follows immediately from the preceding one if the spectrum of $Df_\mu(z_0)$ in the submanifold is strictly inside the unit circle. It is true in full generality—i.e., allowing eigenvalues on the unit circle—by a slightly more delicate argument.) Similar statements hold for periodic points.

We can use these results to extend the analysis of the preceding section to mappings in more than one dimension. Let f_μ be a mapping depending smoothly on a parameter, and assume that 0 is a fixed point for $\mu = 0$ and that the spectrum of $Df_0(0)$ contains 1. We make a first non-degeneracy assumption that 1 is an isolated point of the spectrum of $Df_0(0)$, that the spectral multiplicity of 1 is 1, and that no other point of the unit circle is in the spectrum. (These conditions can always be made to hold by perturbing f slightly in finite dimensions, or if Df_μ is compact, but not in the general infinite-dimensional situation.) We can then apply the analysis outlined above, with $\mathcal{E}_c$ one-dimensional. Each invariant manifold W_μ will be one-dimensional, and, by parametrizing it in any reasonable way, we can think of the restriction of f_μ to W_μ as a mapping of $\mathbb{R}$ to itself, depending smoothly on the parameter μ . This is the situation we analyzed in the preceding section; subject to appropriate non-degeneracy hypotheses, what happens inside W_μ as μ passes through zero in the appropriate direction is that two fixed points collide and disappear; the separation of the fixed points is $O(|\mu|^{1/2})$. Evidently, a fixed point for the restriction of f_μ to W_μ is also a fixed point in the ambient space, and furthermore f_μ has no fixed points near 0 for μ near 0 other than those in W_μ . Thus, in particular, the fixed points really disappear at $\mu = 0$ and don't simply move out of W_μ . Within W_μ , one of the fixed points is attracting and the other repelling. In the ambient space, the situation is a little more complicated, but the fixed point which is attracting in W_μ is attracting in the ambient space for sufficiently small $|\mu|$ if $Df_0(0)$ has no spectrum outside the unit disk.

§4. The period-doubling bifurcation. In the two preceding sections we studied the normal mechanism by which fixed points for maps can disappear as a parameter in the map changes. In this section we investigate the related question of what happens when an attracting fixed point for a map ceases to be attracting as a parameter changes. Suppose f_μ has, for some parameter value $\bar{\mu}$, a fixed point $\bar{x}$ such that the spectral radius of $Df_{\bar{\mu}}(\bar{x})$ is strictly smaller than one. Since the spectrum of $Df_{\bar{\mu}}(\bar{x})$ does not contain one, the fixed point $\bar{x}$ lies on a smooth parametrized family of fixed points. As we move along this family, the fixed point cannot disappear as long as the spectrum stays in the open unit circle. We want to investigate what happens when part of the spectrum passes out of the unit circle. There are three non-degenerate possibilities: Either a single simple eigenvalue goes to 1 (in which case a saddle-node bifurcation as described in the preceding sections can be expected to occur with the consequence that the fixed point disappears), or a single simple eigenvalue goes through -1 , or a complex conjugate pair of complex eigenvalues goes to $\{e^{\pm i\theta}\}$ with θ not a multiple of π . In this section, we investigate the second of these possibilities.

As in the preceding section, we can use the Center Manifold Theorem to reduce to consideration of what happens in a one-dimensional invariant submanifold W_μ and hence to a mapping of $\mathbb{R}$ to itself. Thus, let f_μ be a mapping of $\mathbb{R}$ to itself depending smoothly on a parameter μ , and suppose that f_0 has 0 as a fixed point, with $Df_0(0) = -1$. By the Implicit Function Theorem, there is a smooth function $\mu \mapsto x_\mu$ defined on a neighborhood of 0 such that $x_0 = 0$ and $f_\mu(x_\mu) = x_\mu$ for each μ . By making a μ -dependent shift of origin, we can assume that $x_\mu \equiv 0$, i.e., that 0 is a fixed point for each f_μ .

Since $Df_0(0) = -1$, it follows immediately from the chain rule that $D(f_0 \circ f_0)(0) = +1$. The mapping $f_0 \circ f_0$ is not, however, a *typical* mapping with slope one at a fixed point; it necessarily has $D^2(f_0 \circ f_0)(0) = 0$. To see this, we calculate (for a general mapping f) that

$$D(f \circ f)(x) = (Df)(f(x))Df(x)$$

and

$$D^2(f \circ f)(x) = (D^2f)(f(x))(Df(x))^2 + Df(f(x))(D^2f)(x).$$

If $f(0) = 0$ we get

$$D^2(f \circ f)(0) = (D^2f)(0) [Df(0)^2 + Df(0)],$$

and the right-hand side vanishes if $Df(0) = -1$. On the other hand, differentiating again gives

$$D^3(f \circ f)(0) = (D^3f)(0) [(Df(0))^3 + Df(0)] + 3(D^2f(0))^2 Df(0),$$

which has no special reason to vanish when $Df(0) = -1$. We make the nondegeneracy assumption

$$(4.1) \quad D^3(f_0 \circ f_0)(0) \neq 0.$$

If we write, temporarily, $-b$ for $D^3(f_0 \circ f_0)(0)/3!$, then, for μ near 0, we have roughly

$$f_\mu \circ f_\mu(x) \approx (1 + \epsilon)x - bx^3$$

with ϵ small. For orientation, let us see what the situation would be if this approximate equation were exact. This we can do by very simple computations; we will then have to make a perturbation argument to show that the outcome is qualitatively the same for the original problem.

The fixed point equation for the simplified version of $f_\mu \circ f_\mu$ reads

$$(1 + \epsilon)x - bx^3 = x,$$

or (for non-zero fixed points)

$$\epsilon - bx^2 = 0.$$

For b positive, there are two solutions $\pm\sqrt{\epsilon/b}$ for ϵ positive and none for ϵ negative; for b negative, there are two solutions for ϵ negative and none for ϵ positive. These solutions are fixed points for $f_\mu \circ f_\mu$, so they may be either fixed points or periodic points of period 2 for f itself. For small μ , and small nonzero x , the sign of $f_\mu(x)$ is the opposite of that of x itself, so f_μ cannot have any nonzero fixed points near zero. Thus the two non-zero fixed points for $f_\mu \circ f_\mu$ must be the two points of a period 2 cycle for f_μ itself.

To determine whether this cycle is attracting or not, we compute

$$D(f_\mu \circ f_\mu)(x) = 1 + \epsilon + 3bx^2 = 1 - 2\epsilon \quad \text{at the fixed points.}$$

If b is positive, then, the period 2 cycle, when it exists, is attracting. But positive b also corresponds to the situation where the parameter values for which the period 2 cycle exists are those for which the fixed point at zero is repelling. The general picture is therefore that, when the fixed point at zero passes from being attracting to being repelling, it “emits” an attracting cycle of period 2. We refer to this occurrence as a *normal period-doubling bifurcation*. (Another common name is *flip bifurcation*.)

By contrast, for negative b the period 2 cycle is repelling. Negative b also corresponds to the situation where the fixed point at zero is attracting when the period 2 cycle exists. The picture here is that, when the fixed point at zero passes from being attracting to being repelling, it “absorbs” a repelling cycle of period 2. This is called an *inverted period-doubling bifurcation*.

We will now give a more careful argument to show that the situation for a “general” one-parameter family with $Df_0(0) = -1$ and satisfying the non-degeneracy condition (4.1) is qualitatively the same as that we have found in the simplified example. Our regularity assumption will be that $f_\mu(x)$ is of class C^n , $n \geq 3$, in (μ, x) . We will also assume for simplicity here—although we will drop this restriction later—that $D_\mu f'_\mu(0)$ is non-zero at $\mu = 0$, i.e., that the slope of f_μ at the fixed point is changing at a non-zero rate when it passes through -1 . In this case, it follows from the Inverse Function Theorem that we can make a C^n change of parameter so that

$$(Df)_\mu(0) = -\sqrt{1+\mu}$$

and we will from now on assume that the parameter is chosen in this way. Once this is done, we have

$$D_x^2(f_\mu \circ f_\mu)(0) = (D_x^2 f_\mu)(0)[1 + \mu - \sqrt{1+\mu}] \equiv 2\mu\alpha(\mu),$$

where $\alpha(\mu)$ is of class C^{n-2} .

We now define a function $\beta(\mu, x)$ by

$$(4.2) \quad f_\mu \circ f_\mu(x) = (1 + \mu)x + \mu\alpha(\mu)x^2 - \beta(\mu, x)x^3 \quad \text{for } x \neq 0.$$

It is clear from the definition that β is of class C^n away from $x = 0$. Moreover, by the integral form for the remainder in Taylor's Theorem we can write

$$\beta(\mu, x) = -\frac{1}{2} \int_0^1 (1-t)^2 (D^3 f_\mu)(tx) dt,$$

which shows that $\beta(\mu, x)$ is of class C^{n-3} on a neighborhood of $(0, 0)$; we will see shortly that it is actually a little better than that.

A non-zero fixed point for $f_\mu \circ f_\mu$ is thus a solution of

$$(4.3) \quad \mu + \mu\alpha(\mu)x - \beta(\mu, x)x^2 = 0.$$

Our non-degeneracy assumption (4.1) reads

$$b \equiv \beta(0, 0) \neq 0.$$

The argument from here on depends on the sign of b , and we start by assuming

$$b > 0.$$

Under this assumption we want to show that (4.3) has two solutions near zero for small positive μ and none for small negative μ . We take up first the situation for negative μ . Rather than just proving that the fixed point equation has no solution near zero, we will show that for all small negative μ , and for $\mu = 0$, the fixed point at zero attracts every orbit starting in a sufficiently small *but μ -independent* neighborhood of itself. This will follow if we show that

$$\frac{|f_\mu(f_\mu(x))|}{|x|} < 1$$

for all sufficiently small (μ, x) with μ non-positive and x non-zero. By the formula (4.2), this is the same as showing

$$\mu(1 + \alpha(\mu)x) - \beta(\mu, x)x^2 < 0,$$

and this can be guaranteed by taking x, μ small enough so that $\alpha(\mu)x > -1$ and $\beta(\mu, x) > 0$.

To prove the existence of solutions for positive μ , we rewrite (4.3) as

$$x\sqrt{\beta(\sigma^2, x)} = \pm\sqrt{\mu}\sqrt{1 + \alpha(\sigma^2)x}.$$

We then write σ for $\pm\sqrt{\mu}$ and eliminate μ in favor of σ to get the equation to be solved into the form

$$(4.4) \quad x\sqrt{\beta(\sigma^2, x)} - \sigma\sqrt{1 + \alpha(\sigma^2)x} = 0,$$

where σ can have either sign. This equation has the solution $x = 0$ for $\sigma = 0$, and we want to use the Implicit Function Theorem to produce a solution for small non-zero σ . This is easy if $f_\mu(x)$ is of class C^n with $n \geq 4$. Under this assumption $\beta(\mu, x)$ is of class at least C^1 , and the derivative of the left-hand side with respect to x at $\sigma = 0$

is $\sqrt{\beta(0,0)}$ which is non-zero. Thus, there is a solution $X(\sigma)$ of class C^{n-3} in σ with $x(0) = 0$. A straightforward calculation shows that

$$\frac{dX}{d\sigma}(0) = [\beta(0,0)]^{-1/2} = b^{-1/2} \neq 0.$$

The points $X(\pm\sqrt{\mu})$ (which we will often write simply as $x_{\pm}$) are fixed points for $f_{\mu} \circ f_{\mu}$. As in the simplified example discussed above, f_{μ} cannot have any fixed points near zero other than zero itself, so the $x_{\pm}$ must be the points of a cycle of period 2. Furthermore

$$\begin{aligned} D(f_{\mu} \circ f_{\mu})(x_{\pm}) &= 1 + \mu + 2\mu\alpha(\mu)x_{\pm} - 3\beta(\mu, x_{\pm})x_{\pm}^2 - D_x\beta(\mu, x_{\pm})x_{\pm}^3 \\ &= 1 + \mu\alpha(\mu)x_{\pm} - 2\beta(\mu, x_{\pm})x_{\pm}^2 - D_x\beta(\mu, x_{\pm})x_{\pm}^3 \\ &= 1 - 2\mu + o(\mu), \end{aligned}$$

where we have used the equation (4.3) and the fact that $x_{\pm} \approx \sqrt{\mu/b}$. In particular, the cycle of period 2 is attracting, exactly in the simplified example.

The above was proved under the assumption that $f_{\mu}(x)$ is jointly at least C^4 . We can reduce the differentiability requirement to C^3 , and show that $X(\sigma)$ is of class C^{n-2} , if we can show that $x\sqrt{\beta(\mu, x)}$ is of class C^{n-2} in μ, x , with $n \geq 3$. For this it suffices to show that the first partial derivatives of $x\sqrt{\beta(\mu, x)}$, which are certainly defined (and even of class C^{n-1}) for $x \neq 0$ extend continuously through $x = 0$ and the extensions are of class C^{n-3} . Since we already know that β is of class C^{n-3} and non-zero at $(0,0)$, it suffices to show that $x\beta(\mu, x)$ is of class C^{n-2} .

For example:

$$D_x(x\sqrt{\beta(\mu, x)}) = \frac{x}{2} \frac{D_x\beta(\mu, x)}{\sqrt{\beta(\mu, x)}} + \sqrt{\beta(\mu, x)},$$

which is C^{n-3} at $x = 0$ if and only if $x D_x\beta(\mu, x)$ is, and this expression in turn is C^{n-3} at $x = 0$ if and only if $D_x(x\beta(\mu, x))$ is. The derivative with respect to μ can be treated similarly.

Using the integral formula for β and integration by parts:

$$x\beta(\mu, x) = \frac{1}{2} \int_0^1 (1-t)^2 x g'''(tx) dt = \int_0^1 (1-t)[g''(tx) - g''(0)] dt,$$

where we have written g for $f_{\mu} \circ f_{\mu}$; the final expression is manifestly of class C^{n-2} in (x, μ) . (To apply the Implicit Function Theorem, we also needed to know that the x derivative of $x\sqrt{\beta(\mu, x)}$ at $x = 0$, $\mu = 0$ is $\sqrt{\beta(0,0)}$, but this is an immediate consequence of the continuity of β .)

The regularity conditions can be weakened even a little more by dropping any differentiability assumptions on the parameter dependence; most of the above picture remains true just assuming that

$$\mu \mapsto f_{\mu}$$

is *continuous* into the Banach space of C^3 mappings. To use the Implicit Function Theorem to get a solution $X(\sigma)$ of (4.4) depending continuously on σ , we have only to show that the x derivative of the left-hand side exists, does not vanish at $x = 0$, $\sigma = 0$, and varies continuously with σ, x , and this holds under these weak regularity assumptions. From the equation itself, it follows that

$$\lim_{\sigma \rightarrow 0} \frac{x(\sigma)}{\sigma} = [\beta(0, 0)]^{-1/2},$$

i.e., that even under these hypotheses $X(\sigma)$ is differentiable at zero with strictly positive derivative.

We saw earlier that, for μ, x sufficiently small and $\mu \leq 0$, $f_\mu^n(x)$ converges to the fixed point at zero. We can derive a similarly complete picture of the local orbit structure of f_μ near the fixed point at zero for $\mu > 0$. What we want to show in this case is that there is a μ -independent neighborhood of 0 such that every orbit starting in it (except the trivial orbit of the fixed point at zero) converges to the cycle $\{x_\pm\}$. To see this, we start by noting that there is an $\eta > 0$, independent of μ , such that the left-hand side of (4.3) is strictly positive for $0 < x < x_+$ and strictly negative for $x_+ < x < \eta$, for all sufficiently small positive μ . Putting this inequality back into the equation for $f_\mu \circ f_\mu$ shows that $f_\mu \circ f_\mu(x) > x$ for $0 < x < x_+$ and $f_\mu \circ f_\mu(x) < x$ for $x_+ < x < \eta$. Thus, if x, μ are both small and strictly positive, $f_\mu^{2n}(x)$ must converge to x_+ as n goes to $+\infty$. The situation is analogous for small negative x . We can also see what the backwards orbits do. From the positivity considerations above, it follows easily that the backward orbit of any point in (x_-, x_+) converges to 0 and that there is a neighborhood of 0 from which backward orbits of all points outside $[x_-, x_+]$ eventually escape.

The analysis for $b < 0$ is, formally, almost exactly the same as for $b > 0$. Exactly as before, (4.3) has no solutions near zero for small positive μ , and, furthermore, the fixed point is "macroscopically repelling" in the sense that there is a neighborhood of 0 from which every orbit except that of the fixed point itself eventually escapes. To see that solutions do exist for small negative μ , write $\mu = -\sigma^2$ and reorganize as before; the equation analogous to (4.4) differs from (4.4) only by replacing β by $-\beta$ and σ^2 by $-\sigma^2$. From here on, the analysis goes just as before. Again, there is for small negative μ a cycle of period 2 whose points $x_\pm$ can be written as C^{n-2} functions of $\sqrt{-\mu}$ with derivative non-vanishing at zero. Again the slope of $f_\mu \circ f_\mu$ at $x_\pm$ is $1 - 2\mu + o(\mu)$, but this time μ is negative and so the periodic cycle is repelling. The regularity conditions can be weakened exactly as in the case $b > 0$. The behavior of a general orbit can be worked out as before: Every orbit—other than the fixed point itself—starting between x_+ and x_- converges as n goes to $+\infty$ to the fixed point at the origin and as n goes to $-\infty$ to the cycle $\{x_\pm\}$; an orbit starting outside of $[x_-, x_+]$ converges backwards to the cycle $\{x_\pm\}$ and is driven a macroscopic distance away from zero by forward iteration. It may be noted that the behavior of f_μ in this case as μ increases through zero is the same as behavior of f_μ^{-1} in the case $b > 0$ with μ decreasing through zero. This can also be seen directly: Replacing f_0 by f_0^{-1} replaces b by $-b$, so f_μ is a non-degenerate family undergoing a normal period-doubling bifurcation at $\mu = 0$ if and only if f_μ^{-1} is a non-degenerate family undergoing an inverted period-doubling bifurcation at $\mu = 0$.

By using the Center Manifold Theorem, the above analysis can be extended to mappings in more than one dimension. One way to express the results is as follows: For each μ near zero, let f_μ be a mapping of an open set in the Banach space $\mathcal{E}$ into $\mathcal{E}$. Suppose that $(\mu, x) \mapsto f_\mu(x)$ is of class C^n , $n \geq 3$. Suppose also that 0 is a fixed point for all f_μ ; that -1 is an isolated point of the spectrum of $Df_0(0)$ which is a simple eigenvalue; and that the rest of the spectrum of $Df_0(0)$ lies strictly inside the unit circle. By general spectral theory, the eigenvector of $Df(0)$ with eigenvalue -1 can be followed as the parameter μ is varied; we denote the corresponding eigenvalue by $-(1 + \lambda_\mu)$. We assume also that λ_μ is strictly increasing in μ near zero.

There is an analytic function b of the first three derivatives of f_0 at 0, not identically zero, such that:

- (a) (Normal period-doubling bifurcation) If $b > 0$, there is for small $\mu > 0$ an attracting cycle $x_\pm(\mu)$ of period 2 near the fixed point at zero. If we write $y_\pm(\mu)$ for $x_\pm(\mu)/\sqrt{\lambda}$, then the vector-valued functions $y_\pm(\mu)$ are continuous for $\mu > 0$ and approach non-zero limits as μ decreases to zero. If $d\lambda/d\mu > 0$, the $x_\pm(\mu)$ are C^{n-2} functions of $\sqrt{\mu}$. Every orbit starting near enough to 0 and not in its local stable manifold converges to the cycle $x_\pm(\mu)$. On the other hand, for $\mu \leq 0$, the fixed point $\bar{x}(\mu)$ attracts all nearby orbits; there is a (μ -independent) neighborhood of 0 such that, for all sufficiently small non-positive μ , $f_\mu^n(x)$ converges to 0 for all x in that neighborhood.
- (b) (Inverted period-doubling bifurcation) If $b < 0$, there is for $\mu < 0$ a cycle $x_\pm(\mu)$ of period 2. The dependence of the cycle on μ is as in the preceding case. The cycle is not attracting but is hyperbolic with one-dimensional unstable manifold. For small nonnegative μ , every orbit starting near enough to the fixed point $\bar{x}(\mu)$ and not on the local stable manifold of this fixed point is eventually pushed away from the fixed point; in particular, there are no other fixed or periodic points near $\bar{x}(\mu)$.

The requirement that the rest of the spectrum of $Df_0(0)$ is inside the unit disk is not necessary; as long as there is no spectrum on the unit circle other than the simple isolated eigenvalue -1 , we can conclude that there is a cycle of period 2 near the fixed point either just before or just after the eigenvalue passes through -1 . Furthermore, there is no other recurrence near the fixed point for either sign of μ . In this more general situation the fixed point will not be attracting in either of the two cases; normal and inverted bifurcations differ only in the dimension of the unstable manifold for the cycle.

The requirement that $f_\mu(x)$ be jointly C^3 in μ and x can be weakened to the condition that $\mu \mapsto f_\mu$ is continuous in the C^3 topology; proving this requires a version of the Center Manifold Theorem slightly different from the one we have proved.

It is also possible to say a bit more about the local structure of the dynamical system near the bifurcation point. For simplicity and definiteness we describe this only for a normal (non-inverted) bifurcation in the case where the spectrum of Df_0 is inside the unit circle except for the eigenvalue responsible for the bifurcation. In this case, for small positive μ , the fixed point at zero is hyperbolic with one-dimensional unstable manifold. The unstable manifold is a little piece of one-dimensional arc passing through

zero which terminates at the two points of the attracting cycle; the mapping flips this piece of arc around the fixed point at the origin. This can be seen by reduction to one dimension, where the global unstable manifold is the interval (x_-, x_+) .

It is sometimes useful to adopt a more global geometric view of what is going on in this and similar bifurcations. For simplicity, we describe this global view only in one dimension and only for the normal bifurcation. Consider, for concreteness, the Banach space of C^n ($n \geq 3$) mappings f for $(-1, 1)$ into $\mathbb{R}$ with $f(0) = 0$. Let Σ denote the set of f 's in this space with $f'(0) = -1$ and $(f \circ f)'''(0) < 0$. This is an open subset of a hyperplane; in particular, a smooth codimension-one submanifold of the Banach space of f 's. In this perspective, a one-parameter family of f 's is a curve in the Banach space, and a normal period-doubling bifurcation is what happens when a curve crosses the hypersurface Σ ; accordingly, one says that Σ is the *bifurcation surface* for the normal period-doubling bifurcation. A reworking of the analysis given above of what happens along curves, paying more attention to questions of uniformity, leads to the following result: For any f_0 in Σ , there is a neighborhood $\mathcal{U}$ of f_0 in the Banach space of f 's and a neighborhood $\mathcal{V}$ of the fixed point 0 in $\mathbb{R}$ such that:

- (a) If $f \in \mathcal{U}$ and $Df(0) \geq -1$ (" f is below or on Σ "), then for any $x \in \mathcal{V}$, $f^n(x)$ converges to the fixed point at zero as n goes to $+\infty$ and eventually gets outside $\mathcal{V}$ as n goes to $-\infty$.
- (b) If $f \in \mathcal{U}$ and $Df(0) < -1$ (" f is strictly above Σ "), then f admits an attracting cycle of period 2 $\{x_\pm\}$ in $\mathcal{V}$, with $x_- < 0 < x_+$. For any x in $\mathcal{V}$ other than 0, $f^n(x)$ converges to this cycle as n goes to $+\infty$. As n goes to $-\infty$, $f^n(x)$ converges to the fixed point at zero if $x_- < x < x_+$ and eventually gets out of $\mathcal{V}$ if $x < x_-$ or $x > x_+$. The magnitude of $x_\pm$ is $O(\sqrt{\lambda(f)})$ (where we have written $\lambda(f)$ for $-Df(0) - 1$). More precisely, the $x_\pm/\sqrt{\lambda(f)}$ extend to continuous nowhere vanishing functions of f on the subset of $\mathcal{U}$ where $\lambda(f) \geq 0$ (i.e., where $Df(0) \leq -1$).

This formulation loses some of the smoothness of the dependence of the periodic cycle on $\sqrt{\mu}$ which we proved in the analysis of one-parameter families. We can recover this smoothness, at the expense of a little care in formulating the result: A mapping f may be specified by giving the number $f'(0)$ and the function $f''(x)$, so we can identify the space of f 's with $\mathbb{R} \times C^{n-2}$; to simplify formulas, we take as the first coordinate

$$\lambda = -f'(0) - 1$$

rather than $f'(0)$ itself. The sharp smoothness assertion is that there is a real-valued C^{n-2} mapping $X(\sigma, g)$, defined on a neighborhood of $\{0\} \times C^{n-2}$, such that the pair of points $X(\pm\sqrt{\lambda}, g)$ constitute the attracting cycle of period 2 for the mapping f corresponding to (λ, g) ; furthermore $(D_\sigma X)(0, g)$ does not vanish anywhere. Intuitively: the $x_\pm$ depend in a C^{n-2} way on $\sqrt{\lambda}$ and the "other components" of f .

It is also possible to extend the analysis in another direction, restricting to one-parameter families where the eigenvalue passes through the value -1 at non-zero velocity but dropping the non-degeneracy condition $(f_0 \circ f_0)'''(0) \neq 0$. Again we will only

describe this extension for the bifurcation in one dimension; the extension to higher dimensions goes in the standard way. To start with, we argue on very general grounds that passage of $Df_\mu(0)$ through -1 produces cycles of period 2. Thus, suppose that f_μ is an continuous one parameter family of continuously differentiable mappings with $f_\mu(0) = 0$ and such that

$$Df_\mu(0) > -1 \quad \text{for } \mu = \mu_0 \quad \text{and} \quad Df_\mu(0) < -1 \quad \text{for } \mu = \mu_1.$$

By the Chain Rule and the definition of derivative, there is a neighborhood $\mathcal{V}$ of zero such that

$$\begin{aligned} \frac{f_\mu \circ f_\mu(x)}{x} &< 1 \quad \text{for all non-zero } x \in \mathcal{V}, \text{ when } \mu = \mu_0 \\ &> 1 \quad \text{for all non-zero } x \in \mathcal{V}, \text{ when } \mu = \mu_1 \end{aligned}$$

Thus, for each nonzero $x \in \mathcal{V}$, there must be at least one μ between μ_0 and μ_1 such that x is a fixed point for $f_\mu \circ f_\mu$. For the usual reason, this fixed point must be periodic of period 2 for f_μ . Intuitively, then, under these very general hypotheses there are at least continuously many cycles of period 2 associated with the passage of Df_μ through -1 . These cycles can be organized in many different ways. At one extreme, as we have already seen, in the non-degenerate case the cycles are arranged in (μ, x) -space so that is exactly one of them for each μ of one sign and none for each μ of the other sign. At the other extreme, if $f_0(x) = -x$ (linear case), there are continuously many cycles coexisting at $\mu = 0$.

Now let us assume that $f_\mu(x)$ is jointly C^2 and that

$$Df_\mu(0) = -\sqrt{1 + \mu}.$$

(As noted above, this last condition can be arranged by a smooth reparametrization if we start with a parametrized family for which $D_\mu f'_\mu(0) > 0$ at $\mu = 0$.) We can then write

$$f_\mu \circ f_\mu(x) = (1 + \mu)x - \beta_2(\mu, x)x^2,$$

and thus write the equation for a non-zero fixed point as

$$\mu - x\beta_2(\mu, x) = 0.$$

By arguments similar to ones used above, $x\beta_2(\mu, x)$ is jointly C^1 , and the μ derivative of the left-hand side of this equation at $\mu = 0$, $x = 0$ is $1 \neq 0$. Hence, again by the Implicit Function Theorem, this equation can be solved (uniquely) for μ as a continuous function of x , and the dependence is actually C^1 (or C^{n-1} if $f_\mu(x)$ is jointly C^n). A simple computation shows that $d\mu/dx = 0$ at $x = 0$. In fact: The number of solutions of the fixed point equation for a given μ must be even if it is finite, since solutions are points of period 2, and points of period 2 come in pairs, one positive and one negative. Thus, the first non-vanishing derivative of μ with respect to x , at $x = 0$, if it exists, is of even order.

§5. Flows and the Hopf bifurcation. We have so far discussed bifurcations only for maps. There is a parallel theory for flows generated by differential equations

$$\frac{dx}{dt} = X(x).$$

The analogue of a fixed point of a map is a stationary solution, i.e., a point $\bar{x}$ where $X(\bar{x})$ vanishes. If X depends on a parameter μ , then a stationary solution can be followed as the parameter changes as long as zero is not in the spectrum of $DX(\bar{x})$. If we ask how a stationary solution can lose attractivity as the parameter is changed, there are only two non-degenerate possibilities: Either a single eigenvalue of DX_μ goes to 1, or a complex-conjugate pair goes through $\pm i\phi$ with $\phi \neq 0$. In the first case, the only non-degenerate possibility is that the stationary solution undergoes a saddle-node bifurcation, i.e., collides with a non-attracting fixed point and disappears. The theory of the saddle-node bifurcation for flows can be deduced from the theory for mappings, but in this instance it is probably better to prove the result for flows directly; the argument is simpler than for mappings. We will not repeat the analysis here.

The case of a complex-conjugate pair of eigenvalues is much more interesting. We are going to argue that there are two non-degenerate possibilities, one of which is a bifurcation producing an attracting periodic orbit. In other words, the asymptotic behavior of most orbits starting near the fixed point changes from stationary, with the typical orbit converging to the fixed point, to oscillatory, with the typical orbit converging to the attracting periodic orbit.

We start, as usual, by looking at the bifurcation in the lowest-dimensional situation in which it can occur, i.e., in two dimensions. We will later use the Center Manifold Theorem to extend the analysis to higher dimensions. As usual, we can arrange things so that the origin in $\mathbb{R}^2$ is a stationary solution for all μ near zero and that this stationary solution ceases to be attracting at $\mu = 0$. We write $\lambda(\mu) \pm i\phi(\mu)$ for the eigenvalues of $DX_\mu(0)$, and we assume that $\lambda(0) = 0$ and that $\lambda(\mu)$ is strictly increasing on a neighborhood of 0. For definiteness we also take $\phi(\mu)$ to be positive. It will be convenient to make a linear change of variable in $\mathbb{R}^2$ if necessary so that the linearized solution flow at $\mu = 0$ is a rotation.

Since we have arranged to make the linearized flow at $\mu = 0$ be a rotation, orbits near the origin at $\mu = 0$ are approximately circular with angular velocity $\phi(0)$. For μ near zero, they spiral around the origin, either inward or outward depending on the sign of $\lambda(\mu)$. In particular, every orbit near zero (except for the stationary solution itself) crosses the x axis repeatedly. We will study the return map to the x axis, which we denote by f_μ . (To avoid confusion: We mean by f_μ the mapping which follows the orbit *half-way* around the stationary solution, not all the way around.) Evidently, f_μ has the stationary solution as a fixed point, and the sign of $f'_\mu(x)$ is opposite to that of x for $x \neq 0$. We are going to argue that $f_\mu(x)$ is jointly regular in μ, x and that $f'_\mu(0)$ goes through -1 as μ goes through zero. *Thus we would expect f_μ to undergo either a normal or an inverted period-doubling bifurcation.* We will also show that this is in fact the non-degenerate situation. If, for example, a normal bifurcation occurs, then, after

the bifurcation, f_μ has an attracting cycle of period 2. It is easy to see, however, that a point x is periodic of period 2 for f_μ if and only if it lies on a periodic solution for the differential equation, and that x is attracting (repelling) for f_μ if and only if the periodic solution is attracting (repelling) for the flow.

What comes out, therefore, is that there are two non-degenerate possibilities when a complex conjugate pair of eigenvalues crosses from the left half-plane to the right. In one of them, called the *normal Hopf bifurcation*, the stationary solution emits an attracting periodic solution as the eigenvalue pair crosses the axis. In the other, called the *inverted Hopf bifurcation*, it absorbs a repelling periodic orbit. The detailed analysis given in the preceding section for the period-doubling bifurcation gives correspondingly detailed information about the local dynamics and parameter dependence for the Hopf bifurcation.

To justify the above, we need

- (a) to find regularity conditions on $X_\mu(x)$ which guarantee that f_μ is sufficiently regular for the theory of the preceding section to apply
- (b) to show that $f'_\mu(0)$ passes through -1 as the eigenvalue pair for $DX_\mu(0)$ passes through the imaginary axis.
- (c) to show that, for all but exceptional X_0 's, $(f_0 \circ f_0)'''(0) \neq 0$.

For the first point, it is convenient to work in polar coordinates. We write the rectangular coordinates in $\mathbf{R}^2$ as (x_1, x_2) and the differential equation as

$$\frac{dx_1}{dt} = X_1(x_1, x_2), \quad \frac{dx_2}{dt} = X_2(x_1, x_2).$$

(We suppress, for the moment, the μ dependence.) The mapping from polar coordinates to rectangular is, is usual,

$$x_1 = r \cos \theta, \quad x_2 = r \sin \theta,$$

and a standard computation shows that the differential equation transformed to polar coordinates reads

$$(5.1) \quad \frac{dr}{dt} = R(r, \theta), \quad \frac{d\theta}{dt} = \Theta(r, \theta),$$

where

$$\begin{aligned} R(r, \theta) &= \cos \theta X_1(r \cos \theta, r \sin \theta) + \sin \theta X_2(r \cos \theta, r \sin \theta) \\ \Theta(r, \theta) &= \frac{1}{r} [\sin \theta X_1(r \cos \theta, r \sin \theta) - \cos \theta X_2(r \cos \theta, r \sin \theta)] \end{aligned}$$

Because of the singularity at the origin in this change of coordinates, it is necessary to be careful about exactly what relation is being asserted to hold between the original differential equations and the transformed ones. The fact we want to use is that, if $r(t)$, $\theta(t)$ is any solution of (5.1), and if $x_1(t) = r(t) \cos(\theta(t))$, $x_2(t) = r(t) \sin(\theta(t))$,

then $x_1(t), x_2(t)$ is a solution of the original equation. There is no need to worry about whether r is positive or not.

Evidently, if $X_\mu(x)$ is jointly C^n in μ, x , then $R_\mu(r, \theta)$ is jointly C^n in μ, r, θ . From the same sorts of considerations used in the preceding section, $\Theta_\mu(r, \theta)$ is jointly C^{n-1} . Note also that $R_\mu(0, \theta) = 0$ and that $\Theta_0(0, \theta) = \phi(0)$. Thus, if μ, x are small enough the solution of (5.1) with $r(0) = x$ and $\theta(0) = 0$ will have $d\theta/dt$ strictly positive up to a value of t where $\theta(t) = \pi$. We accordingly write $t_\mu(x)$ for the first positive time when $\theta(t) = \pi$ for this solution. By standard results about smoothness of solutions of differential equations, and the Implicit Function Theorem, $t_\mu(x)$ is jointly C^{n-1} in μ, r . Here, x can be positive or negative. The return mapping which we called f_μ above is given simply by

$$f_\mu(x) = r(t_\mu(x)),$$

and is thus of class C^{n-1} in μ, x .

Remark: I believe that the lost degree of differentiability can be recovered by more careful estimates. It is easy to show by example that $t_\mu(x)$ need not be more than C^{n-1} in x , but it appears that $xt_\mu(x)$ is C^n and similarly that $f_\mu(x)$ is C^n . I have not, however, carried out the estimates in detail.

The next task is to compute $f'_\mu(0)$. Since $f_\mu(0) = 0$, we have only to compute $f_\mu(x)$ for small x to first order in x . For small x , $t_\mu(x) \approx \pi/\phi(\mu)$, and hence

$$r(t_\mu(r)) \approx -r \exp\left(\frac{\lambda(\mu)\pi}{\phi(\mu)}\right),$$

i.e.,

$$f'_\mu(0) = -\exp\left(\frac{\lambda(\mu)\pi}{\phi(\mu)}\right),$$

which does decrease through -1 as $\lambda(\mu)$ increases through zero.

Finally, we have to argue that $(f_0 \circ f_0)'''(0)$ is "generically non-zero". This could be done by direct computation, but it is easier to make an indirect argument: It is evident that there is a formula for $(f_0 \circ f_0)'''(0)$ involving only the first, second, and third derivatives of X_0 at the origin and that this formula is analytic in its finite set of variables. Since it is analytic, we need only show that it is not *identically* zero; an analytic function which is not identically zero is generically non-zero. Thus, all we need to do is to exhibit one differential equation X such that DX has spectrum $\{\pm i\phi\}$ and whose corresponding f satisfies $(f \circ f)'''(0) \neq 0$. We take

$$\frac{dx_1}{dt} = (x_1^2 + x_2^2)x_1 + \phi x_2, \quad \frac{dx_2}{dt} = (x_1^2 + x_2^2)x_2 - \phi x_1.$$

In polar coordinates these equation become simply

$$\frac{d\theta}{dt} = \phi; \quad \frac{dr}{dt} = r^3$$

and in this form it is easy to solve them explicitly; $f \circ f$ is the time- $2\pi/\phi$ mapping for the r differential equation, i.e.,

$$f \circ f(r) = r \left[1 - \frac{4\pi r^2}{\phi} \right]^{-1/2}$$

and hence $(f \circ f)'''(0) \neq 0$ as desired.

The generalization of the preceding analysis to more than two dimensions, using the Center Manifold Theorem, goes exactly as for the period doubling bifurcation. If the spectrum is in the left half-plane except for the eigenvalue pair which is crossing the axis, then the periodic solution produced in a normal bifurcation is attracting; the one which is absorbed in the inverted bifurcation is hyperbolic with one-dimensional unstable manifold. Just after a normal bifurcation takes place, the stationary solution—which still exists—is hyperbolic with two-dimensional unstable manifold. This unstable manifold is a small disk bounded by the attracting periodic solution. All orbits starting near enough to the stationary solution but not on its (codimension two) stable manifold are attracted to the periodic solution. As with the period-doubling bifurcation, it is not necessary for the rest of the spectrum to be in the left half-plane in order that the crossing of an eigenvalue pair into the right half-plane be associated with either the creation or the absorption of a periodic orbit, but this orbit will not be attracting in the general case.

The bifurcation-surface picture developed in the preceding section for the period-doubling bifurcation extends easily to the Hopf bifurcation. Furthermore, in the same way as the passage of an eigenvalue of a map through -1 is accompanied by the appearance of continuously many cycles of period 2, one can show with a minimum of technical hypotheses that the passage of a non-real eigenvalue pair for a differential equation through the imaginary axis is accompanied by the appearance of continuously many periodic solutions.

One point which needs to be stressed about the Hopf bifurcation is that it is easy to recognize (normal) Hopf bifurcations occurring in complicated systems, e.g. in hydrodynamics, even without being able to verify mathematically that the conditions for the above analysis apply. In hydrodynamic systems, for example, one often wants to study how the long-time behavior changes as the strength of some kind of driving is increased. If one observes a situation in which, as the strength of driving is increased slowly past a critical value, the observed behavior changes from stationary to oscillatory, and if the amplitude of the oscillations grows like the square root of the difference between the parameter and the critical parameter value, one can be reasonably confident that this transition is due to a Hopf bifurcation.

§6. The Hopf bifurcation for maps. In §4 we did not complete our analysis of the list of possible non-degenerate ways that a fixed point for a mapping could cease to be attracting; it remains to treat the case where a complex-conjugate pair of non-real eigenvalues cross the unit circle. This transition presents an evident analogy to the Hopf bifurcation discussed in the preceding section. Using the Center Manifold Theorem in

the usual way, we can reduce to the two-dimensional case. The theory of the minimal-dimensional manifestation of the bifurcation turns out, however, to be significantly more difficult than the cases considered previously, and we will only sketch the argument.

Let us first describe what the non-degenerate possibilities turn out to be. As with the period-doubling and Hopf bifurcations, there are only two non-degenerate possibilities, which can be regarded as normal and inverted versions of a single bifurcation. In the normal version, the fixed point emits an attracting circle invariant for the map as it loses attractivity; in the inverted version, it absorbs a repelling invariant circle. What makes the theory for maps much harder than for flows is that the invariant circle is not a single orbit; proving it exists for a given parameter value requires solving a functional equation rather than simply a one or two dimensional fixed point problem. The theory of this bifurcation is relatively recent. The result was announced (without proof) by J. Naimark "On some cases of periodic motions depending on parameters", *Dokl. Akad. Nauk SSR* 129 (1959) pp. 736-739. The first complete proof was apparently given by R. Sacker in his Ph.D. dissertation (1964, unpublished), and the result was rediscovered and published by D. Ruelle and F. Takens in "On the nature of turbulence", *Commun. Math. Phys.* 20 (1971) 167-192.

Probably the most important applications of the theory of this bifurcation is to Poincaré maps of periodic orbits for flows. The invariant circle produced by the bifurcation then lies in a transverse section to the periodic orbit; the set of all orbits passing through this circle is a two-dimensional torus invariant and attracting for the flow. The nature of the motion on this invariant torus is complicated, and we will discuss it in more detail when we come to talk about homeomorphic mappings of the circle to itself. The non-degenerate situation is that, for some but not all parameter values just after the transition, the motion on the torus (in appropriate coordinates) is just constant-velocity irrational flow. Thus, orbits starting near enough to the invariant torus display asymptotic behavior which is doubly-periodic.

Let us now turn to a sketch of the theory of this bifurcation. Consider a mapping f_μ of $\mathbb{R}^2$ to itself, jointly sufficiently smooth, with zero as a fixed point for all μ , and such that the spectrum of $Df_0(0)$ is $\{e^{\pm i\phi}\}$ where $0 < \phi < \pi$. (We will eventually also need to assume that $\phi \neq \pi/2, 2\pi/3$.) To avoid extraneous technical problems, we assume that this pair of eigenvalues crosses the unit circle with non-zero velocity at $\mu = 0$, and we can then reparametrize smoothly so that the spectrum of $Df_\mu(0)$ is $\{(1 + \mu)e^{\pm i\phi(\mu)}\}$. By making a linear change of coordinates, depending smoothly on μ , we arrange further that

$$Df_\mu(0) = (1 + \mu) \begin{pmatrix} \cos \phi(\mu) & -\sin \phi(\mu) \\ \sin \phi(\mu) & \cos \phi(\mu) \end{pmatrix}.$$

The first substantial step in the proof is to make a non-linear change of coordinates to eliminate as many non-linear terms as possible of degree 2 and 3 in f_μ . This is a simple concrete application of the ideas about formal linearizations developed in §V.2. We want to make these coordinate changes for each μ , and they should vary smoothly with μ . Thus, the presence of resonances at $\mu = 0$ prevents us from eliminating corresponding terms at $\mu \neq 0$ (where there are never any resonances). A second-order resonance can

only happen at $\mu = 0$ when

(6.1a)
$$e^{i\phi(0)} = e^{-2i\phi(0)},$$

and we rule this out by assuming that

$$e^{3i\phi(0)} \neq 1, \quad \text{i.e., } \phi(0) \neq 2\pi/3.$$

Third order resonances definitely *do* occur:

$$e^{\pm i\phi(0)} = e^{\pm i\phi(0)} e^{\pm i\phi(0)} e^{\mp i\phi(0)},$$

but we can ensure that there will be no third-order resonances other than these unavoidable ones by requiring

(6.1b)
$$e^{4i\phi(0)} \neq 1.$$

The detailed calculation of what terms can and cannot be eliminated is particularly easy to do in the case we are considering, using complex notation. We need the following elementary algebraic fact: Describe a point $(x_1, x_2) \in \mathbb{R}^2$ by the single complex number $z = x_1 + ix_2$. In this representation, any homogeneous polynomial ϕ of degree n mapping $\mathbb{R}^2$ to itself can be written uniquely as

$$z \mapsto z' = \sum_{j=0}^n a_j z^j (\bar{z})^{n-j},$$

where the a_j are complex coefficients, and where the overbar denotes complex conjugacy.

To prove this, first argue that, since

$$x_1 = \frac{1}{2}(z + \bar{z}), \quad x_2 = \frac{1}{2i}(z - \bar{z}),$$

any real-valued homogeneous polynomial can be written in the indicated form with

$$a_{n-j} = \bar{a}_j.$$

Similarly, any pure-imaginary homogeneous polynomial can be written in this form (with another condition on the coefficients). Thus, any homogeneous polynomial with values in $\mathbb{R}^2$ admits such a representation with some set of coefficients. The uniqueness of the representation can be proved by exhibiting each coefficient as the value at zero of some appropriate linear combination of n -th order partial derivatives of the two components of the polynomial.

We write λ_μ for the eigenvalue $(1 + \mu)e^{i\phi(\mu)}$ of $Df_\mu(0)$. The main advantage of the complex representation is that the action of $Df_\mu(0)$ becomes simply multiplication by λ_μ . Thus, if ϕ is represented as above, the transform $\Lambda^{-1}\phi\Lambda^n$ of ϕ (where Λ denotes $Df_\mu(0)$) is represented by

$$\sum_{j=0}^n \lambda_\mu^{j-1} (\bar{\lambda}_\mu)^{n-j} z^j (\bar{z})^{n-j}.$$

Terms which can be eliminated are those for which

$$\lambda_\mu^{j-1} (\bar{\lambda}_\mu)^{n-j} \neq 1$$

As already noted, all terms with $n = 2$ can be eliminated provided (6.1a) holds. For $n = 3$, the term $a_2|z|^2z$ cannot be eliminated, but all others can provided that (6.1b) holds. By the analysis of §V.2, this shows that there exists a change of coordinates $\Phi_\mu(z) = z + O(z^2)$ such that

$$\Phi_\mu \circ f_\mu(z) = \lambda_\mu \Phi(z) + a_\mu |z|^2 z + O(|z|^4).$$

Since $\Phi(z) = z + \text{higher order terms}$, this equation can be rewritten as

$$\Phi \circ f_\mu \circ \Phi^{-1}(z) = \lambda z + a|z|^2 z + O(|z|^4).$$

Thus, by making the change of coordinates Φ , we can bring f_μ approximately into the normal form

$$z \mapsto \lambda_\mu(1 + a_\mu |z|^2)z.$$

(We have changed the meaning of a_μ slightly by absorbing a factor of λ_μ^{-1} .) The significance of this normal form is clearer if we rewrite the transformed mapping in polar coordinates $(r, \theta) \mapsto (r', \theta')$ where

$$(6.2a) \quad r' = (1 + \mu)|1 + ar^2|r = (1 + \mu)r - br^3 + O(r^4)$$

and

$$(6.2b) \quad \theta' = \theta + \phi(\mu) + \arg(1 + a_\mu r^2) = \theta + \phi(\mu) + cr^2 + O(r^3)$$

(Here, we have introduced two new real μ -dependent quantities b and c , related in a straightforward way to μ and a_μ . Note also that the terms " $O(r^4)$ " (respectively " $O(r^3)$ ") mean jointly C^4 (respectively C^3) functions of r, θ, μ which vanish to fourth (respectively third) order as $r \rightarrow 0$, uniformly in μ (small) and θ . To make all this work, we need to assume that $f_\mu(x)$ is jointly C^4 .)

The mapping obtained by dropping the corrections:

$$(r, \theta) \mapsto ((1 + \mu)r - br^3, \theta + \phi(\mu) + cr^2)$$

has a very simple structure: the r -component of the image does not depend on θ . If we assume

$$b > 0$$

then for $\mu > 0$ the circle

$$r = r_0 = \sqrt{\mu/b}$$

is mapped to itself. Furthermore

$$\left. \frac{d}{dr} ((1 + \mu)r - br^3) \right|_{r=r_0} = 1 - 2\mu$$

so this circle is attracting—any orbit starting near to it converges to it asymptotically. The idea will be to show that the correction can be treated as perturbations and that, when they are taken into account, the mapping still admits an attracting invariant circle near to this one. The argument is a little delicate because the strength of attractivity of the unperturbed circle under the unperturbed map goes to zero linearly with μ . What ultimately makes it work is that, in the vicinity of the attracting circle and with appropriate rescalings, the other terms are smaller still.

We note, first of all that there exist positive constants R_0, A (depending on the map but not on μ provided μ is small) such that

$$0 < r \leq r_0 - A\mu \quad \text{implies} \quad r' > r,$$

and

$$r_0 + A\mu \leq r \leq R_0 \quad \text{implies} \quad r' < r.$$

(Here, (r', θ') denotes the image of the point (r, θ) under the full map including the correction terms.) Thus, the annulus

$$r_0 - A\mu \leq r \leq r_0 + A\mu$$

is mapped into itself and attracts all orbits starting anywhere in the punctured disk

$$0 < r \leq R_0.$$

Next we shift and rescale the radial coordinate, introducing a new radial variable y by

$$r = r_0 + A\mu y.$$

Thus, the attracting invariant annulus becomes the band

$$-1 \leq y \leq 1,$$

and we will focus our attention from now on purely on this band; in particular, in the absence of any specification to the contrary, all estimates are supposed to hold uniformly

in this band. We will write the mapping f_μ , restricted to this domain and expressed in these coordinates, as

$$F : (\theta, y) \mapsto (\theta', y');$$

the μ dependence is suppressed from the notation except when it is really necessary.

In outline, the argument is now going to proceed as follows: We will show that the image under F of the graph of a Lipschitz-continuous mapping $y = w(\theta)$, with an appropriate bound on the Lipschitz constant, is again the graph of a Lipschitz continuous mapping with the same bound on the Lipschitz norm. Thus, we can define an operator $\mathcal{F}$ acting on a space of Lipschitz functions so that the graph of $\mathcal{F}w$ is the image under F of the graph of w . We then show that this operator is contractive in a norm equivalent to the supremum norm, and hence that there is a unique Lipschitz continuous function whose graph is mapped to itself by F . The construction is μ -dependent, but all estimates hold uniformly for small μ . Once the existence of a Lipschitz invariant circle has been proved, the same techniques as were used to prove the smoothness of stable and unstable manifolds can be applied to establish the smoothness of this invariant circle. We will sketch the argument only up to the point of showing the existence of a Lipschitz invariant circle; the rest is laborious but very similar to things we have already seen.

Straightforward computations starting from the approximate normal form (6.2) yield:

$$(6.3) \quad DF(\theta, y) = \begin{pmatrix} 1 + O(\mu^{3/2}) & O(\mu^{3/2}) \\ O(\mu) & 1 - 2\mu + O(\mu^{3/2}) \end{pmatrix}.$$

(Note that the arguments are written in the order (θ, y) , so, for example, the upper right-hand element in the derivative matrix is the partial derivative of θ' with respect to y .) In deriving these estimates, repeated use is made of the fact that $r = O(\sqrt{\mu})$.

The proof of the existence of an invariant Lipschitz circle can be efficiently formulated in geometric terms by using the idea of expanding and contracting wedges. For a positive number B , we define

$$H_B = \{(v_\theta, v_y) \in \mathbb{R}^2 : |v_y| \leq B|v_\theta|\}$$

and

$$V_B = \{(v_\theta, v_y) \in \mathbb{R}^2 : |v_\theta| \leq B^{-1}|v_y|\}.$$

Heuristically, the elements of H_B and V_B should be thought of as displacements between points; the displacements in H_B are predominantly horizontal and those in V_B predominantly vertical.

To show that the image under F of the graph of a function with Lipschitz norm not greater than B is again the graph of such a function, it is enough to prove that, whenever the separation between two points z_1 and z_2 is in H_B , the separation between $F(z_1)$ and $F(z_2)$ is also in H_B . This will follow if we can show that $DF(z)$ maps H_B into itself, preserving the sign of the θ component, for all relevant z . It will suffice therefore to find B such that

$$B \left| \frac{\partial y'}{\partial y} \right| + \left| \frac{\partial y'}{\partial \theta} \right| \leq B \frac{\partial \theta'}{\partial \theta} - B^2 \left| \frac{\partial \theta'}{\partial y} \right|,$$

i.e., such that

$$\left| \frac{\partial y'}{\partial y} \right| + \frac{1}{B} \left| \frac{\partial y'}{\partial \theta} \right| + B \left| \frac{\partial \theta'}{\partial y} \right| \leq \frac{\partial \theta'}{\partial \theta}.$$

In view of (6.3), this inequality holds for any sufficiently large B provided that μ is small enough. (More precisely: For any sufficiently large B , there is $\mu_0 > 0$ such that $DF_\mu(z)$ maps H_B into itself for all $\mu < \mu_0$ and all relevant z .) We choose one such B , say B_1 , and work in the space of Lipschitz continuous mappings w which are periodic with period 2π , have $|w(\theta)| \leq 1$ everywhere, and satisfy $\text{Lip}(w) \leq B_1$. We have just shown that $\mathcal{F}$ is well defined on this space and maps it to itself.

The next step is to show that $\mathcal{F}$ is contractive in a norm equivalent to the supremum norm on this space of Lipschitz mappings. The norm we use is as follows: We take $B_2 > B_1$, and let

$$\|w_1 - w_2\| = \sup\{|(z_1 - z_2)_y| : z_1 \in \Gamma(w_1), z_2 \in \Gamma(w_2), z_1 - z_2 \in V_{B_2}\}.$$

Here, $(z_1 - z_2)_y$ means the y component of $z_1 - z_2$ and $\Gamma(w_i)$ the graph of w_i . This norm is evidently bounded below by the supremum norm, and it is easy to see that it is bounded above by the supremum norm multiplied by $B_2/(B_2 - B_1)$.

It follows easily from (6.3) that by taking B_2 large enough and μ small enough we can guarantee that

$$|(DFv)_y| \leq (1 - \mu)|v_y| \quad \text{for } v \in V_{B_2}.$$

We can simultaneously arrange that DF maps H_{B_2} into itself and, since the complement of H_{B_2} is the interior of V_{B_2} , it follows that DF^{-1} maps V_{B_2} into itself. Thus

$$\begin{aligned} \|\mathcal{F}w_1 - \mathcal{F}w_2\| &= \\ &= \sup\{|(F(z_1) - F(z_2))_y| : z_1 \in \Gamma(w_1), z_2 \in \Gamma(w_2), F(z_1) - F(z_2) \in V_{B_2}\} \\ &\leq \sup\{|(F(z_1) - F(z_2))_y| : z_1 \in \Gamma(w_1), z_2 \in \Gamma(w_2), z_1 - z_2 \in V_{B_2}\} \\ &\leq (1 - \mu) \sup\{|(z_1 - z_2)_y| : z_1 \in \Gamma(w_1), z_2 \in \Gamma(w_2), z_1 - z_2 \in V_{B_2}\} \\ &= (1 - \mu)\|w_1 - w_2\|. \end{aligned}$$

Thus, $\mathcal{F}$ is contractive, so the existence of a Lipschitz invariant circle follows. The Contraction Mapping Principle also shows that the successive images of the graph of any w converge to the unique invariant circle, and this implies in particular that the orbit of any point in the strip $\{|y| \leq 1\}$ converges to the invariant circle. Combining this with remarks made earlier shows that all orbits starting in $\{0 < r \leq R_0\}$ converge to the invariant circle.

Remark: The above analysis doesn't say anything about smoothness of the dependence of the circle on μ . Presumably something along these lines could be proved.

§1. The ω -limit set. Let f be a mapping (which, for present purposes, does not need to be any more than continuous); suppose x is such that $f^n(x)$ is defined for all $n \geq 0$ and that $\{f^n(x)\}$ is relatively compact in the domain of f . Under these assumptions, we define the ω -limit set (of x , or x with respect to f , and accordingly denoted $\omega(x)$ or $\omega(x, f)$) as the set of all limits of convergent sequences of the form $f^{n_j}(x)$, where $n_j \rightarrow \infty$. Alternatively

$$\omega(x, f) = \bigcap_{m \rightarrow \infty} \overline{f^m o^+(x, f)},$$

where the bar denotes closure (in the domain of f) and $o^+(x, f)$ the (forward) orbit of x under f :

$$o^+(x, f) = \{f^n(x) : n \geq 0\}.$$

It is easy to see that the ω -limit set is strictly invariant, i.e., is mapped onto itself by f . Using the relative compactness of the orbit of x , it follows easily that $\omega(x, f)$ is non-empty and compact, and that it "attracts" the orbit of x , i.e., that

$$f^n(x) \xrightarrow{n \rightarrow \infty} \omega(x, f).$$

In fact, the definition makes sense without assuming the relative compactness of the orbit, and the set so defined is strictly invariant, but it may be empty (and need not attract the orbit of x).

We make similar definition for a flow (or semiflow) f^t , i.e., we put

$$o^+(x, f^t) = \{f^t(x) : t \geq 0\} \quad \text{and} \quad \omega(x, f^t) = \bigcap_{s \rightarrow \infty} \overline{f^s o^+(x, f^t)}.$$

In addition to the properties of the ω -limit set formulated above for mappings, the ω -limit set for a point under a flow is connected. The analogous limit set with time going to $-\infty$ is called the α -limit set.

For a mapping f (or a flow) all of whose orbits are relatively compact, each orbit converges asymptotically to its ω -limit set. Thus, the very long-term behavior of all orbits can be determined from knowledge of the behavior of the mapping on a neighborhood of its limit set $L(f)$ defined by

$$L(f) = \overline{\cup \{\omega(x, f)\}}$$

where the union is taken over all initial points x . It is often the case that $L(f)$ is much smaller than the whole space. For example, we saw earlier that, for the Hénon mapping with small b , there is a non-empty open invariant set all orbits in which converge to a certain compact set of Lebesgue measure zero; this set of measure zero thus contains the limit set of the restriction of the Hénon mapping to the open invariant set.

§2. Flows in the plane. The main interest in dynamical systems theory is in systems with complicated limit sets. We will present here a classical result, due to Poincaré and Bendixson, which says that there is no point in looking, from this point of view, at flows in the plane—orbits for such flows have uninteresting asymptotic behavior.

THEOREM 2.1 (POINCARÉ-BENDIXSON THEOREM). *Let $X(x)$ be a differential equation defined in an open subset of the plane $\mathbb{R}^2$, x_0 a point of the domain of X such that the solution curve $f^t(x_0)$ is defined for all $t \geq 0$ and relatively compact in the domain of X . If the ω limit set of x_0 contains no stationary solutions of X , this ω -limit set consists of a single periodic orbit to which $f^t(x_0)$ converges.*

If we drop the assumption that the ω -limit set contains no stationary solutions, the situation becomes more complicated but there still appear to be no interesting possibilities. The simplest thing that can happen is that the orbit converges to a stationary solution. A more general possibility is that it converges to an invariant closed curve on which there are one or more stationary solutions.

PROOF: The heart of the proof is a simple argument which rests essentially on the Jordan Curve Theorem and thus uses in an essential way deep facts about the topology of $\mathbb{R}^2$. We will use the term *transverse arc* to refer to a smooth arc γ in the domain of X which is everywhere transverse to the vector field X , i.e. such that, for every point x on γ , $X(x)$ is not a constant multiple of the tangent vector to γ at x (and in particular $X(x) \neq 0$.)

LEMMA 2.2. *Let X be a vector field defined on an open subset of $\mathbb{R}^2$, γ a transverse arc, $x(t)$ a solution curve for the differential equation defined by X . If $x(t)$ crosses γ more than twice, the successive crossing points form a monotonic sequence in γ .*

PROOF: Let $t_0 < t_1$ be two successive crossing times. If $x(t_0) = x(t_1)$, there is nothing to prove. Otherwise, consider the closed curve Γ made up of the segment of the orbit for t between t_0 and t_1 and the segment of γ between $x(t_0)$ and $x(t_1)$. (In Figure 1, γ runs from γ_0 to γ_1 . The heavy line is the part of Γ within the picture. Γ starts at $x_0 = x(t_0)$, runs off the top of the picture, winds around in some unspecified way, re-enters the picture from below and crosses γ again at $x_1 = x(t_1)$; it is then closed by adding the part of γ between x_1 and x_0 .)

Γ is a Jordan curve; hence, it separates the plane into two components. The part of the orbit for t slightly larger than t_1 may be either inside or outside Γ , depending on the global geometry, but in any event is on the other side from points lying slightly downstream from the part of γ running from $x(t_1)$ to $x(t_0)$ and beyond. (In Figure 1, the dotted arc represents the set of points obtained by following the part of γ from γ_0 to x_1 backwards for some short period of time; the future orbit of $x(t_1)$ thus cannot reach the arc from γ_0 to x_1 without crossing the dotted arc, which is on the opposite side of Γ from the part of the solution curve immediately above x_1 .) Thus, the forward orbit of $x(t_1)$ cannot get to the dotted arc without first crossing Γ . But it certainly cannot cross the part of Γ which is a segment of orbit because orbits cannot cross. It also cannot cross the other part of Γ —the arc from x_0 to x_1 —because it can't get there

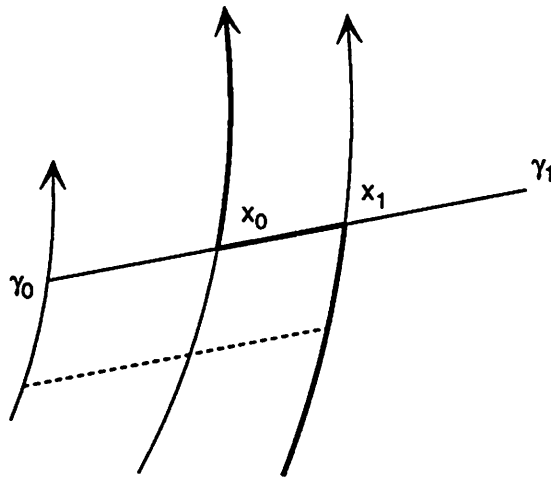


Figure 1

without first crossing the dotted arc. Hence, the next place where the orbit crosses γ , if there is one, must lie between x_1 and γ_1 , as desired.

COROLLARY 2.3. *If $x(t)$ is defined for all $t \geq 0$, then no more than one point of any transverse arc γ can belong to the ω -limit set of $x(t)$*

PROOF: Let y be a point of γ which is in the ω -limit set. Then y is the limit of some sequence $x(t_j)$ with $t_j \rightarrow \infty$. Since the vector field is transverse to γ at y , we can take the t_j to be times when $x(t)$ crosses γ , i.e., y is an accumulation point of the sequence of successive points where $x(t)$ crosses γ . This holds for any point on γ which is in the ω -limit set. But the sequence of all successive crossing points is monotonic on γ ; hence, cannot have more than one accumulation point.

PROOF OF THEOREM: Let $y_0 \in \omega(x_0, f^t)$. By the invariance and compactness of $\omega(x_0, f^t)$, the orbit $f^t y_0$ is defined for all $t \geq 0$ and contained in $\omega(x_0)$; hence is relatively compact. It thus has a non-empty ω -limit set, and

$$\omega(y_0) \subset \omega(x_0).$$

Let y_1 be a point of $\omega(y_0)$. Since, by assumption, there are no stationary solutions in $\omega(x_0)$, $X(y_1) \neq 0$, so there exists a transverse arc γ through y_1 . Because $y_1 \in \omega(y_0)$, the orbit of y_0 must return infinitely often to the neighborhood of y_1 and so must cross γ infinitely often. But the whole orbit of y_0 is in $\omega(x_0)$, so each of the crossing points must be in $\omega(x_0)$. By the Corollary, then, there can only be one crossing point, i.e., y_1 must lie on a periodic solution. From the existence of a sequence t_j such that $x(t_j)$ converges to y_1 , it follows (simply from the continuous dependence of solutions on initial data) that the orbit $x(t)$ converges asymptotically to the periodic solution through y_1 .

Note, incidentally, that the argument does not use the full force of the assumption that $\omega(x_0)$ does not contain a stationary solution, but only that there is some point $y_0 \in \omega(x_0)$ whose orbit is not asymptotic to a stationary solution.

The above argument evidently depends in an essential way on the global topology of the plane. The example of constant-velocity irrational flow on the torus—which has no stationary solutions and for which the ω -limit set of every orbit is the whole torus—shows that the theorem as it stands does not hold if the plane is replaced by a general two-dimensional manifold. There is, however, a partial extension to general two-dimensional manifolds, in which the torus appears as the only case with even mildly interesting dynamics; see §VII.12 of [Hartmann].

Roughly speaking, the situation is that interesting dynamics can occur for non-invertible mappings already in one dimension; for invertible mappings in two and more dimensions, and for flows in three and more dimensions.

CHAPTER VIII. HYPERBOLIC SETS

Dynamical systems often have the property that most orbit pairs which start very near to each other diverge exponentially, at least until their separation becomes "macroscopic," and the investigation of this phenomenon has been a particularly fruitful direction in the study of dynamical systems. Exponential divergence of orbits gives rise, notably, to the sensitive dependence on initial conditions which accounts for the apparently stochastic behavior of deterministic systems. In this and the following chapters, we will discuss the theory of dynamical systems obeying a technically strong version of the exponential separation of nearby orbits, i.e., ones with strong uniformity assumptions on the rate of divergence. We begin by looking at some examples.

§1. Linear torus automorphisms. We return to the example discussed in §IV.8.7, the mapping f of $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ induced by the linear mapping $z \mapsto Fz$ of $\mathbb{R}^2$ to itself, where F denotes the 2×2 matrix

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

In our earlier discussion of this example, we concentrated on the stable and unstable manifolds of its fixed and periodic points. We saw that periodic points are exactly the points with rational coordinates; that all periodic points are hyperbolic; and that the stable (unstable) manifolds for these periodic points are all projections into $\mathbb{T}^2$ of lines in $\mathbb{R}^2$ parallel to the contracting (expanding) eigenvector of F . We will write λ for the contracting eigenvalue:

$$\lambda = \frac{3 - \sqrt{5}}{2};$$

the expanding eigenvalue is λ^{-1} , and we write ϕ_s and ϕ_u for a pair of eigenvectors of F with eigenvalues λ and λ^{-1} respectively.

The distance between two points in $\mathbb{T}^2$ will mean the minimum distance between representatives in $\mathbb{R}^2$. Thus if $z = (x, y)$ and $z' = (x', y')$ are two points in $\mathbb{R}^2$, there is no need to distinguish the distance between z and z' from the distance between their images in $\mathbb{T}^2$ as long as the former is small compared to one. By expanding $z - z'$ in the basis $\{\phi_s, \phi_u\}$, we see that, unless $z - z'$ is exactly parallel to ϕ_s , the distance between $F^n z$ and $F^n z'$ will eventually get to be of order one. Furthermore, the distance from z to z' is small enough, there will be a range of n 's over which the distance between $F^n z$ and $F^n z'$ is small compared to one and growing with increasing n like a constant multiple of λ^{-n} . Hence, separation between most nearby pairs of orbits grows exponentially as the orbits are followed forward until this separation gets to be of order one. On the other hand, if $z - z'$ is exactly parallel to ϕ_s , then the distance between $F^n z$ and $F^n z'$ goes to zero and in fact decreases exponentially like λ^n . Thus, although the projections of the $F^n z$ and $F^n z'$ may wander around $\mathbb{T}^n$ in an uncontrolled way, they nevertheless approach each other. One piece of information to be extracted from this is the following: Let $\eta > 0$ be sufficiently small. (Anything smaller than $\lambda/2$ will do if we use the standard Euclidean norm on $\mathbb{R}^2$.) Then, for any $z \in \mathbb{T}^2$

$$W_\eta^s(z) \equiv \{z' \in \mathbb{T}^2 : \|f^n z - f^n z'\| < \eta \text{ for all } n \geq 0\}$$

is a smooth submanifold passing through z (In fact: It is the image under the quotient mapping of a straight line segment parallel to ϕ_s in $\mathbb{R}^2$.) Furthermore, for any $z' \in W_\eta^s(z)$, $\|f^n z - f^n z'\|$ goes to zero exponentially as n goes to infinity. All this is very reminiscent of the situation for stable manifolds of fixed or periodic points except that the “reference orbit” $f^n z$ may keep moving around the space in an irregular way rather than being stationary or periodic.

The above analysis for forward iteration can be repeated for backward iteration. If the separation between z and z' is not exactly parallel to ϕ_u , the backward orbits of z and z' diverge exponentially until the distance between them is of order one. If, on the other hand, the separation is parallel to ϕ_u , the backward orbits approach each other exponentially. If, as above, we put

$$W_\eta^u(z) \equiv \{z' \in \mathbb{T}^2 : \|f^{-n} z - f^{-n} z'\| < \eta \text{ for all } n \geq 0\},$$

then $W_\eta^u(z)$ is a smooth one-dimensional submanifold passing through z and every point on this submanifold has backward orbit converging exponentially to that of z .

A number of features of this example occur often in more general situations. First: All but exceptional pairs of nearby points have orbits which diverge exponentially both forward and backward in time. Second: For each point z , the set of z' whose forward orbits remain near to that of z for all time is a smooth submanifold; similarly the set of z' whose backward orbits remain near to that of z . These two submanifolds intersect only in z ; their tangent spaces are complementary subspaces of the tangent space to the full state space. Third: Periodic points are dense, and the number of periodic points of each period can be determined by combinatorial means. At this point, all these features have been seen to hold by very special methods based on the explicit form of the mapping. We will see later on, however, that they are in fact qualitative features of the system preserved under any perturbation which is small in the C^1 topology. In fact, as we will also see later, this system is even *structurally stable*; any mapping of $\mathbb{T}^2$ to itself near enough in the C^1 topology to f can be transformed into f by a topological change of coordinates (but usually not by a smooth one.)

§2. The Smale horseshoe. We consider a mapping f of the unit square into the plane obtained as follows: Compress the square horizontally and stretch it vertically to obtain a tall thin rectangle; then bend this tall thin rectangle around into a horseshoe shape and lay it back over the original square. (See Figure 1). Although this mapping does not send the unit square into itself, we can still ask about the structure of the set of orbits (z_n) which stay in the unit square for all n , positive and negative. It is not immediately clear, for example, whether there are any such orbits.

To answer the sorts of questions suggested above, it is enough to know how f acts on the part of the square which is mapped back into itself. The image of the square reintersects the square in two pieces, each a thin vertical strip running across the full height of the square. The preimage of each of these strips is a thin horizontal strip running across the full width of the square. We call these preimages Δ_0 and Δ_1 . To go further, it is necessary to impose some conditions on how f acts on Δ_0 and Δ_1 . What

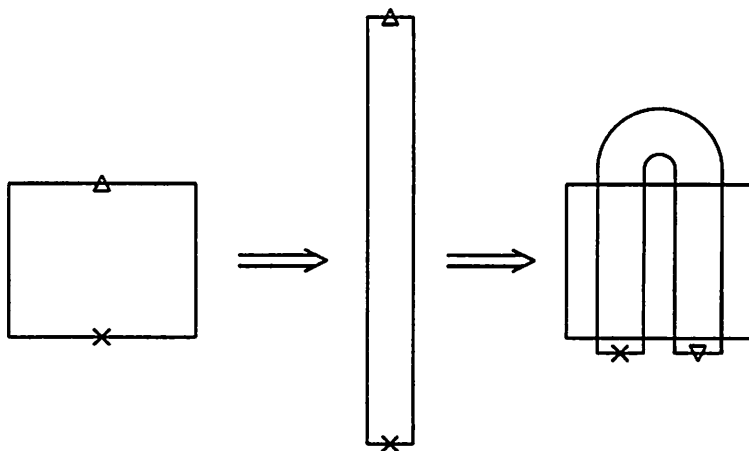


Figure 1

would be most satisfactory would be to base an analysis on qualitative conditions. This can be done, but, in order to arrive quickly at a clear picture of the structure, we will do something less ambitious—we make the very restrictive assumption that f is exactly affine on each of Δ_0 and Δ_1 . More precisely, we assume that f is given on Δ_i , $i = 0, 1$, by

$$z \mapsto M_i z + z_i$$

where the M_i are diagonal matrices and the z_i constant vectors. In accordance with our picture of how the mapping is constructed, we should take the diagonal elements of M_0 to be positive and those of M_1 to be negative. There are some further constraints that the M_i and z_i should satisfy to conform to our picture, but since we have already made the example quite special, we will go on and make an even more specific choice which will satisfy all the necessary conditions. The choice will be more special than necessary, but will make it easier to describe the example.

We thus choose a real number α , a contraction factor, with $0 < \alpha < 1/2$. We also choose four more numbers x_0 , x_1 , y_0 , and y_1 such that

$$0 < x_0 < x_0 + \alpha < 1/2 < x_1 - \alpha < x_1 < 1$$

$$0 < y_0 < y_0 + \alpha < 1/2 < y_1 - \alpha < y_1 < 1.$$

We take

$$\Delta_0 = \{(x, y) : y_0 \leq y \leq y_0 + \alpha, 0 \leq x \leq 1\}$$

$$f : (x, y) \mapsto (x_0 + \alpha x, \alpha^{-1}(y - y_0)) \text{ on } \Delta_0$$

$$\Delta_1 = \{(x, y) : y_1 - \alpha \leq y \leq y_1, 0 \leq x \leq 1\}$$

$$f : (x, y) \mapsto (x_1 - \alpha x, \alpha^{-1}(y_1 - y)) \text{ on } \Delta_1.$$

With these choices,

$$f\Delta_0 = \{(x, y) : x_0 \leq x \leq x_0 + \alpha, 0 \leq y \leq 1\}$$

$$f\Delta_1 = \{(x, y) : x_1 - \alpha \leq x \leq x_1, 0 \leq y \leq 1\},$$

i.e., we take Δ_0 and Δ_1 to be horizontal rectangles with height α and unit width; then $f\Delta_0$ and $f\Delta_1$ are vertical rectangles with width α and unit height.

Now consider a point z in the unit square such that $f^n(z)$ is defined, i.e., such that $f(z)$ is in the unit square (so $f^2(z)$ is defined); $f^2(z)$ is also in the unit square (so $f^3(z)$ is also defined); and so on. In order that $f(z)$ be in the unit square, z must be either in Δ_0 or in Δ_1 , and since these are disjoint, it is in one or the other, not both. Similarly, in order that $f^2(z)$ be in the unit square, $f(z)$ must be either in Δ_0 or in Δ_1 , etc. Thus, we can associate with z a sequence $(i_0, i_1, \dots, i_{n-1})$ of 0's and 1's by the rule that $i_j = 0$ or 1 according as $f^j(z)$ is in Δ_0 or Δ_1 . With this as motivation, we define

$$\Delta_{i_0, \dots, i_n}^{(0, n)} = \{z : f^j(z) \in \Delta_{i_j} \text{ for } j = 0, \dots, n\}$$

Here, n can be any positive integer and $i_0, \dots, i_n$ any n -tuple of 0's and 1's. Up to now, it is conceivable that these considerations are empty, i.e., that there are not any z 's such that $f^n(z)$ is defined. This is not true, however:

CLAIM. For any n , and any $i_0, \dots, i_n$, $\Delta_{i_0, \dots, i_n}^{(0, n)}$ is a closed rectangle of height α^{n+1} and unit width.

We prove this by induction on n ; it is clearly true for $n = 0$. For $n > 0$, the definition implies directly that

$$\Delta_{i_0, \dots, i_n}^{(0, n)} = \{z \in \Delta_{i_0} : f(z) \in \Delta_{i_1, \dots, i_n}^{(0, n-1)}\}$$

Also, f is invertible from Δ_{i_0} to its image, and the inverse mapping acts by scaling down by a factor of α in the vertical direction and up by a factor of α^{-1} in the horizontal direction, reflecting if $i_0 = 1$, and translating appropriately. By the induction hypothesis, $\Delta_{i_1, \dots, i_n}^{(0, n-1)}$ is a horizontal rectangle of height α^n running all the way across the unit square; its intersection with $f\Delta_{i_0}$ is thus mapped by f^{-1} to a horizontal rectangle of height α^{n+1} running all the way across the unit square. This proves the induction step and hence the claim.

By the remarks made above, then, the set of z 's such that $f^n(z)$ is defined and in the unit square is

$$\bigcup_{i_0, \dots, i_{n-1}} \Delta_{i_0, \dots, i_{n-1}}^{(0, n-1)},$$

i.e., is a union of 2^n disjoint horizontal rectangles, each of height α^n and unit width.

Since each application of f contracts horizontally by α and expands vertically by α^{-1} ,

$$f^n \Delta_{i_0, \dots, i_{n-1}}^{(0, n-1)}$$

is a vertical rectangle of width α^n and unit height.

For any pair n, m of positive integers, and any sequence $i_{-m}, \dots, i_n$ of 0's and 1's, we define

$$\begin{aligned} \Delta_{i_{-m}, \dots, i_n}^{(-m, n)} &\equiv \{z : f^j(z) \in \Delta_{i_j} \text{ for } j = -m, \dots, n.\} \\ &= (f^m \Delta_{i_{-m}, \dots, i_{-1}}^{(0, m-1)}) \cap \Delta_{i_0, \dots, i_n}^{(0, n)} \end{aligned}$$

The second formula displays $\Delta_{i_{-m}, \dots, i_n}^{(-m, n)}$ as the intersection of a vertical rectangle of width α^m and unit height with a horizontal rectangle of height α^{n+1} and unit width. The result is therefore a rectangle of height α^{n+1} and width α^m .

Many striking consequences follow from this determination of the $\Delta_{i_{-m}, \dots, i_n}^{(-m, n)}$. For example:

CLAIM. *For each sequence*

$$\mathbf{i} = (\dots, i_{-1}, i_0, i_1, \dots) \in \{0, 1\}^{\mathbb{Z}}$$

there is one and only one z such that $f^j(z)$ is defined for all $j \in \mathbb{Z}$ and

$$f^j(z) \in \Delta_{i_j} \text{ for all } j \in \mathbb{Z}.$$

To see this, we fix $\mathbf{i}$ and consider the decreasing sequence of sets

$$\Delta_{i_{-n}, \dots, i_{n-1}}^{(-n, n-1)}.$$

A point z has the desired property if and only if it belongs to the intersection of this sequence of sets. But by what we have just seen, the n -th set in the sequence is a closed square of side α^n , so the intersection contains exactly one point.

We let π denote the mapping which associates with each $\mathbf{i} \in \{0, 1\}^{\mathbb{Z}}$ the corresponding z . For any sequence $\mathbf{i}$, the point $f(\pi \mathbf{i})$ has the defining property of $\pi \sigma \mathbf{i}$, where σ denote the *left shift mapping*:

$$(\sigma \mathbf{i})_j = i_{j+1}.$$

Thus,

$$f \circ \pi = \pi \circ \sigma.$$

Furthermore, π is a homeomorphism from $\{0, 1\}^{\mathbb{Z}}$ equipped with the product topology into the unit square. Since we already know that π is injective, and since $\{0, 1\}^{\mathbb{Z}}$ is compact, we have only to prove that π is continuous. But if $\mathbf{i}$ and $\mathbf{i}'$ are close enough in the product topology so that i_j and i'_j agree for j from $-n$ to $n-1$, then $\pi \mathbf{i}$ and $\pi \mathbf{i}'$ are both in $\Delta_{i_{-n}, \dots, i_{n-1}}^{(-n, n-1)}$, and this is a square of side α^n ; thus, continuity follows.

We can thus sum up the results so far as follows:

There is a homeomorphism π from $\{0,1\}^{\mathbb{Z}}$ onto the set Ξ of points z in the unit square such that $f^n(z)$ is in the unit square for all $n \in \mathbb{Z}$, and

$$f \circ \pi = \pi \circ \sigma,$$

i.e., the action of f on Ξ is topologically conjugate to the left shift σ on $\{0,1\}^{\mathbb{Z}}$.

Hence any statement about the left shift mapping gives a result about the action of f . For example: f has an infinite number of periodic points; the number of periodic points of period p is the number of periodic sequences of 0's and 1's of period p . The periodic points of f are dense in Ξ . (Proof: Given any $i \in \{0,1\}^{\mathbb{Z}}$ and any positive integer p , we can make a periodic sequence by taking the segment of i from $-p$ to $p-1$ and repeating it with period $2p$. By making p large, we can make this sequence as close as we like to i in the product topology.) There are points $z \in \Xi$ whose forward orbits are dense in Ξ . (Proof: Take an enumeration of all the finite sequences of 0's and 1's, and make a one-sided infinite sequence by concatenating all these finite sequences. Extend to negative j arbitrarily. The result is a sequence whose forward orbit under the left shift is dense.)

The technique of using a correspondence between symbol sequences and orbits to analyze orbit structure is known as *symbolic dynamics*; it is one of the principal tools in the analysis of hyperbolic sets. It must be said that it usually doesn't work quite as smoothly as it does in the present case: The more usual situation is that one can construct a continuous mapping from a compact, shift-invariant subset of a sequence space (rather than from the whole sequence space) onto an invariant set, which transforms the shift to the restriction of a smooth mapping to the image set, and which is almost but not quite one-to-one, i.e., for which "most" points in the image set have only one preimage.

We can also get a more geometric picture of the set Ξ . If we let Ξ_n denote the set of points z whose orbits are in the unit square from time $-n$ to time n , then

$$\Xi_n = \bigcup_{i_{-n}, \dots, i_{n-1}} \Delta_{i_{-n}, \dots, i_{n-1}}^{(-n, n-1)}$$

is a union of 4^n squares of side α^n . We pass from n to $n+1$ by putting inside each level- n square 4 squares smaller by a factor of α and deleting what is left over. Thus, Ξ is built by a two-dimensional analogue of the Cantor middle-third construction; it is a self-similar Cantor set.

The set of points z whose *forward* orbits are in the unit square from time 0 to time n is, as noted above, the union of 2^n rectangles, each of height α^n and unit width. The set of points whose forward orbits are in the unit square for all time is a union of horizontal lines of unit length; its intersection with a vertical line is a binary Cantor set K , so the whole set is in $[0,1] \times K$. Each of the horizontal lines is associated with a one-sided sequence $i_0, i_1, \dots$; it is the set of all z such that $f^j(z) \in \Delta_{i_j}$ for all $j \geq 0$. Two points on the same horizontal line have forward orbits which approach each other exponentially

as j goes to $+\infty$. In fact, given a sufficiently small positive η and given any z in one of the horizontal lines, the set of all z' whose forward orbit always stays within a distance η of that of z is a small horizontal line segment. Thus, each $z \in [0, 1] \times K$ admits a local stable manifold.

If we look instead at backward orbits, we find that the set of z 's whose backwards orbits remain in the unit square for all time is $K \times [0, 1]$, i.e., is a union of vertical lines. Two points on the same vertical line have the same "backward symbolic orbit", and have backward orbits which approach each other exponentially. Local unstable manifolds are defined for points of $K \times [0, 1]$; they are small vertical line segments.

With the above notation, we can recognize Ξ as the intersection of the set of points whose forward orbits stay in the unit square ($[0, 1] \times K$) and those whose backwards orbits stay in the unit square ($K \times [0, 1]$); thus,

$$\Xi = K \times K$$

and each point of Ξ has associated stable and unstable manifolds which are respectively horizontal and vertical line segments. Note that, in this instance, the stable and unstable manifolds are not contained in Ξ (which is totally disconnected) but stick out into the ambient space. Nevertheless, each local unstable manifold intersects Ξ in a linear Cantor set.

Mappings at least qualitatively similar to the horseshoe mapping we have considered occur very frequently in applications, since, as Smale discovered, they accompany transverse homoclinic points. We will sketch this argument heuristically. Let $\bar{x}$ be a hyperbolic fixed point for a smooth invertible mapping f . A point x_0 (different from $\bar{x}$) such that both forward and backward orbits of x_0 converge to $\bar{x}$ is called a *homoclinic point*. The condition that the forward orbit converge to $\bar{x}$ can be reexpressed as saying that x_0 is in the (global) stable manifold of $\bar{x}$; similarly, the convergence of the backward orbit means that $\bar{x}$ is in the unstable manifold. We say that x_0 is a *transverse homoclinic point* if the stable and unstable manifolds cross transversally at x_0 . (This is by no means necessary: It is possible, for example, for the stable and unstable manifolds to coincide; it is also possible for the stable and unstable manifolds to be tangent to each other.) There are a number of possibilities for the global geometry of the stable and unstable manifolds of a hyperbolic fixed point. For definiteness we suppose that the situation is as shown in Fig. ???. We also suppose, again for definiteness, that the eigenvalues of the linearization at the fixed point are positive. Now consider a thin rectangular region Δ bounded below by the part of W^s between $\bar{x}$ and x_0 and on the two ends by pieces of W^u . Iterating f leaves the lower right-hand corner of Δ —the fixed point—where it is; stretches out the right-hand edge of Δ , and pulls in the left-hand end of the rectangle. Thus, after some number n_1 of iterations, the image of Δ looks as shown in Fig. ?????. Applying f more times pushes out the image along W^u until finally it reaches and passes x_0 . Thus, if we choose a large n and if we take Δ to have exactly the right height, the image of Δ under f^n will look as shown in Fig. ???: a thin strip running along W^u from the fixed point $\bar{x}$ to x_0 . This looks very much like the picture we drew of the horseshoe mapping; it can be made to look more like it by extending Δ

very slightly on the two ends and the bottom; the result then will look as shown in Fig ????. One can also convince oneself that f^n on Δ is (strongly) contractive horizontally and expansive vertically. Of course, this does not in itself show that the action of f^n is qualitatively what we found for the exactly linear horseshoe mapping. This can in fact be proved by a combination of careful estimates and a qualitative version of the analysis we gave above. We will present later on a different approach leading to the same conclusion. That is, we will prove that, for any sufficiently large n , there is in a small neighborhood of $\{\bar{x}, x_0\}$ a Cantor set Ξ , invariant for f^n , the restriction of f^n to which is topologically conjugate to the left shift on $\{0, 1\}^{\mathbb{Z}}$. Thus, in particular, there are infinitely many periodic points for f , etc.

§3. Definition and elementary properties. For the theory to be developed here, we will assume that the mappings under consideration are at least continuously differentiable with continuously differentiable inverse, and that the space on which they act is finite dimensional. We will denote the state-space by $\mathcal{M}$, and write the formulas for the case where $\mathcal{M}$ is an open subset of a normed vector space $\mathcal{E}$ (and we will, whenever convenient, suppose that $\mathcal{E}$ is simply $\mathbb{R}^m$.) Extension to general manifolds is straightforward.

Let $x \in \mathcal{M}$. We say that a vector ξ in $\mathcal{E}$ (in the general context, a *tangent vector* to $\mathcal{M}$ at x) is a *contracting* or *stable* vector for f at x if there exist positive constants λ and c with $\lambda < 1$ such that

$$(3.1) \quad \|Df^n(x)\xi\| \leq c\lambda^n \|\xi\|$$

for all $n \geq 0$. For any fixed x , the set of contracting vectors at x forms a linear subspace, which we denote by $\mathcal{E}^s(x)$, the *stable* subspace. It is easy to see, by considering a basis for the finite-dimensional space $\mathcal{E}^s(x)$, that there is a single pair of constants $c, \lambda < 1$ such that (3.1) holds for all $\xi \in \mathcal{E}^s(x)$, i.e., that c, λ can be taken to depend only on x , not on ξ .

Similarly, we say that ξ is an *expanding* or *unstable* vector if there exist constants $c, \lambda < 1$, such that

$$\|Df^{-n}(x)\xi\| \leq c\lambda^n \|\xi\|$$

for all $n \geq 0$, i.e., an expanding vector for f means a contracting vector for f^{-1} . The space of expanding vectors will be denoted by $\mathcal{E}^u(x)$. (The notation $Df^{-n}(x)$ is ambiguous; it is to be interpreted here and in what follows as meaning the derivative at x of f^{-n} , which is equal, by the Inverse Function Theorem, to the inverse of the derivative of f^n at $f^n(x)$.)

Note that, contrary to what the name suggests, we do *not* define an expanding vector to be one which is expanded exponentially by $Df^n(x)$. In the first place, it is not always true that expanding vectors are expanded by $Df^n(x)$. For example, if the forward orbit of x converges to a linearly* attracting fixed point for f , and if the backward

*We say that a fixed point x_0 for a differentiable map f is *linearly* attracting if the spectrum of $Df(x_0)$ is strictly inside the unit circle.

orbit converges to a linearly attracting fixed point for f^{-1} , then *every* vector is both expanding and contracting. We will shortly make a more restrictive definition which rules out such counterintuitive situations and guarantees that expanding vectors are indeed expanded by $Df^n(x)$. The problem then is that there are many more vectors expanded by $Df^n(x)$ than just the expanding vectors. For example, in the simple case where x is a hyperbolic fixed point, any vector which has a non-zero component in $\mathcal{E}^u(x)$ is eventually expanded exponentially, not just vectors actually in $\mathcal{E}^u(x)$. This example reflects accurately the general situation; the correct intuition is that expanding vectors are those which are *purely* expansive, i.e., have no contractive component. Paradoxically, although the intuitive motivation of the study of hyperbolicity is the phenomenon of exponential separation of nearby orbits, what turns out to make a detailed analysis possible is the existence of enough pairs of nearby orbits which converge rather than diverge.

It follows easily from the definition that

$$(3.2) \quad Df(x)\mathcal{E}^s(x) = \mathcal{E}^s(f(x)) \quad \text{and} \quad Df^{-1}(x)\mathcal{E}^s(x) = \mathcal{E}^s(f^{-1}(x)),$$

and similarly for $\mathcal{E}^u$. Furthermore, a λ which works at x will also work at $f(x)$ and $f^{-1}(x)$, but it may be necessary to take c larger at $f(x)$ and $f^{-1}(x)$ than at x .

We now make the fundamental definition: A set $\Lambda \subset \mathcal{E}$ is said to be a *hyperbolic set* (for f) if

- (a) Λ is compact and $f\Lambda = \Lambda$ (i.e., Λ is strictly invariant.)
- (b) For each $x \in \Lambda$, $\mathcal{E}^s(x)$ and $\mathcal{E}^u(x)$ are complementary subspaces of $\mathcal{E}$, i.e.,

$$\mathcal{E} = \mathcal{E}^s(x) \oplus \mathcal{E}^u(x).$$

- (c) there exist positive constants $c, \lambda < 1$, *independent of x* such that

$$(3.3) \quad \begin{aligned} \|Df^{+n}(x)\xi\| &\leq c\lambda^n \|\xi\| \quad \text{for all } x \in \Lambda, \xi \in \mathcal{E}^s(x), n \geq 0 \\ \|Df^{-n}(x)\xi\| &\leq c\lambda^n \|\xi\| \quad \text{for all } x \in \Lambda, \xi \in \mathcal{E}^u(x), n \geq 0 \end{aligned}$$

The essential element in this definition is the uniformity of the rate of decay of the contracting (respectively expanding) vectors under $Df^n(x)$ (respectively $Df^{-n}(x)$). It would therefore be more accurate to call such sets Λ *uniformly hyperbolic* rather than simply hyperbolic. A natural way to weaken these strong uniformity hypotheses is to make them *probabilistic*, i.e., to consider, instead of a compact invariant set Λ , a f -invariant probability measure μ , and to replace the above conditions (b) and (c) by the assumption that

$$\mathcal{E} = \mathcal{E}^s(x) \oplus \mathcal{E}^u(x) \quad \mu\text{-almost everywhere.}$$

In the language of *Lyapunov characteristic exponents*, this amounts to requiring that all characteristic exponents be non-zero almost everywhere. The investigation of the extent to which the—by now classical—theory of uniformly hyperbolic sets can be extended

to the more general probabilistic situation is an active subject of research where there still appears to be much to be done.

In addition to the explicit uniformity assumptions of condition (c), there is more uniformity implicit in the requirement that Λ be compact. We will see shortly that this implies, for example, that (intuitively) the angle between $\mathcal{E}^s(x)$ and $\mathcal{E}^u(x)$ is bounded away from zero on Λ . To a large extent, the compactness requirement is imposed simply as a convenient way of formulating many uniformity conditions; most results about hyperbolic sets can be proved with the assumption of compactness replaced by weaker—but messy—explicit uniformity assumptions.

Historically, the investigation of hyperbolic sets began with the case in which the whole space is hyperbolic (if we leave aside the relatively trivial case of a hyperbolic fixed point.) If the whole space is hyperbolic for f , we say that f is an *Anosov diffeomorphism*. The hyperbolic torus automorphism discussed above is an example.

If x is in a hyperbolic set Λ , then any vector $\xi \in \mathcal{E}^u(x)$ really is expanded by $Df^n(x)$. To see this, we write η for $Df^n(x)\xi$ and observe that, by the covariance (3.2) of $\mathcal{E}^u$, $\eta \in \mathcal{E}^u(f^n(x))$. By the Inverse Function Theorem,

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

so, using (3.3),

$$\|\xi\| = \|Df^{-n}(f^n(x))\eta\| \leq c\lambda^n \|\eta\|,$$

We can rewrite this estimate as

$$(3.4) \quad \|Df^n(x)\xi\| \geq c^{-1}\lambda^{-n}\|\xi\| \quad \text{for all } \xi \in \mathcal{E}^u(x)$$

which shows that ξ is indeed expanded by $Df^n(x)$.

It is easy to see from this that the condition

$$\|Df^n(x)\xi\| \leq c\lambda^n \|\xi\| \quad \text{for all } n \geq 0,$$

which by (3.3) holds for all $\xi \in \mathcal{E}^s(x)$, in fact characterizes elements of $\mathcal{E}^s(x)$. Thus, suppose ξ is not in $\mathcal{E}^s(x)$, i.e., has a non-zero component in $\mathcal{E}^u(x)$. Applying (3.4) to the component of ξ in $\mathcal{E}^u(x)$ and (3.3) to the component of ξ in $\mathcal{E}^s(x)$, we see that $\|Df^n(x)\xi\|$ goes to infinity with n , so ξ certainly does not satisfy the above estimate.

We can reformulate this result as the equation

$$\{(x, \xi) : x \in \Lambda, \xi \in \mathcal{E}^s(x)\} = \bigcap_{n \geq 0} \{(x, \xi) : x \in \Lambda, \|Df^n(x)\xi\| \leq c\lambda^n \|\xi\|\}.$$

The right-hand side is an intersection of closed sets in $\Lambda \times \mathcal{E}$, so the left-hand side is closed. Hence:

$$(3.5) \quad \text{If } x_n \rightarrow \hat{x} \text{ and } \xi_n \in \mathcal{E}^s(x_n) \rightarrow \hat{\xi} \text{ then } \hat{\xi} \in \mathcal{E}^s(\hat{x}).$$

The same statement holds with $\mathcal{E}^s$ replaced with $\mathcal{E}^u$.

It turns out that the subspaces $\mathcal{E}^s(x)$ and $\mathcal{E}^u(x)$ vary in a continuous way with x for x in any hyperbolic set. To make precise the intuitive notion of a continuously varying splitting of a vector space into two complementary subspaces, we let $P^s(x)$ denote the projection onto $\mathcal{E}^s(x)$ along $\mathcal{E}^u(x)$, i.e., the linear operator which is the identity on $\mathcal{E}^s(x)$ and zero on $\mathcal{E}^u(x)$; continuity of the splitting will just mean continuity of the operator valued function $x \mapsto P^s(x)$.

PROPOSITION 3.1. *Let Λ be a hyperbolic set for f . Then $x \mapsto P^s(x)$ is continuous on Λ .*

PROOF: We want to show that, if x_n is a sequence in Λ converging to a limit $\hat{x}$, then $P^s(x_n)$ converges to $P^s(\hat{x})$. To prove this, it suffices—by standard compactness considerations—to prove the following two statements:

- (a) If $x_n \rightarrow \hat{x}$ and $P^s(x_n) \rightarrow \hat{P}$ then $\hat{P} = P^s(\hat{x})$.
- (b) $\|P^s(x)\|$ is bounded on Λ .

The first step in proving (a) is to note that $\hat{P}$ is a projection, i.e., that $\hat{P}^2 = \hat{P}$. This follows simply by taking the limit of the corresponding identity for the $P^s(x_n)$. Next, if ξ is in the range of $\hat{P}$, then

$$\xi = \hat{P}\xi = \lim_{n \rightarrow \infty} P^s(x_n)\xi \in \mathcal{E}^s(\hat{x}),$$

by (3.5). Thus, the range of $\hat{P}$ is contained in $\mathcal{E}^s(\hat{x})$. The same argument applied to the sequence $1 - P^s(x_n)$ of projections onto the $\mathcal{E}^u(x_n)$ shows that the range of the projection $1 - \hat{P}$ is contained in $\mathcal{E}^u(\hat{x})$. But since $\mathcal{E}^s(\hat{x})$ and $\mathcal{E}^u(\hat{x})$ are complementary subspaces, it follows from this pair of inclusions that we must have $\hat{P} = P^s(\hat{x})$, as desired.

To prove (b), we assume the contrary, that $\|P^s(x)\|$ is not bounded, i.e., that there exists a sequence x_n of points in Λ and a sequence ξ_n of non-zero vectors such that

$$\frac{\|P^s(x_n)\xi_n\|}{\|\xi_n\|} \rightarrow \infty.$$

We normalize the ξ_n by making $\|P^s(x_n)\xi_n\| = 1$; then $\xi_n \rightarrow 0$. By passing to a subsequence, we can arrange that the x_n converge to a limit $\hat{x}$ and that the $P^s(x_n)\xi_n$ converge to a limit $\hat{\xi}$; from the normalization condition, $\|\hat{\xi}\| = 1$. By (3.5), $\hat{\xi} \in \mathcal{E}^s(\hat{x})$. On the other hand,

$$(1 - P^s(x_n))\xi_n = \xi_n - P^s(x_n)\xi_n \rightarrow -\hat{\xi},$$

since $\xi_n \rightarrow 0$. Hence, by (3.5) with $\mathcal{E}^s$ replaced by $\mathcal{E}^u$, $\hat{\xi} \in \mathcal{E}^u(\hat{x})$, i.e., $\hat{\xi}$ is a non-zero vector in both $\mathcal{E}^s(\hat{x})$ and $\mathcal{E}^u(\hat{x})$. This contradicts the assumption that $\mathcal{E}^s(\hat{x})$ and $\mathcal{E}^u(\hat{x})$ are complementary subspaces, and so proves that $\|P^s(x)\|$ is bounded. The proof of the continuity of $P^s(x)$ is thus complete.

It can be shown, in fact, that $P^s(x)$ is Hölder continuous (provided that $Df(x)$ is.) This is proved by constructing an operator on a space of matrix-valued functions which has P^s as a fixed point; which is contractive in the supremum norm; and which preserves a bound on some Hölder norm. Somewhat surprisingly, it is not possible to do better in general; there exist, for example, analytic Anosov diffeomorphisms for which $P^s(x)$ is nowhere differentiable.

The presence of the constant c in the condition

$$\|Df^n(x)\xi\| \leq c\lambda^n \|\xi\| \quad \text{for all } x \in \Lambda, \xi \in \mathcal{E}^s(x), n \geq 0$$

is a nuisance; it is easier to work with the stronger condition

$$\|Df(x)\xi\| \leq \lambda \|\xi\| \quad \text{for all } x \in \Lambda, \xi \in \mathcal{E}^s(x)$$

(which is equivalent to the preceding condition with $c = 1$.) It is necessary to use the former condition in the definition to make the definition of hyperbolic set invariant under differentiable changes of coordinates. It is possible, however, to make the latter condition hold by adjusting the norm, at the expense, in general, of making the norm x -dependent.

PROPOSITION 3.2. *Let Λ be a hyperbolic set for f . Then there is a constant $\lambda < 1$ and a Riemannian metric for $\mathcal{M}$ (i.e., an x -dependent positive definite scalar product $(\cdot, \cdot)_x$ with associated norm $\|\xi\|_x = [(\xi, \xi)_x]^{1/2}$) such that, for all $x \in \Lambda$,*

$$(3.6) \quad \begin{aligned} \|Df(x)\xi\|_{f(x)} &\leq \lambda \|\xi\|_x && \text{for } \xi \in \mathcal{E}^s(x), \\ \|Df^{-1}(x)\xi\|_{f^{-1}(x)} &\leq \lambda \|\xi\|_x && \text{for } \xi \in \mathcal{E}^u(x). \end{aligned}$$

Such a metric is called an adapted or Lyapunov metric. To keep things elementary and avoid complicating the notation unnecessarily, we will consider henceforth only the case in which (3.6) holds with the standard (x -independent) norm.

PROOF: The main idea of the proof is essentially the same as that used to show that, by adjusting the norm, the norm of an operator can be made as close as we want to its spectral radius. This part of the construction will nevertheless differ in a detail from the one we gave in §?? because we want to construct a Hilbert norm, i.e., one given by a scalar product, rather than a general one. We start from (3.3):

$$(3.3) \quad \begin{aligned} \|Df^{+n}(x)\xi\| &\leq c\lambda_1^n \|\xi\| && \text{for all } x \in \Lambda, \xi \in \mathcal{E}^s(x), n \geq 0 \\ \|Df^{-n}(x)\xi\| &\leq c\lambda_1^n \|\xi\| && \text{for all } x \in \Lambda, \xi \in \mathcal{E}^u(x), n \geq 0, \end{aligned}$$

where we have written λ_1 rather than λ to distinguish the contraction constant with general c from the contraction constant with $c = 1$. The λ of the Proposition can in fact be taken to be anything strictly larger than λ_1 .

Thus, we start by choosing a λ with $\lambda_1 < \lambda < 1$; then a λ_2 with $\lambda_1 < \lambda_2 < \lambda$. We construct a first approximation $\langle, \rangle_x$ to the desired x -dependent inner product by

$$\begin{aligned} \langle \xi, \eta \rangle_x &= 0 \quad \text{if } \xi \in \mathcal{E}^s(x) \text{ and } \eta \in \mathcal{E}^u(x) \\ &= \sum_{n=0}^{\infty} \frac{(Df^n(x)\xi, Df^n(x)\eta)}{\lambda_2^{2n}} \quad \text{if } \xi, \eta \in \mathcal{E}^s(x), \\ &= \sum_{n=0}^{\infty} \frac{(Df^{-n}(x)\xi, Df^{-n}(x)\eta)}{\lambda_2^{2n}} \quad \text{if } \xi, \eta \in \mathcal{E}^u(x). \end{aligned}$$

We make this construction only for $x \in \Lambda$. The convergence of the infinite sums, the continuity in x of the scalar product so defined, and the fact that this scalar product is uniformly bounded above and below by multiples of the standard scalar product all follow easily from (3.3) (and the inequality $\lambda_2 > \lambda_1$.) By a straightforward computation,

$$\langle Df(x)\xi, Df(x)\xi \rangle_{f(x)} \leq (\lambda_2)^2 \langle \xi, \xi \rangle_x \quad \text{for } \xi \in \mathcal{E}^s(x);$$

$$\langle Df^{-1}(x)\xi, Df^{-1}(x)\xi \rangle_{f^{-1}(x)} \leq (\lambda_2)^2 \langle \xi, \xi \rangle_x \quad \text{for } \xi \in \mathcal{E}^u(x).$$

There are two things wrong with the inner product $\langle, \rangle_x$: It is only defined for $x \in \Lambda$, not for all $x \in \mathcal{M}$, and it is only continuous in x , not smooth. To deal with these problems, we first extend the inner product to all of $\mathcal{M}$, then approximate the extension by a smooth one. If the smoothed inner product is close enough in the unsmoothed one in the supremum norm, the desired estimates (3.6) hold.

To show that the inner product can be extended continuously to all of $\mathcal{M}$, we can argue as follows: Write

$$\langle \xi, \eta \rangle_x = (R_x \xi, \eta)$$

with R_x a positive-definite matrix. For some $\epsilon > 0$,

$$R_x \geq \epsilon 1 \quad \text{for all } x \in \Lambda.$$

We can therefore write

$$R_x = \epsilon 1 + S_x^* S_x,$$

with S_x a continuous matrix-valued function (by taking, for example, S_x as the unique positive square root of $R_x - \epsilon 1$). Now extend S_x to a continuous bounded matrix-valued function defined on all of $\mathcal{M}$ by applying Tietze's Theorem to each of its entries separately; then extend R_x by using the formula $R_x = \epsilon 1 + S_x^* S_x$, with the extended S , on all of $\mathcal{M}$. The approximation by a smooth metric is easy: Just approximate each of the entries in the matrix-valued function R_x by a smooth function in such a way as to preserve symmetry; if the approximation is close enough in the supremum norm, the smooth approximating function will still be positive-definite for each x .

§4. Transverse homoclinic orbits.

We saw, in the first section of this chapter, two examples of hyperbolic sets. Hyperbolic fixed points and periodic cycles are also hyperbolic sets, albeit of a comparatively

trivial kind. We will present in this section another class of examples, with a rather different flavor.

Let $\bar{x}$ be a hyperbolic fixed point of saddle type* for f ; x_0 a transverse homoclinic point, i.e., a point of $W^s(\bar{x}) \cap W^u(\bar{x})$ such that $W^s(\bar{x})$ and $W^u(\bar{x})$ intersect transversally at x_0 . Write x_n for $f^n(x_0)$, $n \in \mathbb{Z}$. Then x_n converges to $\bar{x}$ both as n goes to $+\infty$ and as n goes to $-\infty$, so the set Λ consisting of the homoclinic orbit $\{x_n : n \in \mathbb{Z}\}$ together with the fixed point $\bar{x}$ is compact; it is also evidently invariant. We claim that it is a hyperbolic set.

To prove this, what we need to do is to identify $\mathcal{E}^s(x_n)$ and $\mathcal{E}^u(x_n)$ for each n ; to show that they are complementary subspaces; and to prove uniformity of contraction under the action of Df and Df^{-1} respectively. (The spaces of contracting and expanding vectors at the fixed point $\bar{x}$ have the desired properties by elementary spectral theory.) By the invariance of the global stable and unstable manifolds $W^s(\bar{x})$ and $W^u(\bar{x})$, each x_n is in $W^s(\bar{x}) \cap W^u(\bar{x})$. Also by invariance, $Df^n(x_0)$ carries the tangent spaces of $W^s(\bar{x})$ and $W^u(\bar{x})$ to the respective tangent spaces at x_n ; hence, $W^s(\bar{x})$ and $W^u(\bar{x})$ intersect transversally at each x_n .

CLAIM. *The space of contracting (expanding) vectors at x_n is exactly the tangent space to $W^s(\bar{x})$ ($W^u(\bar{x})$) at x_n .*

It suffices to consider contracting vectors, since expanding vectors are simply contracting vectors for f^{-1} . Furthermore, by covariance, it suffices to prove the assertion for large positive n , i.e., values of n such that x_n is in a small local stable manifold for f at $\bar{x}$. Thus, the claim reduces to a statement about the local behavior of a continuously differentiable mapping on a small neighborhood of a hyperbolic fixed point. We set up the problem as in Chapter IV: We put the fixed point $\bar{x}$ at the origin, and we write $\mathcal{E} = \mathcal{E}^s \oplus \mathcal{E}^u$, where $\mathcal{E}^s$ and $\mathcal{E}^u$ are respectively the contracting and expanding subspaces for $Df(0)$. We write Λ_s (Λ_u) for the restriction of $Df(0)$ to $\mathcal{E}^s$ ($\mathcal{E}^u$); we choose norms on $\mathcal{E}^s$ and $\mathcal{E}^u$ so that $\|\Lambda_s\| < 1$ and $\|\Lambda_u^{-1}\| < 1$; and we use the norm

$$\|(\xi, \eta)\| = \sup\{\|\xi\|, \|\eta\|\}$$

on $\mathcal{E}$. Furthermore, as we saw in §??, we can assume that $W^s(0)$ is simply $\mathcal{E}^s$, at least on a small enough neighborhood of 0. Then for x on the stable manifold and close enough to 0, the derivative of f at x , expressed as a 2×2 matrix relative to the splitting $\mathcal{E} = \mathcal{E}^s \oplus \mathcal{E}^u$, takes the form

$$\begin{pmatrix} \alpha(x) & \beta(x) \\ 0 & \delta(x) \end{pmatrix},$$

where

$$\alpha(x) \approx \Lambda_s, \quad \delta(x) \approx \Lambda_u, \quad \beta(x) \approx 0.$$

Consider now a tangent vector to W^s at x_n , i.e., a vector of the form $(\xi, 0)$. Then

$$Df^j(x_n)(\xi, 0) = (\alpha(x_{n+j-1})\alpha(x_{n+j-2}) \dots \alpha(x_n)\xi, 0).$$

*i.e., whose stable and unstable manifolds both have dimension greater than zero.

If we take n large enough, we can arrange that

$$\|\alpha(x_m)\| \leq \lambda < 1 \quad \text{for all } m \geq n,$$

and hence

$$\|Df^j(x_n)(\xi, 0)\| \leq \lambda^j \|(\xi, 0)\|.$$

Here, λ can be any number smaller than one and larger than $\|\Lambda_s\|$. This shows that every tangent vector to W^s is a contracting vector for f ; it also proves the uniformity of contraction of $Df(x_n)$ on this tangent space for large (positive) n .

We also have to show that any vector which is *not* tangent to W^s at x_n is not a contracting vector. Any such vector ζ can be written as (ξ, η) with $\eta \neq 0$, so the $\mathcal{E}^u$ -component of $Df^j(x_n)\zeta$ is equal to

$$\delta(x_{n+j-1})\delta(x_{n+j-2})\dots\delta(x_n)\eta.$$

By taking n large enough, we can arrange that

$$\|\delta(x_m)^{-1}\| \leq \lambda < 1 \quad \text{for all } m \geq n,$$

and hence

$$\|Df^j(x_n)\zeta\| \geq \lambda^{-j}\|\eta\|,$$

which shows that ζ cannot be a contracting vector. (This time, λ has to be larger than $\|\Lambda_u^{-1}\|$.)

To prove that the homoclinic orbit is a hyperbolic set, it remains to prove that $Df^j(x_n)$ is sufficiently uniformly contractive on the tangent space of $W^s(\bar{x})$ at x_n . It will suffice to prove that, for some $\lambda < 1$ and the norm introduced above,

$$\|Df(x_{\pm n})\xi\| \leq \lambda\|\xi\|,$$

for all tangent vectors ξ to $W^s(\bar{x})$ at $x_{\pm n}$, whenever n is sufficiently large. (Failure of contractivity for a finite number of n 's can be compensated for by choosing the constant c sufficiently large.) We have already dealt with the case of large positive n . The point x_{-n} , for large n , is again near $\bar{x}$, and the problem is again in part reduced to a local question near the fixed point. This time, however, the x_{-n} belong to a local *unstable* manifold, and, although near to the fixed point and on the stable manifold, they are not near to the fixed point in the stable manifold, i.e., a path in $W^s(\bar{x})$ running from the fixed point to x_{-n} cannot be short. Thus, we cannot directly use what we know about the local stable manifold to determine what the tangent space to $W^s(\bar{x})$ at x_{-n} looks like.

CLAIM. *There exists a sequence of positive constants K_{-n} , defined for n sufficiently large and converging to zero, such that, if (ξ, η) is in the tangent space to $W^s(\bar{x})$ at x_{-n} , then*

$$\|\eta\| \leq K_{-n}\|\xi\|.$$

In other words: The tangent space to $W^s(\bar{x})$ at x_{-n} converges to the contracting subspace at $\bar{x}$. This result is part of J. Palis' λ -Lemma. Once the claim is proved, the desired contractivity follows easily, since, for large n ,

$$Df(x_{-n}) \approx \begin{pmatrix} \Lambda_s & 0 \\ 0 & \Lambda_u \end{pmatrix}$$

To prove this claim, we use the same setup and special notation that we used in proving the preceding one; we also arrange to make a small enough local unstable manifold coincide with $\mathcal{E}^u$. We choose n_0 large enough so that x_{-n} is in this local unstable manifold for all $n \geq n_0$. We will assume for the remainder of the proof that $n \geq n_0$. Since $W^s(\bar{x})$ intersects $\mathcal{E}^u = W^u(\bar{x})$ transversally at x_{-n_0} , there is a constant K_{-n_0} such that

$$\|\eta\| \leq K_{-n_0} \|\xi\| \quad \text{for } (\xi, \eta) \text{ tangent to } W^s(\bar{x}) \text{ at } x_{-n_0}.$$

This, together with the fact that $Df^{-1}(x_{-n})$ maps the tangent space to $W^s(\bar{x})$ at x_{-n} to the tangent space at x_{-n-1} , is all we are going to assume about these tangent spaces in the proof of the claim; the idea will be to use what we know about $Df^{-1}(x_{-n})$ to construct inductively a sequence K_{-n} with the desired properties.

From the local invariance of $\mathcal{E}^u$ under f^{-1} , we see that for $x \in \mathcal{E}^u$ and near enough to the fixed point, $Df^{-1}(x)$ takes the form

$$\begin{pmatrix} \alpha^{-1}(x) & 0 \\ \gamma(x) & \delta^{-1}(x) \end{pmatrix},$$

where

$$\alpha(x) \approx \Lambda_s, \quad \delta(x) \approx \Lambda_u, \quad \gamma(x) \approx 0.$$

Let $\zeta = (\xi, \eta)$ satisfy $\|\eta\| \leq K_{-n} \|\xi\|$, and write $(\bar{\xi}, \bar{\eta})$ for $Df^{-1}(x_{-n})\zeta$. Then

$$\|\bar{\eta}\| = \|\gamma(x_{-n})\xi + \delta^{-1}(x_{-n})\eta\| \leq [\|\gamma(x_{-n})\| + K_{-n}\|\delta^{-1}(x_{-n})\|]\|\xi\|$$

and

$$\|\bar{\xi}\| = \|\alpha(x_{-n})\bar{\xi}\| \leq \|\alpha(x_{-n})\| \|\bar{\xi}\|.$$

Hence

$$\|\bar{\eta}\| \leq K_{-(n+1)} \|\bar{\xi}\|$$

where

$$K_{-(n+1)} = \|\alpha(x_{-n})\| \|\delta^{-1}(x_{-n})\| K_{-n} + \|\alpha(x_{-n})\| \|\gamma(x_{-n})\|.$$

From this recursion relation and

$$\|\alpha(x_{-n})\| \longrightarrow \|\Lambda_s\| < 1,$$

$$\|\delta^{-1}(x_{-n})\| \longrightarrow \|\Lambda_u^{-1}\| < 1,$$

$$\|\gamma(x_{-n})\| \longrightarrow 0,$$

it follows that

$$K_{-n} \longrightarrow 0,$$

as desired.

This completes the proof that a transverse homoclinic orbit, together with the corresponding fixed point, forms a hyperbolic set. This set is quite different from, e.g., the horseshoe. For example, periodic points are dense in the horseshoe but certainly not in this example. On the other hand, the horseshoe, as constructed, is an *isolated* hyperbolic set in the sense that every orbit near but not on it eventually escapes. This is not at all the case for a transverse homoclinic orbit, although it is not yet evident. We will see shortly, for example, that every point on a transverse homoclinic orbit is a limit of hyperbolic periodic points. Intuitively, it will turn out that a transverse homoclinic orbit is a small piece of a larger hyperbolic set while the horseshoe is not. This will be made precise later when we introduce the notion of *locally maximal* hyperbolic set.

CHAPTER IX. SHADOWING

§1. Formulation of the principal result. Let $\delta > 0$. A δ *pseudo-orbit* for f means a sequence of points (x_i) in $\mathcal{M}$, indexed by $\mathbf{Z}$ or by a subinterval of $\mathbf{Z}$, such that

$$\|x_{i+1} - f(x_i)\| \leq \delta \quad \text{for all (relevant) } i.$$

In other words, x_{i+1} is obtained from x_i by applying f ; then making a jump which has length less than δ but is otherwise arbitrary.

One of the keys to the analysis of hyperbolic sets is the fact that, if δ is small enough, every δ pseudo-orbit near a hyperbolic set is followed closely by a true orbit. The precise statement is as follows:

THEOREM 1.1 (SHADOWING THEOREM). *Let Λ be a hyperbolic set for f , and let $\epsilon > 0$. Then there exists $\delta > 0$ such that, for every δ pseudo-orbit (x_i) each point of which is within distance δ of Λ , there is a y_0 such that*

$$(1.1) \quad \|f^i(y_0) - x_i\| \leq \epsilon \quad \text{for all } i.$$

Furthermore, if ϵ is small enough and if the pseudo-orbit is indexed by all of $\mathbf{Z}$, then y_0 is uniquely determined.

If (1.1) holds, one says that the orbit $(f^i(y_0))$ ϵ -*shadows* the pseudo-orbit (x_i) . It is not difficult to see that it does not weaken the theorem significantly to assume that the pseudo-orbit (x_i) is actually in Λ rather than simply *near* Λ , since any pseudo-orbit near Λ can be approximated by a pseudo-orbit in Λ , possibly with a slightly larger δ . Once the pseudo-orbit is taken to be in Λ , there is no loss of generality in assuming it to be indexed by all of $\mathbf{Z}$, since a pseudo-orbit in Λ indexed by a segment of $\mathbf{Z}$ can be extended to all of $\mathbf{Z}$ without increasing δ . On the other hand, it is a delicate matter whether the shadowing orbit will be in Λ or not; we will return to this question later.

Before giving the proof of the Shadowing Theorem, we will indicate some of its uses. One immediate consequence of the uniqueness assertion is:

COROLLARY 1.2. *Let Λ be a hyperbolic set for f . There exists $\eta > 0$ such that, if $f^i(x_0)$ and $f^i(\bar{x}_0)$ are two distinct orbits one of which stays within a distance η of Λ for all i , then*

$$\|f^i(x_0) - f^i(\bar{x}_0)\| > \eta \quad \text{for some } i.$$

In other words, two distinct orbits near Λ cannot stay arbitrarily close together for all time. This property is called *expansivity*; a number η with the indicated property is called an *expansivity constant*.

§2. Embedded shifts near homoclinic orbits. A typical use of the Shadowing Theorem is to prove the existence of periodic orbits. The idea is to construct periodic pseudo-orbits with sufficiently small δ so that they have unique shadowing orbits; then the shadowing orbits must also be periodic. For a concrete example, consider the closure of a transverse homoclinic orbit, i.e., a set Λ consisting of a hyperbolic fixed point $\bar{x}$ together with $\{f^i(x_0)\}$, x_0 a point where $W^s(\bar{x})$ and $W^u(\bar{x})$ cross transversally.

PROPOSITION 2.1. *Let $\bar{x}$ be a hyperbolic fixed point for the C^1 diffeomorphism f ; x_0 a point of transverse intersection of $W^s(\bar{x})$ and $W^u(\bar{x})$. Then x_0 is a limit of periodic points for f .*

PROOF: Let Λ be as above, i.e., the orbit of x_0 together with the fixed point $\bar{x}$. For N a large integer, we construct a sequence (z_n) in Λ by putting $z_n = x_n$ for $-N + 1 \leq n \leq N$ and extending to all of $\mathbb{Z}$ by periodicity, i.e., by requiring $z_{n+2N} = z_n$. Thus, $z_{n+1} = f(z_n)$ except for n an odd multiple of N , in which case $f(z_n) = x_{N+1}$ and $z_{n+1} = x_{-N+1}$. Since x_{N+1} and x_{-N+1} can both be made as close as desired to $\bar{x}$ by making N large, we can make $\sup_n \|z_{n+1} - f(z_n)\|$ as small as we like by taking N large.

Now take any $\epsilon > 0$ small enough so that the uniqueness statement in the Shadowing Theorem holds; then take N large enough so that the sequence (z_n) constructed in the preceding paragraph is a δ pseudo-orbit for the δ corresponding to ϵ . The Shadowing Theorem then says that there is a unique exact orbit (y_n) which ϵ -shadows (z_n) , i.e., such that $\|z_n - y_n\| < \epsilon$ for all n . Because $z_{n+2N} = z_n$, (y_{n+2N}) is another ϵ -shadowing orbit, so by uniqueness $y_n = y_{n+2N}$ for all n , so y_0 is a periodic point for f . Because of ϵ -shadowing,

$$\|y_0 - x_0\| = \|y_0 - z_0\| \leq \epsilon.$$

We can extend the above argument to prove the following:

PROPOSITION 2.2. *Let $\epsilon > 0$. For any sufficiently large N , there is a compact set Ξ contained in the union of the balls of radius ϵ about $\bar{x}$ and x_0 , invariant for f^N , such that f^N on Ξ is topologically conjugate to the left shift on $\{0, 1\}^{\mathbb{Z}}$.*

PROOF: We can assume that ϵ is small enough to guarantee uniqueness of ϵ -shadowing orbits near Λ , and let δ correspond to ϵ as in the Shadowing Theorem. Fix a large N , and construct four sequences of length N in Λ as follows:

$$X_{(0,0)} = (x_0, x_1, \dots, x_m, x_{-N+m+1}, x_{-N+m+2}, \dots, x_{-1}),$$

where m is the integer part of $N/2$;

$$X_{(0,1)} = (x_0, x_1, \dots, x_{N-2}, \bar{x});$$

$$X_{(1,0)} = (\bar{x}, x_{-N+2}, \dots, x_{-1});$$

$$X_{(1,1)} = (\bar{x}, \bar{x}, \dots, \bar{x}).$$

If we take N large enough, each of these is a δ pseudo-orbit.

For each sequence $i \in \{0, 1\}^{\mathbb{Z}}$, we construct an infinite sequence $z(i)$ in Λ by concatenating the above sequences in an order determined by i :

$$z(i) = \dots X_{(i_{-1}, i_0)} X_{(i_0, i_1)} X_{(i_1, i_2)} \dots$$

(with the origin placed so that z_{Nn} is x_0 or $\bar{x}$ according as i_n is 0 or 1.) Again, each $z(i)$ is a δ pseudo-orbit. By the Shadowing Theorem, there is a unique exact orbit

ϵ -shadowing $z(i)$. We let πi denote the position of this orbit at time zero, i.e., the shadowing orbit is $f^n(\pi i)$. The correspondence $i \mapsto z$ is constructed in such a way as to make

$$z(\sigma i) = f^N z(i),$$

so by the uniqueness of the shadowing orbit

$$f^N \circ \pi = \pi \circ \sigma.$$

Since $z(i)_{nN}$ is x_0 or $\bar{x}$ according as i_n is 0 or 1, $f^{nN}(\pi i)$ is in the ball of radius ϵ about x_0 or $\bar{x}$ according as i_n is 0 or 1. If ϵ is small enough so that these balls do not overlap, π is injective.

Thus, if we can prove that π is continuous, it will follow that it is a homeomorphism, and this will complete the proof. The proof of continuity uses a device which is frequently useful: We deduce continuity from the uniqueness of the shadowing orbit. Let $i^{(n)}$ be a sequence converging to $i^{(\infty)}$. We need only show that the sequence $\pi i^{(n)}$ cannot have any accumulation point other than $\pi i^{(\infty)}$. But it follows almost immediately from the definitions that the orbit of any accumulation point of $\pi i^{(n)}$ ϵ -shadows $z(i^{(\infty)})$, so, by the uniqueness of the shadowing orbit, the only possible accumulation point is $\pi(i^{(\infty)})$.

§3. Proof of the Shadowing Theorem.

As noted above, it is enough to prove the Shadowing Theorem for pseudo-orbits indexed by $\mathbb{Z}$. The proof we are going to give goes by reducing the Shadowing Theorem to a fixed point problem for a differentiable operator in a Banach space; then applying an approximate Newton's method. Schematically, what we do is to define an operator $\mathcal{A}$ on the space $\mathcal{X}$ of all bounded sequences (z_n) in $\mathcal{M}$ by

$$\mathcal{A}(z)_n = f(z_{n-1}).$$

A sequence z is a fixed point for $\mathcal{A}$ if and only if it is an orbit of f . Similarly, equipping $\mathcal{X}$ with the supremum norm, z is a δ pseudo-orbit for f if and only if

$$\delta \geq \sup_n \|z_{n+1} - f(z_n)\| = \|z - \mathcal{A}z\|,$$

i.e., a pseudo-orbit with small δ is the same as a sequence z such that $\mathcal{A}(z) \approx z$. Thus, in this language, what the Shadowing Theorem says is that near every point moved sufficiently little by $\mathcal{A}$ there is an exact fixed point.

To prove this, we are going to use Proposition 2.5 of Chapter III, which we reformulate in our present notation as follows:

PROPOSITION 3.1. *Let $B_\rho(z)$ denote the closed ball in the Banach space $\mathcal{X}$ with center z and radius ρ , and let $\mathcal{A}$ be differentiable mapping from $B_\rho(z)$ into $\mathcal{X}$. If there exist a linear operator Γ on $\mathcal{X}$ such that $\Gamma - 1$ is invertible with bounded inverse and a positive number $\kappa < 1$ such that*

- (1) $\|(\Gamma - 1)^{-1}(D\mathcal{A}(x) - \Gamma)\| \leq \kappa$ for all $x \in B_\rho(z)$
- (2) $\|(\Gamma - 1)^{-1}(\mathcal{A}(z) - z)\| \leq (1 - \kappa)\rho$

then $\mathcal{A}$ has exactly one fixed point in $B_\rho(z)$.

It is easy to see that the operator $\mathcal{A}$ as defined above is differentiable and that

$$(3.1) \quad (D\mathcal{A}(x)\eta)_n = Df(x_{n-1})\eta_{n-1}.$$

To be entirely precise, we have to be a little careful about the domains: The exact statement is that $\mathcal{A}$ is differentiable at x provided that Df is uniformly continuous on some neighborhood of the closure of the point set $\{x_n\}$. To deal with this and a number of similar technicalities, we will fix once and for all an open neighborhood of Λ with compact closure and consider only sequences taking values in this neighborhood.

Our strategy for applying the above proposition will be as follows: We let z be the pseudo-orbit for which we want to find a shadowing orbit, and we take for Γ simply $D\mathcal{A}(\bar{z})$. Most of the work in the proof is in showing that $\Gamma - 1$ has bounded inverse. Let us admit this for the moment and see how the rest of the argument goes. We can take any κ we like, and we choose $\kappa = 1/2$. From (3.1) it is immediate that

$$\|D\mathcal{A}(x) - D\mathcal{A}(z)\| = \sup_n \|Df(x_n) - Df(z_n)\|.$$

Using the uniform continuity of $Df(x)$ on a neighborhood of Λ , we can find a ρ small enough so that

$$(3.2) \quad \|Df(x) - Df(z)\| \leq \frac{1}{2\|(\Gamma - 1)^{-1}\|}$$

whenever $\|x - z\| \leq \rho$ and z is within distance ρ of Λ . Then condition (1) of the fixed point theorem holds (with, as indicated, $\kappa = 1/2$). Thus, in order to ensure that there is a unique orbit (y_n) with $\|y_n - z_n\| \leq \rho$, we have only to require that

$$\|z_n - f(z_{n-1})\| \leq \frac{\rho}{2\|(\Gamma - 1)^{-1}\|} \quad \text{for all } n,$$

in order to make condition (2) hold.

Although the construction is done separately for each pseudo-orbit, it is essential that all estimates be uniform in the pseudo-orbit. Examination of the argument in the preceding paragraph in this perspective shows that the Shadowing Theorem will follow if we prove

LEMMA 3.2. *Let Λ be a hyperbolic set for the C^1 diffeomorphism f . There exist positive constants ρ and B such that, if z is any ρ pseudo-orbit staying within distance ρ of Λ , and if Γ denotes the operator defined by*

$$(\Gamma\eta)_n = Df(z_{n-1})\eta_{n-1},$$

on the space $\mathcal{X}$ of bounded sequences in $\mathcal{E}$ equipped with the supremum norm, then $\Gamma - 1$ is invertible and

$$\|(\Gamma - 1)^{-1}\| \leq B.$$

Furthermore, we can make the estimates of the Shadowing Theorem more explicit; the lemma together with the argument above proves:

PROPOSITION 3.3 (SHADOWING THEOREM WITH EXPLICIT ESTIMATES). *Let Λ be a hyperbolic set for the C^1 diffeomorphism f , and let ρ, B be as in Lemma 3.2. Suppose also that ρ is small enough so that (3.2) holds. Then two distinct orbits one of which stays within distance ρ of Λ are at some time a distance greater than ρ apart. Also, for any $\epsilon \leq \rho$, we let*

$$\delta(\epsilon) = \frac{\epsilon}{2B};$$

Then any δ pseudo-orbit (defined for all n in $\mathbb{Z}$) staying within distance ρ of Λ is ϵ -shadowed by a true orbit.

The relation between ϵ and δ can thus be taken to be a linear one, with constants of reasonable size.

PROOF OF LEMMA 3.2: While the proof is not difficult, it is necessary to deal with a certain number of technical complications. To show what the main idea is, with a minimum of distracting detail, we will consider first the particularly simple case where z is an exact orbit in Λ . As already indicated, we denote by $\mathcal{X}$ the space of bounded sequences ξ_n in $\mathcal{E}$ equipped with the supremum norm. We write $\mathcal{X}^s$ ($\mathcal{X}^u$) for the subspace of sequences such that ξ_n is in $\mathcal{E}^s$ ($\mathcal{E}^u$) for all n . Then $\mathcal{X}^s$ and $\mathcal{X}^u$ are closed subspaces of $\mathcal{X}$, invariant for Γ , and $\mathcal{X} = \mathcal{X}^s \oplus \mathcal{X}^u$. From the estimate

$$\|Df(x)\xi\| \leq \lambda\|\xi\| \quad \text{for all } \xi \in \mathcal{E}^s(x), \text{ for all } x \in \Lambda,$$

it follows that the norm of the restriction of Γ to $\mathcal{X}^s$ is not greater than λ . Hence, the Neumann series expansion

$$(1 - \Gamma)^{-1} = \sum_{i=0}^{\infty} \Gamma^i$$

converges in norm on $\mathcal{X}^s$, so the restriction $\Gamma - 1$ to $\mathcal{X}^s$ is invertible and the norm of its inverse is bounded by $(1 - \lambda)^{-1}$. Similarly, the restriction of Γ to $\mathcal{X}^u$ is invertible and its inverse has norm not greater than λ . Thus, the formal expansion

$$(\Gamma - 1)^{-1} = ((1 - \Gamma^{-1})\Gamma)^{-1} = \Gamma^{-1} \sum_{i=0}^{\infty} \Gamma^{-i}$$

converges in norm on $\mathcal{X}^u$, so the restriction of $\Gamma - 1$ to $\mathcal{X}^u$ is invertible and its inverse has norm bounded by $\lambda/(1 - \lambda) < (1 - \lambda)^{-1}$.

Since $\Gamma - 1$ is invertible on each of two complementary invariant subspaces of $\mathcal{X}$, it is invertible on $\mathcal{X}$ itself. We have z -independent estimates for the norm of the inverse restricted to each of these subspaces; to get such an estimate on the whole space, we need to pay attention to how the norm on $\mathcal{X}$ is related to its restrictions to $\mathcal{X}^s$ and $\mathcal{X}^u$. We let (as above) $P^s(x)$ denote the projection onto $\mathcal{E}^s(x)$ along $\mathcal{E}^u(x)$ and $P^u(x)$ denote $1 - P^s(x)$. By the continuity of $P^s(x)$, there is a constant D such that

$$\|P^s(x)\| \leq D \quad \text{and} \quad \|P^u(x)\| \leq D \quad \text{for all } x \in \Lambda.$$

The projection P^s onto $\mathcal{X}^s$ along $\mathcal{X}^u$ acts "componentwise":

$$(P^s\xi)_n = P^s(z_n)\xi_n,$$

so $\|P^s\| \leq D$ and similarly $\|P^u\| \leq D$, where P^u denotes $1 - P^s$. Hence, for any $\xi \in \mathcal{X}^s$,

$$\|\xi\| \leq 2 \sup\{\|P^s\xi\|, \|P^u\xi\|\} \leq D\|\xi\|.$$

From this estimate together with the fact that the restrictions of $(\Gamma - 1)^{-1}$ to $\mathcal{X}^s$ and $\mathcal{X}^u$ both have norm not larger than $(1 - \lambda)^{-1}$ it follows easily that

$$\|(\Gamma - 1)^{-1}\| \leq \frac{2D}{1 - \lambda},$$

which is the desired z -independent estimate.

The proof of the lemma in the general case goes by constructing an approximation Γ_0 to Γ to which the above argument can be applied and then using operator-theoretic perturbation theory to estimate $(\Gamma - 1)^{-1}$ in terms of $(\Gamma_0 - 1)^{-1}$. Everything needs to be done in a very systematic way in order to ensure that the estimates obtained are uniform in the pseudo-orbit. The first step in constructing Γ_0 is

LEMMA 3.4. *The projection-valued function $P^s(x)$ can be extended to a continuous projection-valued function defined on a neighborhood of Λ*

The proof of this lemma will be given later. We choose one such extension, which we will continue to denote by $P^s(x)$, and we define an extended $P^u(x)$ by $1 - P^s(x)$. Given z and z' near enough to Λ , we write

$$\Gamma_z^z, \quad \text{for} \quad P^s(z')Df(z)P^s(z) + P^u(z')Df(z)P^u(z).$$

Given a pseudo-orbit z , we are going to define Γ_0 by

$$(\Gamma_0\xi)_n = \Gamma_{z_n}^{z_{n-1}}\xi_{n-1}.$$

Note that on the one hand, by its very construction, Γ_z^z , maps the range of $P^s(z)$ to the range of $P^s(z')$ and the range of $P^u(z)$ to $P^u(z')$, but that, on the other hand, if $z \in \Lambda$ and $z' = f(z)$, then by the covariance of $\mathcal{E}^s(x)$ and $\mathcal{E}^u(x)$, Γ_z^z is equal to $Df(z)$. Hence, by uniform continuity, we can make Γ_z^z , as close as we like to $Df(z)$ by making the distances from z to Λ and from z' to $f(z)$ small. Expressed in terms of pseudo-orbits, this says that, by making ρ small enough, we can guarantee that $\|\Gamma - \Gamma_0\|$ is as small as we like for all ρ pseudo-orbits z staying within distance ρ of Λ .

Furthermore, since the norm of $Df(x)$ on the range of $P^s(x)$ is majorized by λ for $x \in \Lambda$, we can guarantee that the norm of Γ_z^z , on the range of $P^s(z)$ is not much bigger than λ by requiring that the distances from z to Λ and from $f(z)$ to z' be small enough. A similar statement holds for P^u .

We can now begin to collect a set of conditions on ρ which, taken together, will make the lemma hold. We start by choosing ρ small enough so that $P^s(z)$ is defined, uniformly continuous, and bounded on the set of all z within distance ρ of Λ , and let D denote a bound on $\|P^s(z)\|$ and $\|P^u(z)\|$ on this set. Next, we pick λ_1 with $\lambda < \lambda_1 < 1$; then make ρ still smaller so that the norms of Γ_z^z on the range of $P^s(z)$ and $(\Gamma_{z'}^z)^{-1}$ on the range of $P^u(z')$ are both majorized by λ_1 whenever the distance from z to Λ and from z' to $f(z)$ are both not greater than ρ .

With the conditions imposed so far, if z is a ρ pseudo-orbit staying within distance ρ of Λ and if $\mathcal{X}^s$ and $\mathcal{X}^u$ denote, as earlier, the subspaces of sequences with ξ_n in the range of $P^s(z_n)$ and $P^u(z_n)$ respectively for all n , then the norm of Γ_0 restricted to $\mathcal{X}^s$ is majorized by λ_1 as is the norm of the inverse of Γ_0 restricted to $\mathcal{X}^u$. The argument given above in the case where z is an orbit in Λ can be applied to show that

$$\|(\Gamma_0 - 1)^{-1}\| \leq \frac{2D}{1 - \lambda_1}$$

It remains to control the effects of the difference between Γ_0 and Γ . We use the fact, easily checked using the Neumann series, that if A, B are operators with A invertible and $\|B\| \|A^{-1}\| \leq 1/2$, then $A + B$ is invertible and $\|(A + B)^{-1}\| \leq 2\|A^{-1}\|$. In addition to the conditions imposed on ρ above, we also require that it be small enough so that, if the distances from z to Λ and from z' to $f(z)$ are both not larger than ρ , then

$$\|Df(z) - \Gamma_{z'}^z\| \leq \frac{1 - \lambda_1}{4D}.$$

Then whenever z is a ρ pseudo-orbit staying within distance ρ of Λ we have

$$\|\Gamma - \Gamma_0\| \leq \frac{1 - \lambda_1}{4D},$$

and hence

$$\|(\Gamma - 1)^{-1}\| \leq \frac{4D}{1 - \lambda_1}.$$

Except for the proof of Lemma 3.4, this completes the proof of Lemma 3.2 and hence of the Shadowing Theorem.

PROOF OF LEMMA 3.4: . We first apply the Tietze Extension Theorem the individual matrix elements of $P^s(x)$ to extend $P^s(x)$ to a continuous matrix-valued function defined on all of $\mathcal{M}$. We will call the extended matrix-valued function $\hat{P}(x)$. For x in Λ , $\hat{P}(x) = P^s(x)$ is a projection, so the spectrum of $\hat{P}(x)$ is simply $\{0, 1\}$. By the continuity of the spectrum as a function of the operator, there is a neighborhood of Λ on which the spectrum of $\hat{P}(x)$ is contained in the union of two small disks, one centered at 0 and the other at 1. We can take for the extension of $P^s(x)$ to this neighborhood the spectral projection corresponding to the part of the spectrum in the small disk around 1.

§4. Structural stability of hyperbolic sets. There is still more to be extracted from the methods developed in the preceding section. One major extension starts from the observation that these methods imply the existence of shadowing orbits *not just for f itself, but for any $\hat{f}$ near enough to f .*

THEOREM 4.1. *Let Λ be a hyperbolic set for f , and let $\epsilon > 0$. Then there is a neighborhood $\mathcal{V}$ of f in the space of C^1 diffeomorphisms of $\mathcal{M}$ and a $\delta > 0$ such that, if z is a δ pseudo-orbit for f staying within distance δ of Λ , and if $\hat{f} \in \mathcal{V}$, then there is a y_0 such that*

$$\|\hat{f}^n(y_0) - z_n\| \leq \epsilon \quad \text{for all } n \in \mathbb{Z}.$$

If ϵ is small enough, y_0 is uniquely determined.

It might seem more natural to formulate the theorem as asserting that every δ pseudo-orbit for $\hat{f}$ (rather than for f) is shadowed by an exact orbit of $\hat{f}$. In fact, the above formulation is the one that we will find most directly useful. In any event, the alternative formulation follows easily from the one given (with a different relation between ϵ and δ .)

PROOF: Very little in the proof of the Shadowing Theorem has to be changed. For each pseudo-orbit z and each diffeomorphism $\hat{f}$ we construct an operator $\mathcal{A}$ on $\mathcal{X}$ by

$$\mathcal{A}(x)_n = \hat{f}(x_{n-1}).$$

Again, we apply the abstract fixed point theorem with Γ the linear operator

$$(\Gamma\xi)_n = Df(x_{n-1})\xi_{n-1}.$$

(Note that it is Df which appears here, not $D\hat{f}$.) Lemma 3.1 again applies; we let ρ , B denote what they mean in that lemma. To apply the fixed point theorem as we did to prove the Shadowing Theorem, we need to guarantee that, for any sufficiently small ϵ , there is a δ such that

$$(1) \quad \|D\mathcal{A}(x) - \Gamma\| \leq \frac{1}{2B} \quad \text{for all } x \text{ in the ball of radius } \epsilon \text{ about } z.$$

$$(2) \quad \|\mathcal{A}(z) - z\| \leq \frac{\rho}{4B}$$

whenever $\hat{f} \in \mathcal{V}$ and z is a δ pseudo-orbit. These conditions are what the corresponding conditions of the fixed point theorem become when we take into account the results of Lemma 3.1. To start with, we will always take $\epsilon \leq \rho$.

For (1), we first note that

$$\|D\mathcal{A}(x) - \Gamma\| \leq \sup_n \|D\hat{f}(x_n) - Df(x_n)\| + \sup_n \|Df(x_n) - Df(z_n)\|.$$

By making ϵ small, we can insure that

$$\|Df(x) - Df(z)\| \leq \frac{1}{4B}$$

whenever the distances from z to Λ and to x are both not greater than ϵ . Then by making $\mathcal{V}$ small we can ensure that

$$\|D\hat{f}(x) - Df(x)\| \leq \frac{1}{4B}$$

whenever $\hat{f} \in \mathcal{V}$ and the distance from x to Λ is not greater than 2ϵ . Thus, with the restrictions so far, (1) holds.

We are now finished imposing conditions on ϵ ; for the rest of the argument, ϵ is fixed. For (2), we note that

$$\|\mathcal{A}(z) - z\| \leq \sup_n \|\hat{f}(z_n) - f(z_n)\| + \sup_n \|f(z_n) - z_{n+1}\|,$$

and the second supremum is majorized by δ . Hence, if we take $\delta = \frac{\epsilon}{8D}$, and take $\mathcal{V}$ small enough so that $\|\hat{f}(z) - f(z)\| \leq \frac{\epsilon}{8D}$ whenever the distance from z to Λ is not greater than ϵ , then (2) will hold. This completes the proof of the theorem.

We get a striking result if we apply Theorem 4.1 to *exact* f -orbits on Λ . Take a small ϵ , and let $\mathcal{V}$ correspond to ϵ as in the theorem. Then, for any $\hat{f}$ in $\mathcal{V}$ and any x_0 in Λ , there is a unique y_0 such that

$$\|\hat{f}^n(y_0) - f^n(x_0)\| \quad \text{for all } n \in \mathbb{Z}.$$

We write

$$y_0 = \hat{h}(x_0);$$

$\hat{h}$ is a mapping of Λ into $\mathcal{M}$ such that $\|\hat{h}(x) - x\| \leq \epsilon$ for all $x \in \Lambda$. For each x , $\hat{f}(\hat{h}(x))$ has the defining property of $\hat{h}(f(x))$, so

$$\hat{f} \circ \hat{h} = \hat{h} \circ f.$$

We will argue shortly that $\hat{h}$ is continuous and injective, and hence maps Λ homeomorphically onto a near-by compact set $\hat{\Lambda}$. The relation $\hat{f} \circ \hat{h} = \hat{h} \circ f$ implies that $\hat{\Lambda}$ is invariant for $\hat{f}$ and that $\hat{f}$ on $\hat{\Lambda}$ is topologically conjugate to f on Λ . Thus

THEOREM 4.2 (STRUCTURAL STABILITY OF HYPERBOLIC SETS). *Let Λ be a hyperbolic set for the C^1 diffeomorphism f . For any $\hat{f}$ sufficiently near to f in the C^1 topology, there is a homeomorphism $\hat{h}$ of Λ onto a set $\hat{\Lambda}$ invariant for $\hat{f}$ such that*

$$\hat{f} = \hat{h} \circ f \circ \hat{h}^{-1} \quad \text{on } \hat{\Lambda}.$$

Moreover, $\hat{h}$ can be made as close as desired to the identity by making $\hat{f}$ close to f .

PROOF: It remains only to show that $\hat{h}$, as constructed above, is injective and continuous. Injectivity follows easily from the expansivity of f on Λ . Take ϵ small enough

so that two distinct f -orbits on Λ are somewhere separated by a distance of more than 2ϵ , and let $\hat{f}$ belong to the corresponding $\mathcal{V}$. Suppose that x_0 and x'_0 are two distinct points of Λ . Then

$$\|f^n(x_0) - f^n(x'_0)\| > 2\epsilon \quad \text{for some } n.$$

But since

$$\|f^n(x_0) - \hat{f}^n(\hat{h}(x_0))\| \leq \epsilon \quad \text{and} \quad \|f^n(x'_0) - \hat{f}^n(\hat{h}(x'_0))\| \leq \epsilon,$$

it follows that $\hat{f}^n(\hat{h}(x_0)) \neq \hat{f}^n(\hat{h}(x'_0))$, and hence that $\hat{h}(x_0) \neq \hat{h}(x'_0)$, as desired.

Continuity is proved from the uniqueness of the shadowing orbit and compactness in much the same way as it was proved above for the embedding of a two-shift near a transverse homoclinic point. That is: We start by observing that, by compactness, it suffices to prove that, if x_n is a sequence in Λ converging to a limit $\bar{x}$, and if $\hat{h}(x_n)$ converges to $\bar{y}$, then $\bar{y} = \hat{h}(\bar{x})$. For each $j, n \in \mathbb{Z}$,

$$\|\hat{f}^j(\hat{h}(x_n)) - f^j(x_n)\| \leq \epsilon.$$

Letting n go to infinity with j fixed gives

$$\|\hat{f}^j(\bar{y}) - f^j(\bar{x})\| \leq \epsilon.$$

This is true for all j , so the $\hat{f}$ orbit of $\bar{y}$ ϵ -shadows the f orbit of $\bar{x}$, so by the uniqueness of the shadowing orbit $\bar{y} = \hat{h}(\bar{x})$, as desired.

The above theorem says that, for any $\hat{f}$ near f , there is a set $\hat{\Lambda}$ near Λ , invariant for $\hat{f}$, such that the dynamics of $\hat{f}$ on $\hat{\Lambda}$ is topologically the same as the dynamics of f on Λ . We will see in the next chapter that $\hat{\Lambda}$ is in fact a hyperbolic set for $\hat{f}$. It is worth noting that we cannot expect $\hat{h}$ to be smooth, even if f and $\hat{f}$ are. For example, if Λ contains a periodic point x_0 with period p , then $\hat{h}$ carries x_0 to a periodic point $\hat{x}_0$ for $\hat{f}$. If $\hat{h}$ were differentiable at x_0 (i.e., were the restriction to Λ of a differentiable mapping defined on a neighborhood of Λ) with invertible derivative, then the spectrum of $D\hat{f}^p(\hat{x}_0)$ would have to be the same as the spectrum of $Df^p(x_0)$. Generically, however, perturbations of f change this spectrum, so, generically, $\hat{h}$ cannot be differentiable at periodic points. Since periodic points are dense in many hyperbolic sets, this at least suggests that we should not expect $\hat{h}$ to be differentiable anywhere.

We have called the above theorem the structural stability theorem for hyperbolic sets. Strictly speaking, this is not quite in accord with the normal use of the term "structural stability". A diffeomorphism f is said to be *structurally stable* (with respect to some topology on the space of all diffeomorphisms) if, for every $\hat{f}$ near enough to f in that topology, there is a homeomorphism $\hat{h}$ of the whole space to itself such that

$$\hat{f} = \hat{h} \circ f \circ \hat{h}^{-1}.$$

That is, structural stability is a global stability property, and not, as in the case at hand, a relative stability property of an invariant set. If the hyperbolic set in question is the whole space, however, then structural stability in the strict sense does hold:

COROLLARY 4.3 (STRUCTURAL STABILITY OF ANOSOV DIFFEOMORPHISMS). *Let f be a C^1 Anosov diffeomorphism. Then for any $\hat{f}$ close enough to f in the C^1 topology, there is a homeomorphism $\hat{h}$ of $\mathcal{M}$ onto itself such that*

$$\hat{f} = \hat{h} \circ f \hat{h}^{-1}$$

PROOF: The preceding theorem asserts the existence of a homeomorphic mapping $\hat{h}$ of $\mathcal{M}$ into itself conjugating $\hat{f}$ on its image to f on $\mathcal{M}$. We have to show that the image of $\hat{h}$ is all of $\mathcal{M}$. This follows immediately from general—if high-powered—topological considerations, but can also easily be proved using shadowing. What we need is to show is that, for every y_0 , there is an x_0 such that

$$\|f^n(x_0) - \hat{f}^n(y_0)\| \leq \epsilon \quad \text{for all } n.$$

By making $\hat{f}$ close to f (approximation in the C^0 topology will suffice), we can make $\hat{f}^n(y_0)$ be a δ pseudo-orbit for f with as small a δ as desired, uniformly in y_0 ; then the existence of x_0 follows from the Shadowing Theorem.

The ideas developed here can also be used to prove the following result, mentioned without proof in Chapter V:

THEOREM 4.4 (HARTMAN-GROBMAN LINEARIZATION THEOREM). *Let x_0 be a hyperbolic fixed point for the C^1 diffeomorphism f defined on a neighborhood of x_0 in the finite-dimensional vector space $\mathcal{E}$. Then f admits a topological linearization at x_0 , i.e., there is a homeomorphism Φ from a neighborhood of 0 onto a neighborhood of x_0 , with $\Phi(0) = x_0$, such that*

$$f = \Phi \circ Df(x_0) \circ \Phi^{-1} \quad \text{near } x_0.$$

SKETCH OF PROOF: We can assume that $x_0 = 0$. By multiplying the nonlinear part of f at 0 by a smooth cutoff function which is one on a small neighborhood of 0, we can reduce to the case where f is defined on all of $\mathcal{E}$ and $Df(x)$ is uniformly as near as we like to $Df(0)$. If we can show that the linear mapping $Df(0)$ is structurally stable on all of $\mathcal{E}$, the desired result will follow. The idea is now that, intuitively, the whole of $\mathcal{E}$ is like a hyperbolic set for $Df(0)$, i.e., $Df(0)$ is like an Anosov diffeomorphism on $\mathcal{E}$. Technically it is not quite true that $\mathcal{E}$ is a hyperbolic set for $Df(0)$ because hyperbolic sets are required to be compact. Nevertheless, the proof of the structural stability of Anosov diffeomorphisms can be repeated in this simpler situation, with minor and straightforward modifications, to prove the structural stability of $Df(0)$ and hence to prove the theorem.

CHAPTER X. STABILITY OF HYPERBOLICITY

Our objective in this chapter is to prove that the property of hyperbolicity is stable under small perturbations. The main tool for doing this is a theorem which characterizes hyperbolic sets in terms of associated operators on Banach spaces to which operator-theoretic perturbation methods can be applied.

§1. A Banach-space characterization of hyperbolicity. Let Λ denote a compact invariant set for the C^1 diffeomorphism f . We let $\mathcal{X}(\Lambda)$ denote the space of continuous mappings ξ from Λ to $\mathcal{E}$. (In the context of general manifolds, we take for $\mathcal{X}(\Lambda)$ the space of continuous sections of the tangent bundle to $\mathcal{M}$.) We will refer to elements of $\mathcal{X}(\Lambda)$ as (continuous) vector fields on Λ . Equipped with the supremum norm, $\mathcal{X}(\Lambda)$ is a Banach space. We define an operator $f_*(\Lambda)$ on $\mathcal{X}(\Lambda)$ by

$$(f_*(\Lambda)\xi)(x) = Df(f^{-1}(x))\xi(f^{-1}(x)).$$

It will usually be clear what set Λ we have in mind, in which case we will simply write $\mathcal{X}$ and f_* for $\mathcal{X}(\Lambda)$ and $f_*(\Lambda)$ respectively.

With this notation, we have the following characterization of hyperbolic sets:

THEOREM 1.1. Λ is a hyperbolic set for f if and only if $f_*(\Lambda)$ is a hyperbolic linear operator on $\mathcal{X}(\Lambda)$.

REMARK: The result remains true if we replace the $\mathcal{X}$ by the space of all bounded but not necessarily continuous vector fields, and the proof is essentially the same.

PROOF: For the proof of this theorem we need to work with the general, coordinate-independent condition for hyperbolicity

$$\|Df^n(x)\xi\| \leq c\lambda^n\|\xi\| \quad \text{for } \xi \in \mathcal{E}^s(x),$$

(and correspondingly for $\mathcal{E}^u$) rather than the condition with $c = 1$.

It is almost immediate that f_* is hyperbolic if Λ is a hyperbolic set. Assume that Λ is a hyperbolic set, and write

$$(1.1) \quad \begin{aligned} \mathcal{X}^s &= \{\xi \in \mathcal{X} : \xi(x) \in \mathcal{E}^s(x) \text{ for all } x\}, \\ \mathcal{X}^u &= \{\xi \in \mathcal{X} : \xi(x) \in \mathcal{E}^u(x) \text{ for all } x\}. \end{aligned}$$

Then $\mathcal{X}^s$ and $\mathcal{X}^u$ are invariant subspaces for f_* by the covariance of the splitting $\mathcal{E} = \mathcal{E}^s(x) \oplus \mathcal{E}^u(x)$, and $\mathcal{X} = \mathcal{X}^s \oplus \mathcal{X}^u$ by the corresponding pointwise equality. By the chain rule,

$$(1.2) \quad \underline{(f_*^n \xi)(f^n(x)) = Df^n(x)\xi(x)}$$

(where n may be negative as well as positive.) Hence, for $\xi \in \mathcal{X}^s$,

$$\begin{aligned} \|f_*^n \xi\| &= \sup_x \|f_*^n \xi(f^n(x))\| = \sup_x \|Df^n(x)\xi(x)\| \\ &\leq c\lambda^n \sup_x \|\xi(x)\| = c\lambda^n \|\xi\|. \end{aligned}$$

Thus, the norm of the restriction of f_*^n to $\mathcal{X}^s$ is bounded by $c\lambda^n$, and so the spectrum of the restriction of f_* to $\mathcal{X}^s$ is contained in the closed disk of radius λ centered at zero. Similarly, the spectrum of the restriction of f_* to $\mathcal{X}^u$ is outside the open disk with radius λ^{-1} centered at zero, so f_* is a hyperbolic linear operator.

Going in the other direction, we assume that f_* is a hyperbolic linear operator. Then by general spectral theory there is a splitting

$$\mathcal{X} = \mathcal{X}^s \oplus \mathcal{X}^u,$$

into f_* -invariant subspaces, and constants c, λ , with $0 < \lambda < 1$, such that the restrictions of f_*^n and f_*^{-n} to $\mathcal{X}^s$ and $\mathcal{X}^u$ respectively have norms bounded by $c\lambda^n$ for all $n \geq 0$. What we need to show, in essence, is that this splitting has the "pointwise form" (1.1). To prove this, we need to exploit the "pointwise form" of f_* . Specifically, we will exploit systematically the following observation:

CLAIM. *If ψ is a continuous real-valued function on Λ , then pointwise multiplication by ψ maps $\mathcal{X}^s$ ($\mathcal{X}^u$) into itself.*

To see this, first note that, by spectral theory, elements ξ of $\mathcal{X}^s$ are characterized by the property that $\|f_*^n \xi\|$ goes to zero as n goes to $+\infty$. On the other hand, from the formula (1.2) it follows that

$$f_*^n(\psi\xi)(x) = \psi(f^{-n}(x))(f_*^n\xi)(x);$$

hence that $\|f_*^n \psi\xi\| \rightarrow 0$ if $\|f_*^n \xi\| \rightarrow 0$.

Now, for each $x \in \Lambda$, let

$$\mathcal{F}^s(x) = \{\xi(x) : \xi \in \mathcal{X}^s\};$$

$$\mathcal{F}^u(x) = \{\xi(x) : \xi \in \mathcal{X}^u\}.$$

Since $\mathcal{X}^u$ and $\mathcal{X}^s$ span $\mathcal{X}$, $\mathcal{F}^s$ and $\mathcal{F}^u$ span $\mathcal{E}$. For $\xi \in \mathcal{F}^s(x)$, we choose any $\xi \in \mathcal{X}^s$ with $\xi(x) = \xi$; then since $\|f_*^n \xi\|$ goes to zero exponentially, $\|Df^n(x)\xi\|$ also goes to zero exponentially, so $\xi \in \mathcal{E}^s(x)$. Furthermore, by multiplying ξ by an appropriate continuous real-valued function, we can arrange that

$$\|\xi(x)\| = \|\xi\|$$

and we then get that

$$\|Df^n(x)\xi\| \leq c\lambda^n \|\xi\|,$$

i.e., that elements in $\mathcal{F}^s(x)$ decay under the action of $Df^n(x)$ at the uniform rate $c\lambda^n$. Similarly, elements of $\mathcal{F}^u(x)$ decay under the action of $Df^{-n}(x)$ also at the uniform rate $c\lambda^n$.

Finally, from the invariance of the splitting $\mathcal{X} = \mathcal{X}^s \oplus \mathcal{X}^u$ with respect to f_* , it follows that $Df(x)$ maps $\mathcal{F}^s(x)$ to $\mathcal{F}^s(f(x))$ and similarly for $\mathcal{F}^u$. From this covariance property and the decay established above it follows that

$$\|Df^n(x)\xi\| \geq c^{-1}\lambda^{-n}\|\xi\| \quad \text{for } \xi \in \mathcal{F}^u(x);$$

the proof is exactly like that used in ?? to establish corresponding growth rates for unstable vectors at points of hyperbolic sets. From this estimate it follows that $\mathcal{F}^s(x) \cap \mathcal{F}^u(x) = \{0\}$, and hence that $\mathcal{E} = \mathcal{F}^s(x) \oplus \mathcal{F}^u(x)$, and also that $\mathcal{F}^s(x) = \mathcal{E}^s(x)$. (We already know that $\mathcal{F}^s(x) \supset \mathcal{E}^s(x)$. But any vector not in $\mathcal{F}^s(x)$ has a non-zero component in $\mathcal{F}^u(x)$; this component grows exponentially under the action of $Df^n(x)$, so the original vector cannot be in $\mathcal{E}^s(x)$.) Similarly, we prove that $\mathcal{F}^u(x) = \mathcal{E}^u(x)$. We have thus established everything necessary to show that Λ is a hyperbolic set.

§2. Stability of hyperbolicity. In Theorem IX.4.2 we saw that, if Λ is a hyperbolic set for f , and if $\hat{f}$ is close enough to f in the C^1 topology, then there is a set $\hat{\Lambda}$ near Λ , compact and invariant for $\hat{f}$, on which the action of $\hat{f}$ is topologically conjugate to the action of f on Λ . We will show here, using the characterization of hyperbolic set given in the preceding section, that $\hat{\Lambda}$ is a hyperbolic set for $\hat{f}$. In fact, we prove something more comprehensive: If $\hat{f}$ is close enough to f in the C^1 topology, then any compact set, invariant for $\hat{f}$ and near enough to Λ , is a hyperbolic set for $\hat{f}$. There is no need for the structure of $\hat{\Lambda}$ to be at all like the structure of Λ . This result implies, for example, that the embedded two-shifts which we saw in Proposition IX.2.2 could be found near any transverse homoclinic point are hyperbolic sets.

THEOREM 2.1. *Let Λ be a hyperbolic set for f . Then there exist a C^1 neighborhood $\mathcal{V}$ of f and a neighborhood $\mathcal{U}$ of Λ in $\mathcal{M}$ such that, if $\hat{f} \in \mathcal{V}$ and if Ξ is a compact subset of $\mathcal{U}$ invariant for $\hat{f}$, then Ξ is a hyperbolic set for $\hat{f}$.*

PROOF: The proof uses many of the ideas which were used to prove the Shadowing Theorem. We want to use the characterization of hyperbolic sets proved in the preceding section. The idea will be to develop a systematic procedure for constructing, for any set Ξ which is near Λ and invariant for an $\hat{f}$ near f , an operator Γ_0 on $\mathcal{X}(\Xi)$ which can be shown to be hyperbolic by explicitly exhibiting invariant contracting and expanding subspaces and which is near enough to $\hat{f}_*(\Xi)$ so that operator perturbation theory can be used to deduce the hyperbolicity of $\hat{f}_*(\Xi)$.

As in §IX.3, we start by finding a continuous projection-valued extension of $P^s(x)$ to a compact neighborhood* $\mathcal{U}_1$ of Λ , and we let $P^u(x)$ denote $1 - P^s(x)$. We construct a continuous matrix-valued function of two variables

$$\Gamma_x^z = P^s(x')Df(x)P^s(x) + P^u(x')Df(x)P^u(x).$$

*We use the convention that a neighborhood of a set X is not necessarily open; it is a set containing an open set containing X .

By its construction, Γ_x^z maps the range of $P^s(x)$ to the range of $P^s(x')$ and similarly for P^u . Furthermore, Γ_x^z reduces to $Df(x)$ when $x \in \Lambda$ and $x' = f(x)$. Let $\lambda < 1$ be such that

$$\|Df(x)\xi\| \leq \lambda\|\xi\| \quad \text{for } x \in \Lambda \text{ and } \xi \in \mathcal{E}^s(x),$$

$$\|Df^{-1}(x)\xi\| \leq \lambda\|\xi\| \quad \text{for } x \in \Lambda \text{ and } \xi \in \mathcal{E}^u(x),$$

and let $\lambda < \lambda_1 < 1$. Since Γ_x^z reduces to $Df(x)$ for $x \in \Lambda$ and $x' = f(x)$, by uniform continuity there exists a subneighborhood $\mathcal{U}_2$ of $\mathcal{U}_1$ and a $\delta_1 > 0$ such that

$$\|\Gamma_x^z\xi\| \leq \lambda_1\|\xi\| \quad \text{for } x \in \mathcal{U}_2, P^s(x)\xi = \xi, \text{ and } \|x' - f(x)\| \leq \delta_1;$$

$$\|\Gamma_x^z\xi\| \geq \lambda_1^{-1}\|\xi\| \quad \text{for } x \in \mathcal{U}_2, P^u(x)\xi = \xi, \text{ and } \|x' - f(x)\| \leq \delta_1.$$

Now define D by

$$D = \sup_{x \in \mathcal{U}_2} \max\{\|P^s(x)\|, \|P^u(x)\|\},$$

and let B denote $\frac{2D}{1 - \lambda_1}$.

CLAIM. For any $\hat{f}$ such that

$$\sup_{x \in \mathcal{U}_2} \|\hat{f}(x) - f(x)\| \leq \delta,$$

and any compact set $\Xi \subset \mathcal{U}_2$ invariant for $\hat{f}$, the operator Γ_0 defined on $\mathcal{X}(\Xi)$ by

$$(\Gamma_0\xi)(\hat{f}(x)) = \Gamma_{\hat{f}(x)}^z\xi(x)$$

is hyperbolic and

$$\|(\Gamma_0 - e^{i\theta}1)^{-1}\| \leq B \quad \text{for all } \theta \in \mathbb{R}.$$

PROOF: Let $\mathcal{X}^s$ ($\mathcal{X}^u$) denote the subspace of $\mathcal{X}$ consisting of vector fields such that $\xi(x)$ is in the range of $P^s(x)$ ($P^u(x)$) for all $x \in \Xi$. Then $\mathcal{X} = \mathcal{X}^s \oplus \mathcal{X}^u$; this splitting is invariant under Γ_0 ; the norms of the restrictions of Γ_0 to $\mathcal{X}^s$ and of Γ_0^{-1} to $\mathcal{X}^u$ are both bounded by λ_1 . As we saw in section ??, this implies that $\Gamma_0 - 1$ is invertible on both $\mathcal{X}^s$ and $\mathcal{X}^u$ and that its norm on each of these subspaces is bounded by $(1 - \lambda_1)^{-1}$. Again as in section ??, it follows that the norm on $(\Gamma_0 - 1)^{-1}$ on $\mathcal{X}$ is bounded by $\frac{2D}{1 - \lambda_1} = B$. Once all the vector spaces in question have been appropriately complexified, the argument works with Γ_0 replaced by $e^{-i\theta}\Gamma_0$ for any real θ , and this proves the claim.

§1. **Generalities on recurrence.** In analyzing dynamical systems, it is useful to distinguish between orbits which are *transient*, i.e., which return to the vicinity of their initial points at most a finite number of times and those which are in some sense recurrent. Unfortunately, there are a number of ways of making this distinction precise, each with its advantages and disadvantages. These alternative formulations turn out to agree in nice cases, but not in general. We will discuss three useful notions of recurrence here. For the purposes of this section, we have no need to restrict ourselves to differentiable and invertibility mappings; we consider instead arbitrary continuous mappings of metrizable spaces to themselves.* One reason for working with this extra generality is that we want to be able to speak of recurrent and transient points of the restriction of a diffeomorphism to a hyperbolic set which is not necessarily smooth.

Let f be a continuous mapping of a metrizable space X to itself. As in Section ??, we define the ω -limit set $\omega(x, f)$ of a point x under the action of f to be the set of accumulation points of the forward orbit $O_+(x, f) = \{f^n x : n = 0, 1, \dots\}$ of x . Although we can make the definition with no further restrictions, and although $\omega(x, f)$ is closed under these general hypotheses, it need not be true that $\omega(x, f)$ is non-empty. This will be true, however, if the forward orbit is precompact (i.e., has compact closure), which will usually be the case for the situations we consider. We define the *limit set* $L(f)$ by

$$L(f) := \bigcup_{x \in X} \omega(x, f).$$

One useful notion of recurrence is that a point is recurrent if it belongs to the closure $\overline{L(f)}$ of $L(f)$. (Intuitively, it might seem more natural to say that a point is recurrent if it belongs to *its own* ω -limit set, i.e., if it is an accumulation point of its forward orbit. This property is, however, badly behaved and does not seem to be very useful.)

PROPOSITION 1.1. *The set $L(f)$ is mapped into itself by f and contains all periodic points of f . If all forward orbits of f are precompact, then f maps $L(f)$ onto itself. If x is any point of X whose forward orbit is precompact, and if $\mathcal{U}$ be an open neighborhood of $L(f)$, then $f^n(x) \in \mathcal{U}$ for all sufficiently large n .*

PROOF: It is nearly immediate that $L(f)$ contains all periodic points and is mapped into itself by f . To prove that it is mapped onto itself if all forward orbits are precompact, let $x \in L(f)$ be an accumulation point of the forward orbit of z . Then there exist $n_j \rightarrow \infty$ such that $f^{n_j}(z) \rightarrow x$. By precompactness, the sequence $f^{n_j-1}(z)$ has accumulation points; any accumulation point is in $L(f)$ (by definition) and is mapped to x by f .

*In fact, much of what we have to say extends to spaces more general than metrizable ones, but we do not want to obscure the discussion by worrying about giving the most general hypothesis for each result. On the other hand, when we come to the notion of chain recurrence, we will need to consider spaces which are not merely metrizable but actually equipped with a metric (or at least with a metrizable uniform structure.) In any event, the applications we have in mind are all to compact metric spaces.

To prove the last assertion, assume the contrary. Then there exists $n_j \rightarrow \infty$ such that $f^{n_j}(x) \notin \mathcal{U}$. By precompactness, this subsequence has an accumulation point, and this accumulation point is not in $\mathcal{U}$ since the complement of $\mathcal{U}$ is closed. But on the other hand the accumulation point is in $\omega(x, f)$; hence, in $L(f)$, and this contradicts the assumption that $\mathcal{U}$ contains $L(f)$.

If all forward orbits are precompact—e.g., if X is compact—then all forward orbits converge to $L(f)$. WARNING: This does not mean that $L(f)$ is an attracting set. As an example of why it is not, suppose that f is differentiable and has a hyperbolic fixed point x_0 of saddle type (i.e., with nontrivial stable and unstable manifolds), and that the unstable manifold of x_0 is contained in the stable manifold of an attracting hyperbolic fixed point x_1 . It is then easy to see that x_0 is an isolated point of $L(f)$ —the idea is that any orbit passing near enough to x_1 but not actually on the stable manifold of x_0 must leave the vicinity of x_0 near the unstable manifold of x_0 and is therefore attracted to x_1 and never returns near x_0 . Now take an open neighborhood $\mathcal{U}$ of $L(f)$ such that any orbit passing near x_0 but ultimately asymptotic to x_1 must pass outside $\mathcal{U}$ between x_0 and x_1 . If $L(f)$ were an attracting set, there would have to be an open set $\mathcal{V}$ containing at least x_1 and a positive integer n_0 such that $f^n(x) \in \mathcal{U}$ for all $x \in \mathcal{V}$ and $n \geq n_0$. Such a $\mathcal{V}$ and n_0 cannot exist: If there did, we could choose x very near to x_0 —hence, in $\mathcal{V}$ —so that its orbit remains in the part of $\mathcal{U}$ around x_0 up to time n_0 . But if, in addition, we choose x not to lie on the stable manifold of x_0 , its orbit will eventually be attracted to x_1 and hence will get out of $\mathcal{U}$ at some time later than n_0 .

A second candidate for a precise definition of recurrence is the notion of *non-wandering* point. A point x is said to be *wandering* if there exists a neighborhood $\mathcal{U}$ of x and a $j_0 > 0$ such that

$$f^j \mathcal{U} \cap \mathcal{U} = \emptyset \quad \text{for all } j \geq j_0,$$

i.e., if every orbit starting near x eventually gets and stays away from x . If there exists such a pair $\mathcal{U}, j_0$, then any small enough subneighborhood $\mathcal{V}$ of $\mathcal{U}$ will have

$$f^j \mathcal{V} \cap \mathcal{V} = \emptyset \quad \text{for all } j > 0.$$

A point x is said to be *non-wandering* if it is not wandering. The set of wandering points is manifestly open; its complement—called the *non-wandering set* and denoted by $\Omega(f)$ —is therefore closed.

PROPOSITION 1.2. $f\Omega(f) \subset \Omega(f)$ and $L(f) \subset \Omega(f)$. If f is a homeomorphism, $\Omega(f^{-1}) = \Omega(f)$.

PROOF: To prove that f maps $\Omega(f)$ into itself, it suffices to prove that, if $f(x)$ is wandering, so is x . But if $\mathcal{U}$ is an open set containing $f(x)$ such that $f^j \mathcal{U} \cap \mathcal{U} = \emptyset$ for $j > 0$, then $\mathcal{V} = f^{-1}\mathcal{U}$ is an open set containing x such that $f^j \mathcal{V} \cap \mathcal{V} = \emptyset$ for all $j > 1$.

If $x \in L(f)$ then there exist y such that, for any neighborhood $\mathcal{U}$ of x there are infinitely many n 's such that $f^n(y) \in \mathcal{U}$; hence, x cannot be non-wandering.

If f is a homeomorphism and x is wandering for f , then there is an open $\mathcal{U}$ containing x such that $f^n \mathcal{U} \cap \mathcal{U}$ is empty for all $n > 0$. Applying f^{-n} shows that $\mathcal{U} \cap f^{-n} \mathcal{U}$ is

empty for $n > 0$, which means that x is wandering for f^{-1} . Thus, the non-wandering set of f^{-1} is contained in the non-wandering set of f . Interchanging the roles of f and f^{-1} shows that $\Omega(f^{-1}) = \Omega(f)$, as desired.

The difference between $\overline{L(f)}$ and $\Omega(f)$ lies in the fact that, for x in $\overline{L(f)}$ and $\mathcal{U}$ is any neighborhood of x , then there is a single orbit which returns infinitely often to $\mathcal{U}$ whereas, for x in $\Omega(f)$ and $\mathcal{U}$ is a neighborhood of x , there are points in $\mathcal{U}$ which return to $\mathcal{U}$ at arbitrarily large time (but any given orbit may return only finitely many times.) This is a genuine distinction; there do exist mappings for which $\overline{L(f)} \neq \Omega(f)$. One can make an example roughly as follows: Take a mapping defined in particular on a neighborhood of the unit circle in the plane. The unit circle is to be invariant; there is a single fixed point x_0 on the circle; and all orbits $f^n(x)$ on the circle converge to x_0 as n goes to $\pm\infty$. Orbits near but not on the circle spiral slowly away from the circle and never come back. Then every point of the circle is non-wandering, but the only point of $L(f)$ on or near the circle is x_0 .

A third possible definition of recurrence—and the one we will find most useful—is the notion of chain recurrence. To formulate the definition we need to choose a metric $d(.,.)$ on X . We then say that a point x is *chain-recurrent* if for each $\delta > 0$ there is a finite δ -pseudo orbit starting and ending at x , i.e., a sequence $x_0 = x, x_1, \dots, x_n = x$ such that $d(x_{i+1}, f(x_i)) < \delta$ for $i = 0, \dots, n-1$. If we have a finite δ pseudo-orbit starting and ending at x , we can repeat it to get a periodic δ pseudo-orbit passing through x , i.e., the definition of chain recurrence can be reformulated as saying that x lies on a periodic δ pseudo-orbit for arbitrarily small δ . Intuitively, saying that x is *not* chain recurrent means that x is transient in the strong sense that the orbits of points near x go away from x and never return, not only for the motion induced by f itself but for arbitrary small (possibly non-deterministic) perturbations of f .

PROPOSITION 1.3. *$R(f)$ is closed and contains $\Omega(f)$; if f is uniformly continuous, f maps $R(f)$ into itself.*

PROOF: To prove that $R(f)$ is closed, we show that its complement is open. Suppose $x \notin R(f)$. Then there exists $\delta > 0$ such that there is no δ pseudo-orbit starting and ending at x . Find $\delta_1 < \delta/2$ such that $d(x, y) < \delta_1$ implies $d(f(x), f(y)) < \delta/2$. We claim that, if $d(x, y) < \delta_1$, there is no δ_1 pseudo-orbit starting and ending at y . To see this, suppose there is such a pseudo-orbit, say $y = y_0, y_1, \dots, y_n = y$. We want to argue that $x, y_1, \dots, y_{n-1}, x$ is a δ pseudo-orbit; then, since we have assumed that there is no δ pseudo-orbit starting and ending at x , we have arrived at a contradiction. Since $\delta_1 < \delta$, there is no problem with any step except possibly the first and the last. For the first step, we estimate

$$d(f(x), y_1) \leq d(f(x), f(y)) + d(f(y), y_1) < \delta/2 + \delta_1 \leq \delta.$$

For the last step, we estimate

$$d(f(y_{n-1}), x) \leq d(f(y_{n-1}), y) + d(y, x) < \delta_1 + \delta_1 \leq \delta;$$

and this completes the proof that the complement of $R(f)$ is open.

Now let $x \in \Omega(f)$ and let $\delta > 0$. We want to show that there is a δ pseudo-orbit from x to itself. To do this, choose $\delta_1 < \delta$ so that $d(x, y) < \delta_1$ implies $d(f(x), f(y)) < \delta$. From the definition of $\Omega(f)$, there exist y and $n > 0$ such that $d(x, y) < \delta_1$ and $d(x, f^n(y)) < \delta_1$. We claim that $x, f(y), f^2(y), \dots, f^{n-1}(y), x$ is a δ pseudo-orbit. All steps but the first and last are made by exact application of f , so all that needs to be checked is that $d(f(x), f(y)) < \delta$ (which is guaranteed by $d(x, y) < \delta_1$) and that $d(f^n(y), x) < \delta$ (which is how we chose y, n .) Thus, we have proved $\Omega(f) \subset R(f)$.

Finally, assume that f is uniformly continuous and that $x \in R(f)$; we want to prove that $f(x) \in R(f)$. Thus, choose $\delta > 0$; then choose δ_1 such that $d(y_1, y_2) < \delta_1$ implies $d(f(y_1), f(y_2)) < \delta$. Now let $x = y_0, y_1, \dots, y_{n-1}, y_n = x$ be a δ_1 pseudo-orbit; we claim that $f(y_0), \dots, f(y_n)$ is a δ pseudo-orbit. This follows immediately from the construction; for each $i = 0, \dots, n-1$, $d(x_{i+1}, f(x_i)) < \delta_1$, so $d(f(x_{i+1}), f(f(x_i))) < \delta$; this proves the claim and hence proves that $f(x) \in R(f)$.

The theory of chain recurrence is not quite as trivial as the above straightforward argument might suggest. Notably we have

PROPOSITION 1.4. *If X is compact, if $x \in R(f)$, and if $\delta > 0$, then there is a δ pseudo-orbit in $R(f)$ starting and ending at x .*

What is new in the assertion is that the pseudo-orbit can be taken to lie in $R(f)$, not just in the ambient space X . In other words,

$$R(f|_{R(f)}) = R(f).$$

The corresponding identity does not hold for the non-wandering set. This can be seen in the above example of a mapping for which $\Omega(f) \neq \overline{L(f)}$. In this example, all points of the circle are in $\Omega(f)$ and no points near but off the circle are. Thus, points in the circle other than x_0 are *not* in the non-wandering set of the restriction of f to $\Omega(f)$.

PROOF OF PROPOSITION 1.4: We need some notation. For $\delta > 0$ and $x \in X$, we write $R_\delta(x; f)$ for the set of points $y \in X$ such that there exist δ pseudo-orbits from x to y and from y to x . (When we speak of pseudo-orbits here, we always mean sequences with at least two terms; x does not count as a pseudo-orbit from x to itself.) It is immediate from the definition that $x \in R_\delta(x; f)$ if $R_\delta(x; f)$ is non-empty; also, since the definition of δ pseudo-orbit is that $d(y_{j+1}, f(y_j))$ is strictly smaller than δ , $R_\delta(x; f)$ is open. We will further write $R(x; f)$ for the intersection over $\delta > 0$ of $R_\delta(x)$. A point x is then chain recurrent if and only if $R(x; f)$ is not empty, and all points of $R(x)$ are chain recurrent. Also, an argument very similar to the one used to show that $R(f)$ is closed shows that $R(x; f)$ is closed. This argument in fact shows more; it shows that, if $y \notin R_\delta(x; f)$, then there is a neighborhood $\mathcal{V}$ of y and a $\delta_1 > 0$ such that $\mathcal{V}$ is in the complement of $R_{\delta_1}(x; f)$. Hence: $R(x; f)$ is also the intersection of the closures of the $R_\delta(x; f)$.

The idea is now the following: We pick a small open set $\mathcal{U}$ containing $R(x; f)$. By compactness, since the closures of the $R_\delta(x; f)$ decrease to $R(x; f)$, we have $R_{\delta_1}(x; f) \subset$

$\mathcal{U}$ for sufficiently small δ_1 . Any δ_1 pseudo-orbit from x to x has to be in $R_{\delta_1}(x; f)$ (by the definition of this latter set), so there is a δ_1 pseudo-orbit from x to x in $\mathcal{U}$. But now, if δ_1 and $\mathcal{U}$ are small enough, we can perturb such a pseudo-orbit slightly by pushing each of its points into $R(x; f)$ to get a δ pseudo-orbit in $R(x; f) \subset R(f)$.

The details are as follows: Choose $\delta_2 < \delta/3$ so that

$$d(z_1, z_2) < \delta_2 \text{ implies } d(f(z_1), f(z_2)) < \delta/3.$$

Then choose $\delta_1 < \delta_2$ so that $R_{\delta_1}(x; f)$ is contained in the δ_2 -neighborhood of $R(x; f)$ (i.e., the set of points within distance δ_2 of $R(x; f)$.) Let $z_0 = x, z_1, \dots, z_n = x$ be a δ_1 pseudo-orbit from x to x . For $i = 1, \dots, n-1$, choose $y_i \in R(x; f)$ with $d(y_i, z_i) < \delta_2$. We now claim that $y_0 = x, y_1, \dots, y_n = x$ is a δ pseudo-orbit. Consider first $d(f(y_i), y_{i+1})$ for $i = 1, \dots, n-2$. Then

$$\begin{aligned} d(f(y_i), y_{i+1}) &\leq d(f(y_i), f(z_i)) + d(f(z_i), z_{i+1}) + d(z_{i+1}, y_{i+1}) \\ &< \delta/3 + \delta_1 + \delta_2 < \delta/3 + \delta/3 + \delta/3, \end{aligned}$$

as desired.

The following version of the Poincaré Recurrence Theorem says that finite invariant measures are, in a natural sense, concentrated on the non-wandering set:

PROPOSITION 1.5. *Let f be continuous mapping of the topological space X to itself; μ a finite measure on the σ -algebra generated by the open sets of X which is invariant under f (i.e., which satisfies $\mu(f^{-1}B) = \mu(B)$ for all measurable sets B .) If X contains a denumerable dense set, then the complement of the non-wandering set has μ -measure zero.*

PROOF: Let x be a wandering point for f ; $\mathcal{U}$ an open set containing x such that $f^n\mathcal{U} \cap \mathcal{U}$ is empty for all $n > 0$. The sets $f^{-n}\mathcal{U}$, $n = 0, 1, \dots$ all have the same μ -measure, and are also mutually disjoint, since if $y \in f^{-n}\mathcal{U} \cap f^{-m}\mathcal{U}$ with, say, $m > n$, then $f^m(y)$ is in both $\mathcal{U}$ and $f^{m-n}\mathcal{U}$, and this is impossible by assumption. Since μ is a finite measure, this implies that $\mu(\mathcal{U}) = 0$. Thus, we have shown that any point of the complement of $\Omega(f)$ has an open neighborhood of μ -measure zero, and this argument only uses the continuity of f and the fact that open sets are measurable. We want to conclude that the complement of $\Omega(f)$ has measure zero. This requires some further assumption; the existence of a denumerable dense set in X suffices since it implies that the complement of $\Omega(f)$ can be written as a countable union of these open sets of measure zero. Many other conditions—e.g., that X is locally compact and μ inner regular—would also have worked.

COROLLARY 1.6. *If X is a compact manifold, and if there is a finite f -invariant measure equivalent to Lebesgue measure*, then $\Omega(f) = X$.*

*Strictly speaking, it does not make sense to speak of Lebesgue measure on a general manifold. A manifold does have, however, a natural Lebesgue measure class: The null sets of this measure class are those sets whose intersections with all coordinate patches are images of Lebesgue null sets under the corresponding coordinate mappings. When we speak of a measure equivalent to Lebesgue measure, we mean a measure having exactly these sets as sets of measure zero.

PROOF: The complement of $\Omega(f)$ is a Lebesgue null set; hence $\Omega(f)$ is dense. Since $\Omega(f)$ is also closed, it is all of X . Note that this argument does *not* require ergodicity of the measure.

§2. The Anosov closing lemma. If f is any mapping, we write $\text{Per}(f)$ for the set of all periodic points of f .

PROPOSITION 2.1. Let Λ be a hyperbolic set for the C^1 -diffeomorphism f ; assume that Λ has local product structure. Then the chain recurrent set of the restriction of f to Λ is the closure of the set of periodic points in Λ .

PROOF: Clearly, any periodic point is chain-recurrent. Hence, since the chain-recurrent set is closed, it contains the closure of the set of periodic points. What we need to show, therefore, is that any chain-recurrent point can be approximated arbitrarily closely by periodic points. This, however, is a simple consequence of shadowing: Let x be a chain-recurrent point. Then there exist periodic pseudo-orbits with arbitrarily small δ passing through x . If δ is small enough, such a pseudo-orbit is closely shadowed by a unique exact orbit (by the Shadowing Theorem.) This exact orbit is periodic (by uniqueness) and lies in Λ (by local maximality, which follows from local product structure.) Hence, as desired, there are periodic orbits in Λ passing arbitrarily near to x .

COROLLARY 2.2 (ANOSOV CLOSING LEMMA). If f is an Anosov diffeomorphism, periodic points are dense in $R(f)$.

PROOF: The only thing which needs to be noted is that local product structure is immediate if the hyperbolic set in question is the whole space.

One of the most important open questions in this general area is whether there exist Anosov diffeomorphisms for which $R(f)$ is not the whole space. Note that, by Corollary 1.6, volume-preserving Anosov diffeomorphisms—i.e., ones which leave invariant some finite measure equivalent to Lebesgue—necessarily have $R(f)$ the whole space and thus have a dense set of periodic points.

§3. Topological transitivity and mixing. A continuous mapping f of a topological space X to itself is said to be *topologically transitive* if, for any two non-empty open sets $\mathcal{U}$ and $\mathcal{V}$, there exists n such that $f^n\mathcal{U} \cap \mathcal{V}$ is non-empty, and to be *topologically mixing* if, for any two non-empty open sets $\mathcal{U}$ and $\mathcal{V}$, $f^n\mathcal{U} \cap \mathcal{V}$ is non-empty for all sufficiently small n . It is immediate from the definition that a topologically mixing mapping is topologically transitive. An illuminating example of a topologically transitive mapping which is not topologically transitive is a cyclic permutation on a finite set with the discrete topology.

PROPOSITION 3.1. Let f be a continuous mapping from a complete separable metric space X to itself. Then f is topologically transitive if and only if it has a dense orbit.

PROOF: It is easy to see that a mapping with a dense orbit is topologically transitive (without any supplementary topological assumptions.) Thus, suppose that f is topolog-

ically transitive. Let $\mathcal{V}_1, \mathcal{V}_2, \dots$ be a base for the topology of X . It is easy to construct recursively a sequence of non-empty open sets $\mathcal{W}_j$ and indices $i_1, i_2, \dots$, such that

- (1) $\overline{\mathcal{W}_{j+1}} \subset \mathcal{W}_j$.
- (2) $\text{diam}(\mathcal{W}_j) \leq \frac{1}{j}$.
- (3) $f^{i_j} \mathcal{W}_j \subset \mathcal{V}_j$.

Then the $\mathcal{W}_j$ decrease to a single point x , and, for each j , $f^{i_j}(x) \in \mathcal{V}_j$, which implies that the orbit of x is dense.

§4. The Smale spectral decomposition. A hyperbolic set for a diffeomorphism f can fail in a trivial way to be topologically transitive; it can consist of two or more "independent" pieces, i.e., separately invariant sets at non-zero distances from each other. Furthermore, if the set is topologically transitive, it can fail in a trivial way to be topologically mixing; it can consist of a finite number of pieces permuted cyclically by f . We will see in this section that, provided that the set in question has local product structure and that the action of f on it is recurrent, this is the only ways in which topological mixing can fail. Once we assume that the set has local product structure, all reasonable notions of recurrence coincide, and we can therefore assume the form which is formally strongest, the density of periodic points.

THEOREM 4.1 (ANOSOV'S GENERALIZATION OF SMALE'S SPECTRAL DECOMPOSITION THEOREM). *Let Λ be a hyperbolic set for the C^1 diffeomorphism f ; assume that Λ has local product structure and that periodic points are dense in Λ . Then Λ can be decomposed as*

$$\Lambda = \bigcup_{i=1}^m \left(\bigcup_{j=1}^{n_i} \Lambda_{i,j} \right),$$

where

- (1) The $\Lambda_{i,j}$ are closed and mutually disjoint (and hence, by compactness, at non-zero distance from each other.)
- (2) f maps $\Lambda_{i,j}$ to $\Lambda_{i,j+1}$ for $1 \leq j \leq n_i - 1$ and Λ_{i,n_i} to $\Lambda_{i,1}$.
- (3) f^{n_i} , which by (2) maps $\Lambda_{i,j}$ to itself, is topologically mixing on $\Lambda_{i,j}$.

This decomposition is unique up to labelling. If x is any point of $\Lambda_{i,j}$, then $\Lambda_{i,j}$ is the closure of the $W^u(x) \cap \Lambda$ and also of $W^s(x) \cap \Lambda$.

Note that this theorem gives a two-stage decomposition of f . Each

$$\Lambda^{(i)} := \bigcup_{j=1}^{n_i} \Lambda_{i,j}$$

is closed and invariant; it is easy to see, since f^{n_i} is topologically mixing on each $\Lambda_{i,j}$, that f is topologically transitive on $\Lambda^{(i)}$. Thus, the decomposition

$$\Lambda = \bigcup_{i=0}^m \Lambda^{(i)}$$

splits Λ into invariant topologically transitive pieces. The further splitting $\Lambda^{(i)}$ into the $\Lambda_{i,j}$ is a *non-invariant* decomposition into pieces which are cyclically permuted by f . Heuristically, this decomposition extracts a cyclic component of the action of f on $\Lambda^{(i)}$ so that what is left—the action of f^{n_i} on any one of the $\Lambda_{i,j}$ —is topologically mixing.

COROLLARY 4.2. *Let f be an Anosov diffeomorphism of a connected compact manifold; assume that periodic points for f are dense. Then f is topologically mixing, and each $W^s(x)$ and $W^u(x)$ is dense.*

PROOF: Since the whole space is hyperbolic for an Anosov diffeomorphism, local product structure holds trivially. Thus, the theorem applies. But since the space is assumed connected, the decomposition has to be trivial, i.e., f must already be topologically mixing.

PROOF OF THEOREM 4.1: The proof is a little intricate, and we will break it up into a sequence of lemmas. The first important idea is that the topological mixing property is equivalent to the density of all unstable manifolds:

LEMMA 4.3. *Let Λ, f be as in the statement of the theorem. Then f is topologically mixing on Λ if and only if $W^u(x) \cap \Lambda$ is dense in Λ for all periodic points x in Λ .*

PROOF (BEGINNING): We prove the “if” part, which is the more important for our purposes. Thus, we assume the density of the unstable manifolds of all periodic points, and we let $\mathcal{U}, \mathcal{V}$ be non-empty open subsets of Λ . We want to show that $f^n \mathcal{U}$ intersects $\mathcal{V}$ for all sufficiently large n . By assumption, $\mathcal{U}$ contains a periodic point x . We consider first the case where x is a fixed point for f , and we take $\eta > 0$ small enough so that $W_\eta^u(x)$ is contained in $\mathcal{U}$. Since $f^n W_\eta^u(x)$ is an increasing sequence of sets whose union is dense, $f^n W_\eta^u(x)$ intersects $\mathcal{V}$ for all sufficiently large n , and this implies $f^n \mathcal{U}$ intersects $\mathcal{V}$ for all sufficiently large n .

The case where x is periodic with period $p > 1$ is complicated slightly by the fact that $W^u(x)$ is no longer an invariant set. Write x_i for $f^i(x)$; each x_i is a fixed point for f^p . Again choose $\eta > 0$ so that $W_\eta^u(x)$ is contained in $\mathcal{U}$, and note that $f^i W_\eta^u(x)$ contains $W_\eta^u(x_i)$ for each $i = 0, 1, \dots, p-1$. By the argument in the preceding paragraph applied to each of the fixed points x_i of f^p , we see that there is an n_0 such that $f^{np} W_\eta^u(x_i)$ intersects $\mathcal{V}$ when $n \geq n_0$. Then $f^{np+i} W_\eta^u(x) \supset f^{np} W_\eta^u(x_i)$ intersects $\mathcal{V}$ whenever $n \geq n_0$, i.e., $f^m W_\eta^u(x) \supset f^m \mathcal{U}$ intersects $\mathcal{V}$ when $m \geq n_0 p$.

We will return later to prove the “only if” part of the lemma. We first need to develop a second important idea in the proof of the theorem. The preceding lemma suggests trying to construct the $\Lambda_{i,j}$ as closures of unstable manifolds of periodic points of Λ . On the other hand, the $\Lambda_{i,j}$ are both closed and open in Λ , so, to make this idea work, it is necessary to show that closures of unstable manifolds are automatically open as well as closed. This is the content of the following lemma:

LEMMA 4.4. *Let Λ, f be as in the statement of the theorem, and let x be a periodic point for f in Λ . Then $\overline{W^u(x)} \cap \Lambda$ is relatively open in Λ .*

PROOF: Intuitively, the idea is the following: What we have to show is that, whenever $W^u(x)$ passes reasonably near a point y of Λ , it comes arbitrarily near to it. But

whenever $W^u(x)$ passes reasonably near a periodic point y , it must intersect the stable manifold of y (simply because unstable and stable manifolds are transverse to each other) and is then “pulled into” y along the stable manifold.

The details go as follows: Take $\delta > 0$ small enough so that $[y, z]$ is defined and in Λ whenever y, z are points of Λ whose separation is less than δ . To prove the assertion, it will suffice to show that for any $z \in W^u(x) \cap \Lambda$, the open ball in Λ of radius δ and center z is contained in $\overline{W^u(x) \cap \Lambda}$. If this is not true, then the complement of $\overline{W^u(x) \cap \Lambda}$ in this ball is a non-empty relatively open subset of Λ and hence contains a periodic point y . Let $z_1 = [y, z]$; this is defined and in Λ by our choice of δ . Furthermore, z_1 is in a local unstable manifold of z , which in turn is in $W^u(x)$, so z_1 is in $W^u(x)$. Let p be the least common multiple of the periods of x and y . Because p is a multiple of the period of x , f^{np} maps $W^u(x) \cap \Lambda$ to itself for all n . Also, because p is a multiple of the period of y , and z_1 is in $W^s(y)$, $f^{np}(z_1)$ converges to y as $n \rightarrow \infty$, and this contradicts the assumption that $y \notin \overline{W^u(x) \cap \Lambda}$.

We can now complete the proof of Lemma 4.3. Suppose that x is a periodic point of Λ but that $W^u(x) \cap \Lambda$ is not dense in Λ . By Lemma 4.4, both $\overline{W^u(x) \cap \Lambda}$ (which we will denote by $\mathcal{U}$) and $\mathcal{V} = \Lambda \setminus \mathcal{U}$ are relatively open subsets of Λ . If p is the period of x , then $f^p\mathcal{U} = \mathcal{U}$, so $f^{np}\mathcal{U}$ is disjoint from $\mathcal{V}$ for all n , and thus f is not topologically mixing on Λ .

The next major step in the proof is the following:

LEMMA 4.5. *Let Λ, f be as in the statement of the theorem, and let x, y be periodic points for f in Λ . If y is in $\overline{W^u(x) \cap \Lambda}$ then*

$$\overline{W^u(y) \cap \Lambda} = \overline{W^u(x) \cap \Lambda}.$$

Before proving this lemma, we note two immediate consequences: First, if f^n maps $W^u(x) \cap \Lambda$ to itself—which will certainly be true for n the period of x , and perhaps true for some proper divisor of the period—then $\overline{W^u(x) \cap \Lambda}$ is a hyperbolic set for f^n in which the unstable manifolds of all periodic points are dense, and hence f^n is topologically mixing on $\overline{W^u(x) \cap \Lambda}$ by Lemma 4.3. Second: If x and y are periodic points in Λ , and if $\overline{W^u(x) \cap \Lambda}$ and $\overline{W^u(y) \cap \Lambda}$ intersect, then there is a periodic point z in their intersection; then applying Lemma 4.5 to the pairs x, z and y, z we conclude that

$$\overline{W^u(x) \cap \Lambda} = \overline{W^u(z) \cap \Lambda} = \overline{W^u(y) \cap \Lambda};$$

closures of unstable manifolds are either disjoint or equal.

PROOF OF LEMMA 4.5: We first show that

$$W^u(y) \cap \Lambda \subset \overline{W^u(x) \cap \Lambda}.$$

Since $\overline{W^u(x) \cap \Lambda}$ is open and contains y , there exists $\eta > 0$ such that

$$W_\eta^u(y) \cap \Lambda \subset \overline{W^u(x) \cap \Lambda}.$$

If p is the least common multiple of the periods of x and y then f^p maps $\overline{W^u(x) \cap \Lambda}$ onto itself and

$$W^u(y) \cap \Lambda = \bigcup_{n=0}^{\infty} f^{np} W^u_\eta(y);$$

hence

$$W^u_\eta(y) \cap \Lambda \subset \overline{W^u(x) \cap \Lambda},$$

as desired. Since the set on the right-hand side is closed, the inclusion remains true if we replace the set on the left by its closure.

To complete the proof, we need only prove that x is in $\overline{W^u(y) \cap \Lambda}$; we can then interchange the roles of x and y in the above argument to prove equality. Since on the one hand y is in the closure of $W^u(x) \cap \Lambda$ and on the other hand $\overline{W^u(y) \cap \Lambda}$ is an open neighborhood of y , there must exist z in $W^u(x) \cap \overline{W^u(y) \cap \Lambda}$. Again letting p denote the least common multiple of the periods of x and y , we have

$$f^{-np}(z) \rightarrow x \quad \text{since } z \in W^u(x),$$

and

$$f^{-np}(z) \in \overline{W^u(y) \cap \Lambda} \quad \text{since } f^p \text{ maps } \overline{W^u(y) \cap \Lambda} \text{ onto itself.}$$

Since $\overline{W^u(y) \cap \Lambda}$ is closed, this implies that x is in $\overline{W^u(y) \cap \Lambda}$.

All of the major steps in the proof of the theorem are now in place. We know that the collection of all closures of unstable manifolds of periodic points is a partition of Λ into sets which are simultaneously open and closed. Since Λ is compact, there are only finitely many of these sets. Moreover, f permutes the elements of this partition. Grouping them into orbits of the action of f , and labelling them accordingly, gives a partition of the desired form. Furthermore, as we noted in the paragraph following the statement of Lemma 4.5, since f^{n_i} maps $\Lambda_{i,j}$ to itself, then f^{n_i} is topologically mixing on $\Lambda_{i,j}$.

We next show that the decomposition is essentially unique. Suppose we have another decomposition, and let M be one of the components. Then M is an open and closed subset of Λ , invariant for some f^m and on which f^m is topologically mixing. It then follows from Lemma 4.3 that M must be the closure of the unstable manifold of any one of its periodic points, i.e., M must be one of the $\Lambda_{i,j}$ already constructed.

We next note that, if Λ, f satisfy the hypotheses of the theorem, so do Λ, f^{-1} . By uniqueness, the decomposition for f^{-1} must be the same as the decomposition for f (up to labelling). But the elements of the decomposition for f^{-1} are the closures of the stable manifolds of their periodic points. Hence, for any periodic $x \in \Lambda_{i,j}$ $W^s(x) \cap \Lambda$ is dense in $\Lambda_{i,j}$.

Finally, we want to show that the stable and unstable manifolds of a general—not necessarily periodic—point x of $\Lambda_{i,j}$ are dense in $\Lambda_{i,j}$. As in the preceding paragraph, it will suffice to consider the unstable manifold. Since periodic points are dense, it will suffice to show that any periodic point y of $\Lambda_{i,j}$ is in the closure of $W^u(x) \cap \Lambda$. Replacing f by an appropriate power of itself, we can even assume that y is a fixed point of f . The idea will be to take a large enough n and find a z in the intersection of the stable

manifold of y with the unstable manifold of $f^{-n}(x)$; then $f^n(z)$ will be in $W^u(x)$ (by the covariance of unstable manifolds) and near y (because z was on the stable manifold of y and n was large.) To make this work, it is necessary to control how far away z is from y along the stable manifold. What we do is to choose a small positive η ; then a $\delta > 0$ so that, when z_1 and z_2 are two points of Λ with $\|z_1 - z_2\| < \delta$, $W_\eta^s(z_1)$ and $W_\eta^u(z_2)$ have exactly one point $[z_1, z_2]$ in common, and $[z_1, z_2]$ is in Λ . Since the $f^n(W_\eta^s(y) \cap \Lambda)$ increase to $W^s(y) \cap \Lambda$, which is dense in $\Lambda_{i,j}$ by the preceding paragraph, it follows from a simple compactness argument that, if n is large enough, $f^n W_\eta^s(y)$ intersects any ball of radius δ . Take such an n ; then choose $z_1 \in f^{-n}(W_\eta^s(y) \cap \Lambda)$ at a distance less than δ from $f^{-n}(x)$; and let z be $[z_1, f^{-n}(x)]$, the point of intersection of $W_\eta^s(z_1)$ with $W_\eta^u(f^{-n}x)$. Then $f^n(z_1)$ is in $W_\eta^s(y)$, and $f^n(z)$ is in $W_\eta^s(z_1)$; hence, the distance from y to $f^n(z)$ is not greater than 2η . Since $z \in W^u(f^{-n}(x)) \cap \Lambda$, $f^n(z)$ is in $W^u(x) \cap \Lambda$, and since η was any sufficiently small positive number, it follows that $y \in \overline{W^u(x) \cap \Lambda}$, as desired. It is also necessary to show that $\overline{W^u(x) \cap \Lambda}$ is contained in $\Lambda_{i,j}$. This follows from the corresponding fact for periodic x 's by a straightforward approximation argument.

Up until now, we have been studying hyperbolic sets from a somewhat local point of view, i.e., we have been studying the action of a diffeomorphism on a small neighborhood of some hyperbolic set without concerning ourselves with how it acts elsewhere. We now want to change our point of view; the idea here is to assume that the set of all recurrent points—i.e., either $\overline{L(f)}$ or $\Omega(f)$ or $R(f)$ —is hyperbolic, and to see what conclusions can be drawn about the global action of f . The general hypothesis for this chapter is that f is a diffeomorphism, of class at least C^1 , of a compact manifold $\mathcal{M}$.

§1. Periodic points. The first step is to show that hyperbolicity of the set of recurrent points implies density of periodic points in that set. This works for $\overline{L(f)}$ and $R(f)$, but not for $\Omega(f)$.

PROPOSITION 1.1. *Let f be a C^1 diffeomorphism of a compact manifold $\mathcal{M}$. If $\overline{L(f)}$ is a hyperbolic set, then periodic points are dense in $L(f)$.*

PROOF: We use shadowing. Take $\epsilon > 0$ and then a corresponding $\delta < \epsilon$ so that every δ pseudo-orbit in the δ neighborhood of $\overline{L(f)}$ is ϵ -shadowed by a true f orbit. Assume, furthermore, that ϵ is small enough so that the shadowing orbit is unique. We will show that, for any point x in $L(f)$, there is a periodic point y within distance 2ϵ of x . Since $x \in L(f)$, there is a $z \in \mathcal{M}$ such that $x \in \omega(z; f)$. Furthermore, the forward orbit of z is precompact (since we are working in a compact space) and therefore converges to $L(f)$. Thus, replacing z if necessary by a point on its forward orbit, we can assume that the forward orbit of z is contained in the δ neighborhood of $L(f)$, and that the distance from z to x is less than $\delta/2$. If $n > 0$ is such that the distance from $f^n(z)$ to x is again less than $\delta/2$, then

$$\dots, z, f(z), f^2(z), \dots, f^{n-1}(z), z, \dots$$

is a periodic δ pseudo-orbit in the δ -neighborhood of $\overline{L(f)}$. Thus, there is a true periodic orbit $f^n(y)$ which stays for all time within distance ϵ of the pseudo-orbit, and

$$\|y - x\| \leq \|y - z\| + \|z - x\| < \epsilon + \epsilon,$$

as desired. Furthermore, $y \in \overline{L(f)}$ because, by its very definition, $\overline{L(f)}$ contains all periodic points.*

There is an entirely analogous result for the chain recurrent set:

PROPOSITION 1.2. *Let f be a C^1 diffeomorphism of a compact manifold. If $R(f)$ is a hyperbolic set, then periodic points are dense in $R(f)$.*

PROOF: By Proposition ??, there is for any $x \in R(f)$ and any $\delta > 0$ a periodic δ pseudo-orbit in $R(f)$ passing through x . If δ is small enough, this pseudo-orbit is

*Note that we do not yet know that $\overline{L(f)}$ has local product structure and hence cannot argue that all shadowing orbits are necessarily in $\overline{L(f)}$

closely shadowed by a true periodic orbit, i.e., x can be approached as closely as desired by periodic points. These periodic points, as before, are automatically in $R(f)$.

The corresponding assertion for the non-wandering set is *false*; there exists a diffeomorphism f of a compact manifold for which $\Omega(f)$ is hyperbolic but for which periodic points are not dense in $\Omega(f)$. The existence of such a mapping was an open question for many years; a (complicated) example was finally given in A. Danker ("On Smale's Axiom A dynamical systems," *Ann. of Math.* 107 (1978) 517-553.) In his ground-breaking work, out of which the general theory of hyperbolic sets grew, Smale proposed to consider diffeomorphisms for which $\Omega(f)$ is hyperbolic, and for which periodic points are dense in $\Omega(f)$. He was thus led to formulate his *Axiom A*: A diffeomorphism of a compact manifold is said to satisfy Axiom A if its non-wandering set is hyperbolic and if periodic points are dense in its non-wandering set.

§2. Local product structure. It is not at all evident that any one of the three global hyperbolicity hypotheses formulated in the preceding section—hyperbolicity of $\overline{L(f)}$, Axiom A, and hyperbolicity of $R(f)$ —implies that the hyperbolic set in question has local product structure. All three of them do in fact imply local product structure; the tool for showing this is the following theorem, due to Smale:

THEOREM 2.1. *Let f be a C^1 diffeomorphism of a compact manifold, and assume that the closure of the set of periodic points of f is a hyperbolic set. Then this hyperbolic set has local product structure.*

PROOF: By the continuity of the mapping $(x, y) \mapsto [x, y]$, it is enough to prove that $[x, y]$ is a limit of periodic points when x and y are periodic points and $\|x - y\|$ is small. We will prove this by a simple variant on the argument showing that a transverse homoclinic point is a limit of periodic points.

We can first of all, by replacing f by f^n for an appropriate choice of n , assume that x and y are fixed points of f . Recall that $[x, y]$ is the (unique) point of intersection of a reasonable-size local stable manifold of x with a local unstable manifold of y . Thus the forward orbit of $[x, y]$ converges to x , and the backward orbit to y . Furthermore, if the distance from x to y is small enough, $[y, x]$ is also defined; its forward orbit converges to y and its backward orbit to x . We now claim that the set consisting of x, y , the full orbit of $[x, y]$, and the full orbit of $[y, x]$ is a hyperbolic set. We will not give the proof; it is essentially the same as the proof the the set consisting of a transverse homoclinic orbit together with the hyperbolic fixed point to which it converges is hyperbolic. This set certainly does not have local product structure. It is however easy to make periodic pseudo-orbits in this set with arbitrarily small δ : One starts, say, at $[x, y]$ and follows the exact orbit until very close to x ; then makes a small jump to a point near x on the backward orbit of $[y, x]$; then follows the exact orbit through $[y, x]$ until very close to y ; then again makes a small jump to a point on the backward orbit of $[x, y]$, and follows the true orbit back to $[x, y]$. This sequence is then repeated periodically. If the jumps around x and y are made small enough, this periodic pseudo-orbit will be uniquely shadowed by a true orbit, which will have to be periodic by uniqueness, and to pass as near as desired to $[x, y]$ because the pseudo-orbit actually passes *through* $[x, y]$.

COROLLARY 2.2. *Let f be a C^1 diffeomorphism of $\mathcal{M}$*

1. *If $\overline{L(f)}$ is a hyperbolic set, then $\overline{L(f)}$ has local product structure.*
2. *If f satisfies Axiom A (i.e., if $\Omega(f)$ is a hyperbolic set in which periodic points are dense) then $\Omega(f)$ has local product structure.*
3. *If $R(f)$ is a hyperbolic set, then $R(f)$ has local product structure.*

Note one consequence of this corollary: The set $\overline{L(f)}$ or $\Omega(f)$ or $R(f)$ satisfies the hypotheses of Theorem ?? and hence can be split into a finite union of invariant mutually disjoint topologically transitive pieces.

§3. R-stability. We will need to introduce some general terminology to describe various global stability properties of diffeomorphisms. As already noted in §IX.4, a diffeomorphism f of a compact manifold $\mathcal{M}$ is said to be C^r *structurally stable* if, for each g in some C^r neighborhood of f , there is a homeomorphism h_g of $\mathcal{M}$ to itself such that

$$g = h_g \circ f \circ h_g^{-1},$$

and if h_g can be taken to be as close as desired to the identity mapping in the C^0 topology if g is sufficiently close to f in the C^r topology.* Thus, a diffeomorphism is structurally stable if every nearby mapping differs from it only by a continuous change of coordinates, which can furthermore be taken to be near to the identity.

Structural stability is a very natural and powerful property, and we have seen that it holds for Anosov diffeomorphisms (Corollary IX.4.3) was once thought that it might hold for “almost all” diffeomorphisms, but this turned out to be false, and one was therefore led to investigate weaker notions of global stability. One such notion is the following: We say that a C^r diffeomorphism f is C^r *R-stable* if, for each g in some C^r neighborhood of f , there is a homeomorphism h_g from $R(f)$ onto $R(g)$ such that

$$g = h_g \circ f \circ h_g^{-1} \quad \text{on } R(f),$$

and if h_g can be taken to be as close as desired to the identity (as a mapping from $R(f)$ into $\mathcal{M}$) in the C^0 topology if g is sufficiently close to f in the C^r topology. The idea is thus that f is R-stable if the recurrent part of the action of any g near f —understood as meaning the action on its chain-recurrent set—is topologically the same as the recurrent part of the action of f , but that the transient part is allowed to vary topologically as g changes. In the same way, one defines the notion of Ω -stability and $\overline{L}$ stability by replacing in the above definition the chain recurrent set $R(f)$ by the non-wandering set $\Omega(f)$ and the limit set $\overline{L(f)}$ respectively.

One of the high points of the global theory of dynamical systems is the following theorem, which is closely related to the Ω -stability Theorem of Palis and Smale:

*Intuitively, this last condition means that the mapping $g \mapsto h_g$ is continuous from the C^r topology to the C^0 topology. We cannot formulate it in this way, however, since we do not want to put into the definition the requirement that h_g be unique.

THEOREM 3.1. *If f is a C^1 diffeomorphism such that $R(f)$ is hyperbolic, then f is C^1 R -stable.*

PROOF: The main idea in the proof is already contained in Theorem IX.4.2, which we described as asserting the “structural stability of hyperbolic sets.” In the present context, this theorem tells us that, if $R(f)$ is hyperbolic, then any g near enough to f in the C^1 topology admits an invariant set Λ_g near to $R(f)$ on which the action of g is topologically the same as the action of f on $R(f)$. We also know, by Theorem X.2.1, that Λ_g is hyperbolic for g . To prove the theorem, we have only to show that Λ_g is equal to $R(g)$.

Since periodic points for f are dense in $R(f)$ (Proposition 1.2), periodic points for g are dense in Λ_g , and thus Λ_g is contained in $R(g)$. The problem is to show that $R(g)$ is not bigger than Λ_g , i.e., that $R(f)$ does not “explode” when f is perturbed slightly. We break this problem into two parts—we argue separately that g has no chain-recurrent points outside Λ_g

- (1) near $R(f)$.
- (2) away from $R(f)$.

To deal with points near $R(f)$, we have to rework the proof of the structural stability theorem for hyperbolic sets to get a sharper version valid when the initial set has local product structure.

PROPOSITION 3.2. *Let Λ be a hyperbolic set for the C^1 diffeomorphism f ; assume that Λ has local product structure. There then exist*

- (1) *An open neighborhood $\mathcal{U}$ of Λ in $\mathcal{M}$.*
- (2) *A C^1 neighborhood of f in the space of diffeomorphisms of $\mathcal{M}$*

such that, if $g \in \mathcal{V}$,

$$\Lambda_g := \bigcap_{n=-\infty}^{\infty} g^n \mathcal{U}$$

is a closed invariant set for g , and there is a homeomorphism h_g from Λ to Λ_g such that

$$g = h_g \circ f \circ h_g^{-1} \quad \text{on } \Lambda.$$

Furthermore, for any $\epsilon > 0$, if $\mathcal{V}$ is taken small enough, then h_g can be chosen to satisfy

$$\|h_g(x) - x\| < \epsilon \quad \text{for all } x \in \Lambda.$$

REMARKS: The main assertion of the proposition—that Λ_g as defined is an invariant set for g on which the action of g is topologically the same as the action of f on Λ —is perhaps best thought of as divided into two assertions:

- (1) *Any $g \in \mathcal{V}$ admits a closed invariant set Λ_g , contained in $\mathcal{U}$ on which the action of g is topologically the same as the action of f on Λ .*
- (2) *$\mathcal{U}$ is an isolating neighborhood for each such Λ_g .*

We already know the first of these assertions; it is essentially Theorem IX.4.2. We also know that Λ_g is a hyperbolic set, and that it has local product structure (because local product structure is a topological property of the restriction of g to Λ_g , and Λ has local product structure for f .) Thus, each Λ_g has an isolating neighborhood. What is essentially new in this result is thus only a uniformity statement, that we can find a common isolating neighborhood for all the Λ_g 's.

PROOF: The main idea is the same as in the proof of Theorem IX.4.2: We apply the Generalized Shadowing Theorem IX.4.1 to exact orbits $f^n(x)$ for f (with $x \in \Lambda$), regarded as pseudo-orbits for g ; h_g is then constructed by writing the exact g orbit which shadows this one is $g^n(h_g(x))$. We can summarize the results of the proof of Theorem IX.4.1 as follows: There exist $\epsilon_0 > 0$, and, for any $0 < \epsilon < \epsilon_0$ a C^1 neighborhood $\mathcal{V}_\epsilon$ of f such that, if $g \in \mathcal{V}_\epsilon$, then for any x in Λ , there is a unique y in $\mathcal{M}$ such that $\|g^n(y) - f^n(x)\| \leq \epsilon_0$ for all $n \in \mathbb{Z}$, and the y actually satisfies $\|g^n(y) - f^n(x)\| \leq \epsilon$ for all $n \in \mathbb{Z}$. If we define h_g as the function mapping x to the corresponding y , then h_g is a homeomorphism of Λ onto a compact subset Λ_g invariant for g . Note that this construction guarantees that $\|h_g(x) - x\| \leq \epsilon$ for all x in Λ .

Now choose $\epsilon \leq \epsilon_0$, and take a corresponding $\delta > 0$ as in the ordinary Shadowing Theorem, i.e. such that any δ pseudo-orbit for f in the δ neighborhood of Λ is ϵ -shadowed by a unique f -orbit. Since Λ has local product structure, we can arrange by taking ϵ and δ small enough, that this orbit is always in Λ . We then fix δ , and find a C^1 neighborhood $\mathcal{V} \subset \mathcal{V}_\epsilon$ of f small enough so that, for any $g \in \mathcal{V}$, Λ_g is in the δ -neighborhood of Λ and any exact g orbit is a δ pseudo-orbit for f . We then take the δ -neighborhood of Λ for $\mathcal{U}$, and want to show that the assertions of the theorem now hold.

What has to be shown is that, if $g \in \mathcal{V}$ and if $g^n(z) \in \mathcal{U}$ for all $n \in \mathbb{Z}$, then $z \in \Lambda_g$. By the choice of $\mathcal{V}$, $g^n(z)$ is then a δ pseudo-orbit for f , and by the choice of δ , there is an x in Λ such that the f -orbit of x ϵ -shadows $g^n(z)$. But, by the construction of h_g , this implies that $z = h_g(x)$ and thus $z \in \Lambda_g$ as desired.

The above proposition tells us, loosely, that if g is near enough to f , then any g orbit not in $h_g(R(f))$ eventually gets well away from $R(f)$, i.e., outside some fixed neighborhood. We next have to show that points which are not too close to $R(f)$ cannot be chain recurrent for any g near f ; the following proposition provides the tools to do this:

PROPOSITION 3.3. *Let f be a continuous mapping of the metric space (X, d) to itself; x an element of X which is not chain-recurrent for f . Then there exists $\delta > 0$ such that, if g is a mapping of X to itself with $d(f(z), g(z)) < \delta$ for all $z \in X$, and if $d(x, y) < \delta$, then y is not chain-recurrent for g .*

PROOF: Since x is by assumption not chain recurrent, there is a $\delta_1 > 0$ such that there is no δ_1 pseudo-orbit starting and ending at x . Choose $\delta < \delta_1/3$ so the $d(f(x), f(y)) < \delta_1/3$ when $d(x, y) < \delta$. We claim that this δ has the property asserted in the proposition. To see this, suppose that there are g and y as in the proposition such that there is a δ

pseudo-orbit for g from y to itself; let

$$y, y_1, \dots, y_{n-1}, y$$

be such a pseudo-orbit. We claim that

$$x, y_1, y_2, \dots, y_{n-1}, x$$

is then a δ_1 pseudo-orbit for f , and this will contradict the assumption that there is no δ_1 pseudo-orbit for f from x to itself. We check first that

$$d(f(x), y_1) \leq d(f(x), f(y)) + d(f(y), y_1) < \delta_1/3 + \delta < \delta_1.$$

Then, for $i = 1, \dots, n-2$,

$$d(f(y_i), y_{i+1}) \leq d(f(y_i), g(y_i)) + d(g(y_i), y_{i+1}) < 2\delta < \delta_1.$$

Finally,

$$d(f(y_{n-1}), x) \leq d(f(y_{n-1}), g(y_{n-1})) + d(g(y_{n-1}), y_n) + d(y_n, x) < 3\delta < \delta_1.$$

This completes the verification that the indicated sequence is indeed a δ_1 pseudo-orbit for f and hence the proof of the proposition.

We now apply Proposition 3.2 to $R(f)$, which we are assuming hyperbolic; we denote the corresponding $\mathcal{U}$ and $\mathcal{V}$ by $\mathcal{U}_R$ and $\mathcal{V}_R$ respectively. By Proposition 3.3, there is for each $x \in \mathcal{M} \setminus \mathcal{U}_R$ an open neighborhood $\mathcal{W}_x$ of x in $\mathcal{M}$ and a C^0 neighborhood $\mathcal{V}_x$ of f such that, for $g \in \mathcal{V}_x$, no point of $\mathcal{W}_x$ is chain recurrent for g . Since $\mathcal{M} \setminus \mathcal{U}_R$ is compact, we can cover it by a finite number of $\mathcal{W}_{x_i}$. Letting $\mathcal{V}_0$ denote the intersection of the corresponding $\mathcal{V}_{x_i}$, we see that no point outside $\mathcal{U}$ can be chain recurrent for any g in $\mathcal{V}_0$. Now take for $\mathcal{V}$ the intersection of $\mathcal{V}_R$ with $\mathcal{V}_0$; as indicated above, we have only to show that, for $g \in \mathcal{V}$, $R(g) = h_g(R(f))$. Furthermore, since periodic points for f are dense in $R(f)$, periodic points are dense in $h_g(R(f))$ and so $h_g R(f) \subset R(g)$, and all we have to do is to show that no point not in $h_g R(f)$ can be chain recurrent for g . The construction of $\mathcal{V}_0$ shows immediately that no point of $\mathcal{M} \setminus \mathcal{U}_R$ can be chain recurrent for g . Also, by Proposition 3.2, if z is any point of $\mathcal{U}_R$ which is not in $h_g(R(f))$, there must be an n such that $g^n(z) \notin \mathcal{U}_R$ and hence $g^n(z) \notin R(g)$. Since $R(g)$ is invariant for g , $z \notin R(g)$, as desired.

§4. Cycles and Ω -stability. Although the R-stability Theorem of the preceding section is the main global stability result for hyperbolic sets, it is also useful to understand that, for example, Axiom A by itself does not suffice to guarantee Ω -stability but that the addition of a further simple and natural condition does guarantee it.

We begin by sketching how one might construct an example which satisfies Axiom A but is not Ω -stable. The example is two-dimensional and is illustrated in Figure ???. To construct it, we start with a pair of hyperbolic fixed points of saddle type,

which we call A and B . The eigenvalues of the linearization of f at A and B are to be positive so the stable and unstable manifolds of A each consists of two separate curves ("branches") joined at A , and similarly for B . The essential element in the example is that *one branch of the unstable manifold of A is also a branch of the stable manifold of B , and one branch of the unstable manifold of B is also a branch of the stable manifold at A* . In other words, A and B are joined by two *heteroclinic connections*. Note that this is a very "non-generic" situation, but it can certainly be arranged. The two heteroclinic connections, together with the two fixed points, form a simple closed curve which is invariant for f ; the complement of this curve consists of two regions which are individually invariant for f , and which we will refer to as the bounded and unbounded regions.

We further arrange that, in the bounded region, there are two hyperbolic fixed points R and S , the first attracting and the second repelling. The orbit structure in the bounded region is to be as follows: As shown in the figure, the non-heteroclinic branch of the stable manifold of B lies in the bounded region; orbits on this curve converge as $n \rightarrow +\infty$ to B , and they are to converge at $n \rightarrow -\infty$ to A . Similarly, the non-heteroclinic branch of the unstable manifold of B is in the bounded region; orbits on this branch converge to B as $n \rightarrow -\infty$ and to R as $n \rightarrow +\infty$. Every orbit in the bounded region other than the fixed points and those on these two curves is asymptotic to R as $n \rightarrow +\infty$ and to S as $n \rightarrow -\infty$. Similarly, in the unbounded region, there is an attracting-repelling pair P and Q ; all orbits in the unbounded region except Q and those on the stable manifold of A converge to P as $n \rightarrow +\infty$; and all orbits in the unbounded region except P and those on the unstable manifold of A converge to Q as $n \rightarrow -\infty$.

We now want to argue that, with this construction, the non-wandering set of f consists simply of the six fixed points A, B, P, Q, R, S . If this is true, then f certainly satisfies Axiom A. It is more or less clear that the only other points which might turn out to be non-wandering are those on the two heteroclinic connections. We thus want to argue, for example, that points on the arc from A to B —i.e., on the branch of the unstable manifold of A which is also a branch of the stable manifold of B —are wandering. In the first place, the orbit of the point itself certainly goes away and never comes back; it goes "monotonically" to B . Points slightly above the arc are first attracted toward B , but, when they get near enough, they are repelled out again along the upper (non-heteroclinic) branch of the unstable manifold of B and converge to R ; they therefore never return near the arc from A to B , no matter how near to this arc they started out. Points slightly below the arc are more interesting: They are first attracted toward B , but again, when they get near enough, they are repelled out along the unstable manifold. This time, however, they go along the heteroclinic branch of the unstable manifold, which they follow around toward A . Again, if the initial point was close enough to the arc from A to B , the orbit will get very near to A , but again it will eventually be repelled out, this time along the non-heteroclinic branch of the unstable manifold of A , and will therefore be trapped by the attracting fixed point P and never return. Thus, a small enough neighborhood of any point on the arc from A to B will consist entirely of

points whose orbits never return, i.e., each such point is wandering. A similar argument shows that each point on the homoclinic arc from B to A is wandering.

Thus, an f constructed in this way satisfies Axiom A. It is easy to see, however, that such an f can be perturbed arbitrarily little in such a way as to break the homoclinic connections and replace them by transverse intersections of the unstable manifold of A with the stable manifold of B and of the unstable manifold of B with the stable manifold of A . (This can be done, for example, by composing a small "twist" with the action of the mapping in each of two small disks, one along the arc from A to B , the other along the arc from B to A , without changing the mapping outside these disks.) Once f has been perturbed to produce these transverse intersections of stable and unstable manifolds, the argument used in proving Theorem ?? above shows that the intersection points are limits of periodic points for the perturbed mapping. The perturbed mapping thus has infinitely many periodic points and hence a much larger non-wandering set than f itself. In particular, f is not Ω -stable. An occurrence like this is known as an Ω -explosion.

There is a simple condition which rules out Ω -explosions. Let f be a C^1 diffeomorphism satisfying Axiom A.* We can then apply the Spectral Decomposition Theorem (Theorem ??) to decompose

$$\Omega(f) = \Omega^{(1)} \cup \Omega^{(2)} \cup \dots \cup \Omega^{(m)},$$

where the $\Omega^{(i)}$ are mutually disjoint, closed, and invariant, and the restriction of f to each $\Omega^{(i)}$ is topologically transitive. We construct a directed graph with vertices $1, \dots, m$ by drawing an edge from i to j if there is an orbit in $\mathcal{M} \setminus \Omega(f)$ which is asymptotic to $\Omega^{(i)}$ as $n \rightarrow -\infty$ and to $\Omega^{(j)}$ as $n \rightarrow +\infty$. Edges from a vertex to itself are allowed (and can happen). Equivalently, we draw an edge from i to j if there are point $x_- \in \Omega^{(i)}$ and $x_+ \in \Omega^{(j)}$ such that $W^u(x_-)$ and $W^s(x_+)$ intersect (with the point of intersection required to be outside $\Omega^{(i)}$ if $i = j$.) Note that figure ?? is precisely a sketch of the graph associated with f . We say that f has *no cycles* if this graph does not have any cycles (and the example constructed above does have a cycle.) We now have

THEOREM 4.1. *If f satisfies Axiom A, then f has no cycles if and only if $\Omega(f) = R(f)$. If this condition holds, then f is Ω -stable.*

*Another case to which essentially the same analysis applies is that in which the closure of $L(f) \cup L(f^{-1})$ is hyperbolic.

CHAPTER XIV. MARKOV PARTITIONS

In this section, we will discuss Markov partitions, which are the sharpest tools available for analyzing the internal structure of hyperbolic sets. These methods work for hyperbolic sets which have local product structure; it is not necessary that the action of f on Λ be recurrent, i.e., that periodic points be dense in Λ .

§1. Definitions. Let Λ be a hyperbolic set for f , and let $\eta_0 > 0$ be small enough so that the local stable manifolds $W_{\eta_0}^s(x)$ and local unstable manifolds $W_{\eta_0}^u(x)$ are well-behaved. If $x \in \Lambda$, a set of *canonical coordinates* for Λ at x means a homeomorphism from a product space $P_0 \times Q_0$ onto a neighborhood $\mathcal{U}$ of x in Λ such that “horizontal lines”, i.e., subsets of $P_0 \times Q_0$ of the form $P_0 \times \{q\}$, are mapped to local stable manifolds and sets of the form $\{p\} \times Q_0$ to local unstable manifolds.

A little care is needed here, having to do with the distinction between local and global stable and unstable manifolds. Global stable and unstable manifolds tend to wind around Λ in a complicated way; we saw, for example, that if Λ has local product structure and f is topologically mixing on Λ then each stable and unstable manifold is dense in Λ . For canonical coordinates, the idea is that the image of each horizontal line is the full intersection of the image neighborhood $\mathcal{U}$ with a well-behaved local stable manifold of any one of its points. We can make this precise by requiring, in the definition of canonical coordinates, that the image neighborhood $\mathcal{U}$ has diameter less than η_0 , and that, if $(p, q) \in P \times Q$ maps to x , then the image of $P \times \{q\}$ is the intersection of $\mathcal{U}$ with $W_{\eta_0}^s(x)$, and the image of $\{p\} \times Q$ is the intersection of $\mathcal{U}$ with $W_{\eta_0}^u(x)$.

The condition that Λ has local product structure can be reformulated to say that there exist $\delta_0 > 0$ such that Λ admits canonical coordinates at each of its points with image neighborhood containing the ball (in Λ) of radius δ_0 about the point in question.

If Λ is a hyperbolic set with local product structure, a *rectangle* in Λ will mean a subset R of Λ of sufficiently small diameter which is closed under the bracket operation, i.e., such that $x, y \in R$ implies $[x, y] \in R$. This is equivalent to requiring that R has the form $P \times Q$ in canonical coordinates. Note that the image under f of a rectangle of sufficiently small diameter is again a rectangle, taller but narrower than the original one. A rectangle is said to be *proper* if it is the closure of its interior. It will also be useful to have available the following notation: If R is a rectangle with sufficiently small diameter, and if $x \in R$, we write

$$W^s(x, R) := R \cap W_{\eta_0}^s(x) \quad \text{and} \quad W^u(x, R) := R \cap W_{\eta_0}^u(x).$$

In other words, if R corresponds to $P \times Q$ and x to (p, q) in some canonical coordinate system, then $W^s(x, R)$ corresponds to $P \times \{q\}$ and $W^u(x, R)$ to $\{p\} \times Q$.

A *Markov partition* of Λ is a finite covering $\{R_1, \dots, R_p\}$ of Λ by proper rectangles (of sufficiently small diameter) such that

- (1) The interiors of the R_i are mutually disjoint.
- (2) If x is in the interior of R_i and $f(x)$ in the interior of R_j , then

$$f(W^s(x, R_i)) \subset W^s(f(x), R_j); \quad f(W^u(x, R_i)) \supset W^u(f(x), R_j).$$

Condition (2) can be reformulated as follows: In canonical coordinates centered at $f(x)$, write

$$R_j = P \times Q; \quad f(R_i) = P' \times Q';$$

then

$$P' \subset P; \quad Q' \supset Q.$$

Somewhat more loosely, $f(R_i)$ and R_j can overlap as shown in Figure ??a, but not as shown in Figure ??b-d.

We will give one more reformulation of (2) which indicates better than those already given how globally restrictive it is. Given a rectangle R represented by $P \times Q$ in some canonical coordinate system, we will say that a subrectangle of R is an *s-band* if it has the form $P \times Q'$ for some $Q' \subset Q$, and an *u-band* if it has the form $P' \times Q$ for some $P' \subset P$. Now the interiors of the R_i are disjoint open sets whose union is nearly all of Λ , i.e., all but a closed nowhere dense subset. Condition (2) says that

- (a) Intersecting R_j with the interiors of the $f(R_i)$ splits it into a union of u-bands, with a closed nowhere dense subset left over.
- (b) Intersecting $f(R_i)$ with the interiors of the R_j splits it into a union of s-bands, again with a closed nowhere dense subset left over.

Note that (b) is just (a) with f replaced by f^{-1} .

In view of the strength of this condition, it may be surprising that Markov partitions exist. We are going to prove that any hyperbolic set with local product structure admits Markov partitions into rectangles of arbitrarily small diameter. Before doing this, we will show one use that can be made of them.

§2. Topological Markov shifts and symbolic dynamics. A *Markov matrix* will mean a finite square matrix $A = (A_{ij})$ of 0's and 1's such that, for each i , there is at least one j such that $A_{ij} = 1$. If A is an $\ell \times \ell$ Markov matrix, we let $\Sigma(A)$ denote the subset of $\{1, \dots, \ell\}^{\mathbb{Z}}$ consisting of all sequences (i_j) for which $A_{i_j i_{j+1}} = 1$ for all $j \in \mathbb{Z}$. That is: We regard the matrix A as specifying "allowed" and "forbidden" transitions between elements of $\{1, \dots, \ell\}$; the transition from i to j is allowed if $A_{ij} = 1$. Then $\Sigma(A)$ is obtained from the set of all sequences by throwing out all those for which some transition from j th to $j + 1$ st element is forbidden. It is easy to see that $\Sigma(A)$ is a closed subset of $\{1, \dots, \ell\}^{\mathbb{Z}}$ and hence is compact in the product topology. It is also immediate from the definition that $\Sigma(A)$ is invariant for the left shift σ . The topological space $\Sigma(A)$ equipped with the restriction of the left shift is called the *topological Markov shift with transition matrix A*.

The idea that we now want to develop is that the existence of a Markov partition for a hyperbolic set Λ almost permits us to identify the restriction of f to Λ with a topological Markov shift.

PROPOSITION 2.1. *Let Λ be a hyperbolic set with local product structure which admits a Markov partition into rectangles of sufficiently small diameter. Then there exist a Markov matrix A and a continuous mapping π from $\Sigma(A)$ onto Λ such that*

$$f \circ \pi = \pi \circ \sigma$$

which is approximately one-to-one in the sense that the set of x 's in Λ for which $\pi^{-1}\{x\}$ contains more than one point is contained in a countable union of closed nowhere dense sets.

PROOF: Let $\{R_1, \dots, R_p\}$ be a Markov partition of Λ . In the course of the proof, we will need to assume that the diameter of this partition is small, but we will not try to state here how small it has to be. We define a $p \times p$ matrix A by putting A_{ij} equal to 1 if the interior of $f(R_i)$ intersects the interior of R_j and 0 otherwise. Clearly, for each i , there is a j such that $A_{ij} = 1$, so A is a Markov matrix.

We are going to construct the desired π from $\Sigma(A)$ to Λ ; the general idea will be that π maps $i \in \Sigma(A)$ to a point x such that $f^j(x) \in R_{i_j}$ for all j . It is convenient, however, to formulate the construction of π in a slightly more roundabout way. For each $i \in \Sigma(A)$ and each pair r, s of non-negative integers, we put

$$\Delta^{(-r,s)}(i) = \{x \in \Lambda : f^j(x) \in R_{i_j} \text{ for } -r \leq j \leq s.\}$$

We will prove

LEMMA 2.2. *Each $\Delta^{(-r,s)}(i)$ is non-empty, and, as $r, s \rightarrow \infty$ with i fixed, the diameter of $\Delta^{(-r,s)}(i) \rightarrow 0$.*

Before proving this lemma, we show how it implies the proposition. We take $\pi(i)$ to be the unique point in

$$\bigcap_{r,s} \overline{\Delta^{(-r,s)}(i)}.$$

It is immediate that

$$f \circ \pi = \pi \circ \sigma,$$

and the continuity of π follows at once from the fact that the diameter of Δ goes to zero. We let

$$\mathcal{U} := \bigcup_{i=1}^p \overset{\circ}{R}_i;$$

since the R_i cover Λ , $\mathcal{U}$ is dense in Λ . We also let

$$Z := \bigcap_{n=-\infty}^{\infty} f^{-n}(\mathcal{U});$$

Z is a countable intersection of open dense subsets of Λ and, in particular, is dense in Λ . If $x \in Z$, there is a unique i such that $f^j(x) \in R_{i_j}$ for all j , and this certainly implies that $\pi(i) = x$. Thus, the image of π contains a dense set, and this together with the fact that π is continuous and $\Sigma(A)$ compact implies that π is surjective.

We claim, finally, that any $x \in Z$ has only one preimage under π . To see this, we note that, independent of whether $x \in Z$, $x = \pi(i)$ implies $f^j(x) \in R_{i_j}$ for all j . But each $\overset{\circ}{R}_i$ is disjoint from the R_j for $j \neq i$, and hence $x \in Z$ —i.e., each $f^j(x)$ is in some

$\overset{\circ}{R}_i$ —implies that x has only one preimage under π . Thus, the proposition will follow once we have proved the lemma.

PROOF OF LEMMA 2.2: The main thing which has to be proved is that the $\Delta^{(-r,s)}(i)$ are non-empty. Consider, first of all, the case $r = 0$. Then

$$f^j(x) \in \overset{\circ}{R}_{i_j} \quad \text{for } j = 0, 1, \dots, s$$

if and only if

$$x \in \overset{\circ}{R}_{i_0} \quad \text{and} \quad f^j(f(x)) \in \overset{\circ}{R}_{i_{j+1}} \quad \text{for } j = 0, \dots, s-1,$$

i.e.,

$$\Delta^{(0,s)}(i) = \overset{\circ}{R}_{i_0} \cap f^{-1} \left(\Delta^{(0,s-1)}(\sigma i) \right).$$

The idea now is going to be to show by induction on s that $\Delta^{(0,s)}(i)$ is a non-empty s -band in R_{i_0} . This is evident for $s = 0$; the induction step follows from the following lemma (and the fact that $A_{i_0 i_1} = 1$, which means that $f(\overset{\circ}{R}_{i_0})$ intersects $\overset{\circ}{R}_{i_1}$):

LEMMA 2.3. *If $f(R_i)$ intersects R_j , and if Δ is a non-empty s -band in R_j , then*

$$f^{-1}(\Delta) \cap R_i$$

is a non-empty s -band in R_i .

Assuming for the moment Lemma 2.3, we complete the proof of Lemma 2.2. As we have seen, Lemma 2.3 implies that each $\Delta^{(0,s)}(i)$ is a non-empty s -band in $\overset{\circ}{R}_{i_0}$. The same analysis applied to f^{-1} shows that each $\Delta^{(-r,0)}(i)$ is a non-empty u -band in $\overset{\circ}{R}_{i_0}$. In other words, if we write

$$R_{i_0} = P \times Q$$

in local coordinates, then

$$\Delta^{(0,s)}(i) = P \times Q' \quad \text{and} \quad \Delta^{(-r,0)}(i) = P' \times Q$$

for non-empty $P' \subset P$ and $Q' \subset Q$. Thus

$$\Delta^{(-r,s)}(i) = \Delta^{(0,s)}(i) \cap \Delta^{(-r,0)}(i) = P' \times Q' \neq \emptyset.$$

Once we know that the $\Delta^{(-r,s)}$ are non-empty, it follows from the compactness of Λ that

$$\bigcap_{r,s} \overline{\Delta^{(-r,s)}(i)}$$

is non-empty. We want to prove that this intersection actually contains only one point; it will then follow—again by compactness—that the diameter of $\Delta^{(-r,s)}$ goes to zero, and this will complete the proof of Lemma 2.2. Suppose, therefore, that x, x' are both in the intersection. Then $f^j(x)$ and $f^j(x')$ are both in the same R_i , for all j . Thus, if we take the original Markov partition to have a diameter small enough to be an expansion^{vi/hy} constant for f on Λ (i.e, small enough so that any two distinct orbits on Λ get a distance greater than $\max_i\{\text{diam}(R_i)\}$ apart somewhere,) then x and x' must be equal.

To complete the proof of Proposition 2.1, then, it remains only to prove Lemma 2.3.

PROOF OF LEMMA 2.3: We work systematically in canonical coordinates. In one set of these coordinates we write $R_i = P_i \times Q_i$; in (possibly) another set of coordinates, we write $R_j = P_j \times Q_j$. Expressed in these coordinates, f takes the form

$$(p, q) \mapsto (f_s(p), f_u(q)),$$

and Δ the form $P_j \times Q'_j$ with $Q'_j \subset Q_j$. By the definition of Markov partition, $f_u(Q_i) \supset Q_j$ and $f_s(P_i) \subset P_j$. Since $Q'_j \subset Q_j \subset f_u(Q_i)$, $f_u^{-1}(Q'_j) \subset Q_i$, and $f_s^{-1}(P_j) \supset P_i$; hence,

$$f^{-1}(\Delta) \cap R_i = f_u^{-1}(Q'_j) \times P_i$$

and this displays $f^{-1}(\Delta) \cap R_i$ as a non-empty s-band of R_i .

The extent to which the mapping π of Proposition 2.1 fails to be one-to-one can be estimated much more sharply than is indicated in Proposition 2.1

PROPOSITION 2.4. *Let Λ be a hyperbolic set with local product structure. If $\{R_1, \dots, R_p\}$ is a Markov partition of Λ with sufficiently small maximum diameter, and if π denotes the mapping from a Markov shift space to Λ constructed as in Proposition 2.1 from this partition, then no point of Λ has more than p^2 preimages under π .*

PROOF: We use the notation introduced in the proof of Proposition 2.1. The main step in the argument is to prove:

LEMMA 2.5. *Let $i, i' \in \Sigma(A)$ be such that $\pi(i) = \pi(i')$, and assume that $i_n = i'_n$, $i_m = i'_m$, where $n < m$. Then $i_j = i'_j$ for $m \leq j \leq n$.*

PROOF: Without loss of generality, we can assume $n = 0$. By Lemma 2.2, there exist x such that $f^j(x) \in R_{i_j}$ for $0 \leq j \leq m$ and y such that $f^j(y) \in R_{i'_j}$ for $0 \leq j \leq m$.

Since $i_0 = i'_0$, $[x, y] \in R_{i_0}$. Furthermore, since $\pi(i) = \pi(i')$, R_{i_j} and $R_{i'_j}$ must intersect for each j , and thus

$$\|f^j(x) - f^j(y)\| \leq 2 \max_i \{\text{diam}(R_i)\} \quad \text{for } 0 \leq j \leq m.$$

If $\max_i \{\text{diam}(R_i)\}$ is small enough, this implies that $[f^j(x), f^j(y)]$ is defined for $0 \leq j \leq m$. Once we know that these $[f^j(x), f^j(y)]$ are defined, it follows easily from the definition of the bracket operation that $[f^j(x), f^j(y)] = f^j([x, y])$. On the other hand, from the Markov property and $[x, y] \in R_{i_0} \cap W_{\eta_0}^s(x)$ it follows that $f^j([x, y]) \in R_{i_j}$.

It will suffice to prove that $f([x, y]) \in \overset{\circ}{R}_{i_1}$; the proof for general j then follows by induction. The Markov property implies immediately that $f([x, y]) \in R_{i_1}$; what we have to show is that it is actually in the interior of R_{i_1} . This follows from a general fact about Markov partitions: Let j, k be two indices such that $\overset{\circ}{R}_j \cap f^{-1}(\overset{\circ}{R}_k)$ is non-empty. In appropriate canonical coordinates, write $R_j = P_j \times Q_j$, $R_k = P_k \times Q_k$ and $f : (p, q) \mapsto (f_s(p), f_u(q))$. Then f_s is a (local) homeomorphism and, by the definition of Markov partition, maps P_j into P_k . Hence, it maps the interior of P_j into the interior of P_k , i.e.,

$$f(W^s(x, \overset{\circ}{R}_j)) \subset W^s(f(x), \overset{\circ}{R}_k).$$

We now make the same argument with f^{-1} replacing f and working down from the top. That is, we write

$$[f^j(x), f^j(y)] = f^{-(m-j)}[f^m(x), f^m(y)],$$

and we note that $[f^m(x), f^m(y)] \in W^u(f^m(y), \overset{\circ}{R}_{i'_m})$ since $f^m(x)$ and $f^m(y)$ are both in $\overset{\circ}{R}_{i_m} = \overset{\circ}{R}_{i'_m}$. From this we conclude that $[f^j(x), f^j(y)]$ is in $\overset{\circ}{R}_{i'_j}$. Since we have just seen that it is also in $\overset{\circ}{R}_{i_j}$, and since the interiors of the R_k are mutually disjoint, it follows that $i'_j = i_j$. Since j is any integer in the range $0 \leq j \leq m$, the lemma is proved.

We can now easily complete the proof of Proposition 2.4. Suppose $x \in \Lambda$ has more than p^2 preimages. Then, for sufficiently large m , it has more than p^2 preimages which differ pairwise on $-m \leq j \leq m$. But there are only p^2 possible choices for (i_{-m}, i_m) , so some pair of these preimages must agree at both $-m$ and m . Lemma 2.5 then implies that they must agree for all j from $-m$ to m , and this is a contradiction.

§3. Existence of Markov partitions.

THEOREM 3.1 (BOWEN). *A hyperbolic set Λ with local product structure admits Markov partitions of arbitrarily small diameter.*

The proof is fairly long and quite subtle. The construction will involve a number of smallness conditions which will be imposed as we go along. To start with, we choose η_0 small enough so that local stable manifolds $W_{\eta_0}^s(x)$ and $W_{\eta_0}^u(x)$ are smooth for all $x \in \Lambda$; so that $W_{\eta_0}^s(x_1)$ and $W_{\eta_0}^u(x_2)$ never have more than one point in common; and so that, if x_1, x_2 are distinct points of Λ then $\|f^n(x_1) - f^n(x_2)\| > \eta_0$ for some $n \in \mathbb{Z}$. We then choose $\delta_0 \leq \eta_0/2$ small enough so that, at each point x of Λ , there are canonical coordinates at x whose image contains the ball of radius δ_0 about x . All rectangles considered will have diameter much smaller than δ_0 , and in particular smaller than $\delta_0/\text{Lip}(f)$ so that the image under f of any rectangle we consider can be represented in canonical coordinates.

The first step is to use essentially the reverse of the proof of Proposition 2.1, plus shadowing, to construct a covering of Λ by closed rectangles which fit together almost as nicely as do the elements of a Markov partition, but which need not be proper and whose interiors need not be disjoint. The idea is to cover Λ with a finite number of very small balls with centers at $y_1, \dots, y_p$; then to define a $p \times p$ Markov matrix A by putting

$A_{ij} = 1$ if the distance from $f(y_i)$ to y_{i+1} is small enough; then to map $i \in \Sigma(A)$ to the exact orbit which shadows the pseudo-orbit y_i . If things are set up correctly, this mapping is continuous and surjective (but it is usually very far from being injective.) The elements of the desired "Markov-partition like" covering are just the images of the sets $\{i : i_0 = k\}$ under this mapping.

The details go as follows: We start by choosing $\epsilon > 0$ smaller than δ_0 and also smaller than half the maximum diameter desired for rectangles making up the Markov partition to be constructed; then choose $\delta < \epsilon$ small enough so that every δ pseudo-orbit in Λ is ϵ -shadowed by a unique exact orbit in Λ . Then take $\gamma < \delta/2$ so that

$$\|x_1 - x_2\| < \gamma \quad \text{implies} \quad \|f(x_1) - f(x_2)\| < \delta/2.$$

Choose a finite set $\{y_1, y_2, \dots, y_m\}$ of points in Λ such that the balls of radius γ about the y_i cover Λ . Put

$$A_{ij} = \begin{cases} 1 & \text{if } \|y_j - f(y_i)\| < \delta \\ 0 & \text{otherwise.} \end{cases}$$

If $i \in \Sigma(A)$, (y_i) is a δ pseudo-orbit in Λ and so is ϵ -shadowed by a unique true orbit. We write $\theta(i)$ for the time-zero position of the shadowing orbit, so θ maps $\Sigma(A)$ into Λ .

LEMMA 3.2.

- (1) $f \circ \theta = \theta \circ \sigma$
- (2) θ maps $\Sigma(A)$ onto all of Λ
- (3) θ is continuous.

The first assertion is immediate from the construction and the uniqueness of shadowing. To prove (2), we pick $x \in \Lambda$ and show how to find $i \in \Sigma(A)$ which is mapped by θ to x . For each j , choose i_j so that

$$\|f^j(x) - y_{i_j}\| < \gamma.$$

We will show that i is in $\Sigma(A)$, then $\theta(i)$ is certainly equal to x since the sequence y_{i_j} is constructed to ϵ -shadow the orbit of x . In view of the definition of A , what we have to do is to show that

$$\|y_{i_{j+1}} - f(y_{i_j})\| < \delta$$

for all j . But

$$\begin{aligned} \|y_{i_{j+1}} - f(y_{i_j})\| &\leq \|y_{i_{j+1}} - f^{j+1}(x)\| + \|f^{j+1}(x) - f(y_{i_j})\| \\ &< \gamma + \|f(f^j(x)) - f(y_{i_j})\| < \gamma + \delta/2 \leq \delta, \end{aligned}$$

as desired.

The proof of (3) uses a device which we have seen several times before: Uniqueness of the shadowing orbit together with compactness implies continuity. We first note that,

by compactness, it suffices to show that, if $\mathbf{i}^{(n)} \rightarrow \mathbf{i}$ and if $\theta(\mathbf{i}^{(n)}) \rightarrow x$, then $x = \theta(\mathbf{i})$. Next, by the construction of the $\theta(\mathbf{i}^{(n)})$ and passage to limits,

$$\|f^j(x) - y_{i_j}\| \leq \epsilon \quad \text{for all } j.$$

On the other hand, again by the construction of θ ,

$$\|f^j(\theta(\mathbf{i})) - y_{i_j}\| < \epsilon,$$

so

$$\|f^j(\theta(\mathbf{i})) - f^j(x)\| < 2\epsilon \quad \text{for all } j \in \mathbb{Z}.$$

Since we took ϵ small enough so that 2ϵ is an expansivity constant, it follows that $x = \theta(\mathbf{i})$, as desired.

It should be clear that, although θ has many of the nice properties of the mapping π constructed in the proof of Proposition ??1, it will normally be very far from injective. Roughly speaking, it becomes less and less injective as γ is made smaller and smaller, and we may be forced to take γ quite small in order to guarantee surjectivity.

For each $j = 1, \dots, m$, we put

$$T_k := \theta(\{\mathbf{i} \in \Sigma(A) : i_0 = k\}).$$

Since $\{\mathbf{i} \in \Sigma(A) : i_0 = k\}$ is closed and hence compact in the product topology, and since θ is continuous, T_k is compact and hence closed in Λ . Since θ is surjective, the T_k cover Λ .

The T_k are going to turn out to be rectangles. Rather than proving this as directly as possible, however, it is advantageous to obtain it as a consequence of some machinery which will be useful for other purposes as well.

LEMMA 3.3. *Let $x = \theta(\mathbf{i})$; $x' = \theta(\mathbf{i}')$. If*

- (1) $i_j = i'_j$ for $j \geq 0$, then $x' \in W_{\eta_0}^s(x)$
- (2) $i_j = i'_j$ for $j \leq 0$, then $x' \in W_{\eta_0}^u(x)$.

PROOF: Consider (1). By the construction of θ ,

$$\|f^j(x) - y_{i_j}\| < \epsilon; \quad \|f^j(x') - y_{i'_j}\| < \epsilon$$

for all j . If $i_j = i'_j$ for $j \geq 0$, we get

$$\|f^j(x) - f^j(x')\| < 2\epsilon \quad \text{for } j \geq 0,$$

and, since $2\epsilon < \eta_0$, this implies $x' \in W_{\eta_0}^s(x)$, as desired. The proof of (2) is completely analogous.

LEMMA 3.4. Let i, i' be in $\Sigma(A)$ and satisfy $i_0 = i'_0$. Define a sequence $[i, i']$ by

$$[i, i']_j = \begin{cases} i_j & \text{for } j \geq 0 \\ i'_j & \text{for } j \leq 0. \end{cases}$$

Then

- (1) $[i, i']$ is in $\Sigma(A)$.
- (2) $\theta([i, i']) = [\theta(i), \theta(i')]$

This assertion may be expressed by saying that θ is a "bracket homomorphism."

PROOF: The first assertion follows immediately from the way $\Sigma(A)$ is defined together with the fact that $i_0 = i'_0$; the second from the preceding claim.

It then follows immediately that, if $x, y \in T_k$, then $[x, y] \in T_k$, i.e., T_k is a rectangle.

The above argument is an instance of the fact that many ideas about hyperbolic sets have useful and non-trivial analogues in Markov shift spaces (and sometimes also in more general sequence spaces.) The analogue in sequence space of a local stable manifold of a sequence i is the set of all sequences i' such that $i'_j = i_j$ for $j \geq 0$; similarly, the analogue of a local unstable manifold is the set of i' such that $i'_j = i_j$ for $j \leq 0$. To pursue the analogy, we denote these two sets by $W_0^s(i)$ and $W_0^u(i)$ respectively. Note that

- (1) If $i' \in W_0^s(i)$ then the sequences $\sigma^n(i)$ and $\sigma^n(i')$ approach each other* as $n \rightarrow +\infty$; if $i' \in W_0^u(i)$, then these sequences approach each other as $n \rightarrow -\infty$.
- (2) $\sigma W_0^s(i) \subset W_0^s(\sigma i)$; $\sigma W_0^u(i) \supset W_0^u(\sigma i)$.

There are also natural analogues of the global stable and unstable manifolds. The global stable manifold of i , for example, is the set of all sequences i' for which there exists j_0 such that $i_j = i'_j$ for all $j \geq j_0$; it can also be described alternatively as either

- (1) $\bigcup_n \sigma^{-n} W_0^s(\sigma^n i)$
- (2) The set of all i' such that $\sigma^n i$ and $\sigma^n i'$ approach each other as $n \rightarrow +\infty$.

We recall that, if R is a sufficiently small rectangle and $x \in R$, we write $W^s(x, R)$ for $W_{\eta_0}^s(x) \cap R$ (and similarly for unstable manifolds); in canonical coordinates, if $R = P \times Q$ and $x = (p, q)$, then $W^s(x, R) = P \times \{q\}$.

LEMMA 3.5. Let $i \in \Sigma(A)$, $x = \theta(i)$, $i_0 = k$, (so $x \in T_k$.) Then

$$W^s(x, T_k) = \theta(W_0^s(i))$$

and

$$W^u(x, T_k) = \theta(W_0^u(i)).$$

PROOF: Consider the statement about stable manifolds. It follows at once from Lemma 3.3 that $\theta(W_0^s(i)) \subset W^s(x, T_k)$. To get the opposite inclusion, let $x' \in W^s(x, T_k)$. Then,

*in the product topology, or, more precisely, in the unique uniform structure compatible with the product topology.

just because $x' \in T_k$, there is an i' with $i_0 = k$ which is mapped by θ to x' . Let $i'' = [i, i']$. By the definition of the bracket operation on $\Sigma(A)$, $i'' \in W_0^s(i)$. On the other hand, $\theta(i'') = [x, x']$, and $[x, x'] = x'$ since $x' \in W_{\eta_0}^s(x)$ by assumption; thus, $x' = \theta(i'')$ is the desired representation of x' . The proof of the assertion for unstable manifolds is completely analogous.

We are next going to argue that the T_k have a formal property very close to the main formal property of a Markov partition. Let k, ℓ be any pair of indices such that $A_{k\ell} = 1$. Then $T_k \cap f^{-1}T_\ell$ is non-empty since it contains any point of the form $\pi(i)$ with $i_0 = k$ and $i_1 = \ell$. Furthermore, the T_j can be made to have small diameter, so we can find a canonical coordinate system whose image in Λ contains both $f(T_k)$ and T_ℓ ; in such a coordinate system, we write $T_\ell = P_\ell \times Q_\ell$ and $f(T_k) = P'_k \times Q'_k$.

LEMMA 3.6 (SEMI-MARKOV PROPERTY). *If $A_{k\ell} = 1$ then, with the above notation*

$$(3.1) \quad P'_\ell \subset P_\ell; \quad Q'_\ell \supset Q_\ell.$$

If x is any point of T_k such that $f(x) \in T_\ell$, then

$$(3.2) \quad f(W^s(x, T_k)) \subset W^s(f(x), T_\ell) \quad \text{and} \quad f(W^u(x, T_k)) \supset W^u(f(x), T_\ell).$$

Note that the conclusion of this lemma is not a property of the T_i 's alone; it also depends on the matrix A_{jk} . We do not assert that the conclusions hold whenever $f(T_k)$ intersects T_ℓ .

PROOF: The first task is to work out the relation between (3.1) and (3.2). For any $x \in T_k \cap f^{-1}T_\ell$, we have (expressed in the same canonical coordinates as above, and writing $f(x) = (p, q)$)

$$f(W^s(x, T_k)) = P'_k \times \{q\}; \quad f(W^u(x, T_k)) = \{p\} \times Q'_k,$$

and

$$W^s(f(x), T_\ell) = P_\ell \times \{q\}; \quad W^u(f(x), T_\ell) = \{p\} \times Q_\ell.$$

Thus, (3.1) follows if we prove (3.2) for *one* $x \in T_k \cap f^{-1}(T_\ell)$, and implies (3.2) for all such x . As noted above, there exists $i \in \Sigma(A)$ with $i_0 = k$ and $i_1 = \ell$; we choose one such i and write x for $\theta(i)$. It is immediate from the definitions that $x \in T_k \cap f^{-1}(T_\ell)$.

The first assertion of (3.2) for this x is now more or less immediate:

$$\begin{aligned} f(W^s(x, T_k)) &= f\theta W_0^s(i) \quad \text{by Lemma ??} \\ &= \theta\sigma W_0^s(i) = \theta\sigma\{i' : i'_j = i_j, j \geq 0\} \\ &= \theta\{i'' : i''_j = (\sigma i)_j, j \geq -1\} \\ &\subset \theta\{i'' : i''_j = (\sigma i)_j, j \geq 0\} = \theta W_0^s(\sigma i) \\ &= W^s(f(x), T_\ell) \quad \text{again by Lemma ??}. \end{aligned}$$

The proof of the second assertion is completely analogous.

In spite of the fact, implied by the preceding lemma, that the T_k intersect their images under f in almost as well-organized a way as do the elements of a Markov partition, they may fail to constitute a Markov partition in a number of ways:

- (1) The T_k need not be proper rectangles, i.e., need not be the closure of their respective interiors.
- (2) The interiors of the T_k need not be disjoint.
- (3) It can happen that the interior of $f(T_j)$ intersects the interior of T_k even if $A_{jk} = 0$.

We will remedy these defects by cutting up the T_k in a systematic way. This part of the construction is an exercise in the geometry of rectangles; it does not involve f .

Let j, k be a pair of indices such that $T_j \cap T_k \neq \emptyset$. (We need to include here $j = k$, although the construction we are going to give is trivial in this case.) In canonical coordinates, we can write

$$T_j = P_j \times Q_j; \quad T_k = P_k \times Q_k,$$

and define:

$$T_{jk}^{(1)} = (P_j \cap P_k) \times (Q_j \times Q_k) (= T_j \cap T_k);$$

$$T_{jk}^{(2)} = (P_j \setminus P_k) \times (Q_j \times Q_k)$$

$$T_{jk}^{(3)} = (P_j \cap P_k) \times (Q_j \setminus Q_k)$$

$$T_{jk}^{(4)} = (P_j \setminus P_k) \times (Q_j \setminus Q_k).$$

Each of these sets is a rectangle; the $T_{jk}^{(n)}$ with fixed j, k are mutually, disjoint; and

$$T_j = \bigcup_{n=1}^4 T_{jk}^{(n)}.$$

There are alternative characterizations of the $T_{jk}^{(n)}$ which are better adapted to doing some kinds of computations, e.g.,

$$T_{jk}^{(2)} = \{x \in T_j : W_{\eta_0}^s(x) \cap T_k = \emptyset; \quad W_{\eta_0}^u(x) \cap T_k \neq \emptyset\}.$$

The idea is now going to be that the closures of the minimal non-empty intersections of interiors of sets $T_{jk}^{(n)}$ form a Markov partition. At some places in the construction, it is necessary to avoid boundary points, and so we define

$$Z := \Lambda \setminus \left\{ \bigcup_{j,k,n} \partial T_{jk}^{(n)} \right\}.$$

We need to know that Z is, in an appropriate sense, "large".

LEMMA 3.7. Z is an open dense subset of Λ . If R is any rectangle, then $R \setminus Z$ can be covered by a finite union of closed rectangles $\hat{P}_k \times \hat{Q}_k$ where, for each k , at least one of $\hat{P}_k$ or $\hat{Q}_k$ is nowhere dense.

PROOF: We need to use the following elementary topological facts:

1. For any rectangle $P \times Q$,

$$\partial(P \times Q) = (\partial P \times \overline{Q}) \cup (\overline{P} \times \partial Q)$$

2. For any two subsets X, Y of a topological space

$$\begin{aligned}\partial(X \cap Y) &\subset \partial X \cup \partial Y \\ \partial(X \setminus Y) &\subset \partial X \cup \partial Y.\end{aligned}$$

3. If X is a closed subset of a topological space, then ∂X is nowhere dense.

For any pair j, k of indices such that $T_j \cap T_k \neq \emptyset$, we represent as above $T_j = P_j \times Q_j$ and $T_k = P_k \times Q_k$. Then, from 1. and 2. above,

$$\bigcup_{n=1}^4 \partial T_{jk}^{(n)} \subset ((\partial P_j \cup \partial P_k) \times Q_j) \cup (P_j \times (\partial Q_j \cup \partial Q_k)).$$

Since P_j, P_k, Q_j , and Q_k are closed, their boundaries are nowhere dense, and from this all the assertions follow.

Each $x \in \Lambda$ is in some T_j and hence in some $T_{jk}^{(n)}$. (As noted above, we allow $j = k$, in which case $T_{jj}^{(1)} = T_j$.) An $x \in Z$ is in the interior of any $T_{jk}^{(n)}$ which contains it; for such x , we define $R(x)$ to be the intersection of the interiors of all the $T_{jk}^{(n)}$ which contain x . Since $R(x)$ is a finite intersection of open rectangles, it is itself an open rectangle, and there are only finitely many different $R(x)$. As already indicated, we are going to argue that the closures of the set of all the distinct $R(x)$'s, $x \in Z$, form a Markov partition of Λ .

The above description of $R(x)$ can be rephrased as follows: For each j such that T_j contains x , and each k such that $T_j \cap T_k \neq \emptyset$, there is exactly one $n = n(j, k, x)$ such that x is in $T_{jk}^{(n)}$. Hence,

$$(3.3) \quad R(x) = \bigcap_{j: x \in T_j} \bigcap_{k: T_k \cap T_j \neq \emptyset} \overset{\circ}{T}_{jk}^{(n(j, k, x))}.$$

The first thing which needs to be proved is:

LEMMA 3.8. If x, y are in Z , and if $R(x)$ intersects $R(y)$, then $R(x) = R(y)$.

PROOF: We first argue that it will suffice to consider the special case where $y \in R(x)$. To reduce the general case to this one, we note that, if $R(x)$ and $R(y)$ intersect, their intersection is an open set and hence contains a $z \in Z$. Applying the result for the special case twice, we show that $R(x) = R(z)$ and that $R(y) = R(z)$ and hence that $R(x) = R(y)$.

Thus, suppose $y \in R(x)$. From the definition of the $R(x)$, y belongs to all the $T_{jk}^{(n)}$ which contain x ; what we have to show is that it is not contained in any other $T_{jk}^{(n)}$. The first step will be to show that the j 's such that $x \in T_j$ are the same as those such that $y \in T_j$. In one direction, this is easy; y is in $R(x)$, which is contained in the intersection of all the T_j 's which contain x , so any T_j which contains x also contains y . Conversely, we assume that T_k contains y and want to show that $x \in T_k$. Since $y \in R(x)$, T_k must intersect each T_j which contains x ; we choose one such j . The four sets $T_{jk}^{(n)}$, $n = 1, 2, 3, 4$, are mutually disjoint subsets of T_j , and x and y have to be in the same one of them since both x and y are in T_j and y is in $R(x)$. But we know that $y \in T_j \cap T_k = T_{jk}^{(1)}$, so n must equal 1, so $x \in T_{jk}^{(1)} = T_j \cap T_k$, so $x \in T_k$.

The proof of the proposition now follows almost immediately from the the formula (3.3) for $R(x)$. The j 's—and hence also the k 's—appearing in the formulas for $R(x)$ and $R(y)$ are the same, by what we have just shown. But for any such pair j, k , $y \in T_{jk}^{(n(x,j,k))}$ (because $y \in R(x)$), and this implies $n(y, j, k) = n(x, j, k)$, which, when reinserted into (3.3) with y replacing x , gives $R(y) = R(x)$, as desired.

Let $\{R(x_1), \dots, R(x_p)\}$ be an enumeration of the distinct $R(x)$'s for $x \in Z$, and write R_i for the closure of $R(x_i)$, $1 \leq i \leq p$. The union of the R_i is closed and contains the dense open set Z ; hence, is all of Λ . By the preceding lemma, the $R(x_i)$ are mutually disjoint, and this implies that the interiors of the R_i are mutually disjoint.

This is a general fact from point set topology: If U and V are disjoint open sets in some topological space, then the interiors of the closures of U and V , although they may be bigger than U and V , are still disjoint. To see this, we argue that

$$\begin{aligned} U \cap V = \emptyset \quad &\Rightarrow \quad \overline{U} \cap V = \emptyset \quad \Rightarrow \quad \overset{\circ}{\overline{U}} \cap V = \emptyset \\ &\Rightarrow \quad \overset{\circ}{\overline{U}} \cap \overline{V} = \emptyset \quad \Rightarrow \quad \overset{\circ}{\overline{U}} \cap \overset{\circ}{\overline{V}} = \emptyset \end{aligned}$$

Thus: $\{R_1, \dots, R_p\}$ is a finite cover of Λ by proper rectangles with mutually disjoint interiors. To prove that it is a Markov partition, we have to prove that, if $\overset{\circ}{R}_i$ intersects $f^{-1}\overset{\circ}{R}_j$, and if $x \in R_i$, then

(Ms) $f(W^s(x, R_i)) \subset W^s(f(x), R_j),$

and

(Mu) $f(W^u(x, R_i)) \supset W^u(f(x), R_j).$

These two statements require separate proofs, but the proof of (Mu) is essentially the proof of (Ms) with f^{-1} in place of f . We will give the argument only for (Ms).

Because the sets involved are all rectangles, to prove (Ms) it will suffice to prove the apparently weaker statement that, for *one* x in R_i , $f(W^s(x, R_i))$ is contained in R_j . We will choose the x we work on, first of all, to lie in $\overset{\circ}{R}_i \cap f^{-1}(\overset{\circ}{R}_j)$. Since Z is dense, we can also require it to be in $Z \cap f^{-1}(Z)$, in which case we have $R_i = \overline{R(x)}$ and $R_j = \overline{R(f(x))}$. We claim, in fact, the x can also be chosen so that

$$(3.4) \quad W^s(x, R_i) \cap Z \cap f^{-1}(Z) \text{ is dense in } W^s(x, R_i).$$

To see this, we write R_i in canonical coordinates as $P_i \times Q_i$. It follows from Lemma 3.?? that $R_i \setminus (Z \cap f^{-1}(Z))$ is contained in a set of the form $P' \times Q_i \cup P_i \times Q'$, where P' and Q' are closed and nowhere dense; then any $x \notin P_i \times Q'$ satisfies (3.4) since

$$W^s(x, R_i) \cap Z \cap f^{-1}(Z) \supset (P_i \setminus P') \times \{q\} \quad \text{for } x = (p, q) \text{ with } q \notin Q'.$$

Thus, we choose x in

$$\left(\overset{\circ}{R}_i \cap f^{-1}(\overset{\circ}{R}_j) \cap Z \cap f^{-1}(Z) \right) \setminus (P_i \times Q')$$

and this is possible since this latter set is dense in $\overset{\circ}{R}_i \cap f^{-1}(\overset{\circ}{R}_j)$. It then suffices to prove for such an x that

$$f(W^s(x, R_i) \cap Z \cap f^{-1}(Z)) \subset R_j,$$

i.e., to prove

LEMMA 3.9. *If x, y are in $Z \cap f^{-1}(Z)$, if $R(y) = R(x)$ (which is equivalent to $y \in R(x)$), and if $y \in W_{\eta_0}^s(x)$, then $f(y) \in R(f(x))$ (which implies $R(f(x)) = R(f(y))$).*

PROOF: The first step is to show that $f(x)$ belongs to the same T_j 's as does $f(y)$. Because of the symmetry of the hypotheses in x and y , it will suffice to show that

$$f(x) \in T_j \text{ implies } f(y) \in T_j.$$

Choose i such that

$$\theta(i) = x \quad \text{and} \quad i_1 = j.$$

Denote i_0 by k ; then $x \in T_k$ and so, since $R(y) = R(x)$, $y \in T_k$ also. But y is also in $W_{\eta_0}^s(x)$, i.e., $y \in W^s(x, R_k)$. By Lemma 3.?, there is an $i' \in W_0^s(i)$ such that $y = \theta(i')$. But from the definition of W_0^s it follows that $i'_1 = i_1 = j$ and hence that $f(y) \in R_j$.

Now fix j such that T_j contains $f(x)$ (and hence also $f(y)$), and choose k so that $T_j \cap T_k \neq \emptyset$. We want to show that $f(x)$ and $f(y)$ both belong to the same $T_{jk}^{(n)}$. To do this, it suffices to show that

$$W_{\eta_0}^s(f(x)) \cap T_k \neq \emptyset \quad \text{if and only if} \quad W_{\eta_0}^s(f(y)) \cap T_k \neq \emptyset$$

and

$$W_{\eta_0}^u(f(x)) \cap T_k \neq \emptyset \quad \text{if and only if} \quad W_{\eta_0}^u(f(y)) \cap T_k \neq \emptyset.$$

The first of these is immediate: Since $y \in W_{\eta_0}^s(x)$, $f(y) \in W_{\eta_0}^s(f(x))$, so $W_{\eta_0}^s(f(x))$ and $W_{\eta_0}^s(f(y))$ agree on a neighborhood of $f(x)$ which we can take to be large enough to contain T_j and T_k (by taking the diameters of the T_i to be sufficiently small compared to η_0 .)

To prove the second, it will suffice, again by the symmetry of the hypotheses in x and y , to show that

$$W_{\eta_0}^u(f(x)) \cap T_k \neq \emptyset \quad \text{implies} \quad W_{\eta_0}^u(f(y)) \cap T_k \neq \emptyset.$$

Assume, therefore, that $W_{\eta_0}^u(f(x))$ intersects T_k . Writing T_j and T_k in canonical coordinates and using $f(x) \in T_j$, it is then easy to see that $W_{\eta_0}^u(f(x))$ intersects $T_j \cap T_k$. Take a z such that $f(z) \in W_{\eta_0}^u(f(x)) \cap T_j \cap T_k$. We are going to show that $[f(z), f(y)]$ is again in T_k ; since it is also in $W_{\eta_0}^u(f(y))$ (by the definition of the bracket operation), this will show that $W_{\eta_0}^u(f(y))$ intersects T_k , as desired.

Choose i, i' so that

$$x = \theta(i) \quad \text{and} \quad i_1 = j, \text{ and write } r \text{ for } i_0;$$

$$z = \theta(i') \quad \text{and} \quad i'_1 = k, \text{ and write } s \text{ for } i'_0.$$

Such representations are possible since, e.g., $f(x) \in T_j$. Since $f(z) \in W^u(f(x), T_j)$, the semi-Markov property (Lemma 3.6) implies that $z \in W^u(x, T_r)$. But, on the other hand, $z \in T_s$, and hence

$$W^u(x, T_r) \cap T_s \neq \emptyset.$$

Now we use the hypothesis that $R(x) = R(y)$; this implies that x and y are in the same $T_{rs}^{(n)}$, and hence, since $W^u(x, T_r) \cap T_s$ is non-empty, so is $W^u(y, T_r) \cap T_s$. Pick any $z'' \in W^u(y, T_r) \cap T_s$ and put $z' = [z, z'']$. Since both z and z'' are in the rectangle T_s , $z' \in T_s$. Furthermore, since $z'' \in W_{\eta_0}^u(y)$, $z' = [z, y]$ by the definition of the bracket operation. Now we use the semi-Markov property a second time: $z' \in W^s(z, T_s)$, and $A_{sk} = 1$, so $f(z') \in W^s(f(z), T_k)$. Since $[f(z), f(y)] = f([z, y]) = f(z')$, this shows that $[f(z), f(y)] \in W^s(f(z), T_k)$, as asserted.

§4. Zeta functions. One of the main applications of Markov partitions is to prove a striking "arithmetic" result on the number of periodic points in a hyperbolic set which is usually formulated in terms of the *zeta function* of the hyperbolic set. We can make the definition in general: Let X be a set, f a mapping of X to itself, and write $N_p(f)$ for the number of fixed points of f^p .

The N_p are related in a simple way to the number of periodic orbits of f : Let P_q denote the number of periodic orbits of f with period exactly q , i.e., the number of periodic points of period q is qP_q . Since every fixed point for f^p is a periodic point with some period q which divides p , we have

$$N_p = \sum_{q|p} qP_q.$$

If all the N_p are known, this sum rule can be solved to find successively $P_1, P_2, \dots$.

In fact, it can also be solved explicitly using the *Möbius transform*.

If all the $N_p(f)$ are finite, we define the *zeta function* of f to be the formal power series

$$\zeta(f)(z) = \exp \left(\sum_{p=1}^{\infty} \frac{N_p(f)}{p} z^p \right).$$

(The reason for calling this object a zeta function is that it lies at the end of a chain of analogies, starting from the Riemann zeta function and passing through the Weil's zeta functions for varieties over finite fields.)

The case where the power series defining the zeta function has a non-zero radius of convergence, and where the analytic function it defines is actually rational, corresponds to a nice behavior of the sequence $N_p(f)$. We can see this as follows: If the zeta function is rational, then since $\zeta(0) = 1$, it can be written in the form

$$\zeta(z) = \frac{\prod_{i=1}^r (1 - \lambda_i z)}{\prod_{j=1}^s (1 - \mu_j z)},$$

where we can assume that $\lambda_i \neq \mu_j$ for all i, j . Thus, for small z , we can take logarithms to get

$$\sum_{p=1}^{\infty} \frac{N_p}{p} z^p = \sum_{i=1}^r \log(1 - \lambda_i z) - \sum_{j=1}^s \log(1 - \mu_j z).$$

Differentiating both sides and reorganizing in a straightforward way gives

$$\left(\sup_{p=1}^{\infty} N_p \right) z^{p-1} = \sum_{p=1}^{\infty} \left[\sum_{j=1}^s \mu_j^p - \sum_{i=1}^r \lambda_i^p \right] z^{p-1},$$

and hence

$$(4.1) \quad N_p = \sum_{j=1}^s \mu_j^p - \sum_{i=1}^r \lambda_i^p \quad \text{for } p = 1, 2, \dots$$

Conversely, if N_p admits a representation of this form—i.e., as a difference of two sums of exponentials—we can reverse the above computation to show that $\zeta(z)$ is rational. Hence: Asserting that the zeta function of f is rational is equivalent to asserting the existence of a formula of the form (4.1) valid for all the N_p .

We now have

THEOREM 4.1 (GUCKENHEIMER AND MANNING). *If Λ is a hyperbolic set for f with local product structure, then the zeta function for f restricted to Λ is rational.*

We will not give a proof of this theorem; it can be found in Chapter X of Shub's book. It was proved ~~simultaneously~~ ^{independently} by Guckenheimer and Manning, Manning's proof ^{using different methods} used symbolic dynamics built using Markov partitions. The proof would have been easy if the projection π could have been made to be one-to-one above the periodic points for

f : It is not difficult to see that, if σ denotes the left shift on $\Sigma(A)$, then $N_p(\sigma)$, the number of periodic sequences in $\Sigma(A)$ with period dividing p , is simply $\text{tr}(A^p)$. Thus, we get the representation

$$N_p(\sigma) = \sum_{i=0}^r \mu_i^p,$$

where $\mu_1, \dots, \mu_r$ are the eigenvalues ~~and generalized eigenvalues~~ of A , listed with multiplicity, and this does indeed have the form (4.1). To get the result for f on Λ , Manning analyzed very precisely the extent to which π fails to be one-to-one above periodic points.

§1. Preliminaries. Let f be a continuous mapping of a topological space X into itself. Recall that we defined an *attracting set* for f to be a subset Λ of X which is strictly invariant for f (i.e., $f\Lambda = \Lambda$) and which admits a neighborhood $\mathcal{U}$ such that $f^n\mathcal{U}$ shrinks down to Λ in the sense that, for any neighborhood $\mathcal{V}$ of Λ , $f^n\mathcal{U} \subset \mathcal{V}$ for all sufficiently large n . We also saw that, if such a $\mathcal{U}$ exists, there also exists one such that $f\mathcal{U} \subset \mathcal{U}$. Although we did not formally assume X to be closed, this is automatic under very mild assumptions on the topology of X . Also, if Λ has a compact neighborhood—in particular, if X is compact—then Λ is attracting if and only if there is a compact neighborhood $\mathcal{U}$ of Λ such that

$$f\mathcal{U} \subset \mathcal{U} \quad \text{and} \quad \Lambda = \bigcap_{n \rightarrow \infty} f^n\mathcal{U}.$$

If Λ is an attracting set, the *basin of attraction* $B(\Lambda)$ of Λ means the set of all x such that $f^n(x) \rightarrow \Lambda$ as $n \rightarrow \infty$. If $\mathcal{U}$ is an open neighborhood of Λ whose forward images shrink down to Λ , then $x \in B(\Lambda)$ if and only if $f^n(x) \in \mathcal{U}$ for some $n \geq 0$, so

$$B(\Lambda) = \bigcup_{n=0}^{\infty} f^{-n}\mathcal{U},$$

and this shows that $B(\Lambda)$ is open.

If f is a C^1 -diffeomorphism for $\mathcal{M}$, a *hyperbolic attractor* for f means a set Λ which is both hyperbolic and attracting.

PROPOSITION 1.1. *Let Λ be a hyperbolic set for the C^1 diffeomorphism f . The following are equivalent:*

- (1) Λ is a hyperbolic attractor.
- (2) Every neighborhood $\mathcal{V}$ of Λ contains a neighborhood $\mathcal{U}$ such that $f\mathcal{U} \subset \mathcal{U}$.
- (3) For all small $\eta > 0$, $\bigcup_{x \in \Lambda} W_{\eta}^s(x)$ is a neighborhood of Λ .

If one—and hence all—of these conditions hold, then

- (4) For each $x \in \Lambda$, $W^u(x) \subset \Lambda$.
- (5) Λ has local product structure.

PROOF: It is clear that (1) $\Rightarrow$ (2), and that (4) $\Rightarrow$ (5). To prove the proposition, we will show that

- i. (2) $\Rightarrow$ (4),
- ii. (2) together with (5) $\Rightarrow$ (3),
- iii. (3) $\Rightarrow$ (1).

i. (2) $\Rightarrow$ (4). Let $x \in \Lambda$, $y \in W^u(x)$, and let $\mathcal{U}$ be a neighborhood of Λ such that $f\mathcal{U} \subset \mathcal{U}$. Since $\|f^{-n}(y) - f^{-n}(x)\| \rightarrow 0$, $f^{-n}(y) \in \mathcal{U}$ for large enough n . Pick one such n ; then $y = f^n(f^{-n}(y)) \in \mathcal{U}$. This is true for all invariant neighborhoods $\mathcal{U}$, and (2)

asserts that Λ is the intersection of its invariant neighborhoods; hence, $y \in \Lambda$. Since y was any point of $\bigcup_{x \in \Lambda} W^u(x)$, the assertion follows.

ii. (2) together with (5) implies (3). From Proposition ??, it follows from the local product structure of Λ that any orbit $f^n(y)$ which stays near Λ for all positive n is actually in $W^s(x)$ for some $x \in \Lambda$. Examination of the proof of this proposition shows that, for any $\epsilon > 0$, there exists $\delta > 0$ such that, if the distance from $f^n(y)$ to Λ is less than δ for all $n \geq 0$, then $y \in W_\epsilon^s(x)$ for some x in Λ . By (2), there is a neighborhood $\mathcal{U}$ of Λ contained in the δ neighborhood of Λ and such that $f\mathcal{U} \subset \mathcal{U}$. For any $y \in \mathcal{U}$, the forward orbit $f^n(y)$ stays in $\mathcal{U}$; hence, in the δ neighborhood of Λ ; hence, $y \in W_\epsilon^s(x)$ for some $x \in \Lambda$. Thus,

$$\bigcup_{x \in \Lambda} W_\epsilon^s(x)$$

contains the δ neighborhood of Λ , and hence (3) holds.

iii. For sufficiently small $\epsilon > 0$, and any $x \in \Lambda$,

$$f^n W_\epsilon^s(x) \subset W_{\lambda^n \epsilon}^s(f^n(x)),$$

where $0 < \lambda < 1$. Thus

$$f^n \left(\bigcup_{x \in \Lambda} W_\epsilon^s(x) \right)$$

is contained in the $\lambda^n \epsilon$ neighborhood of Λ and hence, for any neighborhood $\mathcal{V}$ of Λ , is contained in $\mathcal{V}$ for sufficiently large n . Thus, if

$$\bigcup_{x \in \Lambda} W_\epsilon^s(x)$$

is a neighborhood of Λ , Λ is an attracting set.

REMARK: It is not difficult to see (4) of Proposition 1.1 is actually equivalent to (1)–(3), i.e., is another characterization of hyperbolic attractors among hyperbolic sets. To prove this seems to require another set of estimates like those used in establishing the existence and properties of local stable and unstable manifolds for hyperbolic sets, so we will not give the argument here.

§2. Examples. Perhaps the simplest non-trivial example of a hyperbolic attractor is the attractor called the *solenoid*. The ambient space for this example in three dimensional—it is easy to make analogous examples in higher dimensions, but there do not seem to be analogues in two dimensions. As we will see later in this section, there do exist hyperbolic attractors in two dimensional spaces, but the examples known seem more complicated and less transparent than the solenoid example.

Intuitively, the solenoid example is constructed as follows: Take a solid torus T . First squeeze it transversally and stretch it longitudinally to roughly twice its original length. Take the resulting long thin solid torus, give one end a half-turn, and then fold it over

once, and put it back inside the original torus so that its image wraps twice around the hole. See fig. ?? Let f denote the mapping which does this.

Thus, f maps T into itself; fT is a thin solid torus inside T wrapping around twice. Since $fT \subset T$, the $f^n T$ are decreasing; they are successively thinner tori wrapping 2^n times around. The attractor itself, which we will denote by Λ , is the intersection of these sets over n ; it is thus made up of lines of zero thickness ("wires") wrapping infinitely often around inside T in a manner reminiscent of the windings on a toroidal electromagnet; hence the name solenoid. (Contrary to intuition, and unlike an electromagnet, Λ is not made up of a single line but of uncountably many of them, as we shall shortly see.)

Of course, in order to make this construction yield a hyperbolic attractor, it is necessary to be more specific about the details than we have been. We will deal with this by writing an explicit formula for a mapping f which has the indicated intuitive form and which does produce a hyperbolic attractor. One of the nice features of the solenoid example is that it can be realized by a mapping with an extremely simple formula. We can write the solid torus as the product of a circle with a disk. We identify the circle with the set of complex numbers of modulus one and the disk with the set of complex numbers of modulus not greater than one.

§1. **Introduction.** A vigorous mathematical development has grown out of computer experiments on the iteration of simple one-dimensional mappings. We begin by describing some of the phenomena observed in the most extensively studied family of examples, the *quadratic family*. There are several essentially equivalent ways of writing this family; we will, for the moment, write it in the form

$$f_r(x) = 4rx(1 - x).$$

Here, r is a parameter with $0 \leq r \leq 1$; for these r 's, f_r maps the unit interval $[0, 1]$ into itself. For r up to $3/4$, it is not hard to determine theoretically the dynamics generated by f_r . To begin with, for $0 \leq r \leq 1/4$, $f_r(x) < x$ for $x \in (0, 1]$ and hence $f^n(x) \rightarrow 0$ for all $x \in [0, 1]$. For $r > 1/4$, f_r has a fixed point x_0 in $(0, 1]$. This fixed point is easy to compute; it satisfies $4r(1 - x_0) = 1$, so

$$x_0 = 1 - \frac{1}{4r}.$$

The derivative of f_r at x_0 is easily computed to be $2 - 4r$, so x_0 is an attracting fixed point for $1/4 < r < 3/4$, and a repelling one for $r > 3/4$. It can in fact be shown that, for $1/4 < r \leq 3/4$, $f^n(x)$ converges to x_0 for every $x \in (0, 1]$. Thus, for r up to $3/4$, the experiment consisting of picking an x at random in $[0, 1]$ and applying f_r to it repeatedly gives an uninteresting result—the computed orbit always converges to a fixed point, 0 for $0 \leq r \leq 1/4$ and x_0 for $1/4 < r \leq 3/4$.

As r passes through $3/4$, the derivative of f_r at x_0 passes through -1 , and it can be verified by a simple calculation that the condition for a normal period doubling bifurcation is fulfilled; hence, for r slightly larger than $3/4$, f_r has an attracting cycle of period 2. This is about the limit of what can be done by simple explicit computation, but it is easy to explore on the computer what happens as r is increased further. What happens next is that the attracting cycle of period 2 itself undergoes a period doubling bifurcation, becoming repelling but producing at the same time an attracting cycle of period 4. This cycle in turn bifurcates to produce a cycle of period 8, which bifurcates to produce a cycle of period 16, and so on through all the powers of 2. This infinite sequence of bifurcations accumulates at $r = r_\infty = .892486\dots$

For most r 's between r_∞ and 1, typical orbits seem to be “chaotic” rather than asymptotically periodic. However, there are exceptions: There is a parameter interval where there is an attracting cycle of period 3; another where there is an attracting cycle of period 4; three more where there is an attracting cycle of period 5, etc. We will refer to these parameter intervals where the mapping has an attracting periodic cycle as *periodic windows*. The order on the parameter axis of the periodic windows above r_∞ with periods up to 6 is shown in Fig. 1. The number of periodic windows with period p grows rapidly with p ; in fact, a theoretical analysis to be given later shows that the number is asymptotically $2^{p-1}/p$ plus a much smaller correction. It appears from this

sort of numerical experiments that periodic windows are dense in $[r_\infty, 1]$, but this is one aspect of the situation which it has not yet been possible to prove. It is known, on the other hand, that the periodic windows do not “fill up” $[r_\infty, 1]$; the complement of the union of all the periodic windows has been shown to have strictly positive Lebesgue measure (Jakobson, Benedicks and Carleson). Although the situation, as we have described it, is complicated, there is one respect in which it is simpler than might have been expected: If a given f_r admits an attracting periodic cycle, it appears that almost all orbits converge to that cycle, and in particular that no f_r seems to have more than one attracting cycle. This observation can be justified theoretically, as we shall see.

It was certainly surprising to the pioneers of this subject that such a simple family of examples could behave in such a complicated way. What was even more surprising was the discovery, made by Metropolis, Stein, and Stein, that many other examples have, qualitatively, exactly the same complicated behavior. Thus, for example, the family of mappings $x \mapsto r \sin(\pi x)$, again on the unit interval and again for r running from 0 to 1, has the same number of periodic windows for each period as the quadratic family, and these windows occur in the same order.

In this and the following chapter, we will develop some simple analytical tools to account for this observed “universality” of the sequence of periodic windows. We have organized these tools into two main groups: In this chapter, we see what can be shown assuming only that we are dealing with mappings with a single turning point (which are sometime further assumed to be differentiable.) In the next chapter, we introduce a very useful restriction on the class of mappings we consider—the negativity of the Schwarzian derivative—which permits us to sharpen up the results considerably or, at least, to arrive at simpler and more natural statements. For example, the analysis accounting for the universality of the sequence of periodic windows is carried out in the general context of differentiable mappings with a single turning point. Such mappings can have more than one attracting cycle, and, as a result, the mathematical formulation of the universality result has to be rather round-about. Negativity of the Schwarzian derivative implies that there is no more than one attracting cycle and thus permits a more comprehensible formulation.

Before starting on this analysis, let us first prove a very simple result showing that *monotone* mappings do not produce interesting dynamics:

PROPOSITION 1.1. (1) *Let f be a continuous non-decreasing mapping of a compact interval I to itself. Then the orbit of every $x \in I$ is asymptotic to a fixed point.*

(2) *Let the hypotheses be as in (1), except that f is assumed to be non-increasing. Then the orbit of every $x \in I$ is asymptotic to either a fixed point or a cycle of period two.*

PROOF: The second statement follows from the first applied to f^2 . To prove the first, let $x \in I$. We consider separately the cases $f(x) \geq x$ and $f(x) \leq x$. If $f(x) \geq x$, then, since f is non-decreasing, $f(f(x)) \geq f(x)$ and so, by induction, the sequence $f^n(x)$ is non-decreasing. Since this sequence stays in the compact interval I , it must converge, and, by the continuity of f , its limit must be a fixed point of f . Similarly, if $f(x) \leq x$,

the sequence $f^n(x)$ is non-increasing; hence, converges, and its limit is again a fixed point.

§2. Notation and normalizations. A continuous function f defined on an interval $[a, b]$ is said to be *piecewise monotone* if there is a finite set of points

$$a = x_0 < x_1 < x_2 \cdots < x_\ell = b$$

such that f is strictly monotone* on each $[x_{i-1}, x_i]$. A *turning point* for a piecewise monotone f means a point $\bar{x} \in (a, b)$ such that f is not monotone on any interval containing $\bar{x}$ in its interior. We can evidently assume that each of the points $x_1, \dots, x_{\ell-1}$ is a turning point (any one which isn't can be omitted); if we make this assumption then each of the maximal monotone pieces of the graph of f obtained by restricting to $[x_i, x_{i+1}]$ is called a *lap* of f , and ℓ is the *number of laps*. We are going to concentrate on piecewise monotone mappings with exactly two laps, i.e., only one turning point. Reorienting, if necessary, the interval on which such a mapping acts, we can assume without loss of generality that the turning point is a *maximum*. In this case, we will also speak of a *unimodal* map.

Thus, let f be a continuous unimodal mapping of the compact interval $[a, b]$ into itself; let x_c denote its turning point. It will often be convenient to leave out some inessential parts of the interval on which f acts. The subinterval $[a, f(x_c)]$ contains the image of f , so every orbit gets and stays in this subinterval after just one iteration. Thus, from the point of view of dynamics, we may as well restrict f to this interval. If $f(x_c) \leq x_c$, the restriction of f to this interval is non-decreasing, so every orbit converges to a fixed point. This is the situation, for example, for the quadratic family $4rx(1-x)$ for $0 < r \leq 1/2$. Since this situation is not very interesting, we may as well assume that $f(x_c) > x_c$ and then also replace the upper bound of the interval by $f(x_c)$, i.e., assume $b = f(x_c)$.

There is a possible complication at the left-hand end of the interval $[a, b]$. A nice situation, which happens reasonably often, is that

$$(2.1) \quad f(x) > x \quad \text{for all } x \in (a, x_c].$$

(We have omitted the left end-point since it is often convenient, as in the quadratic family, to organize things so that a is a fixed point.) If this is the case, then the image $[f(b), b]$ of $[x_c, b]$ under f is mapped into itself by f , and any orbit starting in $(a, f(b))$ gets into this interval eventually and then stays there forever. Hence, in this case, we may as well restrict f to $[f(b), b]$. In particular, if $f(b) \geq x_c$, the restriction of f to $[f(b), b]$ is decreasing, and hence every orbit starting in $(a, b]$ is asymptotic either to a fixed point or to a cycle of period 2. (This last observation applies, for example, for the quadratic family $4rx(1-x)$ for r in the range $1/2 < r \leq (1 + \sqrt{5})/4$.) If (2.1) does

*Many of the considerations of this chapter extend easily to situations in which monotonicity is not required to be strict. We make the assumption of strict monotonicity mostly just to simplify the notation.

not hold, then—since we are assuming $f(x_c) > x_c$ — f has at least one fixed point in $[a, x_c]$, and there is a neighborhood of x_c which contains no fixed point. Let x_+ denote the largest fixed point in $[a, x_c]$. The interval $[a, x_+]$ is then mapped onto itself, and f is increasing on this interval, so this interval provides a trivial part of the action of f . If $f(b) \geq x_+$, then the analysis given earlier in this paragraph applies to the restriction of f to $[x_+, b]$, and the action of f can be split into two essentially independent parts, one—trivial—on $[a, x_+]$, and the other on $[f(b), b]$. The case where $f(b) < x_+$ is more complicated and a little pathological: At least some orbits starting between x_c and b eventually get mapped into $[a, x_+]$ and then stay there forever. We will usually want to rule out this case. We mention at this point that, if f has negative Schwarzian and if either $f(a) > a$ or if a is a repelling fixed point for f then f cannot have a fixed point in $[a, x_c]$.

To sum up: We can, by eliminating trivial cases, assume that $f(x_c) > x_c$ and that $b = f(x_c)$. From now on, unimodal mappings will be assumed to satisfy these conditions unless we explicitly say otherwise. We will often want to assume also that f has no fixed point in (a, x_c) , and, in this case, we can assume $b = f(a)$ (but we will not always choose to do this.)

By making a linear change of variables, we can put x_c at 0 and b at 1. This change of variables brings the quadratic mapping $4rx(1-x)$ (for $1/2 < r \leq 1$) into the form

$$x \mapsto 1 - \mu x^2 \quad \text{where } \mu = 4r(r - 1/2).$$

As r runs from $1/2$ to 1, μ runs from 0 to 2. We will from now on usually write the quadratic mapping in this form, and restrict it to the interval $[-1, 1]$.

A certain number of considerations are also simplified if we can arrange that f is *even* as well as putting x_c at 0 and $f(x_c)$ at 1. This turned out to happen automatically in the above quadratic example. It is possible to make a topological change of coordinates converting f on $[a, b]$ to an even mapping on $[-1, 1]$ if and only if $f(a) = f(b)$. If $f(a) < f(b)$, we can restrict f to $[a', b]$, where $a' \in [a, x_c)$ is determined by $f(a') = f(b)$, to get a mapping satisfying these conditions, and if $f(a) > f(b)$, we can make a topological change of coordinates to put x_c at 0, b at 1, and a somewhere in $(-1, 0)$, and which converts f to a function which is even to the extent that this makes sense. Furthermore, if f is smooth and satisfies some non-degeneracy condition—e.g., f is C^2 , $f''(x_c) < 0$, and $f'(x) \neq 0$ for $x \neq x_c$ —then these changes of coordinates can also be taken to be smooth (but perhaps not quite as smooth as the original f .) We will therefore, wherever convenient, assume that f is even.

Notwithstanding the remarks of the preceding few paragraphs, it is often clearer to call the working interval $[a, b]$ and the turning point x_c , rather than $[-1, 1]$ and 0 respectively. We will choose whichever representation seems better adapted to the topic being discussed. For the remainder of this chapter, we use the more general $[a, b]$ notation.

§3. Superstable cycles. In general terms, our strategy for explaining the “universal structure” of periodic windows in one-parameter families of unimodal mappings is as

follows: In the first place, it needs to be emphasized that what has to explained is not simply the occurrence of periodic cycles but of *attracting* periodic cycles. Existence of periodic cycles without regard to attractivity can be proved under much more general circumstances than those we are considering. Second: If x_0 is a periodic point for a differentiable mapping f , of period p , then the corresponding cycle is attracting if $|(f^p)'(x_0)|$ is strictly smaller than one and not if this quantity is strictly larger than one; if it is exactly equal to one, the cycle may or may not be attracting. Third: If we write x_i for $f^i(x_0)$, then the chain rule gives

$$(f^p)'(x_0) = \prod_{i=0}^{p-1} f'(x_i).$$

If, in particular, $f'(x_0) = 0$, then the cycle is necessarily attracting whatever the values of the derivatives at the other points may be (provided only that these derivatives do exist.) Since, under these circumstances, $(f^p)'(x_i) = 0$, the cycle is very strongly attracting, and we therefore call a cycle passing through a critical point of f a *superstable* cycle.

Intuitively, mappings which have superstable cycles should be exceptional, but the arguments about persistence of attracting periodic points given at the beginning of the chapter on bifurcation theory show that, if a mapping admitting a superstable cycle is perturbed a little in the C^1 topology, the perturbed mapping admits a nearby cycle, of the same period, which is attracting (but not necessarily superstable.) If f_μ is a C^1 -continuous one-parameter family of mappings, and if f_{μ_0} admits a superstable cycle, then there is an interval around μ_0 in the parameter space where f_μ admits an attracting cycle of the same period.

These consideration apply to mappings f more general than unimodal, but they do require differentiability. We put them together with an argument which is purely topological but requires unimodality, and which we can illustrate in the quadratic family written in the form

$$f_\mu(x) = 1 - \mu x^2 \quad \text{on } [-1, 1], \text{ for } 0 \leq \mu \leq 2.$$

For $\mu < 1$, the image of f_μ is contained in $(0, 1]$, so $f_\mu^p(0) > 0$ for all p . On the other hand, for $\mu = 2$, we have $f_\mu(0) = 1$, $f_\mu(1) = -1$, and $f_\mu(-1) = -1$. Hence,

$$f_\mu^p(0) = -1 \quad \text{for } \mu = 2 \text{ and } p = 2, 3, \dots$$

As μ increases from 1^- to 2, each of the continuous functions $f_\mu^p(0)$ goes from a strictly positive value to -1 and must therefore pass through 0. But if $f_\mu^p(0) = 0$, the critical point 0 is periodic (with period dividing p), so this argument shows that the family f_μ necessarily has many periodic windows. We can in fact easily improve on it to show that there must be at least one periodic window with each period $p \geq 2$ as follows: Let μ_p be the largest value of μ for which $f_\mu^p(0) = 0$. Then

$$f_{\mu_p}^{p+1}(0) = f_{\mu_p}(0) = 1;$$

this implies $\mu_{p+1} > \mu_p$; and this (for all p) implies that the period of 0 under f_{μ_p} cannot be smaller than p . Evidently, this argument uses only simple qualitative features of the family f_μ , not the exact functional form. In fact, it applies exactly as it stands to the piecewise linear family $x \mapsto 1 - \mu|x|$ to prove the existence of parameter values where 0 is periodic, but that this does not imply the existence of periodic windows because the mappings are not differentiable (and, in fact, these mappings with $\mu \geq 1$ cannot have attracting periodic orbits.)

The next step is to show that there must be several μ 's for which 0 is periodic with period p for f_μ . This is done by showing that there are several possibilities for the ordering (in the interval $[-1, 1]$) of the points on the orbit of 0, when 0 is periodic, and that all the possible orderings do in fact occur. The machinery for doing this is developed in subsequent sections of this chapter.

§4. Itineraries. We now introduce a version of symbolic dynamics into the investigation of unimodal mappings. Let f be a unimodal mapping of $[a, b]$ to itself with turning point at x_c and with $f(x_c) = b$, and let $x \in [a, b]$. For each $n \geq 0$, put $I_n(x, f)$ equal to L , C , or R according as $f^n(x) < x_c$, $= x_c$, or $> x_c$. Note that, if $I_{n_0}(x, f) = C$ for some n_0 , then $I_n(x, f)$ is determined for all $n > n_0$; it is equal to $I_{n-n_0}(x_c, f)$. This fact complicates the formulation of a number of results; to minimize the resulting inconvenience, it seems to be a good idea to truncate the sequence $I_n(x, f)$ after the first C , if there is one. Thus, we will refer to the sequence $I_n(x, f)$, truncated after the first C if there is one, as the *itinerary of x under f* , and denote it by $\mathbf{I}(x, f)$. Thus, an itinerary is either an infinite sequence of L 's and R 's, or a finite sequence of L 's and R 's followed by C . The itinerary of x_c is a particularly degenerate case; it is simply the sequence C . We will use the term *abstract itinerary* to mean either an infinite sequence of L 's and R 's or a finite (possibly empty) sequences of L 's and R 's followed by a C , and denote the set of all abstract itineraries by $\mathcal{A}$. It is a slightly pedantic but useful observation that the definition of $\mathcal{A}$ does not logically need to refer to unimodal mappings and itineraries.

It is immediate from the definition that

$$\mathbf{I}(f(x), f) = \sigma \mathbf{I}(x, f)$$

where σ , as usual, denotes the left shift (but allowed to act on finite as well as infinite sequences.) We encounter here an annoying consequence of our convention about truncating itineraries after the first C ; this identity holds only for $x \neq x_c$, as the right-hand side does not really make sense for $x = x_c$.

Perhaps the most important observation about the notion of itinerary is that, if $\mathbf{I}(x_1, f) \neq \mathbf{I}(x_2, f)$, we can determine whether $x_1 < x_2$ or not directly from the two itineraries, in a way which is independent of f . We can see this as follows: First of all, if $I_0(x_1, f) \neq I_0(x_2, f)$, then $x_1 < x_2$ if and only if $I_0(x_1, f) < I_0(x_2, f)$ with respect to the ordering $L < C < R$. If $I_0(x_1, f) = I_0(x_2, f)$, then (since $\mathbf{I}(x_1, f) \neq \mathbf{I}(x_2, f)$), it is not possible that this common value is C . We consider separately the two remaining possibilities. If $I_0(x_1, f) = I_0(x_2, f) = L$, then $x_1 < x_2$ if and only if $f(x_1) < f(x_2)$, since f is strictly increasing on $[a, x_c]$. Similarly, if $I_0(x_1, f) = I_0(x_2, f) = R$, then

$x_1 < x_2$ if and only if $f(x_1) > f(x_2)$ (note that the sense of the inequality is reversed) because f is strictly decreasing on $(x_c, b]$. In either case, we have reduced the problem of determining whether $x_1 < x_2$ to determining whether $f(x_1) < f(x_2)$, and the itineraries of the $f(x_i)$ are obtained from the itineraries of the x_i simply by deleting the first term. After repeating this reduction at most a finite number of times, we reduce to the case where the two itineraries differ in their first terms.

We can organize the results of this argument by introducing an order of the space of abstract itineraries as follows: Given two distinct sequences $I^{(1)}$ and $I^{(2)}$, we let n_0 denote the smallest integer such that $I_{n_0}^{(1)} \neq I_{n_0}^{(2)}$. We then write $I^{(1)} < I^{(2)}$ provided that either

- (1) the number of R 's in the common initial segment $(I_0^{(i)}, \dots, I_{n_0-1}^{(i)})$ is even and $I_{n_0}^{(1)} < I_{n_0}^{(2)}$,
- (2) the number of R 's in this common initial segment is odd and $I_{n_0}^{(1)} > I_{n_0}^{(2)}$,

where the order on individual symbols is, as above,

$$R < C < L.$$

It is clear that this prescription defines a linear ordering on $\mathcal{A}$; we will refer to this ordering as the *twisted lexicographic order*. The following proposition now follows easily from the argument already given:

PROPOSITION 4.1. *Let f be a unimodal mapping on $[a, b]$, and let x_1, x_2 be points of $[a, b]$ such that $I(x_1, f) \neq I(x_2, f)$. Then*

$$x_1 < x_2 \quad \text{if and only if} \quad I(x_1, f) < I(x_2, f).$$

Evidently, the utility of this proposition depends on the extent—as yet unclear—to which points are distinguished by their itineraries. As a preliminary remark on this question, note that if f admits an attracting cycle which does not pass through x_c , then all points in a small neighborhood of any point on this cycle have the same itinerary. It may also be worth remarking that this proposition remains true if the assumption that f is *strictly* monotone on $[a, x_c]$ and $[x_c, b]$ is weakened to simple monotonicity, but that, if f has an interval of constancy, all points in that interval have the same itinerary.

The “twist” in the above ordering is sometimes inconvenient to deal with, and it is worth noting that it can easily be eliminated (but at the expense of introducing some other inconveniences . . .) Given an itinerary I , we construct a sequence of ϵ of ± 1 's (perhaps terminated by a 0) by putting $\epsilon_0 = \epsilon(I_0)$ and $\epsilon_n = \epsilon_{n-1}\epsilon(I_n)$, where $\epsilon(R) = -1$, $\epsilon(L) = +1$, and $\epsilon(C) = 0$. It is very easy to see that the mapping $I \mapsto \epsilon$ is bijective (from the set of all abstract itineraries to an evident analogous space of possible ϵ 's) and that it converts the twisted lexicographic order on I 's into the reverse of the ordinary lexicographic order on ϵ 's. Furthermore, the sequence ϵ corresponding to $I(x, f)$ has a natural interpretation; ϵ_n is $+1$ or -1 according as f^n is increasing or decreasing at x ; it is zero if and only if $f^n(x) = x_c$ and n is the first integer for which this happens.

One disadvantage of replacing I by ϵ is that, although this correspondence carries periodic sequences to periodic sequences, it need not preserve periods—if there are an

odd number of R 's in a period, the period of ϵ will be twice the period of I . This is, however, the only way in which the period can change. Furthermore, this is not as bad as it might seem, since the correspondence between the period of a cycle and that of its itinerary have a similar defect: To take a concrete example, for μ slightly smaller than 1, the quadratic mapping $1 - \mu x^2$ has a periodic orbit of period 2 contained in $(0, 1]$, and the itinerary of this orbit is $RRR\dots$, which is periodic with period 1. This is in fact essentially all that can happen for a general unimodal mapping: If x is periodic with period p , (and if its cycle doesn't pass through x_c) then the itinerary of x has either period p or period $p/2$, and, in the latter case, there must be an odd number of R 's per period in the itinerary. (Exercise) Thus, the period of the ϵ -sequence corresponding to a periodic orbit is either the same as the period of the orbit or twice as large; the period of the itinerary of a periodic orbit is either the period or half as large.

§5. Kneading sequences. We will call the itinerary of $b = f(x_c)$ the *kneading sequence** of f and denote it by $K(f)$. It is nearly evident that it is a topological invariant of f , i.e., that if f is a unimodal mapping on $[a, b]$, if h is a continuous strictly increasing mapping from $[a, b]$ to some (other) interval $[\bar{a}, \bar{b}]$, and if $\bar{f}$ is given by hfh^{-1} on $[\bar{a}, \bar{b}]$, then the kneading sequence of $\bar{f}$ is the same as that of f . As we will see after quite a non-trivial analysis, the kneading sequence is, under reasonably general circumstances, a complete topological invariant. Note that the kneading sequence is finite if and only if x_c is a periodic point for f ; in this case, we will say that f is *superstable* (whether or not f is differentiable, although the terminology is really only justified in the differentiable case.)

It is natural to ask what sequences can occur as kneading sequences for unimodal mappings. One obvious necessary condition is that the sequence begin with R , because of our standing hypothesis that $b = f(x_c) > x_c$. There is a second, less-trivial necessary condition: For any n , $f^n(b) \leq b$, so by Proposition ??,

$$I(f^n(b), f) \leq I(b, f),$$

with respect to the twisted lexicographic order, and hence

$$\sigma^n K(f) = I(f^n(b), f) \leq K(f) = I(b, f).$$

(If $K(f)$ is finite, this condition only holds for n less than or equal to the length of $K(f)$, i.e., less than the period of x_c .) Accordingly, we say that a sequence K is *shift maximal* if

$$\sigma^n K \leq K \quad \text{for all } n \text{ less than the length of } K.$$

We are going to see that the condition of shift maximality is not only necessary but also sufficient for K to occur as the kneading sequence of a unimodal f . It turns out in fact that any shift-maximal K occurs in any "reasonable" one-parameter family of differentiable unimodal mappings. Before we formulate this statement precisely, we make a few elementary remarks about shift-maximal sequences.

*It would be better justified historically to use the term "kneading sequence" for the ϵ -sequence of b rather than for its itinerary.

The condition of shift-maximality is not quite so innocent as it looks; it can sometimes be surprisingly restrictive. For example: The only shift-maximal sequence of the form $RR\dots$ is R^∞ . To see this, consider any sequence of the form $I = R^n L\dots$, $n \geq 2$. Then $\sigma^{n-1}I = RL\dots$, which is larger than anything of the form $R^2\dots$ and thus in particular than I , so I cannot be shift maximal. An exactly analogous argument shows that $R^n C$ cannot be shift maximal for $n \geq 2$.

Shift maximal sequences can be compared in the twisted lexicographic order. The smallest shift-maximal sequence (beginning with R) is R^∞ (since on the one hand R^∞ is the only shift maximal sequence of the form $RR\dots$ and on the other hand any sequence of the form $RR\dots$ is earlier in the twisted lexicographic ordering than any one of the form $RL\dots$ and than RC .) Furthermore R^∞ is the kneading sequence of any unimodal f such that $f(x_c) > x_c$, e.g., the mapping $1 - \mu x^2$ for $0 < \mu < 1$. At the other extreme, the largest sequence beginning with R is RL^∞ (since the smallest sequence is L^∞ , and prepending an R reverses the ordering), and it is trivial to verify that RL^∞ is shift maximal; hence RL^∞ is the largest shift-maximal sequence. It is the kneading sequence of any unimodal f with $f(b) = a$ and $f(a) = a$, e.g., for $1 - \mu x^2$ for $\mu = 2$.

We can now formulate

THEOREM 5.1 (INTERMEDIATE VALUE THEOREM). *Let J be a shift maximal sequence beginning with R , and let $\mu \mapsto f_\mu$ be a continuous mapping from the parameter interval $[\mu_0, \mu_1]$ to the space of continuously differentiable unimodal mappings on $[a, b]$ which is continuous in the C^1 topology. Assume that J lies between $K(f_{\mu_0})$ and $K(f_{\mu_1})$ in the twisted lexicographic order. Then there exists $\hat{\mu} \in [\mu_0, \mu_1]$ such that $K(f_{\hat{\mu}}) = J$.*

COROLLARY 5.2. *Let $\mu \mapsto f_\mu$ be a C^1 -continuous one-parameter family of differentiable unimodal mappings defined on $[\mu_0, \mu_1]$, such that $K(f_{\mu_0}) = R^\infty$ and $K(f_{\mu_1}) = RL^\infty$. Then every shift-maximal sequence beginning with R occurs as the kneading sequence of some member of this family.*

We would like to say furthermore that the order in which the shift-maximal sequences occur in families satisfying the hypotheses of Corollary 5.2 is the twisted lexicographic order. This need not be true, however, since it is certainly possible for a family to transverse the possible kneading sequences in a non-monotone way. The above intermediate value theorem does, however, impose strong constraints on the order in which the sequences occur, analogous to the constraints imposed by the ordinary Intermediate Value Theorem on the order in which a continuous function on the line takes on the values between 0 and 1 in the process of running from 0 to 1. If we restrict ourselves to finite shift-maximal sequences J , it makes sense to speak of the *largest* (or *smallest*) value of μ for which $K(f_\mu) = J$. If we call this largest value $\mu(J)$, then the mapping $J \mapsto \mu(J)$ is order preserving.

Combining the above remarks with the argument giving above to show that a parameter value which makes x_c periodic with period p is in the interior of a parameter interval where f_μ has an attracting periodic cycle of period p , and doing just a little more easy technical work, we prove

COROLLARY 5.3. *Let f_μ be as in Corollary 5.2. It is possible to choose, for each finite shift-maximal $\mathbf{J}$, an open subinterval $M_{\mathbf{J}}$ of the parameter interval such that*

- (1) *If $\mathbf{J}^{(1)} < \mathbf{J}^{(2)}$, then every point of $M_{\mathbf{J}^{(1)}}$ lies to the left of every point of $M_{\mathbf{J}^{(2)}}$ (so the intervals are disjoint and occur in the same order as the $\mathbf{J}$'s.)*
- (2) *For $\mu \in M_{\mathbf{J}}$, f_μ has an attracting cycle, with period the length of $\mathbf{J}$, which attracts the orbit of x_c .*

Two remarks: First, as we will see in the next chapter, unimodal mappings with negative Schwarzian derivative cannot have more than one attracting cycle, and, if there is an attracting cycle, it must attract the orbit of x_c . The preceding Corollary can therefore be reformulated in a somewhat more natural way for mappings with negative Schwarzian derivative. Second, to get a really natural formulation, it is tempting to assume that the family in question is monotone with respect to the twisted lexicographic order on the kneading sequence. Checking this for a concretely given family seems, however, to be extremely difficult. The only result I know in this direction is a proof, by Milnor, Thurston, Sullivan, and Douady, that the quadratic family is monotone in this sense. This proof uses very powerful methods and rests on the global behavior of quadratic mappings in the complex plane.

§6. Proof of the Intermediate Value Theorem. There is a natural way to try to prove Theorem 5.1, imitating the proof of the ordinary Intermediate Value Theorem. This turns out to work only partially, but it can be patched in the places where it doesn't work directly. The idea is to try to show that, for $\mathbf{J}$ a shift-maximal sequence, $\{f : K(f) < \mathbf{J}\}$ and $\{f : K(f) > \mathbf{J}\}$ are both open in the C^1 topology. If we could show this, it would imply that $\{K(f_\mu) < \mathbf{J}\}$ and $\{f : K(f_\mu) > \mathbf{J}\}$ are disjoint relatively open subsets of $[\mu_0, \mu_1]$ which are both non-empty since one contains μ_0 and the other μ_1 . (We evidently need only consider the case where $\mathbf{J} \neq K(f_{\mu_0})$ and $\mathbf{J} \neq K(f_{\mu_1})$.) By the connectedness of $[\mu_0, \mu_1]$, their union cannot fill up this interval, so the only possibility would be that there is at least one $\hat{\mu}$ such that $K(f_{\hat{\mu}}) = \mathbf{J}$. When we try to carry out this program, we find that the sets $\{f : K(f) < \mathbf{J}\}$ and $\{f : K(f) > \mathbf{J}\}$ are not open for all shift-maximal $\mathbf{J}$'s, but they do turn out to be open for those $\mathbf{J}$'s which are not of the form $(AR)^\infty$ or $(AL)^\infty$, where A is a finite sequence of L 's and R 's such that AC is shift maximal. We then need to make a special argument to deal with these exceptional $\mathbf{J}$'s.

Note that the $\mathbf{J}$'s which require special treatment are all periodic. In fact, as we shall see in the next section, they are exactly the periodic shift-maximal sequences (beginning with R), but we do not need this fact in carrying out the proof.

The first step, then, is to prove:

LEMMA 6.1. *Let $\mathbf{J}$ be any shift-maximal sequence beginning with R which is not of the form $(AR)^\infty$ or $(AL)^\infty$ with AC shift-maximal. Then $\{f : K(f) < \mathbf{J}\}$ and $\{f : K(f) > \mathbf{J}\}$ are both open subsets of the space of continuously differentiable unimodal mappings on $[a, b]$ equipped with the C^1 topology.*

PROOF: We will give the argument for $\{f : K(f) < J\}$, pointing out the changes necessary to deal with $\{f : K(f) > J\}$ only when they are not evident. Let $\bar{f}$ be such that $K(\bar{f}) < J$, and write $\bar{J}$ for $K(\bar{f})$. We want to find a C^1 neighborhood of $\bar{f}$ on which $K(f) < J$. Let n be the smallest index such that $\bar{J}_n \neq J_n$. If $\bar{J}_n \neq C$, (i.e., it is equal to either R or L) then there is a C^0 neighborhood of $\bar{f}$ on which $K_i(f) = K_i(\bar{f})$ for $i = 0, 1, \dots, n$, and evidently $K(f) < J$ everywhere on this neighborhood. Notice, incidentally, that this argument (a) works in the C^0 rather than C^1 topology, and (b) never uses the hypothesis that J is shift-maximal. This strongly suggests that the real difficulty in the argument must be hidden in the case where $\bar{J}_n = C$. The analysis for this case has two elements, the first analytic and the second combinatorial, which we organize into two lemmas:

LEMMA 6.2. *Let $\bar{f}$ be a continuously differentiable superstable unimodal mapping; write its kneading sequence as AC , where A is a finite (possibly empty) sequence of R 's and L 's. There is a C^1 neighborhood $\mathcal{U}$ of $\bar{f}$ such, for all $f \in \mathcal{U}$, $K(f)$ is one of the three sequences AC , $(AR)^\infty$, $(AL)^\infty$.*

LEMMA 6.3. *Let AC be a finite shift-maximal sequence, J some other shift-maximal sequence (which may be either finite or infinite.)*

- (1) *If $J > AC$, then $J \geq (AR)^\infty$ and $J \geq (AL)^\infty$.*
- (2) *If $J < AC$, then $J \leq (AR)^\infty$ and $J \leq (AL)^\infty$.*

Given these two lemmas, it is easy to finish the proof of Lemma 6.1. Returning to the notation used above, the case remaining to deal with is that where $K(\bar{f})$ is a finite sequence; hence, of the form AC . If $\mathcal{U}$ is the C^1 neighborhood of $\bar{f}$ given by Lemma 6.2, then (1) of Lemma 6.3 implies that $K(f) \leq J$ for all $f \in \mathcal{U}$. But also $K(f)$ cannot equal J for any $f \in \mathcal{U}$; on the one hand, by assumption, $AC < J$, and on the other hand $(AR)^\infty$ and $(AL)^\infty$ cannot equal J because the latter is assumed not to have this form. Thus $K(f) < J$ for all $f \in \mathcal{U}$, as desired. The proof that $\{f : K(f) > J\}$ is open goes in just the same way, except that (2) of Lemma 6.3 takes the place of (1).

PROOF OF LEMMA 6.2: Let p denote the length of AC , so $\bar{f}^p(\bar{x}_c) = \bar{x}_c$ but $\bar{f}^n(\bar{x}_c) \neq \bar{x}_c$ for $n < p$. (Here, $\bar{x}_c$ denotes the turning point of $\bar{f}$.) The idea is going to be the following: Take f very near to $\bar{f}$. Then, simply by C^0 approximation, we can guarantee that $f^i(x_c) - x_c$ has the same sign as $\bar{f}^i(\bar{x}_c) - \bar{x}_c$ for i from 1 to $p-1$, i.e., that $K(f) = A\dots$. If $f^p(x_c) = x_c$ there is nothing to prove, so we assume that this equality does not hold and, for definiteness, that $f^p(x_c) = x_c + \delta$ with $\delta > 0$, i.e., that $K_{p-1}(f) = R$. By making f near $\bar{f}$ we can make δ as small as we like. Let V_f denote the interval $[x_c, x_c + 2\delta]$. Since the derivative of f^p vanishes at x_c , f^p will be strongly contractive on V_f provided δ is small; it is at this point that we need C^1 approximation. If the contractivity constant is strictly less than one-half, it follows from the fact that f^p maps the endpoint x_c of V_f to the midpoint of V_f that it maps V_f into its own interior. This implies that $f^{np}(x_c) \in V_f$, i.e., that $K_{np-1}(f) = R$, for $n = 2, 3, \dots$. We can furthermore arrange, again by making f close to $\bar{f}$ in the C^0 topology, that for all $x \in V_f$ (and not just for $x = x_c$), $f^i(x) - x_c$ has the same sign as $\bar{f}^i(\bar{x}_c) - \bar{x}_c$ for i

from 1 to $p-1$, and from this it follows easily by induction that $K(f) = (AR)^\infty$. The case where $K_{p-1}(f) = L$ is treated in exactly the same way, except that we take V_f to denote $[x_c + 2\delta, x_c]$, where δ again denotes $f^p(x_c) - x_c$, but is negative rather than positive.

Formally, we proceed as follows: Take $\epsilon > 0$ such that

- (1) $|(\bar{f}^p)'(x)| < 1/4$ for all x with $|x - \bar{x}_c| \leq \epsilon$.
- (2) $|\bar{f}^i(x) - \bar{x}_c| \geq \epsilon$ for $|x - \bar{x}_c| \leq \epsilon$ and $i = 1, 2, \dots, p-1$.

Then there exists a C^1 neighborhood of $\bar{f}$ such that, for $f \in \mathcal{U}$,

- (3) $|x_c - \bar{x}_c| \leq \frac{\epsilon}{3}$.
- (4) $|f^p(x_c) - x_c| \leq \frac{\epsilon}{3}$.
- (5) $|f^i(x) - \bar{f}^i(x)| \leq \frac{\epsilon}{3}$ for $|x - \bar{x}_c| \leq \epsilon$ and $i = 1, 2, \dots, p$.
- (6) $|(f^p)'(x)| \leq \frac{1}{3}$ for $|x - \bar{x}_c| \leq \epsilon$.

We are going to show that this $\mathcal{U}$ works.

Now let δ denote $f^p(x_c) - x_c$. As already noted, there is nothing to be proved if $\delta = 0$, so we assume $\delta \neq 0$. We let V_f denote $[x_c, x_c + 2\delta]$ if $\delta > 0$ and $[x_c + 2\delta, x_c]$ if $\delta < 0$. By (3) and (4), V_f is contained in $\{x : |x - \bar{x}_c| \leq \epsilon\}$, and hence, by (6) $|(f^p)'(x)| \leq \frac{1}{3}$ on V_f . Since f^p maps an endpoint of V_f to the midpoint, it maps V_f into its interior, and, as noted, this implies that $K_{np-1}(f) = K_{p-1}(f)$ for $n = 2, 3, \dots$. Finally, it follows from (2) and (5) that the sign of $f^i(x) - x_c$ is the same as the sign of $\bar{f}^i(\bar{x}_c) - \bar{x}_c$ for $x \in V_f$ and $i = 1, \dots, p-1$, and this implies that $K(f) = (AR)^\infty$ if $\delta > 0$ and $K(f) = (AL)^\infty$ if $\delta < 0$.

PROOF OF LEMMA 6.3: The proofs of parts (1) and (2) are surprisingly different. We begin with (1), i.e., we assume $AC < J$. If A contains an odd number of R 's, then $(AR)^\infty < AC < J$, so certainly $(AR)^\infty < J$. On the other hand, if A contains an even number of R 's, then $(AL)^\infty < AC < J$. Thus, in either case, one of $(AL)^\infty$ or $(AR)^\infty$ is trivial. Considering the two cases, we see that what remains to be done is to show that $(AX)^\infty \leq J$, where X is either R or L , chosen so that AX contains an odd number of R 's.

We prove this by assuming the contrary, i.e., that $J < (AX)^\infty$. Then

$$AC < J < (AX)^\infty,$$

which implies that J has A as an initial segment. Furthermore, the following element of J must equal X ; otherwise it would not be true that $AC < J$. We can therefore assume that J has the form $AX\bar{J}$. Since AX contains an odd number of R 's,

$$J = AX\bar{J} < (AX)^\infty = AX(AX)^\infty \quad \text{implies} \quad \bar{J} > (AX)^\infty,$$

by "canceling" the initial AX . On the other hand, by the shift-maximality of $AX\bar{J}$,

$$\bar{J} \leq J.$$

Putting these inequalities together gives

$$(AX)^\infty < \bar{J} \leq J < (AX)^\infty,$$

which is contradictory because some of the intermediate inequalities are strict.

For (2), we have only to show that $J \leq B^\infty$, where $B = AX$, with X chosen this time so that the number of R 's in B is *even*. Again the result is immediate unless B is an initial segment of J , so we can assume that J has the form $B\bar{J}$. Furthermore, if $\bar{J} \leq B^\infty$, then—since B contains an even number of R 's— $J = B\bar{J} \leq BB^\infty = B^\infty$ as desired. We have thus only to show that $\bar{J} > B^\infty$ is impossible. This latter inequality and shift-maximality of $B\bar{J}$ imply

$$B^\infty < \bar{J} \leq B\bar{J}.$$

This chain of inequalities implies that B is an initial segment of $\bar{J}$; if we reinsert this fact into the chain of inequalities we see that $(B)^2$ is an initial segment of $\bar{J}$; and continuing in this way we see that $\bar{J} = (B)^\infty$. But this contradicts the assumption that $\bar{J} > B^\infty$, so this latter assumption is shown to be impossible, as desired.

It is now not difficult to complete the proof of the Intermediate Value Theorem. The argument so far has proved the theorem for all shift-maximal J *except* those of the form $(AL)^\infty$ and $(AR)^\infty$, where A is such that AC is also shift-maximal. We thus choose such an A , and for definiteness we consider the proof that there is a $\hat{\mu}$ such that $K(f_{\hat{\mu}}) = (AR)^\infty$. We also assume, again for definiteness, that $AC < (AR)^\infty < K(f_{\mu_1})$. By the part of the theorem we have already proved, there are parameter values where $K(f_\mu) = AC$, and by continuity there is a largest such parameter value, say $\bar{\mu}$. Then, again by the part of the theorem already proved, $K(f_\mu) > AC$ for all $\mu \in (\bar{\mu}, \mu_1]$. (If this were not true, then there would have to be values of μ in this interval where $K(f_\mu) = AC$, and we have assumed that $\bar{\mu}$ is the largest value for which this happens.) But by Lemma 6.2, the only values of $K(f_\mu)$ possible for μ near to $\bar{\mu}$ are AC , $(AL)^\infty$, and $(AR)^\infty$. For μ to the right of $\bar{\mu}$, the first two are ruled out because $K(f_\mu) > AC$, so any $\hat{\mu}$ sufficiently near to, but larger than, $\bar{\mu}$ must have $K(f_{\hat{\mu}}) = (AR)^\infty$, as desired.

§7. Finite and periodic kneading sequences. As we have seen, each finite shift-maximal sequences gives rise to a periodic window in one-parameter families of differentiable unimodal mappings. In this section, we study these sequences from a combinatorial point of view.

PROPOSITION 7.1. *Let AC be a finite shift-maximal sequence. Then $(AL)^\infty$ and $(AR)^\infty$ are shift maximal.*

PROOF: A quick way to prove this proposition is to note that it follows from Lemma 6.2 and the existence of a family—e.g., $1 - \mu x^2$ —in which the kneading sequence runs from R^∞ to RL^∞ . The statement is a combinatorial one, however, so it seems inappropriate to make a detour through unimodal mappings to prove it. We will therefore give a direct

combinatorial proof. For definiteness, we consider $(AR)^\infty$; the case with L instead of R is exactly analogous. We let p denote the length of AC ; what we need to show is that

$$(7.1) \quad \sigma^n(AR)^\infty \leq (AR)^\infty \quad \text{for } n = 1, \dots, p-1.$$

The major steps in the argument are as follows:

- (1) We show that (7.1) holds *unless* the two sequences $\sigma^n(AR)^\infty$ and $(AR)^\infty$ have a common initial segment of length $p-n$; furthermore, if they do have such a common initial segment, it must contain an *odd* number of R 's.
- (2) In the case not settled by (1), the desired inequality (7.1) is equivalent to the inequality obtained by dropping the first $p-n$ terms on both sides and reversing the sense of the inequality, i.e., to

$$\sigma^{p-n}(AR)^\infty \leq (AR)^\infty.$$

- (3) By (1) (with $p-n$ in place of n), this latter inequality holds unless the two sides have a common initial segment of length n . If this latter condition holds, then the two original sequences $\sigma^n(AR)^\infty$ and $(AR)^\infty$ have a common initial segment of length p and so must be equal (since both are periodic with period p .)

Steps (2) and (3) are immediate. To prove (1), we write out the two sequences in question as

$$\begin{aligned} (AR)^\infty &= A_0 \cdots A_{p-n-2} A_{p-n-1} A_{p-n} \cdots \\ \sigma^n(AR)^\infty &= A_n \cdots A_{p-2} \quad R \quad A_0 \cdots \end{aligned}$$

We will want to look also at

$$\begin{aligned} AC &= A_0 \cdots A_{p-n-2} A_{p-n-1} \\ \sigma^n AC &= A_n \cdots A_{p-2} \quad C \end{aligned}$$

The case where $A_i \neq A_{n+i}$ for some $i = 0, \dots, p-n-2$ is easy to dispose of; the order relation between $(AR)^\infty$ and $\sigma^n(AR)^\infty$ is then the same as that between AC and $\sigma^n AC$, so the shift-maximality of AC implies $\sigma^n(AR)^\infty < (AR)^\infty$ as desired. We can thus assume $A_i = A_{n+i}$ for $i = 0, \dots, p-n-2$. In this case, the shift-maximality of AC implies that A_{p-n-1} must be such as to make

$$A_0 \cdots A_{p-n-2} A_{p-n-1} \cdots > A_0 \cdots A_{p-n-2} C.$$

Considering separately the situations where the number of R 's in $A_0 \cdots A_{p-n-2}$ is odd and even, we see that A_{p-n-1} must be such as to make the number of R 's in $A_0 \cdots A_{p-n-1}$ *odd*.

We now distinguish two cases: If $A_{p-n-1} = L$, then, since the replacing A_{p-n-1} in $A_0 \cdots A_{p-n-1}$ by C gives a smaller sequence, replacing it by R must give a still smaller one (because $L < C < R$), i.e.,

$$A_0 \cdots A_{p-n-2} A_{p-n-1} \cdots > A_0 \cdots A_{p-n-2} R \cdots,$$

so the desired inequality holds. If, on the other hand, $A_{p-n-1} = R$, then the two sequences $\sigma^n(AR)^\infty$ and $(AR)^\infty$ which we want to compare have $A_0 \cdots A_{p-n-1}$ as a common initial segment, and this segment contains an odd number of R 's. Thus, the proof of (1), and hence of the proposition, is complete.

We now take up the converse question: Does, e.g., shift-maximality of $(AR)^\infty$ imply shift-maximality of AC ? There is a complication here arising from the fact that the period of $(AR)^\infty$ need not be the length p of AR , but may be a non-trivial divisor of p . The situation is as follows:

PROPOSITION 7.2. *Suppose $(AX)^\infty$ is shift maximal, where X is either L or R . Let p denote the length of AX . Then AC is shift-maximal if and only if one of the following condition holds:*

- (1) *The period of $(AX)^\infty$ is p .*
- (2) *The period of $(AX)^\infty$ is $p/2$, and the number of R 's per period is odd.*

PROOF: As in the proof of the preceding proposition, we need to determine the order of AC relative to $\sigma^n AC$ for $1 \leq n \leq p-1$. Also as in the proof of that proposition, if the two sequences differ before the $p-n$ th position, the ordering is the same as that of $(AX)^\infty$ relative to $\sigma^n(AX)^\infty$, and this is the correct order for shift-maximality. Thus, we have only to consider the situation where AC and $\sigma^n AC$ have a common initial segment of length at least $p-n-1$. If, furthermore, the sequence $A_0 \cdots A_{p-n-1}$ contains an even number of R 's, then AC is definitely not shift-maximal, since

$$A_0 \cdots A_{p-n-1} \cdots < A_0 \cdots A_{p-n-2}C = \sigma^n AC.$$

This proves the necessity of the conditions of the proposition: If the period of $(AX)^\infty$ is q , we can get a contradiction to shift maximality by taking $p-n = 2q$ (if $q < p/2$) or $= q$ (if $q = p/2$ and the number of R 's per period is even.)

To prove sufficiency of the conditions of the proposition, we continue along the same lines. We have to consider separately the cases where the number of R 's in $A_0 \cdots A_{p-n-1}$ is odd and where it is even. The odd case is easy; it implies

$$AC = A_0 \cdots A_{p-n-2}A_{p-n-1} \cdots C > A_0 \cdots A_{p-n-1}C = \sigma^n AC,$$

as desired.

All that remains, then, is to show that the even case cannot happen. To see this, we first note that, in the even case, the ordering of $\sigma^n(AX)^\infty$ relative to $(AX)^\infty$ is the same as that of the sequences obtained by deleting the first $n-p$ terms from each of the two sequences, i.e., the same as the order of $(AX)^\infty$ relative to $\sigma^{p-n}(AX)^\infty$. (NB: Note the reversal of shifted and unshifted terms.) Thus, under our present hypotheses,

$$\sigma^n(AX)^\infty \leq (AX)^\infty \quad \text{if and only if} \quad (AX)^\infty \leq \sigma^{p-n}(AX)^\infty.$$

Our hypotheses are therefore compatible with the shift-maximality of $(AX)^\infty$ only if

$$\sigma^{p-n}(AX)^\infty = (AX)^\infty,$$

i.e., only if $(AX)^\infty$ is periodic with period $p - n$. The only possibility allowed by the conditions of the proposition is then that $p - n$ is $p/2$. But we are considering the case where the initial segment of A of length $p - n = p/2$ contains an even number of R 's and the conditions of the theorem exclude the case where the period is $p/2$ and there are an even number of R 's per period.

PROPOSITION 7.3. *Let AC be a finite shift-maximal sequence. Let X be L or R according as the number of R 's in A is odd or even (i.e., X is chosen in such a way as to make the number of R 's in AX odd); and let $\check{X}$ be R if $X = L$, L if $X = R$. Then each of the following sequences*

$$\begin{aligned} (A\check{X})^\infty &< AC < (AX)^\infty = (AXAX)^\infty < \\ &< AXAC < (AXA\check{X})^\infty < AXA\check{X}AXAC < \dots \end{aligned}$$

is shift-maximal, and this infinite sequence of shift-maximal sequences is contiguous, i.e., no other shift-maximal sequence lies between any two successive terms.

Note that, if p denotes the length of AC , i.e., the period of the corresponding periodic window, then the lengths of the successive finite sequences are $p, 2p, 2^2p, \dots, 2^n p, \dots$. Thus, immediately following a periodic window of period p , there must occur windows of period $2p, 4p, \dots$. Normally, we would expect these to be produced by successive period-doubling bifurcations, and this can be proved to be the case in families of differentiable unimodal mappings with negative Schwarzian derivative.

PROOF OF PROPOSITION 7.3: From Proposition 7.1 and the assumed shift-maximality of AC , it follows that $(A\check{X})^\infty$ and $(AX)^\infty$ are shift-maximal; from Proposition 7.2, it follows that the period of $(AX)^\infty$ must be p ; the possibility that the period is $p/2$ is ruled out by the fact that the number of R 's in AX is odd. Hence, the sequence $(AXAX)^\infty$ satisfies the conditions of Proposition 7.2, so the finite sequence $AXAC$ is shift-maximal. The argument can now be repeated with AXA in place of A , etc. That the sequence built in this way is contiguous follows from Lemma 6.3.

The preceding results can be used to derive a simple recursive procedure for determining the number of finite shift-maximal sequences of length p . For each finite shift-maximal sequence AC , take X as in the preceding proposition so that the number of R 's in AX is even; then $(AX)^\infty$ is shift-maximal by Proposition 7.1 and has period equal to the length of AC by Proposition 7.2. Conversely, if $(AX)^\infty$ is a shift-maximal sequence whose period is the length of AX and which contains an odd number of R 's per period, then AC is shift-maximal. Hence, the number of shift-maximal sequences of length p is equal to the number of shift-maximal sequences of period p with an odd number of R 's per period. Next we observe that, if J is a sequence of L 's and R 's with period p , one of the p distinct translates of J is largest, i.e., exactly one point on each periodic cycle is shift-maximal. Hence, if n_p denotes the number of shift-maximal sequences of period p , the number of periodic sequences of period p with an odd number of R 's per period must be pn_p .

On the other hand, the number of sequences J of L 's and R 's with $\sigma^p J = J$ and with an odd number of R 's in a segment of length p is 2^{p-1} : The first $p - 1$ terms can

be chosen arbitrarily from $\{L, R\}$; the p th is then determined by the parity condition; and all the remainder by periodicity. Each such sequence must have period dividing p , and the parity condition implies both that the quotient of p by the period cannot be even and that the number of R 's per period is odd. Conversely, any sequence which is periodic with period q , where q is the quotient of p by an odd integer, and which has an odd number of R 's per period, is in this set. Hence we have

PROPOSITION 7.4.

$$2^{p-1} = \sum_{q:p/q \text{ an odd integer}} qn_q = pn_p + \sum_{q < p:p/q \text{ an odd integer}} qn_q.$$

From this it is easy to calculate the first few n_p 's, starting from $n_1 = 1$:

$2^{2-1} = 2 = 2n_2;$	$n_2 = 1$
$2^{3-1} = 4 = 3n_3 + n_1;$	$n_3 = 1$
$2^{4-1} = 8 = 4n_4;$	$n_4 = 2$
$2^{5-1} = 16 = 3n_5 + n_1;$	$n_5 = 3$
$2^{6-1} = 32 = 6n_6 + 2n_2;$	$n_6 = 5$
$2^{7-1} = 64 = 7n_7 + n_1;$	$n_7 = 9$
$2^{8-1} = 128 = 8n_8;$	$n_8 = 16$

and so on.

If p is a prime other than 2, the sum formula gives $pn_p = 2^{p-1} - 1$, and if p is a power of 2 it gives $pn_p = 2^{p-1}$. We can derive a simple asymptotic formula valid in other cases as follows: We start from the observation that the above proposition trivially implies $pn_p < 2^{p-1}$. If we let r denote the smallest odd divisor of p then

$$\sum_{q < p:p/q \text{ an odd integer}} qn_q \leq \sum_{j=1}^{p/r} 2^{j-1} \leq 2^{p/r} - 1,$$

and hence

$$2^{p-1} \geq pn_p \geq 2^{p-1} - 2^{p/r} + 1 \quad \text{with } r \text{ the smallest odd divisor of } p.$$

Note that the correction term $2^{p/r} - 1$ is at most $O(2^{p/3})$ while the main term is 2^{p-1} and hence is much larger for large p . In particular, $2^{p-1}/p$ is an excellent approximation to n_p for large p .

§1. Definition and fundamental properties. Let f be a differentiable function of a (real or complex) variable, defined and three times differentiable on a neighborhood of x_0 , and with $f'(x_0) \neq 0$. The *Schwarzian derivative* $Sf(x_0)$ of f at x_0 is then defined by

$$Sf(x_0) = \left(\frac{f'''(x_0)}{f'(x_0)} \right) - \frac{3}{2} \left(\frac{f''(x_0)}{f'(x_0)} \right)^2.$$

This particular combination of derivatives was introduced by H. A. Schwarz into the theory of conformal mappings, and it was discovered by D. Singer to be very useful in the theory of piecewise monotone interval mappings also. It seems to me not easy to see conceptually why this expression is a useful one to introduce; for this it is necessary to do a little computation. The following two facts about the Schwarzian derivative are proved by straightforward—and not very illuminating—computation, but are of fundamental importance:

LEMMA 1.1.

$$(1) \quad S(f \circ g)(x) = Sf(g(x)) [g'(x)]^2 + Sg(x).$$

$$(2) \quad \frac{d^2 |f'(x)|^{-1/2}}{dx^2} = -\frac{1}{2} |f'(x)|^{-1/2} Sf(x).$$

We omit the proof; as noted, it is a completely straightforward computation. We have also left out of the statement any assertions about domains; the obvious conditions should be imposed. We are always assuming here that f has at least three derivatives near x and that $f'(x) \neq 0$. The first identity of Lemma 1.1 is valid as formulated both for real derivatives and complex derivatives. In the second, we have written a version valid for real derivatives of real-valued functions; it becomes valid for complex derivatives of complex-valued functions if the absolute value signs are suppressed.

COROLLARY 1.2. *If $Sf(x) < 0$ and $Sg < 0$ at all relevant points, then $S(f \circ g) < 0$. In particular, if f is a mapping of class C^3 of some interval into itself, and if $Sf(x) < 0$ wherever $f'(x) \neq 0$, then $Sf^n(x) < 0$ wherever $(f^n)'(x) \neq 0$, for all positive integers n . The same assertions hold with $<$ replaced by $\leq$.*

PROOF: This is immediate by (1) of Lemma 1.1

COROLLARY 1.3. *Let f be a real-valued function, defined and of class C^3 on some open interval (a, b) . Assume that the derivative of f does not vanish anywhere in (a, b) and that $Sf \leq 0$ on (a, b) . Then $|f'(x)|^{-1/2}$ is convex on (a, b) , and, unless $f'(x)$ is constant on (a, b) , $|f'(x)|$ has no local minimum in (a, b)*

PROOF: Convexity follows at once from (2) of Lemma 1.1 and negativity of the Schwarzian derivative. A local minimum for $|f'(x)|$ is a local maximum for $|f'(x)|^{-1/2}$, and non-constant convex functions cannot have local maxima.

REMARK: The preceding two corollaries, taken together, give some feeling for why negativity of the Schwarzian derivative is a useful condition: It is preserved by composition and implies a non-trivial fact, viz. that $|f'(x)|$ has no no-zero local minima.

COROLLARY 1.4. Assume f is C^3 on (a, b) , and that $f'(x)$ does not vanish anywhere in this interval. Then $Sf(x) \equiv 0$ on (a, b) if and only if f is equal on (a, b) to a fractional linear mapping, i.e., one of the form $\frac{ax + b}{cx + d}$.

PROOF: This result is classical, and is part of the reason for the usefulness of the Schwarzian derivative in the theory of conformal mappings: Fractional linear transformations are exactly the solutions of the differential equation $Sf(x) = 0$. It follows easily from (2) of Lemma 1.1: $Sf(x) \equiv 0$ if and only if $|f'(x)|^{-1/2}$ is constant or linear, i.e., if and only if $f'(x)$ is either constant or the reciprocal of the square of a linear function, and this latter condition is easily seen to be equivalent to f 's being fractional linear.

§2. Weakening the differentiability condition. One mildly unsatisfactory feature of the condition of negative Schwarzian derivative is that it requires the existence of three derivatives. We will see that, for most purposes, it can be replaced by the condition of convexity of $|f'(x)|^{-1/2}$, which is equivalent for C^3 functions but whose formulation requires only the existence of one derivative. For this purpose, it is useful to know that the C^1 condition is closed under composition:

PROPOSITION 2.1. Let f_1 and f_2 be continuously differentiable mappings defined on intervals (a_1, b_1) and (a_2, b_2) respectively with f_1 mapping (a_1, b_1) into (a_2, b_2) . Assume that, on these intervals, the $f_i'(x)$ vanish nowhere and that the $|f_i'(x)|^{-1/2}$ are convex. Then $|(f_2 \circ f_1)'(x)|^{-1/2}$ is convex on (a_1, b_1) .

PROOF: This is of course immediate from earlier results if the f_i are of class C^3 : in this case, concavity of $|f_i'|^{-1/2}$ implies negativity of Schwarzian derivative of f_i ; negativity of the Schwarzian derivatives of both f_1 and f_2 implies negativity of the Schwarzian derivative of $f_2 \circ f_1$; and negativity of the Schwarzian derivative of $f_2 \circ f_1$ implies convexity of $|(f_2 \circ f_1)'|^{-1/2}$. We prove the result in the general case by approximating by smooth functions. For simplicity, we assume that the f_i are increasing; extension to the general case is just a matter of inserting appropriate signs.

Let $\rho(x)$ be a positive C^∞ function on $\mathbb{R}$, vanishing outside $(-1, 1)$, with $\int \rho dx = 1$, and write $\rho_\epsilon(x)$ for $\epsilon^{-1}\rho(\epsilon^{-1}x)$. Let $u_i(x)$ denote $(f_i'(x))^{-1/2}$, and write $u_i^{(\epsilon)}$ for $\rho_\epsilon * u_i$. Then $u_i^{(\epsilon)}$ is defined, C^∞ , strictly positive, and convex on $(a_i + \epsilon, b_i - \epsilon)$, and converges uniformly to u_i on any compact subinterval of (a_i, b_i) as ϵ goes to 0^+ . Let $f_i^{(\epsilon)}$ be defined on $(a_i + \epsilon, b_i - \epsilon)$ by requiring that its derivative is equal to $(u_i^{(\epsilon)})^{-2}$ and its value at, say, $(a_i + b_i)/2$ is equal to the value of f_i there. Then the $f_i^{(\epsilon)}$'s have negative Schwarzian derivatives, and $[(f_2^{(\epsilon)} \circ f_1^{(\epsilon)})']^{-1/2}$ is convex and converges pointwise on (a_1, b_1) as ϵ goes to zero to $[(f_2 \circ f_1)']^{-1/2}$. Since a pointwise limit of convex functions is convex, the result follows.

Guckenheimer has observed that negativity of the Schwarzian derivative can in fact be replaced by a condition which requires no derivatives at all. To explain this condition, we recall the notion of the cross-ratio from complex function theory: Given three distinct points x_1, x_2, x_3 in the Riemann sphere, there is a unique fractional linear transformation g carrying x_1 to 0, x_2 to 1, and x_3 to infinity—for finite x_1, x_2 , and x_3 ,

$$g(z) = \frac{x_2 - x_3}{x_2 - x_1} \frac{z - x_1}{z - x_3}.$$

The image under g of a fourth point x_4 is called the *cross-ratio* of the four (distinct) points x_1, x_2, x_3, x_4 ; again assuming finiteness, it is given by

$$\frac{x_2 - x_3}{x_2 - x_1} \frac{x_4 - x_1}{x_4 - x_3}.$$

It is immediate from the definition that two quadruples x_1, x_2, x_3, x_4 and x'_1, x'_2, x'_3, x'_4 have the same cross ratio if and only if there is a fractional linear transformation g such that $g(x_i) = x'_i$ for $i = 1, 2, 3, 4$. For our purposes, it will be convenient to change the sign of the cross-ratio, i.e., to define

$$R(x_1, x_2, x_3, x_4) \equiv \frac{(x_4 - x_1)(x_3 - x_2)}{(x_2 - x_1)(x_4 - x_3)}.$$

Now let f be a real-valued function which is defined, continuous, and monotone on some open interval (a, b) . We say that f satisfies *condition R* on (a, b) if

$$R(f(x_1), f(x_2), f(x_3), f(x_4)) \geq R(x_1, x_2, x_3, x_4) \quad \text{for all } a < x_1 < x_2 < x_3 < x_4 < b.$$

It is immediate from the definition that, if f satisfies condition R on (a, b) and maps (a, b) to (c, d) , and if g satisfies condition R on (c, d) , then $g \circ f$ satisfies condition R on (a, b) . From the invariance of the cross-ratio under fractional linear transformations, it follows that a fractional linear transformation $g(x) = \frac{ax + b}{cx + d}$ satisfies condition R on any interval which does not contain $-d/c$.

PROPOSITION 2.2. *Let f be defined and continuously differentiable on an open interval (a, b) , and suppose that $f'(x) \neq 0$ everywhere on (a, b) . Then $|f'(x)|^{-1/2}$ is convex on (a, b) if and only if f satisfies condition R on (a, b)*

PROOF: We show first that convexity of $|f'(x)|^{-1/2}$ implies condition R, and we do this by assuming the contrary, i.e., that there is a quadruple of points $x_1 < x_2 < x_3 < x_4$ with

$$R(f(x_1), f(x_2), f(x_3), f(x_4)) < R(x_1, x_2, x_3, x_4).$$

It is easy to see that there is a fractional linear transformation g such that $g(f(x_1)) = x_1$, $g(f(x_2)) = x_2$, $g(f(x_4)) = x_4$, and that the pole of g does not lie between $f(x_1)$ and

$f(x_4)$. Replacing f by $g \circ f$ and (a, b) by an interval a little larger than (x_1, x_4) , we can assume that $f(x_i) = x_i$ for $i = 1, 2, 4$. It is easy to see that, for fixed x_1, x_2 , and x_4 ,

$$R(x_1, x_2, x_3, x_4) = \left(\frac{x_4 - x_1}{x_2 - x_1} \right) \left(\frac{x_3 - x_2}{x_4 - x_3} \right) = \left(\frac{x_4 - x_1}{x_2 - x_1} \right) \left(\frac{x_4 - x_2}{x_4 - x_3} - 1 \right)$$

is an increasing function of x_3 for $x_3 < x_4$. Hence, from

$$R(x_1, x_2, f(x_3), x_4) < R(x_1, x_2, x_3, x_4),$$

it follows that $f(x_3) < x_3$. We now use the Mean Value Theorem: Since $f(x_1) = x_1$, $f(x_2) = x_2$, $f(x_3) < x_3$, and $f(x_4) = x_4$, there must exist

- (1) $\xi_1 \in (x_1, x_2)$ with $f'(\xi_1) = 1$,
- (2) $\xi_2 \in (x_2, x_3)$ with $f'(\xi_2) < 1$,
- (3) $\xi_3 \in (x_3, x_4)$ with $f'(\xi_3) > 1$.

These conditions imply that $|f(x)|$ has a local minimum in (x_1, x_4) , and is not constant in this interval; this contradicts the assumed convexity of $|f'(x)|^{-1/2}$.

To prove that condition R implies the convexity of $|f'(x)|^{-1/2}$, it will suffice to consider the case of increasing f , since both conditions are satisfied by $-f$ if they are satisfied by f . We rewrite condition R as

$$\left(\frac{f(x_4) - f(x_1)}{x_4 - x_1} \right) \left(\frac{f(x_3) - f(x_2)}{x_3 - x_2} \right) \geq \left(\frac{f(x_2) - f(x_1)}{x_2 - x_1} \right) \left(\frac{f(x_4) - f(x_3)}{x_4 - x_3} \right),$$

and we let x_2 approach x_1 and x_4 approach x_3 , holding x_1 and x_3 fixed; then write x_2 for x_3 . The result is:

$$\left(\frac{f(x_2) - f(x_1)}{x_2 - x_1} \right)^2 \geq f'(x_2)f'(x_1).$$

Writing u for $(f'(x))^{-1/2}$, and taking square roots, we can rewrite this last inequality as

$$\frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \frac{dx}{u^2} \geq \frac{1}{u(x_1)u(x_2)} \quad \text{whenever } x_1 < x_2.$$

The proof of the proposition is thus reduced to the following lemma:

LEMMA 2.3. *Let u be a strictly positive continuous function defined on the interval (a, b) . Then u is convex on (a, b) if and only if*

$$(2.1) \quad \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \frac{dx}{u^2} \geq \frac{1}{u(x_1)}u(x_2) \quad \text{for all } x_1, x_2 \text{ with } a < x_1 < x_2 < b.$$

PROOF: It is easy to verify that (2.1) holds with equality replacing the inequality if u is a linear function on (a, b) . Hence, (2.1) holds if u is convex.

Conversely, suppose u is not convex on (a, b) . Then there exist y_1, y_2 , and y_3 , with $a < y_1 < y_2 < y_3 < b$, such that $u(y_3) > g(y_3)$, where g denotes the linear function equal to u at y_1 and y_3 . Let x_1 denote the largest $x < y_2$ such that $u(x) \geq g(x)$, and let x_2 denote the smallest such $x > y_2$. Then, on the one hand, $u(x_1) = g(x_1)$ and $u(x_2) = g(x_2)$, so

$$\frac{1}{u(x_1)u(x_2)} = \frac{1}{g(x_1)g(x_2)} = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \frac{dx}{g^2},$$

since (2.1) becomes an identity for linear functions. On the other hand, $u(x) > g(x)$ on (x_1, x_2) , and this, together with the preceding identity, violates (2.1). Hence, non-convexity of u implies that (2.1) does not hold, so the proof of the lemma is complete.

Warning: One might be tempted to try to replace the condition: " f is continuously differentiable and $|f'(x)|^{-1/2}$ is convex on every interval not containing any critical point of f " with the condition " f is continuous and piecewise monotone, and satisfies the cross-ratio condition on any interval not containing a turning point". Indeed, many results which hold for mappings satisfying the first condition have variants which hold under the second condition—typically with "turning point" replacing "critical point" in their formulations. However, the second condition is *not weaker than the first*, i.e., it does not reduce to the first for continuously differentiable mappings. It fails for mappings satisfying the first condition which have critical points which are not turning points, e.g., x^3 on an interval containing 0. The C^1 condition seems to be the more useful one.

§3. D. Singer's Theorem. As a first substantial application of the notion of negative Schwarzian derivative we have:

THEOREM 3.1 (D. SINGER). *Let f be a continuously differentiable mapping of the interval $[a, b]$ to itself; assume that $|f'(x)|^{-1/2}$ is convex on every subinterval of $[a, b]$ not containing any critical point of f . Then any attracting cycle for f must have either a critical point or one of the end-points a, b in its basin of attraction.*

This theorem does not require that f be unimodal. If we add (a slightly strengthened version of) the hypothesis of unimodality, it gives more concrete information. We say that a unimodal mapping f is C^1 -unimodal if it is continuously differentiable and has no critical points other than its turning point; note that the requirement that there are no extra critical points is stronger than the requirement that f is strictly monotone on (a, x_c) and on (x_c, b) . We will further say that a C^1 unimodal mapping is S -unimodal if $|f'(x)|^{-1/2}$ is convex on (a, x_c) and on (x_c, b) .

COROLLARY 3.2. *Let f be a S -unimodal mapping of $[a, b]$ into itself. Assume also that $f(x_c) = b$ and that either $a = f(b)$ or that $f(a) = f(b)$. Then f cannot have more than one attracting cycle.*

PROOF: By Theorem 3.1, any attracting cycle must attract either the turning point x_c or one of the endpoints. Under our assumptions, however, the orbits of the endpoints and the turning point are essentially the same; if one of them converges to some attracting cycle, the other two also converge to that cycle. Thus, there cannot be more

than one attracting cycle. Note that the last condition, concerning the orbit of a , can be replaced by various other conditions, but cannot be eliminated altogether.

PROOF OF THEOREM 3.1: Let x_0 belong to an attracting cycle for f ; denote the period of the cycle by p . We want to argue that we can replace f by f^p , thus reducing to the case of an attracting *fixed point*. To justify this replacement, in addition to observing that, by Proposition 2.2, $|(f^p)'|^{-1/2}$ is convex on each interval where $(f^p)'$ does not vanish, we need to note that the critical points of f^p are exactly the points y such that $f^i(y)$ is a critical point for f for some $i = 0, 1, \dots, p-1$, and hence that, if we show that x_0 attracts, under the action of f^p , the orbit of a critical point of f^p , it will follow that the cycle of x_0 attracts, under the action of f , the orbit of a critical point of f .

Thus, we can assume that x_0 is a fixed point for f . If $f'(x_0) = 0$, we are done, so we assume that this is not the case; then f is monotone on some neighborhood of x_0 . If f is monotone decreasing near x_0 , we can replace it by f^2 and thus reduce to the case where f is monotone increasing near x_0 . Now let J denote the connected component of x_0 in the basin of attraction of x_0 under the action of f . Since x_0 is attracting, this basin of attraction is relatively open in $[a, b]$; hence, the connected component J is a relatively open subinterval of $[a, b]$. If J contains either a or b , we are done, so we assume that this is not the case; then J must be an ordinary open subinterval of (a, b) , i.e., must have the form (α, β) . Note that α and β cannot be in the basin of attraction of x_0 .

Now fJ is an interval, contained in the basin of attraction of x_0 and containing x_0 , so $fJ \subset J$. By continuity, $f\bar{J} \subset \bar{J}$. But f cannot map α into J ; otherwise, $f(\alpha)$ and hence α itself would be in the domain of attraction of x_0 . Thus, $f(\alpha)$ can only equal α or β , and similarly for $f(\beta)$.

We now assume that (α, β) contains no critical point of f and derive a contradiction. First of all, under this assumption, f must be increasing on (α, β) and so, by the preceding paragraph, we must have $f(\alpha) = \alpha$ and $f(\beta) = \beta$. Furthermore, $f'(\alpha) \geq 1$ since otherwise α would be an *attracting* fixed point for f and orbits starting slightly to the right of α —and hence in J —would converge to α , whereas we know that all orbits starting in J converge to x_0 . Similarly, $f'(\beta) \geq 1$. On the other hand, $f'(x_0) \leq 1$; otherwise, x_0 could not be an *attracting* fixed point. But the three inequalities $f'(\alpha) \geq 1$, $f'(x_0) \leq 1$, $f'(\beta) \geq 1$ are only compatible with the fact that $(f'(x))^{-1/2}$ is concave on (α, β) if $f'(x) \equiv 1$ on (α, β) , and this is inconsistent with x_0 's being an attracting fixed point. This is the desired contradiction and shows that (α, β) must contain a critical point.

More can be extracted from the proof of the above theorem. Let f be a continuous mapping defined on an interval $[a, b]$, and let $x_0 \in (a, b)$ be a fixed point for f . We say that x_0 is *attracting from the left* if there exists $\epsilon > 0$ such that, for all $x \in (x_0 - \epsilon, x_0)$, $f^n(x)$ converges to x_0 as $n \rightarrow \infty$ (or equivalently such that $f(x) > x$ on $(x_0 - \epsilon, x_0)$). Similarly, we say that a fixed point in $[a, b]$ is *attracting from the right* if there exists $\epsilon > 0$ such that $f(x) < x$ on $(x_0, x_0 + \epsilon)$, and we say that a fixed point is *semi-attracting* if it is either attracting from the left or attracting from the right. We say that a periodic cycle for f (of period p) is *semi-attracting* if each—or equivalently just one—point on the cycle is semi-attracting for f^p . Note that, although the occurrence

of a cycle which is semi-attracting but not attracting is an exceptional occurrence, it is not *very exceptional*—such cycles can be expected to occur in “unexceptional” one-parameter families of mappings at special parameter values. For example, the quadratic family, at the minimum parameter value for which it admits a cycle of period three, has in fact a semi-attracting cycle of period three. More generally, a fixed or periodic point depending on a parameter and going through a non-degenerate saddle-node bifurcation is semi-attracting but not attracting at the critical parameter value where the bifurcation occurs.

The notion of basin of attraction can be extended to semi-attracting fixed points and cycles. It is desirable to make the definition so that the domain of attraction is open, and this requires a little care. For example, if x_0 is a fixed point which is attracting from the left (but not attracting), we define the *basin of attraction* of x_0 to be the set of x such that $f^n(x) \rightarrow x_0$ and such that $f^n(x) < x_0$ for all sufficiently large n . Similar definitions can be made for fixed points attracting from the right and for semi-attracting cycles. With these definitions, we have

THEOREM 3.3. *Theorem 3.1 and Corollary 3.2 remain true with “attracting” replaced by “semi-attracting”.*

The proof is a straightforward modification of the above arguments.

It follows from Theorems 3.1 and 3.3 that, if f satisfies the hypotheses of Corollary 3.2 (i.e., f is S -unimodal, $b = f(x_c)$, and $f(a)$ is on the forward orbit of $f(b)$) and if it has an attracting or semi-attracting cycle, then b is in the basin of attraction of that cycle. Excluding the superstable case, this implies that the kneading sequence of f , i.e., the itinerary of b , is eventually periodic. We next want to argue that, with one more hypothesis, the kneading sequence is actually periodic and not just eventually periodic. This will follow, essentially, from the fact that there is a critical point or end point in the *connected component* of x_0 in the basin of attraction. As we shall see in the subsequent proposition, the converse is also true (even without the assumption of negative Schwarzian derivative.)

PROPOSITION 3.4. *Let f be a S -unimodal mapping of $[a, b]$ to itself; assume that $f(x_c) = b$; that either $a = f(b)$ or $f(a) = f(b)$; that $f(x) > x$ on (a, x_c) ; and that f has an attracting or semi-attracting cycle. Then either f is superstable or the kneading sequence of f is periodic.*

They hypothesis that $f(x) > x$ on (a, x_c) is necessary; otherwise, it is possible for f to have an attracting fixed point in $[a, x_c)$ which attracts the orbit of b ; then $K(f)$ is eventually equal to L^∞ but is not periodic because its first term is R .

PROOF: We will give the proof under the assumption that there is an attracting cycle; the modifications needed to deal with the semi-attracting case are straightforward. We can assume that x_c is not periodic; otherwise, there is nothing to prove. Let x_0 be a point on the attracting cycle; let p denote the period of the cycle; and, for $i = 0, 1, \dots$, let x_i denote $f(x_0)$ and J_i the connected component of x_i in the basin of attraction of x_i for f^p . Note that the x_i 's and J_i 's are periodic in i with period p ; it is also easy to

see that f maps J_i into J_{i+1} and that the J_i for $i = 0, 1, \dots, p-1$ are mutually disjoint relatively open subintervals of $[a, b]$.

We want to deal first with the possibility that x_c does not belong to any J_i , so assume that this is the case. It then follows that f^p has no critical point in any J_i . Furthermore, because of the assumption that $f(x) > x$ on $(a, x_c]$, some x_i must lie to the right of x_c , so, by relabeling the x_i if necessary, we can assume $x_0 > x_c$. Since we are assuming the x_c is not in any J_i , J_0 must be contained in $(x_c, b]$. But the proof of Theorem 3.1 shows that J_0 must contain a , or b , or a critical point of f^p , and the first and third of these possibilities are ruled out. Hence $b \in J_0$. But since x_c is not in any J_i , all points in J_0 have the same itinerary, so the itinerary of b is the same as that of x_0 and is hence periodic.

It remains to consider the case where x_c is in some J_i and, again by relabelling, we can assume that $x_c \in J_0$. We are going to argue that the itinerary of $b = f(x_c)$ is the same as that of $x_1 = f(x_0)$. A moment's reflection shows that this is the same as showing that there is no $i \geq 1$ such that x_c lies between x_i and $f^i(x_c)$. Since x_i and $f^i(x_c)$ are in J_i , and since $J_i \cap J_0$ is empty unless i is a multiple of p , we need only consider the case where i is a multiple of p . We will assume for definiteness that $x_c < x_0$; it will be easy to see how to modify the argument to treat the opposite case. What we have to argue, then, is that the assumption that $f^{np}(x_c) < x_c$ for some $n \geq 1$ leads to a contradiction. To see this, we first note that for any x near to but slightly smaller than x_0 , $f^{np}(x) > x$: If f^{np} is decreasing near x_0 , this is true because $f^{np}(x)$ is on the opposite side of x_0 from x ; if f^{np} is increasing near x_0 , it is a consequence of the attractivity of x_0 . In view of this observation, the assumption that $f^{np}(x_c) < x_c$ implies that f^{np} has a fixed point in (x_c, x_0) , and this contradicts the fact that every point of this interval is attracted to x_0 under the action of f^p .

Finally, as already indicated, we prove a converse (under weaker assumptions):

PROPOSITION 3.5. *Let f be a continuous unimodal mapping on $[a, b]$ such that $K(f)$ is periodic. Then f has an attracting or semi-attracting cycle.*

PROOF: Let J denote the set of $x \in [a, b]$ such that $I(x, f) = I(b, f) = K(f)$. Evidently $b \in J$. Because of the relation between the twisted lexicographic ordering on sequences and the usual ordering on $[a, b]$, J is an interval (or possibly a single point.) Let p denote the period of $K(f)$. Since $I(f^p(x), f) = \sigma^p I(x, f)$, f^p maps J into itself. Furthermore, b cannot be a periodic point for f since the only preimage of b is x_c , so periodicity of b would imply periodicity of x_c which would imply that $K(f)$ is finite rather than periodic. Thus, $f^p(b) \neq b$, and since $f^p(b) \in J$, this implies that J is a non-trivial interval. Also, f^p has no turning point in J , i.e., is monotone on J , since a turning point for f^p must have forward orbit which hits x_c (in fewer than p steps) and hence has a finite itinerary. Thus, f^{2p} is an increasing mapping of J to itself, which cannot be the identity since b is not periodic. The following very simple lemma then implies that f has an attracting or semi-attracting periodic point in $\bar{J}$:

LEMMA 3.6. *Let f be a continuous increasing mapping of the closed interval $[a, b]$ to itself. Then either f is the identity on $[a, b]$ or f has an attracting or semi-attracting*

fixed point in J .

PROOF: If $f(x) > x$ for some $x \in [a, b]$, then, since $f(b) \leq b$, f has at least one fixed point in $(x, b]$. Let x^- denote the smallest fixed point in $(x, b]$. Then $f(y) > y$ on (x, x^-) , i.e., x^- is attracting from the left. Similarly, if $f(x) < x$ for some $x \in [a, b]$, f has a fixed point which is attracting from the right. If neither $f(x) > x$ nor $f(x) < x$ for any $x \in [a, b]$ then f must be the identity on $[a, b]$.