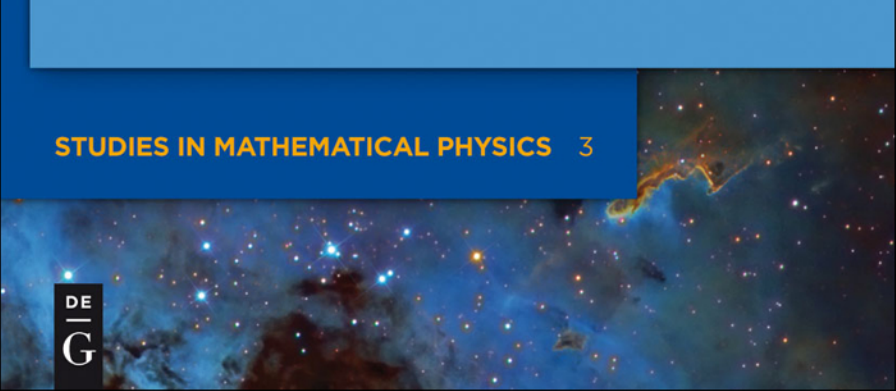
DE GRUYTER

Sergei Yu. Pilyugin

SPACES OF DYNAMICAL SYSTEMS

STUDIES IN MATHEMATICAL PHYSICS 3



De Gruyter Studies in Mathematical Physics 3

Sergei Yu. Pilyugin

Spaces of Dynamical Systems

De Gruyter

Physics and Astronomy Classification 2010: 01.30 mm, 02.30 Hq, 05.45.-a, 05.45 Ac.

ISBN 978-3-11-025595-9
e-ISBN 978-3-11-025841-7

Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

Typesetting: Da-TeX Gerd Blumenstein, Leipzig, www.da-tex.de
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen
∞Printed on acid-free paper

Printed in Germany

www.degruyter.com

To my wife Ira without whose help and patience
this book would not have appeared.

Preface

This book in based on the courses of lectures “Structurally Stable Systems of Differential Equations” and “Spaces of Dynamical Systems” given by the author during the last 30 years to students of Faculty of Mathematics and Mechanics, St. Petersburg State University, specializing in differential equations, geometry, and topology.

As its title indicates, the book is devoted to the theory of dynamical systems (to be exact, to the structure of spaces of dynamical systems with various topologies).

The world mathematical literature contains a lot of books devoted to dynamical systems. First we must mention the classical Birkhoff’s book [1].

The new approaches to theory of dynamical systems related to the problem of structural stability were reflected in the Nitecki monograph [2].

Later, books devoted to dynamical systems were published by Guckenheimer, Moser and Newhouse [3], Palis and di Melo [4], Shub [5], Robinson [6], and other mathematicians.

Finally, let us mention the recent Brin and Stuck’s book [7] and the encyclopedic Katok and Hasselblatt’s monograph [8].

In contrast to the most part of the above-mentioned monographs, the present book is addressed not only to professional mathematicians but also to those who start the study of dynamical systems, especially to students and people working with applications of dynamical systems.

Thus, the main goal of the book is to describe the basic objects of the modern theory of dynamical systems and to formulate its main results.

The first author’s book [9] published in Russian in 1988 served the same purpose. Comparing the book [9] with the present text, an attentive reader will see that the new book does not duplicate the old one; in a sense, they are complementary.

In this book, we mostly work with discrete dynamical systems (and not with flows as in [9]); we describe different approaches to such basic objects as topologies on the considered spaces of dynamical systems and give principally different proofs of some basic results, such as structural stability of Anosov diffeomorphisms. Several important examples of dynamical systems not included into the book [9] are treated in this book; let us mention the Bernoulli shift on the space of two-sided sequences, the hyperbolic toral automorphism, and the Smale horseshoe.

In addition to well-known fields of dynamical systems (such as topological dynamics, theory of structural stability, and chaotic dynamics), the book contains chapters devoted to C^0-generic properties and shadowing of pseudotrajectories (the author’s monographs [10, 11] published in the Springer Lect. Notes in Math. series, vols. 1571 and 1706, were the first monographs in the world mathematical literature devoted to these topics).

The book consists of 12 chapters and two appendices.

In Chapter 1, we define the main objects, dynamical systems with continuous and discrete time. We describe possible types of trajectories and the basic properties of invariant sets. As an example, we consider the Bernoulli shift on the space of two-sided sequences. We study embeddings of discrete dynamical systems into flows and the local Poincaré transformation.

In Chapter 2, we introduce the C^0 topology on the space of homeomorphisms of a compact metric space and the C^1 topology on the space of diffeomorphisms of a smooth closed manifold. For flows generated by autonomous systems of ordinary differential equations, we describe relations between two possible approaches to defining the topology: via estimates of differences between the right-hand sides of the systems and via estimates of closeness of the flows. We consider Baire spaces and generic properties.

In Chapter 3, we study the main equivalence relations on spaces of dynamical systems: topological conjugacy of systems with discrete time and topological equivalence of systems with continuous time. Structural stability and Ω-stability are defined. We introduce the nonwandering set of a dynamical system and prove the Birkhoff theorem: Any trajectory lives only a finite time outside a neighborhood of the nonwandering set.

Chapter 4 is one of the main parts of the book. In this chapter, we define the basic concepts of the theory of structural stability (such as stable and unstable manifolds, fundamental domains, etc.) in the simplest case of a hyperbolic fixed point. We describe properties of hyperbolic linear mappings and prove the Grobman–Hartman theorem on local topological conjugacy of a diffeomorphism near its hyperbolic fixed point and the corresponding linear mapping. A detailed proof of the stable manifold theorem is given; the proof is based on the Perron method. The case of a hyperbolic periodic point is considered as well.

In Chapter 5, we prove analogs of results obtained in Chapter 4 for the case of rest points and closed trajectories of an autonomous system of differential equations. It is shown how to reformulate the definition of hyperbolicity of a closed trajectory in terms of multiplicators of the corresponding periodic solution.

Chapter 6 is devoted to transversality. We define transversality of mappings and submanifolds. The property of transversality of stable and unstable manifolds is introduced. We prove the Palis λ-lemma and describe relations between transversality and hyperbolicity for one-dimensional mappings.

In Chapter 7, the second main part of the book, we study hyperbolic sets. We analyze the definition of a hyperbolic set and give two basic examples of a hyperbolic set: a hyperbolic fixed point and a hyperbolic automorphism of the torus. We formulate the stable manifold theorem, introduce Axiom A, and prove the spectral decomposition theorem. The main results of the theory of structural stability are formulated. Hyperbolic sets of flows are described. We analyze relations between the structural stability theorem and the classical Andronov–Pontryagin theorem on "roughness" of planar autonomous systems.

In Chapter 8, we prove structural stability of an Anosov diffeomorphism.

Chapter 9 is devoted to Smale's horseshoe and chaos. We prove that the horseshoe invariant set is topologically conjugate to the Bernoulli shift. It is shown that the horseshoe invariant set is chaotic. Transverse homoclinic points of planar diffeomorphisms are considered.

We formulate the classical C^1 closing lemma in Chapter 10. The C^0 closing lemma is proven.

In Chapter 11, we study C^0 generic properties of dynamical systems. The Hausdorff metric is defined. The main results of the Takens theory related to the tolerance stability conjecture are proven. The second part of Chapter 11 is devoted to the behavior of attractors under C^0 small perturbations. We prove the Hurley theorem on genericity of stability of attractors in the Hausdorff metric under C^0 small perturbations.

Chapter 12 is devoted to shadowing of pseudotrajectories. We prove that a hyperbolic set has the Lipschitz shadowing property. The Lipschitz inverse shadowing property for a trajectory having (C, λ)-structure is established. The proofs of these results are based on the Tikhonov–Schauder fixed point theorem. Shadowing and inverse shadowing properties of linear mappings are completely characterized.

In Appendix A, we describe a scheme of the proof of Mañé's theorem on the necessity of hyperbolicity for structural stability.

Appendix B is devoted to the history of the theory of differential equations and dynamical systems. Sections: Differential equations and Newton's anagram; Development of the general theory; Linear equations and systems; Stability; Nonlocal qualitative theory. Dynamical systems; Structural stability; Dynamical systems with chaotic behavior. This text is based on lectures on history of mathematics given by the author in the last years to PhD students of the Faculty of Mathematics and Mechanics.

In the text, we do not give references to basic University mathematical courses.

For the author of this book, it was very important to read books and research papers on differential equations and dynamical systems. At the same time, the author is grateful to many mathematicians for the personal contacts.

First, the author wants to thank his teachers in differential equations, dynamical systems, and topology: Yu. N. Bibikov, S. M. Lozinskii, N. N. Petrov, V. A. Pliss, and V. A. Rokhlin.

The author thanks for cooperation his colleagues at the Faculty of Mathematics and Mechanics: L. Ya. Adrianova, A. F. Andreev, V. E. Chernyshev, Yu. V. Churin, Yu. A. Il'in, O. A. Ivanov, S. G. Kryzhevich, G. A. Leonov, N. Yu. Netsvetaev, and A. V. Osipov.

It was very useful to discuss dynamical systems and related fields of mathematics with D. V. Anosov, V. S. Afraimovich, V. I. Arnold, Yu. S. Ilyashenko, V. M. Millionshchikov, Yu. I. Neimark, G. S. Osipenko, N. Kh. Rozov, A. N. Sharkovskii, and L. P. Shilnikov, and also with W.-J. Beyn, B. Fiedler, and P. Kloeden (Germany), G. R. Sell and J. K. Hale (USA), K. Palmer (Taiwan), L. Wen and S. Gan (China), C. Bonatti (France), L. Diaz (Brazil), T. Eirola (Finland), R. Corless (Canada), and K. Sakai (Japan).

The author is happy to mention several of his students who had contributed to the development of the theory of dynamical systems: N. Ampilova, A. Felshtyn, A. Katina, A. Osipov, O. Plamenevskaya, V. Pogonysheva, O. Tarakanov, and S. Tikhomirov.

This book is a translation of the author's book "Prostranstva dinamicheskikh sistem", Izhevsk, R&C Dynamics, 2008. The English text of the book is slightly modified compared to the Russian one. The structure of chapters and appendices is the same, but some details of presentation are improved, for the convenience of the Western reader we give references to books published in the West instead of Russian ones, and so on.

St. Petersburg, November 2011 Sergei Yu. Pilyugin

Nomenclature

$\mathbb{R}^n$	– the Euclidean n-space (we write $\mathbb{R}$ instead of $\mathbb{R}^1$)
$\mathbb{C}^n$	– the complex n-space (we write $\mathbb{C}$ instead of $\mathbb{C}^1$)
$\mathbb{Z}$	– the set of integers
$\mathbb{Z}_+$	– the set of nonnegative integers
$\mathbb{Z}_-$	– the set of nonpositive integers
E	– the identity matrix
$\mathrm{diag}(A_1, \ldots, A_m)$	– a block-diagonal matrix with blocks $A_1, \ldots, A_m$
Id	– the identity mapping
$f \circ g$	– the composition of mappings f and g
$\frac{\partial f}{\partial x}$	– the partial derivative of a mapping f in variable x
$C^k(U, V)$	– the class of continuous mappings from U to V having continuous derivatives up to order k
Df	– the derivative of a mapping f
$T_x M$	– the tangent space of a manifold M at a point x
$\dim M$	– the dimension of a manifold M
$N(a, A)$	– the a-neighborhood of a set A
$\mathrm{Cl}\, A$	– the closure of a set A
$\mathrm{Int}\, A$	– the interior of a set A
∂A	– the boundary of a set A
$\mathrm{card}\, A$	– the cardinality of a finite set A

Contents

Preface vii

Nomenclature xi

1 Dynamical systems 1
1.1 Main definitions 1
1.2 Embedding of a discrete dynamical system into a flow 10
1.3 Local Poincaré diffeomorphism 11
1.4 Time-periodic systems of differential equations 14
1.5 Action of an Abelian group 15

2 Topologies on spaces of dynamical systems 16
2.1 C^0-topology 16
2.2 C^1-topology 17
2.3 Metrics on the space of systems of differential equations 18
2.4 Generic properties 24
2.5 Immersions and embeddings 24

3 Equivalence relations 26
3.1 Topological conjugacy 26
3.2 Topological equivalence of flows 30
3.3 Nonwandering set 30
3.4 Local equivalence 36

4 Hyperbolic fixed point 37
4.1 Hyperbolic linear mapping 37
4.2 The Grobman–Hartman theorem 40
4.3 Neighborhood of a hyperbolic fixed point 48
4.4 The stable manifold theorem 53
4.5 Hyperbolic periodic point 65

5 Hyperbolic rest point and hyperbolic closed trajectory 67
5.1 Hyperbolic rest point 67
5.2 Hyperbolic closed trajectory 72

6 Transversality 78
6.1 Transversality of mappings and submanifolds 78
6.2 Transversality condition 80
6.3 Palis lemma 82
6.4 Transversality and hyperbolicity for one-dimensional mappings 90

7 Hyperbolic sets 92
7.1 Definition of a hyperbolic set 92
7.2 Examples of hyperbolic sets 94
7.3 Basic properties of hyperbolic sets 97
7.4 Stable manifold theorem 101
7.5 Axiom A 103
7.6 Hyperbolic sets of flows 112

8 Anosov diffeomorphisms 119

9 Smale's horseshoe and chaos 126
9.1 Smale's horseshoe 126
9.2 Chaotic sets 131
9.3 Homoclinic points 132

10 Closing Lemma 135

11 C^0-generic properties of dynamical systems 141
11.1 Hausdorff metric 141
11.2 Semicontinuous mappings 142
11.3 Tolerance stability and Takens' theory 143
11.4 Attractors of dynamical systems 147

12 Shadowing of pseudotrajectories in dynamical systems 159
12.1 Definitions and results 159
12.2 Proof of Theorem 12.1 164
12.3 Proof of Theorem 12.2 172
12.4 Proof of Theorem 12.3 175

A Scheme of the proof of the Mañé theorem 181

B Lectures on the history of differential equations and dynamical systems 192
B.1 Differential equations and Newton's anagram 192
B.2 Development of the general theory 194

B.3 Linear equations and systems 198
B.4 Stability 203
B.5 Nonlocal qualitative theory. Dynamical systems 210
B.6 Structural stability 214
B.7 Dynamical systems with chaotic behavior 217

Bibliography 223

Index 227

Chapter 1
Dynamical systems

1.1 Main definitions

Theory of dynamical systems studies two main classes of dynamical systems, systems with discrete time (cascades) and systems with continuous time (flows).

We first define a dynamical system with discrete time.

Let f be a homeomorphism of a topological space M. We define (functional) degrees of f as follows:

Set $f^0 = \text{Id}$, where Id is the identical mapping of M;

if m is natural, we set

$$f^m = \underbrace{f \circ f \circ \cdots \circ f}_{m \text{ times}};$$

finally, if m is a negative integer, we set

$$f^m = \underbrace{f^{-1} \circ f^{-1} \circ \cdots \circ f^{-1}}_{|m| \text{ times}},$$

where f^{-1} is the inverse of f.

Clearly, the mappings f^m are continuous for all $m \in \mathbb{Z}$.

Denote $\phi(m, x) = f^m(x)$. It is easily seen that the mapping $\phi : \mathbb{Z} \times M \to M$ has the following three properties:

(DDS1) $\phi(0, x) = x, x \in M$;

(DDS2) $\phi(l + m, x) = \phi(l, \phi(m, x)), l, m \in \mathbb{Z}, x \in M$;

(DDS3) for any $m \in \mathbb{Z}$, the mapping $\phi(m, \cdot)$ is continuous.

Any mapping $\phi : \mathbb{Z} \times M \to M$ having properties (DDS1)–(DDS3) is called a *(continuous) dynamical system with discrete time* (sometimes such a system is called a *cascade*). The space M is called the *phase space* of the system.

It is easy to understand that if we are given a mapping $\phi : \mathbb{Z} \times M \to M$ having properties (DDS1)–(DDS3), then there exists a homeomorphism f such that $\phi(m, x) = f^m(x)$.

Indeed, set $f(x) = \phi(1, x)$. Let us show that f is a homeomorphism. The mapping f is continuous by property (DDS3); the mapping $g(x) = \phi(-1, x)$ is continuous as well, and properties (DDS2) and (DDS1) imply that

$$f(g(x)) = \phi(1, \phi(-1, x)) = \phi(0, x) = x, \quad x \in M.$$

Thus, $g(f(x)) = x$, and $g = f^{-1}$. The equality $\phi(m, x) = f^m(x)$ is an immediate corollary of property (DDS2).

Thus, it is possible to define a dynamical system with discrete time taking as the initial object either a homeomorphism or a mapping with properties (DDS1)–(DDS3). These two approaches lead to the same result. For this reason, in what follows we do not distinguish a homeomorphism and the dynamical system generated by this homeomorphism.

The basic object which is studied in theory of dynamical systems is defined as follows. Fix a homeomorphism f and a point x of the phase space. The *trajectory* of the point x in the dynamical system generated by f is the set

$$O(x, f) = \{f^m(x) : m \in \mathbb{Z}\}.$$

Sometimes, if the system is fixed, we denote a trajectory by $O(x)$; if the point x is irrelevant, we use the notation $O(f)$.

Clearly, the following statement holds.

Lemma 1.1. *$O(f^m(x), f) = O(x, f)$ for any $m \in \mathbb{Z}$.*

We also apply the following notation:

$$O^+(x, f) = \{f^m(x) : m \in \mathbb{Z}_+\} \quad \text{and} \quad O^-(x, f) = \{f^m(x) : m \in \mathbb{Z}_-\};$$

the sets $O^+(x, f)$ and $O^-(x, f)$ are called the *positive* and *negative semitrajectories* of the point x, respectively.

Similar objects are defined for a subset A of the phase space; the set

$$O(A, f) = \{f^m(A) : m \in \mathbb{Z}\}$$

is called the *trajectory* of a set A in the dynamical system generated by a homeomorphism f, and the sets

$$O^+(A, f) = \{f^m(A) : m \in \mathbb{Z}_+\} \quad \text{and} \quad O^-(A, f) = \{f^m(A) : m \in \mathbb{Z}_-\}$$

are called the *positive* and *negative semitrajectories* of the set A, respectively.

It is easily shown that, for a trajectory of a discrete dynamical system, only one of the following three possibilities can realize (this fact is a corollary of Lemma 1.2 below).

1. $f(x) = x$. In this case, the point x is called a *fixed point*; the trajectory of a fixed point coincides with the fixed point.
2. There exists a number $m \in \mathbb{N}$ such that the points $x, f(x), \dots, f^{m-1}(x)$ are distinct, and $f^m(x) = x$. Such a point x is called *periodic*, the number m is called the *period* of the point x. The trajectory of x consists of m points $x, f(x), \dots, f^{m-1}(x)$.

   Of course, a fixed point is periodic (with period 1); by tradition, fixed and periodic points (with period $m > 1$) are defined separately.

3. The points $f^l(x)$ and $f^m(x)$ are different if $l \neq m$. In this case, the trajectory of x is a countable set.

Denote by $\operatorname{Per}(f)$ the set of periodic points of a homeomorphism f (we include fixed points into this set).

Lemma 1.2. *The set $O(x, f)$ is finite if and only if $x \in \operatorname{Per}(f)$.*

Proof. It was mentioned above that if $x \in \operatorname{Per}(f)$, then the set $O(x, f)$ is finite.

Let us assume that the set $O(x, f)$ is finite. In this case, there exist different integer numbers k and l such that $f^k(x) = f^l(x)$. Let $l > k$; set $n = l - k$. Applying the homeomorphism f^{-k} to the equality $f^k(x) = f^l(x)$, we see that $x = f^n(x)$. If the points $x, f(x), \dots, f^{n-1}(x)$ are distinct, then x is a periodic point of period n.

Otherwise, there exist different integer numbers $k_1, l_1 \in [0, n-1]$ such that $f^{k_1}(x) = f^{l_1}(x)$. The same reasoning as above shows that there exists a number $n_1 \in (0, n)$ such that $x = f^{n_1}(x)$. The least of such numbers n_1 is the period of the point x. □

Now we introduce one more basic notion of theory of dynamical systems. We say that a set $I \subset M$ is *invariant* for the dynamical system generated by a homeomorphism f if $O(x, f) \subset I$ for any point $x \in I$.

Lemma 1.3. *A set I is invariant if and only if $f(I) = I$.*

Proof. Let I be an invariant set. Fix a point $x \in I$. Since $O(x, f) \subset I$, $f(x) \in I$ and $f^{-1}(x) \in I$. Hence, $f(I) \subset I$ and $f^{-1}(I) \subset I$ (thus, $I \subset f(I)$); it follows that $f(I) = I$.

Inverting the reasoning above, we see that if $f(I) = I$, then the set I is invariant. □

It follows from well-known properties of homeomorphisms that if I and J are invariant sets, then the sets $I \cup J$, $I \cap J$, $I \setminus J$, $\operatorname{Cl} I$, and ∂I are invariant as well.

We give an important example of a dynamical system (we refer to this example below several times).

Example 1.1. Let $\mathcal{X}$ be the space whose elements are two-sided, infinite, binary sequences

$$a = \{a_i : a_i \in \{0, 1\},\ i \in \mathbb{Z}\}.$$

We introduce the following metric in the space $\mathcal{X}$: If $a = \{a_i\}$ and $b = \{b_i\}$, we set

$$\operatorname{dist}(a, b) = \sum_{i=-\infty}^{\infty} \frac{|a_i - b_i|}{2^{|i|}}$$

(check that the above formula defines a metric).

Clearly, our definition of the metric dist implies the following statement. For any given $\varepsilon > 0$ there exist numbers $N(\varepsilon)$ and $n(\varepsilon)$ such that if

$$a_i = b_i, \quad |i| \le N(\varepsilon),$$

then $\mathrm{dist}(a, b) < \varepsilon$, and if $\mathrm{dist}(a, b) < \varepsilon$, then

$$a_i = b_i, \quad |i| \le n(\varepsilon).$$

Obviously, $N(\varepsilon), n(\varepsilon) \to \infty$ as $\varepsilon \to 0$.

Let us recall the definition of the Tikhonov product topology (for the particular case of the space $\mathcal{X}$). Fix an element $a = \{a_i\}$ of $\mathcal{X}$ and a finite subset $K \subset \mathbb{Z}$. Consider the cylinder

$$C(a, K) = \{b \in \mathcal{X} : b_i = a_i, i \in K\}.$$

Recall that a *base of neighborhoods* of a point x of a topological space is a family of neighborhoods of x such that any neighborhood of x contains a neighborhood from this family.

The base of neighborhoods of a in the Tikhonov product topology consists of the cylinders $C(a, K)$ corresponding to all finite subsets $K \subset \mathbb{Z}$.

It is easy to show that the metric topology induced by our metric dist coincides with the Tikhonov product topology (which means that the families of open sets in these topologies are the same).

To prove this statement, it is enough to show that if $b \in C(a, K)$ for some $a \in \mathcal{X}$ and a finite set $K \subset \mathbb{Z}$, then there is a small $d > 0$ such that the metric ball

$$N(d, b) = \{b' \in \mathcal{X} : \mathrm{dist}(b', b) < d\}$$

is a subset of $C(a, K)$ and, conversely, if $b \in N(d, a)$ for some $d > 0$, then there is a finite set $K \subset \mathbb{Z}$ such that $C(b, K) \subset N(d, a)$ (we leave details to the reader).

The metric space $(\mathcal{X}, \mathrm{dist})$ is compact. This fact follows from the Tikhonov theorem since the space $\mathcal{X}$ is the countable product of compact spaces $\{0, 1\}$ and, as was said, our metric dist induces on $\mathcal{X}$ the Tikhonov product topology.

Let us give an independent simple proof of the compactness of $(\mathcal{X}, \mathrm{dist})$. It is known that a metric space is compact if and only if any sequence contains a convergent subsequence.

Consider an arbitrary sequence $a^m = \{a_i^m : i \in \mathbb{Z}\}, m \ge 0$.

The elements a_0^m take values 0 and 1; hence, there exists a subsequence

$$m(0) = \{m(0, 1), m(0, 2), \dots\}$$

of $\{0, 1, 2, \dots\}$ such that $0 < m(0, 1) < m(0, 2) < \cdots$ and

$$a_0^{m(0,1)} = a_0^{m(0,2)} = \cdots .$$

Similarly, there exists a subsequence

$$m(1) = \{m(1,1), m(1,2), \dots\}$$

of $m(0)$ such that $0 < m(1,1) < m(1,2) < \cdots$ and

$$a_i^{m(1,1)} = a_i^{m(1,2)} = \cdots, \quad i = -1, 0, 1.$$

Continuing this process, we find subsequences

$$m(k) = \{m(k,1), m(k,2), \dots\}$$

of the sequences $m(k-1)$ such that $m(k,1) < m(k,2) < \cdots$ and

$$a_i^{m(k,1)} = a_i^{m(k,2)} = \cdots, \quad i = -k, \dots, k.$$

Define elements $b^k, k \geq 1$, of the space $\mathcal{X}$ by the equalities $b^k = a^{m(k,1)}$. Clearly, the sequence $\{b^k\}$ is a subsequence of the sequence $\{a^m\}$ with the following property:

$$b_i^k = a_i^{m(k,1)} = a_i^{m(|i|,1)}, \quad i = -k, \dots, k.$$

Consider the element b of the space $\mathcal{X}$ defined by the relations

$$b_i = a_i^{m(|i|,1)}, i \in \mathbb{Z};$$

then

$$b_i = b_i^k, \quad i = -k, \dots, k,$$

and we see that $\text{dist}(b^k, b) \to 0$, $k \to \infty$. Thus, we have shown that the space $(\mathcal{X}, \text{dist})$ is compact.

Consider a mapping σ of the space $\mathcal{X}$ into itself defined as follows: We assign to an element $a = \{a_i\}$ of the space $\mathcal{X}$ the element $\sigma(a) = b = \{b_i\}$ by the following rule:

$$b_i = a_{i+1}, \quad i \in \mathbb{Z}.$$

The mapping σ shifts indices by 1. Clearly, the mapping σ is invertible: $\sigma^{-1}(a) = b$ if and only if

$$b_i = a_{i-1}, \quad i \in \mathbb{Z}.$$

Both mappings σ and σ^{-1} are continuous.

Let us prove that σ is continuous. Take $\varepsilon > 0$ and find the corresponding number $N(\varepsilon)$. Since $n(\delta) \to \infty$ as $\delta \to 0$, there exists a $\delta > 0$ such that $n(\delta) > N(\varepsilon) + 1$. If $\text{dist}(a, b) < \delta$, then

$$b_i = a_i, \quad |i| \leq n(\delta).$$

In this case,

$$b_{i+1} = a_{i+1}, \quad |i| \le N(\varepsilon),$$

and we conclude that $\operatorname{dist}(\sigma(a), \sigma(b)) < \varepsilon$. A similar reasoning is applicable to σ^{-1}.

Thus, σ is a homeomorphism of the space $\mathcal{X}$. The mapping σ (as well as the dynamical system generated by this mapping) is usually called the *shift* on the space of binary sequences. Sometimes it is called the *Bernoulli shift* (though this last term may be applied to more complicated objects).

Let us note several important properties of the shift.

Property 1. *The system σ has infinitely many different periodic points.*

Proof. Clearly, the equality $\sigma^m(a) = b$ is equivalent to the relations

$$b_i = a_{i+m}, \quad i \in \mathbb{Z}.$$

Thus, $\sigma^m(a) = a$ if and only if

$$a_i = a_{i+m}, \quad i \in \mathbb{Z}.$$

This means that the set of periodic points of the shift σ coincides with the set of periodic binary sequences, for which the statement is trivial. □

Property 2. *The set of periodic points of σ is dense in the space $\mathcal{X}$.*

Proof. Fix an arbitrary element a of the space $\mathcal{X}$ and an arbitrary $\varepsilon > 0$. Find for this ε the corresponding number $N(\varepsilon)$ and denote it by N. Let us construct a periodic binary sequence b as follows: represent any index $i \in \mathbb{Z}$ in the form $i = k(2N + 1) + l$, where $k \in \mathbb{Z}$ and $|l| \le N$, and set $b_i = a_l$. Clearly, b is a periodic point of σ and $b_i = a_i$ for $|i| \le N$, i.e., $\operatorname{dist}(a, b) < \varepsilon$. □

Property 3. *There exists an element of the space $\mathcal{X}$ whose positive semitrajectory in the system σ is dense in the space $\mathcal{X}$.*

Proof. Let us construct the desired element a as follows. Take arbitrary $a_i, i < 0$. Set $a_0 = 0$ and $a_1 = 1$. Fix the pairs

$$(a_2, a_3) = (0, 0), \quad (a_4, a_5) = (0, 1), \quad (a_6, a_7) = (1, 0), \quad (a_8, a_9) = (1, 1);$$

thus, a_1 is followed from the right by all possible blocks of zeros and units of length 2. After that, we put to the right of a_9 all possible blocks of zeros and units of length 3, and so on. Clearly, the element a has the following property: For any finite block $(b_1, \dots, b_n)$ of zeros and units there exists an index k such that

$$a_k = b_1, \quad a_{k+1} = b_2, \quad \dots, \quad a_{k+n-1} = b_n.$$

Let us prove that the closure of the semitrajectory $O^+(a,\sigma)$ coincides with the space $\mathcal{X}$. Fix arbitrary $b \in \mathcal{X}$ and $\varepsilon > 0$. Find for the chosen ε the corresponding number $N(\varepsilon)$ and denote it by N.

By the construction of a, there exists an index $k \geq 0$ such that

$$a_k = b_{-N}, \quad \ldots, \quad a_{k+2N} = b_N.$$

Set $a' = \sigma^{-k-N}(a)$; then $a'_i = b_i$ for $|i| \leq N$. This means that $\mathrm{dist}(a', b) < \varepsilon$. Thus, we can find a point of $O^+(a,\sigma)$ in an arbitrary neighborhood of an arbitrary element of $\mathcal{X}$. □

We define the second basic class of dynamical systems axiomatically. Let, as above, M be a topological space.

A mapping $\phi : \mathbb{R} \times M \to M$ is called a *(continuous) dynamical system with continuous time* (a *flow*) if this mapping has the following properties:

(CDS1) $\phi(0,x) = x$, $x \in M$;

(CDS2) $\phi(t+s,x) = \phi(t,\phi(s,x))$, $t,s \in \mathbb{R}$, $x \in M$;

(CDS3) the mapping ϕ is continuous.

In this case, the space M is called the phase space of the system.

Sometimes, property (CDS3) is replaced by a weaker property:

(CDS3′) for any $t \in \mathbb{R}$, the mapping $\phi(t,\cdot)$ is continuous

(such an assumption corresponds to the general notion of action of a group, which we consider in Section 1.5).

In this book, we study flows with properties (CDS1)–(CDS3) (let us note that these properties are satisfied in the case of flows generated by autonomous systems of differential equations, the main class of flows which we study here).

Along with continuous dynamical systems (with continuous and discrete time), we consider smooth dynamical systems, replacing the condition of continuity of the mapping ϕ in (DDS3) and (CDS3) by the condition of smoothness of this mapping (the exact smoothness conditions are stated separately in every particular case).

Similarly to the case of a dynamical system with discrete time, we define the *trajectory of a point x in the flow ϕ* by the equality

$$O(x,\phi) = \{\phi(t,x) : t \in \mathbb{R}\}.$$

It is known (see the basic course of differential equations) that the following three types of trajectories of a flow are possible.

Consider a point $x_0 \in M$.

1. $O(x_0, \phi) = \{x_0\}$. Such a trajectory (and the point x_0 itself) is called a *rest point*.
2. $O(x_0, \phi) \neq \{x_0\}$, and the mapping $\phi(t, x_0)$ is periodic in t. In this case, the trajectory $O(x_0, \phi)$ is called a *closed trajectory* of the flow ϕ.
3. $\phi(t, x_0) \neq \phi(s, x_0)$ for $s \neq t$. In this case, the trajectory $O(x_0, \phi)$ is a one-to-one image of the line.

Similarly to the case of a discrete dynamical system, we say that a subset of the phase space is invariant under a flow if it contains trajectories of all its points. It is useful for the reader to formulate and prove an analog of Lemma 1.3 for flows.

As was mentioned above, we mostly study flows generated by autonomous systems of differential equations.

Consider an autonomous system of differential equations

$$\frac{dx}{dt} = F(x) \tag{1.1}$$

in the Euclidean space $\mathbb{R}^n$. We assume that the vector-function F is of class C^1 in $\mathbb{R}^n$.

Let x_0 be an arbitrary point of the space $\mathbb{R}^n$. It is known from the basic course of differential equations that there exists a number $h > 0$ with the following property: On the interval $(-h, h)$, there exists a unique solution $\phi(t, x_0)$ of system (1.1) with initial data $(0, x_0)$. As usual, the graph of the mapping

$$\phi(\cdot, x_0) : (-h, h) \to \mathbb{R}^n$$

i.e., the set

$$\{(t, \phi(t, x_0)) : t \in (-h, h)\}$$

is called the *integral curve* of the solution $\phi(t, x_0)$.

The projection of the integral curve to the space $\mathbb{R}^n$, i.e., the set

$$\{x = \phi(t, x_0) : t \in (-h, h)\}$$

is called the *trajectory* of the solution $\phi(t, x_0)$.

Let us first assume that every maximally continued solution of system (1.1) is defined for $t \in \mathbb{R}$. In this case, the corresponding mapping $\phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ has properties (CDS1)–(CDS3); thus, this mapping is a flow. Property (CDS1) holds since $\phi(0, x) = x$. Property (CDS2) is the group property of autonomous systems of differential equations (sometimes, this property is called the basic identity of autonomous systems).

Under our assumptions on the smoothness of F, the mapping ϕ is continuous in (t, x) and differentiable in t and x (these statements are corollaries of the definition of a solution of a differential equation and of theorems on continuity and differentiability of a solution with respect to initial values).

It is known that, in general, not every solution of a (nonlinear) system of differential equations can be continued to the whole real line. To avoid this difficulty, the following idea can be used.

Consider, along with system (1.1), the system

$$\frac{dx}{dt} = G(x), \tag{1.2}$$

where

$$G(x) = \frac{F(x)}{1 + |F(x)|^2},$$

and $|x|$ is the Euclidean norm of a vector x (thus, if $F = (F_1, \dots, F_n)$, then

$$|F(x)|^2 = F_1^2 + \cdots + F_n^2).$$

Below we write $F^2(x)$ instead of $|F(x)|^2$.

Clearly, the vector-function G is of class C^1, and the following inequality holds:

$$|G(x)| < 1, \quad x \in \mathbb{R}^n. \tag{1.3}$$

Let us denote by $\psi(t, x)$ the trajectory of system (1.2) with initial condition $\psi(0, x) = x$.

Inequality (1.3) implies that any maximally continued solution of system (1.2) is defined for $t \in \mathbb{R}$ (why?). Thus, system (1.2) generates a flow in $\mathbb{R}^n$.

Let us describe a relation between solutions of systems (1.1) and (1.2). Let $y(t)$ be a solution of system (1.2) defined for $t \in \mathbb{R}$. Consider the function

$$H(\tau) = \int_0^\tau \frac{ds}{1 + F^2(y(s))}$$

defined for $\tau \in \mathbb{R}$. Let V be the range of values of the function H.

Since

$$\frac{dH}{d\tau} > 0,$$

for any $t \in V$, the equation $t = H(\tau)$ has a unique solution $\theta(t)$; clearly, the function θ is of class C^1.

Differentiating the identity $\theta(H(\tau)) \equiv \tau$ in τ, we get the following identity:

$$\frac{d\theta}{dt}\frac{dH}{d\tau} \equiv 1.$$

Hence,

$$\frac{d\theta}{dt} = \left(\frac{1}{1 + F^2(y(\tau))} \right)^{-1} = 1 + F^2(y(\theta(t))).$$

Let us check that the function $z(t) = y(\theta(t))$ is a solution of system (1.1). Indeed,

$$\frac{dz}{dt} = \frac{dy}{d\theta}\frac{d\theta}{dt} = \frac{F(y(\theta(t)))}{1 + F^2(y(\theta(t)))}(1 + F^2(y(\theta(t)))) = F(z(t)).$$

Thus, solutions of systems (1.1) and (1.2) (as well as trajectories of these solutions) differ by parametrization only; the structure of partition of the phase space into trajectories is the same for both systems.

Since the structure of partition of the phase space into trajectories is the main object which we study in this book, in what follows we assume that autonomous systems of differential equations we work with generate flows.

When one studies the global structure of dynamical systems, it is natural to consider systems on manifolds. We treat in detail smooth dynamical systems with discrete time generated by diffeomorphisms of smooth manifolds. For flows generated by smooth vector fields, we formulate analogs of results established for smooth dynamical systems with discrete time. For this reason, we do not give here exact definitions from the theory of smooth vector fields on manifolds (the reader can find them, for example, in the book [12]).

Let us briefly recall that a smooth tangent vector field F on a smooth manifold M is a smooth mapping of the manifold M into its tangent bundle TM.

A smooth curve

$$\gamma = \phi(\cdot, x) : I \to M,$$

where I is an interval of the real line, is called the trajectory of a point $x \in M$ for the field F if $\phi(0, x) = x$ and the tangent vector of γ at the point $\phi(t, x)$ coincides with the vector $F(\phi(t, x))$ for any $t \in I$.

If the manifold M is compact, then any trajectory of a field F can be continued to $\mathbb{R}$; thus, in this case any vector field F generates a flow on M.

Let us indicate several relations between the objects defined above and theories of vector fields and differential equations.

1.2 Embedding of a discrete dynamical system into a flow

Let ϕ be a flow on a topological space M. Fix $T > 0$ and consider the mapping $f : M \to M$ defined by the formula $f(x) = \phi(T, x)$.

Let us show that f is a homeomorphism of M. Indeed, if $g(x) = \phi(-T, x)$, then g is continuous (see property (CDS3)), and

$$g(f(x)) = \phi(-T, \phi(T, x)) = \phi(0, x) = x$$

by properties (CDS2) and (CDS1). A similar reasoning shows that $f(g(x)) = x$. Hence, g is the inverse of f, and f is a homeomorphism.

In this case, we say that the homeomorphisms f is *embedded into the flow* ϕ.

If M is a smooth manifold and the flow ϕ is smooth (which means that the mappings $\phi(t, \cdot)$ are smooth for any t), then the same reasoning as above shows that any homeomorphism f embedded into the flow ϕ is a diffeomorphism.

Clearly, the structure of the set of trajectories of a flow and the structure of the set of trajectories of a homeomorphism embedded into this flow are closely related.

Nevertheless, one must remember that, in general, properties of the corresponding dynamical systems may differ significantly.

Let us consider the following example.

Example 1.2. Let S be the circle of unit length; introduce on S coordinate $x \in [0, 1)$. Define a flow on S by the equality

$$\phi(t, x) = x + t \pmod 1;$$

in this flow, every point moves along the circle into positive direction with unit speed.

Clearly, the flow ϕ has exactly one trajectory; this is a closed trajectory coinciding with the circle S.

As was shown above, for any $T > 0$, the mapping $f(x) = \phi(T, x)$ is a homeomorphism of the circle S. The dynamics of f is different for rational and irrational T.

If $T \in (0, 1)$ equals n/m, where n and m are relatively prime natural numbers, then any point $x \in S$ is a periodic point of f of period m. If the number T is irrational, then the set of periodic points of f is empty, and every trajectory is a countable set of points that is dense in S (check this!)

In addition, one has to remember that there exist diffeomorphisms that cannot be embedded into flows generated by smooth vector fields, and the set of such flows is large enough; this set is residual in the space of all diffeomorphisms (the exact definitions and statement of the result can be found in Section 2.4).

1.3 Local Poincaré diffeomorphism

Consider system (1.1) and assume that a point $p \in \mathbb{R}^n$ is not a rest point. Fix a number $T > 0$ and denote $q = \phi(T, p)$.

Consider two smooth $(n-1)$-dimensional surfaces P and Q in $\mathbb{R}^n$ that contain the points p and q, respectively.

We assume that locally (in neighborhoods of the points p and q) these surfaces are determined by smooth mappings

$$\Phi : \mathbb{R}^{n-1} \to \mathbb{R}^n \quad \text{and} \quad \Psi : \mathbb{R}^{n-1} \to \mathbb{R}^n, \quad \Phi, \Psi \in C^1,$$

P is parametrized by parameter $s \in \mathbb{R}^{n-1}$, Q is parametrized by parameter $\sigma \in R^{n-1}$, and the equalities $\Phi(0) = p$ and $\Psi(0) = q$ hold.

We assume, in addition, that the surfaces P and Q are nondegenerate at the points p and q, respectively, which means that the ranks of the matrices

$$A = \frac{\partial \Phi}{\partial s}(0) \quad \text{and} \quad B = \frac{\partial \Psi}{\partial \sigma}(0)$$

equal $n - 1$.

Denote by $a_1, \dots, a_{n-1}$ and $b_1, \dots, b_{n-1}$ the columns of the matrices A and B, respectively.

In this case, the tangent spaces $T_p P$ of the surface P at the point p and $T_q Q$ of the surface Q at the point q are spanned by the vectors $a_1, \dots, a_{n-1}$ and $b_1, \dots, b_{n-1}$, respectively.

We say that the surfaces P and Q are transverse to the trajectory $\phi(t, p)$ at the points p and q if the tangent vectors $F(p)$ and $F(q)$ of the trajectory do not belong to the spaces $T_p P$ and $T_q Q$, respectively.

Theorem 1.1. *If the surfaces P and Q are transverse to the trajectory $\phi(t, p)$ at the points p and q, then the mapping determined by the shift along trajectories of system* (1.1) *is a diffeomorphism of a neighborhood of the point p in P to a neigborhood of the point q in Q.*

To prove Theorem 1.1, we apply a variant of the implicit function theorem which we formulate below (Theorem 1.2). Consider two Euclidean spaces $\mathbb{R}^l$ and $\mathbb{R}^m$ with coordinates x and y, respectively.

Theorem 1.2. *Let f be a mapping of class C^1 from a neighborhood of a point (a, b) in $\mathbb{R}^l \times \mathbb{R}^m$ to the space R^m. Assume that $f(a, b) = 0$ and $\operatorname{rank} \partial f / \partial y(a, b) = m$. Then there exists a neighborhood U of the point a in $\mathbb{R}^l$ and a mapping g of class C^1 from the neighborhood U to $\mathbb{R}^m$ such that $g(a) = b$ and $f(x, g(x)) = 0$ for $x \in U$.*

Proof of Theorem 1.1. Since the vector $F(p)$ does not belong to the space spanned by the vectors $a_1, \dots, a_{n-1}$,

$$\operatorname{rank}(A, F(p)) = n. \tag{1.4}$$

Similarly,

$$\operatorname{rank}(B, F(q)) = n. \tag{1.5}$$

The trajectory of system (1.1) starting at a point $\Phi(s) \in P$ intersects the surface Q if and only if there exist $t \in \mathbb{R}$ and $\sigma \in \mathbb{R}^{n-1}$ such that $\phi(t, \Phi(s)) = \Psi(\sigma)$.

Consider the function

$$f(s, t, \sigma) = \phi(t, \Phi(s)) - \Psi(\sigma).$$

This function maps a neighborhood of the point $(0, T, 0)$ in $\mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R}^{n-1}$ to the space $\mathbb{R}^n$. Since the solution $\phi(t, x)$ is continuously differentiable in t and x, the function f is of class C^1. In addition,

$$f(0, T, 0) = \phi(T, p) - q = 0.$$

Let us calculate the Jacobi matrix

$$\begin{aligned}\frac{\partial f}{\partial(t,\sigma)}(0,T,0) &= \left(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial \sigma}\right)(0,T,0)\\ &= \left(\frac{\partial \phi(t,\Phi(s))}{\partial t}, -\frac{\partial \Psi(\sigma)}{\partial \sigma}\right)(0,T,0) = (F(q), -B)\,.\end{aligned}$$

Equality (1.5) implies that f satisfies the conditions of Theorem 1.2 with $l = n-1$, $m = n$, $a = 0$, and $b = (T, 0)$. Hence, there exist mappings $t(s)$ and $\sigma(s)$ of class C^1 defined for small $|s|$ such that $f(s, t(s), \sigma(s)) = 0$, i.e.,

$$\phi(t(s), \Phi(s)) = \Psi(\sigma(s)),$$

$t(0) = T$, and $\sigma(0) = 0$. The mapping which assigns to points $\Phi(s) \in P$ with small $|s|$ the points $\phi(t(s), \Phi(s)) = \Psi(\sigma(s)) \in Q$ is differentiable and invertible (the existence and differentiability of the inverse mapping are proved similarly using equality (1.4)). The theorem is proven. □

Remark. We can apply the reasoning used in the proof of Theorem 1.1 to the function

$$f(x, t, \sigma) = \phi(t, x) - \Psi(\sigma),$$

which maps a neighborhood of the point $(p, T, 0)$ in $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{n-1}$ to the space $\mathbb{R}^n$ and show that there exists a neigborhood U of the point p in $\mathbb{R}^n$ and mappings $t(x)$ and $\sigma(x)$ of class C^1 defined in U such that $f(x, t(x), \sigma(x)) = 0$, i.e.,

$$\phi(t(x), x) = \Psi(\sigma(x)),$$

and the following limit relations hold: $t(x) \to T$ and $\sigma(x) \to 0$ as $x \to p$.

Thus, any trajectory that intersects a small neighborhood of the point p intersects the surface Q as well.

The diffeomorphism given by Theorem 1.1 is called the *local Poincaré diffeomorphism* generated by the transverse surfaces P and Q.

The most important particular case of the construction described by Theorem 1.1 arises when the trajectory of the point p corresponds to a nonconstant periodic solution of system (1.1) (thus, it is a closed trajectory) and the surfaces P and Q coincide (precisely this case was studied by Poincaré).

1.4 Time-periodic systems of differential equations

Consider a time-periodic system of differential equations,

$$\frac{dx}{dt} = F(t, x), \tag{1.6}$$

where $x \in \mathbb{R}^n$. We assume that the vector-function F is of class $C^{0,1}_{t,x}$ in $\mathbb{R} \times \mathbb{R}^n$ and

$$F(t + \omega, x) \equiv F(t, x)$$

for some $\omega > 0$.

Denote by $x(t, t_0, x_0)$ the solution of system (1.6) with initial values (t_0, x_0).

It is well known that if $x(t)$ is a solution of system (1.6) and $k \in \mathbb{Z}$, then the function $x(t + k\omega)$ is a solution as well.

For definiteness, we assume that every solution of system (1.6) can be continued to $\mathbb{R}$.

Consider the mapping $T(\xi) = x(\omega, 0, \xi)$. Let us show that T is a diffeomorphism of the space $\mathbb{R}^n$. Let $U(\xi) = x(-\omega, 0, \xi)$. Fix $\xi \in \mathbb{R}^n$ and denote $\xi' = U(\xi)$.

Consider two solutions $x_1(t) = x(t, 0, \xi')$ and $x_2(t) = x(t - \omega, -\omega, \xi')$ of system (1.6) (the function $x_2(t)$ is a solution as the shift by $-\omega$ of the solution $x(t, -\omega, \xi')$).

Since $x_1(0) = \xi'$ and $x_2(0) = x(-\omega, -\omega, \xi') = \xi'$, the solutions $x_1(t)$ and $x_2(t)$ coincide.

By uniqueness, $x_2(\omega) = x(0, -\omega, x(-\omega, 0, \xi)) = \xi$. Since $x_1(\omega) = \xi$, we get the equality $T(U(\xi)) = \xi$, which shows that U is the inverse of T.

The mappings T are U are differentiable; hence, T is a diffeomorphism (called the *Poincaré diffeomorphism* of system (1.6)).

The following statement holds.

Lemma 1.4. *A solution $x(t, 0, x_0)$ of system* (1.6) *has period $m\omega$ if and only if x_0 is a fixed point of the diffeomorphism T^m.*

Proof. If a solution $x(t, 0, x_0)$ has period $m\omega$, then

$$x(t, 0, x_0) \equiv x(t + m\omega, 0, x_0).$$

Set $t = 0$ in the above identity to show that $x_0 = x(m\omega, 0, x_0) = T^m(x_0)$. Hence, x_0 is a fixed point of T^m.

Assume now that x_0 is a fixed point of T^m, i.e., $x_0 = T^m(x_0)$. Consider the solutions $x_1(t) = x(t, 0, x_0)$ and $x_2(t) = x(t + m\omega, 0, x_0)$. Since $x_1(0) = x_0$ and $x_2(0) = x(m\omega, 0, x_0) = T^m(x_0) = x_0$, the solutions coincide, which means that $x(t, 0, x_0)$ is $m\omega$-periodic. □

Thus, the important problem on the existence of periodic solutions of a system of differential equations is reduced to the problem on the existence of a fixed point of a diffeomorphism; the modern mathematics has a wide class of methods for the latter problem.

1.5 Action of an Abelian group

Let G be an Abelian group, i.e., a set with binary operation $*$ that satisfies the following axioms:

(G1) the operation $*$ is associative, i.e., $(a * b) * c = a * (b * c)$ for $a, b, c \in G$;

(G2) there exist a unit, i.e., an element $e \in G$ such that $a * e = e * a$ for $a \in G$;

(G3) there exist inverse elements, i.e., for any $a \in G$ there exists an element $b \in G$ such that $a * b = e$;

(G4) $a * b = b * a$ for any $a, b \in G$.

The *action of the group* G on a topological space M is a mapping $\phi : G \times M \to M$ with the following properties:

(A1) $\phi(e, x) = x$ for $x \in M$;

(A2) $\phi(a * b, x) = \phi(a, \phi(b, x))$ for $a, b \in G$ and $x \in M$.

Usually, continuous actions are considered, i.e., it is assumed that

(A3) the mapping $\phi(a, \cdot)$ is continuous for any $a \in G$.

The *trajectory (orbit)* of a point $x \in M$ under the action of the group G is the set

$$O(x, G) = \{\phi(a, x) : a \in G\}.$$

Clearly, dynamical systems with discrete and continuous time are actions of the groups $\mathbb{Z}$ and $\mathbb{R}$, respectively (where $*$ denotes addition).

Chapter 2

Topologies on spaces of dynamical systems

2.1 C^0-topology

Let (M, dist) be a compact metric space. If f and g are two homeomorphisms of the space M, we set

$$\rho_0(f, g) = \max_{x \in M} \max(\text{dist}(f(x), g(x)), \text{dist}(f^{-1}(x), g^{-1}(x))). \tag{2.1}$$

It is easy to show that ρ_0 is a metric on the space of homeomorphisms of the space M.

We denote by $H(M)$ the space of homeomorphisms of the space M with the metric ρ_0; the topology induced by the metric ρ_0 is called the C^0-*topology*.

Lemma 2.1. *The metric space $H(M)$ is complete.*

Proof. Consider a sequence of homeomorphisms f_m that is fundamental with respect to ρ_0.

This means that for any $\varepsilon > 0$ we can find an index m_0 such that $\rho_0(f_l, f_k) < \varepsilon$ for $k, l > m_0$.

Then

$$\max_{x \in M} \text{dist}(f_l(x), f_k(x)) < \varepsilon$$

and

$$\max_{x \in M} \text{dist}(f_l^{-1}(x), f_k^{-1}(x)) < \varepsilon$$

for $k, l > m_0$.

Thus, the sequences f_m and f_m^{-1} are fundamental with respect to the uniform metric

$$r(f, g) = \max_{x \in M} \text{dist}(f(x), g(x)).$$

Since the space of continuous mappings is complete with respect to the uniform metric r, there exist continuous mappings f and g of the space M such that $r(f_m, f) \to 0$ and $r(f_m^{-1}, g) \to 0$ as $m \to \infty$.

Fix a point $x \in M$. Passing to the limit as $m \to \infty$ in the equality $f_m(f_m^{-1}(x)) = x$, we see that $f(g(x)) = x$. Similarly, $g(f(x)) = x$.

Thus, f is a homeomorphism of the space M, and $g = f^{-1}$. Clearly, $\rho_0(f_m, f) \to 0$ as $m \to \infty$. The proof is complete. □

Remark. It is easy to show that if we replace the metric ρ_0 by the uniform metric r, then the appearing space of homeomorphisms will not be complete.

Indeed, let M be the segment $[0, 1]$. Fix an integer $m > 1$ and consider a continuous mapping $f_m : [0, 1] \to [0, 1]$ defined as follows: $f_m(0) = 0$, $f_m(1) = 1$, $f_m(1/3) = 1/m$, $f_m(2/3) = 1 - 1/m$, and f_m is affine on any of the segments $[0, 1/3]$, $[1/3, 2/3]$, and $[2/3, 1]$. Clearly, f_m is a homeomorphism of the segment $[0, 1]$.

It is easily seen that the inequality

$$|f_m(x) - f_n(x)| \le \max\left(\frac{1}{m}, \frac{1}{n}\right), \quad x \in [0, 1],$$

holds for any $m, n > 1$; thus, the sequence f_m is fundamental with respect to the uniform metric r. Hence, this sequence converges with respect to r, and the limit function f equals 0 on $[0, 1/3]$ and 1 on $[2/3, 1]$. Hence, f is not a homeomorphism.

Now let ϕ and ψ be two flows on a compact metric space (M, dist). It was shown in Section 1.2 that for any $t \neq 0$, the mapping $\phi(t, \cdot)$ is a homeomorphism of the space M (for $t = 0$, the mapping $\phi(0, \cdot) = \text{Id}$ is a homeomorphism as well). Define

$$\rho_0(\phi, \psi) = \max_{t \in [-1,1]} \rho_0(\phi(t, \cdot), \psi(t, \cdot)). \tag{2.2}$$

Clearly, ρ_0 is a metric on the space of flows on M. We denote by $\mathcal{F}^0(M)$ the space of flows on M with the metric ρ_0; similarly to the case of the space of homeomorphisms, the topology induced by the metric ρ_0 is called the C^0-topology.

The same reasoning as in the proof of Lemma 2.1 shows that $\mathcal{F}^0(M)$ is a complete metric space.

2.2 C^1-topology

Let M be a smooth closed (i.e., compact and boundaryless) manifold. To introduce the C^1-topology on the space of diffeomorphisms of M, we assume that M is a submanifold of the Euclidean space $\mathbb{R}^N$ (a different, equivalent, approach to definition of the C^1-topology based on local coordinates is described in [9]).

No generality is lost assuming that M is a submanifold of a Euclidean space since, by the classical Whitney theorem, any smooth closed manifold can be embedded into a Euclidean space of appropriate dimension.

If M is a submanifold of $\mathbb{R}^N$, for any point $x \in M$ we can identify the tangent space $T_x M$ of M at x with a linear subspace of $\mathbb{R}^N$. Consider the metric dist on M induced by the Euclidean metric of the space $\mathbb{R}^N$. For a vector $v \in T_x M$ we denote by $|v|$ its norm as the norm in the space $\mathbb{R}^N$.

Let f and g be two diffeomorphisms of the manifold M. Define the value $\rho_0(f, g)$ by the same formula (2.1) as for homeomorphisms of a compact metric space.

Take a point x of the manifold M and a vector v from the tangent space T_xM. We consider the vectors $Df(x)v \in T_{f(x)}M$ and $Dg(x)v \in T_{g(x)}M$ as vectors of the same Euclidean space $\mathbb{R}^N$. Hence, the following values are defined: $|Df(x)v - Dg(x)v|$ and

$$\|Df(x) - Dg(x)\| = \sup_{v \in T_xM, |v|=1} |Df(x)v - Dg(x)v|.$$

Similarly, one defines the value

$$\|Df^{-1}(x) - Dg^{-1}(x)\| = \sup_{v \in T_xM, |v|=1} |Df^{-1}(x)v - Dg^{-1}(x)v|.$$

Introduce the number

$$\rho_1(f, g) = \rho_0(f, g) + \sup_{x \in M} \|Df(x) - Dg(x)\| + \sup_{x \in M} \|Df^{-1}(x) - Dg^{-1}(x)\|.$$

Clearly, ρ_1 is a metric on the space of diffeomorphisms of the manifold M. We denote by $\mathrm{Diff}^1(M)$ the space of diffeomorphisms of M with metric ρ_1; the topology induced by the metric ρ_1 is called the C^1-*topology*.

The standard reasoning (left to the reader) shows that ($\mathrm{Diff}^1(M)$,ρ_1) is a complete metric space.

Now we consider the space of smooth flows on M. We say that a flow $\phi : \mathbb{R} \times M \to M$ is smooth if for any $t \in \mathbb{R}$, the mapping $\phi(t, \cdot)$ is smooth (for our purposes, it is enough to assume that this mapping is of class C^1; this assumption is satisfied if we consider a flow generated by a vector field of class C^1).

Our reasoning above (see Section 1.2) shows that if a flow $\phi : \mathbb{R} \times M \to M$ is smooth, then the mapping $\phi(t, \cdot)$ is a diffeomorphism of the manifold M for any t.

If ϕ and ψ are two smooth flows on M, we set

$$\rho_1(\phi, \psi) = \max_{t \in [-1,1]} \rho_1(\phi(t, \cdot), \psi(t, \cdot)). \tag{2.3}$$

It is easy to show that ρ_1 is a metric on the space of smooth flows on M; we denote by $\mathcal{F}^1(M)$ the space of smooth flows on M with the metric ρ_1. Similarly to the case of diffeomorphisms, the topology induced by the metric ρ_1 is called the C^1-topology.

2.3 Metrics on the space of systems of differential equations

Considering flows generated by vector fields on smooth closed manifolds, we have introduced two metrics, ρ_0 and ρ_1. Defining these metrics, we estimated differences between trajectories of the flows with the same initial values and between the corresponding "variational flows" on time intervals of fixed length.

Considering flows generated by autonomous systems of differential equations, it is natural to study metrics that are based on differences between right-hand sides of the systems rather than on differences between trajectories.

As was mentioned in Section 1.1, we denote by $|x|$ the Euclidean norm of a vector $x \in \mathbb{R}^n$.

For an $n \times n$ matrix A, let us denote by $\|A\|$ its operator norm generated by $|\cdot|$, i.e., the value

$$\|A\| = \max_{|x|=1} |Ax|.$$

Let us give a rough estimate of the operator norm of a matrix A, which we use below. Assume that the entries of A are a_{ij}, $i, j \in \{1, \dots, n\}$, and

$$|a_{ij}| \le M, \quad i, j \in \{1, \dots, n\}.$$

Let us write vectors $x, y \in \mathbb{R}^n$ as $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, respectively.

If $y = Ax$ and $|x| = 1$, then (by the Cauchy inequality)

$$\begin{aligned} y^2 &= \sum_{i=1}^{n} (a_{i1}x_1 + \dots + a_{in}x_n)^2 \\ &\le \sum_{i=1}^{n} (a_{i1}^2 + \dots + a_{in}^2)(x_1^2 + \dots + x_n^2) = \sum_{i,j=1}^{n} a_{ij}^2 \le n^2M^2, \end{aligned}$$

and we conclude that

$$\|A\| \le nM. \tag{2.4}$$

Consider two systems of differential equations,

$$\frac{dx}{dt} = F(x) \tag{2.5}$$

and

$$\frac{dx}{dt} = G(x), \tag{2.6}$$

in $\mathbb{R}^n$.

We assume that the vector-functions F and G are of class C^1 in $\mathbb{R}^n$. Denote by ϕ and ψ the flows generated by systems (2.5) and (2.6), respectively (as was mentioned above, we assume that every system of differential equations which we consider generates a flow).

In addition, we assume that the Jacobi matrix $\partial F/\partial x$ of the vector-function F is bounded (in particular, this implies that F is globally Lipschitz continuous in $\mathbb{R}^n$; denote by L its global Lipschitz constant) and uniformly continuous in $\mathbb{R}^n$.

The above assumptions on the Jacobi matrix $\partial F/\partial x$ do not look very natural; this is what we "pay" for working in the noncompact space $\mathbb{R}^n$ (for a vector field of class C^1 on a compact manifold, the corresponding assumptions are satisfied automatically).

Finally, we assume that the values

$$r_0(F,G) = \sup_{x\in\mathbb{R}^n} |F(x) - G(x)|$$

and

$$r_1(F,G) = r_0(F,G) + \sup_{x\in\mathbb{R}^n} \left\| \frac{\partial F}{\partial x}(x) - \frac{\partial G}{\partial x}(x) \right\|$$

are finite (in fact, we study the case where these values tend to 0).

Below, we refer to the following elementary estimate of the difference between solutions of two systems of differential equations (to prove this estimate, it is enough to represent the considered solutions of Cauchy problems in the form of solutions of equivalent integral equations and to apply the Gronwall lemma; a similar but more complicated estimate is applied in the following proof of Lemma 2.2).

Consider two systems of differential equations,

$$\frac{dx}{dt} = f(t,x) \tag{2.7}$$

and

$$\frac{dx}{dt} = g(t,x), \tag{2.8}$$

where $x \in \mathbb{R}^n$.

Assume that the vector-functions f and g are continuous in $\mathbb{R}^{n+1}$, f is globally Lipschitz continuous in x with constant L, and the value

$$m = \sup_{(t,x)\in\mathbb{R}^{n+1}} |f(t,x) - g(t,x)|$$

is finite.

If $x(t)$ and $y(t)$ are solutions of systems (2.7) and (2.8), respectively, that are defined on the same segment $[a,b]$ and have the same initial values (t_0, x_0), where $t_0 \in [a,b]$, then

$$|x(t) - y(t)| \le m \exp(L(b-a)), \quad t \in [a,b]. \tag{2.9}$$

Lemma 2.2. (1) *If* $r_0(F,G) \to 0$*, then* $\rho_0(\phi,\psi) \to 0$.

(2) *If* $r_1(F,G) \to 0$*, then* $\rho_1(\phi,\psi) \to 0$.

Proof. Let x be an arbitrary point of the space $\mathbb{R}^n$. Since $\phi(t,x)$ and $\psi(t,x)$ are solutions of systems (2.5) and (2.6), respectively, with the same initial values $(0,x)$, estimate (2.9) implies that

$$\rho_0(\phi,\psi) = \sup_{x\in\mathbb{R}^n} \max_{t\in[-1,1]} |\phi(t,x) - \psi(t,x)| \le r_0(F,G)\exp(L).$$

This proves statement (1).

To estimate the value $\rho_1(\phi,\psi)$, we again fix a point $x \in \mathbb{R}^n$ and consider the derivatives of the flows ϕ and ψ with respect to initial values,

$$Y(t) = \frac{\partial\phi}{\partial x}(t)$$

and

$$Z(t) = \frac{\partial\psi}{\partial x}(t).$$

Recall that these derivatives are matrix-valued solutions of the variational systems

$$\frac{dY}{dt} = \Phi(t,Y), \quad \text{where } \Phi(t,Y) = \frac{\partial F}{\partial x}(t,\phi(t))Y, \tag{2.10}$$

and

$$\frac{dZ}{dt} = \Psi(t,Z), \quad \text{where } \Psi(t,Z) = \frac{\partial G}{\partial x}(t,\psi(t))Z, \tag{2.11}$$

respectively, with the same initial values $Y(0) = Z(0) = E$, where E is the unit matrix of size $n \times n$.

It was assumed that the Jacobi matrix $\partial F/\partial x$ is bounded; in addition, since $r_1(F,G) \to 0$, we may assume, for example, that $r_1(F,G) \le 1$. Hence, there exists a number $N > 0$ such that

$$\left\| \frac{\partial F}{\partial x} \right\| \le N, \quad \left\| \frac{\partial G}{\partial x} \right\| \le N, \quad x \in \mathbb{R}^n. \tag{2.12}$$

First we estimate the value $\|Z(t)\|$ for $t \in [-1,1]$ (we are not going to get an exact estimate of the value $\|Z(t)\|$; it is important for us to estimate this value by a constant depending on N and independent from the initial point x of the trajectory $\psi(t,x)$).

If z is a column of the matrix Z, then

$$\frac{dz}{dt} = \frac{\partial G}{\partial x}(t,\psi(t))z \tag{2.13}$$

and $|z(0)| = 1$.

We take the scalar product of equality (2.13) and the vector $z(t)$:

$$\left\langle \frac{dz}{dt}, z \right\rangle = \left\langle \frac{\partial G}{\partial x}(t, \psi(t)), z \right\rangle, \tag{2.14}$$

where $\langle \cdot \rangle$ denotes scalar product.

As above, we denote by $z^2(t)$ the square of the Euclidean norm of $z(t)$. Relations (2.12) and (2.14) imply that

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}z^2 &= \left\langle \frac{dz}{dt}, z \right\rangle = \left\langle \frac{\partial G}{\partial x}(t, \psi(t))z, z \right\rangle \\ &\le \left| \frac{\partial G}{\partial x}(t, \psi(t))z \right| |z| \le \left\| \frac{\partial G}{\partial x}(t, \psi(t)) \right\| |z||z| \le N z^2. \end{aligned}$$

Since $z(t) \neq 0$, it follows from the last inequality that

$$\frac{d}{dt}(\log z^2) \le 2N.$$

Integrating this inequality and taking into account that $z^2(0) = 1$, we get the estimate

$$z^2(t) \le N_1^2 := \exp(2N), \quad |t| \le 1.$$

Thus,

$$|z(t)| \le N_1,$$

and estimate (2.4) implies that

$$\|Z(t)\| \le N_2 := nN_1, \quad |t| \le 1.$$

Fix an arbitrary $\varepsilon > 0$ and find a $\delta > 0$ such that if $|x - x'| < \delta$, then

$$\left\| \frac{\partial F}{\partial x}(x) - \frac{\partial F}{\partial x}(x') \right\| \le \varepsilon.$$

Clearly, if $r_1(F, G) \to 0$, then $r_0(F, G) \to 0$. Find a $\delta_1 > 0$ such that if $r_1(F, G) < \delta_1$, then $r_0(F, G)\exp(L) < \delta$. The same reasoning as in the proof of statement (1) shows that

$$|\phi(t, x) - \psi(t, x)| \le \delta, \quad |t| \le 1,$$

for any $x \in \mathbb{R}^n$.

Then

$$\left\| \frac{\partial F}{\partial x}(\phi(t, x)) - \frac{\partial F}{\partial x}(\psi(t, x)) \right\| \le \varepsilon, \quad |t| \le 1, \tag{2.15}$$

for any $x \in \mathbb{R}^n$.

Let us write down equivalent integral equations for Y and Z:

$$Y(t) = E + \int_0^t \Phi(s, Y(s))\, ds \tag{2.16}$$

and

$$Z(t) = E + \int_0^t \Psi(s, Z(s))\, ds. \tag{2.17}$$

Relations (2.16) and (2.17) imply that

$$\|Y(t) - Z(t)\| \le \left| \int_0^t \|\Phi(s, Y(s)) - \Psi(s, Z(s))\|\, ds \right|.$$

The integrand is estimated as follows:

$$\begin{aligned} &\|\Phi(s, Y(s)) - \Psi(s, Z(s))\| \\ &\quad \le \|\Phi(s, Y(s)) - \Phi(s, Z(s))\| + \|\Phi(s, Z(s)) - \Psi(s, Z(s))\|. \end{aligned}$$

It follows from estimates (2.12) that

$$\|\Phi(s, Y(s)) - \Phi(s, Z(s))\| = \left\| \frac{\partial F}{\partial x}(\phi(s, x))(Y(s) - Z(s)) \right\| \le N\|Y(s) - Z(s)\|. \tag{2.18}$$

Further, if $r_1(F, G) < \delta_1$, then

$$\begin{aligned} \|\Phi(s, Z(s)) - \Psi(s, Z(s))\| &\le \left\| \left(\frac{\partial F}{\partial x}(\phi(s, x)) - \frac{\partial F}{\partial x}(\psi(s, x)) \right) Z(s) \right\| \\ &\quad + \left\| \left(\frac{\partial F}{\partial x}(\psi(s, x)) - \frac{\partial G}{\partial x}(\psi(s, x)) \right) Z(s) \right\| \\ &\le (\varepsilon + r_1(F, G))N_2. \end{aligned}$$

Combining these inequalities with estimate (2.18), we get the inequality

$$\|Y(t) - Z(t)\| \le \left| \int_0^t (N\|Y(s) - Z(s)\| + (\varepsilon + r_1(F, G))N_2)\, ds \right|.$$

Applying the Gronwall lemma to the above inequality, we conclude that if $r_1(F, G) < \delta_1$, then

$$\|Y(t) - Z(t)\| \le (\varepsilon + r_1(F, G))N_2 \exp(N)$$

for $t \in [-1, 1]$.

Since ε is arbitrary, the last inequality implies statement (2) of our lemma. □

2.4 Generic properties

In the global theory of dynamical systems, it is important to study generic properties.

Let X be a topological space. We say that a set $A \subset X$ is *residual* if there exists a countable family of open subsets $\{A_n : n \in \mathbb{Z}\}$ of the space X such that

$$\bigcap_{n\in\mathbb{Z}} A_n = A. \tag{2.19}$$

A property of elements of the space X is called *generic* if there exists a residual subset of X such that every element of this subset has this property. If a property is generic, we say that a *generic element* of the space X has this property.

The following classical theorem was proven by Baire.

Theorem 2.1. *If X is a complete metric space, then any residual subset is dense in X.*

A residual set is "large" in the topological sense. At the same time, a residual set can be "small" in the sense of measure.

Let us show that there exists a residual subset of the real line $\mathbb{R}$ whose Lebesgue measure is zero. Fix a countable dense subset $\{a_n, n = 0, 1, \dots\}$ of the line (for example, consider the set of rational numbers).

Take a natural number m and consider the set

$$A_m = \bigcup_{n\geq 0} \left(a_n - \frac{1}{m2^n}, a_n + \frac{1}{m2^n}\right).$$

The set A_m is an open and dense subset of the line, and its Lebesgue measure can be estimated as follows:

$$\text{mes}\, A_m \leq \sum_{n=0}^{\infty} \frac{2}{m2^n} = \frac{4}{m}.$$

The set $A = \bigcap_{m>0} A_m$ is residual, and mes $A = 0$.

It was mentioned in Section 1.2 that there exist diffeomorphisms that cannot be embedded into flows. The precise statement of the main result of [13] is as follows.

Theorem 2.2. *Let M be a smooth closed manifold whose dimension is not less than* 2*. Then a generic diffeomorphism in* $\text{Diff}^1(M)$ *cannot be embedded into a flow generated by a Lipschitz continuous vector field on M.*

2.5 Immersions and embeddings

In this book, we consider two main classes of mappings studied in differential topology, immersions and embeddings. We mostly restrict ourselves to immersions and embeddings of smooth manifolds (Euclidean spaces and disks in such spaces) into Euclidean spaces.

Let us recall the basic definitions.

Let f be a mapping of a manifold M to a manifold N. We say that the mapping f is a *topological immersion* if f is a homeomorphism of M and $f(M)$.

We say that f is an *immersion of class* $C^k, k \geq 1$, if f belongs to class C^k and

$$\operatorname{rank} Df(x) = \dim M$$

for any $x \in M$ (let us note that the last condition implies that $\dim N \geq \dim M$).

The mapping f is called an *embedding of class* $C^k, k \geq 1$, if $f(M)$ is a submanifold of N and f is a diffeomorphism of M and $f(M)$.

We will work with smooth disks. By definition, a *smooth disk* is the image of a ball of a Euclidean space under an embedding.

Consider a disk

$$D = \{x \in \mathbb{R}^k : |x|_e < r,\ r > 0\}$$

in the Euclidean space $\mathbb{R}^k$ and a manifold M (as above, we assume that M is a submanifold of a Euclidean space $\mathbb{R}^N$).

For two embeddings h and g of the disk D into the manifold M we set

$$\rho_1(h, g) = \sup_{x \in D} \left(|h(x) - g(x)| + \|Dh(x) - Dg(x)\|\right),$$

where $|\cdot|$ and $\|\cdot\|$ are the distance in $\mathbb{R}^N$ and the corresponding operator norm (to show that all the objects are properly defined, one can use the same reasoning as in Section 2.2).

Chapter 3
Equivalence relations

3.1 Topological conjugacy

Consider two homeomorphisms $f : M \to M$ and $g : N \to N$, where M and N are topological spaces.

We say that the homeomorphisms f are g *topologically conjugate* if there exists a homeomorphism h of the spaces M and N such that

$$g(h(x)) = h(f(x)) \tag{3.1}$$

for any $x \in M$ (in other words, $g \circ h = h \circ f$).

In this case, the homeomorphism h is called *conjugating homeomorphism* (or *topological conjugacy*).

Sometimes, condition (3.1) is formulated in the following (equivalent) form: The diagram

$$\begin{array}{ccc} M & \xrightarrow{f} & M \\ h\big\downarrow & & h\big\downarrow \\ N & \xrightarrow{g} & N \end{array}$$

commutes.

The following simple (but very important) statement holds.

Lemma 3.1. *If $g \circ h = h \circ f$, then*

$$g^m \circ h = h \circ f^m \tag{3.2}$$

for any $m \in \mathbb{Z}$.

Proof. We apply induction to prove the statement for $m \geq 0$ (for $m < 0$, the proof is similar). If $m = 0$, $f^0 = g^0 =$Id, and the equality (3.2) takes the form $h = h$. Assume that equality (3.2) has been proven for some m.

Then

$$\begin{aligned} g^{m+1} \circ h &= g \circ (g^m \circ h) = g \circ (h \circ f^m) \\ &= (g \circ h) \circ f^m = (h \circ f) \circ f^m = h \circ f^{m+1}, \end{aligned}$$

which proves the statement of our lemma. □

Lemma 3.1 implies that a conjugating homeomorphism maps trajectories of the dynamical system generated by the homeomorphism f to trajectories of the dynamical system generated by the homeomorphism g.

Thus, if homeomorphisms f and g are topologically conjugate, then, from the topological point of view, the global structure of the set of trajectories of the dynamical systems generated by the homeomorphisms f and g is the same.

For example, periodic trajectories of the homeomorphism f are mapped to periodic trajectories of the homeomorphism g. Indeed, let p be a periodic point of f of period m, i.e., the points

$$p_0 = p, \quad p_1 = f(p), \quad \dots, \quad p_{m-1} = f^{m-1}(p)$$

are distinct and $f^m(p) = p$. If $r = h(p)$, then

$$r_i = g^i(r) = g^i(h(p)) = h(f^i(p)) = h(p_i)$$

by Lemma 3.1; thus, the points $r_i, i = 0, \dots, m-1$, are distinct and $g^m(r) = h(f^m(p)) = h(p) = r$.

A similar reasoning shows that if a trajectory $O(p, f)$ (a semitrajectory $O^+(p, f)$) is dense M, then the trajectory $O(h(p), g)$ (semitrajectory $O^+(h(p), g)$) is dense in N.

Remark. Sometimes, it is possible to significantly simplify a problem by passing from a homeomorphism to a topologically conjugate homeomorphism (this idea will be applied in Section 9, where we study the Smale horseshoe).

Here we consider as an example two semi-dynamical systems.

Let f be a continuous mapping of a topological space M into itself. We set $\phi(m, x) = f^m(x), m \in \mathbb{Z}_+$, and get a *semi-dynamical system*, i.e., a mapping $\phi : \mathbb{Z}_+ \times M \to M$ whose properties are similar to properties (DDS1)–(DDS3) (one has to replace $\mathbb{Z}$ by $\mathbb{Z}_+$ in properties (DDS2) and (DDS3)).

The trajectory of a point x in the semi-dynamical system ϕ is defined by the equality

$$O(x, \phi) = \{\phi(m, x) : m \in \mathbb{Z}_+\};$$

the definition of a periodic point is literally the same as in the case of a dynamical system.

Two semi-dynamical systems generated by mappings f and g are called *topologically conjugate* if there exists a homeomorphism h that satisfies equality (3.1).

Exercise 3.1. Consider two semi-dynamical systems on the segment [0, 1] generated by the mappings

$$f(x) = 4x(1 - x)$$

and

$$g(x) = \begin{cases} 2x, & x \in [0, 1/2], \\ 2(1-x), & x \in (1/2, 1]. \end{cases}$$

Prove that the mapping

$$h(x) = \frac{1}{2\pi} \arcsin \sqrt{x}$$

is a homeomorphism of the segment [0, 1] that conjugates the semi-dynamical systems generated by the mappings f and g.

Thus, we can reduce the study of the dynamics of the essentially nonlinear mapping f to the similar problem for the piecewise linear mapping g.

In what follows, we consider dynamical systems with the same phase space M.

Lemma 3.2. *Topological conjugacy is an equivalence relation on the space $H(M)$.*

Proof. Since the identity homeomorphism Id conjugates any homeomorphism with itself, topological conjugacy is reflexive.

Topological conjugacy is symmetric. Indeed, if a homeomorphism h conjugates f and g, i.e., $g \circ h = h \circ f$, then h^{-1} conjugates g and f; to show this, apply h^{-1} to the equality $g \circ h = h \circ f$ both from the right and left. As a result, we get the desired equality $f \circ h^{-1} = h^{-1} \circ g$.

Finally, we show that topological conjugacy is transitive. Assume that h_1 conjugates f and g and h_2 conjugates g and k. Then $h = h_2 \circ h_1$ is a homeomorphism of the space M, and

$$h \circ f = h_2 \circ (h_1 \circ f) = h_2 \circ (g \circ h_1) = (k \circ h_2) \circ h_1 = k \circ h,$$

which completes the proof of our lemma. □

Of course, topological conjugacy is an equivalence relation on the space of diffeomorphisms $\mathrm{Diff}^1(M)$ of a smooth manifold M as well.

This relation allows us to give the main definition of the theory of structural stability.

Let M be a smooth closed manifold. A diffeomorphism $f \in \mathrm{Diff}^1(M)$ is called *structurally stable* if there exists a neighborhood W of the diffeomorphism f in the C^1-topology such that any diffeomorphism $g \in W$ is topologically conjugate with f.

The above definition and Lemma 3.2 imply that any diffeomorphism $g \in W$ is structurally stable as well. Denote by $\mathcal{S}(M)$ the set of structurally stable diffeomorphisms in $\mathrm{Diff}^1(M)$. Clearly, the following statement holds (since this statement is very important for us, we call it a theorem).

Theorem 3.1. *The set* $\mathcal{S}(M)$ *is open in* $\mathrm{Diff}^1(M)$.

The property of structural stability was first defined by Andronov and Pontryagin for autonomous systems of differential equations (we discuss this definition below, in Section 7.6).

In fact, this original definition corresponds to a slightly different property which we formulate below.

A diffeomorphism $f \in \mathrm{Diff}^1(M)$ is called *structurally stable in the strong sense* if for any $\varepsilon > 0$ one can find a neighborhood W of the diffeomorphism f in the C^1-topology such that for any diffeomorphism $g \in W$ there exists a homeomorphism h that topologically conjugates g and f and satisfies the inequality

$$\max_{x \in M} \mathrm{dist}(h(x), x) < \varepsilon.$$

It is easy to understand that this definition does not imply immediately that the set of diffeomorphisms that are structurally stable in the strong sense is open.

Let us pass to the case of flows. Consider two flows $\phi : \mathbb{R} \times M \to M$ and $\psi : \mathbb{R} \times N \to N$, where M and N are topological spaces.

The flows ϕ and ψ are called *topologically conjugate* if there exists a homeomorphism h of the spaces M and N such that

$$\psi(t, h(x)) = h(\phi(t, x)) \tag{3.3}$$

for any $t \in \mathbb{R}$ and $x \in M$.

Thus, topological conjugacy of flows means that there exists a homeomorphism of their phase spaces that maps trajectories to trajectories and preserves time t.

Let us show that the notion of topological conjugacy of flows is too fine for the problem of global classification of flows generated by systems of differential equations.

Consider, for example, two autonomous systems of differential equations in the plane $\mathbb{R}^2$ with coordinates (x, y):

$$\frac{dx}{dt} = -2y, \quad \frac{dy}{dt} = 2x, \tag{3.4}$$

and

$$\frac{dx}{dt} = -y, \quad \frac{dy}{dt} = x. \tag{3.5}$$

Let ϕ and ψ be the flows generated by systems (3.4) and (3.5), respectively.

The origin is a rest point of both systems (3.4) and (3.5); the remaining trajectories are concentric circles with center at the origin along which points move in the positive direction as t grows. Fix an initial point $(x_0, 0)$. The trajectories of this point in the flows ϕ and ψ are given by the following formulas:

$$\phi(t, x_0, 0): \quad x = x_0 \cos 2t, \quad y = x_0 \sin 2t,$$

and

$$\psi(t, x_0, 0): \quad x = x_0 \cos t, \quad y = x_0 \sin t,$$

respectively.

Clearly, the sets of trajectories of systems (3.4) and (3.5) are the same from the topological point of view; at the same time, the flows ϕ and ψ are not topologically conjugate. Let us show this.

To get a contradiction, let us assume that there exists a homeomorphism $h : \mathbb{R}^2 \to \mathbb{R}^2$ for which equality (3.3) holds. The trajectory of the point $\phi(t, 1, 0)$ is closed; hence, the homeomorphism h must map this trajectory to a closed trajectory. Thus, if $h(1, 0) = (x_0, y_0)$, then

$$(x_0, y_0) \neq (0, 0). \tag{3.6}$$

The formula defining the flow ϕ implies that $\phi(\pi, 1, 0) = (1, 0)$. Hence,

$$(x_0, y_0) = h(1, 0) = h(\phi(\pi, 1, 0)) = \psi(\pi, h(1, 0)) = \psi(\pi, x_0, y_0) = -(x_0, y_0),$$

and we get a contradiction with inequality (3.6).

In the problem of global classification of flows, a different notion of equivalence is used. We discuss this property in the next section.

3.2 Topological equivalence of flows

Two flows $\phi : \mathbb{R} \times M \to M$ and $\psi : \mathbb{R} \times N \to N$, where M and N are topological spaces, are called *topologically equivalent* if there exists a homeomorphism h of the spaces M and N that maps trajectories of the flow ϕ to trajectories of the flow ψ and preserves the direction of movement along trajectories.

In other words, there exists a function $\tau : \mathbb{R} \times M \to \mathbb{R}$ such that

(1) for any $x \in M$, the function $\tau(\cdot, x)$ increases and maps $\mathbb{R}$ onto $\mathbb{R}$;

(2) $\tau(0, x) = x$ for any $x \in M$;

(3) $h(\phi(t, x)) = \psi(\tau(t, x), h(x))$ for any $(t, x) \in \mathbb{R} \times M$.

Clearly, the flows ϕ and ψ generated by systems of differential equations (3.4) and (3.5) are topologically equivalent; one may take as h the identical mapping of the plane and set $\tau(t, x) = 2t$.

A flow $\phi \in \mathcal{F}^1(M)$ is called *structurally stable* if there exists a neighborhood W of ϕ in $\mathcal{F}^1(M)$ such that any flow $\psi \in W$ is topologically equivalent to ϕ.

3.3 Nonwandering set

Some equivalence relations which are important for the global qualitative theory of dynamical systems are related to the notion of a nonwandering point.

Consider a homeomorphism f of a topological space M and the corresponding dynamical system.

A point $x_0 \in M$ is called *wandering* for f if there exists a neighborhood U of the point x_0 and a number $N > 0$ such that

$$f^n(U) \cap U = \emptyset \quad \text{for } |n| \geq N.$$

A point x_0 is called *nonwandering* if it is not wandering. Clearly, a point x_0 is wandering if for any neighborhood U of x_0 and for any number N there exists a point $x \in U$ and a number n, $|n| > N$, such that $f^n(x) \in U$.

We denote by $\Omega(f)$ *the set of nonwandering points* of a homeomorphism f (sometimes, the set $\Omega(f)$ is called the *nonwandering set*).

Under rather general assumptions on the space M (for example, these assumptions are satisfied for a metric space), we can give a different definition of a nonwandering point.

A topological space M is said to satisfy the *first axiom of countability* if any point of M has a countable base of neighborhoods (we gave a definition of a base of neighborhoods in Section 1 when Example 1.1 was considered). It is well known that any metric space satisfies the first axiom of countability.

Lemma 3.3. *Assume that the space M satisfies the first axiom of countability. A point $x_0 \in M$ is nonwandering for a homeomorphism f if and only if there exist sequences of points $p_k, q_k \in M$ and numbers τ_k, θ_k such that*

$$p_k, q_k, f^{\tau_k}(p_k), f^{\theta_k}(q_k) \to x_0,$$

$\tau_k \to \infty$, and $\theta_k \to -\infty$ as $k \to \infty$.

Proof. Clearly, if such sequences exist, then $x_0 \in \Omega(f)$.

Take a point $x_0 \in \Omega(f)$ and fix a countable base $V_m, m > 0$, of neighborhoods of the point x_0. For any natural m we can find a number $n(m)$ such that $|n(m)| > m$ and

$$f^{n(m)}(V_m) \cap V_m \neq \emptyset.$$

This means that there exist points $r_m \in V_m$ such that $f^{n(m)}(r_m) \in V_m$.

If the sequence $n(m)$ contains a subsequence $n(l_k) \to \infty$, we set $p_k = r_{n(l_k)}$, $\tau_k = n(l_k)$, $q_k = f^{\tau_k}(p_k)$, and $\theta_k = -\tau_k$.

The case of a subsequence $n(l_k) \to -\infty$ is considered similarly. □

Clearly, fixed and periodic points of a homeomorphism f are nonwandering. Indeed, if p is a periodic point of period m, then the points $f^{mk}(p) = p$ belong to any neighborhood of p, while the numbers mk can be arbitrarily large. Sometimes, such nonwandering points are called trivial.

There exist nontrivial nonwandering points.

Let us recall the notions of ω-limit and α-limit points of a trajectory $O(x, f)$. The ω-limit set, $\omega(x, f)$, of a trajectory $O(x, f)$ is, by definition, the set of limit points of all sequences $f^{n(k)}(x)$, where $n(k) \to \infty$ as $k \to \infty$. Similarly, the α-limit set, $\alpha(x, f)$, of a trajectory $O(x, f)$ is, by definition, the set of limit points of all sequences $f^{n(k)}(x)$, where $n(k) \to \infty$ as $k \to -\infty$.

Both sets $\omega(x, f)$ and $\alpha(x, f)$ are closed and invariant.

Lemma 3.4. *$\omega(x, f) \cup \alpha(x, f) \subset \Omega(f)$ for any point $x \in M$.*

Proof. Let us prove that $\omega(x, f) \subset \Omega(f)$; the case of α-limit set is considered analogously.

Take a point $x_0 \in \omega(x, f)$. There exists a sequence $n(k) \to \infty$, $k \to \infty$, such that $f^{n(k)}(x) \to x_0$.

Let U be an arbitrary neighborhood of the point x_0 and let N be an arbitrary number. There exists an index k_0 such that $f^{n(k)}(x) \in U$ for $k \geq k_0$. In addition, there exists an index $k_1 > k_0$ such that $n_1 := n(k_1) - n(k_0) > N$.

In this case, $f^{n(k_1)}(x) = f^{n_1}(f^{n(k_0)}(x)) \in U$, i.e., $f^{n_1}(U) \cap U \neq \emptyset$.

This means that $x_0 \in \Omega(f)$. □

One can show that there exist dynamical systems for which the nonwandering set contains points that are not ω-limit or α-limit points of individual trajectories. Below we give an example of a flow having this property (see Example 3.1); let us mention that some notions and constructions are more "visible" in the case of a flow.

Let us describe the basic properties of nonwandering sets.

Theorem 3.2. *The set $\Omega(f)$ is closed and invariant. If the space M is compact, then $\Omega(f) \neq \emptyset$.*

Proof. First we show that the set $\Omega(f)$ is closed. It follows from the definition that if x_0 is a wandering point, then any point of the neighborhood U mentioned in the definition is wandering as well. Thus, the set of wandering points is open; its complement $\Omega(f)$ is closed.

Now we prove that the set $\Omega(f)$ is invariant. Consider an arbitrary point $x_0 \in \Omega(f)$, an arbitrary neighborhood U of the point $x' = f(x_0)$, and an arbitrary number N. Since the mapping f is continuous, the set $U_1 = f^{-1}(U)$ is a neighborhood of the point x_0. Hence, there exists a point $x_1 \in U_1$ and a number n, $|n| > N$, such that $f^n(x_1) \in U_1$. Let $x = f(x_1)$. Then $x \in U$ and $f^n(x) = f(f^n(x_1)) \in f(U_1) = U$. Thus, $x' \in \Omega(f)$; it follows that $f(\Omega(f)) \subset \Omega(f)$. A similar reasoning shows that $f^{-1}(\Omega(f)) \subset \Omega(f)$. Hence, $f(\Omega(f)) = \Omega(f)$. Thus, the set $\Omega(f)$ is invariant by Lemma 1.3.

Now let us assume that the space M is compact. In this case, the ω-limit set of any trajectory is nonempty, and the last statement of our theorem is a corollary of Lemma 3.4. □

It is easy to understand that if the phase space of a dynamical system is not compact, then the nonwandering set may be empty.

As an example, consider the homeomorphism $f(x) = x + 1$ of the line $\mathbb{R}$.

In a sense, the global dynamics is characterized by the behavior of a dynamical system near its nonwandering set. In fact, for any trajectory, only a finite number of its points does not belong to a neighborhood of the nonwandering set. More precisely, the following theorem was proven by Birkhoff (the constant T whose existence is established in Theorem 3.3 is usually called the *Birkhoff constant* for a neighborhood U of the set $\Omega(f)$).

Theorem 3.3. *Assume that the phase space M of a dynamical system generated by a homeomorphism f is compact. Let U be an arbitrary neighborhood of the set $\Omega(f)$. There exists a number $T > 0$ such that*

$$\operatorname{card}\{k : f^k(x) \notin U\} \le T$$

for any point $x \in M$.

Proof. Fix a neighborhood U of the set $\Omega(f)$. For any point $x \in M \setminus U$ we can find a number $t(x)$ and neighborhood $V(x)$ such that

$$f^k(V(x)) \cap V(x) = \emptyset, \quad |k| \ge t(x).$$

Since the set $M \setminus U$ is compact, the covering $\{V(x)\}$ of the set $M \setminus U$ contains a finite subcovering $V_1, \dots, V_n$ with the following property: There exist numbers $t_1, \dots, t_n \ge 1$ such that

$$f^k(V_i) \cap V_i = \emptyset, \quad k \ge t_i, i = 1, \dots, n.$$

Set $t = \max t_i$ and $T = (n + 1)t$.

Let us prove that T has the required property. To get a contradiction, assume the contrary. Then there exists a point x and a set of integers

$$L = \{l(0) < l(1) < \cdots < l(m)\}$$

with $\operatorname{card} L = m + 1 \ge T$ such that $f^{l(i)}(x) \notin U$ for $i = 0, \dots, m$.

Note that if $i, j \in \{0, \dots, m\}$ and $j \ge i$, then $l(j) - l(i) \ge j - i$.

Set $y_0 = f^{l(0)}(x)$; let W_0 be a neighborhood from V_i, that contains the point y_0 (if such a neighborhood is not unique, take as W_0 any of them).

Note that $f^k(y_0) \notin W_0$ for $k \ge t$. Set $y_1 = f^{l(t)}(x)$; let W_1 be a neighborhood from the family V_i that contains the point y_1. Since $y_1 = f^{l(t)-l(0)}(y_0)$, it follows from the choice of t and from the inequality $l(t)-l(0) \ge t$ that $W_1 \ne W_0$. In addition, $f^k(y_1) \notin W_0$ for $k \ge 0$ and $f^k(y_1) \notin W_0 \cup W_1$ for $k \ge t$.

The inequality $m + 1 \geq (n + 1)t$ implies that $m \geq nt$; hence, the set L contains the number $l(nt)$.

The same reasoning as applied above to the points y_0 and y_1 shows that if $y_j = f^{l(jt)}(x)$, $j = 1, \dots, n$, and W_j are elements of the family $\{V_1, V_2, \dots, V_n\}$ containing the points y_j, then the sets W_j are pairwise disjoint, and

$$f^k(y_j) \notin W_0 \cup W_1 \cup \dots \cup W_{j-1}, \quad k \geq 0, j = 1, \dots, n.$$

Hence,

$$f^k(y_n) \notin W_0 \cup W_1 \cup \dots \cup W_{n-1}, \quad k \geq 0.$$

Since the sets $W_0, W_1, \dots, W_{n-1}$ are pairwise disjoint elements of the family $\{V_1, V_2, \dots, V_n\}$, the families $\{V_1, V_2, \dots, V_n\}$ and $\{W_0, W_1, \dots, W_{n-1}\}$ coincide (up to the numbering of their elements). Hence,

$$f^k(y_n) = f^{l(nt)+k}(x) \notin M \setminus U, \quad k \geq 0.$$

If we set $k = 0$ in the above relation, we get a contradiction with the inclusion $l(nt) \in L$. This completes the proof. □

Consider two homeomorphisms $f : M \to M$ and $g : N \to N$. We say that the homeomorphisms f and g are *Ω-conjugate* if there exists a homeomorphism h of $\Omega(f)$ and $\Omega(g)$ such that $g(h(x)) = h(f(x))$ for $x \in \Omega(f)$ (let us explain that in this case h is a one-to-one mapping of $\Omega(f)$ onto $\Omega(g)$ such that both h and h^{-1} are continuous with respect to topologies induced on the sets $\Omega(f)$ and $\Omega(g)$ by the topologies of the spaces M and N).

Lemma 3.5. *If h topologically conjugates f and g, then $h(\Omega(f)) = \Omega(g)$.*

Proof. Consider an arbitrary point $x_0 \in \Omega(f)$ and fix an arbitrary neighborhood U of the point $y_0 = h(x_0)$ and an arbitrary number N. Since h is continuous, the set $V = h^{-1}(U)$ is a neighborhood of the point x_0. Hence, there exists a point $x \in V$ and a number n, $|n| > N$, such that $f^n(x) \in V$. Denote $y = h(x)$. Then $y \in U$ and $g^n(y) = g^n(h(x)) = h(f^n(x)) \in U$. Thus, $y \in \Omega(g)$; hence, $h(\Omega(f)) \subset \Omega(g)$. A similar reasoning shows that $h^{-1}(\Omega(g)) \subset \Omega(f)$. We conclude that $h(\Omega(f)) = \Omega(g)$. □

Corollary. *If homeomorphisms f and g are topologically conjugate, then they are Ω-conjugate.*

The same reasoning as in the proof of Lemma 3.2 shows that Ω-conjugacy is an equivalence relation on the space of homeomorphisms $H(M)$.

Let M be a smooth closed manifold. A diffeomorphism $f \in \mathrm{Diff}^1(M)$ is called Ω-*stable* if there exists a neighborhood W of the diffeomorphism f in the C^1-topology such that any diffeomorphism $g \in W$ is Ω-conjugate with f.

The corollary of Lemma 3.5 implies that if a diffeomorphism is structurally stable, then it is Ω-stable.

Nonwandering points of flows are defined similarly to nonwandering points of cascades.

Let ϕ be a flow on a topological space M. A point $x_0 \in M$ is called a *nonwandering point* of the flow ϕ if for any neighborhood U of the point x_0 and for any number N there exists a point $x \in U$ and a number t, $|t| > N$, such that $\phi(t, x) \in U$.

The set of nonwandering points of a flow has the same basic properties as the set of nonwandering points of a cascade.

The definition of an Ω-stable flow repeats that of an Ω-stable diffeomorphism (with a natural replacement of Ω-conjugacy by Ω-equivalence).

Let us give an example of a flow whose nonwandering set does not coincide with the union of the sets of ω-limit and α-limit points of individual trajectories.

Example 3.1. Consider the following autonomous system of differential equations in the plane $\mathbb{R}^2$ with coordinates (x, y):

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = x(1 - x^2). \tag{3.7}$$

System (3.7) has an integral

$$U(x, y) = y^2 - x^2 + x^4/2;$$

hence, trajectories of the system belong to the sets

$$y = \pm\sqrt{C + x^2 - x^4/2}.$$

The set of trajectories is shown in Figure 1.

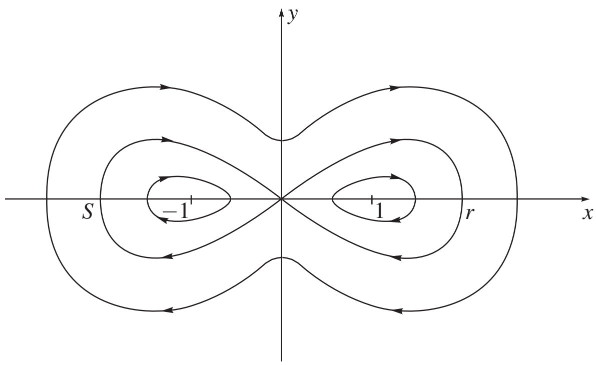


Figure 1. The structure of trajectories of system (3.7).

The origin is a saddle rest point of system (3.7); the rest points $(-1, 0)$ and $(1, 0)$ are centers; all other trajectories, different from the mentioned rest points and from the trajectories of the points $s = (-\sqrt{2}, 0)$ and $r = (\sqrt{2}, 0)$, are closed.

Denote by ϕ the flow generated by system (3.7).

Points of closed trajectories are nonwandering points of the flow ϕ. Since the union of closed trajectories is dense in the plane and the nonwandering set is closed (this fact is proven by the same reasoning as in the case of cascades), every point of the plane is a nonwandering point of the flow ϕ.

Clearly, if p is a rest point or a point belonging to a closed trajectory, then the ω-limit and α-limit points of the trajectory of p belong to the trajectory of p. For the points r and s, the set of ω-limit and α-limit points is the rest point $(0, 0)$.

Thus, the nonwandering points r and s of the flow ϕ do not belong to the set of ω-limit and α-limit points of trajectories of the flow.

In addition, let us note that, moving along the trajectory of the flow ϕ, the point r leaves its small neighborhood when time increases or decreases and does not return to this neighborhood. This point is nonwandering since any its neighborhood contains points of closed trajectories (and these points return to the neighborhood infinitely many times).

3.4 Local equivalence

Often, a researcher considers small perturbations of the system and studies the problem of preservation of the topological structure of the set of trajectories not in the whole phase space but on some its subsets (usually, on neighborhoods of invariant sets). In this case, the notions of topological conjugacy and equivalence are replaced by their "local" variants.

Let us give one of the possible definitions.

Let $f : M \to M$ and $g : N \to N$ be homeomorphisms of topological spaces. Assume that I and J are invariant sets of f and g, respectively.

Let X be a subset of M. Fix a point $x \in X$ and define the set

$$K(x, X) = \{n > 0 : f^k(x) \in X,\ 0 < k \le n\} \cup \{n < 0 : f^k(x) \in X,\ n \le k < 0\}.$$

Thus, the set

$$\{f^k(x) : k \in K(x, X)\}$$

is a "component" of the intersection of the trajectory $O(x, f)$ with the set X.

We say that the invariant sets I and J are *locally topologically conjugate* if there exist neighborhoods U and V of I and J, respectively, and a homeomorphism h mapping U to V and having the following properties:

(LC1) $h(I) = J$;

(LC2) $h(f^k(x)) = g^k(h(x))$ for $x \in U$ and $k \in K(x, U)$.

Chapter 4
Hyperbolic fixed point

4.1 Hyperbolic linear mapping

The notion of hyperbolicity is one of the basic notions in the theory of structural stability of dynamical systems.

We first consider the simplest system with hyperbolic behavior, a hyperbolic linear mapping of the Euclidean space $\mathbb{R}^n$.

Let L be a nondegenerate linear mapping of the space $\mathbb{R}^n$. It is well known that if we fix a basis of the space, then we assign to the mapping L an $n \times n$ matrix A such that the mapping L is given by the formula $y = Ax, x \in \mathbb{R}^n$.

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of the matrix A. The matrix A is called *hyperbolic* if

$$|\lambda_i| \neq 1, \quad i = 1, \dots, n. \tag{4.1}$$

If we fix another basis of the space $\mathbb{R}^n$ (i.e., we perform a nonsingular change of variables $z = Sx$), then , in this new basis, we assign to the mapping L the matrix $A' = S^{-1}AS$ that is similar to A. Since the sets of eigenvalues of similar matrices are the same, the matrices A' and A are hyperbolic or nonhyperbolic simultaneously.

A linear mapping L is called *hyperbolic* if the eigenvalues of its matrix in some (and then in any) basis satisfy inequalities (4.1).

Lemma 4.1. *If L is a hyperbolic linear mapping of the space $\mathbb{R}^n$, then there exists a basis of the space in which the matrix A of the mapping L is block-diagonal, $A = \operatorname{diag}(B, C)$, and the following inequalities hold: $\|B\| < 1$ and $\|C^{-1}\| < 1$.*

Proof. Fix a basis of the space $\mathbb{R}^n$; let A' be the matrix of the mapping L in this basis.

We consider the general case assuming that the matrix A' has both eigenvalues λ_i with $|\lambda_i| < 1$ and eigenvalues λ_i with $|\lambda_i| > 1$.

Thus, we assume that there is an index $m \in (0, n)$ such that

$$|\lambda_1| \leq \cdots \leq |\lambda_m| < 1 < |\lambda_{m+1}| \leq \cdots \leq |\lambda_n|. \tag{4.2}$$

Find a number $b \in (0, 1)$ such that $|\lambda_i| \leq b$ for $i = 1, \dots, m$ and $b|\lambda_i| > 1$ for $i = m + 1, \dots, n$. Fix a number $a > 0$ such that $b + a < 1$.

By the Jordan Canonical Form Theorem (see, for example, [14]), there exists a real-valued nonsingular matrix S such that $A := S^{-1}A'S = G + F$ and the structure of the matrices G and F is described as follows.

The matrix G is a block-diagonal matrix of the form

$$G = \operatorname{diag}(\Lambda_1, \dots, \Lambda_l, \Lambda_{l+1}, \dots \Lambda_n),$$

where any of the submatrices Λ_j is either a scalar value (1×1 matrix) equal a real eigenvalue of the matrix A' or a 2×2 matrix

$$\Lambda_j = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

corresponding to a pair of complex conjugate eigenvalues $\lambda = \alpha \pm i\beta$ of the matrix A'.

In addition, the eigenvalues of the matrices $\Lambda_1, \dots, \Lambda_l$ satisfy the inequalities $|\lambda| < 1$, and the eigenvalues of the matrices $\Lambda_{l+1}, \dots \Lambda_n$ satisfy the inequalities $|\lambda| > 1$ (thus, the matrix

$$A_1 = \operatorname{diag}(\Lambda_1, \dots, \Lambda_l)$$

has size $m \times m$, and the matrix

$$A_2 = \operatorname{diag}(\Lambda_{l+1}, \dots, \Lambda_n)$$

has size $(n - m) \times (n - m)$).

The matrix F is a block-diagonal matrix of the form $F = \operatorname{diag}(F_1, F_2)$, where the submatrices F_1 and F_2 have size $m \times m$ and $(n-m) \times (n-m)$, respectively, and the norm of the matrix F satisfies the inequality

$$\|F\| < a. \tag{4.3}$$

Comment. In the standard statement of the Jordan Canonical Form Theorem, entries of the matrix F are zeros and units. Analyzing the proof of this theorem, it is easy to understand that for any $\varepsilon > 0$ one can find the corresponding matrix S for which the nonzero entries of the matrix F equal ε instead of 1. Since the size of F is $n \times n$, we can take ε small enough to guarantee inequality (4.3) (see estimate (2.4) in Section 2.3).

Thus, the matrix A has the desired block-diagonal form with the matrices $B = A_1 + F_1$ and $C = A_2 + F_2$ of size $m \times m$ and $(n - m) \times (n - m)$, respectively.

Let $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$, $y = Bx$, and $y' = A_1 x$. Take an index $j \in \{1, \dots, m\}$; then either

$$y'_j = \lambda_k x_j$$

(if the corresponding eigenvalue λ_k is real) or

$$y'_{j-1} = \alpha_k x_{j-1} - \beta_k x_j \quad \text{and} \quad y'_j = \beta_k x_{j-1} + \alpha_k x_j$$

(if the corresponding eigenvalues are $\lambda_k = \alpha_k \pm i\beta_k$).

In the first case,

$$(y_j')^2 = \lambda_k^2 x_j^2 \le b^2 x_j^2;$$

in the second case,

$$(y_{j-1}')^2 + (y_j')^2 = (\alpha_k^2 + \beta_k^2)(x_{j-1}^2 + x_j^2) \le b^2(x_{j-1}^2 + x_j^2).$$

Adding these inequalities, we get the inequality

$$(y_1')^2 + \cdots + (y_m')^2 \le b^2(x_1^2 + \cdots + x_m^2),$$

which implies that $|y'| \le b|x|$. Thus, $\|A_1\| \le b$, and

$$\|B\| = \|A_1 + F_1\| < a + b < 1.$$

Now let $x = (x_{m+1}, \ldots, x_n)$, $y = (y_{m+1}, \ldots, y_n)$, $y = Cx$, and $x' = A_2^{-1}y$.

The matrix A_2^{-1} is block-diagonal, and its blocks are either scalars $1/\lambda_k$ (if the corresponding eigenvalue λ_k is real) or

$$\frac{1}{\alpha_k^2 + \beta_k^2}\begin{pmatrix} \alpha_k & \beta_k \\ -\beta_k & \alpha_k \end{pmatrix}$$

(if the corresponding eigenvalue is $\lambda_k = \alpha_k \pm i\beta_k$). In the second case, the inequalities

$$\alpha_k^2 + \beta_k^2 = |\lambda_k|^2 \ge b^{-2}$$

hold. Thus, the same reasoning as in the first case shows that $|x'| \le b|y|$.

Applying the matrix A_2^{-1} to the equality $y = A_2x + F_2x$, we get the equality $x' = A_2^{-1}y = x + A_2^{-1}F_2x$. Since

$$b|y| \ge |x'| \ge |x| - a|x|$$

(we take into account that $\|A_2^{-1}\| < 1$), the inequality

$$|x| \le \frac{b}{1-a}|y|$$

holds.

This means that $\|C^{-1}\| \le b/(1-a)$. We note that $b/(1-a) < 1$ due to the inequality $a + b < 1$. Thus, $\|C^{-1}\| < 1$. This completes the proof. □

Since we study dynamical systems up to topological conjugacy, we may assume that, working with a dynamical system generated by a hyperbolic linear mapping L, we fix coordinates in which the matrix A of the mapping L has the properties described in Lemma 4.1.

Let us represent a vector $x \in \mathbb{R}^n$ in the form $x = (y, z)$ according to the block-diagonal structure of the matrix A. In this case, L maps a vector (y, z) to (By, Cz). It was assumed that L is an invertible mapping; thus, the matrix A is nondegenerate, and the equalities $L^k(y, z) = (B^k y, C^k z)$ hold for $k \in \mathbb{Z}$.

Consider the linear subspaces $\mathcal{S} = \{z = 0\}$ and $\mathcal{U} = \{y = 0\}$ of the space $\mathbb{R}^n$. If relations (4.2) hold, then $\dim\mathcal{S} = m$ and $\dim\mathcal{U} = n - m$. Clearly, both subspaces $\mathcal{S}$ and $\mathcal{U}$ are invariant under the mapping L.

Fix a number $\lambda \in (0, 1)$ such that $\|B\| \leq \lambda$ and $\|C^{-1}\| \leq \lambda$.

If $x = (y, 0) \in \mathcal{S}$, then $|L^k x| = |B^k y| \leq \lambda^k |y|, k \geq 0$, and $L^k x \to 0, k \to \infty$. If $x \in \mathcal{U}$, then $L^k x \to 0, k \to -\infty$.

Let $x = (0, z) \in \mathcal{U}$ and $Lx = (0, z')$. Since $z' = Cz$, $z = C^{-1} z'$; hence, $|z'| \geq \lambda^{-1}|z|$. Thus, if $x \in \mathcal{U}$, then $|L^k x| \geq \lambda^{-k}|x|, k \geq 0$, and $|L^k x| \to \infty$ as $k \to \infty$. Similarly, if $x \in \mathcal{S}$, then $|L^k x| \geq \lambda^k |x|, k \leq 0$, and $|L^k x| \to \infty$ as $k \to -\infty$.

Clearly, if $x \notin \mathcal{S} \cup \mathcal{U}$, then $|L^k x| \to \infty$ as $|k| \to \infty$.

If $m \neq 0, n$, then the fixed point $x = 0$ of the mapping L is called a *saddle point*. If $m = n$, i.e./ $|\lambda_i| < 1, i = 1, \dots, n$, then the fixed point $x = 0$ is called *attractive* (in this case, $\mathcal{S} = \mathbb{R}^n$); if $m = 0$, i.e., $|\lambda_i| > 1, i = 1, \dots, n$, the fixed point $x = 0$ is called *repelling* (in this case, $\mathcal{U} = R^n$).

4.2 The Grobman–Hartman theorem

Let p be a fixed point of a diffeomorphism f of the space $\mathbb{R}^n$ (it is clear from the reasoning below that, without loss of generality, we may work either with a diffeomorphism f that maps the space $\mathbb{R}^n$ to itself or with a diffeomorphism f that maps a neighborhood of the point p to a domain in $\mathbb{R}^n$; thus, the statement which we prove below is applicable to a neighborhood of a point of a smooth manifold as well). We assume that f is a diffeomorphism of class C^1.

The fixed point p is called *hyperbolic* if the Jacobi matrix $Df(p)$ is hyperbolic.

We show that a hyperbolic fixed point p of a diffeomorphism f is locally topologically conjugate with the fixed point $x = 0$ of the linear mapping $L : x \to Df(p)x$. This statement was proven independently by D. M. Grobman and P. Hartman; it plays a fundamental role in the global qualitative theory of dynamical systems. In the proof, Grobman and Hartman used principially different approaches. Grobman constructed the topological conjugacy analyzing the geometrical pattern of trajectories; Hartman solved a functional equation for the conjugating homeomorphism.

Our proof below follows the idea of Hartman (the later development of the theory showed that Hartman's approach is applicable to a wide class of problems; we apply it in this book in the proof of structural stability of an Anosov diffeomorphism).

Theorem 4.1. A hyperbolic fixed point p of a diffeomorphism f is locally topologically conjugate with the fixed point $x = 0$ of the linear mapping $L : x \to Df(p)x$.

Without loss of generality, we assume in the proof of Theorem 4.1 that $p = 0$ and f is a diffeomorphism of a neighborhood U of the origin of $\mathbb{R}^n$ to a domain containing the origin. In this case, we can write f as follows:

$$f(x) = Ax + F(x), \tag{4.4}$$

where the Jacobi matrix A of the diffeomorphism f at the origin has the properties described in Lemma 4.1, and the nonlinearity F vanishes at the origin together with its Jacobi matrix.

First we extend the diffeomorphism f to the whole space $\mathbb{R}^n$; in fact, we construct a diffeomorphism G of the space $\mathbb{R}^n$ to itself that coincides with f in a neighborhood W of the origin; after that, we prove that G is topologically conjugate with the linear mapping L in the whole space. Clearly, if h topologically conjugates G and L in $\mathbb{R}^n$, then the restriction $h|_W$ is the desired local topological conjugacy of the zero fixed points of the diffeomorphism f and the linear mapping L.

The matrix A is the Jacobi matrix of the diffeomorphism f at $x = 0$; hence, A is invertible. It follows that there exists a number $\delta > 0$ such that $|Ax| \geq \delta|x|$ for all x. Fix a number $\varepsilon > 0$ such that

$$\begin{aligned} &(1) \quad \varepsilon < \delta; \\ &(2) \quad c := \max(\|B\|, \|C^{-1}\|) + \varepsilon + \varepsilon\|C^{-1}\| < 1. \end{aligned} \tag{4.5}$$

We need the following auxiliary statement.

Lemma 4.2. *Consider a scalar-valued function g of class $C^r(V, \mathbb{R})$, where V is a neighborhood of the origin in $\mathbb{R}^n$. Assume that*

$$g(0) = 0, \quad Dg(0) = 0. \tag{4.6}$$

Then for any $\varepsilon > 0$ there exists a neighborhood V_0 of the origin, $V_0 \subset V$, and a function $\widetilde{g} \in C^r(\mathbb{R}^n, \mathbb{R})$ such that

$$\begin{aligned} &(1) \quad \widetilde{g}(x) = g(x) \quad \text{for any } x \in V_0, \\ &(2) \quad \left|\frac{\partial \widetilde{g}(x)}{\partial x_i}\right| < \varepsilon, \quad i = 1, \ldots, n, \quad \text{for any } x \in \mathbb{R}^n. \end{aligned} \tag{4.7}$$

Proof. Denote by $\eta(t)$ a scalar-valued function of class $C^\infty(\mathbb{R}, \mathbb{R})$ having the following properties:

(1) $\eta(t) = 1$ for $t \leq 1$;

(2) $\eta(t) = 0$ for $t \geq 2$;

(3) $0 < \eta(t) < 1$ for $1 < t < 2$;

(4) $-2 \le \eta'(t) \le 0$ for $t \in R$.

The existence of such a function is a standard fact from the calculus.

Take a number $\delta > 0$ such that the ball $x^2 < 2\delta$ is a subset of V (recall that $x^2 = |x|^2 = x_1^2 + \cdots + x_n^2$). Consider a function $g_\delta(x)$ defined in $\mathbb{R}^n$ by the following formulas:

$$g_\delta(x) = g(x)\eta(x^2/\delta) \quad \text{for } x^2 < 2\delta,$$
$$g_\delta(x) = 0 \quad \text{for } x^2 \ge 2\delta.$$

Clearly, $g_\delta \in C^r(\mathbb{R}^n, \mathbb{R})$ and $g_\delta(x) = g(x)$ for $x^2 < \delta$. Fix $\varepsilon > 0$; we show that if δ is small enough, then

$$\left| \frac{\partial g_\delta}{\partial x_i}(x) \right| \le \varepsilon, \quad i = 1, \ldots, n, \tag{4.8}$$

in $\mathbb{R}^n$. The corresponding function g_δ is the desired $\widetilde{g}$. If $x^2 \ge 2\delta$, then the partial derivative $\partial g_\delta/\partial x$ vanishes; hence, it is enough to prove estimate (4.8) in the domain $x^2 < 2\delta$. Let us write

$$\frac{\partial g_\delta}{\partial x_i}(x) = \frac{\partial g}{\partial x_i}(x)\eta\left(\frac{x^2}{\delta}\right) + g(x)\eta'\left(\frac{x^2}{\delta}\right)\frac{2x_i}{\delta}.$$

There exists $\delta_1 > 0$ such that if $x^2 < 2\delta_1$, then

$$\left| \frac{\partial g}{\partial x_i}(x) \right| < \frac{\varepsilon}{2}.$$

Take a number $\delta \le \delta_1$. If $x^2 < 2\delta$, then the inequality $|\eta| \le 1$ implies that

$$\left| \frac{\partial g}{\partial x}(x)\eta\left(\frac{x^2}{\delta}\right) \right| < \frac{\varepsilon}{2}.$$

To estimate the second summand, we multiply and divide it by $|x|$ (assuming that $x \ne 0$):

$$\left| \frac{g(x)\eta'(x^2/\delta)2x_i}{\delta} \right| = \frac{|g(x)|}{|x|} \left| \eta'\left(\frac{x^2}{\delta}\right) \right| \frac{2|x_i||x|}{\delta}.$$

By condition (4.6), there exists $\delta_2 > 0$ such that

$$\frac{|g(x)|}{|x|} < \frac{\varepsilon}{16}$$

if $0 < x^2 < 2\delta_2$. Clearly, $|x_i| \le |x|$; hence,

$$2|x_i||x| \le 2x^2.$$

Taking into account that $|\eta'| \le 2$, we conclude that if $0 < x^2 < 2\delta$, where $\delta \le \delta_2$, then

$$\frac{|g(x)|}{|x|} \left|\eta'\left(\frac{x^2}{\delta}\right)\right| \frac{2|x_i||x|}{\delta} \le \frac{\varepsilon}{2}.$$

Thus, δ has the desired property. The lemma is proven. □

Let us apply Lemma 4.2 to construct a mapping F' of class C^1 in $\mathbb{R}^n$ such that

$$F' = F \quad \text{in a neighborhood } V \text{ of the origin} \tag{4.9}$$

and

$$\|DF'(x)\| < \varepsilon, \quad x \in \mathbb{R}^n. \tag{4.10}$$

Consider the mapping

$$G(x) = Ax + F'(x).$$

We show that G is a homeomorphism of $\mathbb{R}^n$ onto itself. One can show that G is a diffeomorphism; we do not prove this latter statement due to the following reasons. First, we do not need this statement in what follows; second, we apply the same reasoning once more to show that the mapping h is a homeomorphism (and h is not a diffeomorphism!).

We show that the statement which we need (G is a homeomorphism of $\mathbb{R}^n$ onto itself) is a corollary of the following three properties of G:

(G1) G is continuous;

(G2) G is injective;

(G3) $|G(x)| \to \infty$ as $|x| \to \infty$.

Property (G1) follows from the definition of G. Let us establish property (G2). Indeed, if $G(x_1) = G(x_2)$, then

$$Ax_1 + F'(x_1) = Ax_2 + F'(x_2), \quad \text{or} \quad A(x_1 - x_2) = F'(x_2) - F'(x_1).$$

By the choice of δ, $|A(x_1 - x_2)| \ge \delta|x_1 - x_2|$. By the Lagrange theorem,

$$|F'(x_1) - F'(x_2)| \le \max_{x \in \mathbb{R}^n} \|DF'(x)\||x_1 - x_2| \le \varepsilon|x_1 - x_2|;$$

hence, $\delta|x_1 - x_2| \leq \varepsilon|x_1 - x_2|$, and it follows from condition (4.5) that $x_1 = x_2$. Property (G3) is an obvious corollary of the invertibility of the matrix A and the fact that the function F' vanishes outside a ball of finite radius.

Since G is continuous and injective, G maps homeomorphically every compact subset of $\mathbb{R}^n$ onto its image.

Now the Brower theorem implies that G maps bounded open sets to open sets; it follows that the image $G(\mathbb{R}^n)$ is open. Let us show that $G(\mathbb{R}^n)$ is closed. Assume that $G(x_k) \to y$. The sequence x_k is bounded; otherwise, the existence of an unbounded subsequence x_{k_l} would imply, by property (G3), that $|G(x_{k_l})| \to \infty$.

Since the sequence x_k is bounded, it contains a convergent subsequence $x_{k_m} \to x$. Since the mapping G is continuous, $G(x_{k_m}) \to G(x)$; by the uniqueness of the limit, $y = G(x) \in G(\mathbb{R}^n)$. Hence, the set $G(\mathbb{R}^n)$ is closed, and it follows that $G(\mathbb{R}^n) = \mathbb{R}^n$ (of course, we take into account that $G(\mathbb{R}^n)$ is nonempty since $G(0) = 0$).

Now we prove that the homeomorphism G and the linear mapping L are topologically conjugate in the whole space.

We find a homeomorphism h conjugating G and L in the form

$$h(x) = x + g(x), \tag{4.11}$$

where g is a continuous mapping $\mathbb{R}^n \to \mathbb{R}^n$ that satisfies the following additional condition:

$$|g(x)| \to 0 \tag{4.12}$$

as $|x| \to \infty$. Denote by Id the identity mapping of the space $\mathbb{R}^n$. Let us write equality (4.11) without arguments:

$$h = \mathrm{Id} + g.$$

The homeomorphism h conjugates G and L if and only if

$$G \circ h = h \circ L, \tag{4.13}$$

or

$$G \circ (\mathrm{Id} + g) = (\mathrm{Id} + g) \circ L. \tag{4.14}$$

We solve equation (4.14) using the fixed point approach. Consider the space

$$H = \{g \in C(R^n, R^n) : g(0) = 0, |g(x)| \to 0 \text{ as } |x| \to \infty\}.$$

According to the decomposition $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$ (which corresponds to our representation of the matrix A in the block-diagonal form), we write elements $g \in H$

in the form $g = (g_1, g_2)$ (we also denote by f_1, f_2 the corresponding components of the vector-function F'). For $g, j \in H$ we consider the value

$$\varrho(g, j) = \sup_{x \in \mathbb{R}^n} |g_1(x) - j_1(x)| + \sup_{x \in \mathbb{R}^n} |g_2(x) - j_2(x)|.$$

Clearly, ϱ is a metric on H, and the space H with this metric is complete.

Let us substitute the expression $G(x) = Lx + F'(x)$ into equality (4.14):

$$(L + F') \circ (\mathrm{Id} + g) = (\mathrm{Id} + g) \circ L,$$

or

$$L + L \circ g + F' \circ (\mathrm{Id} + g) = L + g \circ L. \tag{4.15}$$

Since the operator L is linear, $L \circ (\mathrm{Id} + g) = L + L \circ g$; at the same time, $F' \circ (\mathrm{Id} + g) \neq F' \circ \mathrm{Id} + F' \circ g$. Cancelling L on the right and left in (4.15), we get the equation

$$L \circ g + F' \circ (\mathrm{Id} + g) = g \circ L. \tag{4.16}$$

Let us write equation (4.16) calculating all its terms at a point x:

$$Ag(x)) + F'(x + g(x)) = g(Ax).$$

The operator L is invertible; hence, equation (4.16) is equivalent to the following equation:

$$g = L^{-1} \circ (g \circ L - F' \circ (\mathrm{Id} + g)). \tag{4.17}$$

Taking into account our decomposition of the space $\mathbb{R}^n$ and the equality $A^{-1} = \mathrm{diag}(B^{-1}, C^{-1})$, we can write equation (4.17) componentwise as follows:

$$g_1 = B^{-1} \circ [g_1 \circ L - f_1 \circ (\mathrm{Id} + g)], \tag{4.18}$$

$$g_2 = C^{-1} \circ [g_2 \circ L - f_2 \circ (\mathrm{Id} + g)] \tag{4.19}$$

(note that the arguments of f_1 and f_2 contain the expression $I + g$; thus, at a point x, the arguments equal $x + g(x)$). The linear operator C^{-1} is contracting; the operator B^{-1} does not have this property. Let us transform equation (4.18) as follows: we apply the operator B to both sides, apply the operator L^{-1} from the right, and get the equality

$$g_1 = B \circ g_1 \circ L^{-1} + f_1 \circ (\mathrm{Id} + g) \circ L^{-1}.$$

Consider an operator T defined as follows: A function $g \in H$ is mapped to

$$T(g) = \begin{pmatrix} B \circ g_1 \circ L^{-1} + f_1 \circ (\mathrm{Id} + g) \circ L^{-1} \\ C^{-1} \circ [g_2 \circ L - f_2 \circ (\mathrm{Id} + g)] \end{pmatrix}.$$

Clearly, the function $T(g)$ is continuous in $\mathbb{R}^n$ and $T(g)(0) = 0$. If $|x| \to \infty$, then $|(\mathrm{Id} + g)(x)| = |x + g(x)| \to \infty$ since $|g(x)| \to 0$. By construction, $F'(x) = 0$ for $|x|$ large enough; hence,

$$|f_2 \circ (\mathrm{Id} + g)(x)| \to 0$$

as $|x| \to \infty$. If $|x| \to \infty$, then $|Lx| \to \infty$, hence $|g_2(Lx)| \to 0$ and $|C^{-1}g_2(Lx)| \to 0$; thus, the norm of the second component of $T(g)$ tends to 0 as $|x| \to \infty$. We apply a similar reasoning to the first component to show that $|T(g)(x)| \to 0$ as $|x| \to \infty$. Thus, $T(g) \in H$, i.e., T is a mapping $H \to H$. Let us show that this mapping is contracting. We take $g, j \in H$ and estimate

$$\begin{aligned} &\varrho(T(g), T(j)) \\ &\quad = \sup_{x\in\mathbb{R}^n} \{|Bg_1(L^{-1}x) - Bj_1(L^{-1}x) \\ &\qquad\qquad + f_1(L^{-1}x + g(L^{-1}x)) - f_1(L^{-1}x + j(L^{-1}x))|\} \\ &\qquad + \sup_{x\in\mathbb{R}^n} \{|C^{-1}[g_2(Lx) - j_2(Lx) + f_2(x + j(x)) - f_2(x + g(x))]|\}. \end{aligned} \tag{4.20}$$

Note that

$$\sup_{x\in\mathbb{R}^n} |B(g_1(L^{-1}x) - j_1(L^{-1}x))| = \sup_{x\in\mathbb{R}^n} |B(g_1(x) - j_1(x))|$$

since the mapping L^{-1} is invertible, and if a point x runs over the whole space $\mathbb{R}^n$, then $L^{-1}x$ runs over the whole space $\mathbb{R}^n$ as well, and vice versa. Hence, we can estimate the first term in (4.20) by the expression

$$\begin{aligned} &\|B\| \sup_{x\in\mathbb{R}^n} |g_1(x) - j_1(x)| + \sup_{x\in\mathbb{R}^n} \left\| \frac{\partial f_1}{\partial x} \right\| |g(x) - j(x)| \\ &\qquad \le \|B\| \sup_{x\in\mathbb{R}^n} |g_1(x) - j_1(x)| + \varepsilon \sup_{x\in\mathbb{R}^n} |g(x) - j(x)|. \end{aligned}$$

In the above estimation, we apply the Lagrange theorem to the expression

$$f_1(L^{-1}x + g(L^{-1}x)) - f_1(L^{-1}x + j(L^{-1}x))$$

and take inequality (4.10) into account. Similarly, the second term in (4.20) does not exceed

$$\|C^{-1}\| \sup_{x\in\mathbb{R}^n} |g_2(x) - j_2(x)| + \varepsilon\|C^{-1}\| \sup_{x\in\mathbb{R}^n} |g(x) - j(x)|.$$

The obvious inequality

$$|g(x) - j(x)| \le |g_1(x) - j_1(x)| + |g_2(x) - j_2(x)|$$

implies the estimate

$$\varrho(T(g), T(j)) \le [\max(\|B\|, \|C^{-1}\|) + \varepsilon(1 + \|C^{-1}\|)]\varrho(g, j).$$

It follows from (4.5) that the operator T is contracting. Hence, in the space H there exists (a unique) fixed point g^* of the operator T. The definition of the operator T implies that $g^* = (g_1^*, g_2^*)$ is a solution of system (4.18), (4.19); hence, g^* satisfies equation (4.14). The mapping $h(x) = x + g^*(x)$ satisfies relation (4.13); thus, to complete the proof of Theorem 4.1 it is enough to show that h is a homeomorphism. Clearly, the mapping h is continuous. Let us show that h is injective. To get a contradiction, let us assume that there exist points $x_1, x_2 \in \mathbb{R}^n$ such that $x_1 \neq x_2$ and $h(x_1) = h(x_2)$. Clearly, equality (4.13) implies that

$$G^k \circ h = h \circ L^k, \quad k \in \mathbb{Z}.$$

The equalities

$$G^k(h(x_1)) = G^k(h(x_2)), \quad k \in \mathbb{Z},$$

imply that

$$h(L^k x_1) = h(L^k x_2), \quad k \in \mathbb{Z}. \tag{4.21}$$

By construction, $h(x) = x + g^*(x)$, and it follows from (4.21) that

$$L^k x_1 + g^*(L^k x_1) = L^k x_2 + g^*(L^k x_2), \quad k \in \mathbb{Z},$$

or

$$L^k(x_1 - x_2) = g^*(L^k x_2) - g^*(L^k x_1) \quad k \in \mathbb{Z}. \tag{4.22}$$

Let $x_1 - x_2 = v = (v_1, v_2)$. By our assumption, $v \neq 0$; for definiteness, let $v_1 \neq 0$. It was shown in Section 4.1 that

$$|B^k v_1| \to \infty \quad \text{as } k \to -\infty;$$

hence, the norm of $|L^k(x_1 - x_2)|$ is unbounded. At the same time, the definition of the space H implies that the right-hand side of (4.22) is bounded. The contradiction obtained proves that the mapping h is injective.

Finally, we note that if $|x| \to \infty$, then $|g^*(x)| \to 0$; hence, $|h(x)| \to \infty$ as $|x| \to \infty$.

Now the same reasoning as that applied to the mapping G after Lemma 4.2 shows that h is a homeomorphism of the space $\mathbb{R}^n$ to itself. This completes the proof of the Grobman–Hartman theorem. □

4.3 Neighborhood of a hyperbolic fixed point

The Grobman–Hartman theorem and the statements of Section 4.1 allow us to describe the behavior of trajectories of a diffeomorphism in a neighborhood of a hyperbolic fixed point.

Let p be a hyperbolic fixed point of a diffeomorphism f (we treat both cases where the phase space is a smooth manifold or the Euclidean space).

By the Grobman–Hartman theorem, there exists a neighborhood U of the point p, a neighborhood V of the origin in $\mathbb{R}^n$, and a homeomorphism $h : V \to U$ that maps the origin to p and intersections of trajectories of the linear mapping $Ly = Df(p)y$ with V to intersections of trajectories of the diffeomorphism f with U.

Let us state several corollaries of statements of Section 4.1.

If 0 is an attracting fixed point of L (i.e., the eigenvalues λ_j of the matrix of the operator L satisfy the inequalities $|\lambda_j| < 1$, $j = 1, \dots, n$) and $x \neq p$ is an arbitrary point of the neighborhood U close enough to p, then $f^k(x) \to p$ as $k \to \infty$, and the trajectory $f^k(x)$ leaves the neighborhood U as k decreases (in this case, the fixed point p of the diffeomorphism f is called *attracting*).

If 0 is a repelling fixed point of L (i.e., the eigenvalues λ_j of the matrix of the operator L satisfy the inequalities $|\lambda_j| > 1$, $j = 1, \dots, n$) and $x \neq p$ is an arbitrary point of the neighborhood U close enough to p, then $f^k(x) \to p$ as $k \to -\infty$, and the trajectory $f^k(x)$ leaves the neighborhood U as k increases (in this case, the fixed point p of the diffeomorphism f is called *repelling*).

Let us treat in more detail the case of a saddle fixed point of the mapping L (i.e., the case where there exists an index $m \in (0, n)$ such that eigenvalues λ_j of the matrix of the operator L satisfy inequalities (4.2); in this case, the fixed point p of the diffeomorphism f is called *saddle*).

Let $\mathcal{S}$ and $\mathcal{U}$ be the linear subspaces defined in Section 4.1 in the case of a saddle fixed point of L.

Without loss of generality, we may assume that the neighborhood V is a ball centered at the origin, and the matrix A of L has the properties described in Lemma 4.1 (then $|Av| \le |v|$ for $v \in \mathcal{S}$ and $|A^{-1}v| \le |v|$ for $v \in \mathcal{U}$).

Consider the set $W^s_U(p) := h(\mathcal{S} \cap V)$; clearly, this set has the following properties: If $x \in W^s_U(p)$, then $f^k(x) \to p$ as $k \to \infty$, and if $x \in U \setminus W^s_U(p)$, then the trajectory $f^k(x)$ leaves the neighborhood U as k decreases. The set $W^s_U(p)$ is called the *local stable manifold* of the saddle fixed point p of the diffeomorphism f in the neighborhood U.

It follows from the Grobman–Hartman theorem that the set $W^s_U(p)$ is a topological m-dimensional disk (i.e., a homeomorphic image of the unit ball of the m-dimensional Euclidean space).

Consider the set $W^u_U(p) := h(\mathcal{U} \cap V)$; clearly, this set has the following properties: If $x \in W^u_U(p)$, then $f^k(x) \to p$ as $k \to -\infty$, and if $x \in U \setminus W^u_U(p)$, then the trajectory $f^k(x)$ leaves the neighborhood U as k increases. The set $W^u_U(p)$ is called the *local unstable manifold* of the saddle fixed point p of the diffeomorphism f in the neighborhood U.

It follows from the Grobman–Hartman theorem that the set $W_U^u(p)$ is a topological $(n-m)$-dimensional disk (i.e., a homeomorphic image of the unit ball of the $(n-m)$-dimensional Euclidean space).

It is also clear that if

$$x \in U \setminus (W_U^s(p) \cup W_U^u(p)),$$

then the trajectory $f^k(x)$ leaves the neighborhood U as $|k|$ increases.

In the next subsection, we show that, in fact, the sets $W_U^s(p)$ and $W_U^u(p)$ are smooth disks.

It follows from the above-mentioned properties of the sets $W_U^s(p)$ and $W_U^u(p)$ that these sets are not invariant. Let us define the main invariant sets related to a hyperbolic saddle fixed point p, its stable and unstable manifolds (these sets and their analogs play the main role in the theory of structural stability).

Let f be a diffeomorphism of a Euclidean space or of a smooth manifold (in both cases, we denote the phase space by M) and let p be a hyperbolic fixed point of f.

Define the sets

$$W^s(p) = \{x \in M : f^k(x) \to p,\ k \to \infty\}$$

and

$$W^u(p) = \{x \in M : f^k(x) \to p,\ k \to -\infty\}.$$

The sets $W^s(p)$ and $W^u(p)$ are called the *stable* and *unstable manifolds* of the fixed point p, respectively.

Let us fix a neighborhood U of the point p in M in which f has the above-described behavior.

Lemma 4.3. (1) *The sets $W^s(p)$ and $W^u(p)$ are invariant;*

(2) *$x \in W^s(p)$ if and only if $O(x, f) \cap W_U^s(p) \neq \emptyset$;*

(3) *$x \in W^u(p)$ if and only if $O(x, f) \cap W_U^u(p) \neq \emptyset$.*

Proof. (1) Let $y = f^m(x)$, where $m \in \mathbb{Z}$. Since

$$f^k(y) = f^k(f^m(x)) = f^m(f^k(x)) \to f^m(p) = p, \quad k \to \infty,$$

the inclusion $y \in W^s(p)$ holds. A similar reasoning shows that the set $W^u(p)$ is invariant.

(2) If $x \in W^s(p)$, then there exists an index m such that $f^k(x) \in U$ for $k \geq m$. The inclusion $f^m(x) \in W_U^s(p)$ follows from the above-mentioned properties of the neighborhood U.

If $O(x, f) \cap W^s_U(p)$, then there exists an index m such that $y := f^m(x) \in W^s_U(p)$. Hence,

$$f^k(y) \to p, \quad k \to \infty.$$

Now our claim follows from the inclusion $y \in W^s(p)$ and from the invariance of the set $W^s(p)$.

A similar reasoning proves statement (3). □

For the linear mapping L, the stable manifold of the origin coincides with the m-dimensional linear subspace $\mathcal{S}$ introduced in Section 4.1 (if the eigenvalues λ_j of the matrix of the operator L satisfy inequalities (4.2)).

Let us show that, from the topological point of view, the stable manifold of the hyperbolic fixed point p has the same structure.

Fix a positive number a and consider the m-dimensional ball

$$D_a = \{|x| \le a\} \cap \mathcal{S}.$$

Let $c := \|B\| < 1$ (where the matrix B is given by Lemma 4.1). Since $|Lx| \le c|x|$ for $x \in \mathcal{S}$, the following inclusions hold:

$$L(D_a) \subset \operatorname{Int} D_a \subset D_a \tag{4.23}$$

(here $\operatorname{Int} D_a$ is the interior in the subspace $\mathcal{S}$). Consider the sets

$$G_k = L^k(D_a) \setminus L^{k+1}(D_a), \quad k \ge 0. \tag{4.24}$$

Lemma 4.4. *For any nonzero $x \in \mathcal{S}$, the intersection $O(x, L) \cap G_0$ is nonempty and consists of a single point.*

Proof. First we show that

$$D_a \setminus \{0\} = \bigcup_{k \ge 0} G_k. \tag{4.25}$$

Inclusions (4.23) imply that any set G_k is a subset of $D_a \setminus \{0\}$; hence, the set on the right in (4.25) is a subset of the set on the left.

To prove the inverse inclusion, consider a point $x \in D_a \setminus \{0\}$.

If $x \in G_0$, we have nothing to prove. If $x \notin G_0$, then equality (4.24) with $k = 0$ implies that $x \in L(D_a)$; if $x \notin G_1$ in this case, then equality (4.24) with $k = 1$ implies that $x \in L^2(D_a)$, and so on.

Thus, if $x \notin G_0 \cup G_1 \cup \dots \cup G_m$, then $x \in L^{m+1}(D_a)$.

Now equality (4.25) is a corollary of the following fact: If m is so large that $c^m a < |x|$, then the inclusion $x \in L^m(D_a)$ is impossible since

$$L^m(D_a) \subset \{|x| \le c^m a\} \cap \mathcal{S}.$$

For any nonzero $x \in \mathcal{S}$, its trajectory $O(x, L)$ intersects the set $D_a \setminus \{0\}$. Hence, there exist indices m and k such that $L^m x \in G_k$. Since $G_k = L^k(G_0)$, the inclusion $L^{m-k} x \in G_0$ holds.

It follows from the definition of G_0 that if $x \in G_0$, then $Lx \notin G_0$; hence, the intersection of any trajectory of the system L with G_0 cannot contain more than one point. □

Take a number $a > 0$ such that

$$L^{-1}(D_a) \subset V \tag{4.26}$$

(here V is a neighborhood of the point $x = 0$ in which we apply the Grobman–Hartman theorem). Denote $D = h(D_a)$ and $F = h(G_0)$. The set F is called a *fundamental neighborhood* of the fixed point p in its stable manifold $W^s(p)$.

For any point $x \in W^s(p)$, its trajectory reaches the set D and does not leave D as time increases; it follows from Lemma 4.4 that such a trajectory intersects the set F; the equality

$$F = h(D_a \setminus L(D_a)) = D \setminus f(D)$$

implies that the intersection consists of a single point.

Lemma 4.5. *The stable manifold $W^s(p)$ is the image of the space $\mathcal{S}$ under a topological immersion.*

Proof. Take a number $a > 0$ such that inclusion (4.26) holds. Define a mapping of $\mathcal{S}$ into the stable manifold $W^s(p)$ as follows. By Lemma 4.4, for any point $x \in \mathcal{S}$ there exists a (unique) index $m(x)$ such that $L^{m(x)} x \in G_0$. Set

$$\alpha(x) = f^{-m(x)}(h(L^{m(x)} x))$$

and $\alpha(0) = p$.

The mapping α is surjective. Indeed, if $y \in W^s(p) \setminus \{p\}$, then there exists an index m such that $f^m(y) \in F$; then $y = \alpha(L^{-m} h^{-1}(y))$. By definition, $p = \alpha(0)$.

To prove that the mapping α is injective, consider two different points y_1 and y_2 of the manifold $W^s(p)$; let $y_i = \alpha_i(x_i)$, $i = 1, 2$. For definiteness, we consider the case where $y_1, y_2 \neq p$.

If $y_2 \in O(y_1, f)$, then $y_2 = f^n(y_1)$ for some $n \neq 0$. Let $y = O(y_1, f) \cap F$. If $y_1 = f^m(y)$, then $y_2 = f^{m+n}(y)$. Clearly, the points $x_1 = L^{-m} h^{-1}(y)$ and $x_2 = L^{-m-n} h^{-1}(y)$ are different.

If $y_2 \notin O(y_1, f)$, then the points $O(y_1, f) \cap F$ and $O(y_2, f) \cap F$ are different. Hence, the points x_1 and x_2 belong to different trajectories of L (and it follows that $x_1 \neq x_2$).

Finally, let us show that the mapping α is continuous.

If $x \in D_a \setminus \{0\}$ and $x' = O(x, L) \cap G_0$, then $x = L^m x'$ for some $m \geq 0$; in this case, the set $\{L^k x' : 0 \leq k \leq m\}$ is a subset of D_a (and, hence, a subset of V). Then $h(L^{-m}x) = f^{-m}(h(x))$, and we conclude that

$$\alpha(x) = f^m(h(L^{-m}x)) = h(x).$$

Since $p = \alpha(0)$, on the set D_a, the mapping α coincides with the conjugating homeomorphism h (hence, α is continuous on this set).

Now we consider a point $x \in \mathcal{S} \setminus D_a$ and a sequence $x_k \in \mathcal{S}$ such that $x_k \to x, k \to \infty$.

If $x' = L^{m(x)}x$ is an interior point of the set G_0 (with respect to the topology of the subspace $\mathcal{S}$), then $L^{m(x)}x_k \in G_0$ for k large enough; hence,

$$\alpha(x_k) = f^{-m(x)}(h(L^{m(x)}x_k)) \to \alpha(x), \quad k \to \infty.$$

If x' is not an interior point of the set G_0, then it follows from the definition of the set G_0 that the point x' has a neighborhood in $\mathcal{S}$, that belongs to the union $G_0 \cup L^{-1}(G_0)$. Hence, if k is large enough, then either $m(x_k) = m(x)$ or $m(x_k) = m(x) + 1$. In the latter case,

$$L^{m(x)}x_k, L^{m(x)+1}x_k \in V,$$

and the equality

$$h(L^{m(x_k)}x_k) = h(L^{m(x)+1}x_k) = f(h(L^{m(x)}x_k))$$

implies that

$$\alpha(x_k) = f^{-m(x)-1}(f(h(L^{m(x)}x_k))) \to \alpha(x), \quad k \to \infty.$$

Thus, we have established the continuity of the mapping α at the point x' in the second case.

We have shown that the mapping α is surjective, injective, and continuous. The same reasoning as that applied to the mapping G after Lemma 4.2 shows that α maps homeomorphically $\mathcal{S}$ onto $W^s(p)$. This completes the proof of Lemma 4.5. □

Remark. Let $x \in \mathcal{S}, x \neq 0$, and let $x' = Lx$. In this case, $m(x') = m(x) - 1$. Then

$$\begin{aligned}\alpha(Lx) = \alpha(x') &= f^{-m(x)+1}(h(L^{m(x)-1}x')) \\ &= f(f^{-m(x)}(h(L^{m(x)}x))) = f(\alpha(x)).\end{aligned}$$

Thus, the mapping α conjugates the mapping L on $\mathcal{S}$ and the diffeomorphism f on $W^s(p)$.

Remark. Let us note that a stable manifold is not necessarily a (topological) submanifold of the phase space, i.e., the topology induced by the topology of the phase space may not coincide with the topology induced by the topological immersion of the space $\mathcal{S}$ described above; as an example, consider the shift by time 1 along trajectories of the system of differential equations described in Example 3.1, Section 3.

4.4 The stable manifold theorem

In this section, we show that local stable and unstable manifolds of a hyperbolic fixed point are smooth disks. This statement (usually called the stable manifold theorem) was proven by A. M. Lyapunov (for analytic systems of differential equations; in this case, Lyapunov worked with special representations of solutions), by J. Hadamard (the Hadamard approach was based on the study of transformations of graphs; one can find a detailed description of Hadamard's proof in the book [8]), and by O. Perron (our proof of Theorem 4.2 below uses Perron's method).

Since the stable manifold theorem is a local statement, we consider a diffeomorphism f of class C^1 that can be represented in the form (4.4) in a neighborhood of the origin of the space $\mathbb{R}^n$. We assume that the Jacobi matrix A of the diffeomorphism f at the origin has the properties described in Lemma 4.1, and the nonlinear term F vanishes at the origin together with its Jacobi matrix. We denote by y and z coordinates in the subspaces $\mathcal{S}$ and $\mathcal{U}$ introduced in Section 4.1. Fix a number $\lambda \in (0, 1)$ such that

$$\|B\| \leq \lambda \quad \text{and} \quad \|C^{-1}\| \leq \lambda.$$

Theorem 4.2. *Assume that $x = 0$ is a hyperbolic fixed point of a diffeomorphism f that has the form* (4.4) *in a neighborhood U of the origin. Then*

(1) *there exists a number $\Delta > 0$ such that*

(1.1) *there exists a mapping*

$$\gamma : \{y : |y| < \Delta\} \to \mathcal{U}$$

of class C^1 such that

$$\gamma(0) = 0, \quad \frac{\partial \gamma}{\partial y}(0) = 0, \tag{4.27}$$

and the surface

$$\Gamma = \{(y, \gamma(y)) : |y| < \Delta\}$$

has the following property: If $x \in \Gamma$, then $|f^k(x)| \to 0, k \to \infty$;

(1.2) *if $|f^k(x)| < \Delta$ for $k \geq 0$, then $x \in \Gamma$;*

(2) *for any numbers* $b > 1$ *and* $\nu \in (\lambda, 1)$ *there exists a number* $\delta = \delta(b, \nu)$ *such that if* $x_0 = (y_0, z_0) \in \Gamma$ *and* $|y_0| < \delta$, *then*

$$|f^k(x_0)| \le b\nu^k |y_0|, \quad k \ge 0. \tag{4.28}$$

Remark. Reducing Δ, if necessary, we may assume that the following inclusions hold:

$$\{(y, \gamma(y)) : |y| < \Delta\} \subset \{|x| < \Delta\} \subset U$$

(we take relations (4.27) into account). Clearly,

$$W^s_U(0) \cap \{|y| < \Delta\} = \Gamma$$

in this case. Thus, the local stable manifold of the fixed point $x = 0$ is a C^1-smooth disk in a small neighborhood of this point.

Remark. Relations (4.27) imply that the subspace $\mathcal{S}$ is the tangent space of the surface Γ at the origin.

Remark. If the diffeomorphism f belongs to the class $C^k, k \ge 1$ (or f is analytic), then the disk Γ has the same smoothness (see, for example, [15]).

Remark. Since $A^k x_0 = (B^k y_0, C^k z_0)$, the rate of convergence of a trajectory of the linear mapping L to the origin along the stable subspace $\mathcal{S}$ is estimated by the inequalities $|B^k y_0| \le \lambda^k |y_0|$.

Statement (2) of Theorem 4.2 shows that the rate of convergence of a trajectory of the diffeomorphism f to the hyperbolic fixed point $x = 0$ along the local stable manifold $W^s_U(0)$ is arbitrarily close to the corresponding rate for the linear mapping (if the initial point is close enough to the fixed point).

The proof of Theorem 4.2 is long; we divide it into several lemmas. First we note that the following two auxiliary statements hold.

Let $\nu \in (\lambda, 1)$. Then

$$\sum_{i=0}^{k-1} \lambda^{k-1-i} \nu^i < \nu^k \frac{1}{\nu - \lambda} \tag{4.29}$$

and

$$\sum_{i=k}^{\infty} \lambda^{i+1-k} \nu^i = \nu^k \frac{\lambda}{1 - \nu\lambda} \tag{4.30}$$

for any $k > 0$.

We prove inequality (4.29) (equality (4.30) is established similarly):

$$\sum_{i=0}^{k-1} \lambda^{k-1-i} \nu^i = \lambda^{k-1} \sum_{i=0}^{k-1} \left(\frac{\nu}{\lambda}\right)^i = \lambda^{k-1} \frac{(\nu/\lambda)^k - 1}{\nu/\lambda - 1}$$
$$< \nu^k \frac{1/\lambda}{\nu/\lambda - 1} = \nu^k \frac{1}{\nu - \lambda}.$$

First we fix numbers $b_0 > 1$ and $\nu_0 \in (\lambda, 1)$. Take a number Δ such that the set $D_\Delta = \{|x| \le b_0 \Delta\}$ belongs to the domain of definition of the diffeomorphism f and the Lipschitz constant l_0 of the function F on the set D_Δ satisfies the inequalities

$$l_0^* := \max\left(\frac{l_0}{\nu_0 - \lambda}, \frac{l_0}{1 - \nu_0 \lambda}\right) < 1 \tag{4.31}$$

and

$$1 + \frac{l_0 b_0}{\nu_0 - \lambda} < b_0. \tag{4.32}$$

Our main tool in the proof of Theorem 4.2 is the so-called *Perron operator* defined as follows. Fix a vector y_0 and assign to a sequence $X = \{x_k \in \mathbb{R}^n : k \ge 0\}$ with $x_0 = (y_0, z_0)$ the sequence $T(X) = \Xi = \{\xi_k \in \mathbb{R}^n : k \ge 0\}$, where

$$\eta_k = B^k y_0 + \sum_{i=0}^{k-1} B^{k-1-i} f_1(x_i) \quad \text{and} \quad \zeta_k = -\sum_{i=k}^{\infty} C^{k-1-i} f_2(x_i) \tag{4.33}$$

(we represent the vector ξ_k in the form (η_k, ζ_k) according to the representation $x = (y, z)$; f_1 and f_2 are the corresponding components of F). We show that the considered objects are well defined (i.e., that the series representing ζ_k is convergent) using a separate proof in any particular case.

Fix numbers $d > 0$, $b > 1$, and $\nu \in (\lambda, 1)$. Denote by $G(d, b, \nu)$ the set of sequences $X = \{x_k \in \mathbb{R}^n : k \ge 0\}$ that satisfy the inequalities

$$|x_k| \le b d \nu^k, \quad k \ge 0. \tag{4.34}$$

Lemma 4.6. *For any numbers $b > 1$ and $\nu \in (\lambda, 1)$ there exists a number $\delta > 0$ such that if $|y_0| \le \delta$, then there exists a positive semitrajectory $X = \{x_k : k \ge 0\}$ of the diffeomorphism f with the following properties: $x_0 = (y_0, z_0)$ and $X \in G(|y_0|, b, \nu)$.*

Proof. Take a number $\delta \le \Delta$ such that the Lipschitz constant l of the function F on the set $D_\delta = \{|x| \le b\delta\}$ satisfies the inequalities

$$l^* := \max\left(\frac{l}{\nu - \lambda}, \frac{l}{1 - \nu\lambda}\right) < 1 \tag{4.35}$$

and

$$1 + \frac{lb}{\nu - \lambda} < b. \tag{4.36}$$

Fix a vector y_0 with $|y_0| \leq \delta$ and assign to a sequence $X \in G(|y_0|, b, \nu)$ with $x_0 = (y_0, z_0)$ the sequence $T(X) = \Xi = \{\xi_k : k \geq 0\}$. We perform the required estimates and prove simultaneously that the series in the definition of ζ_k converges, i.e., the operator T is well defined.

If $X \in G(|y_0|, b, \nu)$, then $x_k \in D_\delta$; hence, we obtain the following inequalities:

$$|f_j(x_i)| \leq l|x_i| \leq lb|y_0|\nu^i, \quad i \geq 0, j = 1, 2.$$

It follows that

$$|\eta_k| \leq \lambda^k|y_0| + \sum_{i=0}^{k-1} \lambda^{k-1-i} lb|y_0|\nu^i \leq \lambda^k|y_0| + \frac{lb}{\nu - \lambda}|y_0|\nu^k \leq b|y_0|\nu^k$$

(we take into account the inequalities $\lambda < \nu$, (4.29), and (4.36)) and

$$|\zeta_k| \leq \sum_{i=k}^{\infty} \lambda^{i+1-k} lb|y_0|\nu^i \leq \frac{lb\lambda}{1 - \nu\lambda}|y_0|\nu^k \leq b|y_0|\nu^k$$

(we take into account inequality (4.30) and inequality (4.35) which implies that

$$\frac{lb\lambda}{1 - \nu\lambda} \leq b).$$

Thus, $T(X) \in G(|y_0|, b, \nu)$.

Let us construct a sequence $\{X^m \in G(|y_0|, b, \nu) : m \geq 0\}$ as follows: We set $X^0 = \{x_k = 0 : k \geq 0\}$ (clearly, $X^0 \in G(|y_0|, b, \nu)$) and $X^m = T(X^{m-1})$ (as was shown above, in this case $X^m \in G(|y_0|, b, \nu)$ for any m).

Let $X^m = \{x_k^m : k \geq 0\}$, where $x_k^m = (y_k^m, z_k^m)$. Let us show (using induction on m) that

$$|x_k^{m+1} - x_k^m| \leq (l^*)^m b|y_0|\nu^k, \quad m \geq 0. \tag{4.37}$$

Since $X^1 \in G(|y_0|, b, \nu)$, estimates (4.37) are valid for $m = 0$. Assume that

$$|x_k^m - x_k^{m-1}| \leq (l^*)^{m-1} b|y_0|\nu^k.$$

Then

$$\begin{aligned} |y_k^{m+1} - y_k^m| &= \Big|\sum_{i=0}^{k-1} B^{k-1-i} f_1(x_i^m) - \sum_{i=0}^{k-1} B^{k-1-i} f_1(x_i^{m-1})\Big| \\ &\leq \sum_{i=0}^{k-1} \lambda^{k-1-i} l(l^*)^{m-1} b|y_0|\nu^i \leq (l^*)^m b|y_0|\nu^k \end{aligned} \tag{4.38}$$

and

$$|z_k^{m+1} - z_k^m| = \Big| \sum_{i=k}^{\infty} C^{k-1-i} f_2(x_i^m) - \sum_{i=k}^{\infty} C^{k-1-i} f_2(x_i^{m-1}) \Big|$$
$$\leq \sum_{i=k}^{\infty} \lambda^{i+1-k} l (l^*)^{m-1} b |y_0| \nu^i \leq (l^*)^m b |y_0| \nu^k. \qquad (4.39)$$

Combining estimates (4.38) and (4.39), we complete the proof of estimates (4.37).

Since $l^* < 1$ (see inequality (4.35)), estimates (4.37) imply that the limit $\lim_{m\to\infty} x_k^m = x_k$ exists for any $k \geq 0$.

We claim that the sequence $X = \{x_k\}$ is a positive semitrajectory of the diffeomorphism f.

Indeed, the equalities

$$y_k^{m+1} = B^k y_0 + B^{k-1} f_1(x_0^m) + \cdots + f_1(x_{k-1}^m)$$

imply that

$$y_{k+1}^{m+1} = B^{k+1} y_0 + B^k f_1(x_0^m) + \cdots + B f_1(x_{k-1}^m) + f_1(x_k^m)$$
$$= B y_k^{m+1} + f_1(x_k^m); \qquad (4.40)$$

similarly, the equalities

$$z_k^{m+1} = -(C^{-1} f_2(x_k^m) + C^{-2} f_2(x_{k+1}^m) + C^{-3} f_2(x_{k+2}^m) + \cdots)$$

imply that

$$z_{k+1}^{m+1} = -(C^{-1} f_2(x_{k+1}^m) + C^{-2} f_2(x_{k+2}^m) + \cdots) = C z_k^{m+1} + f_2(x_k^m). \quad (4.41)$$

It follows from equalities (4.40) and (4.41) that

$$x_{k+1}^{m+1} = A x_k^{m+1} + F(x_k^m). \qquad (4.42)$$

Passing in equalities (4.42) to the limit as $m \to \infty$, we conclude that $x_{k+1} = Ax_k + F(x_k)$; thus, the sequence $X = \{x_k\}$ is a positive semitrajectory of the diffeomorphism f.

In addition, passing to the limit as $m \to \infty$ in the inequalities $|x_k^m| \leq b|y_0|\nu^k$, we see that $X \in G(|y_0|, b, \nu)$.

To complete the proof, it is enough to note that the definition of the operator T implies the equalities $x_0^m = (y_0, z_0^m), m \geq 1$. Hence, $x_0 = (y_0, z_0)$. □

Remarks. 1. Note that inequalities (4.31) and (4.32), similar to inequalities (4.35) and (4.36), hold. Hence, the statement of Lemma 4.6 is valid for $|y_0| \leq \Delta$; in this case, there exists a positive semitrajectory $X = \{x_k : k \geq 0\}$ of the diffeomorphism f such that $x_0 = (y_0, z_0)$ and $X \in G(|y_0|, b_0, \nu_0)$.

2. The same reasoning as that used in the proof of (4.42) shows that if a sequence $X = \{x_k\}$ with $|x_k| \leq b_0\Delta$ is a fixed point of the operator T (i.e., $T(X) = X$), then the sequence X is a positive semitrajectory of the diffeomorphism f.

Lemma 4.7. *For any $|y_0| \leq \Delta$, the Perron operator T has not more than one fixed point $X = \{x_k\}$ with $|x_k| \leq b_0\Delta$.*

Proof. Let $X = \{x_k\}$ and $X' = \{x'_k\}$ be fixed points of the operator T with $|x_k|, |x'_k| \leq b_0\Delta$. Set

$$\mu = \sup_{k \geq 0} |x_k - x'_k|;$$

let $x_k = (y_k, z_k)$ and $x'_k = (y'_k, z'_k)$.

The estimates

$$|y_k - y'_k| \leq \Big| \sum_{i=0}^{k-1} B^{k-1-i} f_1(x_i) - \sum_{i=0}^{k-1} B^{k-1-i} f_1(x'_i) \Big|$$
$$\leq \sum_{i=0}^{k-1} \lambda^{k-1-i} l_0 |x_i - x'_i| \leq \frac{l_0 \mu}{1-\lambda} \leq l_0^* \mu$$

and

$$|z_k - z'_k| \leq \Big| \sum_{i=k}^{\infty} C^{k-1-i} f_2(x_i) - \sum_{i=k}^{\infty} C^{k-1-i} f_2(x'_i) \Big|$$
$$\leq \frac{l_0 \mu \lambda}{1-\lambda} \leq l_0^* \mu$$

imply that $\mu \leq l_0^* \mu$.

Since $l_0^* < 1$, $\mu = 0$. This completes the proof. □

Lemma 4.8. *Assume that $X = \{x_k\}$ is a positive semitrajectory of the diffeomorphism f with $|x_k| \leq b_0\Delta$ and $x_0 = (y_0, z_0)$. Then X is a fixed point of the operator T corresponding to y_0.*

Proof. The equalities

$$y_1 = By_0 + f_1(x_0), \quad y_2 = By_1 + f_1(x_1) = B^2 y_0 + Bf_1(x_0) + f_1(x_1), \quad \dots$$

imply that

$$y_k = B^k y_0 + \sum_{i=0}^{k-1} B^{k-1-i} f_1(x_i).$$

Fix an index $k \geq 0$. It follows from the equality

$$z_{k+1} = Cz_k + f_2(x_k)$$

that

$$z_k = C^{-1}z_{k+1} - C^{-1}f_2(x_k).$$

Further, it follows from the equality

$$z_{k+2} = Cz_{k+1} + f_2(x_{k+1})$$

that

$$z_k = C^{-2}z_{k+2} - C^{-2}f_2(x_{k+1}) - C^{-1}f_2(x_k).$$

Continuing this chain of equalities, we conclude that

$$z_k = C^{-m}z_{k+m} - \sum_{i=k}^{k+m-1} C^{k-1-i} f_2(x_i) \tag{4.43}$$

for any natural m.

Since $|z_{k+m}| \leq b_0\Delta$ and $|f_2(x_i)| \leq l_0 b_0 \Delta$, the first summand in (4.43) tends to 0 as $m \to \infty$, the second summand is a partial sum of the convergent series $\sum_{i=k}^{\infty} C^{k-1-i} f_2(x_i)$.

Passing in (4.43) to the limit as $m \to \infty$, we get the equality

$$z_k = -\sum_{i=k}^{\infty} C^{k-1-i} f_2(x_i).$$

This completes the proof. □

Let us summarize the conclusions of Lemmas 4.6–4.8 as follows:

There exists a mapping

$$\gamma : \{y : |y| \leq \Delta\} \to \mathcal{U}$$

such that

- if $x_0 = (y_0, \gamma(y_0))$, $|y_0| \leq \Delta$, and $X = \{x_k : k \geq 0\}$ is the positive semitrajectory of the point x_0 for the diffeomorphism f, then $X \in G(|y_0|, b_0, \nu_0)$;
- if $x_0 = (y_0, z_0)$, $|y_0| \leq \Delta$, and $z_0 \neq \gamma(y_0)$, then the semitrajectory $\{f^k(x_0) : k \geq 0\}$ leaves the set D_Δ as k grows.

The mapping γ can be defined as follows: Take a point y_0 with $|y_0| \leq \Delta$; let X be the unique fixed point of the Perron operator T corresponding to y_0. Then

$$\gamma(y_0) = z_0 = -\sum_{i=0}^{\infty} C^{k-1-i} f_2(x_i). \tag{4.44}$$

Now we show that the mapping γ is of class C^1. First we show that γ is Lipschitz continuous. To be more precise, the following statement holds.

Lemma 4.9. *For any $L > 0$ there exists a $\delta_1 > 0$ such that*

$$|\gamma(y) - \gamma(y')| \leq L|y - y'| \tag{4.45}$$

if $|y|, |y'| \leq \delta_1$.

Proof. Fix numbers $b > \max(1, L)$ and $\nu \in (\lambda, 1)$ and find a number $\delta_2 \leq \Delta$ such that the Lipschitz constant l of the function F on the set $D_{\delta_2} = \{|x| \leq b\delta_2\}$ satisfies inequalities (4.35) and (4.36) and also the inequality

$$\frac{lb\lambda}{\nu - \lambda} < L. \tag{4.46}$$

Set $\delta_1 = \delta_2(1 - l^*)$. Let $|y|, |y'| \leq \delta_1$.

Using the same reasoning as in the proof of Lemma 4.6, let us construct successive approximations X^m and Ξ^m which converge to the fixed points of the operator T for y and y': $X^0 = \{x_k = 0\}$, $\Xi^0 = \{\xi_k = 0\}$, and $X^m = T(X^{m-1})$ and $\Xi^m = T(\Xi^{m-1})$ for $m \geq 1$.

Let $X^m = \{x_k^m = (y_k^m, z_k^m)\}$ and $\Xi^m = \{\xi_k^m = (\eta_k^m, \zeta_k^m)\}$. Inequalities (4.37) imply that

$$|x_k^m|, |\xi_k^m| \leq \frac{1}{1 - l^*} b\delta_1 \leq b\delta_2.$$

We show (using induction on m) that the following inequalities hold:

$$|y_k^m - \eta_k^m| \leq b\nu^k |y - y'|, \quad k \geq 0, \tag{4.47}$$

and

$$|z_k^m - \zeta_k^m| \leq L\nu^k |y - y'|, \quad k \geq 0. \tag{4.48}$$

The choice of b implies that if inequalities (4.47) and (4.48) are satisfied, then

$$|x_k^m - \xi_k^m| \leq b\nu^k |y - y'|, \quad k \geq 0.$$

Clearly, inequalities (4.47) and (4.48) are satisfied for $m = 0$. Assume that they are satisfied for $m - 1$.

Then

$$
\begin{aligned}
|y_k^m - \eta_k^m| &\le |B^k(y - y')| + \Big| \sum_{i=0}^{k-1} B^{k-1-i}(f_1(x_i^{m-1}) - f_1(\xi_i^{m-1})) \Big| \\
&\le \lambda^k |y - y'| + \sum_{i=0}^{k-1} \lambda^{k-1-i} l b \nu^k |y - y'| \\
&\le \left(\lambda^k + lb \frac{1}{\nu - \lambda} \nu^k \right) |y - y'| \le b\nu^k |y - y'|
\end{aligned}
$$

(we take inequality (4.36) into account) and

$$
\begin{aligned}
|z_k^m - \zeta_k^m| &\le \Big| \sum_{i=k}^{\infty} C^{k-1-i}(f_2(x_i^{m-1}) - f_2(\xi_i^{m-1})) \Big| \\
&\le lb \frac{\lambda}{1 - \nu\lambda} \nu^k |y - y'| \le L\nu^k |y - y'|
\end{aligned}
$$

(we take inequality (4.46) into account). Thus, we have proven inequalities (4.47) and (4.48).

Now we pass to the limit as $m \to \infty$ in the inequality

$$|z_0^m - \zeta_0^m| \le L|y - y'|$$

and get the required estimate (4.45). □

Remark. The same reasoning as in the proof of Lemma 4.9 shows that the mapping γ is Lipschitz continuous with constant $L_0 = b_0$ on the set $\{|y| \le \Delta\}$.

For this purpose, one has to replace inequality (4.47) by the inequality

$$|y_k^m - \eta_k^m| \le L_0 \nu^k |y - y'|, \quad k \ge 0,$$

and inequality (4.48) by the inequality

$$|z_k^m - \zeta_k^m| \le L_0 \nu^k |y - y'|, \quad k \ge 0.$$

In this case, inequalities (4.31) and (4.32) imply the inequality

$$|x_k^m - \xi_k^m| \le L_0 \nu^k |y - y'|, \quad k \ge 0,$$

and its limit (as $m \to \infty$) variant,

$$|x_k - \xi_k| \le L_0 \nu^k |y - y'|, \quad k \ge 0. \tag{4.49}$$

Lemma 4.9 shows that the Lipschitz constant of the mapping γ can be done arbitrarily small by the choice of a small neighborhood of the point $y = 0$ in $\mathcal{S}$.

Lemma 4.10. *The mapping γ is continuously differentiable on the set $|y| < \Delta$, and the second equality in* (4.27) *holds.*

Proof. Let e_i be the ith unit basic vector of the space $\mathbb{R}^m$. Take a vector y with $|y| < \Delta$ and a scalar h and set $y' = y + h_i$, where $h_i = he_i$ (we assume that the value $|h|$ is small enough, so that $|y'| < \Delta$). Let $X = \{x_k\}$ and $X' = \{x'_k\}$ be the fixed points of the Perron operator T corresponding to y and y', respectively.

Denote $P(h) = \{p_k(h) : k \geq 0\} = X - X'$. Let $p_k(h) = (y_k(h), z_k(h))$ according to the representation $x = (y, z)$. Then

$$y_k(h) = B^k h_i + \sum_{i=0}^{k-1} B^{k-1-i} \left[f_1(x_i) - f_1(x'_i) \right]$$

and

$$z_k(h) = -\sum_{i=k}^{\infty} C^{k-1-i} \left[f_2(x_i) - f_2(x'_i) \right].$$

Since the functions f_1 and f_2 are continuously differentiable,

$$y_k(h) = B^k h_i + \sum_{i=0}^{k-1} B^{k-1-i} \left[\frac{\partial f_1}{\partial x}(x_i) p_i(h) + d_{1,i} \right] \tag{4.50}$$

and

$$z_k(h) = -\sum_{i=k}^{\infty} C^{k-1-i} \left[\frac{\partial f_2}{\partial x}(x_i) p_i(h) + d_{2,i} \right], \tag{4.51}$$

where the order of smallness of the values $d_{1,i}$ and $d_{2,i}$ as $h \to 0$ is higher than the order of smallness of p_i.

Inequality (4.49) implies that

$$|p_k(h)| \leq L_0 |h| \nu^k, \quad k \geq 0, \tag{4.52}$$

where L_0 is a Lipschitz constant of the mapping γ on the set $\{|y| \leq \Delta\}$ (see the remark after Lemma 4.9).

The Jacobi matrix $\partial F / \partial x$ is uniformly continuous on the compact set D_Δ. Hence, it follows from inequalities (4.52) that there exists a function $\chi(h)$ with the following properties: $\chi(h) \to 0$ as $h \to 0$ and

$$|d_{j,k}| \leq \chi(h) |h| \nu^k, \quad k \geq 0, j = 1, 2. \tag{4.53}$$

Fix a number μ, $|\mu| \leq \Delta$, and consider the operator Θ that assigns to a sequence $G = \{g_k \in \mathbb{R}^n : k \geq 0\}$ the sequence $\Theta(G) = Q = \{q_k \in \mathbb{R}^n : k \geq 0\}$, where

$$r_k = B^k \mu e_i + \sum_{i=0}^{k-1} B^{k-1-i} \frac{\partial f_1}{\partial x}(x_i) g_i \quad \text{and} \quad t_k = -\sum_{i=k}^{\infty} C^{k-1-i} \frac{\partial f_2}{\partial x}(x_i) g_i$$

(as above, we represent the vector q_k in the form (r_k, t_k) according to the representation $x = (y, z)$).

The operator Θ has precisely the same form as the operator T. Our proof of existence of a fixed point of the operator T (see Lemma 4.6 and the remark after it) used the estimates of the matrices B and C listed before the statement of Theorem 4.2 and the estimates of Lipschitz constants of the functions f_1 and f_2 on the set D_Δ.

Since the multipliers at B^{k-1-i} and C^{k-1-i} in the definition of the operator Θ are linear with respect to the multipliers g_i, and the estimates

$$\left\| \frac{\partial f_1}{\partial x}(x) \right\|, \left\| \frac{\partial f_2}{\partial x}(x) \right\| \leq l_0$$

are valid in D_Δ, one can prove the existence of a fixed point of the operator Θ using the same reasoning as that used in the proof of existence of a fixed point of the operator T.

Let G be a fixed point of the operator Θ; clearly, the sequence $U = \{u_k = g_k/\mu\}$ satisfies the equalities

$$v_k = B^k e_i + \sum_{i=0}^{k-1} B^{k-1-i} \frac{\partial f_1}{\partial x}(x_i) u_i \quad \text{and} \quad w_k = -\sum_{i=k}^{\infty} C^{k-1-i} \frac{\partial f_2}{\partial x}(x_i) u_i, \tag{4.54}$$

where $u_k = (v_k, w_k)$.

Divide equalities (4.50) and (4.51) by h and subtract form the results equalities (4.54); we get the equalities

$$\frac{y_k(h)}{h} - v_k = \sum_{i=0}^{k-1} B^{k-1-i} \left[\frac{\partial f_1}{\partial x}(x_i) \left(\frac{p_i(h)}{h} - u_i \right) + \frac{d_{1,i}}{h} \right]$$

and

$$\frac{z_k(h)}{h} - w_k = -\sum_{i=k}^{\infty} C^{k-1-i} \left[\frac{\partial f_2}{\partial x}(x_i) \left(\frac{p_i(h)}{h} - u_i \right) + \frac{d_{2,i}}{h} \right].$$

These equalities imply that

$$\left| \frac{y_k(h)}{h} - v_k \right| \leq \sum_{i=0}^{k-1} \lambda^{k-1-i} \left(l_0 \left| \frac{p_i(h)}{h} - u_i \right| + \left| \frac{d_{1,i}}{h} \right| \right) \tag{4.55}$$

and

$$\left|\frac{z_k(h)}{h} - w_k\right| \le \sum_{i=k}^{\infty} \lambda^{i-k+1} \left(l_0 \left|\frac{p_i(h)}{h} - u_i\right| + \left|\frac{d_{2,i}}{h}\right| \right). \tag{4.56}$$

Let

$$\theta = \sup_{k \ge 0} \left|\frac{p_k(h)}{h} - u_k\right|.$$

Inequalities (4.55), (4.56), and (4.53) imply that

$$\theta \le l_0^* \theta + \frac{l_0^*}{l_0} \chi(h).$$

Since $l_0^* < 1$, it follows from the last inequality that $\theta \to 0$ as $h \to 0$.

Thus, there exists the partial derivative

$$\frac{\partial \gamma}{\partial y} y_i = \frac{\partial z_0}{\partial y_i} = w_0$$

(we take equality (4.44) into account).

Let us show that this partial derivative is continuous (i.e., w_0 depends on y continuously).

Take a vector y with $|y| < \Delta$ and an arbitrary $\varepsilon > 0$.

Since the matrix $\partial F/\partial x$ is continuous (and, hence, uniformly continuous on the set D_Δ), there exists a $\delta > 0$ such that if $y' \in D_\Delta$, $|y - y'| < \delta$, and $X = \{x_k\}$ and $X' = \{x'_k\}$ are the fixed points of the Perron operator T corresponding to y and y', respectively, then

$$\left\| \frac{\partial F}{\partial x}(x_k) - \frac{\partial F}{\partial x}(x'_k) \right\| < \varepsilon$$

(see the Remark after Lemma 4.9). In what follows, we assume that if $|y - y'| < \delta$, then $y' \in D_\Delta$.

Let $U = \{u_k = (v_k, w_k)\}$ and $U' = \{u'_k = (v'_k, w'_k)\}$ be the sequences that satisfy equalities (4.54) and the analogs of these equalities for y'.

Since the sequence $G = \{g_k = \mu u_k\}$ (the fixed point of the operator Θ) belongs to the set $G(|\mu|, b_0, \nu_0)$, there exists a constant $c > 0$ such that $|u_k| \le c\nu_0^k$.

Set $M = \sup_{k\geq 0} |u_k - u'_k|$ (clearly, this value is finite). Let us estimate (we take into account that $\|\partial F/\partial x\| \leq l_0$)

$$
\begin{aligned}
|v_k - v'_k| &= \left| \sum_{i=0}^{k-1} B^{k-1-i} \left(\frac{\partial f_1}{\partial x}(x_i)u_i - \frac{\partial f_1}{\partial x}(x'_i)u'_i \right) \right| \\
&\leq \sum_{i=0}^{k-1} \lambda^{k-1-i} \left(\left| \left[\frac{\partial f_1}{\partial x}(x_i) - \frac{\partial f_1}{\partial x}(x'_i) \right] u_i \right| + \left| \frac{\partial f_1}{\partial x}(x'_i)(u_i - u'_i) \right| \right) \\
&\leq \sum_{i=0}^{k-1} \lambda^{k-1-i} (\varepsilon c \nu_0^i + l_0 M) \\
&\leq \varepsilon c \frac{\nu_0^k}{\nu_0 - \lambda} + \frac{l_0 M}{1-\lambda} \leq \varepsilon c \frac{1}{\nu_0 - \lambda} + l_0^* M.
\end{aligned}
$$

Similar estimates show that

$$
|w_k - w'_k| \leq \varepsilon c \frac{1}{1 - \nu_0 \lambda} + l_0^* M.
$$

Thus,

$$
M \leq \varepsilon c \frac{l_0^*}{l_0} + l_0^* M.
$$

Hence,

$$
|w_0 - w'_0| \leq M \leq \frac{c\varepsilon l_0^*}{l_0(1 - l_0^*)},
$$

which establishes the required continuous dependence of w_0 on y.

The second equality in (4.27) follows from the Remark after Lemma 4.9. The proof of Theorem 4.2 is complete. □

4.5 Hyperbolic periodic point

Let p be a periodic point of a diffeomorphism f of period $m > 1$. Denote $p_i = f^i(p), i = 0, \ldots, m-1$.

The point p is called a *hyperbolic periodic point* if p is a hyperbolic fixed point of the diffeomorphism f^m.

Let us show that in this case the points p_i are hyperbolic periodic points as well.

Consider two points p_i and p_j with $0 \leq i < j \leq m-1$. Denote $g = f^{j-i}$; clearly, $g^{-1} = f^{i-j}$. Thus, $f^m = g^{-1} \circ f^m \circ g$.

Then

$$Df^m(p_i) = Dg^{-1}(p_j)\, Df^m(p_j)\, Dg(p_i).$$

Since $g(p_i) = p_j$, the matrices $Dg^{-1}(p_j)$ and $Dg(p_i)$ are mutually inverse. It follows that the sets of eigenvalues of the matrices $Df^m(p_i)$ and $Df^m(p_j)$ are the same, and these matrices are (or are not) hyperbolic simultaneously, which proves our claim.

We consider the stable manifold of the fixed point p of the diffeomorphism f^m and call it the stable manifold $W^s(p)$ of the hyperbolic periodic point p. Thus, $x \in W^s(p)$ if and only if $f^{mk}(x) \to p$ as $k \to \infty$.

Let $x' = f^j(x)$. Clearly, $f^{mk}(x) \to p$ if and only if $f^{mk}(x') = f^{mk+j}(x) \to f^j(p)$ as $k \to \infty$. Thus,

$$f^j(W^s(p)) = W^s(p_j). \tag{4.57}$$

Note that

$$W^s(p_i) \cap W^s(p_j) = \emptyset, \quad 0 \leq i < j \leq m-1. \tag{4.58}$$

Indeed, if $x \in W^s(p_i) \cap W^s(p_j)$, then $f^{mk}(x) \to p_i$ and $f^{mk}(x) \to p_j$ as $k \to \infty$, which is impossible.

We define the stable manifold $W^s(O(p, f))$ of the trajectory of the periodic point p by the equality

$$W^s(O(p, f)) = \bigcup_{i=0}^{m-1} W^s(p_i).$$

Relations (4.57) imply that $W^s(O(p, f))$ is an invariant set of the diffeomorphism f. Relations (4.58) show that $W^s(O(p, f))$ is the union of m disjoint copies of a Euclidean space of the same dimension.

The unstable manifold $W^u(O(p, f))$ of the trajectory of the periodic point p is defined similarly.

Chapter 5
Hyperbolic rest point and hyperbolic closed trajectory

5.1 Hyperbolic rest point

Consider an autonomous system of differential equations

$$\frac{dx}{dt} = F(x), \quad F \in C^1(\mathbb{R}^n). \tag{5.1}$$

As above, we assume that system (5.1) generates a flow ϕ. Let p be a rest point of system (5.1). We say that the point p is a *hyperbolic rest point* if the eigenvalues λ_j of the Jacobi matrix

$$P = \frac{\partial F}{\partial x}(p)$$

satisfy the condition

$$\operatorname{Re}\lambda_j \neq 0, \quad j = 1, \dots, n. \tag{5.2}$$

In the case of a hyperbolic rest point, the following two analogs of Theorems 4.1 and 4.2 hold.

Theorem 5.1. *A hyperbolic rest point p of system* (5.1) *is locally topologically conjugate with the rest point $u = 0$ of the linear system*

$$\frac{du}{dt} = Pu. \tag{5.3}$$

Let us assume for definiteness that

$$\operatorname{Re}\lambda_j < 0,\ j = 1, \dots, m, \quad \operatorname{Re}\lambda_j > 0,\ j = m+1, \dots, n$$

(the cases $m = 0$ and $m = n$ are not excluded).

We also assume that p is the origin of $\mathbb{R}^n$ and that coordinates are chosen so that

$$P = \operatorname{diag}(P_1, P_2), \tag{5.4}$$

where P_1 is an $m \times m$ matrix whose eigenvalues satisfy the inequalities $\operatorname{Re}\lambda_j < 0$, and P_2 is an $(n-m) \times (n-m)$ matrix whose eigenvalues satisfy the inequalities $\operatorname{Re}\lambda_j > 0$. As in Section 4, we represent vectors x and u in the form $x = (y, z)$ and $u = (v, w)$ according to representation (5.4). Consider the subspaces $\mathcal{S} = \{z = 0\}$ and $\mathcal{U} = \{y = 0\}$.

Theorem 5.2. *Assume that* $x = 0$ *is a hyperbolic rest point of system* (5.1). *Then there exists a number* $\Delta > 0$ *and a mapping*

$$\gamma : \{y : |y| < \Delta\} \to \mathcal{U}$$

of class C^1 *such that*

(1)

$$\gamma(0) = 0, \quad \frac{\partial \gamma}{\partial y}(0) = 0,$$

and the surface

$$\Gamma = \{(y, \gamma(y)) : |y| < \Delta\}$$

has the following property: If $x \in \Gamma$*, then* $\phi(t, x) \to 0, t \to \infty$*;*

(2) *if* $|\phi(t, x)| < \Delta$ *for* $t \geq 0$*, then* $x \in \Gamma$*.*

Let us show that Theorems 5.1 and 5.2 are reduced to Theorems 4.1 and 4.2.

Denote by $\Psi(t)$ the fundamental matrix of system (5.3) that equals the unit matrix at $t = 0$. Representation (5.4) implies that

$$\Psi(t) = e^{Pt} = \operatorname{diag}(e^{P_1 t}, e^{P_2 t}).$$

It is shown in the basic course of differential equations that if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of P, then

$$\mu_1(t) = e^{\lambda_1 t}, \dots, \mu_n(t) = e^{\lambda_n t}$$

are the eigenvalues of $\Psi(t)$. Hence, there exists a number $T > 0$ such that

$$|\mu_j(T)| < 1, \ j = 1, \dots, m, \quad |\mu_j(T)| > 1, \ j = m+1, \dots, n.$$

Thus, the linear mapping L with the matrix

$$A = \operatorname{diag}(e^{P_1 T}, e^{P_2 T})$$

is hyperbolic.

In addition, it is known that

$$\|e^{P_1 t}\| \to 0, \quad t \to \infty,$$

and

$$\|e^{P_2 t}\| \to 0, \quad t \to -\infty.$$

Hence, we may assume that

$$\|e^{P_1 T}\| < 1 \quad \text{and} \quad \|e^{-P_2 T}\| < 1. \tag{5.5}$$

Consider the diffeomorphism $f(x) = \phi(T, x)$. Let us show that f is representable in the form

$$f(x) = Ax + G(x), \tag{5.6}$$

where the nonlinearity G vanishes at the origin together with its Jacobi matrix. Hence, representation (5.6) is similar to representation (4.4).

In a neighborhood of the point $x = 0$, we can write the right-hand side of system (5.6) in the form

$$F(x) = Px + H(x), \tag{5.7}$$

where the nonlinearity H vanishes at the origin together with its Jacobi matrix.

According to representation (5.4), let us write

$$x = (y, z), \quad \phi(t, x_0) = (y(t, y_0, z_0), z(t, y_0, z_0)), \quad H = (H_1, H_2).$$

Then the functions $y(t, y_0, z_0)$ and $z(t, y_0, z_0)$ satisfy the system

$$\dot{y} = P_1 y + H_1(y, z), \quad \dot{z} = P_2 z + H_2(y, z).$$

Thus, the function $y(t, y_0, z_0)$ is a solution of the nonhomogeneous linear system of differential equations

$$\dot{y} = P_1 y(t, y_0, z_0) + H_1(y(t, y_0, z_0), z(t, y_0, z_0)).$$

By the Lagrange formula,

$$y(T, y_0, z_0) = e^{P_1 T} y_0 + \int_0^T e^{P_1(T-s)} H_1(y(s, y_0, z_0), z(s, y_0, z_0)) ds. \tag{5.8}$$

Set

$$G_1(y_0, z_0) = \int_0^T e^{P_1(T-s)} H_1(y(s, y_0, z_0), z(s, y_0, z_0)) ds.$$

Since $(y(t, 0, 0), z(t, 0, 0)) \equiv (0, 0)$, formula (5.8) implies that $G_1(0, 0) = 0$.

Further,

$$\frac{\partial G_1}{\partial y_0} = \int_0^T e^{P_1(T-s)} \left(\frac{\partial H_1}{\partial y} \frac{\partial y}{\partial y_0} + \frac{\partial H_1}{\partial z} \frac{\partial z}{\partial y_0} \right) ds.$$

Hence, it follows from properties of the nonlinearity H that

$$\frac{\partial G_1}{\partial y_0}(0,0) = 0.$$

A similar reasoning shows that

$$\frac{\partial G_1}{\partial z_0}(0,0) = 0.$$

Applying a similar reasoning to the function

$$G_2(y_0, z_0) = z(T, y_0, z_0) - e^{P_2 T} z_0,$$

we complete the proof of formula (5.6).

Thus, Theorem 4.1 is applicable to the hyperbolic linear mapping L and the diffeomorphism f.

Denote by h_0 a homeomorphism that conjugates the mapping L in a neighborhood of the fixed point $u = 0$ with the diffeomorphism f in a neighborhood of the fixed point $x = 0$ (to avoid unnecessary technical complications, we assume that h_0 conjugates the mapping L and the modified diffeomorphism f on the whole space $\mathbb{R}^n$).

Denote by

$$\psi(t, u) = \Psi(t)u = e^{Pt}u$$

the flow generated by the linear system (5.3) (in this case, $Lu = \psi(T, u)$).

Set

$$h(x) = \int_0^T \psi(-s, h_0(\phi(s, x)))\, ds.$$

Let us show that h is a homeomorphism that conjugates the flows of systems (5.1) and (5.3).

For any fixed $t \in \mathbb{R}$, the following equalities hold:

$$\begin{aligned} \psi(t, h(x)) &= \psi\left(t, \int_0^T \psi(-s, h_0(\phi(s, x)))\, ds\right) \\ &= e^{Pt} \int_0^T e^{-Ps} h_0(\phi(s, x))\, ds = \int_0^T e^{P(t-s)} h_0(\phi(s, x))\, ds \\ &= \int_0^T \psi(t - s, h_0(\phi(s - t, \phi(t, x))))\, ds. \end{aligned} \tag{5.9}$$

We have applied the group property of autonomous systems of differential equations:

$$\phi(s, x) \equiv \phi(s - t, \phi(t, x)).$$

The change of variables $\sigma = s - t$ on the right in equality (5.9) gives us the following relations:

$$\int_{-t}^{T-t} \psi(-\sigma, h_0(\phi(\sigma, \phi(t,x)))) \, d\sigma$$
$$= \int_{-t}^{0} \psi(-\sigma, h_0(\phi(\sigma, \phi(t, x_0)))) \, d\sigma + \int_{0}^{T-t} \psi(-\sigma, h_0(\phi(\sigma, \phi(t, x_0)))) \, d\sigma. \tag{5.10}$$

Let us transform the integrand in (5.10):

$$\begin{aligned} \psi(-\sigma, h_0(\phi(\sigma, \cdot))) &= \psi(-\sigma - T, \psi(T, h_0(\phi(\sigma, \cdot)))) \\ &= \psi(-\sigma - T, h_0(\phi(T, \phi(\sigma, \cdot)))). \end{aligned}$$

We have used the equality

$$\psi(T, h_0(\cdot)) = h_0(\phi(T, \cdot)).$$

Set $\tau = T + \sigma$ in the first integral on the right in equality (5.10) and note that

$$\begin{aligned} \int_{-t}^{0} \psi(-\sigma, h_0(\phi(\sigma, \cdot))) \, d\sigma &= \int_{-t}^{0} \psi(-\sigma - T, h_0(\varphi(T + \sigma, \cdot))) \, d\sigma \\ &= \int_{T-t}^{T} \psi(-\tau, h_0(\phi(\tau, \cdot))) \, d\tau. \end{aligned} \tag{5.11}$$

Substituting $\phi(t,x)$ into equalities (5.11), we conclude that

$$\begin{aligned} \int_0^T \psi(t - s, h_0(\phi(s - t, \phi(t,x)))) \, ds &= \int_0^T \psi(-s, h_0(\phi(s, \phi(t,x)))) \, ds \\ &= h(\varphi(t,x)). \end{aligned}$$

Thus, relation (5.9) implies that $\psi(t, \cdot) \circ h = h \circ \phi(t, \cdot)$. To prove that h is a homeomorphism, one can apply literally the same reasoning as in the proof of Theorem 4.1.

Reduction of Theorem 5.2 to Theorem 4.2 is left as an exercise for the reader.

Let p be a hyperbolic rest point of system (5.1). The same reasoning as in Section 4.3 gives us the following description of the behavior of trajectories in a small neighborhood of the point p.

We say that the point p is *attracting* if all the eigenvalues of the matrix P satisfy the inequalities $\operatorname{Re} \lambda_j < 0$. In this case, the point p has a neighborhood U such that if $x \neq p$ is an arbitrary point of U, then $\phi(t,x) \to p$ as $t \to \infty$, and the trajectory $\phi(t,x)$ leaves the neighborhood U as t decreases.

Let us note that this statement is close to the famous Lyapunov theorem on asymptotic stability by the first approximation.

We say that the point p is *repelling* if all the eigenvalues of the matrix P satisfy the inequalities $\operatorname{Re}\lambda_j > 0$. In this case, the point p has a neighborhood U such that if $x \neq p$ is an arbitrary point of U, then $\phi(t,x) \to p$ as $t \to -\infty$, and the trajectory $\phi(t,x)$ leaves the neighborhood U as t increases.

Finally, we say that p is a *saddle rest point* if there a natural number $m \neq 0, n$ such that $\operatorname{Re}\lambda_j < 0$, $j = 1,\dots,m$, and $\operatorname{Re}\lambda_j > 0$, $j = m+1,\dots,n$.

In this case, there exists a neighborhood U of the point p and topological disks $W_U^s(p)$ and $W_U^u(p)$ of dimension $m \times m$ and $(n-m)\times(n-m)$, respectively, such that

- if $x \in W_U^s(p)$, then $\phi(t,x) \to p$ as $t \to \infty$, and the trajectory $\phi(t,x)$ leaves the neighborhood U as t decreases;
- if $x \in W_U^u(p)$, then $\phi(t,x) \to p$ as $t \to -\infty$, and the trajectory $\phi(t,x)$ leaves the neighborhood U as t increases.

The disks $W_U^s(p)$ and $W_U^u(p)$ are called the *local stable* and *local unstable manifolds* of the rest point p, respectively; by Theorem 5.2, these disks are smooth (of class C^1).

Similarly to the case of a diffeomorphism, we define the *stable manifold* of the rest point p,

$$W^s(p) = \{x : \phi(t,x) \to p,\ t \to \infty\}$$

and the *unstable manifold* of the rest point p,

$$W^u(p) = \{x : \phi(t,x) \to p,\ t \to -\infty\}.$$

The same reasoning as in Lemma 4.3 shows that the sets $W^s(p)$ and $W^u(p)$ are invariant with respect to the flow ϕ and that the inclusion $x \in W^s(p)$ (the inclusion $x \in W^u(p)$) holds if and only if the intersection $\phi(t,x) \cap W_U^s(p)$ is nonempty (respectively, the intersection $\phi(t,x) \cap W_U^u(p)$ is nonempty).

Repeating the proof of Lemma 4.5, one can show that the sets $W^s(p)$ and $W^u(p)$ are images of Euclidean spaces of the corresponding dimension under topological immersions.

5.2 Hyperbolic closed trajectory

Consider system (5.1); to simplify notation, we assume that $x \in \mathbb{R}^{n+1}$. As above, we denote by ϕ the flow generated by system (5.1).

Assume that system (5.1) has a nonconstant periodic solution ψ; let $\omega > 0$ be the minimal period of the solution ψ and let γ be the closed trajectory of the flow ϕ corresponding to the solution ψ.

We fix coordinates in $\mathbb{R}^{n+1}$ so that $0 \in \gamma$ and $F(0) = (0,\dots,0,a)$, where $a > 0$.

Denote by S the coordinate hyperplane $\{x_{n+1} = 0\}$. By our choice of coordinates, the vector $F(0)$ is orthogonal to the hyperplane S. It follows from Theorem 1.1 that there exists a diffeomorphism T of a neighborhood of the origin in S to a neighborhood of the origin in S generated by shift along trajectories of the flow ϕ (the Poincaré diffeomorphism of the closed trajectory γ).

It was shown in the proof of Theorem 1.1 that two functions $t(s)$ and $\sigma(s)$ of class C^1 are defined for $s = (x_1, \dots, x_n) \in S$ with $|s|$ small enough; these functions have the following properties:

$$t(0) = \omega, \quad \sigma(0) = 0, \quad T(s) = \sigma(s) = \phi(t(s), (s, 0)) \cap S.$$

We say that the closed trajectory γ is *hyperbolic* if $0 \in S$ is a hyperbolic fixed point of the diffeomorphism T.

Let ϕ^n be the vector of the first n coordinates of the flow ϕ. Then

$$T(s) = \phi^n(t(s), (s, 0)).$$

It follows from our choice of coordinates that

$$\frac{\partial \phi^n}{\partial t}(\omega, 0) = (0, \dots, 0) \quad \text{and} \quad DT(0) = \frac{\partial T}{\partial s}(0) = \frac{\partial \phi^n}{\partial s}(\omega, 0).$$

Thus, the matrix $DT(0)$ is the matrix of the first n coordinates of the solution ϕ with respect to the first n coordinates of the initial point calculated at $t = \omega$ and $s = 0$.

It is known from the basic course of differential equations that the derivative of a solution with respect to initial values is calculated as follows.

We consider the variational system along the solution ψ,

$$\dot{y} = \frac{\partial F}{\partial x}(\phi(t, 0))y = \frac{\partial \phi F}{\partial x}(\psi(t))y; \tag{5.12}$$

let $\Phi(t)$ be the fundamental matrix of this system such that $\Phi(0) = E$ (recall that we denote by E the unit matrix of any dimension). Then

$$\frac{\partial \phi}{\partial x}(\omega, 0) = \Phi(\omega).$$

The eigenvalues of the matrix $\Phi(\omega)$ are called the *multipliers* of the closed trajectory γ.

Note that the vector-function $\xi(t) = \dot{\psi}(t)$ is a solution of system (5.12). Indeed, let us differentiate the equality $\dot{\psi}(t) = F(\psi(t))$:

$$\dot{\xi} = \frac{\partial F}{\partial x}(\psi(t))\xi.$$

Hence,

$$\xi(t) = \Phi(t)\xi(0).$$

Since

$$\xi(0) = \dot{\psi}(0) = F(0) \quad \text{and} \quad \xi(\omega) = F(\psi(\omega)) = F(0) = \xi(0),$$

the following equality holds:

$$F(0) = \Phi(\omega) F(0).$$

Hence, the matrix $\Phi(\omega)$ always has an eigenvalue 1 corresponding to the eigenvector $F(0)$. This eigenvalue is called the *standard multiplier*.

The equality

$$\Phi(\omega) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a \end{pmatrix}$$

implies that the last column of the matrix $\Phi(\omega)$ has the form

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Hence,

$$\frac{\partial \phi}{\partial x}(\omega, 0) = \Phi(\omega) = \begin{pmatrix} & 0 \\ \frac{\partial \phi^n}{\partial s}(\omega, 0) & \vdots \\ & 0 \\ \cdots & 1 \end{pmatrix} = \begin{pmatrix} & 0 \\ DT(0) & \vdots \\ & 0 \\ \cdots & 1 \end{pmatrix}.$$

It follows that the eigenvalues of the matrix $DT(0)$ are precisely the multipliers of the closed trajectory γ that differ from the standard unit multiplier.

Thus, we get the following equivalent form of the definition of hyperbolicity for a closed trajectory: For all its multipliers different from the standard unit one, the absolute values differ from 1.

In our reasoning above, the transverse section S to the closed trajectory γ has been chosen in a special way. Let us show that the eigenvalues of the derivative of the Poincaré diffeomorphism at the origin of a transverse section do not depend on the choice of the section.

Let S_1 and S_2 be two transverse sections, and let T_1 and T_2 be the corresponding Poincaré diffeomorphisms. Denote by χ the diffeomorphism of a neighborhood of the point $\gamma \cap S_1$ in the section S_1 to a neighborhood of the point $\gamma \cap S_2$ in the section S_2 generated by the shift along trajectories of the flow ϕ.

Take a point $x \in S_1$; let $\chi(x) = x_1 \in S_2$, $T_1(x) = x_2 \in S_1$, and $\chi(x_2) = x_3 \in S_2$. The equality $x_3 = T_2(x_1)$ implies that

$$\chi(T_1(x)) = T_2(\chi(x))$$

for points $x \in S_1$ that are close to $\gamma \in S_1$. Hence, $T_1(x) = \chi^{-1}(T_2(\chi(x)))$, and

$$DT_1(0) = G^{-1} DT_2(0) G, \tag{5.13}$$

where $G = D\chi(0)$.

Relation (5.13) implies that the matrices $DT_1(0)$ and $DT_2(0)$ are conjugate; hence, their spectra coincide.

Let us apply results of Section 4.3 to the hyperbolic point $0 \in S$ of the diffeomorphism T. Assume that the eigenvalues λ_j of the matrix $DT(0)$ satisfy relations (4.2).

By Theorems 4.1 and 4.2, the point $0 \in S$ is contained in two smooth disks D^s and D^u that belong to S and have the following properties:

- the dimensions of the disks D^s and D^u equal m and $n - m$, respectively;
- if $x \in D^s$, then $T^k(x) \in D^s$ as $k \geq 0$ and $T^k(x) \to 0, k \to \infty$;
- if $x \in D^u$, then $T^k(x) \in D^u$ as $k \leq 0$ and $T^k(x) \to 0, k \to -\infty$;
- there exists a neighborhood U of the point 0 in S such that if

  $$x \in U \setminus (D^s \cup D^u),$$

  then the trajectory $T^k(x)$ leaves U as $|k|$ grows.

Consider a point $x \in D^s$ and its trajectory $\phi(t, x)$. The disk D^s is positively invariant under the diffeomorphism T; it follows from the definition of T that, after a turn around γ, the trajectory $\phi(t, x)$ hits the disk D^s at the point $T(x)$.

Let us consider segments of trajectories with endpoints $x \in D^s$ and $T(x)$ for all points $x \in D^s$; the union of such segments is an $(m + 1)$-dimensional manifold. We denote this manifold by $W_l^s(\gamma)$ and call it the *local stable manifold* of the hyperbolic closed trajectory γ. Clearly, the manifold $W_l^s(\gamma)$ is positively invariant with respect to the flow ϕ, i.e., $\phi(t, x) \in W_l^s(\gamma)$ for $x \in W_l^s(\gamma)$ and $t \geq 0$.

Similarly (considering negative semitrajectories of points of the disk D^u), we construct the $(n - m + 1)$-dimensional *local unstable manifold* $W_l^u(\gamma)$ of the hyperbolic closed trajectory γ; this manifold is negatively invariant with respect to the flow ϕ.

We define the *stable manifold* of the hyperbolic closed trajectory γ by the equality

$$W^s(\gamma) = \{x : \phi(t, x) \to \gamma,\ t \to \infty\}$$

and the *unstable manifold* of the hyperbolic closed trajectory γ by the equality

$$W^u(\gamma) = \{x : \phi(t, x) \to \gamma,\ t \to -\infty\}.$$

In these definitions, the relation $\phi(t, x) \to \gamma$ means that

$$\mathrm{dist}(\phi(t, x), \gamma) \to 0.$$

The same reasoning as in Lemma 4.3 shows that the sets $W^s(\gamma)$ and $W^u(\gamma)$ are invariant with respect to the flow ϕ.

Let us show that the inclusion $x \in W^s(\gamma)$ (the inclusion $x \in W^u(\gamma)$) holds if and only if the intersection $\phi(t, x) \cap W_l^s(\gamma)$ (respectively, the intersection $\phi(t, x) \cap W_l^u(\gamma)$) is nonempty. We consider the case of $W^s(\gamma)$; for $W^u(\gamma)$, the reasoning is similar.

If $\phi(t, x) \cap W_l^s(\gamma) \neq \emptyset$, then $\phi(\tau, x) \in W_l^s(\gamma)$ for some $\tau \in \mathbb{R}$. Our construction of $W_l^s(\gamma)$ implies that the point $\phi(\tau, x)$ belongs to a segment of a trajectory of the flow ϕ whose endpoints lie in D^s. Thus, there exists a number $\theta \geq 0$ such that $y := \phi(\theta, x) \in D^s$.

Take an arbitrary $\varepsilon > 0$. The theorem on continuous dependence of the solution on its initial point implies that there exists a $\delta > 0$ such that if $|z| < \delta$, then

$$|\phi(t, z) - \psi(t)| < \varepsilon, \quad t \in [0, 2\omega] \tag{5.14}$$

(recall that ω is the period of the solution ψ and $\psi(0) = 0$).

It was shown in the proof of Theorem 1.1 that if $z \in D^s$, then $T(z) = \phi(t(z), z) \cap S$, where $t(z) \to \omega$ as $z \to 0$.

Since $T^k(y) \to 0$ as $k \to \infty$, there exists an index m such that $|T^k(y)| < \delta$ and $0 < t(T^k(y)) < 2\omega$ for $k \geq m$.

If $z = T^k(y), k \geq m$, then relation (5.14) implies that

$$T(z) = \phi(t(z), z) \cap S \subset D^s$$

and

$$\mathrm{dist}(\phi(t, z), \gamma) \leq |\phi(t, z) - \psi(t)| < \varepsilon, \quad t \in [0, t(z)].$$

Hence,

$$\mathrm{dist}(\phi(t, z), \gamma) < \varepsilon, \quad t \geq 0.$$

Since ε is arbitrary, we conclude that

$$\mathrm{dist}(\phi(t, y), \gamma) \to 0, \quad t \to \infty. \tag{5.15}$$

Since $y = \phi(\theta, x)$, relation (5.15) implies that

$$\mathrm{dist}(\phi(t, x), \gamma) \to 0, \quad t \to \infty. \tag{5.16}$$

Now we assume that $x \in W^s(\gamma)$, i.e., relation (5.16) holds.

By the Remark after Theorem 1.1, every point $y \in \gamma$ has a neighborhood U_y such that if $z \in U_y$, then the positive semitrajectory of the point z intersects the transverse section S at a point belonging to the domain of definition of the Poincaré diffeomorphism T. Set

$$U = \bigcup_{y \in \gamma} U_y.$$

There exists a number $\tau \in \mathbb{R}$ such that

$$\phi(t, x) \in U, \quad t \geq \tau.$$

Consider the point $y = \phi(\tau, x)$. It follows from the choice of τ that the positive semitrajectory of the point y intersects the transverse section S at points $z_1, z_2, \dots$ such that $z_{k+1} = T(z_k)$. Properties of the section S listed above imply that $z_1 \in D^s$. Clearly,

$$z_1 \in \phi(t, x) \cap W_l^s(\gamma).$$

When one constructs stable and unstable manifolds of a hyperbolic closed trajectory of a flow, the following phenomenon is possible. Let us explain it by giving a simple example. Consider system (5.1) in the space $\mathbb{R}^3$ with coordinates (x, y, z). Assume that the plane $S = \{x = 0\}$ is a transverse section for a closed trajectory γ and that the corresponding local Poincaré diffeomorphism T has the form

$$(y, z) \mapsto \left(\frac{-y}{2}, -2z\right)$$

for small $|y|$ and $|z|$.

In this case (under a proper choice of the neighborhood U), the disk D^s is a segment of the axis $\{x = z = 0\}$. Take a small oriented segment l of this axis such that the origin is one of the endpoints of this segment. Clearly, T maps l to a segment of the same axis with the opposite orientation. Thus, the appearing local stable manifold $W_l^s(\gamma)$ of the closed trajectory γ is a Möbius band for which γ is a middle circle.

One can show that the stable and unstable manifolds $W^s(\gamma)$ and $W^u(\gamma)$ of a hyperbolic closed trajectory γ are images under topological immersions either of cylinders $\mathbb{R}^m \times S^1$ and $\mathbb{R}^{n-m} \times S^1$ or of nonorientable foliations over S^1 with leaves $\mathbb{R}^m$ and $\mathbb{R}^{n-m}$.

Chapter 6
Transversality

6.1 Transversality of mappings and submanifolds

Let K and M be smooth manifolds, let L be a submanifold of the manifold M, and let f be a smooth mapping of K into M.

We say that the mapping f is *transverse* to L at a point $x \in K$ if either $f(x) \notin L$ or

$$Df(x)T_xK + T_{f(x)}L = T_{f(x)}M. \tag{6.1}$$

The left-hand side of equality (6.1) is the sum of two linear subspaces of the space $T_{f(x)}M$, the image of T_xK under the action of the derivative of f at the point x and the tangent space $T_{f(x)}L$.

Let us note that, in general,

$$Df(x)T_xK \neq T_{f(x)}f(K)$$

(consider the following example: $K = \mathbb{R}$ with coordinate t, $M = \mathbb{R}^2$ with coordinates (x, y), and $f(t) = (0, t^3)$; in this case,

$$Df(0)R \neq T_{(0,0)}f(K)).$$

We say that the mapping f is *transverse* to L if f is transverse to L at every point $x \in K$.

If K and L are two submanifolds of the same manifold M, we say that K is *transverse* to L at a point $x_0 \in K$ if the mapping Id that takes any point $x \in K$ to the same point $x \in M$ is transverse to L at x_0.

In other words, K is transverse to L at a point $x_0 \in K \cap L$ if

$$T_{x_0}K + T_{x_0}L = T_{x_0}M, \tag{6.2}$$

i.e.,

$$T_{x_0}M = \{a + b : a \in T_{x_0}K, b \in T_{x_0}L\}.$$

We say that submanifolds K and L are *transverse* if the above-mentioned mapping Id is transverse to L.

One can reformulate condition (6.2) as follows:

$$\dim K + \dim L - \dim(T_{x_0}K \cap T_{x_0}L) = \dim M. \tag{6.3}$$

This statement is a corollary of the following simple algebraic lemma.

Lemma 6.1. *If X_1 and X_2 are two linear subspaces of a finite-dimensional linear space X, then*

$$\dim X_1 + \dim X_2 - \dim(X_1 \cap X_2) = \dim(X_1 + X_2). \tag{6.4}$$

Proof. Consider the linear operator $A : X_1 \times X_2 \to X$ defined by the formula

$$A(x_1, x_2) = x_1 + x_2.$$

Let R and K be the range and kernel of the operator A, respectively.

Clearly,

$$K = \{(x, -x) : x \in X_1 \cap X_2\}.$$

Hence, $\dim K = \dim(X_1 \cap X_2)$.

Since $R = X_1 + X_2$, equality (6.4) is a corollary of the relation

$$\dim R + \dim K = \dim(X_1 \times X_2) = \dim X_1 + \dim X_2. \qquad \square$$

Equality (6.3) follows from equality (6.4) in which we take $X_1 = T_{x_0}K$ (in this case, $\dim X_1 = \dim K$) and $X_2 = T_{x_0}L$ (in this case, $\dim X_2 = \dim L$) since $X_1 + X_2 = T_{x_0}M$ by (6.2).

Let us note that the following simple statement holds (this statement is very important for us).

Lemma 6.2. *If K and L are two submanifolds of a manifold M, f is a diffeomorphism of M, and K and L are transverse at a point $x_0 \in K \cap L$, then $f(K)$ and $f(L)$ are transverse at the point $f(x_0)$.*

Proof. Since equality (6.2) holds, there exists a basis of the tangent space $T_{x_0}M$ such that every vector of this basis belongs either to $T_{x_0}K$ or to $T_{x_0}K$. Since f is a diffeomorphism, its derivative $Df(x_0)$ is a linear isomorphism between the spaces $T_{x_0}M$ and $T_{f(x_0)}M$. Hence, $Df(x_0)$ takes the above-mentioned basis to a basis of the space $T_{f(x_0)}M$ such that every vector of the latter basis belongs either to $T_{f(x_0)}K$ or to $T_{f(x_0)}L$. $\square$

Corollary. *If K and L are two submanifolds of a manifold M, ϕ is a flow on M generated by a smooth vector field, K and L are transverse at a point $x_0 \in K \cap L$, and $T \neq 0$, then $\phi(T, K)$and $\phi(T, L)$ are transverse at the point $\phi(T, x_0)$.*

Proof. As was shown in Section 1.2, $\phi(T, \cdot)$ is a diffeomorphism of M. $\square$

We formulate one more important statement without a proof (the reader can find a proof in [16]).

Let K and L be two smooth disks in a manifold M with Riemannian metric dist. Assume that the disks K and L are images of disks $D_1 \in \mathbb{R}^k$ and $D_2 \in \mathbb{R}^l$ under smooth embeddings h_1 and h_2, respectively.

Lemma 6.3. *Assume that the disks K and L have a common point x_0 at which they are transverse. For any $\varepsilon > 0$ there exists a $\delta > 0$ such that if H_1 and H_2 are embeddings of the disks D_1 and D_2 into M for which the inequalities*

$$\rho_1(h_i, H_i) < \delta, \quad i = 1, 2,$$

hold, then the disks $H_1(D_1)$ and $H_2(D_2)$ have a point x of transverse intersection such that

$$\operatorname{dist}(x, x_0) < \varepsilon.$$

6.2 Transversality condition

Let p and q be two hyperbolic fixed points of a diffeomorphism f of a manifold M (we do not exclude the case where $p = q$). Assume that a point x is a point of intersection of the unstable manifold $W^u(p)$ and the stable manifold $W^s(q)$.

It was noted above that the unstable manifold $W^u(p)$ and the stable manifold $W^s(q)$ are not necessarily submanifolds of the manifold M. At the same time, it follows from Theorem 4.2 (applied both to f and to f^{-1}) that there exist neighborhoods U and V of the points p and q, respectively, such that the local unstable manifold $W^u_U(p)$ and the local stable manifold $W^s_V(q)$ are smooth disks.

Let us first assume that $x \neq p$ and $x \neq q$. By Lemma 4.3, there exist indices m and n such that $y = f^m(x) \in W^u_U(p)$ and $z = f^n(x) \in W^s_V(q)$. Fix smooth disks $K \subset W^u_U(p)$ and $L \subset W^s_V(q)$ such that $y \in K$, $z \in L$, $\dim K = \dim W^u(p)$, and $\dim L = \dim W^s(q)$.

We say that the unstable manifold $W^u(p)$ and the stable manifold $W^s(q)$ are *transverse* at the point x if the disks $f^{-m}(K)$ and $f^{-n}(L)$ are transverse at x.

Lemma 6.2 implies that the introduced property of transversality of $W^u(p)$ and $W^s(q)$ at the point x does not depend on the choice of the indices m and n and the disks K and L; in addition, $W^u(p)$ and $W^s(q)$ are transverse at a point x if and only if $W^u(p)$ and $W^s(q)$ are transverse at the point $f^l(x)$ for some (for any) $l \in \mathbb{Z}$.

If $x = p$ (then $x = q$ as well), the smooth disks $W^u_U(p)$ and $W^s_U(p)$ are transverse at the point p (the tangent spaces of these disks at p are complementary linear subspaces of the space T_pM); in this case, it is natural to consider p as a point at which $W^u(p)$ and $W^s(p)$ are transverse.

Similarly one defines transversality of stable and unstable manifolds for hyperbolic periodic points (recall that if p is a periodic point of a diffeomorphism f of period m, then, by definition, its stable manifold is the stable manifold of the fixed point p of the diffeomorphism f^m) and also of stable and unstable manifolds for hyperbolic rest points and closed trajectories of flows generated by autonomous systems of differential equations.

Let us mention a simple corollary of the transversality condition.

Lemma 6.4. (1) *If p and q are hyperbolic periodic points of a diffeomorphism such that the manifolds $W^u(p)$ and $W^s(q)$ have a common point at which they are transverse, then*

$$\dim W^u(p) \geq \dim W^u(q). \tag{6.5}$$

(2) *If p and q are hyperbolic rest points or closed trajectories of a smooth flow ϕ such that the manifolds $W^u(p)$ and $W^s(q)$ have a common point x_0 at which they are transverse and x_0 is not a rest point, then inequality* (6.5) *holds. In this case,* (6.5) *may turn into an equality only if q is a closed trajectory.*

Proof. (1) If $\dim M = n$ and q is a hyperbolic periodic point of a diffeomorphism such that $\dim W^s(q) = m$, then $\dim W^u(q) = n - m$.

Assume that $W^u(p)$ and $W^s(q)$ intersect transversally at a point x_0.

Relation (6.3) with $K = W^u(p)$ and $L = W^s(q)$ implies that

$$\begin{aligned}\dim W^u(p) &= n - \dim W^s(q) - \dim(T_{x_0}W^u(p) \cap T_{x_0}W^s(q)) \\ &= n - m - \dim(T_{x_0}W^u(p) \cap T_{x_0}W^s(q)) \geq n - m = \dim W^u(q).\end{aligned}$$

(2) Let $W^u(p)$ and $W^s(q)$ have a common point x_0 at which they are transverse; assume that x_0 is not a rest point,

Since the manifolds $W^u(p)$ and $W^s(q)$ are invariant under the flow ϕ, the trajectory $\phi(t, x_0)$ belongs to both manifolds. Hence, the tangent vector $F(x_0) \neq 0$ of this trajectory at the point x_0 belongs to the intersection $T_{x_0}W^u(p) \cap T_{x_0}W^s(q)$. It follows that

$$\dim(T_{x_0}W^u(p) \cap T_{x_0}W^s(q)) \geq 1.$$

The same reasoning as above shows that

$$\dim W^u(p) \geq \dim M - \dim W^s(q) - 1.$$

To complete the proof, we note that if q is a rest point,then

$$\dim W^s(q) + \dim W^u(q) = \dim M,$$

and if q is a closed trajectory, then

$$\dim W^s(q) + \dim W^u(q) = \dim M + 1.$$ □

Statement (2) of Lemma 6.4 implies that if p is a hyperbolic rest point of a smooth flow, then the manifolds $W^u(p)$ and $W^s(p)$ have no points of transverse intersection different from p.

At the same time, the manifolds $W^u(p)$ and $W^s(p)$ may have points of transverse intersection different from p (in the case of a hyperbolic periodic point p of a diffeomorphism) or not belonging to the trajectory p (in the case of a hyperbolic closed trajectory p of a smooth flow).

Following Poincaré, such points are called *double-asymptotic* to p, or *homoclinic*. The behavior of trajectories in a neighborhood of the trajectory of a transverse homoclinic point is very complicated (see Section 9.3).

Let us define an important class of dynamical systems.

A diffeomorphism f of a smooth closed manifold M is called *Kupka–Smale* if all periodic points of f are hyperbolic, and the stable and unstable manifolds of periodic points are transverse at any point of intersection.

The following statement had been proven independently by I. Kupka and S. Smale (see also [9]).

Theorem 6.1. *The set* KS *of Kupka–Smale diffeomorphisms is residual in* $\text{Diff}^1(M)$.

Kupka–Smale flows are defined in a similar way: a flow ϕ generated by a smooth vector field is called *Kupka–Smale* if all rest points and closed trajectories of ϕ are hyperbolic, and the stable and unstable manifolds of rest points and closed trajectories are transverse at any point of intersection.

In the case of flows, a statement similar to Theorem 6.1 is valid.

6.3 Palis lemma

We will prove an important statement related to the transversality condition (this statement, usually called the λ-lemma, was proven by J. Palis).

Let p be a hyperbolic fixed point of a diffeomorphism f (since the statement which we prove is local, the cases of a manifold and Euclidean space are treated similarly).

For definiteness, we assume that p is the origin of the space $\mathbb{R}^n$ and that the diffeomorphism f is represented in the form (4.4) in a neighborhood U of the point $x = 0$, where the Jacobi matrix A of f at zero has the form described in Lemma 4.1, while the nonlinear term F vanishes at the origin together with its Jacobi matrix.

We consider coordinates $x = (y, z)$ related to the representation of the matrix A in block-diagonal form; let $\mathcal{S}$ and $\mathcal{U}$ be the coordinate subspaces corresponding to the parts of the spectrum of the matrix A inside and outside the unit disk, respectively.

Fix a number $\Delta > 0$ such that the local stable manifold $W_V^s(0)$ of the fixed point $x = 0$ in the neighborhood $V = \{|x| < \Delta\}$ has the structure described by Theorem 4.2.

Let γ be the mapping which determines $W_V^s(0)$.

Let us assume that the local unstable manifold $W_V^u(0)$ is determined by the mapping

$$\alpha : \{|z| < \Delta\} \to \mathcal{S},$$

with the following properties:

$$\alpha(0) = 0 \quad \text{and} \quad \frac{\partial \alpha}{\partial z}(0) = 0$$

(the existence of such a mapping follows from Theorem 4.2 applied to the diffeomorphism f^{-1}).

Lemma 6.5 (λ-lemma). *Let N be a smooth disk having a point of transverse intersection with $W_V^s(0)$. Then there exists a smooth disk $Q \subset W_V^u(0)$ containing the fixed point $x = 0$ (let the disk Q be the image of a smooth embedding h of the disk $D \subset \mathbb{R}^{n-m}$ into $\mathbb{R}^n$) and having the following property: For any $\varepsilon > 0$ there is an index $m(\varepsilon)$ such that for any $m \geq m(\varepsilon)$ there exists an embedding h_m of the disk D into $\mathbb{R}^n$ such that $h_m(Q) \subset f^m(N)$ and $\rho_1(h, h_m) < \varepsilon$.*

Proof. To prove Lemma 6.5, we perform a change of coordinates under which the local stable and unstable manifolds become disks that belong to the coordinate subspaces $\mathcal{S}$ and $\mathcal{U}$, respectively.

Since the statement on the existence of such a change of variables is of independent interest, we formulate it as a separate lemma.

Lemma 6.6. *There exist neighborhoods $V_1 \subset V$ and V_2 of the origin of $\mathbb{R}^n$ and a diffeomorphism H that maps V_1 onto V_2 and such that the images of $W_{V_1}^s(0)$ and $W_{V_1}^u(0)$ under H are disks belonging to the coordinate subspaces $\mathcal{S}$ and $\mathcal{U}$, respectively.*

Proof of Lemma 6.6. Consider the mapping

$$H : V \to \mathbb{R}^n$$

defined by the equalities

$$\xi = y - \alpha(z), \quad \eta = z - \gamma(y) \tag{6.6}$$

(here ξ and η are coordinates in $\mathcal{S}$ and $\mathcal{U}$, respectively).

It follows from properties of the mappings α and γ that

$$H(0) = 0 \quad \text{and} \quad DH(0) = \text{Id}.$$

By the Inverse Mapping Theorem, there exist neighborhoods $V_1 \subset V$ and V_2 of the origin in $\mathbb{R}^n$ such that H is a diffeomorphism of V_1 onto V_2; we may assume that the neighborhood V_1 is so small that the sets $W_{V_1}^s(0)$ and $W_{V_1}^u(0)$ are smooth disks.

Equalities (6.6) imply that the diffeomorphism H maps $W_{V_1}^s(0)$ and $W_{V_1}^u(0)$ to disks that belong to the coordinate subspaces $\mathcal{S}$ and $\mathcal{U}$, respectively. □

Let r be a point of transverse intersection of the disk N with $W_V^s(0)$. Since $f^k(r) \to 0$ as $k \to \infty$, there exists an index k_0 such that $r_0 = f^{k_0}(r) \in V_1$. By Lemma 6.2, the disk $f^{k_0}(N)$ is transverse to $W_{V_1}^s(0)$ at the point r_0. Clearly, the disk $f^{k_0}(N)$ contains a smaller disk N_0 that belongs to V_1, contains the point r_0, and is transverse to $W_{V_1}^s(0)$ at this point.

The same reasoning as in the proof of Lemma 6.2 shows that the disk $N' = H(N_0)$ is transverse to $\mathcal{S}$ at the point $\pi_0 = H(r_0)$.

Decreasing the neighborhood V_1, we may assume that the diffeomorphism H is defined on $f(V_1)$. Consider the mapping

$$g = H \circ f \circ H^{-1};$$

clearly, g maps V_2 diffeomorphically onto $H(f(V_1))$.

The equalities $DH(0) = E$ and $DH^{-1}(0) = E$ imply that

$$Dg(0) = DH(0)Df(0)DH^{-1}(0) = Df(0). \tag{6.7}$$

Hence, $x = 0$ is a hyperbolic fixed point of the diffeomorphism g.

We reduce the proof of Lemma 6.5 to the following auxiliary statement.

Lemma 6.7. *There exists a number $\delta > 0$ having the following property: For any $\varepsilon > 0$ one can find an index $m(\varepsilon)$ such that if $m \geq m(\varepsilon)$, then there exists a family of smooth disks $\nu_0, \dots, \nu_m$ such that*

$$\pi_0 \in \nu_0 \subset N', \quad \pi_k := g^k(\pi_0) \in \nu_k \subset g(\nu_{k-1}) \cap V_2, \quad 1 \leq k \leq m-1, \tag{6.8}$$

and the disk ν_m contains a smooth disk μ_m given by

$$y = \beta(z), \quad |z| < \delta,$$

where

$$|\beta(z)| < \varepsilon \quad \text{and} \quad \left\| \frac{\partial \beta}{\partial z} \right\| < \varepsilon \tag{6.9}$$

for $|z| < \delta$.

Let us show that Lemma 6.5 is a corollary of Lemma 6.7. Consider the smooth disks $N_k = H^{-1}(\nu_k)$, $1 \leq k \leq m$. Relations (6.8) imply that

$$H^{-1} \circ g|_{\nu_k} = f \circ H^{-1}|_{\nu_k};$$

hence,

$$N_{k+1} = H^{-1}(\nu_{k+1}) \subset H^{-1} \circ g(\nu_k) = f \circ H^{-1}(\nu_k) = f(N_k). \tag{6.10}$$

Denote by D the disk $\{y = 0, |z| < \delta\}$ (a ball in the space $\mathcal{U}$). Set $Q = H^{-1}(D)$. The disk Q is the image of D under the embedding H^{-1}, and the disk N_m is the image of D under the embedding $H^{-1} \circ b$, where $b : (0, z) \to (\beta(z), z)$.

Relations (6.10) imply that $N_m \subset f^m(N_0) \subset f^{m+k_0}(N)$. Since H^{-1} is a diffeomorphism and we can guarantee that inequalities (6.9) hold with arbitrarily small ε, the value $\rho_1(H^{-1}, H^{-1} \circ b)$ can be done arbitrarily small as well. Thus, the statement of Lemma 6.4 is fulfilled. □

Now let us prove Lemma 6.7. We consider the diffeomorphism g in a small neighborhood V of the origin (we reduce the neighborhood V several times to provide the properties of g which we need in the proof of Lemma 6.7). Equality (6.7) implies that

$$g(x) = Ax + G(x).$$

Let, as above, $A = \operatorname{diag}(B, C)$, and let g_1 and g_2 be the components of the nonlinearity G with respect to the representation $x = (y, z)$. If $x \in \mathcal{S}$, then $x = (y, 0)$ and $g(x) \in \mathcal{S}$. The equality

$$g(y, 0) = (By + g_1(y, 0), g_2(y, 0))$$

implies that $g_2(y, 0) = 0$. Hence,

$$g_2(y, 0) = 0 \quad \text{and} \quad \frac{\partial g_2}{\partial y}(y, 0) = 0, \quad (y, 0) \in V. \tag{6.11}$$

A similar reasoning shows that if the neighborhood V is small enough, then

$$g_1(0, z) = 0 \quad \text{and} \quad \frac{\partial g_1}{\partial z}(0, z) = 0, \quad (0, z) \in V. \tag{6.12}$$

Let

$$a_0 = \|B\| < 1 \quad \text{and} \quad \frac{1}{b_0} = \|C^{-1}\| < 1, \tag{6.13}$$

where the operator norms are induced by the Euclidean norm. One can repeat the proof of Lemma 4.1 to prove that there exist coordinates for which estimates (6.13) hold.

Take a number $k > 0$ for which the following inequalities are valid:

$$a = a_0 + k < 1, \quad b = b_0 - k > 1, \quad k < \frac{(b-1)^2}{8}. \tag{6.14}$$

Relations (6.11) and (6.12) imply that if the neighborhood V is small enough, then the following inequalities hold in V:

$$\left\| \frac{\partial g_i}{\partial (y, z)}(y, z) \right\| < k, \quad i = 1, 2. \tag{6.15}$$

As above, let us fix a disk N_0 that belongs to the intersection $f^{k_0}(N) \cap V'$ (where $V' = H^{-1}(V)$), contains the point r_0, and is transverse to $W^s_{V'}(0)$ at this point.

Let $\dim \mathcal{S} = m$. It is geometrically obvious that the disk $N' = H(N_0)$ contains an $(n-m)$-dimensional disk ν_0 such that the point $\pi_0 = H(r_0)$ lies in ν_0 and ν_0 is transverse to $\mathcal{S}$ at this point.

Consider a tangent vector v_0 of ν_0 at the point π_0 with $|v| = 1$. According to the representation $x = (y, z)$, we write $v = (v^s, v^u)$. Since the disk ν_0 is transverse to $\mathcal{S}$ at the point π_0, formula (6.3) implies that $v^u \neq 0$.

Define the inclination of the vector v to the subspace $\mathcal{S}$ by the formula

$$\lambda = \frac{|v^s|}{|v^u|}.$$

The main contents of the following proof is estimation of the inclinations of images of the vector v under the action of derivatives $Dg^k(\pi_0)$; this explains the term "λ-lemma".

Since the unit sphere of the tangent space$T_{\pi_0}\nu_0$ is compact, there exists a number $L > 0$ such that the inclination of any vector $v \in T_{\pi_0}\nu_0$ satisfies the inequality $\lambda \leq L$.

Let us write

$$Dg(\pi_0) = \begin{pmatrix} B + \frac{\partial g_1}{\partial y} g_1(\pi_0) & \frac{\partial g_1}{\partial z}(\pi_0) \\ \frac{\partial g_2}{\partial z}(\pi_0) & C + \frac{\partial g_2}{\partial z}(\pi_0) \end{pmatrix}.$$

Since $\pi_0 \in \mathcal{S}$, relations (6.11) imply that

$$\frac{\partial g_2}{\partial y}(\pi_0) = 0.$$

If $v_1 = Dg(\pi_0)v = (v_1^s, v_1^u)$, then

$$v_1^s = \left(B + \frac{\partial g_1}{\partial y}(\pi_0)\right) v^s + \frac{\partial g_1}{\partial z}(\pi_0) v^u \quad \text{and} \quad v_1^u = \left(C + \frac{\partial g_2}{\partial z}(\pi_0)\right) v^u. \tag{6.16}$$

It follows from inequalities (6.15) and (6.13) that

$$|v_1^s| \leq (a_0 + k)|v^s| + k|v^u| = a|v^s| + k|v^u|. \tag{6.17}$$

Let us apply the operator C^{-1} to the second equality in (6.16):

$$v^u + C^{-1}\frac{\partial g_2}{\partial z}(\pi_0) = C^{-1} v_1^u.$$

Clearly, the following estimates hold:

$$\frac{1}{b_0}|v^u| \geq \|C^{-1}\||v_1^u| \geq |v^u| - \left|C^{-1}\frac{\partial g_2}{\partial z}(\pi_0)v^u\right| \geq |v^u| - \frac{k}{b_0}|v^u|$$

and

$$|v_1^u| \geq (b_0 - k)|v^u| = b|v^u|. \tag{6.18}$$

Let us estimate the inclination λ_1 of the vector v_1 in terms of the inclination λ_0 of the vector v using inequalities (6.17) and (6.18):

$$\lambda_1 = \frac{|v_1^s|}{|v_1^u|} \leq \frac{a|v^s| + k|v^u|}{b|v^u|} \leq \frac{a}{b}\lambda_0 + \frac{k}{b}. \tag{6.19}$$

Denote by $\lambda_2, \dots, \lambda_m$ the inclinations of the vectors $v_2 = Dg^2(\pi_0)v, \dots, v_m = Dg^m(\pi_0)v$. Iterating estimate (6.19), we get the following chain of inequalities:

$$\lambda_2 \leq \left(\frac{a}{b}\right)^2 \lambda_0 + \frac{k}{b^2} + \frac{k}{b},$$

$$\vdots$$

$$\lambda_m \leq \left(\frac{a}{b}\right)^m \lambda_0 + \frac{k}{b^m} + \dots + \frac{k}{b} \leq \left(\frac{a}{b}\right)^m \lambda_0 + \frac{k}{b-1}. \tag{6.20}$$

Find an index m_1 such that

$$\left(\frac{a}{b}\right)^{m_1} L < \frac{b-1}{8}.$$

Relations (6.14) and (6.20) imply that if $m \geq m_1$, then

$$\lambda_m \leq \frac{b-1}{4}, \tag{6.21}$$

where λ_m is the inclination of the vector $Dg^m(\pi_0)v$ for any vector $v \in T_{\pi_0}\nu_0$.

We construct disks $\nu_1, \dots, \nu_{m_1}$ applying the following procedure. Assume that a disk ν_k with $0 \leq k \leq m_1 - 1$ is constructed; we take a small disk ν_{k+1} belonging to $g(\nu_k) \cap V$ and containing the point π_{k+1}.

Since ν_{m_1} is a smooth disk containing the point π_{m_1} and inequalities (6.21) hold, we can select in ν_{m_1} a subdisk (denoted again ν_{m_1}) with the following properties: ν_{m_1} contains the point π_{m_1}, is transverse to $\mathcal{S}$ at this point, and the inclination λ of any unit tangent vector $v \in T_x\nu_{m_1}$ at any point ν_{m_1} satisfies the inequality

$$\lambda < \frac{b-1}{2}. \tag{6.22}$$

(Let us note that now we consider inclinations of tangent vectors at all points of the disks, and not only at points belonging to $\mathcal{S}$.)

Set

$$b_1 = \frac{b+1}{2}.$$

Find a number $\delta > 0$ such that if $D = \{(0, z) : |z| < \delta\}$, then the closed disk $\text{Cl}\, D$ belongs to the neighborhood V. Relations (6.12) imply that

$$\frac{\partial g_1}{\partial z} = 0$$

at points of the disk D.

Fix an arbitrary $\varepsilon > 0$; we assume that ε is small enough, so that the following inequalities hold:

$$b_2 := \frac{b_1}{\sqrt{1+\varepsilon^2}} > 1 \tag{6.23}$$

and

$$\left(\frac{a}{b_1}\right)\left(\frac{b-1}{2}\right) + \frac{\varepsilon(b_1 - 1)}{2b_1} < \frac{b-1}{2} \tag{6.24}$$

(we take into account that $a < b_1$).

Fix a neighborhood $V_0 \subset V$ of D in which the following inequality is satisfied:

$$\left\| \frac{\partial g_1}{\partial z} \right\| < k_1 := \frac{\varepsilon(b_1 - 1)}{2}. \tag{6.25}$$

Without loss of generality, we may assume that the index m_1 chosen above is large enough, so that $\pi_{m_1} \in V_0$.

We decrease the disk once more ν_{m_1} and take its subdisk (denoted again ν_{m_1}) such that this subdisk belongs to V_0, contains the point π_{m_1}, and is transverse to $\mathcal{S}$ at this point.

Take an arbitrary point $x \in \nu_{m_1}$ and an arbitrary unit tangent vector v in $T_x \nu_{m_1}$. Represent v in the form (v^s, v^u) and denote by λ the inclination of the vector v.

Let us estimate the number λ_1, the inclination of the vector $v_1 = (v_1^s, v_1^u) = Dg(x)v$. The equalities

$$v_1^s = \left(B + \frac{\partial g_1}{\partial y}(x)\right) v^s + \frac{\partial g_1}{\partial z}(x) v^u$$

and

$$v_1^u = \frac{\partial g_2}{\partial y}(x) v^s + \left(C + \frac{\partial g_2}{\partial z}(x)\right) v^u$$

and estimates (6.14), (6.15), and (6.25) imply the following inequalities (as above, we multiply the second equality by C^{-1} from the left):

$$|v_1^s| \le a|v^s| + k_1|v^u|,$$

$$\frac{1}{b_0}|v^u| \ge \|C^{-1}\| |v_1^u| \ge \left(1 - \frac{k}{b_0}\right)|v^u| - \frac{k}{b_0}|v^s|,$$

and

$$|v^u| \geq b|v^u| - k|v^s|. \tag{6.26}$$

Hence,

$$\lambda_1 \leq \frac{a|v^s| + k_1|v^u|}{b|v^u| - k|v^s|} = \frac{a\lambda + k_1}{b - k\lambda}.$$

Since $0 < k < 1$, inequality (6.22) implies that

$$b - k\lambda > b - \frac{b-1}{2} = b_1. \tag{6.27}$$

Taking into account inequality (6.24) and the definition of k_1, we get the inequalities

$$\lambda_1 \leq \frac{a}{b_1}\lambda + \frac{k_1}{b_1} < \frac{b-1}{2}.$$

Assume that a point $x \in v_{m_1}$ is such that $g(x), \dots, g^{m-1}(x) \in V_0$.

Let $\lambda_1, \dots, \lambda_m$ be the inclinations of the vectors $v_1 = Dg(x)v, \dots, v_m = Dg^m(x)v$. Iterating the above estimate, we obtain the inequality

$$\lambda_m \leq \left(\frac{a}{b_1}\right)^m \lambda + \frac{k_1}{b_1^m} + \dots + \frac{k_1}{b_1} \leq \left(\frac{a}{b_1}\right)^m \lambda + \frac{k_1}{b_1 - 1}. \tag{6.28}$$

Inequalities (6.25) and (6.28) imply that there exists an index m_2 such that $\lambda_m < \varepsilon$ for $m \geq m_2$.

If $m \geq m_2$, $g^m(x) \in V_0$, and $v_{m+1} = Dg(g^m(x))v_m$, then

$$\frac{|v_{m+1}|}{|v_m|} = \frac{\sqrt{(v^s_{m+1})^2 + (v^u_{m+1})^2}}{\sqrt{(v^s_m)^2 + (v^u_m)^2}} = \frac{|v^u_{m+1}|}{|v^u_m|}\frac{\sqrt{1 + \lambda^2_{m+1}}}{\sqrt{1 + \lambda^2_m}}. \tag{6.29}$$

It follows from relations (6.26) and (6.28) that

$$|v^u_{m+1}| \geq b_1|v^u_m|.$$

In addition,

$$\frac{\sqrt{1 + \lambda^2_{m+1}}}{\sqrt{1 + \lambda^2_m}} \geq \frac{1}{\sqrt{1 + \lambda^2_m}} \geq \frac{1}{\sqrt{1 + \varepsilon^2}}.$$

Hence, equalities (6.29) imply the estimate

$$|v_{m+1}| \geq b_2|v_m|.$$

Thus, if we construct successively disks ν_{m_1+k+1} that belong to the intersections of the images $g(\nu_{m_1+k})$ with the neighborhood V_0, then, for $k \geq m_2$, these disks are uniformly expanded under the action of diffeomorphism g. At the same time, the inclination of tangent spaces of such disks to the subspace $\mathcal{U}$ is uniformly small (the smallness of the inclination is determined by the number ε). Hence, if m is large enough, then projection onto $\mathcal{U}$ parallel to $\mathcal{S}$ defines mappings

$$\beta_m : D \to \nu_m.$$

The Grobman–Hartman theorem implies that if the neighborhood V_0 is small enough, then the diffeomorphism g in V_0 is topologically conjugate with a contraction along $\mathcal{S}$. Hence,

$$|\beta_m(x) - x| < \varepsilon, \quad x \in D,$$

for m large enough.

It follows from our reasoning that if m is large enough, then the mapping $(0, z) \mapsto (\beta_m(z), z)$ is an embedding of the disk D having the properties described by Lemma 6.7.

It is suggested to the reader to formulate an analog of Lemma 6.5 in the case of flows.

6.4 Transversality and hyperbolicity for one-dimensional mappings

Let us describe a relation between transversality and hyperbolicity for one-dimensional mappings.

Let f be a diffeomorphism of the line $\mathbb{R}$.

Consider the following mapping related to f, the graph of f, $\mathrm{grf} : \mathbb{R} \to \mathbb{R}^2$:

$$\mathrm{grf}(x) = (x, f(x)).$$

Denote by Δ the diagonal of $\mathbb{R}^2$: $\Delta = \{(x, x) : x \in \mathbb{R}\}$.

Clearly, p is a fixed point of f if and only if $\mathrm{grf}(p) \in \Delta$.

Lemma 6.8. *If p is a hyperbolic fixed point of f, then the mapping* grf *is transverse to Δ at the point p.*

Proof. Clearly, the spectrum of the one-dimensional linear mapping $Df(p)$ consists of a single number equal to the derivative $f'(p)$. By the definition of hyperbolicity of the fixed point p, $|f'(p)| \neq 1$. Hence, the line

$$\{y = f'(p)x\} = Df(p)\mathbb{R}$$

does not coincide with $T_{(p,p)}\Delta = \Delta$, which means that the mapping grf is transverse to Δ at the point p. □

Remark. The example of the diffeomorphism $f(x) = -x$ shows that the inverse statement is not always true.

Consider the following two sets of diffeomorphisms of the line $\mathbb{R}$; denote by $\mathcal{F}_1$ the set of diffeomorphisms f such that any periodic point of f is hyperbolic and by $\mathcal{F}_2$ the set of diffeomorphisms f for which the mappings grf^m are transverse to Δ for all $m \geq 1$.

Lemma 6.9. $\mathcal{F}_1 = \mathcal{F}_2$.

Proof. The inclusion $\mathcal{F}_1 \subset \mathcal{F}_2$ follows from Lemma 6.8 applied to the family of diffeomorphisms $\{f^m\}$. To prove the inverse inclusion, let us assume that there exists a diffeomorphism $f \in \mathcal{F}_2 \setminus \mathcal{F}_1$ that has a nonhyperbolic periodic point p. If m is the period of p, then $|(f^m)'(p)| = 1$.

If $(f^m)'(p) = 1$, then the line

$$\{y = (f^m)'(p)x\} = Df(p)\mathbb{R}$$

coincides with $T_{(p,p)}\Delta = \Delta$; hence, the mapping grf^m is not transverse to Δ at the point p, which contradicts the inclusion $f \in \mathcal{F}_2$. If $(f^m)'(p) = -1$, then

$$(f^{2m})'(p) = (f^m \circ f^m)'(p) = (f^m)'(f^m(p))(f^m)'(p) = 1,$$

and we conclude that the mapping grf^{2m} is not transverse to Δ at the point p, which again contradicts the inclusion $f \in \mathcal{F}_2$. $\square$

Chapter 7
Hyperbolic sets

7.1 Definition of a hyperbolic set

In Sections 4 and 5, we defined and studied hyperbolic fixed and periodic points of diffeomorphisms and hyperbolic rest points and closed trajectories of smooth flows generated by autonomous systems of differential equations.

Now we give a general definition of a hyperbolic set, one of the basic objects in the theory of structural stability.

We start with the case of a diffeomorphism f of a smooth closed manifold M.

Let dist be a Riemannian metric on M and let $|v|$ be the corresponding norm of a tangent vector $v \in T_x M$.

We say that a set $I \subset M$ is a *hyperbolic set* of a diffeomorphism f if the following conditions hold:

(HS1) the set I is compact and f-invariant;

(HS2) there exist numbers $C > 0$ and $\lambda \in (0, 1)$ with the following property: For any point $p \in I$, two linear subspaces $S(p)$ and $U(p)$ of the tangent space $T_p M$ are defined such that

(HS2.1) $S(p) + U(p) = T_p M$;

(HS2.2) $Df(p)S(p) = S(f(p))$ and $Df(p)U(p) = U(f(p))$;

(HS2.3) if $v \in S(p)$, then $|Df^k(p)v| \leq C\lambda^k|v|$ for $k \geq 0$;

(HS2.4) if $v \in U(p)$, then $|Df^k(p)v| \leq C\lambda^{-k}|v|$ for $k \leq 0$.

The numbers $C > 0$ and $\lambda \in (0, 1)$ are usually called *hyperbolicity constants* of the set I; the families $S(p)$ and $U(p)$ are called the *hyperbolic structure* on I.

Remark. In the above definition of a hyperbolic set, it is assumed that the hyperbolic structure is invariant with respect to the derivative Df of the diffeomorphism f (property (HS2.2)).

Let us show that property (HS2.2) can be replaced by the following one: For any point $p \in I$, the dimensions of the spaces $S(f^k(p))$ and $U(f^k(p))$ are the same for all $k \in \mathbb{Z}$.

Assume that this condition is fulfilled and

$$Df^{\nu} S(p) \neq S(q)$$

for some $p \in I$ and $q = f^{\nu}(p)$ (the case of the subspace U is considered similarly). Let $\dim S(f^k(p)) = l$ for $k \in \mathbb{Z}$.

Since Df is a nondegenerate linear mapping,

$$Q := Df^{-\nu} S(q) \neq S(p).$$

For any vector $v \in Q$, the inclusion

$$Df^{\nu} v \in S(q)$$

holds, hence, condition (HS2.3) implies the inequalities

$$|Df^k(p)v| \leq C\lambda^{k-\nu}|Df^{\nu}(p)v|, \quad k \geq \nu.$$

Thus, there exists a constant $C_1 \geq C$ such that

$$|Df^k(p)v| \leq C_1\lambda^k|v|, \quad k \geq \nu, \tag{7.1}$$

for all $v \in Q$.

By our assumption, the subspaces Q and $S(p)$ of the space T_pM have the same dimension l and do not coincide. Fix a vector v_0 of unit length that belongs to Q and does not belong to $S(p)$.

Denote by L the one-dimensional space spanned by v_0; set $P = L + S(p)$.

The angle between the vector v_0 and the subspace $S(p)$ is nonzero. Hence, there exists a number $a > 0$ such that if a vector $w \in P$, $|w| = 1$, is represented in the form

$$w = a_0v_0 + a_1v_1, \quad v_1 \in S(p),\ |v_1| = 1,$$

then $|a_i| \leq a$, $i = 0, 1$.

Condition (HS2.3) and inequalities (7.1) imply that there exists a number $C_2 \geq C_1$ such that if $w \in P$, then

$$|Df^k(p)w| \leq C_2\lambda^k|w|, \quad k \geq 0. \tag{7.2}$$

Find a natural number m such that $C_2\lambda^m < 1/2$.

Let $r = f^m(p)$. Property (HS2.1) and the fact that the dimensions $U(f^k(p))$ are constant imply that

$$\dim Df^{-m}U(r) = \dim U(p) \geq \dim M - l.$$

Since $\dim P = l + 1$, there exists a vector

$$w_0 \in Df^{-m}U(r) \cap P$$

of unit length.

This vector satisfies the following relations:

$$1 = |w_0| = |Df^{-m}(r)Df^m(p)w_0| \leq \frac{1}{2}|Df^m(p)w_0| \leq \frac{1}{4}|w_0| = \frac{1}{4}$$

(in the first inequality, we refer to property (HS2.4) and the inequality $C\lambda^m < 1/2$; in the second one, we refer to inequality (7.2)). We get a contradiction; our statement is proven.

7.2 Examples of hyperbolic sets

Let us give two important examples of hyperbolic sets.

Example 7.1 (Hyperbolic fixed point). The simplest example of a hyperbolic set is a hyperbolic fixed point.

Let p be a hyperbolic fixed point of a diffeomorphism f of a smooth manifold M. Let us check the above definition of a hyperbolic set using coordinates whose existence has been established in Lemma 4.1.

Let L be the matrix corresponding to the linear mapping

$$Df(p) : T_pM \to T_pM.$$

There exists a nonsingular matrix T such that the matrix $A = T^{-1}LT$ has the properties described in Lemma 4.1 (with $\mathbb{R}^n$ replaced by T_pM).

Consider the representation $x = (y, z)$ corresponding to the block-diagonal structure of the matrix $A = \text{diag}(A_+, A_-)$ in which $\max(\|A_+\|, \|A_-\|) < 1$.

Denote by S' and U' the subspaces of the space T_pM given by th equalities $z = 0$ and $y = 0$, respectively.

Set $S = TS'$ and $U = TU'$ (we identify the matrix T and the corresponding nonsingular mapping of T_pM).

The obvious equality $S' \oplus U' = T_pM$ implies that $S \oplus U = T_pM$. Thus, property (HS2.1) is fulfilled.

Let us check property (HS2.2). Take a vector $v \in S$ and find the vector $v' \in S'$ such that $v = Tv'$.

Since the subspace S' is invariant with respect to A, the equality

$$Lv = TAT^{-1}Tv' = TAv'$$

implies that $Lv \in S$. Thus, the subspace S is invariant with respect to $Df(p)$. The same reasoning shows that U is Df-invariant.

Let us show that property (HS2.3) is fulfilled with $C = \|T\|\|T^{-1}\|$ and $\lambda = \max(\|A_+\|, \|A_-\|)$.

If $v \in S$ and $v = Tv'$, then v' satisfies the inequalities $|A^k v'| \leq \lambda^k |v'|$ for $k \geq 0$. Hence,

$$|L^k v| = |TA^k v'| \leq \|T\| \lambda^k |v'|, \quad k \geq 0,$$

and the desired estimate follows from the inequality $|v'| \leq \|T^{-1}\| |v|$.

A similar reasoning shows that property (HS2.4) holds.

Example 7.2 (Hyperbolic automorphism of the two-dimensional torus). Consider a linear mapping L of the plane $\mathbb{R}^2$ given by the formula $x \mapsto Ax$, where

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

Note that

$$\det A = 1. \tag{7.3}$$

We define the two-dimensional torus T^2 as the quotient $\mathbb{R}^2/\mathbb{Z}^2$, where $\mathbb{Z}^2$ is the two-dimensional integer lattice; thus, we identify points $x = (x_1, x_2)$ and $x' = (x_1', x_2')$ of the plane if $x_1 - x_1' \in \mathbb{Z}$ and $x_2 - x_2' \in \mathbb{Z}$.

Entries of the matrix A are integer. Hence, if $x = (x_1, x_2) \in \mathbb{R}^2$ and $Ax = y = (y_1, y_2)$, then for any vector $x' = (x_1 + m_1, x_2 + m_2)$ with integer m_1 and m_2, $Ax' = (y_1 + n_1, y_2 + n_2)$, where the numbers n_1 and n_2 are integer. Thus, the mapping L generates a mapping f of the torus T^2. Equality (7.3) implies that entries of the matrix A^{-1} are integer as well. Hence, the linear mapping L^{-1} generates a mapping of T^2; clearly, this mapping gives us the inverse mapping f^{-1} of f. Both mappings f and f^{-1} are smooth (in local coordinates near any point of T^2, these mappings are linear).

The diffeomorphism f is called the *hyperbolic automorphism of the torus*. The diffeomorphism f was introduced in the theory of dynamical systems by R. Thom. The proof of structural stability of f (see Theorem 8.1) was an important step in the development of the theory of structural stability.

Let us consider a point $p \in T^2$; we identify in the natural way the tangent space T_p^2 and the plane $\mathbb{R}^2$. It is easily seen that the derivative $Df(p)$ coincides with the linear mapping L under this identification.

The eigenvalues of the matrix A are

$$\lambda_{1,2} = \frac{3 \pm \sqrt{5}}{2}.$$

Clearly,

$$\lambda := \lambda_1 = \frac{3 - \sqrt{5}}{2} \in (0, 1) \quad \text{and} \quad \lambda_2 = \frac{1}{\lambda} = \frac{3 + \sqrt{5}}{2} > 1.$$

Denote by S and U the one-dimensional subspaces spanned by the eigenvectors v^s and v^u corresponding to the eigenvalues λ_1 and λ_2, respectively.

For any point $p \in T^2$ we set $S(p) = S$ and $U(p) = U$.

Since the vectors v^s and v^u are linearly independent, property (HS2.1) is fulfilled. The equalities $Av^s = \lambda_1 v^s$ and $Av^u = \lambda_2 v^u$ imply property (HS2.2).

If $v \in S(p)$, then

$$Df^k(p)v = A^k v = \lambda^k v, \quad k \geq 0;$$

if $v \in U(p)$, then

$$Df^k(p)v = A^k v = \lambda_2^k v = \lambda^{-k} v, \quad k \leq 0.$$

Thus, the whole torus T^2 is a hyperbolic set of the diffeomorphism f (with hyperbolicity constants λ and $C = 1$).

Let us prove that the nonwandering set $\Omega(f)$ of the diffeomorphism f coincides with the whole torus T^2.

Lemma 7.1. *A point of the torus is a periodic point of the diffeomorphism f if and only if the coordinates of this point are rational.*

Proof. Fix a natural number n and consider the set Q_n of points of the torus T^2 that correspond to the set of points

$$\left\{\left(\frac{m_1}{n}, \frac{m_2}{n}\right),\ 0 \leq m_1, m_2 \leq n-1\right\}$$

of the plane $\mathbb{R}^2$.

The entries of the matrix A are integer; hence, $f(Q_n) \subset Q_n$. Since the set Q_n is finite, the trajectory of each point of this set is finite. It follows from Lemma 1.2 that any point of the set Q_n is a periodic point of the diffeomorphism f.

Any point of the torus with rational coordinates belongs to one of the sets Q_n. Thus, any point of the torus with rational coordinates is a periodic point of the diffeomorphism f.

Let us prove the converse statement. Let p be a periodic point of f of period m.

If the point p corresponds to a point x of the plane $\mathbb{R}^2$, the equality $f^m(p) = p$ implies the equality $A^m x = x + y$, where y is a point with integer coordinates.

Then $(A^m - E)x = y$. Since the eigenvalues of the matrix A^m equal $\lambda^{\pm m}$ (hence, they are different from 1), the matrix $A^m - E$ is nondegenerate, and we conclude that

$$x = (A^m - E)^{-1} y.$$

It remains to note that the entries of the matrix $A^m - E$ are integer; hence, the entries of the matrix $(A^m - E)^{-1}$ are rational, and the coordinates of the vector x are rational as well. □

Lemma 7.1 implies that any point of the torus T^2 with rational coordinates belongs to the nonwandering set $\Omega(f)$. Since the set of points with rational coordinates is dense in T^2 and the set $\Omega(f)$ is closed (see Theorem 3.2), $\Omega(f) = T^2$.

We have considered two examples of hyperbolic set with opposite properties; in the first case, the hyperbolic set is a point, while in the second case, the hyperbolic set coincides with the phase space.

7.3 Basic properties of hyperbolic sets

Let us note several basic properties of hyperbolic sets, which we use below.

Let I be a hyperbolic set of a diffeomorphism f of a smooth closed manifold M with hyperbolic structure $\{S(p), U(p)\}$ and hyperbolicity constants C, λ.

Take a vector $v \in U(p)$, $p \in I$, and an integer $k > 0$. Let $q = f^k(p)$. Write the vector v in the form

$$v = Df^{-k}(q)Df^k(p)v.$$

Property (HS2.2) of the hyperbolic structure implies that $Df^k(p)v \in U(q)$. By property (HS2.4),

$$|v| \leq C\lambda^k|Df^k(p)v|.$$

Hence,

$$|Df^k(p)v| \geq \frac{\lambda^{-k}}{C}|v|, \quad v \in U(p), k \geq 0. \tag{7.4}$$

A similar reasoning shows that

$$|Df^k(p)v| \geq \frac{\lambda^{k}}{C}|v|, \quad v \in S(p), k \leq 0. \tag{7.5}$$

First we show that the subspaces $S(p)$ and $U(p)$ are complementary subspaces of T_pM.

Lemma 7.2. *If $p \in I$, then $T_pM = S(p) \oplus U(p)$.*

Proof. It is enough to show that $S(p) \cap U(p) = \{0\}$.

Consider a vector $v \in S(p) \cap U(p)$. Property (HS2.3) and inequalities (7.4) imply that if $k > 0$, then

$$\frac{\lambda^{-k}}{C}|v| \leq |Df^k(p)v| \leq C\lambda^k|v|.$$

Hence,

$$\left(C\lambda^k - \frac{\lambda^{-k}}{C}\right)|v| \geq 0. \tag{7.6}$$

Since

$$C\lambda^k - \frac{\lambda^{-k}}{C} < 0$$

for large k, inequalities (7.6) imply that $v = 0$. □

By Lemma 7.2, any vector $v \in T_pM$ is uniquely representable in the form

$$v = v^s + v^u,$$

where $v^s \in S(p)$ and $v^u \in U(p)$. This defines in T_pM projections $P(p)$ to $S(p)$ parallel to $U(p)$ and $Q(p)$ to $U(p)$ parallel to $S(p)$, respectively: $v^s = P(p)v$ and $v^u = Q(p)v$.

Let us show that the norms of these projections are bounded from above by a constant depending only on hyperbolicity constants and on an estimate of the derivative Df on the set I.

Let L and K be complementary subspaces of T_pM. Define the value

$$\angle(L, K) = \min |v - w|,$$

where the minimum is taken over all pairs of vectors $v \in L, w \in K$ such that $|v| = |w| = 1$.

Let $N = \max_{p \in I} \|Df(p)\|$.

Lemma 7.3. *There exists a number $\alpha = \alpha(C, \lambda, N) > 0$ such that if $p \in I$, then $\angle(S(p), U(p)) \geq \alpha$.*

Proof. Take vectors $v \in S(p), w \in U(p)$ with $|v| = |w| = 1$ and a number $k \geq 0$ and consider the vector $a(k) = Df^k(p)(v - w)$.

By the definition of N,

$$|a(k)| \leq N^k |a(0)|. \tag{7.7}$$

Property (HS2.3) and inequalities (7.4) imply that

$$|a(k)| \geq \frac{\lambda^{-k}}{C} - C\lambda^k \to \infty, \quad k \to \infty.$$

Hence, there exists a number $l = l(C, \lambda)$ such that $|a(l)| \geq 1$. By inequality (7.7) with $k = l$,

$$|v - w| = |a(0)| \geq \alpha(C, \lambda, N) = N^{-l(C,\lambda)}.$$ □

Take a vector $v \in T_p M$ and represent it in the form

$$v = P(p)v + Q(p)v.$$

If $P(p)v = 0$, then $|Q(p)v| = |v|$; if $Q(p)v = 0$, then $|P(p)v| = |v|$.

Consider the case where $P(p)v \neq 0$ and $Q(p)v \neq 0$; let

$$v^s = \frac{P(p)v}{|P(p)v|}, \quad v^u = \frac{Q(p)v}{|Q(p)v|}, \quad \text{and} \quad w^u = -v^u.$$

If β is the angle between the vectors $P(p)v$ and $Q(p)v$, then

$$|v^s - v^u| = 2\sin(\beta/2) \geq \alpha \quad \text{and} \quad |v^s - w^u| = 2\cos(\beta/2) \geq \alpha.$$

Hence, $\sin(\beta) \geq b := \alpha^2/2$.

The Law of Sines applied to the triangle with sides v, $P(p)v$, and $Q(p)v$ implies that

$$|P(p)v| \leq \frac{|v|}{\sin(\beta)} \leq \frac{|v|}{b}.$$

Clearly, a similar inequality holds for the vector $Q(p)v$.

Thus, the following statement is a corollary of Lemma 7.3.

Corollary. *There exists a number $R = R(C, \lambda, N)$ such that*

$$\|P(p)\|, \|Q(p)\| \leq R, \quad p \in I. \tag{7.8}$$

Finally, we show that the hyperbolic structure is continuous on a hyperbolic set.

Consider a subset $M_0 \subset M$ and assume that to any point $p \in M_0$ we assign a linear subspace $L(p) \subset T_p M$. The tangent bundle TM of the manifold M is a manifold as well (see [16]); hence, for a sequence of points $p_k \in M_0$ converging to a point $p \in M_0$, we can define the limit

$$\lim_{k\to\infty} L(p_k) = \{v \in T_p M : v = \lim_{k\to\infty} v_k, v_k \in T_{p_k} M\}.$$

We say that the family $\{L(p)\}$ is continuous on M_0 if

$$\lim_{k\to\infty} L(p_k) = L(p)$$

for any point $p \in M_0$ and for any sequence of points p_k converging to p.

Lemma 7.4. *The families $\{S(p)\}$ and $\{U(p)\}$ are continuous on the hyperbolic set I.*

Proof. Consider a sequence of points $p_k \in I$ converging to a point $p \in I$. First we show that if $v_k \in S(p_k)$ is a sequence of unit vectors and $v_k \to v$, then

$$v \in S(p). \tag{7.9}$$

Assume that inclusion (7.9) is not valid. Since $v \in T_pM$, property (HS2.1) of the hyperbolic structure implies that

$$v = v^s + v^u, \quad v^s \in S(p),\ v^u \in U(p),\ v^u \neq 0.$$

Take a number m such that

$$C\lambda^m < \frac{1}{3}, \quad |Df^m(p)v^s| < \frac{1}{3}, \quad |Df^m(p)v^u| > 1. \tag{7.10}$$

Due to the first inequality in (7.10),

$$|Df^m(p_k)v| < \frac{1}{3}.$$

The mapping $Df^m(r)w$ is continuous in r and w; hence,

$$|Df^m(p)v| \leq \frac{1}{3}.$$

The second and third inequalities in (7.10) imply that

$$|Df^m(p)v| \geq |Df^m(p)v^u| - |Df^m(p)v^s| \geq \frac{2}{3}.$$

The contradiction obtained proves inclusion (7.9).

Inclusion (7.9) implies that

$$\mathcal{S}(p) = \lim_{k\to\infty} S(p_k) \subset S(p). \tag{7.11}$$

A similar reasoning shows that

$$\mathcal{U}(p) = \lim_{k\to\infty} U(p_k) \subset U(p).$$

Let s be a number that is met in the sequence $\dim S(p_k)$ infinitely many times. Consider a subsequence p_{k_l} such that $\dim S(p_{k_l}) = s$.

Fix in any subspace $S(p_{k_l})$ an orthonormal basis $v^1_{k_l}, \dots, v^s_{k_l}$. Without loss of generality, we may assume that there exist the limits

$$v^i = \lim_{k_l\to\infty} v^i_{k_l}, \quad i = 1, \dots, s.$$

Clearly, the vectors $v^1, \dots, v^s$ are pairwise orthogonal unit vectors in $\mathcal{S}(p)$. Relation (7.11) implies that

$$\dim S(p) \geq s.$$

Since $\dim U(p_{k_l}) \geq n - s$, a similar reasoning gives us the inequality

$$\dim U(p) \geq n - s.$$

Since

$$\dim S(p) + \dim U(p) = n,$$

we conclude that

$$\dim S(p) = s \quad \text{and} \quad \dim U(p) = n - s.$$

It follows that the number s is uniquely defined, and $\mathcal{S}(p)$ and $\mathcal{U}(p)$ are linear subspaces of $S(p)$ and $U(p)$, respectively, whose dimensions coincide with the dimensions of $S(p)$ and $U(p)$. Clearly,

$$\mathcal{S}(p) = S(p) \quad \text{and} \quad \mathcal{U}(p) = U(p).$$

This proves Lemma 7.4. □

Remarks. 1. It follows from the proof of Lemma 7.4 that the dimensions of the subspaces $S(p)$ and $U(p)$ are locally constant at points of a hyperbolic set.

2. Modifying the proof of Lemma 7.4, it is easy to show that the projections $P(p)$ and $Q(p)$ defined after Lemma 7.2 are continuous as well: If

$$p_k \in I, p_k \to p \in I, w_k \in T_{p_k}M, \quad \text{and} \quad w_k \to w \in T_p M,$$

then

$$P(p_k)w_k \to P(p)w \quad \text{and} \quad P(p_k)w_k \to P(p)w. \tag{7.12}$$

7.4 Stable manifold theorem

Points of hyperbolic sets have stable and unstable manifolds whose properties are parallel to properties of stable and unstable manifolds of hyperbolic fixed and periodic points.

Let us give the corresponding definitions. Let p be a point of a hyperbolic set I for a diffeomorphism f of a smooth closed n-dimensional manifold M.

The stable and unstable manifolds of the point p are defined by the equalities

$$W^s(p) = \{x \in M : \text{dist}(f^k(x),\ f^k(p)) \to 0,\ k \to \infty\}$$

and

$$W^u(p) = \{x \in M : \text{dist}(f^k(x),\ f^k(p)) \to 0,\ k \to -\infty\}.$$

Theorem 7.1 below describes properties of these sets. Usually, this theorem is called the stable manifold theorem.

Denote by $N(a, p)$ the open ball of radius $a > 0$ centered at p (with respect to the Riemannian metric dist).

Theorem 7.1. *Assume that f is a diffeomorphism of class $C^r, r \geq 1$; let I be a hyperbolic set of the diffeomorphism f. There exists a number $\Delta > 0$ with the following properties. If $p \in I$ and* $\dim S(p) = l$*, then:*

(1) *there exists immersions $b^s : \mathbb{R}^l \to M$ and $b^u : \mathbb{R}^{n-l} \to M$ of class C^r such that*

$$b^s(0) = p, \quad b^s(\mathbb{R}^l) = W^s(p), \quad b^u(0) = p, \quad \textit{and} \quad b^s(\mathbb{R}^{n-l}) = W^u(p);$$

(2) *for any $a \in (0, \Delta)$, the point p belongs to smooth (of class C^r) disks $W^s(a, p)$ and $W^u(a, p)$ that are components of intersection of $W^s(p)$ and $W^u(p)$ with $N(a, p)$, respectively, and*

(2.1) *$T_pW^s(a, p) = S(p)$ and $T_pW^u(a, p) = U(p)$;*

(2.2) *if $x \in N(a, p) \setminus W^s(a, p)$, then there exists an index $m > 0$ such that*

$$\text{dist}(f^m(x), f^m(p)) \geq \Delta;$$

(2.3) *if $x \in N(a, p) \setminus W^u(a, p)$, then there exists an index $m < 0$ such that*

$$\text{dist}(f^m(x), f^m(p)) \geq \Delta.$$

The proof of statement (2) of Theorem 7.1 for the set $W^s(a, p)$ is practically the same as the proof of Theorem 4.2 above. Proving Theorem 4.2, we applied the following estimates in the study of the Perron operator:

$$|B^{k-1-i} f_1(x_i)| \leq \lambda^{k-1-i}|f_1(x_i)| \quad \text{and} \quad |C^{k-1-i} f_2(x_i)| \leq \lambda^{i+1-k}|f_2(x_i)|. \tag{7.13}$$

In the proof of the corresponding statement of Theorem 7.1, one works with the representation

$$f(p_k + v) = f(p_k) + Df(p_k)v + F(p_k, v) \tag{7.14}$$

in local coordinates near a point $p_k = f^k(p)$, where $v \in T_{p_k}M$.

To construct the local stable manifold, one uses the Perron operator that assigns to a sequence $V = \{v_k \in T_{p_k} M\}$ the sequence $W = \{w_k \in T_{p_k} M\}$ by the formulas

$$P(p_k)w_k = Df^k(p)P(p)v_0 + \sum_{i=0}^{k-1} Df^{k-1-i}(p)P(p_i)F(p_i, v_i)$$

and

$$Q(p_k)w_k = -\sum_{i=k}^{\infty} Df^{k-1-i}(p)Q(p_i)F(p_i, v_i),$$

where P and Q are the projections defined after Lemma 7.2.

Properties (HS2.3) and (HS2.4) of a hyperbolic set and estimates (7.8) imply the estimates

$$|Df^{k-1-i}(p)P(p_i)F(p_i, v_i)| \le CR\lambda^{k-1-i}|F(p_i, v_i)|$$

and

$$|Df^{k-1-i}(p)Q(p_i)F(p_i, v_i)| \le CR\lambda^{i+1-k}|F(p_i, v_i)|,$$

that are similar to estimates (7.13) (up to the factor CR).

It remains to note that the value Δ (the size of the local stable and unstable manifolds in Theorem 4.2) is determined by the smallness of the Lipschitz constant of the nonlinear term F in a neighborhood of the hyperbolic fixed point p, and the radius of this neighborhood is proportional to Δ.

The derivative Df on the compact set I is uniformly continuous. Hence, one can choose the size Δ of neighborhoods of points of the hyperbolic set I small enough, so that the Lipschitz constants of the nonlinear terms $F(p_k, v)$ with respect to v in representations (7.14) are so small that the required estimates of the Perron operator in the proof of Theorem 7.1 are similar to the corresponding estimates in the proof of Theorem 4.2.

The reader can find the remaining details of the proof of Theorem 7.1, for example, in the book [15].

7.5 Axiom A

S. Smale introduced the following condition on a diffeomorphism f of a smooth closed manifold M.

Axiom A. (1) *The nonwandering set* $\Omega(f)$ *is hyperbolic.*

(2) *The set of periodic points of* f *is dense in* $\Omega(f)$.

This condition played a very important role in the development of the theory of structural stability. First we describe the structure of the nonwandering set of a diffeomorphism that satisfies Axiom A. Smale proved the following statement.

Theorem 7.2 (spectral decomposition theorem). *If a diffeomorphism f satisfies Axiom* A, *then its nonwandering set can be represented in the form*

$$\Omega(f) = \Omega_1 \cup \cdots \cup \Omega_m, \tag{7.15}$$

where the Ω_i are disjoint, compact, invariant sets such that each of these sets contains a dense positive semitrajectory. Representation of the form (7.15) *is unique.*

The sets Ω_i in representation (7.15) are called *basic*.

We have shown in Lemma 7.4 that the spaces $S(p)$ and $U(p)$ of the hyperbolic structure of the set $\Omega(f)$ are continuous with respect to p. The smooth disks in the stable and unstable manifolds of a point p described in Theorem 7.2 are tangent at the point p to the spaces $S(p)$ and $U(p)$, respectively; analyzing the proof of the continuous differentiability of the mappings b^s and b^u (and using the same reasoning as in the proof of Lemma 4.10), one can prove the following statement.

Lemma 7.5. *Any point $p \in \Omega(f)$ has a neighborhood U_p such that if $x, y \in U_p \cap \Omega(f)$, then the manifolds $W^u(x)$ and $W^s(y)$ have a point of transverse intersection.*

To prove Theorem 7.2, we need the following lemma.

Lemma 7.6. *Let p and q be hyperbolic periodic points of the diffeomorphism f. If x_1 and x_2 are points of transverse intersection for the pairs of manifolds $W^u(p), W^s(q)$ and $W^u(q), W^s(p)$, respectively, then $x_1, x_2 \in \Omega(f)$.*

Proof. Assume, for definiteness, that p and q are hyperbolic fixed points of f (to consider the general case, it is enough to pass from f to f^k, where k is large enough so that p and q are fixed points of f^k; clearly, if $x \in \Omega(f^k)$, then $x \in \Omega(f)$).

Take an arbitrarily small open ball V containing the point x_1. The ball V is a smooth disk; clearly, this disk is transverse to $W^s(p)$ at the point x_1. By λ-lemma (Lemma 6.5) there exists a smooth disk $Q^u \subset W^u(q)$ containing the point q and having the following property. If Q^u is the image of a ball D_0 under an embedding g_u, then there exists a sequence of embeddings g_u^m of the disk D_0 such that

$$g_u^m(D_0) \subset f^m(V) \quad \text{and} \quad \rho_1(g_u^m, g_u) \to 0,\ m \to \infty.$$

Consider smooth disks P^s and P^u belonging to $W^s(p)$ and $W^u(q)$, respectively, and such that x_2 is a point of transverse intersection of P^s and P^u. Assume that the disks P^s and P^u are the images of balls D_1 and D_2 under embeddings h_s and h_u, respectively. Apply Lemma 6.3 to find a number $\delta > 0$ such that if H_s and

H_u are embeddings of D_1 and D_2 with $\rho_1(H_s, h_s) < \delta$ and $\rho_1(H_u, h_u) < \delta$, then $H_s(D_1) \cap H_u(D_2) \neq \emptyset$.

There exists a negative number l such that $f^l(P_u) \subset Q^u$.

Clearly, the disks $g_u^m(D_0)$ contain disks ν_m that are images of D_2 under embeddings h_u^m such that $\rho_1(h_u^m, f^l \circ h_u) \to 0$ as $m \to \infty$.

Set $H_u^m = f^{-l} \circ h_u^m$. Then

$$H_u^m(D_2) = f^{-l}(\nu_m) \subset f^{m+l}(V).$$

Since the number l is fixed, $\rho_1(H_u^m, h_u) \to 0$ as $m \to \infty$.

Hence, there exists a number m_1 such that $\rho_1(H_u^m, h_u) < \delta$ for $m \geq m_1$.

A similar reasoning shows that there exist numbers $k > 0$ and m_2 and embeddings H_s^m of the disk D_1 such that $\rho_1(H_s^m, h_s) < \delta$ for $m \geq m_2$, and the disks $H_s^m(D_1)$ are subsets of $f^{-k-m}(V)$.

The choice of the number δ implies that

$$f^{m+l}(V) \cap f^{-k-m}(V) \neq \emptyset, \quad \text{i.e.,} \quad f^{2m+l+k}(V) \cap V \neq \emptyset \tag{7.16}$$

for $m \geq \max(m_1, m_2)$.

Since relations (7.16) are valid for arbitrarily large m and the neighborhood V can be taken arbitrarily small, $x_1 \in \Omega(f)$. Similarly one shows that $x_2 \in \Omega(f)$. □

Let us pass to the proof of Theorem 7.2. Take a point $x_0 \in \Omega(f)$ and consider a neighborhood U of this point having the property formulated in Lemma 7.5 (we take the set $\Omega(f)$ as the hyperbolic set I).

Set

$$\Xi(x_0) = \mathrm{Cl}\, O(U \cap \Omega(f), f)$$

(recall that $O(A, f)$ is the trajectory of a set A in the dynamical system f). The set $\Xi(x_0)$ is the closure of an invariant subset of $\Omega(f)$; hence, the set $\Xi(x_0)$ is a closed, invariant subset of $\Omega(f)$.

We claim that the set $\Xi(x_0)$ depends only on the point x_0 and not on the choice of the neighborhood U (let us formulate the exact statement as a separate lemma).

Lemma 7.7. *If V is an open subset of U such that $V \cap \Omega(f) \neq \emptyset$, then the set*

$$\Xi_1(x_0) = \mathrm{Cl}\, O(V \cap \Omega(f), f)$$

coincides with $\Xi(x_0)$.

Proof. Fix a point $y \in V \cap \Omega(f)$. Since periodic points are dense in $\Omega(f)$, there exists a periodic point p that belongs to V. Let q be an arbitrary periodic point in U.

By the choice of the neighborhood U, there exist points x_1, x_2 of transverse intersection for the pairs $W^u(p), W^s(q)$ and $W^u(s), W^s(p)$. By Lemma 7.6, $x_1, x_2 \in \Omega(f)$.

Since $x_1 \in W^u(p)$, the trajectory of the point x_1 intersects the set V; hence, $x_1 \in \Xi_1(x_0)$. In addition, $x_1 \in W^s(q)$; it follows that an arbitrary neighborhood of the point q contains points of the trajectory $O(x_1, f)$. Hence, $q \in \mathrm{Cl}\, \Xi_1(x_0)$; since the set $\Xi_1(x_0)$ is closed, we conclude that $q \in \Xi_1(x_0)$.

Consider now an arbitrary point $x \in U \cap \Omega(f)$. Any neighborhood of this point contains a periodic point $q \in U$. As was shown above, $q \in \Xi_1(x_0)$; hence, $x \in \Xi_1(x_0)$. Thus,

$$O(U \cap \Omega(f), f) \subset \Xi_1(x_0).$$

Passing to closures in the above inclusion, we get the inclusion

$$\Xi(x_0) \subset \Xi_1(x_0).$$

The inverse inclusion is obvious. □

Corollary. *For any point $x \in \Xi(x_0)$, $\Xi(x_0) = \Xi(x)$.*

Proof. Fix an arbitrary point $x \in \Xi(x_0)$ and its arbitrary neighborhood W. Consider the set

$$\Xi(x) = \mathrm{Cl}\, O(W \cap \Omega(f), f).$$

It follows from the definition of the set $\Xi(x_0)$ that there exists a point $x_1 \in U \cap \Omega(f)$ and an index k such that $f^k(x_1) \in W$. In this case, there exists an open set $V \subset U$ such that $f^k(V) \subset W$.

By Lemma 7.7,

$$\Xi(x_0) = \mathrm{Cl}\, O(V \cap \Omega(f), f) \subset \mathrm{Cl}\, O(W \cap \Omega(f), f) = \Xi(x). \tag{7.17}$$

Relations (7.17) imply that $x_0 \in \Xi(x)$, and the same reasoning shows that $\Xi(x) \subset \Xi(x_0)$. □

Let us construct the sets Ω_x for all points $x \in \Omega(f)$. The corollary to Lemma 7.7 implies that if $x, y \in \Omega(f)$, then the sets Ω_x and Ω_y either coincide or do not intersect. Indeed, if $z \in \Omega_x \cap \Omega_y$, then

$$\Omega_z = \Omega_y = \Omega_x.$$

Let us show that the number of different sets Ω_x is finite. Otherwise, there exists a countable family of distinct sets $\Omega_{x_1}, \Omega_{x_2}, \dots$. Let y be a limit set of the sequence $\{x_k\}$ (such a point exists since the manifold M is compact). Since $x_k \in \Omega(f)$ and the set $\Omega(f)$ is closed, $y \in \Omega(f)$. Let V be the neighborhood used in the construction of the set Ω_y. There exists an index k_0 such that $x_k \in V$ for $k \geq k_0$. By the construction of the set Ω_y, the inclusions $x_k \in \Omega_y$ hold for $k \geq k_0$, which implies that the sets Ω_{x_k}, $k \geq k_0$, coincide with Ω_y.

The contradiction obtained shows that the number of different sets Ω_x is finite. Denote these sets $\Omega_1, \dots, \Omega_m$.

To prove that every set Ω_i contains a dense positive semitrajectory, we apply a construction invented by Birkhoff (see [1]).

Take a set Ω_i and fix a countable dense subset $p_k, k = 1, 2, \dots,$ of the set Ω_i. We assume that the set $\{p_k\}$ has the following property: For any $k_0 > 0$, the set $\{p_k : k \geq k_0\}$ is still dense in Ω_i (since the set Ω_i is compact, such a dense set obviously exists). Take a number $d \in (0, 1)$ and assume that d is small enough, so that if x and y are points of different basic sets, then

$$\operatorname{dist}(x, y) \geq d.$$

Clearly, if a sequence of points $q_k \in \Omega_i, k = 1, 2, \dots,$ satisfies the inequalities

$$\operatorname{dist}(p_k, q_k) < d^k,$$

then the set $\{q_k\}$ is dense in Ω_i. As above, we denote by $N(a, p)$ the ball of radius a centered at a point p. Let $D_k = N(d^k, p_k)$.

The set D_1 is open and contains the point $p_1 \in \Omega_i$; Lemma 7.6 and its corollary imply that

$$\Omega_i = \operatorname{Cl} O(D_1 \cap \Omega(f), f).$$

Hence, there exists a point $r_2 \in D_1 \cap \Omega(f)$ and a number k_2 such that

$$f^{k_2}(r_2) \in D_2.$$

Since the mapping f^{k_2} is continuous, there exists a neighborhood Q_2 of the point r_2 such that $\operatorname{Cl} Q_2 \subset D_1$ and

$$f^{k_2}(\operatorname{Cl} Q_2) \in D_2.$$

We apply a similar reasoning to show that there exists a point $r_3 \in Q_2 \cap \Omega(f)$ and a number k_2 such that

$$f^{k_3}(r_3) \in D_3.$$

Find a neighborhood Q_3 of the point r_3 such that $\operatorname{Cl} Q_3 \subset Q_2$ and

$$f^{k_3}(\operatorname{Cl} Q_3) \in D_3.$$

By construction,

$$f^{k_2}(\mathrm{Cl}\, Q_3) \in D_2,$$

which means that the positive semitrajectory of any point of the set Q_3 intersects both D_2 and D_3.

We continue this procedure and construct sequences of points r_j, numbers k_j, and neighborhoods Q_j such that

$$r_j \in Q_j \cap \Omega(f), \quad f^{k_j}(\mathrm{Cl}\, Q_j) \in D_j,$$

and

$$D_1 \supset \mathrm{Cl}\, Q_1 \supset Q_1 \supset \mathrm{Cl}\, Q_2 \supset Q_2 \supset \cdots .$$

We require, in addition, that

$$\operatorname{diam} Q_j \to 0, \quad j \to \infty. \tag{7.18}$$

The sequence of embedded compact sets $\mathrm{Cl}\, Q_j$ has a nonempty intersection; relation (7.18) implies that this intersection is a point (which we denote by q).

It follows from (7.18) that $r_j \to q$, $j \to \infty$. Since $r_j \in \Omega_i$, $q \in \Omega_i$ as well. By our construction $q_j = f^{k_j}(q) \in D_j$; hence, the positive semitrajectory of the point q is dense in Ω_i.

To complete the proof of Theorem 7.2, it remains to establish the uniqueness of representation (7.15).

Let

$$\Omega(f) = \Xi_1 \cup \cdots \cup \Xi_l,$$

where the Ξ_i are disjoint, compact, invariant sets each of which contains a dense positive semitrajectory.

Take a point $x \in \Xi_1$ such that

$$\Xi_1 = \mathrm{Cl}\, O^+(x, f).$$

There exists an index $i \in \{1, \dots, m\}$ such that $x \in \Omega_i$.

Then $O^+(x, f) \subset \Omega_i$. Passing to closures in this inclusion, we conclude that $\Xi_1 \subset \Omega_i$.

The same reasoning shows that there exists an index j such that $\Omega_i \subset \Xi_j$. Clearly, $j = 1$, and $\Omega_i = \Xi_j$.

Hence, any of the sets Ω_i coincides with one of the sets Ξ_k. This completes the proof of Theorem 7.2. □

Let Ω_i be a basic set; define the sets

$$W^s(\Omega_i) = \{x \in M : \operatorname{dist}(f^k(x), \Omega_i) \to 0,\ k \to \infty\}$$

and

$$W^u(\Omega_i) = \{x \in M : \mathrm{dist}(f^k(x), \Omega_i) \to 0,\ k \to -\infty\}.$$

These sets are analogs of stable and unstable manifolds for individual trajectories.

Theorem 7.3. *If a diffeomorphism f satisfies Axiom* A*, then*

$$M = \bigcup_{i=1}^{m} W^s(\Omega_i) = \bigcup_{i=1}^{m} W^u(\Omega_i). \tag{7.19}$$

Proof. Consider a point $x \in M$. By Lemma 3.3, the sets $\alpha(x, f)$ and $\omega(x, f)$ are subsets of the nonwandering set $\Omega(f)$.

Let us show that the set $\omega(x, f)$ intersects not more than one basic set. To get a contradiction, let us assume that there exist two different basic sets Ω_i and Ω_j such that

$$\omega(x, f) \cap \Omega_i \neq \emptyset \quad \text{and} \quad \omega(x, f) \cap \Omega_j \neq \emptyset.$$

Find neighborhoods $U_1, \dots, U_m$ of the basic sets $\Omega_1, \dots, \Omega_m$ such that

$$(U_k \cup \mathrm{Cl}\, f(U_k)) \cap U_l = \emptyset, \quad k \neq l,$$

and set $U = U_1 \cup \dots \cup U_m$.

There exist sequences of indices $l_k, m_k \to \infty, k \to \infty$, such that $l_k < m_k < l_{k+1}$,

$$\mathrm{dist}(f^{l_k}(x), \Omega_i) \to 0, \quad \text{and} \quad \mathrm{dist}(f^{m_k}(x), \Omega_j) \to 0.$$

If k is large enough, then

$$f^{l_k}(x) \in U_i \quad \text{and} \quad f^{m_k}(x) \in U_j.$$

Hence, there exists a sequence of indices n_k such that $l_k < n_k < m_k$ and

$$f^{n_k}(x) \in \mathrm{Cl}\, f(U_i) \setminus U_i.$$

A limit point y of the sequence $f^{n_k}(x)$ belongs to the compact set $\mathrm{Cl}\, f(U_i) \setminus U_i$. The choice of the neighborhoods U_k implies that $y \notin U$, and we get a contradiction with the inclusion

$$y \in \omega(x, f) \subset \Omega(f).$$

Hence, the set $\omega(x, f)$ is a subset of a unique basic set (let this set be Ω_i). Clearly, $x \in W^s(\Omega_i)$.

A similar reasoning proves the second equality in (7.19). □

Theorem 7.3 means that if a diffeomorphism satisfies Axiom A, then any trajectory of this diffeomorphism tends to a basic set both for $k \to \infty$ and for $k \to -\infty$.

In fact the following, more exact, statement holds (we do not give a proof here and refer the reader to the original paper [17]).

Theorem 7.4. *If a diffeomorphism f satisfies Axiom* A, *then*

$$M = \bigcup_{p \in \Omega(f)} W^s(p) = \bigcup_{p \in \Omega(f)} W^u(p). \tag{7.20}$$

Now we give definitions which we need to formulate necessary and sufficient conditions of Ω-stability and structural stability of diffeomorphisms.

Let Ω_i and Ω_j be two different basic sets of a diffeomorphism that satisfies Axiom A. We write $\Omega_i \to \Omega_j$ if

$$W^u(\Omega_i) \cap W^s(\Omega_i) \neq \emptyset.$$

We say that a diffeomorphism has a 1-*cycle* if there exists a basic set Ω_i such that

$$\big(W^u(\Omega_i) \cap W^s(\Omega_i)\big) \setminus \Omega_i \neq \emptyset.$$

We say that a diffeomorphism has a k-*cycle* ($k > 1$) if there exist k different basic sets $\Omega_{i_1}, \dots, \Omega_{i_k}$ such that

$$\Omega_{i_1} \to \dots \to \Omega_{i_k} \to \Omega_{i_1}.$$

We say that a diffeomorphism satisfies the *no cycle condition* if it does not have k-cycles with $k \geq 1$.

The following theorem states necessary and sufficient conditions of Ω-stability.

Theorem 7.5. *A diffeomorphism f is Ω-stable if and only if f satisfies Axiom* A *and the no cycle condition.*

We say that a diffeomorphism f satisfies the *strong transversality condition* if for any points $p, q \in \Omega(f)$, the stable manifold $W^s(p)$ and the unstable manifold $W^u(q)$ are transverse (for a diffeomorphism satisfying Axiom A, the sets $W^s(p)$ and $W^u(q)$ are not necessarily submanifolds of the phase space; their transversality is understood in the sense explained in Section 6.2).

Theorem 7.6. *A diffeomorphism f is structurally stable if and only if f satisfies Axiom* A *and the strong transversality condition.*

The history of the proof of Theorems 7.5 and 7.6 is complicated; we analyze it in Appendix B. The necessity of conditions of Theorem 7.6 had been established by R. Mañé. Appendix A of this book is devoted to a scheme of the proof of the Mañé result.

In a particular but important case of Anosov diffeomorphisms, we prove structural stability in the next section (to simplify presentation, we consider the case of a hyperbolic automorphism of the two-dimensional torus).

Let us also note that in the case of diffeomorphisms with finite nonwandering set, an analog of Theorem 7.6 had been proven by Smale and Palis [18]. It follows from Lemma 1.2 that the nonwandering set of a diffeomorphism is finite if and only if this set coincides with the set of periodic points (and this latter set is finite).

The corresponding analog of Theorem 7.6 is as follows.

A diffeomorphism f with finite nonwandering set is structurally stable if and only if the following two conditions are fulfilled:

(MS1) *the nonwandering set consists of a finite number of hyperbolic periodic points;*

(MS2) *stable and unstable manifolds of periodic points are transverse.*

Diffeomorphisms satisfying conditions (MS1) and (MS2) are called *Morse–Smale diffeomorphisms.* These diffeomorphisms are analogs of structurally stable (rough) systems introduced by A. A. Andronov and L. S. Pontryagin (see the next subsection).

Let us mention (without proofs) several important results related to description of the sets of Ω-stable and structurally stable diffeomorphisms.

It was shown by Smale [19] that the set of structurally stable diffeomorphisms is not dense in the space $\mathrm{Diff}^1(M)$ if $\dim M \geq 3$; later, R. Williams showed that structurally stable diffeomorphisms are not dense in the space $\mathrm{Diff}^1(T^2)$, where T^2 is the two-dimensional torus [20].

R. Abraham and Smale had constructed an open subset U of the space $\mathrm{Diff}^1(M)$, where M is a 4-dimensional manifold, such that any diffeomorphism $f \in U$ is not Ω-stable and does not satisfy Axiom A [21].

The following two statements are widely used when one wants to prove that a particular set of diffeomorphisms consists of structurally stable or Ω-stable diffeomorphisms.

If X is a subset of $\mathrm{Diff}^1(M)$, we denote by $\mathrm{Int}^1(X)$ the C^1-interior of the set X.

Denote by $\mathcal{H}$ the set of diffeomorphisms f such that any periodic point of f is hyperbolic.

S. Hayashi and N. Aoki [22, 23] independently proved that the set $\mathrm{Int}^1(\mathcal{H})$ coincides with the set of Ω-stable diffeomorphisms.

In addition, it was shown by Aoki [22] that the set $\mathrm{Int}^1(\mathrm{KS})$ coincides with the set of structurally stable diffeomorphisms (recall that KS denotes the set of Kupka–Smale diffeomorphisms, see Section 6.2).

Theorem 6.2 and the above-mentioned Smale's theorem on the nondensity of structurally stable diffeomorphisms imply that there exist Kupka–Smale diffeomorphisms that are not structurally stable.

7.6 Hyperbolic sets of flows

Let us formulate analogs of the above definitions and statements for the case of flows generated by smooth vector fields on a smooth manifold.

Let us start with the simplest case of a flow $\phi(t, x)$ generated by an autonomous system of differential equations of the form (1.1) in the Euclidean space $\mathbb{R}^n$.

We say that a set $I \subset M$ is a *hyperbolic set* of the flow ϕ if it has the following properties:

(HSF1) the set I is compact and ϕ-invariant;

(HSF2) there exist numbers $C > 0$ and $\lambda > 0$ and linear subspaces $S(p)$ and $U(p)$ of the space $\mathbb{R}^n$ defined for any point $p \in I$ such that

(HSF2.1) $S(p) + U(p) + \{F(x)\} = \mathbb{R}^n$, where $\{F(x)\}$ is the subspace spanned by the vector $F(x)$;

(HSF2.2) if $\Phi(t)$ is the fundamental matrix of the variational system

$$\frac{dy}{dt} = \frac{\partial F}{\partial x}(\phi(t, x)y$$

along the trajectory $\phi(t, p)$ such that $\Phi(0) = E$, then

$$\Phi(t)S(p) = S(\phi(t, p)) \quad \text{and} \quad \Phi(t)U(p) = U(\phi(t, p)), \quad t \in \mathbb{R};$$

(HSF2.3) if $v \in S(p)$, then $|\Phi(t, p)v| \leq C \exp(-\lambda t)|v|$ for $t \geq 0$;

(HSF2.4) if $v \in U(p)$, then $|\Phi(t, p)v| \leq C \exp(\lambda t)|v|$ for $t \leq 0$.

The main difference in the definitions of a hyperbolic set for a diffeomorphism and a flow is as follows: The representation $S(p) + U(p) + \{F(x)\} = \mathbb{R}^n$ in the case of a flow contains the subspace $\{F(x)\}$ such that the images of vectors from this space do not increase or decrease exponentially under the action of the operator $\Phi(t)$.

In the case of a flow on a smooth manifold generated by a smooth vector field, the definition of a hyperbolic set is literally the same (with the natural replacement of $\mathbb{R}^n$ by the corresponding tangent space). In what follows, we consider flows ϕ on smooth manifolds (unless otherwise explicitly stated).

The following condition introduced by Smale is an analog of Axiom A for the case of flows.

Axiom A′. (1) *The nonwandering set $\Omega(\phi)$ of the flow ϕ is hyperbolic;*

(2) *the set $\Omega(\phi)$ is the union of two disjoint compact ϕ-invariant sets Q_1 and Q_2, where Q_1 consists of a finite number of rest points, while Q_2 does not contain rest points and points of closed trajectories are dense in Q_2.*

If a flow ϕ satisfies Axiom A$'$, then the following analog of Theorem 7.2 holds: The nonwandering set $\Omega(\phi)$ has a unique representation of the form

$$\Omega(\phi) = \Omega_1 \cup \cdots \cup \Omega_m,$$

where the Ω_i are disjoint, compact, ϕ-invariant sets such that each of these sets contains a dense positive semitrajectory.

As in the case of a diffeomorphism, the sets Ω_i are called basic. A basic set is either a rest point or a closed invariant set that does not contain rest points and such that points of closed trajectories are dense in it.

Let Ω_i and Ω_j are two different basic sets of a flow ϕ, that satisfies Axiom A$'$. We write $\Omega_i \to \Omega_j$ if there exists a point x such that

$$\phi(t, x) \to \Omega_i, t \to -\infty, \quad \text{and} \quad \phi(t, x) \to \Omega_j, t \to \infty.$$

The no cycle condition for a flow ϕ literally repeats the corresponding condition for a diffeomorphism.

The following statement is an analog of Theorem 7.5: A flow ϕ is Ω-stable if and only if ϕ satisfies Axiom A$'$ and the no cycle condition.

If a flow ϕ satisfies Axiom A$'$, then hyperbolic trajectories $\phi(t, p)$, $p \in \Omega(\phi)$, have stable and unstable manifolds $W^s(\phi(t, p))$ and $W^u(\phi(t, p))$ whose properties are similar to properties of stable and unstable manifolds of hyperbolic trajectories of a diffeomorphism described in Theorem 7.1.

We say that a flow ϕ satisfies the strong transversality condition if for any points $p, q \in \Omega(\phi)$, the manifolds $W^s(\phi(t, p))$ and $W^u(\phi(t, q))$ are transverse.

An analog of Theorem 7.6 is stated as follows: A flow ϕ is structurally stable if and only if ϕ satisfies Axiom A$'$ and the strong transversality condition.

As was said above, the notion of a structurally stable (rough) system was first introduced by Andronov and Pontryagin [24] for the case of two-dimensional autonomous systems of differential equations and for flows on the two-dimensional sphere generated by smooth vector fields.

To simplify presentation, let us consider the case of an autonomous system of differential equations. Consider a system of the form (1.1) in a closed two-dimensional disk D bounded by a smooth closed curve Γ. We assume that at points of the curve Γ, the vector field of the system is not tangent to the curve and is directed inside the disk D.

Consider a perturbed system

$$\frac{dx}{dt} = G(x). \tag{7.21}$$

Clearly, if the number

$$r_{1,D}(F, G) = \max_{x \in D} \left(|F(x) - G(x)| + \left\| \frac{\partial F}{\partial x}(x) - \frac{\partial G}{\partial x}(x) \right\| \right)$$

is small enough, then the vector field of system (7.21) is not tangent to Γ at points of this curve and is also directed inside the disk D.

In this case, the disk D is positively invariant for the flows ϕ and ψ generated by systems (1.1) and (7.21), respectively, i.e.,

$$\phi(t, x) \in D \quad \text{and} \quad \psi(t, x) \in D \quad \text{for } t \in D, t \geq 0.$$

Andronov and Pontryagin call system (1.1) *rough* if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if system (7.21) satisfies the inequality

$$r_{1,D}(F, G) < \delta,$$

then there exists a homeomorphism h of the disk D to itself that maps intersections of trajectories of system (1.1) with D to intersections of trajectories of system (7.21) with D, preserves the direction of movement along trajectories, and satisfies the inequality

$$\max_{x \in D} |h(x) - x| < \varepsilon$$

(clearly, this definition correspond to the definition of structural stability in the strong sense introduced in Section 3.1).

Before formulating the Andronov–Pontryagin theorem, let us introduce one useful notion.

Let p be a saddle hyperbolic rest point of system (1.1). Our description of stable and unstable manifolds of hyperbolic rest points (see Section 5.1) implies that the stable manifold $W^s(p)$ of the point p is the union of three trajectories: the point p itself and two different trajectories $g_1^s = O(x_1, \phi)$ and $g_2^s = O(x_2, \phi)$ such that $x_i \neq p, i = 1, 2$, and both $\phi(t, x_1)$ and $\phi(t, x_2)$ tend to p as $t \to \infty$. The trajectories g_1^s and g_2^s are called the *stable separatrices* of the saddle rest point p (sometimes, the stable manifold $W^s(p)$ is called the stable separatrix).

Similarly one defines the unstable separatrices g_1^u and g_2^u of the saddle rest point p.

We say that system (1.1) has a *separatrix joining saddles* if there exist saddle hyperbolic rest points p and q (not necessarily distinct) and their stable and unstable separatrices g^s and g^u, respectively, such that $g^s = g^u$.

Now we can state the Andronov–Pontryagin theorem using the terminology of this book.

Theorem 7.7. *System* (1.1) *is rough if and only if the following three conditions are satisfied:*

(AP1) *all the rest points are hyperbolic;*

(AP2) *all the closed trajectories are hyperbolic;*

(AP3) *the system does not have separatrices joining saddles.*

The first complete proof of Theorem 7.7 had been given by de Baggis in [25].

Let us comment relations between Theorems 7.6 and 7.7.

First we note that the relation $r_{1,D}(F,G) \to 0$ implies that $\rho_1(\phi,\psi) \to 0$ (see Lemma 2.2; of course, when we define the corresponding value $\rho_1(\phi,\psi)$, we take the maximum over the disk D).

Denote by Ω the nonwandering set of the flow ϕ in the disk D; let P denote the set of its rest points and closed trajectories in D.

In this case, an analog of Axiom A$'$ is stated as follows:

(AA$'$1) The set Ω is hyperbolic;

(AA$'$2) the set Ω is the union of two disjoint compact ϕ-invariant sets Q_1 and Q_2, where Q_1 consists of a finite number of rest points, while Q_2 does not contain rest points and points of closed trajectories are dense in Q_2.

Clearly, condition (AA$'$1) implies conditions (AP1) and (AP2).

Let us assume that p and q are two saddle hyperbolic rest points of system (1.1) in the disk D and that a stable separatrix g^s of the saddle rest point p coincides with an unstable separatrix g^u of the saddle rest point q. If $x \in g^s \cap g^u$, then the stable manifold $W^s(p)$ and the unstable manifold $W^u(q)$ are not transverse at x. Indeed, both tangent spaces $T_x W^s(p)$ and $T_x W^u(q)$ coincide with the one-dimensional space spanned by the vector $F(x)$. This shows that the condition of transversality of stable and unstable manifolds for trajectories from the set Ω (the natural analog of the strong transversality condition) is violated.

Thus, the strong transversality condition implies condition (AP3). We have shown that, for an autonomous system of differential equations in a two-dimensional disk, the conditions of the Andronov–Pontryagin theorem follow from the conditions of an analog of Theorem 7.6 for flows.

Let us show that the converse statement is true. We start with a simple lemma.

Lemma 7.8. *Let p be a saddle hyperbolic rest point of system* (1.1). *For any sequence of points $p_k \to p$, $k \to \infty$, such that $p_k \notin W^s(p)$ there exists a point $r \in W^u(p) \setminus \{p\}$ and a sequence of times t_k such that $\phi(t_k, p_k) \to r$.*

Proof. We apply the Grobman–Hartman theorem and find a neighborhood V of the origin in $\mathbb{R}^2$, numbers $a, b > 0$, and a homeomorphism h that maps the neighborhood V onto a neighborhood U of p that conjugates the flow ψ of the system

$$\dot{y} = -ay, \quad \dot{z} = bz \tag{7.22}$$

in V and the flow of system (1.1) in U. Let $q_k = h^{-1}(p_k) = (y_k, z_k)$. Since $p_k \notin W^s(p)$, $z_k \neq 0$.

Find a number $c > 0$ such that

$$\{(y,z) : |y|, |z| \le c\} \subset V.$$

Select in the sequence $\{z_k\}$ an infinite subsequence whose elements have the same sign (for definiteness, we assume that $z_k > 0$).

If k is large enough, then $0 < z_k < c$. Since

$$\psi(t, q_k) = (y_k \exp(-at), z_k \exp(bt)),$$

for k large enough there exist numbers t_k such that $z_k \exp(bt_k) = c$. Clearly, $t_k \to \infty$ as $k \to \infty$.

Then $y_k \exp(-at_k) \to 0$, and $\psi(t_k, q_k) \to (0, c)$.

Hence, $\phi(t_k, p_k) \to h^{-1}(0, c) \in W^u(p) \setminus \{p\}$. □

Let us prove one more auxiliary statement.

Lemma 7.9. *If system* (1.1) *satisfies the conditions of Theorem 7.7, then the set P is finite, and the union of trajectories from P coincides with the set Ω.*

Proof. Denote by $\mathcal{P}$ the union of trajectories from P.

Any trajectory belonging to the set $\mathcal{P}$ consists of nonwandering points of the flow ϕ; hence, $\mathcal{P} \subset \Omega$. Let us prove the converse inclusion.

Every rest point p of system (1.1) in the disk D that is not saddle is either attracting or repelling. If p is an attracting or repelling rest point, then p has a neighborhood $U \subset D$ such that if $x \in U$, then $\phi(t, x) \to p$ either as $t \to \infty$ or as $t \to -\infty$. It follows from the Grobman–Hartman theorem that we can find a neighborhood U having the following property: If $x \in U$, then $\phi(t, x) \in U$ for $t \geq 0$ if the rest point p is attracting, and $\phi(t, x) \in U$ for $t \leq 0$ if the rest point p is repelling.

If γ is a hyperbolic closed trajectory of a planar autonomous system of differential equations, then the local Poincaré diffeomorphism T of γ is a diffeomorphism of an interval to an interval. In this case, the hyperbolic fixed point of T corresponding to the closed trajectory γ is either attracting or repelling. It follows that γ has a neighborhood U whose properties are similar to the above-mentioned properties of neighborhoods of attracting and repelling rest points.

The classical Poincaré–Bendixson theorem (see, for example, [26]) states that for any point $x \in \mathbb{R}^2$, its α-limit and ω-limit sets, $\alpha(x, \phi)$ and $\omega(x, \phi)$, are sets of one of the following three types:

(i) a rest point,

ii) a closed trajectory,

(iii) a contour Γ that consists of rest points and of trajectories that tend to these rest points as $t \to \pm\infty$.

Our reasoning above implies that if a contour Γ of type (iii) belongs to the disk D, then any rest point on the contour Γ is a saddle point. Hence, such a contour must consist of saddle rest points and their separatrices. It follows from condition (AP3) of Theorem 7.7 that the third type of sets $\alpha(x, \phi)$ and $\omega(x, \phi)$ cannot realize in the disk D.

We take a point $x \in \Omega$ and show that $x \in \mathcal{P}$. The following two cases are possible.

Case 1. $O(x, \phi) \subset D$.

Since the disk D is closed, $\mathrm{Cl}\, O(x, \phi) \subset D$. Hence, both sets $\alpha(x, \phi)$ and $\omega(x, \phi)$ are subsets of D.

We have shown above that

$$\alpha(x, \phi) \cup \omega(x, \phi) \subset \mathcal{P}.$$

Condition (AP3) of Theorem 7.7 implies that at least one of the sets $\alpha(x, \phi)$ and $\omega(x, \phi)$ is not a saddle rest point; for definiteness, we assume that $\omega(x, \phi)$ is not a saddle rest point. Then $p = \omega(x, \phi)$ is either an attracting rest point or an attracting closed trajectory. Find a neighborhood U of the trajectory p having the above-mentioned properties.

If $x \in p$, then $x \in \mathcal{P}$.

Let us show that the relation $x \notin p$ cannot realize. If this relation holds, we may assume that $x \notin U$ (reducing the neighborhood U, if necessary).

Find a number $T > 0$ such that $\phi(T, x) \in U$. Since ϕ is continuous, there exists a neighborhood V of the point x such that $V \cap U = \emptyset$ and $\phi(T, V) \subset U$.

An analog of Lemma 3.3 for flows implies that there exist sequences $x_k \to x$ and $t_k \to \infty$ as $k \to \infty$ such that $\phi(t_k, x_k) \to x$.

If k is large enough, then $x_k \in V$. The above-mentioned properties of V imply that in this case, $\phi(T, x_k) \in U$, and we deduce from properties of U that $\phi(t, x_k) \in U, t \geq T$.

Since $t_k \geq T$ for k large enough, it follows from the relation $V \cap U = \emptyset$ that $\phi(t_k, x_k) \notin V$ for such k.

The contradiction obtained proves that $x \in \mathcal{P}$.

Case 2. There exists a number T such that $\phi(T, x) \notin D$.

In this case, there exists a neighborhood U of the point $\phi(T, x)$ such that $D \cap U = \emptyset$. It was assumed that the disk D is bounded by a closed curve Γ such that, at points of Γ, the vector field of the system is directed inside D. Hence, if $y \in \Gamma$ and $t \leq 0$, then $\phi(t, y) \notin U$.

An analog of Lemma 3.3 for flows implies that there exist sequences $x_k \to x$ and $t_k \to -\infty$ as $k \to \infty$ such that $\phi(t_k, x_k) \to x$.

If k is large enough, then $\phi(T, x_k) \in U$, and the same reasoning as in the first case leads to a contradiction. This shows that Case 2 is impossible.

Thus, $\mathcal{P} = \Omega$. Since the set Ω is closed, the set $\mathcal{P}$ is closed as well.

Let p be an attracting or repelling rest point or closed trajectory. We have shown above that there exists a neighborhood U_p of p such that

$$U_p \cap \mathcal{P} = \{p\}. \tag{7.23}$$

If p is a saddle rest point, then it follows from the Grobman–Hartman theorem that p has a neighborhood containing no rest points different from p. Let us show that p has a neighborhood containing no points belonging to closed trajectories. To get a contradiction, assume that there exists a sequence of points p_k belonging to closed trajectories and such that $p_k \to p$ as $k \to \infty$. If $q \in W^s(p)$, then $\phi(t,q) \to p$, $t \to \infty$, and it follows that q cannot be a point of a closed trajectory. Hence, $p_k \in W^s(p) \setminus \{p\}$. By Lemma 7.8, there exists a sequence of numbers t_k and a point $r \in W^u(p) \setminus \{p\}$ such that $\phi(t_k, p_k) \to r$. Since $p_k \in \mathcal{P}$ and the set $\mathcal{P}$ is closed and invariant, $r \in \mathcal{P} = \Omega$. Thus, we get a contradiction with the structure of the set Ω described above.

Hence, any trajectory $p \in P$ has a neighborhood with property (7.23).

Since the set $\mathcal{P}$ is closed and the disk D is compact, we conclude that the set P is finite. □

Lemma 7.9 implies that the set Ω satisfies conditions (AA′1) and (AA′2).

If p is an attracting rest point or closed trajectory, then $W^u(p) = \{p\}$, and the stable manifold $W^s(p)$ is two-dimensional. Since $W^u(p)$ cannot intersect stable manifolds of trajectories from the set P that are different from p, $W^u(p)$ are $W^s(p)$ transverse to any $W^u(q)$ and $W^s(q)$, respectively.

The same holds in the case of a repelling rest point or closed trajectory.

If p is a saddle rest point, then condition (AP3) implies that the manifold $W^u(p)$ ($W^s(p)$) can intersect only two-dimensional stable (respectively, unstable) manifolds of trajectories from the set P that are different from p.

Thus, an analog of the strong transversality condition is satisfied.

We have shown that in the case of a planar autonomous system of differential equations, the conditions of the Andronov–Pontryagin theorem are equivalent to the general necessary and sufficient conditions of structural stability for smooth flows.

Chapter 8
Anosov diffeomorphisms

A diffeomorphism f of a closed smooth manifold M is called *Anosov* if the manifold M is a hyperbolic set of f (defining such diffeomorphisms in his pioneering book [27], Anosov called them Y-diffeomorphisms). The hyperbolic automorphism of the two-dimensional torus (see Section 7.2) is an example of an Anosov diffeomorphism.

The basic result on Anosov diffeomorphisms mentioned in this book is as follows.

Theorem 8.1. *An Anosov diffeomorphism is structurally stable.*

In fact, we prove that an Anosov diffeomorphism is structurally stable in the strong sense (see Section 3.1).

To avoid unessential technical difficulties, we consider the particular case of a hyperbolic automorphism of the two-dimensional torus. We explain in the proof how to modify it in the general case.

There exist several principally different approaches to proving the structural stability of an Anosov diffeomorphism. The method which we use in this book is, in a sense, a generalization of Hartman's method applied in the proof of Theorem 4.1.

Let us recall that if we fix a Riemannian metric on the manifold M, then we induce the exponential mapping of a neighborhood U of the section $M \times \{0\}$ of the tangent bundle TM to M; this mapping assigns to a pair $(x, v) \in U$ a point $\exp_x(v) \in M$ as follows. Let V be a unit tangent vector in $T_x M$ and let $\gamma(t)$ be a geodesic with natural parameter t such that

$$\gamma(0) = x \quad \text{and} \quad \frac{d\gamma}{dt}(0) = V. \tag{8.1}$$

If $v = tV$, set

$$\exp_x(v) = \gamma(t).$$

It is known that $\exp_x(\cdot)$ is a diffeomorphism of a neighborhood of the origin in $T_x M$ to a neighborhood of the point x in M; it follows from the definition that

$$\exp_x(0) = x \quad \text{and} \quad D\exp_x(0) = \text{Id}. \tag{8.2}$$

We assume that a Riemannian metric on $M = T^2$ (for brevity, we write M instead of T^2) is induced by the Euclidean metric on the plane (recall that we get the torus as the quotient $\mathbb{R}^2/\mathbb{Z}^2$, where $\mathbb{Z}^2$ is the two-dimensional integer lattice).

For any $x \in M$ we naturally identify the tangent space $T_x M$ with the plane $\mathbb{R}^2$.

Let $x \in M$ and let y be the corresponding point of the square

$$[0,1] \times [0,1] \subset \mathbb{R}^2,$$

(if a point x corresponds to a point of the boundary of the square, we take as y any of such points).

We naturally identify $N(1/4, x) \subset M$ (recall that $N(a, A)$ is the a-neighborhood of a set A) with $N(1/4, y) \subset \mathbb{R}^2$.

Let $V \in T_x M$ be a unit tangent vector. Clearly, under the considered identifications, the geodesic $\gamma(t)$ that satisfies conditions (8.1) is identified with the part of the line through the point y and parallel to the vector V that belongs to $N(1/4, y) \subset \mathbb{R}^2$. This line is given by $\gamma(t) = x + Vt$.

Hence, if $|v| < 1/4$, then we can identify the point $\exp_x(v)$ of the torus with the point $x + v$ of the plane.

This means that the equality

$$\exp_x(v) = x + v \tag{8.3}$$

(in the sense of the above identifications) is valid for any point $x \in M$ and any vector $v \in T_x M$ with $|v| < 1/4$.

We use this property of the torus in our proof of Theorem 8.1 considering small tangent vector fields v and treating points $x + v(x)$ as points of the torus that are close to x.

Thus, let f be an Anosov diffeomorphism, let $\{S(x), U(x)\}$ be the corresponding hyperbolic structure on M, and let $C > 0$ and $\lambda \in (0, 1)$ be hyperbolicity constants.

Let g be a diffeomorphism that is C^1-close to f; we search for a homeomorphism h of the manifold M that conjugates f and g.

Let us write the equation

$$h \circ f = g \circ h$$

in the following equivalent form:

$$f^{-1} \circ h \circ f = f^{-1} \circ g \circ h. \tag{8.4}$$

We search for a homeomorphism h of the form $h(x) = x + v(x)$, where v is a small tangent vector field on M.

Remark. In the proof of Theorem 8.1 in the general case, the required homeomorphism h has the form $h(x) = \exp_x(v(x))$, where v is a small tangent vector field on M (and the necessary estimates use uniform estimates of the exponential mapping which follow from relations (8.2)).

Let X be the space of continuous vector fields on M with the uniform norm

$$\|v\| = \max_{x \in M} |v(x)|.$$

Clearly, X is a complete space.

The diffeomorphism $f^{-1} \circ g$ is C^1-close to the identity if the value $\rho_1(f, g)$ is small. Hence, we can write

$$f^{-1} \circ g(x) = x + w_1(x), \tag{8.5}$$

where w_1 is a vector field such that the values

$$\|w_1\| \quad \text{and} \quad \max_{x \in M} \left\| \frac{\partial w_1}{\partial x}(x) \right\|$$

are small if the value $\rho_1(f, g)$ is small.

Define on the space X an operator Φ that assigns to a vector field v a vector field u by the formula

$$u(x) = Df^{-1}(f(x))v(f(x));$$

we can write $\Phi v = Df^{-1}v(f)$.

Let us apply Taylor's formula to the left-hand side of equation (8.4):

$$f^{-1} \circ h \circ f(x) = f^{-1}(f(x) + v(f(x))) = x + \Phi v(x) + s(v)(x).$$

Clearly, $s(0) = 0$ (in this equality, 0 denotes the zero element of the space X).

Since the derivative Df is continuous (and hence, uniformly continuous) on M, there exists a function $G(a)$ such that

$$\text{if } \|v\|, \|v'\| \le a, \quad \text{then } \|s(v) - s(v')\| \le G(a)\|v - v'\|,$$

and $G(a) \to 0$ as $a \to 0$.

Let us transform the right-hand side of (8.4) using representation (8.5):

$$f^{-1} \circ g \circ h(x) = (\mathrm{Id} + w_1)(x + v(x)) = x + v(x) + w(v)(x),$$

where $w(v)(x) = w_1(x + v(x))$.

Since

$$\|w(v) - w(v')\| \le \max_{x \in M} \left\| \frac{\partial w_1}{\partial x}(x) \right\| \|v - v'\|,$$

the Lipschitz constant of the field w in v can be done arbitrarily small if the value $\rho_1(f, g)$ is small enough.

Hence, we can write equation (8.4) in the following form:

$$(\mathrm{Id} - \Phi)(v)(x) = r(v)(x), \tag{8.6}$$

where $r(v) = s(v) - w(v)$.

Note that $r(0) = -w(0)$ (hence, $\|r(0)\| \to 0$ as $\rho_1(f, g) \to 0$) and that the Lipschitz constant of the field r in v on the ball $\{\|v\| \le a\}$ of the space X can be done arbitrarily small if the values a and $\rho_1(f, g)$ are small enough.

Set

$$\mu = RC\frac{1+\lambda}{1-\lambda},$$

where C and λ are hyperbolicity constants of the diffeomorphism f, and the number R is given by the Corollary to Lemma 7.3.

Remark. Of course, it is possible to find the exact value of the constant R for the case of the hyperbolic automorphism of the torus, but we prefer to use the general result of Lemma 7.3 as this is done in the proof of Theorem 8.1 in the general case.

Fix an arbitrary $\varepsilon > 0$; we assume that ε is small enough, so that

$$\|r(v) - r(v')\| \le \frac{1}{2\mu}\|v - v'\| \tag{8.7}$$

if $\|v\|, \|v'\| \le \varepsilon$ and $\rho_1(f, g) \le \varepsilon$.

In addition, we assume that ε is so small that the following three statements are valid:

(a) $2\varepsilon < \Delta$, where the number Δ has the property stated in item (2) of Theorem 7.1; thus, if $x \in M$ and $W^s(2\varepsilon, x)$ and $W^u(2\varepsilon, x)$ are the smooth disks described in the above-mentioned item (2), then

$$W^s(2\varepsilon, x) \cap W^u(2\varepsilon, x) = \{x\};$$

(b) $\varepsilon < 1/4$; hence, equality (8.3) holds for tangent vectors v with $|v| \le \varepsilon$;

(c) if h is a continuous mapping of M into itself such that

$$\max_{x \in M} \mathrm{dist}(x, h(x)) \le \varepsilon,$$

then $h(M) = M$ (such an ε exists due to index theory).

Find a number $\delta \in (0, \varepsilon)$ such that if $\rho_1(f, g) < \delta$, then

$$\|r(0)\| \le \frac{\varepsilon}{2\mu}. \tag{8.8}$$

The main step in solving equation (8.6) is construction of the operator $(\mathrm{Id} - \Phi)^{-1}$.

Consider a vector field $v \in X$ and represent it in the form $v = (v_s, v_u)$, where $v_s(x)$ and $v_u(x)$ are the projections of the vector $v(x)$ to the subspace $S(x)$ parallel to the subspace $U(x)$ and to the subspace $U(x)$ parallel to the subspace $S(x)$, respectively.

Property (HS2) of the hyperbolic structure implies that, according to representation $v = (v_s, v_u)$, the derivative $Df(x)$ is block-diagonal:

$$Df(x) = \mathrm{diag}(Df_s(x), Df_u(x)).$$

It is easily seen that the operator Φ is block-diagonal as well: $\Phi = \mathrm{diag}(\Phi_s, \Phi_u)$.

Define an operator Ψ on the space X by the equality

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

Clearly, $\Psi v(f(x)) = Df(x)v(x)$. The operator Ψ is also block-diagonal: $\Psi = \mathrm{diag}(\Psi_s, \Psi_u)$.

Fix a vector field $v \in X$; let $u = \Psi v$ and $v' = \Phi u$.

The equalities

$$u(f(x)) = Df(x)v(x)$$

and

$$v'(x) = Df^{-1}(f(x))u(f(x)) = Df^{-1}(f(x))Df(x)v(x) = v(x)$$

imply that the operator Φ is the inverse of Ψ. Clearly,

$$\Phi_s\Psi_s = \mathrm{Id} \quad \text{and} \quad \Phi_u\Psi_u = \mathrm{Id} \tag{8.9}$$

(we use the same symbol Id to denote the identity operator in various spaces).

Let us construct the operator $L = (\mathrm{Id} - \Phi)^{-1}$ in block-diagonal form: $L = \mathrm{diag}(L_s, L_u)$, where

$$L_s = -\sum_{k=1}^{\infty} \Psi_s^k \quad \text{and} \quad L_u = \sum_{k=0}^{\infty} \Psi_u^{-k}.$$

Let us prove that the series introduced above are convergent.

Take a vector field v such that $v(x) \in S(x)$ for any $x \in M$. Let $k \geq 0$. By definition,

$$\|\Psi_s^k v\| = \max_{x \in M} |\Psi_s^k v(x)|.$$

There exists a point $x_0 \in M$ such that

$$\|\Psi_s^k v\| = |\Psi_s^k v(x_0)| = |Df^k(f^{-k}(x_0))v(f^{-k}(x_0))|.$$

Property (HS3) of the hyperbolic structure implies that

$$|Df^k(f^{-k}(x_0))v(f^{-k}(x_0))| \le C\lambda^k |v(f^{-k}(x_0))| \le C\lambda^k \max_{x\in M} |v(x)| = C\lambda^k \|v\|.$$

Thus,

$$\|\Psi_s^k\| \le C\lambda^k, \quad k \ge 0. \tag{8.10}$$

Similar estimates show that

$$\|\Psi_u^{-k}\| \le C\lambda^k, \quad k \ge 0. \tag{8.11}$$

The estimates obtained imply that if $m > l$, then

$$\begin{aligned}\Big\| \sum_{k=1}^{m} \Psi_s^k - \sum_{k=1}^{l} \Psi_s^k \Big\| &\le \sum_{k=l+1}^{m} \|\Psi_s^k\| \\ &\le C(\lambda^{l+1} + \cdots + \lambda^m) < C\frac{\lambda^{l+1}}{1-\lambda} \to 0, \quad l \to \infty;\end{aligned}$$

thus, the sequence of partial sums for the series defining the operator L_s is a Cauchy sequence (hence, this series converges since the corresponding space of operators is complete). One can apply the same reasoning to the series defining the operator L_u.

Equalities (8.9) imply that

$$\begin{aligned}(\mathrm{Id} - \Phi_s)\Big(-\sum_{k=1}^{\infty} \Psi_s^k\Big) &= -\sum_{k=1}^{\infty} \Psi_s^k + \Phi_s(\Psi_s + \Psi_s^2 + \cdots) \\ &= -\sum_{k=1}^{\infty} \Psi_s^k + \mathrm{Id} + \sum_{k=1}^{\infty} \Psi_s^k = \mathrm{Id}\end{aligned}$$

and

$$\begin{aligned}(\mathrm{Id} - \Phi_u)\Big(\sum_{k=0}^{\infty} \Psi_u^{-k}\Big) &= \sum_{k=0}^{\infty} \Psi_u^{-k} - \Phi_u(\mathrm{Id} + \Psi_u^{-1} + \Psi_u^{-2} + \cdots) \\ &= \mathrm{Id} + \Psi_u^{-1} + \Psi_u^{-2} + \cdots - \Psi_u^{-1} - \Psi_u^{-2} - \cdots = \mathrm{Id};\end{aligned}$$

hence, $L = (\mathrm{Id} - \Phi)^{-1}$.

Let us estimate the norm of the operator L. Consider a vector field $v \in X$; let $v = (v_s, v_u)$, where $v_s(x)$ and $v_u(x)$ are the projections of the vector $v(x)$ to the subspace $S(x)$ parallel to the subspace $U(x)$ and to the subspace $U(x)$ parallel to the subspace $S(x)$, respectively.

By the corollary to Lemma 7.3, $|v_s(x)| \le R|v(x)|$ and $|v_u(x)| \le R|v(x)|$. Hence, $\|v_s\| \le R\|v\|$ and $\|v_u\| \le R\|v\|$.

Estimates (8.10) and (8.11) imply that

$$\|L_s v_s\| \le \sum_{k=1}^{\infty} C\lambda^k \|v_s\| \le RC \frac{\lambda}{1-\lambda}\|v\|$$

and

$$\|L_u v_u\| \le \sum_{k=0}^{\infty} C\lambda^k \|v_u\| \le RC \frac{1}{1-\lambda}\|v\|.$$

Since $Lv = (L_s v_s, 0) + (0, L_u v_u)$,

$$\|L\| \le RC\left(\frac{\lambda}{1-\lambda} + \frac{1}{1-\lambda}\right) = \mu. \tag{8.12}$$

Now we show that if $\rho_1(f, g) < \delta$, then there exists a continuous vector field $v \in X$ with $\|v\| \le \varepsilon$ that satisfies equation (8.6).

Applying the operator L to equation (8.6), we get an equivalent equation,

$$v = Lr(v). \tag{8.13}$$

Estimates (8.12), (8.7), and (8.8) imply that if $\rho_1(f, g) < \delta$ and $\|v\| \le \varepsilon$, then

$$\|Lr(v)\| \le \mu(\|r(0)\| + \|r(v) - r(0)\|) \le \mu\left(\frac{\varepsilon}{2\mu} + \varepsilon\frac{1}{2\mu}\right) \le \varepsilon,$$

and if $\|v\|, \|v'\| \le \varepsilon$, then

$$\|Lr(v) - Lr(v')\| \le \mu\frac{1}{2\mu}\|v - v'\| \le \frac{1}{2}\|v - v'\|.$$

Thus, the operator Lr maps the ball $S = \{\|v\| \le \varepsilon\}$ of the space X into itself and contracts in this ball. It was noted above that the space X is complete; it follows that the ball S contains (a unique) solution v of equation (8.13) (hence, of equation (8.6)).

Set $h(x) = x + v(x)$. Since the vector field v is continuous, h is a continuous mapping of M into itself. Our choice of ε implies that $h(M) = M$.

Let us show that the mapping h is injective. Assume that $h(p) = h(r)$. Since

$$h \circ f^k(p) = g^k \circ h(p) \quad \text{and} \quad h \circ f^k(r) = g^k \circ h(r), \quad k \in \mathbb{Z},$$

the following equalities hold:

$$h \circ f^k(p) = h \circ f^k(r), \quad k \in \mathbb{Z}.$$

Hence,

$$\text{dist}(f^k(p), f^k(r)) < 2\varepsilon, \quad k \in \mathbb{Z}.$$

We conclude that

$$r \in W^s(2\varepsilon, p) \cap W^u(2\varepsilon, p).$$

The choice of ε implies that $p = r$. Hence, h is a homeomorphism of the manifold M.

□

Chapter 9
Smale's horseshoe and chaos

9.1 Smale's horseshoe

In 1961, S. Smale gave the first example of a structurally stable diffeomorphism having an infinite number of periodic points. The Smale construction was based on the so-called *horseshoe* (the history of this construction is described in Appendix B).

Let us describe this construction in its simplest, linear, variant.

Let $Q = [0, 1] \times [0, 1]$ be a square in the plane of variables (x, y). We construct a mapping of the square Q into the plane in two steps. In the first step, we apply to Q a hyperbolic linear mapping

$$f_0(x, y) = (x/4, 4y).$$

The image of Q under this mapping is a vertical rectangle Q' with height 4 and width $1/4$.

In the second step, we apply a mapping f_1 of the plane that bends the rectangle Q' and places the image over the initial square Q as shown in Figure 2. Figure 2 explains the origin of the term "horseshoe."

We will study the mapping $f = f_1 \circ f_0$. It is easy to show that one can construct a diffeomorphism of the plane whose restriction to Q coincides with f (or a diffeomorphism of the two-dimensional sphere with a similar behavior in a coordinate neighborhood).

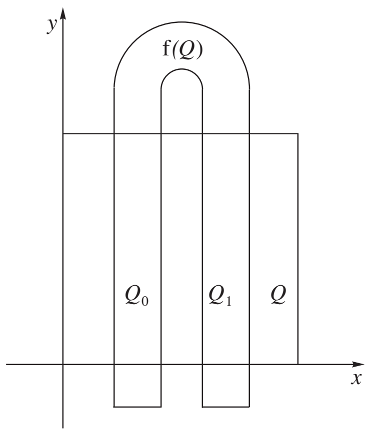


Figure 2. The horseshoe.

Let us treat in detail the behavior of the mapping f_1.

Denote by Q_0 and Q_1 the components of the intersection of $f_1(Q')$ with the square Q (see Figure 2). We call Q_0 and Q_1 vertical rectangles. Denote by Q'_0 and Q'_1 the preimages of Q_0 and Q_1 under the mapping f_1, respectively; we call Q'_0 and Q'_1 horizontal rectangles in Q'.

The sets

$$R_0 = f_0^{-1}(Q'_0) = f^{-1}(Q_0) \quad \text{and} \quad R_1 = f_0^{-1}(Q'_1) = f^{-1}(Q_1)$$

are horizontal rectangles in Q.

We assume that under the action of the mapping f_1 on the rectangles Q'_0 and Q'_1, vertical and horizontal segments are mapped to vertical and horizontal segments of the same length.

By construction, the intersection $f(Q) \cap Q$ is the union of two vertical rectangles Q_0 and Q_1 with height 1 and width $1/4$; we call them first rank vertical rectangles.

It is easy to understand that the intersection $f^2(Q) \cap Q$ consists of four vertical rectangles with height 1 and width $1/16$ (it is useful for the reader to draw the image of the initial square Q under the action of f^2); we call them second rank vertical rectangles.

Let us note that for any of the rectangles Q_0 and Q_1, the intersection $f(Q_i) \cap Q$ consists of two second rank rectangles that lie in different first rank rectangles.

Continuing this process, we conclude that for any natural k, the set $f^k(Q) \cap Q$ is the union of 2^k vertical rectangles with height 1 and width 4^{-k}; we call them rank k vertical rectangles.

Clearly, for any rank k vertical rectangle S, the intersection $f(S) \cap Q$ consists of two rank $k+1$ vertical rectangles that lie in different first rank vertical rectangles.

Passing to actions of negative degrees of the mapping f, we call the rectangles R_0 and R_1 first rank horizontal rectangles.

We note that for any natural k, the intersection $f^{-k}(Q) \cap Q$ consists of 2^k horizontal rectangles with width 1 and height 4^{-k}; we call them rank k horizontal rectangles.

For any rank k horizontal rectangle S, the intersection $f^{-1}(S) \cap Q$ consists of two rank $k+1$ horizontal rectangles that lie in different first rank horizontal rectangles.

Our main object of study is the set

$$\Lambda = \{p \in Q : O(p, f) \subset Q\}.$$

This definition immediately implies that Λ is an invariant set of the mapping f.

Consider a point $p \in Q$. Clearly, if $f^{-1}(p) \in Q$, then $p \in Q_0 \cup Q_1$, i.e., the point p belongs to one of the two first rank vertical rectangles. Further, since $f^{-2}(p) \in Q$ and $f^{-1}(p) \in Q$, the point p belongs to one of the four second rank vertical rectangles, and so on.

Thus, if the negative semitrajectory of a point p lies in Q, then the point p belongs to the intersection of a countable family of vertical rectangles $\{S_k\}$ of all positive ranks in which any rank k vertical rectangle S_k contains the rectangle S_{k+1}.

Clearly, the inverse statement holds as well: If a point p belongs to the intersection of a countable family of vertical rectangles $\{S_k\}$ of all positive ranks with the above property, then the negative semitrajectory of p lies in Q.

This allows us to give the following description of the geometric structure of the set Λ.

Let us call the intersection of a vertical rectangle with the line $\{y = 0\}$ the lower base of the rectangle. The union of the lower bases of rank 1 vertical rectangles is obtained when we exclude three parts from the segment $[0, 1] \times \{0\}$ (see Figure 3).

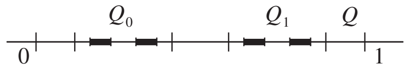


Figure 3. Lower bases of vertical rectangles.

The union of the lower bases of rank 2 vertical rectangles is obtained when we exclude three parts from each of the lower bases of rank 1 vertical rectangles (see Figure 3, where lower bases of rank 2 vertical rectangles are represented by thick lines).

Continuing this process, we construct on the segment $[0, 1] \times \{0\}$ a classical Cantor set (denoted Λ_x). The properties of the mapping f described above imply that the negative semitrajectory of a point $p \in Q$ belongs to the set Q if and only if p is a point of a vertical segment in Q that intersects the segment $[0, 1] \times \{0\}$ at a point of the set Λ_x.

A similar reasoning shows that the segment $\{0\} \times [0, 1]$ contains a Cantor set Λ_y such the positive semitrajectory of a point $p \in Q$ belongs to the set Q if and only if p is a point of a horizontal segment in Q that intersects the segment $\{0\} \times [0, 1]$ at a point of the set Λ_y.

Thus, the set Λ is the product of two one-dimensional Cantor sets: $\Lambda = \Lambda_x \times \Lambda_y$.

The dynamics of f on the set Λ is described by the following Smale theorem. Recall that we denoted by $\mathcal{X}$ the space of binary sequences and by σ the homeomorphism of shift on $\mathcal{X}$ (see Example 1.1).

Theorem 9.1. *There exists a homeomorphism*

$$h : \Lambda \to \mathcal{X}$$

that conjugates f on Λ and σ on $\mathcal{X}$.

Proof. The construction of the homeomorphism h is very visual. Take a point $p \in \Lambda$. Clearly, $f^k(p) \in Q_0 \cup Q_1$ for any k (if $f^k(p) \notin Q_0 \cup Q_1$, then $f^{k-1}(p) \notin Q$).

Assign to the point p a binary sequence $a = \{a_k\}$ as follows:

$$f^k(p) \in Q_{a_k}, \quad k \in \mathbb{Z}$$

(thus, $a_k = 0$ if $f^k(p) \in Q_0$ and $a_k = 1$ if $f^k(p) \in Q_1$). The mapping h is well defined since $Q_0 \cap Q_1 = \emptyset$.

Let us show that $h(f(p)) = \sigma(h(p))$ for all $p \in \Lambda$. Consider points $p \in \Lambda$ and $q = f(p)$; let $a = \{a_k\} = h(p)$ and $b = \{b_k\} = h(q)$.

Since

$$f^k(p) = f^{k-1}(q) \in Q_{b_{k-1}} \quad \text{and} \quad f^k(p) \in Q_{a_k},$$

$a_k = b_{k-1}$, which means that $b = h(f(p)) = \sigma(a) = \sigma(h(p))$.

Let us show that h is injective (as we show below, almost the same reasoning proves that h maps Λ onto $\mathcal{X}$ and is a homeomorphism).

Assume that $h(p) = h(q) = \{a_k\}$ for two points $p, q \in \Lambda$.

Let us analyze the structure of preimages of the rectangles Q_0 and Q_1. Recall that the sets $f^{-1}(Q_0)$ and $f^{-1}(Q_1)$ are rank 1 horizontal rectangles.

The set $f^{-k}(Q_0)$ is the union of rank k horizontal rectangles each of which lies in one of rank $k-1$ horizontal rectangles, and two different horizontal rectangles that lie in $f^{-k}(Q_0)$ cannot lie in one rank $k-1$ horizontal rectangle (the same is true for the set $f^{-k}(Q_1)$).

Since $f(p), f(q) \in Q_{a_1}$, the points p and q belong to one rank 1 horizontal rectangle, namely, to R_{a_1}.

Since $f^2(p), f^2(q) \in Q_{a_2}$, the points p and q belong to $f^{-2}(Q_{a_2})$. The set $f^{-2}(Q_{a_2})$ has as a subset a unique rank 2 horizontal rectangle lying in R_{a_1}; hence, the points p and q belong to one rank 2 horizontal rectangle.

Repeating the same reasoning and using the inclusions $f^k(p), f^k(q) \in Q_{a_k}$, we conclude that the points p and q belong to the intersection of a countable family of horizontal rectangles of all ranks. This implies that p and q belong to a horizontal segment in Q.

A similar reasoning based on the inclusions $f^k(p), f^k(q) \in Q_{a_k}$ with $k \leq 0$ shows that p and q belong to a vertical segment in Q. Hence, $p = q$, and the mapping h is injective.

We apply the same scheme as above to show that the mapping h is surjective. Consider a sequence $a = \{a_k\} \in \mathcal{X}$. As above, we construct a sequence of horizontal rectangles of all positive ranks as follows. We start with the rank 1 horizontal rectangle R_{a_1}; the set $f^{-2}(Q_{a_2})$ contains a unique rank 2 horizontal rectangle that is a subset of $f^{-2}(Q_{a_2})$ which we select as the rank 2 horizontal rectangle of our sequence, and so on.

The intersection L of this sequence of horizontal rectangles is a horizontal segment in Q with the following property: if $p \in L$, then $f^k(p) \in Q_{a_k}$ for all $k > 0$.

We apply the same reasoning to construct a vertical segment L' in Q such that if $p \in L'$, then $f^k(p) \in Q_{a_k}$ for all $k \leq 0$.

Clearly, if $p \in L \cap L'$, then $h(p) = a$.

It is known from the basic course of topology that if g is an injective continuous mapping of a metric compact set K onto $g(K)$, then g is a homeomorphism of K onto $g(K)$.

We have shown that h^{-1} is an injective mapping of $\mathcal{X}$ onto Λ. Thus, to complete the proof of Theorem 9.1 it remains to show that the mapping h^{-1} is continuous.

Take an arbitrary $\varepsilon > 0$ and find a natural number n such that $\sqrt{2} \cdot 4^{-n} < \varepsilon$. There exists a positive number δ such that if two sequences $a, b \in \mathcal{X}$ satisfy the inequality $\mathrm{dist}(a, b) < \delta$, then $a_k = b_k, |k| \leq n$.

Let us show that if $\mathrm{dist}(a, b) < \delta$, then

$$|h^{-1}(a) - h^{-1}(b)| < \varepsilon. \tag{9.1}$$

Indeed, let $p = h^{-1}(a)$ and $q = h^{-1}(b)$. Since $a_k = b_k, 1 \leq k \leq n$, the reasoning applied in the proof of injectivity of the mapping h shows that p and q belong to the same rank n horizontal rectangle.

A similar reasoning based on the equalities $a_k = b_k, 0 \geq k \geq -n$, shows that p and q belong to the same rank n vertical rectangle.

Thus, p and q belong to a square with side 4^{-n}. This proves inequality (9.1). The proof of Theorem 9.1 is complete. □

Remark. Usually, the homeomorphism h constructed in the proof of Theorem 9.1 is called *coding*. The sense of this terminology is as follows: We have selected two disjoint vertical rectangles Q_0 and Q_1 in the square Q; any point of the invariant set Λ is coded by the sequence of indices of vertical rectangles consecutively visited by the trajectory of the point.

Remark. In the proof of Theorem 9.1, we have used not all our assumptions concerning the mapping f.

In fact, we have used only the following properties (a)–(d) of f.

(a) The intersection $f(Q) \cap Q$ consists of two disjoint sets (as above, we call them rank 1 vertical rectangles); the intersection $f^{-1}(Q) \cap Q$ consists of two disjoint sets (as above, we call them rank 1 horizontal rectangles).

Of course, in this case the term "rectangle" is conditional; geometrically, the corresponding sets may significantly differ from usual rectangles.

(b) There exist vertical and horizontal rectangles of all positive ranks; for any rank k vertical rectangle S, the intersection $f(S) \cap Q$ consists of two rank $k + 1$ vertical rectangles belonging to different rank 1 vertical rectangles (and a similar statement holds for horizontal rectangles).

(c) If $S_1 \supset S_2 \supset \cdots$ is a sequence of vertical rectangles of all positive ranks and $S'_1 \supset S'_2 \supset \cdots$ is a sequence of horizontal rectangles of all positive ranks, then the intersection of the sets

$$\bigcap_{k=1}^{\infty} S_k \quad \text{and} \quad \bigcap_{k=1}^{\infty} S'_k$$

is a single point.

(d) For any $\varepsilon > 0$ there exists an index n such that if S is a rank n vertical rectangle and S' is a rank n horizontal rectangle, then the diameter of the set $S \cap S'$ is less than ε.

Under the conditions formulated before Theorem 9.1, the invariant set Λ is hyperbolic with hyperbolicity constants $C = 1$ and $\lambda = 1/4$. To prove this fact, take as spaces of the hyperbolic structure the spaces $S(p) = \{y = 0\}$ and $U(p) = \{x = 0\}$ for all $p \in \Lambda$.

9.2 Chaotic sets

At present, the mathematical literature contains a lot of different definitions of invariant sets of dynamical systems with chaotic behavior (sometimes, a shorter term "chaotic sets" is applied).

The definition of a chaotic set used in this book is close to that given by Devaney in [28] (the original definition by Devaney was given for the case of semi-dynamical systems, see Section 3.1; we define chaotic sets for dynamical systems generated by homeomorphisms).

First we introduce some terminology. Let Λ be a compact invariant set of a dynamical system generated by a homeomorphism f of a metric space (M, dist).

We say that f has *sensitive dependence on initial conditions* on a set Λ if there exists a number $a > 0$ having the following property: Any neighborhood of any point $p \in \Lambda$ contains a point q such that $\text{dist}(f^k(p), f^k(q)) \geq a$ for some $k \in \mathbb{Z}$.

The sense of the above-formulated property is as follows: An arbitrarily small error in the choice of the initial point of a trajectory starting at the set Λ can result in a significant divergence of trajectories.

We say that f is *topologically mixing* on a set Λ if for any couple U, V of open sets such that $U \cap \Lambda \neq \emptyset$ and $V \cap \Lambda \neq \emptyset$ there exists a positive index k such that $f^k(U) \cap V \neq \emptyset$.

Clearly, if a set Λ contains a point whose positive semitrajectory is dense in Λ, then f is topologically mixing on this set.

Finally, we say that a compact invariant set Λ is *chaotic* if

(a) the set Λ contains infinitely many periodic points of f, the set of periods of these points is unbounded, and periodic points are dense in Λ;

(b) f is topologically mixing on the set Λ;

(c) f has sensitive dependence on initial conditions on the set Λ.

Remark. In the course of development of the theory of chaotic sets, it was shown that condition (c) in the above definition is a corollary of conditions (a) and (b) (see, for example, [29]). Nevertheless, it seems reasonable to define a chaotic set formulating conditions (a), (b), and (c) since these three conditions indicate the characteristic properties of chaotic behavior of a dynamical system.

Let us show that the invariant set Λ of Smale's horseshoe is chaotic in the sense of the above definition.

It was shown in Example 1.1 that the shift σ on the space $\mathcal{X}$ satisfies conditions (a) and (b) on the whole space $\mathcal{X}$.

We claim that Theorem 9.1 implies that the diffeomorphism f generating the horseshoe satisfies conditions (a) and (b) on the set Λ.

It was shown in Section 3.1 that a topological conjugacy maps periodic points to periodic points.

Since the shift σ has an infinite set of periodic points, the same is true for the restriction of f to Λ.

It is useful for the reader to check the remaining properties mentioned in conditions (a) and (b).

Let us show that f has sensitive dependence on initial conditions on the set Λ.

The vertical rectangles Q_0 and Q_1 are disjoint compact sets. Let $a > 0$ be the minimal distance between points of these sets. We take this a as the constant in the definition of sensitive dependence on initial conditions.

Let p be an arbitrary point of the set Λ and let U be an arbitrary neighborhood of p in the plane. Set $a = h(p)$.

It follows from Theorem 9.1 that there exists a sequence $b \in \mathcal{X}$ such that $b \neq a$ and $q := h^{-1}(b) \in U$ (why?).

Since the sequences a and b are different, there exists an index $k \in \mathbb{Z}$ such that $a_k \neq b_k$. This means that the points $f^k(p)$ and $f^k(p)$ belong to different rank 1 vertical rectangles. Hence,

$$|f^k(p) - f^k(q)| \geq a,$$

which completes the proof.

9.3 Homoclinic points

We have introduced the notion of a homoclinic point in Section 6.2.

Let us study such points in the simplest case of a planar diffeomorphism.

Thus, let f be a diffeomorphism of the plane. Assume that a point p is a hyperbolic saddle fixed point of f. Denote by $W^s(p)$ and $W^u(p)$, respectively, the one-dimensional stable and unstable manifolds of the point p.

As was said in Section 6.2, homoclinic points are points of intersection of the manifolds $W^s(p)$ and $W^u(p)$ that are different from the fixed point p.

Since the manifolds $W^s(p)$ and $W^u(p)$ are invariant, the trajectory of a homoclinic point consists of homoclinic points.

We say that a point q is a *transverse homoclinic point* of f if the manifolds $W^s(p)$ and $W^u(p)$ are transverse at the point q.

One can show that the existence of a transverse homoclinic point implies the existence of an invariant set whose properties are similar to properties of the invariant set generated by the horseshoe (in particular, this implies that an arbitrary neighborhood of a transverse homoclinic point contains an infinite number of periodic points).

We do not prove this statement; instead, we give a visual explanation.

Consider a transverse homoclinic point q that belongs to a local stable manifolds of the fixed point p (see Figure 4). Take a rectangle Q containing the segment of the local stable manifold with endpoints p and q .

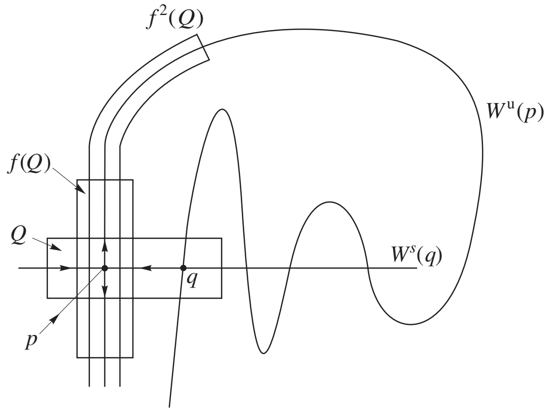


Figure 4. Horseshoe generated by a transverse homoclinic point.

Since the diffeomorphism f contracts in the direction of the stable manifold and expands in the direction of the unstable manifold (in properly chosen coordinates), the images $f(Q), f^2(Q), \ldots$ behave as shown in Figure 4. Under the action of large positive degrees of f, the image of the rectangle Q expands along the unstable manifold $W^u(p)$ and crosses the rectangle Q; thus, we get a horseshoe.

It follows that if a diffeomorphism has a transverse homoclinic point, then it has chaotic invariant sets.

Let us mention that the case of a nontransverse homoclinic point is significantly more complicated. Here we give an example of a diffeomorphism that has a nontransverse homoclinic point while the set of its periodic points is finite.

Consider the flow shown in Figure 5. We assume that if a trajectory does not belong to the set bounded by the closure of the trajectory of the point q, then this trajectory tends to infinity if time goes either to $+\infty$ or to $-\infty$.

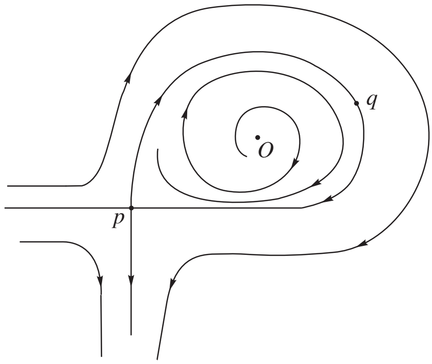


Figure 5. A flow with a nontransverse homoclinic point.

Let f be the shift at time 1 along trajectories of the flow. In this case, q is a nontransverse homoclinic point for the fixed point p; clearly, f does not have periodic points different from o and p.

Chapter 10
Closing Lemma

The following statement is usually called the Closing Lemma.

Theorem 10.1. *Let p be a nonwandering point of a diffeomorphism f of a smooth closed manifold M. Any neighborhood of f in* $\mathrm{Diff}^1(M)$ *contains a diffeomorphism g for which p is a periodic point.*

An analog of Theorem 10.1 has been proven by C. Pugh [30] for the case of flows generated by smooth vector fields. Later, it was shown that the original proof by Pugh contained some inaccuracies, and a corrected proof was published by Pugh and C. Robinson [31].

An important corollary of Theorem 10.1 is the so-called density theorem stated below.

Theorem 10.2. *For a generic diffeomorphism in* $\mathrm{Diff}^1(M)$*, the set of periodic points is dense in the nonwandering set.*

Several modifications of the Closing Lemma have been published; at present, all the existing proofs are very complicated.

Let us state one of recent results on the possibility of connection of close trajectories under a C^1-small perturbation (precisely this technology is the base of the Closing Lemma and its generalizations).

Consider a diffeomorphism f of a smooth closed manifold M and fix a point $z \in M$, a positive number d, and a natural number L.

Consider the set

$$D(z, d, L, f) = \bigcup_{m=1}^{L} f^{-m}(N(d, z)).$$

Theorem 10.3 (a uniform C^1 connecting lemma [32]). *Given an arbitrary neighborhood V of the diffeomorphism f in* $\mathrm{Diff}^1(M)$*, there exist numbers $R > 1$, $d_0 > 0$, a natural number L, and a neighborhood W of the diffeomorphism f in* $\mathrm{Diff}^1(M)$ *having the following property. Let points $p, q, z \in M$, a number $d \in (0, d_0)$, and a diffeomorphism $h \in W$ satisfy the following conditions:*

(1) *The inclusions $h^{-m}(N(z, d)) \subset N(d_0, h^{-m}(z)), m = 1, \dots, L$, hold;*

(2) *the sets $h^{-m}(N(z, d)), m = 1, \dots, L$, are pairwise disjoint;*

(3) $p, q \notin D(z, d, L, f)$;

(4) *the positive semitrajectory* $O^+(p, h)$ *and the negative semitrajectory* $O^-(q, h)$ *intersect the set* $N(d/R, z)$.

Then there exists a diffeomorphism $g \in V$ *coinciding with* h *outside the set* $D(z, d, L, f)$ *and such that* $q \in O^+(p, g)$.

Let us note that an analog of the Closing Lemma in which C^1 topology is replaced by C^0 topology is a rather simple result (we prove it below); at the same time, it is not known whether it is possible to prove an analog of the Closing Lemma in which C^1 topology is replaced by C^2 topology – this is one of the main open problems of the modern theory of dynamical systems (Smale included this problem into his survey "Mathematical problems for the next century" [33]).

Let us prove an analog of Theorem 10.1 for the case of C^0 topology; to simplify presentation, we consider a homeomorphism of a Euclidean space (since our construction is performed in a small neighborhood of a nonwandering point p, almost the same reasoning is applicable in the case of a homeomorphism of a manifold).

Thus, we prove the following statement.

Theorem 10.4. *Let* p *be a nonwandering point of a homeomorphism* f *of the Euclidean space* $\mathbb{R}^n$. *For any* $\varepsilon > 0$ *there exists a homeomorphism* g *of the space* $\mathbb{R}^n$ *such that* p *is a periodic point of* g *and the following inequality holds:*

$$\sup_{x \in \mathbb{R}^n} \max(|f(x) - g(x)|, |f^{-1}(x) - |g^{-1}(x)|) < \varepsilon. \tag{10.1}$$

Let us start with a simple lemma.

Lemma 10.1. *Let* l *be a segment in the Euclidean space* $\mathbb{R}^n$ *with endpoints* a, b. *For any neighborhood* U *of the segment* l *there exists a homeomorphism* h *of the space* $\mathbb{R}^n$ *such that*

(h1) $h(a) = b$;

(h2) $h = \mathrm{Id}$ *outside* U;

(h3)

$$\sup_{x \in \mathbb{R}^n} \max(|h(x) - x|, |h^{-1}(x) - x|) \leq |b - a|. \tag{10.2}$$

Proof. Consider a scalar-valued function $\alpha(x)$ of class C^∞ in $\mathbb{R}^n$ such that

(a1) $\alpha(x) = 1, x \in l$;

(a2) $0 \leq \alpha(x) \leq 1, x \in \mathbb{R}^n$;

(a3) $\alpha(x) = 0$ outside U.

To construct such a function, one can take a neighborhood of the segment l belonging to U and having the form of a parallelepiped; in this neighborhood, one can consider a function $\alpha(x)$ equal to product of functions similar to the function $\eta(t)$ applied in the proof of Lemma 4.2.

Consider the system of differential equations

$$\frac{dx}{dt} = \alpha(x)(b - a) \tag{10.3}$$

and denote by $\phi(t, x)$ the trajectory of system (10.3) with initial point $(0, x)$.

Since

$$a + t(b - a) \in l, \quad t \in [0, 1],$$

the following equalities hold:

$$\phi(t, a) = a + t(b - a),\ t \in [0, 1], \quad \text{and} \quad \phi(1, a) = b. \tag{10.4}$$

Set $h(x) = \phi(1, x)$. It was shown in Section 1.2 that h is a diffeomorphism (hence, h is a homeomorphism) and $h^{-1}(x) = \phi(-1, x)$.

The second equality in (10.4) proves property (h1) of the homeomorphism h.

If $x \notin U$, then x is a rest point of system (10.3); hence, $h(x) = x$, which proves property (h2).

Finally, the inequality $|\alpha(x)(b - a)| \le |b - a|$ implies property (h3) of the homeomorphism h. □

Let us pass to the proof of Theorem 10.4. If p is a fixed point of f, we have nothing to prove. Assume that p is not a fixed point.

There exists a small ball neighborhood V of the point p such that the sets $V' = f^{-1}(V)$, V, and $f(V)$ are pairwise disjoint.

Fix a natural number k (to simplify notation, we do not indicate the dependence of our constructions on k).

By Lemma 3.3, there exists a point r and a number ν such that

$$\max(|r - p|, |q - p|) < \frac{1}{k}, \tag{10.5}$$

where $q = f^{\nu}(r)$; we assume, in addition, that the number k is large enough, so that $r, q \in V$.

Let $y = f(r)$. Consider the set

$$Y = \{f^k(y) : 0 \le k \le \nu - 1\}.$$

Assume that $p \notin Y$ (if $p \in Y$, then we set $f_1 = f$ in the first step of construction of g).

Let $z = f^m(y)$ be the point of the set Y closest to p (if there are several such points, we take as z the point with the minimal value m).

Since $q \in Y$, inequality (10.5) implies that

$$|z - p| < \frac{1}{k}.$$

Let

$$Y_1 = \{f^k(y) : 0 \le k \le m\}.$$

Denote by l_1 the segment with endpoints p and z; the choice of z implies that $Y_1 \cap l_1 = \{z\}$.

Hence, there exists a neighborhood V_1 of the segment l_1 such that $Y_1 \cap V_1 = \{z\}$. It follows from our constructions that $l_1 \subset V$; hence, we may assume that $V_1 \subset V$.

Apply Lemma 10.1 to find a homeomorphism h_1 such that $h_1(z) = p$, $h_1(x) = x$ outside V_1, and

$$\sup_{x \in \mathbb{R}^n} \max(|h_1(x) - x|, |h_1^{-1}(x) - x|) \le \frac{1}{k}. \tag{10.6}$$

Set $f_1 = h_1 \circ f$ (it was mentioned that if $p \in Y$, then we set $f_1 = f$). Note that $f_1(x) = f(x)$ for $x \notin f^{-1}(V_1)$.

Since $Y_1 \cap l_1 = \{z\}$, $f^k(y) \notin f^{-1}(V_1)$ for $0 \le k \le m-1$; hence, $f_1^k(y) = f^k(y), 0 \le k \le m-1$, and $f_1^m(y) = p$ (the last equality is true in the case where $p = f^m(y)$ as well).

Let l_2 be the segment with endpoints p and r.

First we assume that $Y_1 \cap l_2 = \emptyset$. In this case l_2 has a neighborhood V_2 that belongs to V and does not intersect Y_1. Apply Lemma 10.1 to find a homeomorphism h_2 such that $h_2(p) = r$, $h_2(x) = x$ outside V_2, and

$$\sup_{x \in \mathbb{R}^n} \max(|h_2(x) - x|, |h_2^{-1}(x) - x|) \le \frac{1}{k}. \tag{10.7}$$

Set $g = f_1 \circ h_2$. Note that $g(x) = f_1(x)$ for $x \notin V_2$. The relations $f^{-1}(V_1) \subset V'$, $V_2 \subset V$, and $V' \cap V = \emptyset$ imply that $g^k(y) = f^k(y), 0 \le k \le m-1$, $g^m(y) = p$, and $g(p) = y$. Hence, p is a periodic point of g.

To prove that it is possible to guarantee inequality (10.1), let us note that

$$|f_1(x) - f(x)| = |h_1(f(x)) - f(x)|$$

and

$$|f_1^{-1}(x) - f^{-1}(x)| = |f^{-1}(h_1(x)) - f^{-1}(x)|,$$

and we have to estimate the above values on the sets $f^{-1}(V_1)$ and V_1, respectively.

Similarly,

$$|g(x) - f_1(x)| = |f_1(h_2(x)) - f_1(x)|,$$

and we have to estimate this value for $x \in V_2$ only, but if $x \in V_2$, then $f_1(x) = f(x)$.

Hence,

$$|g(x) - f_1(x)| = |f(h_2(x)) - f(x)|$$

for $x \in V_2$.

One can apply a similar reasoning in the estimation of the value $|g^{-1}(x) - f_1^{-1}(x)|$.

We modify the homeomorphism f on a bounded set, where both f and f^{-1} are uniformly continuous. Since k is arbitrary, it follows from inequalities (10.6) and (10.7) that we can achieve inequality (10.1) with an arbitrary $\varepsilon > 0$.

If $Y_1 \cap l_2 \neq \emptyset$, then we construct a homeomorphism h_2 that maps the point p not to r but to the closest to p point of the intersection $Y_1 \cap l_2$ (we leave details of the corresponding proof to the reader). □

Now we formulate one more variant of the Closing Lemma belonging to Mañé [34]. Mañé used this result in the proof of necessity of hyperbolicity of the nonwandering set for structural stability (see Appendix A).

Let f be a diffeomorphism of a smooth closed manifold M.

Consider a measure μ on the set of Borel subsets of M. As usual, μ is called an *f-invariant probability measure* if $\mu(M) = 1$ and $\mu(f^{-1}(A)) = \mu(A)$ for any Borel subset A of M.

Fix a point $x \in M$ and a number $\varepsilon > 0$; let $B(\varepsilon, x, f)$ be the set of points $y \in M$ with the following property: There exist an integer n such that $\mathrm{dist}(f^n(x), y) \leq \varepsilon$ (in other words, $B(\varepsilon, x, f)$ is the closure of the ε-neighborhood of the trajectory of the point x).

Finally, we define a set $\Sigma(f)$ as follows. A point $x \in M$ belongs to $\Sigma(f)$ if for any $\varepsilon > 0$ and for any neighborhood V of the diffeomorphism f in $\mathrm{Diff}^1(M)$ we can find a diffeomorphism $g \in V$, a point $y \in M$, and a natural number m such that

(1) y is a periodic point of g of period m;

(2) $g = f$ on the set $M \setminus B(\varepsilon, x, f)$;

(3) $\mathrm{dist}(f^k(x), g^k(y)) \leq \varepsilon$ for all $0 \leq k \leq m$.

Theorem 10.5. *The equality $\mu(\Sigma(f)) = 1$ holds for any f-invariant probability measure μ.*

The main difference between Theorems 10.1 and 10.5 is as follows. The classical Closing Lemma (Theorem 10.1) states that there exists a diffeomorphism g that is C^1-close to f and such that a point $p \in \Omega(f)$ is a periodic point of g; the trajectories of the point p in dynamical systems generated by f and g are not necessarily close.

In Theorem 10.5, the trajectory of the periodic point y for the perturbed diffeomorphism g is close to the trajectory of the point x for the original diffeomorphism f over the period of y (and the set of points x for which such a closing is possible is full with respect to any f-invariant measure).

Chapter 11
C^0-generic properties of dynamical systems

11.1 Hausdorff metric

In the study of C^0-generic properties of dynamical systems, an important role is played by the Hausdorff metric on the set of compact subsets of the phase space.

Let (M, dist) be a compact metric space. Denote by $\mathcal{C}(M)$ the set of nonempty closed subsets of the space M. Since the space M is compact, any its closed subset is compact as well.

Fix two sets $A, B \in \mathcal{C}(M)$.

The number

$$\text{dev}(A, B) = \max_{x \in A} \text{dist}(x, B)$$

is called the *deviation* of the set A from the set B.

The *Hausdorff distance* between the sets A and B is defined as follows:

$$\text{dist}_H(A, B) = \max(\text{dev}(A, B), \text{dev}(B, A)).$$

Let us show that dist_H is a metric on $\mathcal{C}(M)$.

Clearly, $\text{dist}_H(A, A) = 0$, and if $\text{dist}_H(A, B) = 0$, then $A = B$. Let us prove the triangle inequality. Take $A, B, C \in \mathcal{C}(M)$.

If $x \in A$ and $y \in B$, then

$$\text{dist}(x, C) \leq \text{dist}(x, y) + \text{dist}(y, C) \leq \text{dist}(x, y) + \text{dist}_H(B, C).$$

Hence,

$$\text{dist}(x, C) \leq \min_{y \in B} \text{dist}(x, y) + \text{dist}_H(B, C) = \text{dist}(x, B) + \text{dist}_H(B, C)$$

and

$$\text{dist}(x, C) \leq \text{dist}_H(A, B) + \text{dist}_H(B, C).$$

Hence,

$$\text{dev}(A, C) \leq \text{dist}_H(A, B) + \text{dist}_H(B, C).$$

A similar reasoning shows that

$$\text{dev}(C, A) \leq \text{dist}_H(A, B) + \text{dist}_H(B, C),$$

and we get the required triangle inequality.

Everywhere below, we consider the space $\mathcal{C}(M)$ with the topology induced by the Hausdorff metric.

11.2 Semicontinuous mappings

Let X be a topological space and let (M, dist) be a metric space. Consider a mapping

$$F : X \to \mathcal{C}(M).$$

We say that the mapping F is *upper semicontinuous* if for any point $x \in X$ and any number $\varepsilon > 0$ there exists a neighborhood W of x in X such that

$$F(y) \subset N(\varepsilon, F(x)), \quad y \in W$$

(recall that $N(a, A)$ is the a-neighborhood of a set A).

We say that the mapping F is *lower semicontinuous* if for any point $x \in X$ and any number $\varepsilon > 0$ there exists a neighborhood W of x in X such that

$$F(x) \subset N(\varepsilon, F(y)), \quad y \in W.$$

Fix a number $d > 0$. We say that the mapping F is *d-continuous* at a point $x \in X$ if there exists a neighborhood W of the point x in X such that

$$\text{dist}_H(F(y), F(z)) < d$$

for any $y, z \in W$.

Clearly, a mapping F is continuous at a point x with respect to the Hausdorff metric if and only if F is d-continuous at x for any $d > 0$.

Denote by $V_d(F)$ the set of points of d-continuity for a mapping F.

Lemma 11.1 ([35]). *If the space M is compact and a mapping F is upper or lower semicontinuous, then the set $V_d(F)$ is open and dense in X for any $d > 0$.*

Proof. We prove our lemma for the case of an upper semicontinuous mapping F; the case of a lower semicontinuous mapping is treated similarly.

The set $V_d(F)$ is open by definition. Let us prove that this set is dense.

Since the space M is compact, there exists a finite open covering $U_1, \dots, U_k$ of M such that

$$\text{diam}\, U_i < d, \quad i = 1, \dots, k.$$

Consider the set $K = \{1, \dots, k\}$ as a topological space with discrete topology; in this case, the set $K^* = \mathcal{C}(K)$ coincides with the set of all nonempty subsets of K.

Fix an element $L \in K^*$ and define a set $N_L \subset \mathcal{C}(M)$ as follows: $A \in N_L$ if and only if

(1) $A \cap U_i \neq \emptyset$ for any $i \in L$;

(2) $A \subset \cup_{i \in L} U_i$.

Clearly, the set N_L is open in $\mathcal{C}(M)$ for any L.

Let W be an arbitrary open set in X. We have to show that $W \cap V_d(F) \neq \emptyset$. Consider the set

$$B_W = \{L \in K^* : \text{there exists a point } x \in W \text{ such that } F(x) \in N_L\}.$$

Since the set K^* is finite and partially ordered (with respect to inclusion), the set B_W contains a minimal (with respect to inclusion) element L_0. The minimality of L_0 means that we cannot find an element $L \in B_W$ such that $L \subset L_0$ and $L \neq L_0$.

Take a point $x_0 \in W$ such that $F(x_0) \in N_{L_0}$.

We claim that $x_0 \in V_d(F)$, i.e., the mapping F is d-continuous at the point x_0.

Since the mapping F is upper semicontinuous at x_0, there exists a neighborhood V of x_0 such that

$$F(x) \subset \bigcup_{i \in L_0} U_i, \quad x \in V.$$

We assume that $V \subset W$.

Since L_0 is a minimal element of B_W, the relations

$$F(x) \cap U_i \neq \emptyset, \quad i \in L_0,$$

hold for any $x \in V$.

Consider points $x, y \in V$; let $x' \in F(x)$. There exists an index $i \in L_0$ such that $x' \in U_i$. Find a point $y' \in F(y)$ such that $y' \in U_i$. The inequality $\operatorname{diam} U_i < d$ implies that $\operatorname{dist}(x', y') < d$. Thus,

$$\operatorname{dist}_H(F(x), F(y)) < d,$$

which proves the required density of the set $V_d(F)$. □

Corollary 11.1. *If the space M is compact and a mapping F is upper or lower semi-continuous, then a generic point of the space X is a point of continuity of F.*

11.3 Tolerance stability and Takens' theory

Let $H(M)$ be the space of homeomorphisms of a compact metric space (M, dist) with metric ρ_0 (see Section 2.1). As above, we denote by $\mathcal{C}(M)$ the set of nonempty, closed subsets of the space M. We denote by $\mathcal{C}(\mathcal{C}(M))$ the set of nonempty, closed subsets of the space $(\mathcal{C}(M), \operatorname{dist}_H)$.

For a homeomorphism $f \in H(M)$ and a point $x \in M$, the set $\operatorname{Cl} O(x, f)$ is an element of the space $\mathcal{C}(M)$.

The equality

$$F(f) = \operatorname{Cl}\{\operatorname{Cl} O(x, f) : x \in M\} \tag{11.1}$$

defines a mapping $F : H(M) \to \mathcal{C}(\mathcal{C}(M))$.

Fix $\mathcal{A}, \mathcal{B} \in \mathcal{C}(\mathcal{C}(M))$.

As done in Section 11.1, consider the deviation

$$\mathrm{Dev}(\mathcal{A}, \mathcal{B}) = \max_{A \in \mathcal{A}} \min_{B \in \mathcal{B}} \mathrm{dist}_H(A, B)$$

and the Hausdorff distance

$$\mathrm{Dist}_H(\mathcal{A}, \mathcal{B}) = \max(\mathrm{Dev}(\mathcal{A}, \mathcal{B}), \mathrm{Dev}(\mathcal{A}, \mathcal{B})).$$

The same reasoning as in Section 11.1 shows that Dist_H is a metric on $\mathcal{C}(\mathcal{C}(M))$.

An exercise for the reader: show that if the space M is compact, then both spaces $\mathcal{C}(M)$ and $(\mathcal{C}(\mathcal{C}(M)), \mathrm{Dist}_H)$ are compact as well.

Let D be a subset of $H(M)$. We consider a topology on D that is not coarser than the topology induced by the metric ρ_0 (this means that convergence in the topology considered implies convergence in the standard topology of $H(M)$).

If M is a smooth closed manifold, one may take as D, for example, the space of diffeomorphisms $\mathrm{Diff}^1(M)$ with its standard C^1 topology.

We say that a homeomorphism $f \in H(M)$ is *tolerance D-stable* if f is a continuity point of the restriction $F|_D$.

In the paper [35], F. Takens formulated the following conjecture (and called it *Zeeman's tolerance stability conjecture*): *Any subset $D \subset H(M)$ contains a residual subset D_0 such that any homeomorphism $f \in D_0$ is tolerance D-stable.*

In its general form, the tolerance stability conjecture is not true; a counterexample had been constructed by W. White in [36] (see also [10]).

Nevertheless, Takens obtained in [35] several interesting results related to this conjecture. Let us pass to some of these results.

Consider two homeomorphisms $f, g \in H(M)$; we denote by $O(f)$ and $O(g)$ trajectories of these homeomorphisms, respectively (in this case, the initial points of trajectories are irrelevant for us).

Fix $\varepsilon > 0$. We say that two homeomorphisms f and g are *orbitally ε-equivalent* if the following conditions are satisfied:

(OE1) for any trajectory $O(f)$ there exists a trajectory $O(g)$ such that

(OE1.1) $O(f) \subset N(\varepsilon, O(g))$

(OE1.2) $O(g) \subset N(\varepsilon, O(f))$;

(OE2) for any trajectory $O(g)$ here exists a trajectory $O(f)$ such that

(OE2.1) $O(g) \subset N(\varepsilon, O(f))$

(OE2.2) $O(f) \subset N(\varepsilon, O(g))$.

It is easy to understand that a homeomorphism $f \in D$ is tolerance D-stable if and only if for any $\varepsilon > 0$ there exists a neighborhood V of f in D such that any homeomorphism in V is orbitally ε-equivalent to f.

Takens considered two weakened variants of the property of tolerance D-stability.

As above, fix $\varepsilon > 0$. We say that two homeomorphisms f and g are *minimally ε-equivalent* (*maximally ε-equivalent*) if conditions (OE1.1) and (OE2.1) (respectively, conditions (OE1.2) and (OE2.2)) are omitted in the above definition of orbital ε-equivalence.

Thus, homeomorphisms f and g are maximally ε-equivalent if any trajectory of f belongs to the ε-neighborhood of some trajectory of g and, vice versa, any trajectory of g belongs to the ε-neighborhood of some trajectory of f.

Let, as above, D be a subset of the space $H(M)$. Denote by $D^{\max}$ the subset of D which consists of homeomorphisms f having the following property: For any $\varepsilon > 0$ there exists a neighborhood W of f in D such that any two homeomorphisms $h, g \in W$ are maximally ε-equivalent (recall that we consider the set D with a topology that is not coarser than the standard C^0 topology).

Similarly, we denote by $D^{\min}$ the subset of D which consists of homeomorphisms f having the following property: For any $\varepsilon > 0$ there exists a neighborhood W of f in D such that any two homeomorphisms $h, g \in W$ are minimally ε-equivalent.

Takens had proven the following statement.

Theorem 11.1. *Any of the sets $D^{\max}$ and $D^{\min}$ is residual in D.*

Proof. We give a complete proof for the set $D^{\max}$; the case of the set $D^{\min}$ is treated similarly (we give a comment concerning the difference between the two cases).

Fix $\varepsilon > 0$ and consider the subset Q_ε of the set D consisting of homeomorphisms f with the following property: There exists a neighborhood W of f in D such that any two homeomorphisms $h, g \in W$ are maximally ε-equivalent.

Clearly, any of the sets Q_ε is open in D and the following equality holds:

$$D^{\max} = \bigcap_{\varepsilon > 0} Q_\varepsilon.$$

Thus, to complete the proof we have to show that any of the sets Q_ε is dense in D.

Since the space M is compact, there exists a finite covering of M by open sets $U_1, \dots, U_k$ such that

$$\operatorname{diam} U_i < \varepsilon, \quad i = 1, \dots, k.$$

Similarly to the proof of Lemma 11.1, we consider the set $K = \{1, \dots, k\}$ as a topological space with discrete topology; treating the set $K^* = \mathcal{C}(K)$ as a space with discrete topology as well, we define the set $K^{**} = \mathcal{C}(\mathcal{C}(K))$.

Consider a mapping

$$F^{\max} : D \to K^{**}$$

defined as follows: A subset $L \subset K$ (i.e., an element of the set K^*) is an element of the set $F^{\max}(f)$ if and only if there exists a trajectory $O(f)$ of the homeomorphism f such that

$$O(f) \cap U_i \neq \emptyset, \quad i \in L$$

(this means that the trajectory $O(f)$ intersects every element U_i of our covering with $i \in L$; at the same time, it is not assumed that the trajectory $O(f)$ belongs to the union of elements of the covering with indices $i \in L$).

Let us show that the mapping $F^{\max}$ is lower semicontinuous. Take $f \in D$ and let $L = \{l_1, \dots, l_m\} \subset F^{\max}(f)$.

In this case, there exists a point $x \in M$ and integers $n_1, \dots, n_m$ such that

$$f^{n_i}(x) \in U_{l_i}, \quad i = 1, \dots, m.$$

Since the sets U_i are open, there exists a neighborhood W of f in D such that if $g \in W$, then

$$g^{n_i}(x) \in U_{l_i}, \quad i = 1, \dots, m$$

(we take into account here that convergence in the topology of D implies convergence in the standard topology of $H(M)$).

Thus, $L \subset F^{\max}(f)$. Since the set K^* is finite, it follows that the mapping $F^{\max}$ is lower semicontinuous.

The set K^{**} is discrete; hence, it follows from Lemma 11.1 that there exists an open and dense subset D' of D on which the mapping $F^{\max}$ is locally constant. Let us show that $D' \subset Q_\varepsilon$.

For this purpose, we show that if $F^{\max}(g) = F^{\max}(h)$, then the homeomorphisms g and h are maximally ε-equivalent. Consider a trajectory $O(g)$ and find an element $L \in K^*$ such that

$$O(g) \cap U_i \neq \emptyset, \quad i \in L, \quad \text{and} \quad O(g) \subset \bigcup_{i \in L} U_i.$$

Since $L \in F^{\max}(g)$, $L \in F^{\max}(h)$. Hence, there exists a trajectory $O(h)$ such that

$$O(h) \cap U_i \neq \emptyset, \quad i \in L.$$

Clearly, $O(f) \subset N(\varepsilon, O(h))$.

We see that the homeomorphisms g and h satisfy analogs of conditions (OE1.1) and (OE2.1) in the definition of orbital ε-equivalence. This proves our theorem in the case of the set $D^{\max}$. □

Let us comment the difference between the proofs of Theorem 11.1 for the sets $D^{\max}$ and $D^{\min}$.

Treating the case of the set $D^{\min}$, we again fix $\varepsilon > 0$ and consider a finite covering of M by closed sets $U_1, \dots, U_k$ such that

$$\operatorname{diam} U_i < \varepsilon, \quad i = 1, \dots, k.$$

Consider a mapping

$$F^{\min} : D \to K^{**}$$

defined as follows: A subset $L \subset K$ is an element of the set $F^{\min}(f)$ if and only if there exists a trajectory $O(f)$ of the homeomorphism f such that

$$O(f) \subset \bigcup_{i \in L} U_i.$$

Let us show that the mapping $F^{\min}$ is upper semicontinuous. Take $f \in D$ and consider a set $L \subset K$ that is not an element of $F^{\min}(f)$. In this case, for any point $x \in M$ there exists an index $n(x)$ such that

$$f^{n(x)}(x) \notin \bigcup_{i \in L} U_i.$$

Both sets M and $\cup_{i \in L} U_i$ are compact; hence, there exist numbers $T > 0$ and $d > 0$ such that for any point $x \in M$ there is an index $n(x)$, $|n(x)| \le T$, such that

$$\operatorname{dist}\Big(f^{n(x)}(x), \bigcup_{i \in L} U_i\Big) > d. \tag{11.2}$$

We can find a neighborhood W of f such that, for any homeomorphism $g \in W$, an analog of inequality (11.2) holds with d replaced by $d/2$. Hence, $L \notin F^{\min}(g)$ for $g \in W$. Since the set K^* is finite, we conclude that the mapping $F^{\min}$ is upper semicontinuous.

The rest of the proof repeats the reasoning applied in the case of the set $D^{\max}$.

11.4 Attractors of dynamical systems

Let f be a homeomorphism of a topological space M.

We say that a set $I \subset M$ is *Lyapunov stable* for the dynamical system generated by f if

(LS1) the set I is compact and f-invariant;

(LS2) for any neighborhood U of the set I there exists a neighborhood V of I such that $O^+(V, f) \subset U$ (recall that $O^+(V, f)$ is the positive semitrajectory of the set V in the system f).

We say that a set $I \subset M$ is an *attractor* of the dynamical system f if

(A1) the set I is Lyapunov stable;

(A2) there exists a neighborhood W of the set I such that if $x \in W$, then

$$f^k(x) \to I, \quad k \to \infty. \tag{11.3}$$

Sometimes, the term attractor is replaced by the term *attracting set*.

Consider the set of points $x \in M$ for which relation (11.3) is satisfied; this set is called the *basin* of the attractor I and denoted $D(I)$.

Let us first describe several simple properties of attractors.

Lemma 11.2. *The set $D(I)$ is open.*

Proof. Consider a point $x \in D(I)$. Relation (11.3) implies that there exists an index m such that $f^m(x) \in W$ (here W is the neighborhood mentioned in the definition of the attractor I). There exists a neighborhood N of the point x such that $f^m(N) \subset W$. Clearly, $N \subset D(I)$. □

Lemma 11.3. *If K is a compact subset of $D(I)$ and U is a neighborhood of I, then there exists an index m_0 such that $f^m(K) \subset U$ for $m \geq m_0$.*

Proof. To get a contradiction, let us assume that there exists a compact subset K of the basin $D(I)$ and a neighborhood U of the attractor I with the following property: For any m_0 there exists an index $m \geq m_0$ such that

$$f^m(K) \setminus U \neq \emptyset. \tag{11.4}$$

Find a sequence of points $x_m \in K, m \geq 0$, such that $f^m(x_m) \notin U$. Let y be a limit point of the sequence x_m (for simplicity, we assume that $y = \lim_{m\to\infty} x_m$).

Then $y \in K$; hence, $y \in D(I)$. Since the set I is Lyapunov stable, there exists a neighborhood V of I such that $f^k(V) \subset U$ for $k \geq 0$.

Find an index k_0 such that $f^{k_0}(y) \in V$. If m is large enough, then $f^{k_0}(x_m) \in V$. Hence,

$$f^{k+k_0}(x_m) \in U, \quad k \geq 0.$$

Take $m > k_0$ and $k = m - k_0$ to get a contradiction with relation (11.4). This completes the proof. □

In what follows, we assume that M is a metric space (in particular, we will use the following simple property of metric spaces: If a point x does not belong to a compact set K, then there exists a neighborhood of the set K that does not contain x).

Corollary 11.2. *If K is a compact subset of the basin $D(I)$ and $x \notin I$, then there exists an index m_0 such that $f^m(x) \notin K$ for $m \leq -m_0$.*

Proof. Since the point x does not belong to the compact set I, there exists a neighborhood U of the attractor I that does not contain x. By Lemma 11.3, there exists an index m_0 such that $f^m(K) \subset U$ for $m > m_0$. Clearly, m_0 has the desired property. □

Let us describe a general construction which allows one to establish the existence of an attractor.

Theorem 11.2. *Assume that U is an open subset of M for which there exists a natural number N such that*

$$f^N(\operatorname{Cl} U) \subset U \quad \textit{and} \quad f^{N+1}(\operatorname{Cl} U) \subset U. \tag{11.5}$$

Then the set

$$I = \bigcap_{k \geq 0} f^{kN}(\operatorname{Cl} U) \tag{11.6}$$

is an attractor of the dynamical system f, and $\operatorname{Cl} U \subset D(I)$.

Proof. The first condition in (11.5) implies that

$$\operatorname{Cl} U \supset f^N(\operatorname{Cl} U) \supset f^{2N}(\operatorname{Cl} U) \supset \cdots . \tag{11.7}$$

Hence, the set I defined by equality (11.6) is nonempty and compact.

Note that if l is a natural number, then

$$(l+1)(N+1)N = l(N+1)N + N^2 + N.$$

Hence,

$$f^{(l+1)(N+1)N}(\operatorname{Cl} U) = f^{l(N+1)N+N^2}(f^N(\operatorname{Cl} U) \subset f^{l(N+1)N+N^2}(\operatorname{Cl} U).$$

A similar reasoning shows that

$$\begin{aligned} f^{(l+1)(N+1)N}(\operatorname{Cl} U) &\subset f^{l(N+1)N+N(N-1)}(f^N(\operatorname{Cl} U) \\ &\subset \cdots \subset f^{l(N+1)N}(\operatorname{Cl} U). \end{aligned} \tag{11.8}$$

Relations (11.7) and (11.8) imply that

$$I = \bigcap_{l > 0} f^{l(N+1)N}(\operatorname{Cl} U).$$

Applying the second condition in (11.5), let us define the compact set

$$I_1 = \bigcap_{k \geq 0} f^{k(N+1)}(\operatorname{Cl} U).$$

The same reasoning as above proves that

$$I_1 = \bigcap_{l>0} f^{l(N+1)N}(\mathrm{Cl}\,U).$$

Hence, $I = I_1$.

Now we note that

$$f^N(I) = \bigcap_{k \geq 1} f^{kN}(\mathrm{Cl}\,U) = I.$$

Similarly,

$$f^{N+1}(I) = f^{N+1}(I_1) = I_1 = I.$$

Hence,

$$f^{N+1}(I) = f^N(I), \quad \text{and} \quad f(I) = I.$$

It follows that the set I is f-invariant.

Let us prove that the set I is Lyapunov stable. Let Y be an arbitrary neighborhood of the set I. Since I is a compact f-invariant set, there exists a neighborhood V_0 of the set I such that

$$f^m(V_0) \subset Y, \quad m = 0, 1, \dots, N-1.$$

Relations (11.6) and (11.7) imply the existence of an index m_0 such that

$$f^{mN}(\mathrm{Cl}\,U) \subset V_0, \quad m \geq m_0. \tag{11.9}$$

Set $V = f^{m_0 N}(U)$ and consider an arbitrary $k \geq 0$. Representing k in the form

$$k = lN + l_1, \quad l \geq 0, 0 \leq l_1 \leq N-1,$$

we get the relations

$$f^k(V) = f^{l_1}(f^{lN}(V)) = f^{l_1}(f^{(m_0+l)N}(U)) \subset f^{l_1}(V_0) \subset Y,$$

which imply that I is Lyapunov stable.

A similar reasoning applying the fact that for any neighborhood V_0 of the set I there exists an index m_0 for which relations (11.9) are valid shows that the set I has property (A2). We can take as W any open set containing $\mathrm{Cl}\,U$ and such that $f^N(W) \subset U$ (clearly, any small enough neighborhood of the set $\mathrm{Cl}\,U$ has the latter property).

This implies the desired inclusion $\mathrm{Cl}\,U \subset D(I)$. □

In what follows, we assume that M is a compact metric space.

We say that an attractor I of a dynamical system f is *stable in* $H(M)$ *with respect to the Hausdorff metric* if for any $\varepsilon > 0$ there exists a neighborhood W of f in $H(M)$ such that any dynamical system $g \in W$ has an attractor I' such that

$$\operatorname{dist}_H(I, I') < \varepsilon.$$

M. Hurley studied in [37] stability of attractors in $H(M)$ with respect to the Hausdorff metric.

Theorem 11.3 ([37]). *Any attractor of a generic dynamical system is stable in* $H(M)$ *with respect to the Hausdorff metric.*

We do not prove the Hurley theorem here; we get it as a corollary of a more general statement.

Let us introduce one more metric on the set $\mathcal{C}(M)$; let

$$R_0(A, B) = \max(\operatorname{dist}_H(A, B), \operatorname{dist}_H(\operatorname{Cl}(M \setminus A), \operatorname{Cl}(M \setminus B))$$

(we preserve the original notation introduced by the author in [10]). It is easily shown that R_0 is a metric.

Exercise 11.1. Show that the metrics dist_H and R_0 generate different topologies on the set $\mathcal{C}(D^2)$, where D^2 is the two-dimensional disk.

We say that an attractor I of a dynamical system f is *stable in* $H(M)$ *with respect to the metric* R_0 if for any $\varepsilon > 0$ there exists a neighborhood W of f in $H(M)$ such that any dynamical system $g \in W$ has an attractor I' such that

$$R_0(I, I') < \varepsilon.$$

Theorem 11.4. *Any attractor of a generic dynamical system is stable in* $H(M)$ *with respect to the metric* R_0.

Clearly, $\operatorname{dist}_H(A, B) \le R_0(A, B)$; hence, Theorem 11.3 is a corollary of Theorem 11.4.

First we prove a simple lemma.

Lemma 11.4. *If* $I \in \mathcal{C}(M)$, *then*

$$\lim_{d \to 0} \operatorname{dist}_H(M \setminus N(d, I)), \operatorname{Cl}(M \setminus I)) = 0.$$

Proof. Clearly,

$$M \setminus N(d, I) \subset \operatorname{Cl}(M \setminus I).$$

Thus, we have to show that

$$\lim_{d \to 0} \operatorname{dev}(M \setminus N(d, I)), \operatorname{Cl}(M \setminus I)) = 0. \tag{11.10}$$

Assuming the converse, we can find a number $a > 0$ and a sequence of points $x_k \in \operatorname{Cl}(M \setminus I)$ such that

$$\operatorname{dist}(x_k, N(1/k, I)) \geq a.$$

Let x be a limit point of the sequence x_k (such a point exists since the space M is compact). If k is large enough, then

$$\operatorname{dist}(x, M \setminus N(1/k, I)) \geq a/2. \tag{11.11}$$

If $x \notin I$, then $x \notin N(1/k, I)$ for large k, which shows that inequalities (11.11) cannot hold.

If $x \in I$, then $x \in \partial I$. Find a sequence of points $y_l \in M \setminus I$ such that $y_l \to x$ as $l \to \infty$. There exists an index m such that

$$\operatorname{dist}(y_m, x) < a/2. \tag{11.12}$$

Since $y_m \notin I$, $y_m \notin N(1/k, I)$ for large k, and relations (11.11) and (11.12) are contradictory.

This proves equality (11.10). □

Now we pass to the proof of Theorem 11.4.

Let us extend the Hausdorff metric dist_H from the set $\mathcal{C}(M)$ to the set $\mathcal{C}(M) \cup \{\emptyset\}$ by setting

$$\operatorname{dist}_H(\emptyset, A) = \operatorname{diam} M$$

for $A \neq \emptyset$. It is easily seen (check!) that Lemma 11.1 and Corollary 11.1 remain valid for mappings

$$F : X \to \mathcal{C}(M) \cup \{\emptyset\}.$$

Let U be a nonempty open subset of M. We assign to this set two mappings

$$G^+, G^- : H(M) \to \mathcal{C}(M) \cup \{\emptyset\}$$

as follows. If a dynamical system f has an attractor I such that

$$I \subset U \subset \operatorname{Cl} U \subset D(I), \tag{11.13}$$

then we set

$$G^+(f) = I \quad \text{and} \quad G^-(f) = \operatorname{Cl}(M \setminus I).$$

Otherwise, we set

$$G^+(f) = M \quad \text{and} \quad G^-(f) = \emptyset.$$

To show that the mappings G^+ and G^- are well defined, we prove that a dynamical system f has not more than one attractor that satisfies conditions (11.13).

To get a contradiction, let us assume that a dynamical system f has two attractors $I_1 \neq I_2$ such that

$$I_1 \subset U \subset \mathrm{Cl}\, U \subset D(I_1)$$

and

$$I_2 \subset U \subset \mathrm{Cl}\, U \subset D(I_2).$$

Let $I_1 \setminus I_2 \neq \emptyset$; consider a point $x \in I_1 \setminus I_2$. Since the set I_1 is f-invariant, $O(x, f) \subset I_1 \subset \mathrm{Cl}\, U$. We apply Corollary 11.2 to x and I_2 to show that $f^m(x) \notin \mathrm{Cl}\, U$ for some m. The contradiction obtained shows that the mappings G^+ and G^- are well defined.

Now we prove that the mapping G^+ is upper semicontinuous and the mapping G^- is lower semicontinuous.

We start with the mapping G^+. Fix a dynamical system $f \in H(M)$. If $G^+(f) = M$, then the mapping G^+ is obviously upper semicontinuous at f. Now we assume that f has an attractor I satisfying condition (11.13), so that $G^+(f) = I$. We select a small $d > 0$ such that $N(d, I) \subset U$.

By Lemma 11.3, there exists a natural number n such that

$$f^n(\mathrm{Cl}\, U) \subset N(d, I) \quad \text{and} \quad f^{n+1}(\mathrm{Cl}\, U) \subset N(d, I).$$

We can find a neighborhood W of the dynamical system f in $H(M)$ such that if $g \in W$, then

$$g^n(\mathrm{Cl}\, U) \subset N(d, I) \quad \text{and} \quad g^{n+1}(\mathrm{Cl}\, U) \subset N(d, I).$$

By Theorem 11.2, there exists an attractor I' of g such that

$$I' \subset U \subset \mathrm{Cl}\, U \subset D(I').$$

Clearly, $G^+(g) = I'$, which proves that the mapping G^+ is upper semicontinuous at f.

Now we consider the mapping G^- and a dynamical system f. If $G^-(f) = \emptyset$, then the mapping G^- is obviously lower semicontinuous at f. Now we assume that f has an attractor I satisfying condition (11.13), so that $G^-(f) = \mathrm{Cl}(M \setminus I)$.

Fix a small $d > 0$. By Lemma 11.4, there exists a small $b > 0$ such that $N(b, I) \subset U$ and

$$\operatorname{dist}_H(M \setminus N(b, I), \operatorname{Cl}(M \setminus I)) < d. \tag{11.14}$$

The same reasoning as in the case of the mapping G^+ shows that there exists a neighborhood W of f in $H(M)$ such that any dynamical system $g \in W$ has an attractor I' such that

$$I' \subset N(b, I) \subset U \subset \operatorname{Cl} U \subset D(I').$$

We prove that in this case, the inclusion

$$\operatorname{Cl}(M \setminus I) \subset N(d, \operatorname{Cl}(M \setminus I')) \tag{11.15}$$

holds, which implies that

$$G^-(g) \subset N(d, G^-(f)).$$

Consider a point $x \in \operatorname{Cl}(M \setminus I)$. Since $I' \subset N(d, I)$,

$$M \setminus N(d, I) \subset M \setminus I' \subset \operatorname{Cl}(M \setminus I'),$$

and it follows that

$$\operatorname{dist}(x, \operatorname{Cl}(M \setminus I')) \leq \operatorname{dist}(x, M \setminus N(d, I)) < d$$

(here we refer to inequality (11.14)). This proves inclusion (11.15); since $d > 0$ is arbitrary, we conclude that the mapping G^- is lower semicontinuous at f.

It was assumed that M is a compact metric space. Hence, the topology of M has a countable base. It follows that we can find a countable family $\mathcal{U} = \{U_m, m \in \mathbb{Z}\}$ of open subsets of M having the following property: For any pair K_1, K_2 of disjoint compact subsets of M there exists a set U_m from the family $\mathcal{U}$ such that

$$K_1 \subset U_m \quad \text{and} \quad K_2 \cap \operatorname{Cl} U_m = \emptyset.$$

To any set U_m from the family $\mathcal{U}$ we assign the corresponding mappings G_m^+ and G_m^-; let C_m^+ and C_m^- be the sets of continuity points of the mappings G_m^+ and G_m^-, respectively.

Since the mappings G_m^+ and G_m^- are semicontinuous, Corollary 11.1 implies that the set

$$C^* = \bigcap_{m \in \mathbb{Z}} (C_m^+ \cap C_m^-)$$

is a residual subset of $H(M)$.

Let us show that if $f \in C^*$, then any attractor I of the dynamical system f is stable in $H(M)$ with respect to the metric R_0.

Let I be an attractor I of a dynamical system $f \in C^*$ (we only consider the case $I \neq M$ and leave the easy case $I = M$ to the reader). The sets I and $M \setminus D(I)$ are disjoint compact sets; hence, there exists a set U_m from the family $\mathcal{U}$ such that

$$I \subset U_m \subset \operatorname{Cl} U_m \subset D(I).$$

Fix an arbitrary $\varepsilon > 0$. Since the mappings G_m^+ and G_m^- are continuous at f, there exists a neighborhood W of f in $H(M)$ such that if $g \in W$, then

$$\operatorname{dist}_H(G_m^+(g), G_m^+(f)) < \varepsilon \quad \text{and} \quad \operatorname{dist}_H(G_m^-(g), G_m^-(f)) < \varepsilon. \tag{11.16}$$

By the definition of the mappings G_m^+ and G_m^-, the set $I' = G_m^+(g)$ is an attractor of the dynamical system g, and $G_m^-(g) = \operatorname{Cl}(M \setminus I')$. Inequalities (11.16) imply that $R_0(I, I') < \varepsilon$. The theorem is proven. □

Attractors that are stable in $H(M)$ with respect to the metric R_0 have several important properties. Let us mention one of such properties (to simplify presentation, we assume that the phase space is a smooth closed manifold, though an analog of the result which we prove is valid for more general phase spaces).

Theorem 11.5. *Let M be a smooth closed manifold. If an attractor I of a dynamical system $f \in H(M)$ is different from M and stable in $H(M)$ with respect to the metric R_0, then its boundary $J = \partial I$ is Lyapunov stable.*

Proof. To get a contradiction, assume that the boundary J of the attractor I is not Lyapunov stable. In this case, there exists a number $a > 0$ and sequences of points x_m and numbers k_m such that

$$\operatorname{dist}(x_m, J) \to 0, \ m \to \infty; \quad \operatorname{dist}(f^{k_m}(x_m), J) \geq a.$$

Clearly, in this case the attractor I does not coincide with its boundary J (since the attractor itself is Lyapunov stable); thus, its interior $\operatorname{Int} I$ is not empty.

In addition, the above-mentioned points x_m must belong to $\operatorname{Int} I$, and the numbers k_m must tend to ∞ (since J is a compact f-invariant set).

Let $x \in J$ be a limit point of the sequence x_m (for simplicity, we assume that $x_m \to x$, $m \to \infty$). Since the point x belongs to the boundary of I, there exists a sequence of points $y_m \in M \setminus I$ such that $y_m \to x$, $m \to \infty$.

Since the set $D(I)$ is open, there exists a number $b > 0$ such that

$$\operatorname{Cl}(N(b, I)) \subset D(I).$$

By Corollary 11.2, there exist numbers $l_m < 0$ such that $f^{l_m}(y_m) \notin \operatorname{Cl}(N(b, I))$.

Set $\eta_m = f^{l_m}(y_m)$ and $\xi_m = f^{k_m}(x_m)$. Consider the sets

$$Y_m = \{f^l(y_m) : l_m \le l \le 0\}$$

and

$$X_m = \{f^k(x_m) : 0 \le l \le k_m\}.$$

Take a small coordinate neighborhood V of the point x and assume that this neighborhood is convex in local coordinates (further, considering points of the neighborhood V, we work with local coordinates of these points; thus, we may assume that we work in a domain of a Euclidean space, which allows us to apply Lemma 10.1).

We take m so large that the points x_m and y_m belong to V; let L_m be the segment with endpoints x_m and y_m.

First we assume that

$$X_m \cap L_m = \{x_m\} \quad \text{and} \quad Y_m \cap L_m = \{y_m\}. \tag{11.17}$$

In this case, there exists a neighborhood U_m of the segment L_m such that $X_m \cap U_m = \{x_m\}$ and $Y_m \cap U_m = \{y_m\}$.

Apply Lemma 10.1 to find a homeomorphism h_m of the manifold M such that $h_m(y_m) = x_m$ and h_m coincides with Id outside the neighborhood U_m. Set $g_m = h_m \circ f$.

Clearly, we can find homeomorphisms h_m such that

$$\rho_0(g_m, f) \to 0, \quad m \to \infty. \tag{11.18}$$

Since $g_m(y) = y$ for $y \notin (X_m \cup Y_m) \setminus \{y_m, f^{-1}(y_m)\}$ and $g_m(f^{-1}(y_m)) = f(x_m)$, the following inclusion holds:

$$\xi_m \in O(\eta_m, g_m). \tag{11.19}$$

By our assumption, the attractor I is stable in $H(M)$ with respect to the metric R_0; hence, it follows from relation (11.18) that the dynamical systems g_m have attractors I_m such that

$$R_0(I, I_m) \to 0, \quad m \to \infty. \tag{11.20}$$

If $\eta_m \in I_m$, then

$$\text{dist}_H(I_m, I) \ge b. \tag{11.21}$$

If $\eta_m \notin I_m$, then inclusion (11.19) implies that $\xi_m \in M \setminus I_m$, and

$$\text{dist}_H(\text{Cl}(M \setminus I), \text{Cl}(M \setminus I_m)) \ge a. \tag{11.22}$$

Relation (11.20) implies that none of inequalities (11.21) and (11.22) is satisfied for large m. The contradiction obtained completes the proof of Theorem 11.5 in the case where condition (11.17) is valid.

If condition (11.17) is not valid, we find in the segment L_m closest points y' and x' of the sets Y_m and X_m. After that, we repeat the above reasoning (replacing the sets Y_m and X_m by the "segments" of trajectories with endpoints η_m and y' for the first segment and endpoints x' and ξ_m for the second segment, respectively). □

We agree that the empty set (the boundary of the attractor that coincides with the whole manifold) is Lyapunov stable. Then we can formulate an important corollary of Theorems 11.4 and 11.5.

Theorem 11.6. *Let M be a smooth closed manifold. A generic dynamical system $f \in H(M)$ has the following property: The boundary of any attractor is Lyapunov stable.*

Exercise 11.2. Give an example of a dynamical system that has an attractor I such that I is stable in $H(M)$ with respect to the Hausdorff metric and its boundary is not Lyapunov stable.

Let $I \subset M$ be a nonempty compact set. We say that I is a *quasiattractor* of a dynamical system f if there exists a countable family I_m, $m \geq 0$ of attractors of f such that

$$I = \bigcap_{m \geq 0} I_m.$$

Clearly, any quasiattractor is f-invariant.

Define the basin $D(I)$ of a quasiattractor I as the set of points $x \in M$ for which relation (11.3) holds.

Note that, in contrast to the case of an attractor, the basin of a quasiattractor is not necessarily open.

Example 11.1. Let $M = S^1$ with coordinate $x \in [0, 1)$. Consider a smooth real-valued function $F(x)$ that has the following properties:

$$F(x) = 0, \quad x \in \{0, 1/2, 1/3, \dots\},$$
$$F(x) > 0, \quad x \in (1/2, 1),$$

and

$$F(x) < 0, \quad x \in (1/3, 1/2) \cup (1/4, 1/3) \cup \cdots .$$

Let $\phi(t, x)$ be the flow on S^1 generated by the system of differential equations

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

and let $f(x) = \phi(1, x)$.

Clearly, any of the sets

$$I_m = [0, 1/(m+3)], \quad m \geq 0,$$

is an attractor of f, and the set

$$I = \bigcap_{m \geq 0} I_m$$

is a quasiattractor.

The basin of I is the set $(1/2, 1) \cup \{0\}$ that is not open.

Exercise 11.3. Show that a quasiattractor is Lyapunov stable.

It is easy to understand that the converse statement is not true in general. Consider as an example the dynamical system generated by the identity mapping of a compact metric space M. Clearly, any compact subset of M is Lyapunov stable; at the same time, only the space M itself is the unique attractor (and quasiattractor).

Such a statement is true for C^0-generic systems.

Theorem 11.7. *Let M be a smooth closed manifold. Any Lyapunov stable set of a generic system $f \in H(M)$ is a quasiattractor.*

The reader can find a proof of Theorem 11.7 in the book [10].

It follows from theorems 11.6 and 11.7 that if M is a smooth closed manifold, then the boundary of any attractor of a generic dynamical system $f \in H(M)$ is a quasiattractor.

Chapter 12
Shadowing of pseudotrajectories in dynamical systems

12.1 Definitions and results

Let f be a homeomorphism of a metric space (M, dist).

Fix a $d > 0$. A sequence

$$\xi = \{x_k \in M : k \in \mathbb{Z}\} \tag{12.1}$$

is called a *d-pseudotrajectory* of the dynamical system f if the following inequalities hold:

$$\mathrm{dist}(x_{k+1}, f(x_k)) < d, \quad k \in \mathbb{Z}. \tag{12.2}$$

The basic property of dynamical systems related to the notion of a pseudotrajectory is called shadowing.

We say that a dynamical system f has the *shadowing property* if for any $\varepsilon > 0$ we can find a $d > 0$ such that for any d-pseudotrajectory ξ of f there exists a point $x \in M$ such that

$$\mathrm{dist}(x_k, f^k(x)) < \varepsilon, \quad k \in \mathbb{Z} \tag{12.3}$$

(in this case, we say that the pseudotrajectory ξ is ε-shadowed by the exact trajectory of the point x).

If a dynamical system has the shadowing property, then, for any its approximate trajectory (obtained, for example, by a numerical method with good accuracy) there is a close exact trajectory. In this case, an approximate pattern of trajectories given by numerical modeling reflects the exact structure of trajectories.

In parallel to the shadowing property defined above, one more property (called the inverse shadowing property) is studied. In this case, the problem is formulated as follows. Assume that we are given a family of pseudotrajectories. Given an exact trajectory, can we find a close pseudotrajectory in this family?

Of course, in this case the answer depends not only on the dynamical system but also on the given family of pseudotrajectories. At present, the property of inverse shadowing (first posed in the author's paper [38]) is studied for a wide class of families of pseudotrajectories.

In this book, we restrict ourselves to consideration of one of such families introduced in [39].

A family of continuous mappings $\psi_k : M \to M$, $k \in \mathbb{Z}$, is called a *d-method* if the following inequalities hold:

$$\operatorname{dist}(\psi_k(x), f(x)) < d, \quad x \in M, k \in \mathbb{Z}. \tag{12.4}$$

We say that a pseudotrajectory of the form (12.1) is generated by a d-method $\{\psi_k\}$ if

$$x_{k+1} = \psi_k(x_k), \quad k \in \mathbb{Z}.$$

Finally, we say that a dynamical system f has the *inverse shadowing property* if for any $\varepsilon > 0$ we can find a $d > 0$ such that for any point $x \in M$ and any d-method $\{\psi_k\}$ there exists a pseudotrajectory ξ generated by this method for which inequalities (12.3) hold.

The problem of shadowing first studied by Anosov and Bowen became one of the important problems in the global qualitative theory of dynamical systems (let us mention the monographs [11] and [40] devoted to this problem).

Really less is known about the inverse shadowing property (let us mention the papers [41, 39]).

In parallel to the above-formulated properties, their Lipschitz variants are studied.

We say that a dynamical system f has the *Lipschitz shadowing property* if there exist positive constants $\mathcal{L}$ and d_0 such that for any d-pseudotrajectory ξ with $d \le d_0$ there exists a point $x \in M$ such that

$$\operatorname{dist}(x_k, f^k(x)) \le \mathcal{L}d, \quad k \in \mathbb{Z}. \tag{12.5}$$

We say that a dynamical system f has the *Lipschitz inverse shadowing property* if there exist positive constants $\mathcal{L}$ and d_0 such that for any point $x \in M$ and any d-method $\{\psi_k\}$ with $d \le d_0$ there exists a pseudotrajectory ξ generated by this method for which inequalities (12.5) hold.

In this book, we prove the following basic statements of the shadowing theory (to simplify presentation, we work with diffeomorphisms of Euclidean spaces).

Theorem 12.1. *If Λ is a hyperbolic set of a diffeomorphism f, then f has the Lipschitz shadowing property in a neighborhood of the set Λ.*

In the proof of Theorem 12.1, we use a construction suggested in the paper [41].

To formulate our theorem on inverse shadowing, we introduce one important notion from the theory of structural stability. Let f be a diffeomorphism of a smooth manifold M.

Fix numbers $C > 0$ and $\lambda \in (0, 1)$ and a point $p \in M$. We say that f has *(C, λ)-structure* at the trajectory $O(p, f)$ if for any point $p_i = f^i(p)$ there exist linear subspaces $S(p_i)$ and $U(p_i)$ of the tangent space $T_{p_i}M$ such that

(C1) $T_{p_i}M = S(p_i) \oplus U(p_i)$;

(C2) $Df(p_i)S(p_i) \subset S(p_{i+1})$ and $Df^{-1}(p_i)U(p_i) \subset U(p_{i-1})$;

(C3.1) $|Df^k(p_i)v| \le C\lambda^k|v|$ for $v \in S(p_k)$ and $k \ge 0$;

(C3.2) $|Df^{-k}(p_i)v| \le C\lambda^k|v|$ for $v \in U(p_k)$ and $k \le 0$;

(C4) if $P(p_i)$ and $Q(p_i)$ are the projections in $T_{p_i}M$ onto $S(p_i)$ parallel to $U(p_i)$ and onto $U(p_i)$ parallel to $S(p_i)$, respectively, then $\|S(p_i)\|$, $\|U(p_i)\| \le C$.

An important step in the proof of Theorem 7.6 (Axiom A and the strong transversality condition imply structural stability) was the proof of existence of a (C, λ)-structure (with the same C, λ) at any trajectory (let us note that our definition differs from the original definition introduced in [42]).

Let us note the main difference between the definition of hyperbolicity (see Section 7.1) and that of (C, λ)-structure. Property (C1) is equivalent to property (HS2.1) (taking Lemma 7.2 into account); inequalities (C3.1)–(C3.2) have the same form as inequalities (HS2.2)–(HS2.3). The main difference is as follows: equalities (HS2.2) are replaced by inclusions (C2).

The corollary to Lemma 7.3 shows that in the case of a hyperbolic set, the norms of the projections $P(p)$ and $Q(p)$ are bounded by a constant that depends on the hyperbolicity constants C, λ in inequalities (HS2.2)–(HS2.3) and on the estimate of the norm of $\|Df(x\|$.

In the case of a structurally stable diffeomorphism (for which there exist (C, λ)-structures with the same C, λ at any trajectory), it is possible to find (C, λ)-structures with an arbitrarily large constant C in property (C4) even if the constants in inequalities (C3.1) and (C3.2) are fixed.

Example 12.1. Consider a diffeomorphism f of the two-dimensional sphere S^2 such that the "north pole" P_0 and the "south pole" P_1 are hyperbolic fixed points (P_1 is an attracting fixed point, and P_0 is a repellor), while trajectories of all points different from P_0 and P_1 tend to P_1 as $k \to \infty$ and to P_0 as $k \to -\infty$.

Clearly, $\Omega(f) = P_0 \cup P_1$; thus, f satisfies Axiom A. Since the stable manifold $W^s(P_1)$ and the unstable manifold $W^u(P_0)$ are two-dimensional, they are transverse at any point of intersection; thus, the strong transversality condition is satisfied. By Theorem 7.6, f is structurally stable.

We can construct the diffeomorphism f so that it is given by

$$f(x_1, x_2) = (2x_1, 2x_2)$$

in local coordinates $x = (x_1, x_2)$ in a neighborhood U_0 of the point P_0 and by

$$f(y_1, y_2) = (y_1/2, y_2/2)$$

in local coordinates $y = (y_1, y_2)$ in a neighborhood U_1 of the point P_1. We also may assume that $U_0 = \{|x| < a\}$ and $U_1 = \{|y| < a\}$ for some $a > 0$.

Set $S(p) = \{0\}$ and $U(p) = T_p S^2$ for $p \in U_0$ and $S(p) = T_p S^2$ and $U(p) = \{0\}$ for $p \in U_1$. Set $\lambda = 1/2$.

Clearly, if $p \in U_0$ and $v \in U(p)$, then

$$|Df^{-k}(p)v| \leq \lambda^k |v|, \quad k \geq 0;$$

similarly, if $p \in U_1$ and $v \in S(p)$, then

$$|Df^{k}(p)v| \leq \lambda^k |v|, \quad k \geq 0.$$

Let

$$N = \max_{p \in S^2} \max(\|Df(p)\|, \|Df^{-1}(p)\|).$$

By Theorem 3.3, there exists a number $n > 0$ such that if $p \in S^2 \setminus (U_0 \cup U_1)$, then $f^k(p) \in U_0$ for $k \leq -n$ and $f^k(p) \in U_1$ for $k \geq n$.

Take a point $p \in S^2 \setminus (U_0 \cup U_1)$ and find for it numbers $n_0 < 0$ and $n_1 > 0$ such that $f^{n_0}(p) \in U_0$, $f^{n_0+1}(p) \notin U_0$, $f^{n_1}(p) \in U_1$, and $f^{n_1-1}(p) \notin U_1$ (clearly, in this case $f^k(p) \in U_0$ for $k \leq n_0$ and $f^k(p) \in U_1$ for $k \geq n_1$).

It follows from our choice that $|n_0|, n_1 \leq n$.

Take as the subspaces $S(p)$ and $U(p)$ any two transverse one-dimensional subspaces of $T_p S^2$ and set

$$S(f^i(p)) = Df^i S(p) \quad \text{and} \quad U(f^i(p)) = Df^i U(p), \quad n_0 + 1 \leq i \leq n_1 - 1$$

(at points $f^i(p)$ with $i \leq n_0$ and $i \geq n_1$, the subspaces S and U have been previously defined).

Let us show that for the defined subspaces, properties (C1)-(C3) from the definition of a (C, λ)-structure hold with $\lambda = 1/2$ and $C = (2N)^n$.

If $v \in S(p)$, then the following inequalities hold:

$$|Df^k(p)v| \leq N^k |v|, \quad 0 \leq k \leq n_1,$$

and

$$|Df^k(p)v| \leq N^{n_1} 2^{-(k-n_1)}, \quad n_1 < k.$$

Now inequalities (C3.1) follow from the obvious estimates

$$N^{n_1} 2^{n_1-k} \leq (2N)^n 2^{-k}, \quad 0 \leq k \leq n_1.$$

The same reasoning proves inequalities (C3.2).

Clearly, one can construct similar subspaces S and U for any point of the sphere S^2.

At the same time, $S(p)$ and $U(p)$ are one-dimensional subspaces of the tangent space $T_p S^2$ such that the angle between these subspaces can be arbitrarily small (in which case, the norms of the projections in $T_p S^2$ onto $S(p)$ parallel to $U(p)$ and onto $U(p)$ parallel to $S(p)$, respectively, are arbitrarily large).

Thus, f has trajectories with (C, λ)-structure for which the constant C in property (C4) is arbitrarily large, while the constants $\lambda = 1/2$ and $C = (2N)^n$ in inequalities (C3.1) and (C3.2) are fixed.

We formulate and prove a theorem on conditions of inverse shadowing in a local variant (i.e., for a single trajectory of a diffeomorphism); it is easy to understand that if f has (C, λ)-structure with the same C, λ at any trajectory and there exists a number $r > 0$ such that condition (12.6) below is satisfied at any point $q \in \mathbb{R}^n$, then f has the Lipschitz inverse shadowing property on the whole space.

Theorem 12.2. *Assume that a diffeomorphism f of the space $\mathbb{R}^n$ has (C, λ)-structure at a trajectory $O(p, f)$. Assume also that there exists a number $r > 0$ such that*

$$|f(q + v) - f(q) - Df(q)v| \le \frac{1}{2\mathcal{L}_0}|v|, \quad |v| \le r, \tag{12.6}$$

for any point $q = f^i(p), i \in \mathbb{Z}$, where

$$\mathcal{L}_0 = C^2 \frac{1 + \lambda}{1 - \lambda}.$$

Set

$$d_0 = \frac{r}{2\mathcal{L}_0}.$$

Then any d-method $\{\psi_k\}$ with $d \le d_0$ generates a pseudotrajectory $\{x_k\}$ for which inequalities (12.5) *hold with $x = p$ and $\mathcal{L} = 2\mathcal{L}_0$.*

In the last theorem of this section, we characterize shadowing properties of linear diffeomorphisms of the space $\mathbb{R}^n$ (note that it is possible to get necessary and sufficient conditions for shadowing only for linear diffeomorphims).

Theorem 12.3. *For a linear diffeomorphism L of the space $\mathbb{R}^n$ generated by a mapping $x \mapsto Ax$, the following five statements are equivalent.*

(1) *L has the shadowing property;*

(2) *L has the Lipschitz shadowing property;*

(3) *L has the inverse shadowing property;*

(4) *L has the Lipschitz inverse shadowing property;*

(5) *the matrix A is hyperbolic.*

12.2 Proof of Theorem 12.1

In the proof of Theorem 12.1, we refer to the existence of a so-called *Lyapunov norm* in a neighborhood of a hyperbolic set (with respect to this norm, the constant C in inequalities (HS2.3) and (HS2.4) in the definition of a hyperbolic set equals 1). Working with such a norm, we can simplify many proofs in the theory of hyperbolic sets.

Lemma 12.1. *Let Λ be a hyperbolic set of a diffeomorphism f with hyperbolicity constants C, λ_0. For any $\varepsilon > 0$ and $\lambda \in (\lambda_0, 1)$ we can find a neighborhood W of the set Λ having the following property. There exist positive constants N, δ, a C^∞ norm $|\cdot|_x$ for $x \in W$, and continuous (but not necessarily Df-invariant) extensions S' and U' of the families S and U of the given hyperbolic structure to the neighborhood W such that*

(1) $S'(x) \oplus U'(x) = \mathbb{R}^n, \quad x \in W;$

(2) *if $x, y \in W$, $|f(x) - y| \le \delta$, and P_x is the projection onto $S'(x)$ parallel to $U'(x)$, then the mapping $P_y Df(x)$ is a linear isomorphism between $S'(x)$ and $S'(y)$ (respectively, if $Q_x = \mathrm{Id} - P_x$, then the mapping $Q_y Df(x)$ is a linear isomorphism between $U'(x)$ and $U'(y)$) and the following inequalities hold:*

$$|P_y Df(x)v|_y \le \lambda |v|_x, \quad |Q_y Df(x)v|_y \le \varepsilon |v|_x, \quad v \in S'(x), \tag{12.7}$$

and

$$\lambda |Q_y Df(x)v|_y \ge |v|_x, \quad |P_y Df(x)v|_y \le \varepsilon |v|_x, \quad v \in U'(x); \tag{12.8}$$

(3)

$$\frac{1}{N}|v|_x \le |v| \le N|v|_x, \quad x \in W, v \in \mathbb{R}^n. \tag{12.9}$$

Proof. First we construct a continuous norm $|\cdot|_x$ having the desired properties. Fix a number $\mu \in (\lambda_0, \lambda)$ and find a natural number ν such that

$$C\left(\frac{\lambda_0}{\mu}\right)^{\nu+1} < 1. \tag{12.10}$$

Consider a point $p \in \Lambda$ and a vector $v \in \mathbb{R}^n$; represent $v = v^s + v^u \in S(p) \oplus U(p)$ and set

$$|v|_p^2 = |v^s|^2 + |v_u|^2,$$

where

$$|v^s|_p = \sum_{j=0}^{\nu} \mu^{-j} |Df^j(p)v^s| \quad \text{and} \quad |v^u|_p = \sum_{j=0}^{\nu} \mu^{-j} |Df^{-j}(p)v^u|.$$

The vector v^s satisfies the following estimate:

$$\begin{aligned}|Df(p)v^s|_{f(p)} &= \sum_{j=0}^{\nu} \mu^{-j}|Df^j(f(p))Df(p)v^s| \\ &= \mu\Big(\sum_{j=1}^{\nu} \mu^{-j}|Df^j(p)v^s| + \mu^{-\nu-1}|Df^{\nu+1}(p)v^s|\Big) \\ &\le \mu\Big(\sum_{j=1}^{\nu} \mu^{-j}|Df^j(p)v^s| + \mu^{-\nu-1}C\lambda_0^{\nu+1}|v^s|\Big) \le \mu|v^s|_p\end{aligned}$$

(in the last inequality, we refer to inequality (12.10)).

A similar reasoning shows that

$$\begin{aligned}|Df(p)v^u|_{f(p)} &= \sum_{j=0}^{\nu} \mu^{-j}|Df^{-j}(f(p))Df(p)v^u| \\ &= \mu^{-1}\Big(\sum_{j=0}^{\nu-1} \mu^{-j}|Df^{-j}(p)v^u| + \mu|Df(p)v^u|\Big).\end{aligned}$$

Since

$$w := Df^{-\nu}(p)v^u = Df^{-\nu-1}(f(p))Df(p)v^u,$$

it follows from property (HS2.4) that

$$|w| \le C\lambda_0^{\nu+1}|Df(p)v^u|,$$

which gives us the inequality

$$\begin{aligned}|Df(p)v^u|_{f(p)} &\ge \mu^{-1}\Big(\sum_{j=0}^{\nu-1} \mu^{-j}|Df^{-j}(p)v^u| + \mu^{-\nu}\frac{\mu^{\nu+1}}{C\lambda_0^{\nu+1}}|Df^{-\nu}(p)v^u|\Big) \\ &\ge \mu^{-1}|v^u|_p.\end{aligned}$$

By construction, the norm $|\cdot|_x$ is continuous on the set Λ (recall that the families of subspaces S and U are continuous by Lemma 7.4).

Let us extend the hyperbolic structure from the set Λ by continuous (but not necessarily Df-invariant) families of subspaces S', U' to a small closed neighborhood W_0 (i.e., the closure of a neighborhood of Λ) so that statement (1) of our lemma holds in W_0.

After that, we extend to W_0 (reducing W_0, if necessary) the constructed norm $|\cdot|_x$ and find a constant N such that inequality (12.9) is satisfied in W_0.

For points $x \in \Lambda$ and $y = f(x)$, the mapping $P_y Df(x)$ (the mapping $Q_y Df(x)$) is a linear isomorphism between $S(x)$ and $S(y)$ (between $U(x)$ and $U(y)$, respectively). The following relations hold:

$$\|P_y Df(x)|_{S(x)}\| \le \mu, \quad Q_y Df(x)|_{S(x)} = 0,$$
$$\mu\|Q_y Df(x)|_{U(x)}\| \ge 1, \quad \text{and} \quad P_y Df(x)|_{U(x)} = 0$$

(in these relations, the operator norms are generated by the norm $|\cdot|_x$).

We have selected a $\mu < \lambda$. Consider a number $\lambda' \in (\mu, \lambda)$.

The mappings P_x, Q_x, f, and Df are uniformly continuous, The first two relations above and the inequality $\mu > \lambda'$ imply that, given an arbitrary $\varepsilon > 0$, we can find a neighborhood $W = W(\varepsilon, \lambda')$ and a number $\delta > 0$ such that if $x, y \in W$ and $|f(x) - y| \le \delta$, then

$$\|P_y Df(x)|_{S'(x)}\| \le \lambda' \quad \text{and} \quad \|Q_y Df(x)|_{S'(x)}\| \le \varepsilon,$$

i.e., analogs of inequalities (12.7) hold (with λ replaced by λ). The same reasoning shows that we can guarantee that analogs of inequalities (12.8) are satisfied.

To complete the proof of our lemma, it remains to smooth out the norm $|\cdot|_x$ (reducing W, if necessary) preserving all the required estimates (and replacing λ' by λ). □

Let us pass to the proof of Theorem 12.1.

As we said before stating this theorem, we consider a hyperbolic set $\Lambda \subset \mathbb{R}^n$ of a diffeomorphism f with hyperbolicity constants C, λ_0.

Apply Lemma 12.1 to find a neighborhood W of the set Λ having all the properties stated in this lemma. We assume that the neighborhood W is bounded.

Since the projections P_x and Q_x are bounded, there exists a number $M > 0$ such that

$$\|P_x\|, \|Q_x\| \le M, \quad x \in W. \tag{12.11}$$

Reducing the neighborhood W and the number δ, we may assume that Lemma 12.1 holds with as small ε as we need. Since the numbers λ and M do not decrease as we reduce W, we may assume that

$$\lambda + \varepsilon(1 + 2M) < 1. \tag{12.12}$$

Let us show that f has the Lipschitz shadowing property in W.

Consider a sequence $\xi = \{x_k\} \subset W$ that satisfies the inequalities

$$|f(x_k) - x_{k+1}| \le d. \tag{12.13}$$

Estimates (12.9) imply that

$$|f(x_k) - x_{k+1}|_{x_{k+1}} \le N|f(x_k) - x_{k+1}| \le Nd. \tag{12.14}$$

Assume that we can find numbers D_0 and $\mathcal{L}_0$ such that if a sequence ξ satisfies inequalities (12.14) with $Nd \le D_0$, then there exists a point x such that

$$|f^k(x) - x_k|_{x_{k+1}} \le \mathcal{L}_0 Nd.$$

In this case,

$$|f^k(x) - x_k| \le \mathcal{L}_0 N^2 d;$$

thus, f has the Lipschitz shadowing property in the neighborhood W with constants $d_0 = D_0/N$ and $\mathcal{L} = \mathcal{L}_0 N^2$.

Hence, without loss of generality, we may assume that the standard Euclidean norm has all the properties of the norm $|\cdot|_x$ described in Lemma 12.1. To simplify notation, we write S, U instead of S', U'.

Consider a point $x \in W$ and a vector $v \in \mathbb{R}^n$ and represent

$$f(x + v) = f(x) + Df(x)v + h(x, v).$$

Clearly, there exists a number $a > 0$ (depending on ε and not on the point $x \in W$) such that

$$|h(x, v)| \le \varepsilon |v| \quad \text{as } |v| \le a. \tag{12.15}$$

Set

$$\mathcal{L} = \frac{2M}{1 - \lambda - \varepsilon(1 + 2M)} \tag{12.16}$$

and

$$d_0 = \min\left(\delta, \frac{a}{\mathcal{L}}\right). \tag{12.17}$$

We claim that f has the Lipschitz shadowing property in the neighborhood W with constants d_0 and $\mathcal{L}$.

Thus, we consider a sequence $\xi \subset W$ that satisfies inequalities (12.13) with $d \le d_0$. In addition, we consider a sequence $\eta = \{y_k \in \mathbb{R}^n\}$ and set $y_k = x_k + v_k$. The sequence η is a trajectory of the diffeomorphism f if and only if

$$v_k + x_k = f(x_{k-1} + v_{k-1}).$$

Let us write these equalities in the form

$$v_k = Df(x_{k-1})v_{k-1} + (f(x_{k-1}) - x_k) + h(x_{k-1}, v_{k-1}). \tag{12.18}$$

To find a shadowing trajectory η satisfying the required estimates, we want to apply the Schauder fixed point theorem.

Let $\mathcal{V}$ be the space of bounded sequences $V = \{v_k \in \mathbb{R}^n : i \in \mathbb{Z}\}$ with the norm

$$\|V\| = \sum_{k=-\infty}^{\infty} \frac{|v_k|}{2^{|k|}}.$$

With respect to this norm, $\mathcal{V}$ is a Banach space. We introduce on this space a metric

$$\operatorname{dist}(V, V') = \|V - V'\|.$$

Set $b = d\mathcal{L}/2$ and denote by $\mathcal{B}$ the subset of $\mathcal{V}$ consisting of sequences $V = \{v_k \in \mathbb{R}^n\}$ that satisfy the inequalities

$$|P_{x_k} v_k|, |Q_{x_k} v_k| \le b. \tag{12.19}$$

Since the projections P_{x_k} and Q_{x_k} are complementary,

$$P_{x_k} v_k + Q_{x_k} v_k = v_k,$$

Hence, if $V \in \mathcal{B}$, then

$$|v_k| \le 2b = d\mathcal{L} \le a. \tag{12.20}$$

Let us introduce the Tikhonov product topology on $\mathcal{B}$ (compare with Example 1.1).

Take an element $V \in \mathcal{B}$, a finite subset $I = \{k_1, \dots, k_m\}$ of the set of integers, and a family $c = \{c_1, \dots, c_m\}$ of positive numbers. The set

$$C(V, I, c) = \{V' \in \mathcal{B} : |v'_k - v_k| < c_k,\ k \in I\}$$

is called a cylinder. The base of neighborhoods of an element $V \in \mathcal{B}$ in the Tikhonov product topology consists of all cylinders $C(V, I, c)$ with all possible I and c.

It is easy to show that the metric topology induced by the above metric dist on $\mathcal{B}$ coincides with the Tikhonov product topology on $\mathcal{B}$.

Indeed, consider an open ball

$$B = \{V' \in \mathcal{B} : \operatorname{dist}(V', V) < \varepsilon\}.$$

There exists a finite subset $I \subset \mathbb{Z}$ such that

$$2 \sum_{k \in \mathbb{Z} \setminus I} \frac{a}{2^{|k|}} < \frac{\varepsilon}{2}$$

(when we apply this estimate below, we take into account that $|v'_k - v_k| \le 2a$ for $V', V \in \mathcal{B}$ due to inequality (12.20)).

If we take positive numbers $c_k \in I$ such that

$$\sum_{k \in I} \frac{c_k}{2^{|k|}} < \frac{\varepsilon}{2},$$

then $\text{dist}(V', V) < \varepsilon$ for any $V' \in C(V, I, c)$, which means that the cylinder $C(V, I, c)$ is a subset of the ball B.

On the other hand, if we take any cylinder $C(V, I, c)$, then we can find a positive ε (depending, of course, on I and c) such that the inequality $\text{dist}(V', V) < \varepsilon$ implies the inequalities

$$|v'_k - v_k| < c_k, \quad k \in I,$$

which means that the ball

$$\{V' \in \mathcal{B} : \text{dist}(V', V) < \varepsilon\}$$

is a subset of $C(V, I, c)$.

The Tikhonov theorem implies that $\mathcal{B}$, a countable product of compact sets given by inequalities (12.19), is compact in the Tikhonov product topology.

Hence, $\mathcal{B}$ is a compact subset of the Banach space of sequences with norm $\|\cdot\|$ (and, clearly, the set $\mathcal{B}$ is convex).

Let us recall the Schauder fixed point theorem (Schauder principle): A continuous operator that maps a convex and compact subset of a Banach space into itself has a fixed point in this set (see [43]).

To find a solution of equations (12.8), we define on $\mathcal{B}$ an operator T that assigns to a sequence $V \in \mathcal{B}$ a sequence $W = \{w_k \in \mathbb{R}^n\}$ as follows.

We define the projections of elements w_k of the sequence $W = T(V)$ to the subspaces $S(x_k)$ and $U(x_k)$ (clearly, this defines the sequence W).

To simplify notation, denote by $x = x_{k-1}$ and $y = x_k$ two consecutive points of the pseudotrajectory ξ.

Set

$$P_y w_k = P_y \left[Df(x) v_{k-1} + (f(x) - y) + h(x, v_{k-1}) \right]. \tag{12.21}$$

Projections to the subspaces $U(x_k)$ are defined in a more complicated way. Consider the linear mapping $G(w) = Q_y Df(x) w$ on the subspace $U(x)$. Statement (2) of Lemma 12.1 implies that G maps the subspace $U(x)$ to the subspace $U(y)$ and the following inequality holds:

$$|G(w) - G(w')| \geq \frac{1}{\lambda} |w - w'|, \quad w, w' \in U(x). \tag{12.22}$$

Consider the closed balls

$$B(x) = \{w \in U(x) : |w| \leq b\} \quad \text{and} \quad B(y) = \left\{ z \in U(y) : |z| \leq \frac{b}{\lambda} \right\}.$$

It follows form inequality (12.22) that $B(y) \subset G(B(x))$ and that a mapping Γ, inverse to G, is defined on $B(y)$. In addition,

$$|\Gamma(z) - \Gamma(z')| \leq \lambda |z - z'|, \quad z, z' \in B(y). \tag{12.23}$$

Now we define the projections of elements w_k to the subspaces $U(x_k)$ (to be exact, we define the projections of elements w_{k-1} to the subspaces $U(x_{k-1})$) by the equalities

$$Q_x w_{k-1} = \Gamma\left(Q_y\left[v_k - Df(x)P_x v_{k-1} - (f(x) - y) - h(x, v_{k-1})\right]\right). \quad (12.24)$$

Since $|y - f(x)| \le d \le \delta$, we can apply the statement of Lemma 12.1 to the pair (x, y). In addition, the choice of a and inequality (12.15) imply that if $V \in \mathcal{B}$, then

$$|h(x_k, v_k)| \le \varepsilon |v_k|, \quad k \in \mathbb{Z}.$$

Let us estimate the right-hand side of formula (12.21). By statement (2) of Lemma 12.1 (see inequalities (12.7) and (12.8)),

$$\begin{aligned} |P_y[Df(x)v_{k-1}]| &= |P_y[Df(x)(P_x v_{k-1} + Q_x v_{k-1})]| \\ &\le \lambda |P_x v_{k-1}| + \varepsilon |Q_x v_{k-1}|. \end{aligned}$$

Since $|f(x) - y| \le d$, estimates (12.11) imply that

$$|P_y[f(x) - y]| \le Md.$$

Finally, the following estimates hold:

$$|P_y h(x, v_k)| \le M\varepsilon |v_k| \le 2M\varepsilon b.$$

These estimates imply that if $V \in \mathcal{B}$, then

$$|P_y w_k| \le \lambda(1 + \varepsilon + 2M\varepsilon)b + Md = \lambda(1 + \varepsilon + 2M\varepsilon)b + \frac{2Mb}{\mathcal{L}},$$

and now it follows from the choice of $\mathcal{L}$ and d_0 that

$$|P_y w_k| \le b.$$

Now we estimate the argument of the mapping Γ in the right-hand side of formula (12.24). The same reasoning as above establishes the estimate

$$\begin{aligned} &|Q_y[v_k - Df(x)P_x v_{k-1} - (f(x) - y) - h(x, v_{k-1})]| \\ &\quad \le |Q_y v_k| + \varepsilon |P_x v_{k-1}| + 2M\varepsilon b + Md \le (1 + \varepsilon + 2M\varepsilon)b + Md. \end{aligned}$$

Clearly,

$$\lambda((1 + \varepsilon + 2M\varepsilon)b + Md) < \lambda(1 + \varepsilon + 2M\varepsilon)b + Md.$$

Property (12.23) of the mapping Γ, the choice of the number $\mathcal{L}$, and the above estimate imply that

$$|Q_x w_{k-1}| \le b.$$

Thus, with our choice of the numbers $\mathcal{L}$ and d, the operator T is defined on the set $\mathcal{B}$ and maps this set to itself.

To prove that the operator T has a fixed point in $\mathcal{B}$, it remains to show that T is continuous.

By the definition of the operator T, the element w_k of the sequence $W = T(V)$ is determined by the elements v_{k-1}, v_k, v_{k+1} of the sequence V. Clearly, T is continuous with respect to the Tikhonov product topology. It was shown above that the product topology on $\mathcal{B}$ is equivalent to the topology induced by the metric dist of our Banach space $\mathcal{V}$. Thus, T is continuous, and we can apply the Schauder principle.

Now we show that a fixed point V of the operator T gives us a trajectory that shadows the pseudotrajectory ξ.

Indeed, if V is a fixed point of the operator T, then the following equalities hold:

$$P_y v_k = P_y \left[Df(x)v_{k-1} + (f(x) - y) + h(x, v_{k-1})\right] \tag{12.25}$$

and

$$Q_x v_{k-1} = \Gamma \left(Q_y \left[v_k - Df(x)P_x v_{k-1} - (f(x) - y) - h(x, v_{k-1})\right]\right).$$

Let us apply the mapping G to the second equality:

$$\begin{aligned} G(Q_x v_{k-1}) &= Q_y Df(x) Q_x v_{k-1} \\ &= Q_y[v_k - Df(x)P_x v_{k-1} - (f(x) - y) - h(x, v_{k-1})]. \end{aligned}$$

Transforming this equality, we see that

$$Q_y v_k = Q_y[Df(x)v_{k-1} + (f(x) - y) + h(x, v_{k-1})]. \tag{12.26}$$

Adding equalities (12.25) and (12.26), we conclude that the sequence V satisfies equalities (12.18). The inequalities $|v_k| \le 2b = \mathcal{L}d$ imply that the point $x = x_0 + v_0$ generates the desired shadowing trajectory. □

Remark. It is easy to show that if the number d is small enough, then a d-pseudotrajectory belonging to a neighborhood of a hyperbolic set is shadowed by a unique exact trajectory.

This property of uniqueness of a shadowing trajectory is a corollary of the following property of hyperbolic sets (usually called *expansivity*): There exists a neighborhood W of a hyperbolic set Λ and a number $c > 0$ such that if $O(x, f), O(y, f) \subset W$ and

$$\operatorname{dist}(f^k(x), f^k(y)) \le c, \quad k \in \mathbb{Z}, \tag{12.27}$$

then $x = y$.

Exercise 12.1. Show that the above-formulated property follows from Lemma 12.1.
Hint. Let $x_k = f^k(x)$ and $y_k = f^k(y)$. Represent

$$x_k - y_k = v_k + w_k,$$

where $v_k \in S(x_k), w_k \in U(x_k)$.

Since

$$x_{k+1} - y_{k+1} = f(x_k) - f(y_k) = Df(x_k)(v_k + w_k) + o(c),$$

it follows from Lemma 12.1 that

$$|v_{k+1}| \leq \lambda |v_k| + \varepsilon |w_k| + o(c)$$

and

$$|w_{k+1}| \geq \lambda^{-1} |w_k| - \varepsilon |w_k| + o(c).$$

Show that if c is small enough and inequalities (12.27) hold, then $v_0 = 0$ and $w_0 = 0$.

12.3 Proof of Theorem 12.2

Consider a trajectory $O(p, f)$ of a diffeomorphism f at which f has (C, λ)-structure.

Fix a d-method $\{\psi_k\}$ with $d \leq d_0$; let us find a pseudotrajectory $\xi = \{x_k\}$ generated by the method $\{\psi_k\}$ and close to $O(p, f)$ in the form

$$x_k = p_k + v_k, \quad k \in \mathbb{Z}, \tag{12.28}$$

where $p_k = f^k(p)$.

The equalities $x_{k+1} = \psi_k(x_k)$ are equivalent to the equalities

$$p_{k+1} + v_{k+1} = \psi_k(p_k + v_k). \tag{12.29}$$

Since $p_{k+1} = f(p_k)$ and

$$f(p_k + v_k) = f(p_k) + Df(p_k)v_k + g(p_k, v_k),$$

we can write equalities (12.29) in the form

$$v_{k+1} = Df(p_k)v_k + G_{k+1}(v_k), \tag{12.30}$$

where

$$G_{k+1}(v) = g(p_k, v) + (\psi_k(p_k + v) - f(p_k + v)).$$

Since $\{\psi_k\}$ is a d-method,

$$|\psi_k(p_k + v) - f(p_k + v)| < d. \tag{12.31}$$

Condition (12.6) implies that

$$|g(p_k, v)| \leq \frac{1}{2\mathcal{L}_0}|v| \tag{12.32}$$

for $|v| \leq r$.

Similarly to the proof of Theorem 12.1, we apply the Schauder principle to solve equations (12.30) (but the operator which we use now is essentially different from that in the proof of Theorem 12.1).

We denote by $\mathcal{E}$ the set of bounded sequences

$$V = \{v_k \in R^n : k \in \mathbb{Z}\}$$

that satisfy the inequalities

$$\|V\|_\infty = \sup_{k\in\mathbb{Z}} |v_k| \leq 2\mathcal{L}_0 d.$$

We assign to a sequence $V \in \mathcal{E}$ the sequence $Z(V) = \{z_k(V)\}$, where $z_{k+1}(V) = G_{k+1}(v_k)$.

The inequality $2\mathcal{L}_0 d \leq r$ and inequalities (12.31) and (12.32) imply that

$$\|Z(V)\|_\infty \leq \frac{1}{2\mathcal{L}_0}\|V\|_\infty + d, \quad V \in \mathcal{E}. \tag{12.33}$$

We define on the set $\mathcal{E}$ an operator R: $R(V) = W = \{w_k\}$, where

$$w_k = \sum_{i=-\infty}^{k} Df^{k-i}(p_i)P(p_i)z_i(V) - \sum_{i=k+1}^{\infty} Df^{k-i}(p_i)Q(p_i)z_i(V) \tag{12.34}$$

(compare this operator and the Perron operator which we used in Section 4.4 in the proof of the existence of the stable manifold for a hyperbolic fixed point of a diffeomorphism).

Since $P(p_i)z_i(V) \subset S(p_i)$ and $Q(p_i)z_i(V) \subset U(p_i)$, it follows from the definition of (C, λ)-structure that

$$|Df^{k-i}P(p_i)z_i(V)| \leq C^2\lambda^{k-i}|z_i(V)|, \quad k \geq i,$$

and

$$|Df^{k-i}Q(p_i)z_i(V)| \leq C^2\lambda^{i-k}|z_i(V)|, \quad i \geq k.$$

These inequalities imply that the series defining the operator R converge and the following estimate holds:

$$\|R(V)\|_\infty \le \mathcal{L}_0 \|Z(V)\|_\infty.$$

Inequality (12.33) shows that R maps the set $\mathcal{E}$ into itself.

Let us show that R has a fixed point in $\mathcal{E}$.

Similarly to the proof of Theorem 12.1, we introduce on $\mathcal{E}$ the metric

$$\operatorname{dist}(V, V') = \sum_{i=-\infty}^{\infty} \frac{|v_i - v_i'|}{2^{|i|}}.$$

Literally the same reasoning as in the proof of Theorem 12.1 shows that the metric topology on $\mathcal{E}$ coincides with the Tikhonov product topology. Since $\mathcal{E}$ is a countable product of the compact balls $\{|v| \le 2\mathcal{L}_0 d\}$, the set $\mathcal{E}$ is compact. Obviously, this set is convex.

Thus, it remains to prove that the operator R is continuous on $\mathcal{E}$. In this case, this part of the proof is more complicated.

Let m be a natural number. Consider the set $\mathcal{E}_m$ of finite sequences

$$V = \{v_k \in \mathbb{R}^n : |k| \le m\}$$

that satisfy the inequalities

$$\|V\|_m := \max_{|k| \le m} |v_k| \le 2\mathcal{L}_0 d.$$

Denote by π_m and $\pi_m^l, m \le l$, the natural projections of $\mathcal{E}$ to $\mathcal{E}_m$ and of $\mathcal{E}_l$ to $\mathcal{E}_m$, respectively (the projections π_m and π_m^l just "cut off" the "tails" of the corresponding sequences).

Define operators $R_m : \mathcal{E} \to \mathcal{E}_m$ by analogy with the operator R: $R_m(V) = W = \{w_k : |k| \le m\}$, where

$$w_k = \sum_{i=-m}^{k} Df^{k-i} P(p_i) z_i(V) - \sum_{i=k+1}^{m} Df^{k-i} Q(p_i) z_i(V).$$

Since the values $z_i(V), |i| \le m$, are determined by the values v_k with $|k| \le m+1$, the operator R_m is obviously continuous (with respect to the metrics dist on $\mathcal{E}$ and $\operatorname{dist}_m(V', V) = \|V' - V\|_m$ on $\mathcal{E}_m$).

The operator $\pi_m R$ maps a sequence $V \in \mathcal{E}$ to the sequence $\{w_k : |k| \le m\}$, where the values w_k are given by formulas (12.34).

Take two natural numbers $l > m$ and consider the operator $\pi_m^l R_l$; this operator maps a sequence $V \in \mathcal{E}$ to the sequence $\{w_k' : |k| \le m\}$, where

$$w_k' = \sum_{i=-l}^{k} Df^{k-i} P(p_i) z_i(V) - \sum_{i=k+1}^{l} Df^{k-i} Q(p_i) z_i(V).$$

Let us estimate

$$\begin{aligned}\|\pi_m R(V) - \pi_m^l R_l(V)\|_m &= \max_{|k|\le m} |w_k - w_k'| \\ &\le 2\mathcal{L}_0 C^2 d \max_{|k|\le m} \Big(\sum_{i=-\infty}^{-l-1} \lambda^{k-i} + \sum_{i=-l+1}^{\infty} \lambda^{i-k} \Big) \\ &\le \frac{4\mathcal{L}_0 C^2 d \lambda^{1-m}}{1-\lambda} \lambda^l .\end{aligned}$$

This estimate implies that the operator $\pi_m R$ is the uniform (with respect to $V \in \mathcal{E}$) limit as $l \to \infty$ of the continuous operators $\pi_m^l R_l$. Hence, the operator $\pi_m R$ is continuous. Using the definition of the product topology, we easily see that the operator R is continuous.

Thus, we conclude that the operator R has a fixed point $V \in \mathcal{E}$.

The fixed point V satisfies the equalities

$$\begin{aligned} v_{k+1} &= \sum_{i=-\infty}^{k+1} Df^{k+1-i}(p_i)P(p_i)z_i(V) - \sum_{i=k+2}^{\infty} Df^{k+1-i}(p_i)Q(p_i)z_i(V) \\ &= P(p_{k+1})z_{k+1}(V) + Df(p_k)P(p_k)z_k(V) \\ &\quad + Df^2(p_{k-1})P(p_{k-1})z_{k-1}(V) + \cdots \\ &\quad - Df^{-1}(p_{k+2})Q(p_{k+2})z_{k+2}(V) - Df^{-2}(p_{k+3})Q(p_{k+3})z_{k+3}(V) - \cdots . \end{aligned}$$

Since

$$\begin{aligned} Df(p_k)v_k &= Df(p_k)P(p_k)z_k(V) + Df^2(p_{k-1})P(p_{k-1})z_{k-1}(V) + \cdots \\ &\quad - Q(p_{k+1})z_{k+1}(V) - Df^{-1}(p_{k+2})Q(p_{k+2})z_{k+2}(V) \\ &\quad - Df^{-2}(p_{k+3})Q(p_{k+3})z_{k+3}(V) - \cdots \end{aligned}$$

and

$$P(p_{k+1})z_{k+1}(V) + Q(p_{k+1})z_{k+1}(V) = z_{k+1}(V),$$

the sequence V is a solution of equations (12.30). □

12.4 Proof of Theorem 12.3

First we note that the following simple statement is valid (we leave the proof to the reader).

Exercise 12.2. If two matrices A and B are similar, then the mappings $x \mapsto Ax$ and $x \mapsto Bx$ have the shadowing property (and the inverse shadowing property) simultaneously. The same is true for the Lipschitz shadowing and inverse shadowing properties.

The implications (2) $\Rightarrow$ (1) and (4) $\Rightarrow$ (3) are obvious.

Let us prove the implication (5) $\Rightarrow$ (2).

For this purpose, we can modify the reasoning applied in the proof of Theorem 12.1; we use here a simpler method that gives us a closed-form expression of the shadowing trajectory.

By Lemma 4.1, the hyperbolic matrix A is similar to a block-diagonal matrix $\operatorname{diag}(B, C)$ such that $\|B\| \le \lambda$ and $\|C^{-1}\| \le \lambda$, where $\lambda \in (0, 1)$. Applying Exercise 12.2, we assume that $A = \operatorname{diag}(B, C)$.

Represent a vector $x \in \mathbb{R}^n$ in the form $x = (y, z)$ according to the representation $A = \operatorname{diag}(B, C)$. Clearly,

$$\left| A^k \begin{pmatrix} y \\ 0 \end{pmatrix} \right| = |B^k y| \le \lambda^k |y|, \quad k \ge 0, \tag{12.35}$$

and

$$\left| A^k \begin{pmatrix} 0 \\ z \end{pmatrix} \right| = |C^k z| \le \lambda^{-k} |z|, \quad k \le 0. \tag{12.36}$$

Consider a d-pseudotrajectory $\xi = \{x_k\}$ of the mapping L.

We represent the trajectory $O(p, L)$ of the mapping L that shadows the d-pseudotrajectory $\xi = \{x_k\}$ in the form

$$p_k = A^k p = x_k + v_k, \quad k \in \mathbb{Z}. \tag{12.37}$$

Clearly, equalities (12.37) are equivalent to the equalities

$$x_{k+1} + v_{k+1} = A(x_k + v_k),$$

or

$$v_{k+1} = A v_k + z_{k+1}, \quad k \in \mathbb{Z}, \tag{12.38}$$

where $z_{k+1} = A x_k - x_{k+1}$. Since ξ is a d-pseudotrajectory,

$$|z_k| < d, \quad k \in \mathbb{Z}. \tag{12.39}$$

Let P and Q be the complementary projections in $\mathbb{R}^n$ to the subspaces of variables y and z, respectively. Clearly, $\|P\| = 1$ and $\|Q\| = 1$.

Let us show, by analogy with the proof of Theorem 12.2, that the sequence

$$v_k = \sum_{i=-\infty}^{k} A^{k-i} P z_i - \sum_{i=k+1}^{\infty} A^{k-i} Q z_i$$

satisfies equalities (12.38), i.e., the trajectory $O(p_0, L)$ of the point $p_0 = x_0 + v_0$ is the desired shadowing trajectory.

Inequalities (12.35), (12.36), and (12.39) imply that the series defining the sequence v_k converge, and

$$|v_k| \leq d \frac{1+\lambda}{1-\lambda}. \tag{12.40}$$

Now we apply the same reasoning as in the proof of Theorem 12.2 to show that the sequence V satisfies relations (12.38).

The equalities

$$\begin{aligned} v_{k+1} &= P z_{k+1} + AP z_k + A^2 P z_{k-1} + \cdots - A^{-1} Q z_{k+2} - A^2 Q z_{k+3} - \cdots \\ &= P z_{k+1} + Q z_{k+1} + A v_k = A v_k + z_{k+1} \end{aligned}$$

imply that the sequence V has the desired property.

Thus, the mapping L has the Lipschitz shadowing property with constant

$$\mathcal{L} = \frac{1+\lambda}{1-\lambda}$$

and an arbitrary $d_0 > 0$.

The implication (5) $\Rightarrow$ (2) is proven.

The implication (5) $\Rightarrow$ (4) is a direct corollary of Theorem 12.2 since, at any trajectory $\{p_k\}$, the mapping L has $(1, \lambda)$-structure with the subspaces $S(p_k) = P\mathbb{R}^n$ and $U(p_k) = Q\mathbb{R}^n$, while the left-hand side of inequality (12.6) vanishes in the case of the linear diffeomorphism $f(x) = Ax$.

Now we prove the implication (1) $\Rightarrow$ (5). Assume the converse; let a linear mapping L with a nonhyperbolic matrix A have the shadowing property.

First we assume that A has a real eigenvalue λ such that $|\lambda| = 1$, i.e., $\lambda = \pm 1$. In this case, we assume that A coincides with its Jordan form, $A = \operatorname{diag}(B, C)$, and B is a Jordan block of size $m \times m$,

$$B = \begin{pmatrix} \lambda & & & 0 \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix}$$

(of course, we do not exclude the case $m = 1$).

Take $\varepsilon = 1$; let d be the corresponding constant from the definition of the shadowing property.

Consider a sequence of vectors $\xi = \{x_k\}$ defined as follows. The first component x_k^1 has the form $x_k^1 = k\lambda^k/2$, the components $x_k^2, \dots, x_k^m$ satisfy the relations

$$x_{k+1}^i = \lambda x_k^i + x_k^{i-1}, \quad i = 2, \dots, m, k \in \mathbb{Z},$$

and the components $x_k^i, i > m$, equal zero.

Then

$$x_{k+1} - Ax_k = (\lambda^{k+1} d/2, 0, \dots, 0);$$

thus, the sequence ξ is a d-pseudotrajectory of the mapping L.

For any $y = (y^1, \dots, y^n) \in \mathbb{R}^n$, the first component of the vector $A^k y$ equals $\lambda^k y^1$; hence,

$$|A^k y - x_k| \geq |y^1 - kd/2|.$$

Since the right-hand side of this inequality is unbounded for $k \in \mathbb{Z}$ for any y, L does not have the shadowing property.

To consider the case of a complex eigenvalue λ, let us note that if a mapping $L : \mathbb{R}^n \to \mathbb{R}^n$ with a real-valued matrix A has the shadowing property, then the mapping $L_1 : \mathbb{C}^n \to \mathbb{C}^n$ with the same matrix A has the shadowing property as well.

Indeed, take $\varepsilon > 0$ and find the corresponding d (for the mapping L). Let $\{z_k = v_k + iw_k\}$ be a complex-valued d-pseudotrajectory of the mapping L_1, i.e.,

$$|Az_k - z_{k+1}| < d.$$

Since the matrix A is real-valued, this means that

$$|Av_k - v_{k+1}| < d \quad \text{and} \quad |Aw_k - w_{k+1}| < d.$$

By our choice of d, there exist vectors $\psi, \chi \in \mathbb{R}^n$ such that

$$|A^k \psi - v_k| < \varepsilon \quad \text{and} \quad |A^k \chi - w_k| < \varepsilon.$$

Then

$$|A^k(\psi + i\chi) - z_k| < 2\varepsilon,$$

which implies that the mapping L_1 has the shadowing property.

Now we prove the implication (1) $\Rightarrow$ (5) in the case of a complex eigenvalue λ of the matrix A with $|\lambda| = 1$ using the same reasoning as in the case of eigenvalues $\lambda = \pm 1$: we consider a complex Jordan form of the matrix A similar to that in the real case and construct a (complex) pseudotrajectory of the same form.

Let us prove the implication (3) $\Rightarrow$ (5).

We again assume the converse. Let the mapping L with a nonhyperbolic matrix A have the inverse shadowing property.

First we consider the case where the matrix A has a real eigenvalue λ such that $|\lambda| = 1$.

Similarly to our proof of the implication (1) $\Rightarrow$ (5), we assume that A is block-diagonal: $A = \mathrm{diag}(B, C)$, and B is the same Jordan block of size $m \times m$ as above.

Take $\varepsilon = 1$; let d be the corresponding constant from the definition of the inverse shadowing property.

Consider the mappings

$$\psi_k(x) = A^k x + (d\lambda^k/2, 0, \dots, 0);$$

clearly, the sequence $\{\psi_k\}$ is a d-method.

Let $\xi = \{x_k\}$ be a pseudotrajectory of the method $\{\psi_k\}$; represent $x_k = (x_k^1, \dots, x_k^n)$.

It is easy to see that

$$x_k^1 = x_0^1 + k\lambda^{k-1} d/2.$$

Consider the vector $p = (1, \dots, 0)$. The first component p_k^1 of the vector $A^k p$ equals λ^k.

If $\xi = \{x_k\}$ is a pseudotrajectory generated by the method $\{\psi_k\}$, then the expression

$$|x_k^1 - p_k^1| = |\lambda^k (x_0^1 - 1) + k\lambda^{k-1} d/2|$$

is unbounded for $k \in \mathbb{Z}$. Hence, L does not have the inverse shadowing property.

If the matrix A has a complex eigenvalue λ with $|\lambda| = 1$, then it has the so-called real Jordan form for which the first two rows are as follows (see, for example, [14]):

$$\begin{pmatrix} \cos\phi & -\sin\phi & 0 & \cdots & 0 \\ \sin\phi & \cos\phi & 0 & \cdots & 0 \end{pmatrix},$$

where $\phi \in (-\pi, \pi)$.

Take $\varepsilon = 1$; let d be the corresponding constant from the definition of the inverse shadowing property.

Consider the mappings

$$\psi_k(x) = A^k x + \left(\frac{d\cos(k+1)\phi}{2}, \frac{d\sin(k+1)\phi}{2}, 0, \dots, 0 \right);$$

clearly, the sequence $\{\psi_k\}$ is a d-method.

It is easy to see that if $\xi = \{x_k\}$ is a pseudotrajectory generated by the method $\{\psi_k\}$ and $x_k = (x_k^1, \dots, x_k^n)$, then

$$x_k^1 = x_0^1 \cos k\phi - x_0^2 \sin k\phi + k\frac{d\cos k\phi}{2}$$

and

$$x_k^2 = x_0^1 \sin k\phi + x_0^2 \cos k\phi + k\frac{d \sin k\phi}{2}$$

for $k \geq 0$.

Consider the vector $p = (1, \dots, 0)$. The first two components p_k^1 and p_k^1 of the vector $A^k p$ equal $\cos k\phi$ and $\sin k\phi$, respectively.

Since

$$k^2 \cos^2 k\phi + k^2 \sin^2 k\phi = k^2,$$

for any pseudotrajectory $\xi = \{x_k\}$ generated by the method $\{\psi_k\}$, the sum of the values

$$|x_k^1 - p_k^1| = \left| x_0^1 \cos k\phi - x_0^2 \sin k\phi + k\frac{d \cos k\phi}{2} - \cos k\phi \right|$$

and

$$|x_k^2 - p_k^2| = \left| x_0^1 \sin k\phi - x_0^2 \cos k\phi + k\frac{d \sin k\phi}{2} - \sin k\phi \right|$$

is unbounded for $k \in \mathbb{Z}$. Hence, L does not have the inverse shadowing property.

This completes the proof of Theorem 12.3. □

Appendix A
Scheme of the proof of the Mañé theorem

In 1987, R. Mañé had completed the proof of the main theorem in the theory of structural stability; he proved that a structurally stable diffeomorphism satisfies Axiom A [44] (it is relatively easy ot prove that a structurally stable diffeomorphism satisfies the strong transversality condition, see [45]).

The proof published by Mañé in the paper [44] is quite complicated; here we describe a scheme of this proof omitting many technical details.

Let f be a diffeomorphism of a smooth closed n-dimensional manifold M. Denote by $P(f)$ the set of periodic points of f. If every periodic point of f is hyperbolic, we denote by $P_i(f), 0 \leq i \leq n$, the set of periodic points p for which $\dim W^u(p) = i$; let $R_i(f) := \mathrm{Cl}\, P_i(f)$.

Recall (see Section 7.5) that $\mathcal{F}(M)$ denotes the set of diffeomorphisms f having the following property: f has a neighborhood U in $\mathrm{Diff}^1(M)$ such that any periodic point of a diffeomorphism $g \in U$ is hyperbolic.

It is useful to the reader to prove the following simple statement.

Exercise A.1. If a diffeomorphism f is structurally stable, then $f \in \mathcal{F}(M)$.

Hint. Assume that a diffeomorphism f is structurally stable and does not belong to the set $\mathcal{F}(M)$. Since the set of structurally stable diffeomorphisms is open in $\mathrm{Diff}^1(M)$ (see Theorem 3.1), f has a neighborhood V in $\mathrm{Diff}^1(M)$ such that any diffeomorphism $g \in V$ is structurally stable. Since $f \notin \mathcal{F}(M)$, the neighborhood V contains a diffeomorphism g having a nonhyperbolic periodic point p. Let m be the period of p. Find a diffeomorphism $h \in V$ that is a C^1-small perturbation of g having the following properties: p is a periodic point of h of period m, the diffeomorphism h^m is linear in a neighborhood of the point p, and the derivative $Dh^m(p)$ has an eigenvalue λ with $|\lambda| = 1$. Show that such a diffeomorphism h is not structurally stable (which contradicts the inclusion $h \in V$).

A large part of the proof in [44] uses only the inclusion $f \in \mathcal{F}(M)$ (it was noticed by Mañé that the assumption that f is structurally stable is used only in the last step of the proof).

First Mañé proves the following statement (in fact, this statement is a corollary of the Closing Lemma, see Theorem 10.1).

Theorem A.1. *If* $f \in \mathcal{F}(M)$*, then* $\Omega(f) = \mathrm{Cl}\, P(f)$.

Thus, to prove that a structurally stable diffeomorphism f satisfies Axiom A, it is enough to show that any of the sets $R_i(f)$ is hyperbolic.

Let us note that for $i = 0$ and $i = n$ (recall that n is the dimension of the manifold M), the corresponding statement had been proven by Pliss [46].

Theorem A.2 ([46]). *If $f \in \mathcal{F}(M)$, then the sets $P_0(f)$ and $P_n(f)$ are finite.*

Clearly, a finite union of hyperbolic periodic points coincides with its closure and is a hyperbolic set.

Let $S(p)$ and $U(p)$ be the stable and unstable subspaces of the hyperbolic structure for a hyperbolic set that is the trajectory of a hyperbolic periodic point p. The main goal of the following steps of the proof is to extend the decomposition $S(p) \oplus U(p) = T_p M$ from the sets $P_i(f)$ to the sets $R_i(f)$.

Let us define a property of invariant sets which is very important for the global theory of dynamical systems.

Let Λ be a compact invariant set of a diffeomorphism f. We say that f has a *dominated splitting* on Λ if there exist continuous families of linear subspaces $E(x)$ and $F(x)$ of the tangent spaces $T_x M$ for $x \in \Lambda$ such that

(1) $E(x) \oplus F(x) = T_x M$, $x \in \Lambda$;

(2) the subspaces $E(x)$ and $F(x)$ are Df-invariant (i.e., analogs of equalities (HS2.2) from the definition of a hyperbolic structure hold);

(3) there exist numbers $C > 0$ and $\lambda \in (0, 1)$ such that

$$\|Df^n|_{E(x)}\| \cdot \|Df^{-n}|_{F(f^n(x))}\| \leq C\lambda^n$$

for all $x \in \Lambda$, $n \geq 0$.

Remark. It is possible that a diffeomorphism has a dominated splitting on an invariant set that is not hyperbolic. As an example, consider the fixed point $p = (0, 0)$ of the linear planar diffeomorphism $f(x_1, x_2) = (2x_1, x_2)$.

Show that f has on p a dominated splitting with the subspaces $E(p) = \{x_1 = 0\}$ and $F(p) = \{x_2 = 0\}$ and constants $C = 1$ and $\lambda = 1/2$.

The following statement had been independently proven by several authors (see [47, 48, 49]). This statement shows that if $f \in \mathcal{F}(M)$, then all diffeomorphisms g, C^1-close to f, have dominated splittings on the sets $R_i(g)$ with the same characteristics. As usual, $[a]$ denotes the integer part of a number a.

Theorem A.3. *If $f \in \mathcal{F}(M)$, then there exist numbers $C > 0$, $\lambda \in (0, 1)$, and $m > 0$, and a neighborhood U of the diffeomorphism f in* $\text{Diff}^1(M)$ *having the following property. If $g \in U$ and $0 < i < n$, then there exists a dominated splitting $E_i \oplus F_i$ for g on the set $R_i(g)$ such that*

(1) $\|Dg^m|_{E_i(x)}\| \cdot \|Dg^{-m}|_{F(g^m(x))}\| \le \lambda$ *for all* $x \in R_i(g)$*;*

(2) $E_i(x) = S(x)$ *and* $F_i(x) = U(x)$ *for* $x \in P_i(g)$ *(here* $S(x)$ *and* $U(x)$ *are the subspaces of the hyperbolic structure on the trajectory of a periodic point* x *of the diffeomorphism* g*);*

(3) *if* $x \in P_i(g)$ *and* $l > m$ *is the period of a periodic point* x*, then*

$$\prod_{j=0}^{[l/m]-1} \|Dg^m|_{S(g^{mj}(x))}\| \le C\lambda^{[l/m]}$$

and

$$\prod_{j=0}^{[l/m]-1} \|Dg^{-m}|_{U(g^{mj}(x))}\| \le C\lambda^{[l/m]};$$

(4) *if* $x \in P_i(g)$*, then*

$$\lim_{k\to+\infty} \frac{1}{k} \sum_{j=0}^{k-1} \log \|Dg^m|_{S(g^{mj}(x))}\| \le \log \lambda$$

and

$$\lim_{k\to+\infty} \frac{1}{k} \sum_{j=0}^{k-1} \log \|Dg^{-m}|_{U(g^{mj}(x))}\| \le \log \lambda.$$

Theorem A.3 establishes the existence of a Df-invariant splitting $E_i \oplus F_i$ for f on the set $R_i(f)$. The remaining (and, in fact, main) part of the proof shows that this splitting is a hyperbolic structure.

To prove this fact, it is enough to show that the derivative Df of the diffeomorphism f contracts on the subspaces E_i, i.e., there exist constants $C > 0$ and $\lambda \in (0, 1)$ such that

$$\|Df^k|_{E_i(x)}\| \le C\lambda^k$$

for $x \in R_i(f)$, $k \ge 0$.

Theorem A.4. *If* $f \in \mathcal{F}(M)$, $0 < i < n$, *and the derivative* Df *contracts on the subspaces* E_i*, then there exist numbers* $C_1 > 0$ *and* $\lambda_1 \in (0, 1)$ *such that*

$$\|Df^{-k}|_{F_i(x)}\| \le C_1\lambda_1^k$$

for $x \in R_i(f)$, $k \ge 0$ *(i.e., the splitting* $E_i \oplus F_i$ *is a hyperbolic structure).*

Precisely in the proof of the fact that the derivative Df of the diffeomorphism f contracts on the subspaces E_i, Mañé starts to refer to methods of the theory of invariant measures (at the first glance, such methods have no relation to the theory of structural stability).

Let Λ be an invariant set of a diffeomorphism f; consider the set of f-invariant probability measures on the algebra of Borel subsets of Λ (see Section 10).

An invariant probability measure μ is called *ergodic* if, for any invariant subset A of the set Λ, either $\mu(A) = 0$ or $\mu(A) = 1$.

Denote by $\mathcal{M}(f|_\Lambda)$ the space of invariant probability measures for the restriction $f|_\Lambda$ with the weak topology, i.e., with the topology in which

$$\mu_k \to \mu \quad \Leftrightarrow \quad \int \phi \, d\mu_k \to \int \phi \, d\mu$$

for any continuous function $\phi : \Lambda \to \mathbb{R}$.

Theorem A.5. *Let Λ be a compact invariant set of the diffeomorphism f. Assume that*

$$\{E(x) \subset T_x M : x \in \Lambda\}$$

is a continuous family of Df-invariant subspaces. If there exists an index $m > 0$ such that

$$\int_\Lambda \log \|Df^m|_E\| \, d\mu < 0$$

for any ergodic measure $\mu \in \mathcal{M}(f|_\Lambda)$, then the derivative Df of the diffeomorphims f contracts on the subspaces $\{E(x)\}$.

Now the proof goes by induction. Theorem A.2 implies that if a diffeomorphism f is structurally stable (hence, $f \in \mathcal{F}(M)$), then the set $P_0(f)$ is hyperbolic. It is assumed that the sets $R_k(f)$ are hyperbolic for $0 \le k \le j$; our goal is to show that the set $R_{j+1}(f)$ is hyperbolic as well.

For this purpose, the following statement is proved (in its proof, Mañé applies the ergodic closing lemma (see Theorem 10.5).

Theorem A.6. *Let $f \in \mathcal{F}(M)$ and let $m > 0$ be the number given by Theorem* A.3. *There exists a number $\lambda_0 \in (0,1)$ such that if the sets $R_k(f)$ are hyperbolic for $0 \le k \le i$ and some measure $\mu \in \mathcal{M}(f^m|_{R_i(f)})$ satisfies the inequality*

$$\int_{R_i(f)} \log \|Df^m|_{E_i}\| \, d\mu \ge \log \lambda_0, \tag{A.1}$$

then

$$\mu\Big(\bigcup_{0 \le k < i} R_k(f) \Big) > 0. \tag{A.2}$$

To complete the induction step, it is enough to show that

$$\mu\Big(\bigcup_{0\le k\le j} R_k(f)\Big) = 0 \tag{A.3}$$

for any measure $\mu \in \mathcal{M}(f^m|_{R_{j+1}(f)})$. Indeed, by Theorem A.6, this fact implies that no measure $\mu \in \mathcal{M}(f^m|_{R_{j+1}(f)})$ satisfies relation (A.1). Hence,

$$\int_{R_{j+1}(f)} \log \|Df^m|_{E_{j+1}}\|\, d\mu < \log\lambda_0 < 0$$

for any $\mu \in \mathcal{M}(f^m|_{R_{j+1}(f)})$. Then Df contracts on the subspaces E_{j+1} by Theorem A5, and the set R_{j+1} is hyperbolic by Theorem A.4.

To prove that relation (A.3) is satisfied for any measure $\mu \in \mathcal{M}(f^m|_{R_{j+1}(f)})$, it is not enough to assume that $f \in \mathcal{F}(M)$. Now we assume that the diffeomorphism f is structurally stable.

Mañé introduces the notion of a basic set for f as follows: A hyperbolic set Λ is called a basic set of f if Λ contains a dense semitrajectory (i.e., there is a point $x \in \Lambda$ such that $\omega(x, f) = \Lambda$) and is isolated (i.e., there exists a neighborhood U of the set Λ such that if $V = \operatorname{Cl} U$, then

$$\bigcap_k f^k(V) = \Lambda;$$

in other words, if a trajectory of f belongs to V, then this trajectory belongs to Λ).

Note that basic sets defined after Theorem 7.2 have the above-formulated properties.

Since the hyperbolic structure $\{S(x), U(x) : x \in \Lambda\}$ is continuous on the set Λ, the existence of a dense semitrajectory implies that the dimensions of the stable subspaces $S(x)$ are the same for all points $x \in \Lambda$; this common value is called the index of the basic set and denoted $\operatorname{Ind}(\Lambda)$.

The sets $W^s(\Lambda)$ and $W^u(\Lambda)$ are defined similarly to the sets $W^s(\Omega_i)$ and $W^u(\Omega_i)$.

Let g be a diffeomorphism that coincides with f in a neighborhood of a basic set Λ (clearly, Λ is a basic set of g as well). Analogs of the sets $W^s(\Lambda)$ and $W^u(\Lambda)$ for g are denoted $W^s_g(\Lambda)$ and $W^u_g(\Lambda)$, respectively.

In the following statement, we consider a compact invariant set Λ having a dominated splitting and such that there exists a basic set Λ_i for which the complement $\Lambda \setminus \Lambda_i$ is not closed.

Theorem A.7. *Let Λ be a compact invariant set of f having the following properties: $\Omega(f|_\Lambda) = \Lambda$ and f has a dominated splitting $\{E(x), F(x)\}$ on Λ.*

Assume, in addition, that there exist basic sets $\Lambda_1, \dots, \Lambda_s$ of the diffeomorphism f and numbers $c, m > 0$ and $\lambda \in (0, 1)$ such that

(1) $\mathrm{Ind}(\Lambda_i) < \dim E(x)$ *for all* $1 \le i \le s$ *and* $x \in \Lambda$;

(2) *there exists a* C^1*-neighborhood* U *of* f *such that if a diffeomorphism* $g \in U$ *coincides with* f *in a neighborhood of the set*

$$\bigcup_{1 \le k \le s} \Lambda_k,$$

then

$$W^s_g(\Lambda_i) \cap W^u_g(\Lambda_i) = \Lambda_i$$

for all $1 \le i \le s$;

(3) *if a measure* $\mu \in \mathcal{M}(f^m|_\Lambda)$ *satisfies the inequality*

$$\int_\Lambda \log \|Df^m|_E\| \, d\mu \ge -c,$$

then

$$\mu\Big(\bigcup_{1 \le k \le s} \Lambda_k\Big) > 0;$$

(4) $\|Df^m|_{E(x)}\| \cdot \|Df^{-m}|_{F(f^m(x))}\| \le \lambda$ *for all* $x \in \Lambda$.

Then the following statement holds: If the set $\Lambda \setminus \Lambda_i$ *is not closed for some* $1 \le i \le s$, *then any* C^1*-neighborhood of the diffeomorphism* f *contains a diffeomorphism* g *such that* g *coincides with* f *in a neighborhood of the set*

$$\bigcup_{1 \le k \le s} \Lambda_k$$

and there exists an index $1 \le r \le s, r \ne i$ *such that the set* $\Lambda \setminus \Lambda_r$ *is not closed and*

$$W^s_g(\Lambda_i) \cap W^u_g(\Lambda_r) \ne \emptyset.$$

It is relatively easy to prove the following statement. Let $g \in \mathcal{F}(M)$; denote by $N(i, m, g)$ the number of fixed points of the diffeomorphism g^m belonging to $P_i(g)$.

Theorem A.8. *If* $f \in \mathcal{F}(M)$, *then there exists a* C^1*-neighborhood* U *of the diffeomorphism* f *such that*

(1) $N(i, m, g) = N(i, m, g')$ *for any* $g, g' \in U, m > 0$, *and any* $0 \le i \le n$;

(2) *if a diffeomorphism* $g \in U$ *coincides with* f *in a neighborhood of the set* $R_i(f)$ *for some* $0 \le i \le n$, *then* $R_i(g) = R_i(f)$.

Let us return to the induction process described before Theorem A.6.

We have to show that if a diffeomorphism f is structurally stable (hence, $f \in \mathcal{F}(M)$) and the sets $R_0, \dots, R_j$ are hyperbolic, then the set R_{j+1} is hyperbolic as well. It was mentioned above that to prove the last statement it is enough to show that relation (A.3) holds for any measure $\mu \in \mathcal{M}(f^m|_{R_{j+1}(f)})$.

To get a contradiction, let us assume that there exists a measure $\mu_0 \in \mathcal{M}(f^m|_{R_{j+1}(f)})$ such that

$$\mu_0\Big(\bigcup_{1\le k\le j} R_k(f)\Big) > 0. \tag{A.4}$$

Note that we can represent the hyperbolic set $\bigcup_{1\le k\le j} R_k(f)$ in which periodic points of f are dense in the form

$$\bigcup_{1\le k\le j} R_k(f) = \Lambda_1 \cup \dots \cup \Lambda_s, \tag{A.5}$$

where $\Lambda_1, \dots, \Lambda_s$ are disjoint basic sets (to prove this fact, one can apply literally the same reasoning as that used in the proof of Theorem 7.2).

Note, in addition, that if the intersections $\Lambda_k \cap R_{j+1}(f)$ are empty for all $1 \le k \le s$, then relation (A.4) cannot hold since the set R_{j+1} is the carrier of the measure μ_0. Hence, it follows from relation (A.4) that at least one of the intersections $\Lambda_k \cap R_{j+1}(f)$ is not empty. In this case, it is easy to show that one of the sets $R_{j+1}(f) \setminus \Lambda_k$ is not closed. Thus, the following statement holds.

Theorem A.9. *If there exists a measure $\mu_0 \in \mathcal{M}(f^m|_{R_{j+1}(f)})$ satisfying relation* (A.4)*, then there exists a basic set Λ_k in decomposition* (A.5) *such that the set $R_{j+1}(f) \setminus \Lambda_k$ is not closed.*

Now we check that Theorem A.7 is applicable to the set $\Lambda = R_{j+1}(f)$, dominated splitting $E_{j+1} \oplus F_{j+1}$ on $R_{j+1}(f)$, basic sets $\Lambda_1, \dots, \Lambda_s$, and numbers $\lambda \in (0, 1)$, $m > 0$ (see Theorem A.3), and $c = -\log \lambda_0$ (see Theorem A.6).

Since $\operatorname{Ind}(\Lambda_i) \le j$ for all i and $\dim E_{j+1}(x) = j + 1$ for $x \in R_{j+1}(f)$, assumption (1) of Theorem A.7 is satisfied. The equality

$$\Omega(f|_{R_{j+1}(f)}) = R_{j+1}(f)$$

holds since periodic points are dense in the set $R_{j+1}(f)$. Condition (4) is a corollary of Theorem A.3.

In addition, it follows from Theorem A.6 that if a measure $\mu \in \mathcal{M}(f^m|_{R_{j+1}(f)})$ satisfies the inequality

$$\int_{R_{j+1}(f)} \log \|Df^m|_{E_{j+1}}\| \, d\mu \ge -c = \log \lambda_0,$$

then

$$\mu(\Lambda_1 \cup \cdots \cup \Lambda_s) = \mu\Big(\bigcup_{0\le k\le j} R_k(f)\Big) > 0.$$

Hence, assumption (3) is fulfilled.

Thus, it remains to check assumption (2).

Assume that a diffeomorphism g belongs to the neighborhood U of the diffeomorphism f described in Theorem A.8, coincides with f in a neighborhood of the set $\Lambda_1 \cup \cdots \cup \Lambda_s$, and, at the same time,

$$W_g^s(\Lambda_q) \cap W_g^u(\Lambda_q) \neq \Lambda_q$$

for some $1 \le q \le s$.

A statement similar to Theorem 7.4 implies that there exist points $x, y \in \Lambda_q$ and p such that p belongs to the set

$$\big(W_g^s(\Lambda_q) \cap W_g^u(\Lambda_q)\big) \setminus \Lambda_q$$

and is a point of transverse intersection of the manifolds $W_g^u(x)$ and $W_g^s(y)$ (possibly, we have to slightly perturb the diffeomorphism g to make the intersection of the manifolds $W_g^u(x)$ and $W_g^s(y)$ transverse).

Let X and Y be compact disks in the manifolds $W_g^u(x)$ and $W_g^s(y)$ such that the point p belongs to the intersection of the interiors of these disks.

Since periodic points of the diffeomorphism g are dense in the set Λ_q, there exist sequences of periodic points $r_k, \rho_k \in \Lambda_q$ such that $r_k \to x$ and $\rho_k \to y$.

It follows from known properties of stable and unstable manifolds of points belonging to hyperbolic sets (see, for example, [15]) that the stable and unstable manifolds $W_g^u(r_k)$ and $W_g^s(\rho_k)$ contain smooth disks X_k and Y_k, respectively, that converge to X and Y with respect to C^1 topology.

Lemma 6.3 implies that if k is large enough, then the disks X_k and Y_k contain points p_k of transverse intersection such that $p_k \to p$.

The same reasoning as in the proof of Lemma 7.6 shows that the manifolds $W_g^u(r_k)$ and $W_g^s(\rho_k)$ have a point of transverse intersection; now it follows from the λ-lemma (Lemma 6.4) that the manifolds $W_g^s(\rho_k)$ contain sequences of smooth disks that converge to X_k with respect to C^1 topology.

Applying Lemma 6.3 once more, we see that there exist transverse homoclinic points π_k of the periodic points ρ_k such that $\pi_k \to p$.

An arbitrarily small neighborhood of a transverse homoclinic point for a periodic point ρ with $\dim W^s(\rho) = m$ contains periodic points π with $\dim W^s(\pi) = m$ (see [50, 51]). Hence, it follows from our reasoning above that any neighborhood of the point p contains periodic points π of the diffeomorphism g such that the dimension of the stable manifolds of these periodic points coincides with $\mathrm{Ind}(\Lambda_q)$ (denote this dimension by l).

Note that the point p does not belong to the union $\Lambda_1 \cup \cdots \cup \Lambda_s$. Indeed, if $p \in \Lambda_j$, then the whole trajectory $O(p, g)$ belongs to Λ_j as well, but since the points $g^k(p)$ tend to Λ_q as $k \to \pm\infty$, we conclude that $j = q$, which is impossible.

The diffeomorphism g coincides with f in a neighborhood of the set $\Lambda_1 \cup \cdots \cup \Lambda_s$ which contains all the periodic points of f for which the dimension of the stable manifold equals l. Since g has periodic points that are arbitrarily close to p (hence, such points are outside $\Lambda_1 \cup \cdots \cup \Lambda_s$), there exists a ν such that $N(l, \nu, g) > N(l, \nu, f)$.

The obtained contradiction with Theorem A.8 shows that condition (2) of Theorem A.7 is fulfilled.

Let us apply Theorem A.7 to the sets $\Lambda_1, \dots, \Lambda_s$ and R_{j+1}.

By Theorem A.9, there exists a set Λ_k such that the set $R_{j+1} \setminus \Lambda_k$ is not closed. Let t be the minimal index of the basic sets Λ_k for which the differences $R_{j+1} \setminus \Lambda_k$ are not closed. Let us select a basic set Λ_i having the following properties:

- The difference $R_{j+1} \setminus \Lambda_i$ is not closed;
- $\operatorname{Ind} \Lambda_i = t$;
- if the set $R_{j+1} \setminus \Lambda_k$ is not closed and $k \neq i$, then

$$W_f^s(\Lambda_i) \cap W_f^u(\Lambda_k) = \emptyset.$$

Let us show that such a basic set Λ_i exists. Otherwise, we can find different basic sets $\Lambda_{i_1}, \dots, \Lambda_{i_p}$ such that the indices of these sets equal t and there exist points $x_1, \dots, x_p$ such that

$$x_j \in W_f^s(\Lambda_{i_j}) \cap W_f^u(\Lambda_{i_{j+1}}), \quad j = 1, \dots, p-1,$$

and

$$x_p \in W_f^s(\Lambda_{i_p}) \cap W_f^u(\Lambda_{i_1}).$$

An analog of Theorem 7.4 for basic sets Λ_k implies that there exist points $r_j, \rho_j \in \Lambda_{i_j}$ such that

$$x_j \in W_f^s(\rho_{i_j}) \cap W_f^u(r_{i_{j+1}})$$

(we agree that $\Lambda_{i_{p+1}} = \Lambda_{i_1}$).

Since the diffeomorphism f is structurally stable, the manifolds $W_f^s(\rho_{i_j})$ and $W_f^u(r_{i_{j+1}})$ are transverse at the points x_j (it had been shown long ago that if a diffeomorphism is structurally stable, then stable and unstable manifolds of its nonwandering points are transverse, see [45]).

Applying a reasoning similar to the proof of Theorem 7.2 and based on the existence of a dense semitrajectory in the set Λ_k, one can show that the manifolds $W_f^s(\rho_{i_j})$ and $W_f^u(r_{i_j})$ have points of transverse intersection; thus, there appears a contour of hyperbolic trajectories that are connected by trajectories along which their stable and unstable manifolds intersect transversally.

The behavior of trajectories in a neighborhood of such a contour is similar to that in a neighborhood of a transverse homoclinic trajectory; hence, an arbitrary neighborhood of the point x_1 contains periodic points for which the dimensions of stable manifolds equal t. This implies that $x_1 \in R_j$. Hence, x_1 belongs to one of the sets Λ_k, but in this case, the sets Λ_{i_1} and Λ_{i_2} coincide with Λ_k.

We have got a contradiction; this shows that there exists a basic set Λ_i having the desired properties.

We apply to this basic set the conclusion of Theorem A.7 and find in an arbitrary C^1-neighborhood of the diffeomorphism f a diffeomorphism g such that g coincides with f in a neighborhood of the set

$$\bigcup_{1 \le k \le s} \Lambda_k$$

and there exists an index $1 \le r \le s, r \ne i$, such that the set $\Lambda \setminus \Lambda_r$ is not closed and

$$W_g^s(\Lambda_i) \cap W_g^u(\Lambda_r) \ne \emptyset. \tag{A.6}$$

Since any point belonging to the intersection (A.6) is a point of transverse intersection of stable and unstable manifolds of hyperbolic nonwandering points, the inequality $\mathrm{Ind}(\Lambda_i) \ge \mathrm{Ind}(\Lambda_i)$ holds (compare with Lemma 6.4).

The definition of the number t implies that $\mathrm{Ind}(\Lambda_i) = \mathrm{Ind}(\Lambda_i) = t$.

Since we can find such a diffeomorphism g in an arbitrary C^1-neighborhood of the diffeomorphism f, we may assume that g and f are topologically conjugate (let h be a homeomorphism of M such that $g \circ h = h \circ f$).

Clearly, $h(P_i(f)) = P_i(g)$ for all $0 \le i \le n$; then $h(R_i(f)) = R_i(g)$ for all $0 \le i \le n$.

Hence,

$$h\Big(\bigcup_{1 \le k \le s} \Lambda_k\Big) = h\Big(\bigcup_{1 \le k \le j} R_k(f)\Big) = \bigcup_{1 \le k \le j} R_k(g). \tag{A.7}$$

Since the diffeomorphism g coincides with f in a neighborhood of the set

$$\bigcup_{1 \le k \le s} \Lambda_k,$$

statement (2) of Theorem A.8 implies that $R_i(f) = R_i(g)$ for all $0 \le i \le j$.

Hence, it follows from equalities (A.7) that

$$h\Big(\bigcup_{1 \le k \le s} \Lambda_k\Big) = \bigcup_{1 \le k \le s} \Lambda_k.$$

It is easily seen that if $1 \leq k \leq s$, then $h(\Lambda_k)$ is one of the sets of the family $\Lambda_1, \ldots, \Lambda_s$, and its index is the same as the index of Λ_k.

Define $T(f)$ as the set of pairs (m, q) such that $m \neq q$, $\mathrm{Ind}(\Lambda_m) = \mathrm{Ind}(\Lambda_q) = t$, and

$$W^s_f(\Lambda_m) \cap W^u_f(\Lambda_q) \neq \emptyset. \tag{A.8}$$

The set $T(g)$ is defined similarly to $T(f)$.

Since the homeomorphism h maps any set of the family $\Lambda_1, \ldots, \Lambda_s$ to a set of this family with the same index,

$$\operatorname{card} T(f) = \operatorname{card} T(g). \tag{A.9}$$

In addition, all the intersections of stable and unstable manifolds in (A.8) are transverse. Hence, if the diffeomorphism g is close enough to f, then all such intersections are preserved. It follows from (A.9) that

$$T(f) = T(g). \tag{A.10}$$

Now we note that $(i, r) \notin T(f)$ due to the choice of the basic set Λ_i. At the same time, it follows from (A.6) that $(i, r) \in T(g)$. We have obtained a contradiction with (A.10). This completes the proof.

Appendix B
Lectures on the history of differential equations and dynamical systems

Of course, it is impossible to tell a more or less complete history of the theory of differential equations and dynamical systems in several lectures. The goal of these notes is different. We just want to highlight several most important (from the author's point of view) moments in the development of the theory of differential equations and to demonstrate that this development was not a rectilinear movement along a smooth road (of course, the same is true for any serious field of mathematics and for the whole science). We will see that the development was accompanied with errors, misunderstandings, etc.

We include several nonmathematical comments into the text; these comments are indicated as follows: $\bullet \cdots \bullet$.

B.1 Differential equations and Newton's anagram[1]

When we speak about theory of differential equations as a separate domain of mathematics, the first name which comes into our mind is the name of Isaac Newton.[2]

The modern calculus as a whole developed from the theory of infinitesimals created at the end of the 17th century by Newton and G. Leibniz.[3]

It is well known that the creation of the theory of infinitesimals was accompanied with long discussions concerning priority.

For the main part of the modern specialists in history of mathematics, there is no problem of priority; it is clear that Newton and Leibniz developed the theory independently – it is enough to look at their terminology.

The flavor of physics permeates the terminology used by Newton (in fact, many researchers believe that for Newton, mathematics was just a tool applied to the development of physics). The main object of his study, the value which changes when the independent variable changes, was called fluenta (this word comes from Latin and means flow); the rate of change of the fluenta was called fluxia, etc.

At the same time, Leibniz used the terminology which we use in the modern calculus: function, differential, integral.

Of course, Newton created his methods having strong predecessors. He had mentioned that one of the benchmarks of his methods was the Fermat[4] method of tangents.

---

[1] In this section, we mostly follow the book [52].
[2] Sir Isaac Newton, 1642–1727.
[3] Baron Gottfried Wilhelm von Leibniz, sometimes Leibnitz, 1646–1716.
[4] Pierre de Fermat, 1601–1665.

Before Newton, the Fermat method was developed by Gregory,[5] Newton's teacher Barrow,[6] Wallis,[7] and others.

Newton was of high opinion concerning the merits of his predecessors; this opinion was reflected by his famous phrase "If I have seen further it is only by standing on the shoulders of giants" written in a letter to Robert Hooke.[8]

• Some historians of science think that this phrase may have an additional, not so lofty, sense – in this way, Newton could express his enmity to Hooke, who was undersized. •

Newton created his method of fluxions as a mathematical apparatus for studying the structure of the world around us. In his opinion, derivation and solution of differential equations were the main tools in the study of any process of change. This is confirmed by the famous Newton's anagram:[9]

6a cc d æ 13e ff 7i 3l 9n 4o 4q rr 4s 8t 12u x.

In 1676, Leibniz visited London. The 2nd Newton's letter to Leibniz (via Henry Oldenburg, the Secretary of the Royal Society) contained the above anagram. This anagram formulated, in Newton's opinion, the basic principle of the calculus and his own main discovery:

> Data æquatione quotcunque fluentes quantitates involvente, fluxiones invenire; et vice versa.[10]

The literal translation of this Latin phrase (included into the main Newton's book, "Philosophiæ Naturalis Principia Mathematica" [53]) is as follows: "Given an equation involving any number of fluent quantities to find the fluxions, and vice versa."

The great Russian mathematician V. I. Arnold[11] gave a short modern translation of the Newton's phrase: "It is useful to solve differential equations."

When Newton wrote "Principia," the method of fluxions had been created. Nevertheless, Newton did not use the new mathematical language in his description of such great discoveries as the fundamental laws of mechanics (Newton's laws). One of possible explanations is that for Newton, mathematics played an auxiliary role in the development of physical theories. There is a different explanation; Newton could apprehend misunderstanding due to the novelty of the language.

---

[5] James Gregory, 1638–1675.

[6] Isaac Barrow, 1630–1677.

[7] John Wallis, 1616–1703.

[8] Robert Hooke, 1635–1703.

[9] In an anagram, the characters of a phrase are arranged in alphabetical order. Natural philosophers of that time often transposed their discoveries into Latin anagrams to establish their priority.

[10] An attentive reader may note that the phrase contains nine t's instead of eight; in addition, u's and v's are counted as the same character. Some possible explanations can be found, for example, at http://www.mathpages.com/home/kmath414/kmath414.htm.

[11] Vladimir Igorevich Arnold, 1937–2010.

Such apprehensions were not groundless; it is known, for example, that Huygens[12] wrote to Leibniz in 1692 that he cannot understand why the differential calculus is better than the old methods.

Newton did not publish his work on calculus until after Leibniz had published his.

Publishing "Optics" (in 1704), Newton added to this book two appendices, "De quadratura curvarum" and "Enumeratio linearum tertii ordinis," related to his mathematical discoveries.

A systematical treatment of the new Newton's methods (containing solutions of some simple differential equations) had been published posthumously, in 1736, in the book "The method of fluxions and infinite series."

B.2 Development of the general theory

The main Newton's method for solving differential equations was based on representation of a solution in the form of a power series and search for coefficients of the series. At that time, nobody worried about the convergence of the appearing series (though it is easy to give examples of differential equations for which such series diverge everywhere with the exception of the initial point).

The first rigorous result concerning the convergence of power series representing solutions of differential equations belongs to Cauchy.[13]

The Cauchy theorem was the first one in the long line of existence theorems obtained by mathematicians in the process of development of theory of differential equations.

Cauchy was an outstanding representative of the mathematical movement whose goal was to improve the level of rigority of proofs. Such movements appear in mathematics periodically; they specify new standards of rigor (which sometimes do not suit the next generations of mathematicians . . .). One does not have to exaggerate the influence of such processes on the development of mathematics as a whole (many modern mathematicians do not relate their own research and, for example, the Gödel[14] incompleteness theorems).

Cauchy had published more than 800 scientific papers. In one of them, he had proven that if the right-hand side of the system of differential equations

$$\frac{dx}{dt} = f(t, x) \tag{B.1}$$

is analytic in a neighborhood of a point (t_0, x_0), then there exists a unique solution of this system that satisfies the condition

$$x(t_0) = x_0; \tag{B.2}$$

[12] Christian Huygens, sometimes Huyghens, 1629–1695.
[13] Augustin Louis Cauchy, 1789–1857.
[14] Kurt Gödel, 1906–1978.

this solution is given by a power series,

$$x(t) = \sum_{k=0}^{\infty} c_k (t - t_0)^k, \tag{B.3}$$

that converges in a neighborhood of the point t_0. Problem (B.1)–(B.2) is called the initial (or, more often, Cauchy) problem.

The proof given by Cauchy consists of two parts. First, it is shown that one can uniquely determine the coefficients c_k by substituting series (B.3) into system (B.1) and equating coefficients of the same degrees of $(t - t_0)^k$ on the right and left. After that, Cauchy constructs a special (majorizing) system of differential equations that has a solution given by a converging series and shows that this solution gives a majorant for the obtained series (B.3).

The further development of mathematics demonstrated that the Cauchy method of majorants and its modifications are very powerful tools applicable to a wide range of problems.

The next important step was related to the application of the Lipschitz[15] condition.

This condition (Lipschitz continuity) had been introduced by Lipschitz in 1864 to study the convergence of Fourier[16] series. Picard[17] showed that it is possible to establish the unique solvability of the Cauchy problem under conditions that are significantly weaker than analyticity.

Picard showed that if the right-hand side of system (B.1) is continuous and Lipschitz continuous in x in a neighborhood (t_0, x_0), then the successive approximations defined as follows:

$$g_0(t) \equiv x_0, \quad g_{k+1}(t) = x_0 + \int_{t_0}^{t} f(s, g_k(s))ds, \quad k \geq 0, \tag{B.4}$$

uniformly converge to a solution of the Cauchy problem (B.1)–(B.2) on some interval containing the point t_0.

The Picard existence theorem (sometimes called the Cauchy–Picard or Cauchy–Lipschitz theorem) is, in fact, the basic existence theorem in the modern theory of ordinary differential equations.

Similarly to the Cauchy method of majorants, the Picard method of successive approximations (sometimes called Picard iterations) based on analogs of the iteration process (B.4) is applicable to a wide range of problems in which one has to prove the solvability of some operator equations.

Finally, let us mention the Peano[18] existence theorem. Peano (known also due to his axioms for the natural numbers) had shown that to prove the solvability of

[15]Rudolf Lipschitz, 1832–1903.

[16]Jean Baptiste Joseph Fourier, 1768–1830.

[17]Charles Émile Picard, 1856–1941.

[18]Giuseppe Peano, 1858–1932.

the Cauchy problem (B.1)–(B.2), it is enough to assume that the right-hand side of system (B.1) is continuous in a neighborhood of the point (t_0, x_0). Note that in this case, the solution of problem (B.1)–(B.2) is not necessarily unique (in contrast to the uniqueness guaranteed by the Picard conditions).

The proof of Peano's theorem is based on the simplest piecewise-linear approximations of a solution introduced by Euler.[19] The formulas defining these approximations are extremely simple; for this reason, they are widely applied in the development of software for modern computers.

The mathematical literature contains existence theorems for differential equations with discontinuous right-hand sides, but such theorems are related to control theory rather than to theory of differential equations.

In fact, some mathematicians solved differential equations before the creation of the fundament of calculus by Newton and Leibniz; we know that several geometric problems reducing to differential equations had been solved in the 17th century.

These geometric problems were reduced to differential equations solvable by quadratures (this means that solutions could be expressed by formulas involving one or more operations of integration). Usually, one had to apply more or less complicated changes of variables, which allowed to solve by quadratures various classes of differential equations (a lot of progress in this direction was due to Euler).

Up to now, some mathematicians continue to search for new classes of differential equations solvable by quadratures though such an activity is of limited interest due to the Liouville[20] theorem, see below.

The development of the theory of solving differential equations by quadratures repeated the general scheme of the development of the theory of solving algebraic equations (as we show below, the connection between the two theories was not formal).

It was known from the ancient time how to solve quadratic equations. The problem of solving general cubic and quartic equations had been solved by Italian mathematicians Dal Ferro,[21] Tartaglia,[22] and Ferrari[23] at the first half of the 17th century.

A lot of effort had been applied to solve the general fifth-degree equation by radicals (i.e., to obtain a formula that expresses the roots of an arbitrary algebraic equation of degree 5 in terms of its coefficients). Only at the beginning of the 19th century, Abel[24] had proven that it is impossible to solve a general algebraic equation of degree n with $n \geq 5$ by radicals.

The works of Abel and Galois[25] related to the problem of solvability of algebraic equations by radicals had formed the fundament of the modern algebra.

---

[19] Leonhard Euler, 1707–1783.
[20] Joseph Liouville, 1809–1882.
[21] Scipione Dal Ferro, 1465–1526.
[22] Nicolo Fontana Tartaglia, 1499–1557.
[23] Lodovico Ferrari, 1522–1565.
[24] Niels Henrik Abel, 1802–1829.
[25] Évariste Galois, 1811–1832.

• Galois died from wounds suffered in a duel at the age of 20. The main part of his mathematical research was left in the form of manuscripts. Galois' mathematical contributions were published only in 1846 due to Liouville who reviewed the Galois manuscripts and declared them sound. •

As one of the most important consequences of this Liouville activity, we must mention his creation of the so-called differential algebra (a theory in a sense parallel to the Galois theory of solvability of algebraic equations). One of the famous results of this theory is the Liouville nonintegrability theorem (1841).

An Italian engineer and mathematician Riccati[26] considered the first-order differential equation

$$\frac{dx}{dt} + ax^2 = bt^\alpha \tag{B.5}$$

(usually called the special Riccati equation).

Daniel Bernoulli[27], a representative of the large Bernoulli mathematical family, found in 1824–1825 a solution by quadratures for equation (B.5) for all $\alpha = -4k/(2k-1)$, $k \in \mathbb{Z}$.

• Let us note that this was done in the first published paper of D. Bernoulli; this paper was written to support his father Johann Bernoulli[28] and his uncle Jacob[29] in their discussions with Italian mathematicians. •

Many mathematicians tried to extend the set of exponents α for which equation (B.5) is solvable by quadratures; Liouville found the final solution of this problem.

We do not state here the Liouville theorem. Instead, we describe (on intuitive level) one of its corollaries (for the case of the Riccati equation).

Consider a class of functions $\mathcal{L}$ that is constructed as follows. We first include all the elementary functions (such as polynomials, exponentials, trigonometric functions) into this class. After that, we make the class $\mathcal{L}$ closed with respect to taking any finite compositions of its elements, their inverse functions, and primitives. Clearly, such a class $\mathcal{L}$ will contain all the functions representable by finite closed-form expressions.

It follows from the Liouville theorem that if the exponent α in equation (B.5) differs from one of the Bernoulli exponents $\alpha = -4k/(2k-1)$ (and also $\alpha \neq -2$), then equation (B.5) does not have solutions belonging to the class $\mathcal{L}$ (thus, no solution can be expressed by a finite formula).

This results indicates that it is impossible to solve by quadratures an arbitrary differential equation (thus, the situation is parallel to that with solving algebraic equations by radicals).

As a result of the appearance of the Liouville theorem, specialists in differential equations understood that it is reasonable to study properties of solutions not by their

[26] Jacopo Francesco Riccati, 1676–1754.
[27] Daniel Bernoulli, 1700–1782.
[28] Johann, sometimes Jean, Bernoulli, 1667–1748.
[29] Jacob, sometimes Jacques, Bernoulli, 1654–1705.

closed-form expressions but by properties of the right-hand sides of the equations (the corresponding branch of the theory is called qualitative theory of differential equations).

One of the first representatives of this approach was the great French mathematician Poincaré[30] whose role in qualitative theory of differential equations is described later.

B.3 Linear equations and systems

The input of L. Euler into theory of differential equations is priceless. He had invented multiple changes of variables which allowed to solve by quadratures various nonlinear differential equations, integrated the Riccati equation using continuous fractions, developed the method of integrating factors.

A Russian mathematician N. N. Luzin[31] said that for 150 years, mathematicians could not breach the ring of integrations forged by Euler.

Nevertheless, the most important Euler's achievement in theory of differential equations is his general method for solving linear differential equations with constant coefficients.

Euler considered differential equations of the form

$$0 = Ay + B\frac{dy}{dx} + C\frac{d^2y}{dx^2} + \cdots + N\frac{d^ny}{dx^n}$$

and applied to these equations his own method for reduction of order based on consecutive differentiating of the function

$$y = e^{\int p\,dx}.$$

Euler wrote: "If p is constant, then its derivatives vanish, and we get the algebraic equation

$$0 = A + Bp + Cp^2 + \cdots + Np^n.$$

A solution p of this equation gives us a particular solution $y = e^{px}$."

We have to note that originally, the method had not achieved its final form. First, treating the case of complex roots of the characteristic equation, Euler separately solved equations for the real and imaginary parts of the solution; this is precisely where the famous Euler formula

$$e^{ikx} = \cos kx + i\sin kx$$

[30] Jules Henri Poincaré, 1854–1912.

[31] Nikolay Nikolaevich Luzin, 1883–1950.

had been introduced into mathematics. The preliminary Euler's approach to the case of multiple roots of the characteristic equation was also more complicated than that used now.

The method was not accepted immediately by Euler's contemporaries. Euler first wrote about the new method in a letter to Johann Bernoulli dated September 15, 1739. As Bernoulli's answer shows, he could not estimate the importance of the new approach.

Euler applied his method to linear systems of differential equations with constant coefficients.

It is well known that to find the set of solutions of a linear system

$$\frac{dx}{dt} = P(t)x, \quad x \in \mathbb{R}^n \text{ (or } x \in \mathbb{C}^n\text{)}, \tag{B.6}$$

with a continuous matrix of coefficients $P(t)$, it is enough to know a fundamental matrix $\Phi(t)$ of this system (i.e., a matrix whose columns are n linearly independent solutions of system (B.6)); in this case, the formula

$$x(t) = \Phi(t)c$$

establishes a linear isomorphism between the space of solutions of system (B.6) and the space of constant n-dimensional vectors c.

In its modern form, Euler's method represents a fundamental matrix of the system with constant coefficients

$$\frac{dx}{dt} = Ax \tag{B.7}$$

in the form $\Phi(t) = \exp(At)$.

To be more exact, applying the Euler method to system (B.7), we get a fundamental matrix in the form $\Phi(t) = S\exp(Jt)$, where J is a Jordan form of the matrix A and S is the reducing matrix, i.e., a matrix satisfying the equality $J = S^{-1}AS$.

In the general case (where $n > 1$ and the coefficient matrix $P(t)$ of system (B.6) is not constant), there is no general method for finding a fundamental matrix via the matrix $P(t)$.

Indeed, consider the linear differential equation

$$\frac{d^2y}{dt^2} + t^\alpha y = 0. \tag{B.8}$$

It is easily seen that if $y(t)$ is a nonzero solution of equation (B.8), then the function

$$x(t) = \frac{1}{y}\frac{dy}{dt}$$

is a solution of the Riccati equation

$$\frac{dx}{dt} + x^2 = -t^\alpha.$$

It follows from the Liouville theorem discussed above that if $\alpha \neq -4k/(2k-1)$ and $\alpha \neq -2$, then any nonzero solution of equation (B.8) does not belong to the class $\mathcal{L}$ (note that if $y(t) \in \mathcal{L}$, then $dy/dt \in \mathcal{L}$ and $x(t) \in \mathcal{L}$).

It follows that if

$$\begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$$

is a column of a fundamental matrix of the system

$$\frac{dy}{dt} = z, \quad \frac{dz}{dt} = -t^\alpha y$$

corresponding to equation (B.8), then the function $y(t)$ does not belong to the class $\mathcal{L}$.

The general theory of linear systems of differential equations had been completed long ago (and, in fact, this theory is relatively simple). At the same time, we know really less about linear differential equations and systems with discontinuous coefficients.

A lot of effort was devoted to study of the so-called Fuchs[32] equations and systems. It is natural to study such equations and systems for unknown functions of a complex variable (denoted z below).

A differential equation

$$\frac{d^n y}{dz^n} + p_1(z)\frac{d^{n-1} y}{dz^{n-1}} + \cdots + p_n(z)y = 0 \tag{B.9}$$

belongs to the Fuchs class if

$$p_j(z) = \prod_{m=1}^{k} (z - a_m)^{-j} q_j(z),$$

where $a_1, \ldots, a_k$ are distinct points, and q_j are polynomials of degree not exceeding $j(k-1)$.

A system of differential equations

$$\frac{dy}{dz} = A(z)y, \quad y \in \mathbb{C}^n,$$

belongs to the Fuchs class if

$$\frac{dy}{dz} = \sum_{m=1}^{k} \frac{A_m}{z - a_m} y, \tag{B.10}$$

where $a_1, \ldots, a_k$ are distinct points, and A_m are nonzero constant $n \times n$ matrices.

[32]Immanuel Lazarus Fuchs, 1833–1902.

We pay special attention to equations and systems of the Fuchs class since such equations and systems are mentioned in the 21st problem in the famous list of problems posed by Hilbert[33] in 1900. This list of problems selected several most important mathematical problems of that time and significantly influenced the development of mathematics in the 20th century.

In these lectures, we mention two Hilbert problems related to differential equations (it would be strange not to mention them speaking about the history of differential equations, especially taking into account that in both cases, the history was quite nontrivial). We discuss the second problem (16th in the Hilbert list) later.

Considering equation (B.9), Riemann[34] posed the following problem (later, this problem was called Riemann–Hilbert).

Let us describe the Riemann–Hilbert problem in the case of system (B.10) of the Fuchs class.

Fix a base point a of the complex plane and let γ_j be a simple contour containing the point a and surrounding precisely one pole a_j (see Figure 6).

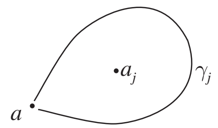


Figure 6. Contour γ_j.

Since coefficients of system (B.10) are continuous in a neighborhood of the point a, we can define in this neighborhood a fundamental matrix $\Phi(z)$ of system (B.10). Continuing this matrix along the contour γ_j and returning to the base point a, we get a fundamental matrix $\Psi(z)$ of system (B.10) (that may differ from $\Phi(z)$). Since $\Phi(z)$ and $\Psi(z)$ are fundamental matrices of the same system, there exists a nonsingular constant matrix G_j such that $\Psi(z) = \Phi(z)G_j$. This matrix (independent of the particular choice of the contour γ_j) is called the monodromy matrix (or a monodromy).

It is assumed that

$$A_1 + \cdots + A_k = 0$$

(this is the so-called regularity condition at infinity for system (B.10)). In this case,

$$G_1 \cdots G_k = \mathrm{Id}. \tag{B.11}$$

The 21st Hilbert problem (Riemann–Hilbert problem): Given a series of points in a complex plane and prescribed monodromies around these points, is there a Fuchsian equation (system) with these singularities and monodromies?

[33] David Hilbert, 1862–1943.

[34] Georg Friedrich Bernhard Riemann, 1826–1866.

Thus, we are given points $a_1, \dots, a_k$ and matrices $G_1, \dots, G_k$ that satisfy condition (B.11).

Can we find matrices $A_1, \dots, A_k$ such that the given matrices $G_1, \dots, G_k$ are the monodromy matrices for the Fuchs system (B.10)?

An affirmative answer to the Riemann–Hilbert problem had been given by Plemelj[35] in 1908.

Let us mention here one outstanding mathematician whose life was related to the St. Petersburg-Leningrad University.

• I. A. Lappo-Danilevskii[36] was a son of a famous Russian historian, Academician A. S. Lappo-Danilevskii. In 1914, I. A. Lappo-Danilevskii became a student of St. Petersburg University. Due to his serious heart disease and the complicated after-Revolution life, the active mathematical research by I. A. Lappo-Danilevskii lasted only for several years. In January of 1931, he was elected Corresponding Member of the Academy of Sciences of USSR, and in March of the same year, he died in Giessen (Germany). •

I. A. Lappo-Danilevskii was a student of V. I. Smirnov.[37] According to traditions of the St. Petersburg-Leningrad mathematical school, the main goal in the study of the Riemann–Hilbert problem was not proving abstract existence theorems but constructing algorithms which would allow to find solutions of the corresponding equations and systems in the form of power series involving coefficient matrices; in addition, coefficients of the series must be calculated by recurrent formulas.

A solution of the Riemann–Hilbert problem given by Lappo-Danilevskii followed these traditions. It was shown that if the given monodromy matrices $G_1, \dots, G_k$ are close enough to the identity matrix, then there exists a unique solution $A_1, \dots, A_k$ of the Riemann–Hilbert problem; in addition, this solution has the form

$$A_j = \frac{1}{2\pi i}(G_j - \mathrm{Id}) + \sum_{1 \le l,m \le k} c_{l,m,j}(a_1, \dots, a_k)(G_l - \mathrm{Id})(G_m - \mathrm{Id}) + \cdots,$$

where the series on the right converge for $\|G_j - \mathrm{Id}\| < b$, where $b > 0$.

This solution of the Riemann–Hilbert problem was obtained as an application of a general theory of analytic functions of matrices developed by Lappo-Danilevskii. The monograph [54] devoted to this theory had been published only in 1957; it contains a lot of ideas which still have to be introduced into the modern theory of differential equations.

The history of the 21st Hilbert problem had not been completed by the works of Plemelj and his followers though it was written in some books (see, for example, [55]) that due to Plemelj's theorem, we have a more or less complete general theory of Fuchs systems of the form (B.10).

[35] Josip Plemelj, 1873–1967.

[36] Ivan Aleksandrovich Lappo-Danilevskii, 1896–1931.

[37] Vladimir Ivanovich Smirnov, 1887–1974.

A monograph containing a full variant of Plemelj's proof was published in 1964. In the 1980s, this proof was analyzed by a young Moscow mathematician A. A. Bolibruch.[38] He understood that, in fact, the Plemelj's proof uses the following (hidden) assumption: At least one of the prescribed monodromy matrices $G_1, \dots, G_k$ has a diagonal Jordan form. For some time, Bolibruch tried to give a proof for the general case, but in 1989 he constructed a counterexample to the 21st problem. Bolibruch showed that there exist families of monodromy matrices $G_1, \dots, G_k$ that do not correspond to Fuchs systems of the form (B.10).

B.4 Stability

Of course, from the point of view of applications, stability theory is the most important branch of theory of differential equations. We do not attempt to tell here a detailed story of stability theory (let us only mention that the book [56] published in 1949 and devoted to this history contains more than 600 pages).

Euler was the first who tried to give a definition of stability. In the book "Scientia navalis seu tractatus de construendis ac dirigendis navibus" ([57]), published in 1749, Euler treats the possible behavior of a body submerged in water and slightly deviated from equilibrium.

Euler defines positive stability (the body returns to the initial equilibrium), negative stability (the body moves away from the equilibrium), and zero stability (the body rests at the deviated position).

To study the stability of a body in water, Euler considered a mathematical pendulum whose behavior models the behavior of the body (i.e., has the same period and amplitude). The corresponding mathematical model was a second-order linear differential equation with constant coefficients (thus, Euler created the foundation of the analytic theory of small oscillations).

• In 1759, Euler obtained the prize for the theory of shipbuilding at the Academy of Sciences, Paris. Academician A. N. Krylov[39] wrote that, as a result of Euler's research, the military sailer got the form which it preserved for 100 years, until 1850s, when sailors were replaced by steamers and, later, by ironclads. •

It is important to note the Lagrange[40] works on stability of mechanical systems.

In his fundamental book on analytical mechanics (1788), Lagrange formulated his main result on stability of an equilibrium for a conservative mechanical system as follows. If the function (potential energy) has minimum, then the equilibrium is stable, so that if the system is slightly deviated from the initial equilibrium, then it tends to return to the equilibrium performing infinitely small oscillations. On the contrary, if the same function has maximum, then the equilibrium is not stable, and, after a devi-

[38] Andrey Andreevich Bolibruch, 1950–2003.
[39] Aleksey Nikolaevich Krylov, 1863–1945.
[40] Compte Joseph Louis Lagrange, 1736–1813.

ation, the system performs oscillations that are not small and move it away from the initial equilibrium.

Note that Lagrange formulates the definition of stability inside the statement of his stability theorem and mentions only small deviations but not small perturbations of velocity (which is necessary when one treats mechanical systems described by second-order differential equations).

The Lagrange proof of the above-mentioned theorem is not complete. Lagrange represents the potential energy in a neighborhood of the equilibrium in the form of a power series and omits terms whose order is greater than two; he explains that since the variables considered are small, it is enough to take into account only second-order terms with respect to these variables.

Thus, Lagrange took into account only the linear approximation of the right-hand side of the considered system of differential equations (which corresponds to second-order terms in the representation of energy).

Such an approach to stability problems was generally accepted preceding the works of the great Russian mathematician A. M. Lyapunov[41] who created mathematical foundations of stability theory.

Let us mention that a rigorous proof of the Lagrange stability theorem was given by Minding[42] in 1838.

It is relatively easy to formulate conditions of asymptotic stability, stability, or instability of the zero solution (and then of an arbitrary solution) for a linear system of differential equations with constant coefficients of the form (B.7). Let λ_j be the eigenvalues of the coefficient matrix A.

The zero solution is asymptotically stable if and only if

$$\operatorname{Re}\lambda_j < 0 \tag{B.12}$$

for all the eigenvalues.

The zero solution is stable if and only if

$$\operatorname{Re}\lambda_j \le 0$$

for all the eigenvalues and if there exists an eigenvalue λ_j with $\operatorname{Re}\lambda_j = 0$, then all Jordan blocks (in a complex Jordan form) corresponding to this eigenvalue must be simple.

In the remaining cases, the zero solution is unstable.

One can check that condition (B.12) is satisfied for all the eigenvalues of the coefficient matrix A by the matrix A itself.

The corresponding criterion was found by Routh[43] in 1877 and then modified by

---

[41] Aleksandr Mikhailovich Lyapunov, 1857–1918.
[42] Ernst Ferdinand Minding, 1806–1885.
[43] Edward John Routh, 1831–1907.

Hurwitz[44] in 1895. Usually, a matrix A whose eigenvalues satisfy condition (B.12) is called Hurwitz (or stable).

As was said above, stability theory as a rigorously developed branch of mathematics was created by Lyapunov.

• Aleksandr Mikhailovich Lyapunov was born in Yaroslavl' in 1857. He became a student of Faculty of Physics and Mathematics of the St. Petersburg University in 1876. First he wanted to study chemistry under supervision of D. I. Mendeleev, but soon moved to the mathematical department. In 1880, Lyapunov graduated with the degree of candidate (his supervisor was D. K. Bobylev). In 1885–1902, Lyapunov was professor of the Kharkov University, where he wrote his famous doctoral thesis "The General Problem of the Stability of Motion." This work contained the basic definitions of the modern stability theory as well as some fundamental results. In 1901, Lyapunov was elected Academician of the Russian Academy of Sciences. He moved to St. Petersburg in 1902. In the summer of 1917, Lyapunov's family moved to Odessa. On October 31, 1918, the day when his wife died, Lyapunov shot himself. He died three days later in hospital. •

We can mention a lot of deep and important results obtained by Lyapunov. He had developed a powerful method based on study of behavior of special functions along integral curves of a differential equation (later, such functions were called Lyapunov functions).

We will speak in detail about the branch of stability theory which originated from the famous theorem characterizing stability by the linear approximation (undoubtedly, this theorem is the mostly often used in applications result of theory of differential equations).

Usually, this theorem is formulated as follows. Consider a system of differential equations of the form

$$\frac{dx}{dt} = Ax + g(t, x), \tag{B.13}$$

where $g(t, 0) = 0$ for $t \geq t_0$ and

$$\frac{|g(t, x)|}{|x|} \to 0, \quad |x| \to 0, \quad \text{uniformly in } t \geq t_0.$$

If A is a Hurwitz matrix, then the zero solution of system (B.13) is asymptotically stable. If the matrix A has an eigenvalue λ with $\operatorname{Re}\lambda > 0$, then the zero solution of system (B.13) is unstable.

Thus, Lyapunov gave an exact answer to the question under which conditions the character of stability of the zero solution is the same for a nonlinear system of differential equations and for its linearization. Precisely this question was the main problem which Lyapunov wanted to solve creating stability theory. Let us cite his phrase from

[44] Adolf Hurwitz, 1859–1919.

the Preface to "The General Problem of the Stability of Motion:" "The problem which I had in mind starting this work can be formulated as follows: Indicate the cases where the first approximation solves the question of stability and develop methods to answer this question at least in some of the cases where the first approximation is not a criterion of stability."

The cases considered by the theorem on stability in the first approximation were called ordinary by Lyapunov. The remaining cases were called critical. Clearly, a critical case corresponds to a matrix A such that all the real parts of its eigenvalues are nonnegative and some of the eigenvalues have zero real parts.

In this case, we can represent the matrix A (after a nonsingular linear change of coordinates) in a block-diagonal form:

$$A = \operatorname{diag}(B, C), \tag{B.14}$$

where B is a matrix whose eigenvalues have zero real parts and C is a Hurwitz matrix.

Lyapunov gave complete answers for two critical cases: the case where the matrix B is the zero scalar and the case where B is a 2×2 matrix with purely imaginary eigenvalues (in the latter case, the answer was given up to the center-focus problem mentioned by Poincaré).

In 1963, V. I. Smirnov found in the archives of the Academy of Sciences of USSR and published the Lyapunov's manuscript "Study of a special case of the problem of the stability of motion." In this manuscript, Lyapunov studied the critical case where B is a 2×2 matrix with zero eigenvalues and the matrix C is nonzero. The study had not been completed; it remained to treat the case where B is a nontrivial Jordan block of size 2×2.

• Analyzing this text, one can see that Lyapunov had very high standards concerning publication of mathematical papers. The results obtained in the text (not perfect enough for publication, in Lyapunov's opinion) were really stronger than many results in this field published by various authors up to the beginning of the 1960s. •

The final solution in the critical case of a nontrivial Jordan block B of size 2×2 was given by V. A. Pliss[45] who applied a new approach using the so-called locally-invariant manifolds and the reduction principle (1964).

In the particular critical case of stability theory mentioned above, the main idea of the reduction principle can be formulated as follows.

Represent the phase variable x in the system

$$\frac{dx}{dt} = Ax + G(x), \quad G(x) = o(|x|), \tag{B.15}$$

which we consider as $x = (y, z)$ according to the block-diagonal form (B.14) of the matrix A. There exists a function $f(y)$ such that, in a neighborhood of the origin, the

[45] Victor Aleksandrovich Pliss, b. 1932.

manifold $z = f(y)$ is invariant for system (B.15), and the character of stability of the zero solution for the reduced system

$$\frac{dy}{dt} = By + G(y, f(y))$$

is the same as that for the initial system (B.15).

The reduction principle (now usually called the center manifold theorem – the title comes from the term "central" often applied to the locally-invariant manifold $z = f(y)$) is now one of the fundamental tools in qualitative theory of differential equations.

Many researchers continued to study the character of stability for systems of the form (B.15) in terms of their right-hand sides (usually, in terms of conditions on coefficients of the series representing the analytic nonlinear term $G(x)$).

In a sense, the logic of development of stability theory in this direction repeated the logic of development of the theory of solving differential equations by quadratures; it was shown that, in general, the stability problem is algebraically unsolvable.

Let us first give the necessary definitions.

Let f be a function of the class $C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ such that $f(0) = 0$. Fix a natural number N; the set

$$jf^{(N)} = \{g \in C^\infty(\mathbb{R}^n, \mathbb{R}^n) : g(0) = 0 \text{ and } |f(x) - g(x)| = o(|x|^N), |x| \to 0\}$$

is called the N-jet of the function f at 0.

Clearly, a function g belongs to $jf^{(N)}$ if and only if the Taylor expansions of the functions f and g at 0 coincide up to terms of degree N.

Let $J^{(N)}(n)$ be the space of N-jets of functions $C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ with the natural coordinates - to any N-jet we assign the coefficients of the vector polynomial of degree N that is determined by terms of degree not exceeding N in the Taylor expansion of a representative of the jet.

A subset of the space $J^{(N)}(n)$ is called semialgebraic if this subset is a union of a finite number of sets each of which is defined by a finite number of equations $P = 0$ and inequalities $Q > 0$, where P and Q are polynomials.

Consider a system of differential equations

$$\frac{dx}{dt} = f(x), \tag{B.16}$$

where $f \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ and $f(0) = 0$. We say that a jet $jf^{(N)}$ is stable (unstable) if the zero solution of the system

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

with any function $g \in jf^{(N)}$ is stable (respectively, unstable).

The sense of this terminology is as follows: If a jet $jf^{(N)}$ is stable (or unstable), then the character of stability of the zero solution of system (B.16) is determined by Taylor coefficients of the function f of degrees not exceeding N and does not depend on Taylor coefficients of higher degrees.

A jet is called neutral if it is neither stable nor unstable.

Consider, for example, system (B.15) and let $f(x) = Ax + G(x)$. Clearly, $jf^{(1)}$ is determined by entries of the matrix A.

The Lyapunov theorem on stability in the first approximation implies that if the matrix A is Hurwitz, then $jf^{(1)}$ is stable, and if A has an eigenvalue with a positive real part, then $jf^{(1)}$ is unstable.

It is relatively easy to show that if the matrix A corresponds to a critical case (i.e., all the real parts of its eigenvalues are nonnegative and some of the eigenvalues have zero real parts), then, by a proper choice of the nonlinearity $G(x)$, we can make the zero solution stable or unstable. This means that in this case, $jf^{(1)}$ is neutral.

The conditions under which a given matrix A is Hurwitz, has an eigenvalue with a positive real part, or corresponds to a critical case can be written in the form of a finite number of polynomial conditions on entries of the matrix A.

Thus, the sets of stable, unstable, and neutral 1-jets are semialgebraic for any dimension of the phase space.

It was shown by V. I. Arnold in 1970 that, in general, an analog of the above statement concerning 1-jets does not hold for N-jets with large N (which means that the stability problem is not algebraically solvable). We do not formulate the corresponding Arnold theorem (this would require more definitions); instead, we formulate one of results in this direction obtained later (see [58]).

Consider in the space $\mathbb{R}^4$ system (B.15) such that the matrix A has eigenvalues $\pm i, \pm 3i$. It is shown in [58] that the intersection of the subset of the space $J^{(3)}(4)$ corresponding to such systems with the set of stable jets is not semialgebraic.

This result means that we cannot impose a finite number of conditions having the form of polynomial equalities and inequalities on coefficients of terms of degree 2 and 3 in the representation of the function $G(x)$ to select the set of systems for which the zero solution is stable for any choice of terms of higher degrees.

We complete this section with description of some stability results for Hamiltonian systems (this class of systems of differential equations, very important for applications, was considered by Hamilton[46]).

First we consider the simplest Hamiltonian systems defined as follows. Introduce in the space $\mathbb{R}^{2n}$ of even dimension coordinates (p, q), where $p, q \in \mathbb{R}^n$. Fix a smooth function $H : \mathbb{R} \times \mathbb{R}^{2n} \to \mathbb{R}$ and assign to it the system of differential equations

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad i = 1, \dots, n. \tag{B.17}$$

[46] Sir William Rowen Hamilton, 1805–1865.

System (B.17) is called a Hamiltonian system with n degrees of freedom (and the function H is called the Hamiltonian of this system).

One of the main problems of the natural science of all times was to create a realistic model of the Solar System.

Undoubtedly, precisely this problem stimulated Newton when he formulated the basic laws of mechanics and the law of gravitation.

Newton considered as a model of the Solar System a system of n bodies moving in the space under the influence of mutual force fields. The problem of description of dynamics in such a system is called the n body problem.

The corresponding system of differential equations is Hamiltonian. Denote by x_i the vector of position of the ith body in the three-dimensional Euclidean space and by m_i the mass of this body.

Introduce coordinates q_i and impulses p_i by the formulas

$$q_i = x_i, \quad p_i = m_i \frac{dx_i}{dt}.$$

If F_i is the force applied to the ith body, then it follows form Newton's laws that

$$m_i \frac{d^2 x_i}{dt^2} = F_i.$$

Let us write these equations as a system:

$$\frac{dp_i}{dt} = m_i \frac{d^2 x_i}{dt^2} = F_i, \quad \frac{dq_i}{dt} = \frac{p_i}{m_i}. \tag{B.18}$$

It is easy to show (check!) that equalities (B.17) hold, i.e., (B.18) is a Hamiltonian system with Hamiltonian

$$H = \sum_{i=1}^{n} \frac{p_i^2}{2m_i} - \gamma \sum_{i \neq j} \frac{m_i m_j}{|q_j - q_i|},$$

where γ is the gravitational constant, and

$$F_i = \gamma \sum_{i \neq j} m_i m_j \frac{x_j - x_i}{|x_j - x_i|^3},$$

which means that the ith body is attracted by the other bodies according to Newton's gravitation law.

In fact, Kepler[47] knew that two bodies that attract each other move along conic sections.

[47] Johann Kepler, 1571–1630.

If $n = 3$, then the dimension of the corresponding system (B.18) equals 18, and we have no hope to give a complete description of its dynamics by tools of the modern mathematics (later, we mention several results on the three body problem).

Deep and interesting results concerning stability of Hamiltonian systems belong to A. N. Kolmogorov.[48]

Consider a Hamiltonian system of the form (B.17) in which coordinates q_i belong to the n-dimensional torus, and impulses p_i are points of a domain in the space $\mathbb{R}^n$. If the Hamiltonian H depends only on impulses, then the phase space is decomposed into invariant n-dimensional tori covered with quasiperiodic integral curves.

In 1954, Kolmogorov proved that if some nondegeneracy condition is satisfied, then the perturbed system with Hamiltonian

$$H(p) + \varepsilon H_1(p, q, \varepsilon)$$

also has invariant tori covered with dense integral curves, and the measure of the complement of the union of such invariant tori is small for small ε.

The Kolmogorov theory had been developed by Arnold and Moser[49]; now it is called Kolmogorov–Arnold–Moser (or KAM) theory.

Arnold is the author of one of the famous applications of KAM theory in stability theory. He considered the n body problem assuming that the mass of one of the bodies ("Sun") is very large compared to masses of the remaining bodies ("planets"). In the nonperturbed problem, the position of the "Sun" is fixed, and "planets" do not influence each other (in this case, the system is decomposed into independent Kepler subsystems).

It was shown by Lagrange that it is possible to average the perturbed system so that the averaged system has a stable equilibrium which corresponds to motion of all the "planets" in the same plane along circular orbits. The motion of "planets" corresponding to small oscillations in the linearization of the averaged system at this equilibrium is called the Lagrange motion.

Arnold proved that if the masses of the "planets" are small enough, then the phase space contains a domain of a large measure covered with trajectories that are close to Lagrange motions (which means the stability of the evolution of the planet system that starts at the mentioned domain).

B.5 Nonlocal qualitative theory. Dynamical systems

As was said above, the qualitative theory of differential equations studies properties of solutions of differential equations without finding the solutions themselves.

In a broad sense of this term, the qualitative theory includes existence and uniqueness theorems, local qualitative theory which studies the behavior of trajectories in

[48] Andrey Nikolaevich Kolmogorov, 1903–1987.
[49] Juergen Moser, 1928–1999.

a small neighborhood of a fixed trajectory (we do not mention such results in these lectures), stability theory, and so on.

In this section, we mostly speak about the development of nonlocal qualitative theory that describes the structure of the set of trajectories either in the whole phase space or in parts of the phase space that we cannot treat as small ones.

The nonlocal qualitative theory was founded by Poincaré in four books with the same title "Mémoire sur les courbes définiés par une équation différentielle" [59].

Beginning the first memoir, Poincaré emphasizes that his goal is to study solutions (functions defined by differential equations) by themselves and not to reduce them to simpler functions.

In these memoires, Poincaré classifies the main types of rest points for planar autonomous systems (nodes, saddles, foci, and centers), studies limit cycles, treats differential equations on the two-dimensional torus (he introduces the rotation number and relates it to the structure of trajectories), considers the center-focus problem.

He uses the method of contactless curves and the index theory which are applied later in various studies on the nonlocal qualitative theory. It is difficult to overestimate the influence of the ideas introduced by Poincaré (let us also mention that Poincaré created topology mainly as a tool devoted to study the structure of trajectories of multidimensional differential equations).

Later, we will speak about one more very important achievement of Poincaré – he had discovered the possibility of finding homoclinic points in problems of celestial mechanics.

Let us mention two deep results which develop the global qualitative theory created by Poincaré.

In 1901, Bendixson[50] studied a possible structure of the limit set of a positive semitrajectory of a planar autonomous system of differential equations in the case where the closure of this semitrajectory belongs to a planar domain containing only a finite set of rest points. His theorem (usually called the Poincaré–Bendixson theorem) states that the limit set has one of the following structures:

- a rest point;
- a closed trajectory;
- a contour consisting of rest points and trajectories that tend to these rest points as time goes to $\pm\infty$

(we refer to the Poincaré–Bendixson theorem in Section 7.6).

Denjoy[51] studied the structure of trajectories of a differential equation on the two-dimensional torus in the case where a meridian C of the torus is a contactless cycle and the Poincaré rotation number is irrational.

[50] Ivar Bendixson, 1861–1935.
[51] Arnaud Denjoy, 1884–1974.

It was shown by Poincaré that in this case, for any trajectory, the intersection P of its ω-limit set with the meridian C either coincides with C or is a perfect Cantor subset of C (and the set P does not depend on the choice of a trajectory).

Denjoy constructed an example of a differential equation of class C^1 for which the second possibility is realized. He also showed that if the equation is of class C^2, then $P = C$.

One of the most important problems of the qualitative theory of planar autonomous systems of differential equations is the Hilbert's 16th problem (to be more exact, the part of this problem related to differential equations): What may be said about number and location of limit cycles of a planar polynomial vector field with components of degree n?

This problem still stays unsolved.

Recall that a closed trajectory of a planar autonomous system is called a limit cycle if it has a neighborhood that does not contain other closed trajectories.

The first essential result concerning the 16th problem had been published by Dulac[52] in 1923.

The Dulac theorem [60] (sometimes called the finiteness theorem) says that a planar autonomous system of differential equations of the form

$$\frac{dx}{dt} = P(x, y), \quad \frac{dy}{dt} = Q(x, y) \tag{B.19}$$

(with polynomial or holomorphic right-hand sides P and Q) has a finite number of limit cycles.

In 1980, the Dulac book containing a complete proof of the finiteness theorem was published in Russian. A Moscow mathematician Yu. S. Ilyashenko[53] analyzed the proof and found that the proof contained a gap.

The history of the Hilbert 16th problem differs from the history of the 21st problem; after long efforts, Ilyashenko had invented a new method for proving the finiteness theorem [61]; independently, several other proofs had been published.

One of the modern formulations of the 16th problem is as follows: Does there exist a function $N(n)$ (Hilbert's number) defined for natural $n \geq 2$ such that the number of limit cycles of system (B.19) whose right-hand sides are polynomials of degree n does not exceed $N(n)$?

At present, it is not known whether $N(2)$ exists. Of course, it follows from the finiteness theorem that any given quadratic system (B.19) has a finite number of limit cycles; are these numbers uniformly bounded?

For a long time, it was believed that

$$N(2) = 3. \tag{B.20}$$

[52] Henry Dulac, 1870–1955.

[53] Yulij Sergeevich Ilyashenko, b. 1943.

In 1955, I. G. Petrovskii[54] and E. M. Landis[55] attempted to prove equality (B.20). Unfortunately, their proof contained errors.

A counterexample to equality (B.20) had been found in 1980. A Chinese mathematician Shi Songling[56] considered a quadratic system of the form (B.19) with

$$P(x, y) = -y + \lambda x - 10x^2 + (5 + \delta)xy + y^2,$$
$$Q(x, y) = x + x^2 - (25 + 9\delta - 8\varepsilon)xy,$$

and showed that this system has four limit cycles for $\delta = -10^{-13}, \varepsilon = -10^{-52}, \lambda = -10^{-200}$.

Thus, though systems (B.19) have a relatively simple form, the qualitative theory of such systems is very far from being complete.

The general theory of dynamical systems originates in Poincaré works.

A Russian translation of the Poincaré memoires [59] contained an appendix, "General Qualitative Theory," written by Russian mathematicians E. A. Leontovich[57] and A. G. Mayer.[58]

Leontovich and Mayer analyze the Bendixson's paper containing the proof of the Poincaré–Bendixson theorem. They show that the basic Poincaré results on the structure of trajectories of planar autonomous systems of differential equations (as well as Bendixson's results) follow from the fact that such systems generate flows (in the terminology accepted in this book).

Dynamical systems with discrete time were also first treated by Poincaré. We must mention his method of study of neighborhoods of closed trajectories (using the Poincaré transformation) and the so-called Poincaré's last geometric theorem on fixed points of measure preserving homeomorphisms of a planar annulus (this theorem has deep applications in the three body problem).

The modern theory of dynamical systems has its roots in the Birkhoff's[59] book [1] published in 1927.

Birkhoff defined nonwandering and recurrent trajectories, described the basic properties of the nonwandering set and the center of a dynamical system (we prove one of Birkhoff's theorems in Section 3.3), studied the property of transitivity, proved the ergodic theorem, generalized the above-mentioned Poincaré geometric theorem, studied existence and stability of periodic motions, proved the existence of a countable set of periodic points in a neighborhood of a homoclinic points of a two-dimensional diffeomorphism.

The further development of the global qualitative theory of dynamical systems was very intensive; of course, we cannot describe it in several lectures.

[54] Ivan Georgievich Petrovskii, 1901–1973.

[55] Evgeny Mikhailovich Landis, 1921–1987.

[56] Shi Songling, b. 1939.

[57] Evgeniya Aleksandrovna Andronova-Leontovich, 1905–1997.

[58] Artemii Grigoryevich Mayer, 1905–1951.

[59] George David Birkhoff, 1884–1944.

The completing sections of our lectures are devoted to the theory of structural stability and the theory of system with chaotic behavior (thus, we do not speak about such fields as the theories of asymptotic methods, of systems with invariant measure, of bifurcations, of complex systems).

B.6 Structural stability

Precisely 250 years passed after the publication of Newton's "Philosophiæ Naturalis Principia Mathematica" (1687) when A. A. Andronov[60] and L. S. Pontryagin[61] published in 1937 the paper [24].

The novelty of the Andronov and Pontryagin's approach can be formulated as follows: They suggested to study not a single system of differential equations but the preservation (or change) of the global structure of the set of trajectories when one considers a C^1-small perturbation of the right-hand side of the system.

Thus, a principally new object had been introduced – the space of systems of differential equations (or dynamical systems). The main problem was to select elements of this space to which there corresponds stability or instability of the global structure of trajectories. In this book, we mostly study this object (to be precise, we consider several spaces of dynamical systems with various topologies).

We have formulated the Andronov–Pontryagin theorem in Section 7.6.

It was mentioned that a structurally stable autonomous system of differential equations (in a two-dimensional disk or on a two-dimensional sphere) has a finite number of rest points and closed trajectories.

At the end of the 1950s, a young American mathematician S. Smale,[62] known for his proof of the Poincaré conjecture for $n > 4$, moved to the theory of dynamical systems. First he tried to extend the Andronov–Pontryagin theorem to the multidimensional case. As Smale wrote, he believed that a structurally stable diffeomorphism of a manifold of arbitrary dimension has a finite number of periodic points.

N. Levinson[63] attracted Smale's attention to the papers by M. Cartwright[64] and J. Littlewood[65] [62, 63] in which they studied the Van der Pol[66] equation

$$\frac{d^2x}{dt^2} + k(x^2 - 1)\frac{dx}{dt} + x = b\lambda k \cos(\lambda t) \tag{B.21}$$

with large parameter k and some generalizations of this equation (we mention the history of this equation in the next section).

[60] Aleksandr Aleksandrovich Andronov, 1901–1952.
[61] Lev Semenovich Pontryagin, 1908–1988.
[62] Stephen Smale, b. 1930.
[63] Norman Levinson, 1912–1975.
[64] Mary Cartwright, 1900–1998.
[65] John Edensor Littlewood, 1885–1977.
[66] Balthasar Van der Pol, 1889–1959.

Cartwright and Littlewood had shown that the Van der Pol equation has sets of solutions of a quite strange structure, and this structure is preserved under small perturbations of coefficients of the equation. In particular, equation (B.21) may have infinite families of periodic solutions with unbounded periods.

The proofs of Cartwright and Littlewood are very complicated. It was shown by Levinson in the paper [64] that one can discover a similar behavior of solutions in a simple analog of equation (B.21) in which the nonlinearity $x^2 - 1$ is replaced by a piecewise-constant function.

Analyzing the behavior of solutions of Levinson's equation, Smale invented his horseshoe (see Section 9). This gave the first example of a structurally stable diffeomorphism with an infinite set of periodic points [65].

It was conjectured by Smale that the hyperbolic torus automorphism considered by R. Thom[67] (see Section 7) is structurally stable, as well as a geodesic flow on a Riemannian manifold of negative curvature.

The first proof of structural stabilty of the hyperbolic torus automorphism had been published by V. I. Arnold and Ya. G. Sinai[68] in 1962, but later it was shown that the smoothness of some constructions used in the proof was not sufficient for application of the Denjoy theory. The structural stabilty of the hyperbolic torus automorphism had been established by Smale.

D. V. Anosov[69] proved a general result on structural stability of the so-called (Y)-systems (later called Anosov systems). As a corollary of this result, it was shown that a geodesic flow on a Riemannian manifold of negative curvature is structurally stable.

In the famous Anosov's book [27], he introduced condition (Y) (later called the hyperbolicity condition). It was shown that any (Y)-system (a flow or diffeomorphism) is structurally stable. Anosov mentioned that, in fact, condition (Y) had been applied by Hopf[70] in his study of ergodicity of geodesic flows.

In the same 1967, Smale published a large paper [66]. In this paper, he introduced Axioms A and A$'$, proved the spectral decomposition theorem and the first variant of the Ω-stability theorem (in this variant, he used a condition that was stronger than the no cycle condition). Later, Smale proved the Ω-stability theorem in its present form [67].

In fact, the paper [66] (and several later papers published by Smale himself and jointly with J. Palis[71]) contained a complete program of creation of theories of Ω-stability and structural stability. The main conjectures of these theories (for the case of diffeomorphisms) were stated by Smale and Palis [18] at the end of the 1960s (see Theorems 7.5 and 7.6 of this book).

---

[67] René Thom, 1923–2002.

[68] Yakov Grigor'evich Sinai, b. 1935.

[69] Dmitrii Victorovich Anosov, b. 1936.

[70] Eberhard Hopf, 1902–1983.

[71] Jacob Palis, b. 1940.

This program had been completed (both for diffeomorphisms and for flows generated by vector fields) 30 years later.

Let us note that one of the keystones of the theory of structural stability was the stable manifold theorem (in its present form, Theorem 7.1).

Anosov mentions that, in the analytic case, some variants of theorems on invariant manifolds were known to Darboux,[72] Poincaré, and Lyapunov.

Lyapunov studied in the book "The General Problem of the Stability of Motion" the so-called problem of conditional stability (i.e., stability with respect to perturbations of initial values belonging to some subsets of the whole space). Lyapunov gave conditions (in the analytic case) under which there exists a smooth manifold containing the initial point of the nonperturbed motion and such that motions that start at points of this manifold tend to the nonperturbed one exponentially as time grows. The Lyapunov method used the representation of the manifold by special series.

Some geometric ideas related to the existence of a stable manifold belong to P. Bohl.[73]

At present, two methods are used in the proof of the stable manifold theorem and its analogs: the method of J. Hadamard[74] and the method of O. Perron[75] (and the theorem itself is often called the Hadamard–Perron theorem).

The Hadamard method is based on transformation of graphs (one can find a detailed description of this method in the book [8]).

Perron applied a special method of successive approximations using the so-called Perron operator; this operator is a very powerful tool in study of various problems of the theory of differential equations and dynamical systems. In this book, we follow Perron's approach.

Let us also note the results of D. Grobman[76] and P. Hartman[77] who had proven that, in a neighborhood of a hyperbolic fixed point, a dynamical system is topologically conjugate to a linear mapping (see Sections 4 and 5).

Let us now pass to the history of the proof of Theorems 7.5 and 7.6. We have mentioned that the sufficiency part of Theorem 7.5 (Axiom A and the no cycle condition imply Ω-stability) had been proven by Smale.

The sufficiency part of Theorem 7.6 (Axiom A and the strong transversality condition imply structural stability) was first proven by J. Robbin[78] for diffeomorphisms of class C^2 [42] and then by C. Robinson[79] in [68] in the general case, for diffeomorphisms of class C^1.

[72] Gaston Darboux, 1842–1917.
[73] Pirs Bohl, 1865–1921.
[74] Jacques Salomon Hadamard, 1865–1963.
[75] Oskar Perron, 1880–1975.
[76] David Matveevich Grobman, b. 1922.
[77] Philip Hartman, b. 1915.
[78] Joel W. Robbin, b. 1941.
[79] Clark Robinson, b. 1943.

After that, almost for 20 years, mathematicians tried to prove the necessity of conditions of Theorems 7.5 and 7.6 (in fact, the main problem was to prove that structural stability implies the hyperbolicity of the nonwandering set). Finally, this was done by R. Mañé[80] in 1987 [44]. To the surprise of many specialists, the proof heavily used the theory of invariant measures. In Appendix A of this book, we describe a scheme of the Mañé proof.

Using practically the same reasoning, Palis had shown that conditions of Theorem 7.5 are necessary [69].

The case of flows generated by smooth vector fields is technically more complicated than the case of diffeomorphisms (and some of the differences are not purely technical). For that reason, in the case of flows, analogs of Theorems 7.5 and 7.6 were published later (complete proofs appeared in 1996).

An analog of Theorem 7.5 is as follows: a flow ϕ is Ω-stable if and only if ϕ satisfies Axiom A$'$ and the no cycle condition.

Sufficiency of these conditions for Ω-stability of a smooth flow had been established by C. Pugh[81] and M. Shub[82] in [70]; the necessity of these conditions follows from the results of L. Wen[83] and S. Hayashi[84] mentioned below.

The sufficiency part of an analog of Theorem 7.6 (a flow ϕ is structurally stable if and only if ϕ satisfies Axiom A$'$ and the strong transversality condition) had been proven by Robinson [71], and the necessity of these conditions had been established by Wen and Hayashi [72, 73].

B.7 Dynamical systems with chaotic behavior

In 1927, a Dutch physicist, mathematician, and engineer Balthasar Van der Pol (renowned not only for his theoretical achievements but also for the construction of the telegraph line that joined the Netherlands and its colonies in South-East Asia) studied oscillations in an electric circuit containing a vacuum tube and a periodic source of current. He changed the amplitude of the source and listened to the changing sound in a telephone receiver (the oscillograph had not been invented at that time).

Van der Pol noticed that a regular passage from one stationary frequency to another was sometimes replaced by a chaotic noise which did not vanish after a small change of amplitude.

This Van der Pol's discovery led to one of the most important steps in the development of science, the introduction of chaos as a model of natural phenomena.

As was mentioned above, Cartwright and Littlewood detected the appearance of a very strange structure of solutions in the Van der Pol equation (B.21) and proved that

[80]Ricardo Mañé, 1948–1995.
[81]Charles Chapman Pugh, b. 1940.
[82]Michael Shub, b. 1943.
[83]Lan Wen, b. 1946.
[84]Shuhei Hayashi, b. 1962.

this structure is preserved under small variation of parameters. After that, analyzing the behavior of solutions of the Levinson equation, a simplified variant of the Van der Pol equation, Smale had invented the horseshoe.

Studying the dynamics of the invariant set of a diffeomorphism with a horseshoe, Smale understood that a similar dynamics is generated by a transverse homoclinic point.

Homoclinic points were discovered by Poincaré (who also called them doubly asymptotic). Analyzing the dynamics generated by homoclinic points, Poincare wrote the following famous words in his fundamental book "Les Méthodes Nouvelles de la Mécanique Céleste" [74]: "If one attempts to imagine the figure formed by these two curves[85] and their infinitely many intersections, each of which corresponds to a doubly asymptotic solution, these intersections form something like a lattice, or fabric, or a net with infinitely tight loops. None of these loops can intersect itself, but it must wind around itself in a very complicated fashion in order to intersect all the other loops of the net infinitely many times. One is struck by the complexity of this figure, which I shall not even attempt to draw.

Nothing gives us a better idea of the complicated nature of the three body problem and the problems of dynamics in general."

It was also mentioned that Birkhoff knew that a homoclinic point of a two-dimensional diffeomorphism can produce an infinite set of periodic points.

The Smale theorem on the dynamics of a horseshoe (see Theorem 9.1) implies that any neighborhood of a transverse homoclinic points contains an invariant set on which the diffeomorphism is topologically conjugate to the shift homeomorphism; in particular, this means that any neighborhood of a transverse homoclinic point contains an infinite set of periodic points.

The problem of complete description of the dynamics of a system in a neighborhood of a transverse homoclinic point had been solved by Yu. I. Neimark[86] [50] and L. P. Shilnikov[87] [51].

Note that the problem of complete description of the dynamics of a system in a neighborhood of a nontransverse homoclinic point seems unsolvable.

At present, the mathematical literature contains many different definitions of chaos; in this book, we work with the definition cited in Section 9.2. As was shown in Section 9.3, a dynamical system having a transverse homoclinic point exhibits chaotic behavior.

Another famous system with chaotic behavior is the model studied by E. Lorenz.[88]

In 1963, almost simultaneously with Smale's invention of the horseshoe, Lorenz

[85] The curves are the stable and unstable manifolds of the hyperbolic saddle fixed point whose intersection creates the homoclinic point.

[86] Yurii Isaakovich Neimark, 1920–2011.

[87] Leonid Pavlovich Shilnikov, 1934–2011.

[88] Edward Norton Lorenz, 1917–2008.

published results of computer modeling of the system of differential equations

$$\dot{x} = -10x + 10y,$$
$$\dot{y} = 28x - y - xz,$$
$$\dot{z} = -8/3z + xy.$$

This system appears in modeling the convective motion of a liquid layer between two parallel planes [75]. Lorenz noted that numerically obtained trajectories with almost the same initial values diverge very significantly as time grows (Figure 7 shows a trajectory of the Lorenz system).

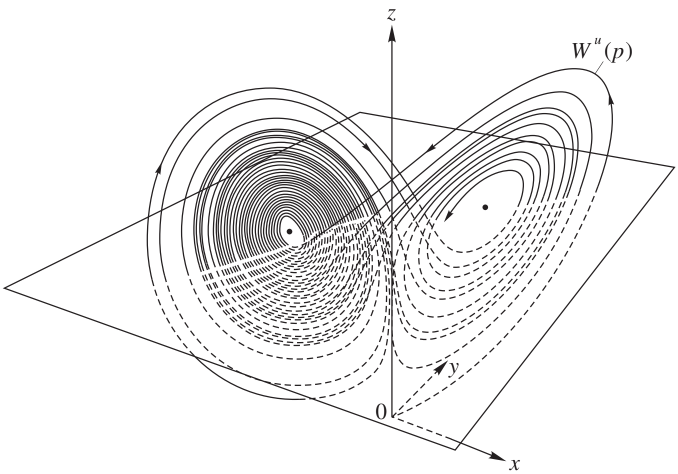


Figure 7. One trajectory of the Lorenz system.

In the study of chaotic dynamical systems similar to the Lorenz system, mathematicians invented the term "strange attractor."

A strange attractor is a limit set of a dynamical system on which the behavior of the system is chaotic. Now we know that strange attractors exist in models appearing in a lot of domains (mechanics, aerohydrodynamics, laser theory, population dynamics, etc.). In the study of such models, scientists apply the techniques of the theory of structural stability.

Though the right-hand sides of the Lorenz system are of a very simple structure, its complete investigation meets a lot of difficulties.

For a long time, the existence of chaotic structures was known not for the Lorenz system but for some of its geometric models.

Rigorous results were obtained using the so-called interval arithmetic, or interval analysis (which allows to estimate the errors of computer modeling and to obtain information concerning the geometry of images of sets under shifts along trajectories).

At the end of the 20th century, methods of interval arithmetic combined with the theory of Conley[89] index allowed K. Mischaikov[90] with a group of his colleagues to prove that the Lorenz system is chaotic.

In 1998, S. Smale published a list of problems that would prove challenging to mathematicians of the 21st century (in the spirit of Hilbert's problems).

In 2002, W. Tucker[91] has received the first R. E. Moore Prize for Applications of Interval Analysis. He had proven that the Lorenz system possesses a strange attractor (this was the contents of the 14th Smale problem).

Let us describe two more results related to chaotic dynamical systems.

One of them was published by A. N. Sharkovskii[92] [76] in 1964. Sharkovskii studied the problem of coexistence of periodic points of semi-dynamical systems (see Section 3.1) generated by continuous mappings of the segment [0, 1] into itself.

Let us order the natural numbers as follows:

$$\begin{aligned} 1 &\ll 2 \ll 2^2 \ll 2^3 \ll \cdots \ll 2^2 \cdot 7 \ll 2^2 \cdot 5 \ll 2^2 \cdot 3 \\ &\ll \cdots \ll 2 \cdot 7 \ll 2 \cdot 5 \ll 2 \cdot 3 \ll \cdots \ll 9 \ll 7 \ll 5 \ll 3 \end{aligned}$$

(thus, from the left we start with powers of two in increasing order, and from the right we start with the odd numbers in increasing order, then we put 2 times the odds, 4 times the odds, etc.). We agree that the relation $\ll$ is transitive, i.e., if $k \ll l \ll m$, then $k \ll m$.

By analogy with the case of a dynamical system, we say that p is a periodic point of f of period m if the points $p, f(p), \dots, f^{m-1}(p)$ are different and $f^m(p) = p$.

Sharkovskii showed that if a semi-dynamical system generated by a continuous mapping $f : [0, 1] \to [0, 1]$ has a periodic point of period m and $k \ll m$, then f has a periodic point of period k.

In addition, for any natural number m there exists a continuous mapping f that has a periodic point of period m and does not have periodic points of period n with $m \ll n$.

As often happens, the Sharkovskii theorem was unnoticed in the West, and only when a weaker result was published, the Sharkovskii ordering became generally acknowledged in the theory of dynamical systems.

In the end, we discuss the results of V. M. Alekseev[93] on the restricted three body problem published in 1968–69 [77].

Let us start with a degenerate three body problem in which two bodies of the same mass move in a plane periodically along ellipses with the same center C (this is one of the known Kepler configurations in the two body problem), while the third body of

[89] Charles C. Conley, 1933–1984.
[90] Konstantin Mischaikov, b. 1957.
[91] Warwick Tucker, b. 1970.
[92] Aleksandr Nikolaevich Sharkovskii, b. 1936.
[93] Vladimir Mikhailovich Alekseev, 1932–1980.

zero mass moves along the straight line that is perpendicular to the plane and intersects the plane at the point C.

The differential equation describing the motion of the third body is obtained from the standard equations of the three body problem in the limit as the mass of the third body tends to zero.

If $z(t)$ is the coordinate of the third body at time moment t (we assume that the origin of coordinates is fixed at the joint center of elliptic orbits of the massive bodies), then the differential equation for $z(t)$ has the form

$$\frac{d^2z}{dt^2} = -\frac{z}{(z^2 + r(t))^{3/2}}, \tag{B.22}$$

where $r(t)$ is the distance of any of the massive bodies to the center C. The right-hand side of equation (B.22) is ω-periodic in time, where ω is the joint period of motion of the massive bodies. Let T be the Poincaré transformation of the plane generated by the shift at time ω along integral curves of the second-order system corresponding to equation (B.22).

Alekseev showed that the diffeomorphism T generates a horseshoe (this horseshoe has a very specific spiral form); thus, the dynamics of equation (B.22) is chaotic.

It is very important to note that the same result (the existence of a horseshoe and chaotic dynamics) holds for the nondegenerate version of the restricted three body problem (if the mass of the third body is very small compared to the masses of the massive bodies).

Thus, chaotic dynamics is possible in the mostly classical problem of the theory of differential equations ascending to Newton.

Bibliography

[1] Birkhoff G, Dynamical systems. Colloquium Publ, 9. New York, Amer Math Soc, 1927.

[2] Nitecki Z, Differentiable dynamics. Cambridge, Massachusetts, and London, MIT Press, 1971.

[3] Guckenheimer J, Moser J, Newhouse S. Dynamical systems. Boston, MA, Birkhäuser-Verlag, 1980.

[4] Palis J, di Melo W. Geometric theory of dynamical systems. An introduction. Berlin, Heidelberg, Springer-Verlag, 1982.

[5] Shub M. Global stability of dynamical systems. Berlin, Heidelberg, Springer-Verlag, 1987.

[6] Robinson C. Dynamical systems. London, CRC Press, 1995.

[7] Brin M, Stuck G. Introduction to dynamical systems. Cambridge, Cambridge Univ Press, 2002.

[8] Katok A, Hasselblatt B. Introduction to the modern theory of dynamical systems. Cambridge, Cambridge Univ Press, 1995.

[9] Pilyugin SYu. Introduction to structurally stable systems of differential equations. Basel, Boston, Berlin, Birkhäuser-Verlag, 1992. Translation of Pilyugin SYu. Vvedeniye v grubye sistemy differentsial'nykh uravnenii. Leningrad, Leningrad Univ Press, 1988 (in Russian).

[10] Pilyugin SYu. The space of dynamical systems with the C^0 topology. Lect Notes Math, 1571. Berlin, Heidelberg, New York, Springer-Verlag, 1994.

[11] Pilyugin SYu. Shadowing in dynamical systems. Lect Notes Math, 1706. Berlin, Heidelberg, New York, Springer-Verlag, 1999.

[12] Abraham R, Robbin J. Transversal mappings and flows. New York, Benjamin Press, 1967.

[13] Brin MI. The inclusion of a diffeomorphism into a flow. Izv Vyssh Uchebn Zaved Mat 1972, 8, 19–25 (in Russian).

[14] Horn RA, Johnson CR. Matrix analysis. Cambridge, Cambridge Univ Press, 1990.

[15] Pliss VA. Integral sets of periodic systems of differential equations (Integral'nye mnozhestva periodicheskikh sistem differentsial'nylkh uravnenii). Moscow, Nauka, 1977 (in Russian).

[16] Fuks DB, Rokhlin VA. Beginner's course in topology: geometric chapters (Universitext). Berlin, Heidelberg, Springer-Verlag, 1984.

[17] Hirsch M, Palis J, Pugh C, Shub M. Neighborhoods of hyperbolic sets. Invent math 1970, 9, 133–163.

[18] Palis J, Smale S. Structural stability theorems. In: Global Analysis, Symp Pure Math, New York, Amer Math Soc, 1970, 14, 223–231.

[19] Smale S. Structurally stable systems are not dense. Amer J Math, 1966, 88, 491–496.

[20] Williams R. The "DA" maps of Smale and structural stability. In: Global Analysis, Symp Pure Math, New York, Amer Math Soc, 1970, 14, 329–334.

[21] Abraham R, Smale S. Non-genericity of Ω-stability. In: Global Analysis, Symp Pure Math, New York, Amer Math Soc, 1970, 14, 5–8.

[22] Aoki N. The set of Axiom A diffeomorphisms with no cycle. Bol Soc Brasil Mat (NS) 1992, 23, 21–65.

[23] Hayashi S. Diffeomorphisms in $\mathscr{F}^1(M)$ satisfy Axiom A. Ergod Theory Dyn Syst 1992, 12, 233–253.

[24] Andronov A, Pontryagin L. Systèmes grossiers. Comp Rend Acad Sci URSS 1937, 14, 247–250.

[25] De Baggis H. Dynamical systems with stable structure. In: Contrib Theory Nonlin Oscill, Princeton, NJ, 1952, 19–36.

[26] Hartman P. Ordinary differential equations. Cambridge, Cambridge Univ Press, 2002.

[27] Anosov DV. Geodesic flows on Riemann manifolds of negative curvature. Proc Steklov Inst Math, 90. Providence, RI, Amer Math Soc, 1967.

[28] Devaney RL. An introduction to chaotic dynamical systems. Reading, MA, Addison-Wesley, 1989.

[29] Glasner W, Weiss B. Sensitive dependence on initial conditions. Nonlinearity 1993, 6, 1067–1075.

[30] Pugh C. The closing lemma. Amer J Math 1967, 89, 956–1009.

[31] Pugh C, Robinson C. The C^1 closing lemma, including Hamiltonians. Erg Theory Dyn Syst 1983, 3, 261–313.

[32] Wen L. A uniform C^1 connecting lemma. Discrete Contin Dyn Syst 2002, 8, 257–265.

[33] Smale S. Mathematical problems for the next century. Math Intelligencer 1998, 20, 7–15.

[34] Mañé R. An ergodic closing lemma. Ann Math 1982, 116, 503–540.

[35] Takens F. On Zeeman's tolerance stability conjecture. In: Manifolds – Amsterdam 1970. Lect Notes Math, 197, Berlin, Heidelberg, New York, Springer-Verlag, 1971, 209–219.

[36] White W. On the tolerance stability conjecture. Salvador Symp Dyn Syst 1971, Univ of Bahia, 663–665.

[37] Hurley M. Attractors: persistence, and density of their basins. Trans Amer Math Soc 1982, 269, 247–271.

[38] Corless RM, Pilyugin SYu. Approximate and real trajectories for generic dynamical systems. J Math Anal Appl 1995, 189, 409–423.

[39] Pilyugin SYu. Inverse shadowing by continuous methods. Discrete Cont Dyn Syst 2002, 8, 29–38.

[40] Palmer K. Shadowing in dynamical systems. Theory and applications. Dordrecht, Kluwer, 2000.

[41] Al-Nayef A, Diamond P, Kloeden P, Kozyakin V, Pokrovskii A. Bi-shadowing and delay equations. Dynamics Stab Syst 1996, 11, 121–134.

[42] Robbin J. A structural stability theorem. Ann Math 1971, 94, 447–493.

[43] Goebel K, Kirk WA. Topics in metric fixed point theory. Cambridge, Cambridge Univ Press, 1990.

[44] Mañé R. A proof of the C^1 stability conjecture. Publ Math IHES 1987, 66, 161–210.

[45] Robinson C. C^r structural stability implies Kupka–Smale. Salvador Symp Dyn Syst 1971, Univ. of Bahia, 443–449.

[46] Pliss VA. A conjecture of Smale. Differents Uravn 1972, 8, 268–282 (in Russian).

[47] Pliss VA. Analysis of necessity of Smale and Robbin's roughness conditions for periodic systems of differential equations. Differents Uravn 1972, 8, 972–983 (in Russian).

[48] Mañé R. Expansive diffeomorphisms. In: Dynamical Systems – Warwick 1974, Lect Notes Math, 468, Berlin, Heidelberg, New York, Springer-Verlag, 1975, 162–174.

[49] Liao ST. On the stability conjecture. Chinese Ann Math 1980, 1, 9–30.

[50] Neimark YuI. On motions close to a double-asymptotic motion. Dokl Akad Nauk SSSR 1967, 172, 1021–1024 (in Russian).

[51] Shilnikov LP. On a Poincaré-Birkhoff problem. Mat Sb 1967, 74, 378–397 (in Russian).

[52] Vavilov SI. Isaac Newton (Isaac Newton). Moscow, Leningrad, Acad Sci USSR, 1945 (in Russian).

[53] Newton I. Philosophiae Naturalis Principia Mathematica, London, 1687.

[54] Lappo-Danilevskii IA. Application of functions of matrices to theory of linear differential equations (Primeneniye funktsii ot matrits k teorii lineinykh differentsial'nykh uravnenii). Moscow, Gos Izd Tekhn Lit, 1957 (in Russian).

[55] Hilbert problems (Problemy Hilberta). Moscow, Nauka, 1969 (in Russian).

[56] Moiseev ND. Essays on development of stability theory (Ocherki razvitiya teorii ustoichivosti). Moscow, Leningrad, Gos Izd Tekhn Lit, 1949 (in Russian).

[57] (Euler L.) Scientia navalis seu tractatus de construendis ac dirigendis navibus, auctore L. Eulero. Petropoli, 1749.

[58] Shnol' EE, Khazin LG. Stability of equilibria in critical cases and in cases close to critical ones. Prikl Mat Mekh 1981, 45, 595–604 (in Russian).

[59] Poincaré H. Mémoire sur les courbes définies par les équations differentiélles, I–IV. J Math Pures Appl, 3 série, 1881, 7, 375–422; 1882, 8, 251–286; 4 série, 1885, 1, 167–244; 1886, 2, 151–217.

[60] Dulac H. Sur les cycles limite. Bull Soc Math France 1923, 51, 45–188.

[61] Ilyashenko Yu. Finiteness theorem for limit cycles. Providence, RI, Amer Math Soc, 1991.

[62] Cartwright ML, Littlewood JE. On nonlinear equations of the second order: I. The equation $\ddot{y} - k(1 - y^2\dot{y})\dot{y} + y = b\lambda k \cos(\lambda t + \alpha), k$ large. J London Math Soc 1945, 20, 180–189.

[63] Cartwright ML, Littlewood JE. On nonlinear equations of the second order: II. The equation $\ddot{y} + kf(y\dot{y} + g(y,k) = p(t) = p_1(t) + kp_2(t)$. Ann Math 1947, 48, 474–494; 1949, 50, 504–505.

[64] Levinson N. A second order differential equation with singular solutions. Ann Math 1949, 50, 126–153.

[65] Smale S. Structurally stable diffeomorphisms with infinitely many periodic points. In: Proc Intern Conf Nonl Oscill, 2, Kiev, 1963, 365–366.

[66] Smale S. Differentiable dynamical systems. Bull Amer Math Soc 1967, 73, 747–817.

[67] Smale S. The Ω-stability theorem. In: Global Analysis, Symp Pure Math, New York, Amer Math Soc, 1970, 14, 289–297.

[68] Robinson C. Structural stability of C^1 diffeomorphisms. J Diff Equat 1976, 22, 28–73.

[69] Palis J. On the C^1 Ω-stability conjecture. Publ Math IHES 1987, 66, 211–215.

[70] Pugh C, Shub M. Ω-stability theorem for vector fields. Invent math 1970, 11, 150–158.

[71] Robinson C. Structural stability of vector fields. Ann Math 1974, 99, 154–175.

[72] Wen L. On the C^1 stability conjecture for flows. J Diff Equat 1996, 129, 334–357.

[73] Hayashi S. Connecting invariant manifolds and the solution of the C^1 stability and Ω-stability conjecture. Ann Math 1997, 145, 81–137.

[74] Poincaré H. Les méthodés nouvelles de la mécanique céleste. Paris, 1892–1899.

[75] Lorenz EN. Deterministic non-periodic flow. J Atmos Sci 1963, 20, 130–141.

[76] Sharkovskii AN. Coexistence of cycles of a continuous map of the line into itself. Ukrain Mat Zh 1964, 16, 61–71 (in Russian).

[77] Alekseev VM. Quasirandom dynamical systems I, II, III. Mat Sb 1968, 76, 72–134; 1968, 77, 545–601; 1969, 78, 3–50 (in Russian).

Index

action of a group 15
Andronov-Pontryagin theorem 114
Anosov diffeomorphism 119
attracting fixed point 40, 48
attracting rest point 71
attractive set 148
attractor 148
Axiom A 103
Axiom A′ 112

base of neighborhoods of a point 4
basic set 104
basin 148
Birkhoff constant 33
Birkhoff theorem 33

(C, λ)-structure 160
C^0 topology 16
C^1 topology 18
cascade 1
chaotic invariant set 131
closed trajectory 8
Closing Lemma 135
coding 130
conjugating homeomorphism 26

d-continuity 142
d-pseudotrajectory 159
d-method 160
deviation 141
dominated splitting 182
double-asymptotic point 82
dynamical system with discrete time 1
dynamical system with continuous time 7

embedding of a diffeomorphism into a flow 11
embedding of class C^k 25
ergodic closing lemma 139
ergodic invariant measure 184
expansivity 171

first axiom of countability 31
fixed point 2
flow 7
fundamental neighborhood 51

generic element 24
generic property 24
Grobman-Hartman theorem 41

Hausdorff distance 141
homoclinic point 82
horseshoe 126
hyperbolic closed trajectory 73
hyperbolic fixed point 40
hyperbolic linear mapping 37
hyperbolic matrix 37
hyperbolic periodic point 65
hyperbolic rest point 67
hyperbolic set of a diffeomorphism 92
hyperbolic set of a flow 112
hyperbolic structure 92
hyperbolic toral automorphism 95
hyperbolicity constants 92

immersion of class C^k 25
integral curve 8
invariant probability measure 139
invariant set 3
inverse shadowing property 160

k-cycle 110
Kupka-Smale diffeomorphism 82
Kupka-Smale flow 82

λ-lemma 83
Lipschitz inverse shadowing property 160
Lipschitz shadowing property 160
local Poincaré diffeomorphism 13
local stable manifold of a hyperbolic closed trajectory 75
local stable manifold of a hyperbolic fixed point 48

local stable manifold of a hyperbolic rest point 72
local topological conjugacy 36
local unstable manifold of a hyperbolic closed trajectory 75
local unstable manifold of a hyperbolic fixed point 49
local unstable manifold of a hyperbolic rest point 72
lower semicontinuous mapping 142
Lyapunov norm 164
Lyapunov stable set 147

maximal ϵ-equivalence 145
minimal ϵ-equivalence 145
Morse-Smale diffeomorphism 111
multipliers of a closed trajectory 73

negative semitrajectory 2
no cycle condition 110
nonwandering point 31
nonwandering set 31

Ω-conjugacy 34
Ω-stability 35
Ω-stability theorem 110
1-cycle 110
orbital ϵ-equivalence 144

Perron operator 55
phase space 1
periodic point 2
Poincaré diffeomorphism 14
positive semitrajectory 2

quasiattractor 157

repelling fixed point 40, 48
repelling rest point 72
residual set 24
rest point 8
rough system 114

saddle fixed point 40, 48
saddle rest point 72
semi-dynamical system 27
sensitive dependence on initial conditions 131
separatrix joining saddles 114
shadowing property 159
shift homeomorphism 6
Smale theorem on horseshoe 128
smooth disk 25
spectral decomposition theorem 104
stability of an attractor with respect to the Hausdorff metric 151
stability of an attractor with respect to the metric R_0 151
stable manifold of a hyperbolic closed trajectory 75
stable manifold of a hyperbolic fixed point 49
stable manifold of a hyperbolic rest point 72
stable manifold theorem for a hyperbolic fixed point 53
stable manifold theorem for a hyperbolic set 102
stable separatrix 114
standard multiplier 74
strong transversality condition 110
structural stability 28
structural stability in the strong sense 29
structural stability theorem 110
structurally stable diffeomorphism 28
structurally stable flow 30

Takens theorem 145
theorem on structural stability of an Anosov diffeomorphism 119
tolerance D-stability 144
topological conjugacy 26
topological immersion 25
topologically conjugate flows 29
topologically conjugate homeomorphisms 26
topologically conjugate semi-dynamical systems 27
topologically equivalent flows 30
topologically mixing homeomorphism 131
trajectory 2
trajectory of a point in a flow 7
trajectory of a solution 8

trajectory (orbit) of a point under the action of a group 15
transversality of mappings 78
transversality of stable and unstable manifolds 80
transversality of submanifolds 78
transverse homoclinic point 133

uniform connecting lemma 135
unstable manifold of a hyperbolic closed trajectory 75
unstable manifold of a hyperbolic fixed point 49
unstable manifold of a hyperbolic rest point 72
unstable separatrix 114
upper semicontinuous mapping 142

wandering point 31