

# HANDBOOK OF DYNAMICAL SYSTEMS

VOLUME 3

*Edited by*  
*B. Hasselblatt*  
*H. W. Broer*  
*F. Takens*



North-Holland is an imprint of Elsevier  
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK  
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands

Copyright © 2010 Elsevier B.V. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: [permissions@elsevier.com](mailto:permissions@elsevier.com). Alternatively you can submit your request online by visiting the Elsevier web site at <http://elsevier.com/locate/permissions>, and selecting *Obtaining permission to use Elsevier material*

#### Notice

No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

#### Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

#### British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13: 978-0-444-53141-4

|                                                                                                                                       |
|---------------------------------------------------------------------------------------------------------------------------------------|
| For information on all North-Holland publications<br>visit our web site at <a href="http://books.elsevier.com">books.elsevier.com</a> |
|---------------------------------------------------------------------------------------------------------------------------------------|

Printed and bound in USA

10 11 12 10 9 8 7 6 5 4 3 2 1

Working together to grow  
libraries in developing countries

[www.elsevier.com](http://www.elsevier.com) | [www.bookaid.org](http://www.bookaid.org) | [www.sabre.org](http://www.sabre.org)

ELSEVIER

BOOK AID  
International

Sabre Foundation

## Preface

The discipline of Nonlinear Dynamical Systems has expanded tremendously over the past five decades and the last couple of years several surveys and encyclopaedia have seen the light. The present book is the third volume of a Handbook that covers a wide range of subjects in the area. The chapters are written in such a way that the handbook forms an extensive entrance to the literature.

The present volume focuses on the conceptual, and often geometric, aspects of the theory in the spirit of Poincaré and Lyapunov and later of Andronov, Thom and Smale. Important notions of dynamical systems often are framed in terms of persistence under variation of initial positions and of system parameters. A particular case of this is that of structural stability under various equivalence relations like topological conjugacy or equivalence, amounting to the stable system being an interior point of its equivalence class. Often structural stability is a far too restrictive notion of persistence. A broader notion is provided by the idea of a generic property, which is associated with a residual set of systems. Often genericity is associated somehow to the notion of transversality on a suitable jet space. To express this we also need a suitable topology on the space of systems at hand, preferably possessing the Baire property. In a preliminary chapter the authors Broer and Takens elaborate these ideas, also taking into account structure-preserving cases like, for instance, the symplectic or volume-preserving setting; this is meant to provide a suitable background for the rest of the book. In the second chapter Hunt and Kaloshin deal with the notion of prevalence, which forms an alternative to that of genericity. Another collection of background tools is next introduced by Takens and Vanderbauwhede who wrote the third chapter on local invariant manifolds and normal forms. In the fourth chapter Devaney deals with iterations on the complex plane, in particular with complex exponential dynamics. [Chapter 5](#), by Levi, deals with some applications of Moser's Twist Theorem, the 2-dimensional instance of KAM Theory. This theory gener(ic)ally concerns quasi-periodicity in dynamical systems and is surveyed in a sixth chapter by Broer and Sevryuk. [Chapter 7](#), by Takens, gives an account of reconstruction theory for nonlinear time series. The volume concludes with [Chapter 8](#) by Homburg and Sandstede which deals with homoclinic and heteroclinic bifurcations for vector fields.

We are well aware of the fact that these volumes only cover a part of the discipline of Nonlinear Dynamical Systems as it stands today and that the chapters in the present volume do not nearly cover everything left out in the preceding volumes. We nonetheless feel that this volume fits well into the Handbook of Dynamical Systems, which accounts for a fair amount of the discipline and aims to be a valuable guide for further reading.

Henk Broer, Boris Hasselblatt, and Floris Takens

## List of Contributors

- Broer, H.W., *Johann Bernoulli Institute for Mathematics and Computer Science, P.O. Box 407 9700 AK Groningen, The Netherlands; Bernoulliborg, Building 5161, Nijenborgh 9, 9747 AG Groningen, The Netherlands* (Ch. 1, 6)
- Devaney, R.L., *Department of Mathematics & Statistics, Boston University, 111 Cummington St., Boston, MA 02215 USA, 617-353-4560* (Ch. 4)
- Homburg, A.J., *Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands* (Ch. 8)
- Hunt, B.R., *Department of Mathematics, University of Maryland, College Park, MD 20742, USA* (Ch. 2)
- Kaloshin, V. Yu., *Department of Mathematics, University of Maryland, College Park, MD, 20740, United States; Department of Mathematics, Penn State University, University Park, PA, 16801, United States* (Ch. 2)
- Levi, M., *Department of Mathematics, Pennsylvania State University, 218 McAllister Building, University Park, PA 16802, USA* (Ch. 5)
- Sandstede, B., *Division of Applied Mathematics, Brown University, Providence, RI 02906, USA* (Ch. 8)
- Sevryuk, M.B., *Institute of Energy Problems of Chemical Physics, The Russia Academy of Sciences, Leninskiĭ; prospect 38, Bldg. 2, Moscow 119334, Russia* (Ch. 6)
- Takens, F., *Johann Bernoulli Institute for Mathematics and Computer Science, P.O. Box 407 9700 AK Groningen, The Netherlands; Bernoulliborg, Building 5161, Nijenborgh 9, 9747 AG Groningen, The Netherlands* (Ch. 1, 3, 7)
- Vanderbauwhede, A., *Department of Mathematics, Ghent University, Krijgslaan 281, B-9000 Gent, Belgium* (Ch. 3)

# Contents

|                                                                                                      |     |
|------------------------------------------------------------------------------------------------------|-----|
| <i>Preface</i>                                                                                       | vii |
| <i>List of Contributors</i>                                                                          | ix  |
| 1. Preliminaries of Dynamical Systems Theory<br><i>H.W. Broer and F. Takens</i>                      | 1   |
| 2. Prevalence<br><i>B.R. Hunt and V. Yu. Kaloshin</i>                                                | 43  |
| 3. Local Invariant Manifolds and Normal Forms<br><i>F. Takens and A. Vanderbauwhede</i>              | 89  |
| 4. Complex Exponential Dynamics<br><i>R.L. Devaney</i>                                               | 125 |
| 5. Some Applications of Moser's Twist Theorem<br><i>M. Levi</i>                                      | 225 |
| 6. KAM Theory: Quasi-periodicity in Dynamical Systems<br><i>H.W. Broer and M.B. Sevryuk</i>          | 249 |
| 7. Reconstruction Theory and Nonlinear Time Series Analysis<br><i>F. Takens</i>                      | 345 |
| 8. Homoclinic and Heteroclinic Bifurcations in Vector Fields<br><i>A.J. Homburg and B. Sandstede</i> | 379 |
| <i>Author Index</i>                                                                                  | 525 |
| <i>Subject Index</i>                                                                                 | 539 |

## CHAPTER 1

# Preliminaries of Dynamical Systems Theory

H.W. Broer and F. Takens

*Johann Bernoulli Institute for Mathematics and Computer Science,  
P.O. Box 407 9700 AK Groningen, The Netherlands  
Bernoulliborg, Building 5161, Nijenborgh 9, 9747 AG Groningen, The Netherlands*

### *Contents*

|                                                                                                       |    |
|-------------------------------------------------------------------------------------------------------|----|
| 1. General definition of a dynamical system . . . . .                                                 | 3  |
| 1.1. The state space . . . . .                                                                        | 4  |
| 1.2. The time set . . . . .                                                                           | 5  |
| 1.3. The evolution map . . . . .                                                                      | 7  |
| 1.4. Relations between the different classes of dynamical systems . . . . .                           | 11 |
| 2. Transversality and generic properties . . . . .                                                    | 13 |
| 2.1. The notion of genericity . . . . .                                                               | 13 |
| 2.2. Sard's theorem and Thom's transversality lemma . . . . .                                         | 14 |
| 2.3. Generic properties of dynamical systems based on transversality (Kupka-Smale theorems) . . . . . | 19 |
| 3. Generic properties which are not based on transversality: the Closing Lemma . . . . .              | 26 |
| 4. Generic local bifurcations . . . . .                                                               | 28 |
| 4.1. Generic local bifurcations of functions (Catatrophe Theory) . . . . .                            | 28 |
| 4.2. Centre manifolds and generic local bifurcations of differential equations . . . . .              | 30 |
| 4.3. Centre manifolds and local bifurcations of fixed points of diffeomorphisms . . . . .             | 34 |
| 4.4. Concluding remarks . . . . .                                                                     | 37 |
| 5. Structural stability and moduli . . . . .                                                          | 38 |
| References . . . . .                                                                                  | 40 |



In this volume we present a collection of surveys on various aspects of the theory of bifurcations of differentiable dynamical systems and related topics. In the choice of the subjects we have tried to focus on those developments in which we think research will be active in the coming years. The surveys are intended to orientate the reader towards the recent literature in these subjects.

The purpose of this chapter is the presentation of preliminary material: basic definitions (in Section 1) and brief compilations of results, which are not any longer central themes of ongoing research, but which are still important as tools. As such we discuss in Section 2 transversality and generic properties like the various forms of the so-called Kupka-Smale theorem, in Section 3 the Closing Lemma and in Section 4 generic local bifurcations of functions (so-called catastrophe theory) and generic local bifurcations in 1-parameter families of dynamical systems. Finally, in Section 5 we discuss the notions of structural stability and moduli.

We should point out also that the various volumes of the Encyclopaedia of Mathematical Sciences (English translation of the original Russian edition) dealing with dynamical systems contain much valuable information on these topics. For the subjects discussed in this chapter, we mention in particular the volumes 6 and 39 (Dynamical Systems VI, respectively, VIII) [6], dealing with singularity theory and its applications, and volume 5 (Dynamical Systems V) [5], dealing with bifurcations and catastrophe theory. These volumes are recommended as a reference for the theory as developed up to the mid-1980s.

Finally we mention also the surveys in [13] on the various topics discussed here, though they are on a more elementary level. For general reference also see the textbook [16] and its guided bibliography.

## 1. General definition of a dynamical system

We first give an abstract and general definition of a dynamical system, which will then be given a more concrete content in the form of a number of specific classes of dynamical systems.

In this general sense, a dynamical system consists of a *state space*  $X$ , which is a set, a *time set*  $T \subset \mathbb{R}$ , which is an additive semi-group, i.e.  $0 \in T$  and for  $t_1, t_2 \in T$  also  $t_1 + t_2 \in T$ , and an *evolution map*  $\Phi : X \times T \rightarrow X$  having the (semi-)group property, i.e. satisfying  $\Phi(\Phi(x, t_1), t_2) = \Phi(x, (t_1 + t_2))$  and  $\Phi(x, 0) = x$ .

These structures are used to construct mathematical models for the deterministic evolution in time of various systems. The state space  $X$  is to be considered as the set of all possible states of the system – the notion of state should be interpreted in such a way that it contains all the information which is relevant for the prediction of the future of the system. The time set  $T$  consists of those real numbers  $t$  for which it is possible, from knowing the state  $x(\tau)$  at time  $\tau$ , to determine the state  $x(\tau + t)$  at time  $(\tau + t)$ . The evolution map assigns to  $x(\tau)$  and  $t$  the state  $x(\tau + t) = \Phi(x(\tau), t)$  at time  $(\tau + t)$ . So possible *evolutions*, or *orbits*, of a dynamical system, in terms of its evolution map  $\Phi$ , are of the form  $T \ni t \mapsto \Phi(x_0, t) \in X$ ;  $x_0 \in X$  is called the *initial state*.



### 1.1. The state space

In our considerations, state spaces always have some extra structure: at least a topological structure, possibly with a Borel (probability) measure or a differentiable structure. The most important case is where the state space  $X$  is a finite dimensional manifold, with a finite dimensional vector space as an important special case. We note here that we shall assume all manifolds to be sufficiently smooth, except if there is an explicit statement about the differentiability of a particular (sub)manifold. This type of state space is used whenever the state of a system can be specified by finitely many real numbers. Important examples of this are:

- chemical reactions, where the state, at a given moment, is specified by the concentrations of the various chemical substances, at that moment;
- electrical circuits, where the state (at a given moment) is specified by the voltages at the various nodes and the currents in the various branches (at that moment) – especially in the case of nonlinear circuits – the set of *realizable* states, i.e. the states which are compatible with the Kirchhoff laws and the resistor characteristics, may form a (nonlinear) sub-manifold of the vector space of all possible voltages and currents;
- simple mechanical systems (simple in the sense that they contain only a finite number of point masses and rigid bodies, apart from springs etc. which create interaction forces), where the state is specified by the various positions and velocities of the point masses and rigid bodies involved – also, here, the state space will in general be a manifold and not just a vector space: for the specification of the position of a rigid body we not only need three space coordinates (as for a point mass) but also its orientation, given by an element of the special orthogonal group  $SO(3)$ .

Also (infinite dimensional) function spaces often are a natural choice as state spaces in models for time evolution: they appear where systems are described by partial differential equations, such as the dynamics of

- fluids or gases (Navier-Stokes-equation), where the state space consists of velocity fields;
- heat (heat equation), where the state space consists of functions which specify the temperature as a function of position;
- electromagnetic fields (Maxwell-equations), where an element of the state space specifies the electromagnetic field.

In this volume we mainly restrict ourselves to systems with finite dimensional state spaces. We note however that it is often possible to reduce systems with infinite dimensional state spaces to ones with finite dimensional state spaces. This is in particular the case for dissipative systems like the ones describing fluid/gas dynamics or heat conduction, e.g. see [23].

There are various other types of state space, like sequence spaces (for symbolic dynamics) or manifolds with extra structure, e.g. symplectic manifolds. We come back to this when discussing evolution maps.

## 1.2. The time set

The restrictions imposed on the time set leave only a few interesting possibilities which we list:

1.  $T = \mathbb{R}$ ;
2.  $T = \mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$ ;
3.  $a\mathbb{Z} = \{an | n \in \mathbb{Z}\}$  for some  $a > 0$ ;
4.  $a\mathbb{Z}_+ = \{an | 0 \leq n \in \mathbb{Z}\}$  for some  $a > 0$ .

Without loss of generality we may assume that the number  $a$  in the items 3. and 4. is equal to one, in which case we denote these time sets by  $\mathbb{Z}$ , respectively  $\mathbb{Z}_+$ . These time sets correspond to classes of dynamical systems which we discuss below.

**1.2.1. Time set  $\mathbb{R}$ .** Dynamical systems with this time set are typically defined by differential equations, at least in the case of a finite dimensional manifold as a state space and a differentiable evolution map. The *differential equation* corresponding to the dynamical system with evolution map  $\Phi : X \times \mathbb{R} \rightarrow X$  is then  $x' = f(x)$ , where  $f$  is the *vector field* on  $X$  defined by  $f(x) = \partial_t \Phi(x, 0)$ . The evolution map can then be considered as the *general solution* of this differential equation: for each  $x \in X$ ,  $t \mapsto \Phi(x, t)$  is the (unique) solution of the differential equation with initial state  $x$ . Note that we often use vector fields and (first-order autonomous) differential equations as synonyms, though formally a vector field is just the ‘right-hand side’ of such a differential equation.

Apart from some technical conditions, like  $\Phi$  being differentiable and the solutions of  $x' = f(x)$  being defined for all  $t \in \mathbb{R}$  and solutions, for a given initial state, being unique, there is a one-to-one correspondence between differential equations and dynamical systems with time set  $\mathbb{R}$ , at least in the case of a finite dimensional manifold as state space.

We define the time  $t$  map as  $\Phi^t(x) = \Phi(x, t)$ . For dynamical systems given by a differential equation  $x' = f(x)$  with  $f$  smooth, it is a time dependent *diffeomorphism*, i.e. a differentiable map with a differentiable inverse: the inverse of  $\Phi^t$  is  $\Phi^{-t}$ . The map  $\mathbb{R} \ni t \mapsto \Phi^t$  is called the *flow* of  $f$ . Sometimes we use the term *flow* also to refer to a dynamical system with time set  $\mathbb{R}$ .

For differential equations which have no unicity of solutions (for some initial states) these maps  $\Phi^t$  cannot be defined; however, if the ‘right-hand side’  $f$  of the differential equation is  $C^1$  (or locally Lipschitz) there is unicity of solutions. If there is unicity of solutions, but if some solutions disappear to ‘infinity’ for finite  $t$ , i.e. if there are solutions whose domain of definition cannot be extended to all of  $\mathbb{R}$ , then the maps  $\Phi^t$  are only partially defined. Even in this case the maps  $\Phi^t$  have locally a differentiable inverse, i.e. they are local diffeomorphisms. Also, where they are defined, they have the group property, i.e.  $\Phi^{t_1}(\Phi^{t_2}(x)) = \Phi^{t_1+t_2}(x)$ , wherever defined. We sometimes call such a system with an only partially defined evolution map a *local dynamical system*. They occur e.g. on so-called centre manifolds, to be introduced in Section 4.2.1. If the state space is a *compact* manifold, then no solutions can disappear to infinity and the maps  $\Phi^t$  are globally defined (for all  $t$ ).

Also some partial differential equations define dynamical systems with time set  $\mathbb{R}$ . In that case the state space is an infinite dimensional function space. The partial differential equation should then be of a type which can also be solved backwards in time. An

example of this is the system of equations for electromagnetic fields in vacuum (the Maxwell equations).

**1.2.2. Time set  $\mathbb{R}_+$ .** For each dynamical system with time set  $\mathbb{R}$ , we can trivially obtain a system with time set  $\mathbb{R}_+$  by simply restricting the evolution map. The most important non-trivial examples where the time set is  $\mathbb{R}_+$ , are dynamical systems defined by partial differential equations which can only be solved in the positive time direction, like the heat equation and other parabolic equations. Since we are mainly concerned with dynamical systems whose state space is finite dimensional, this time set is not of much interest to us.

**1.2.3. Time set  $\mathbb{Z}$ .** In this case the evolution map  $\Phi$  is completely determined by the map  $\varphi : X \rightarrow X$ , defined by  $\varphi(x) = \Phi(x, 1)$ : we have  $\Phi(x, n) = \varphi^n(x)$  for all  $n \in \mathbb{Z}$ , so that the evolution map consists of all iterations of  $\varphi$ . This state of affairs makes the above notation  $\Phi^t$  less useful. The map  $\varphi$  is invertible: its inverse is given by  $\varphi^{-1}(x) = \Phi(x, -1)$ . Conversely, any invertible map, or *automorphism*  $\varphi : X \rightarrow X$  determines a dynamical system with time set  $\mathbb{Z}$ . In the case of a differentiable dynamical system, the evolution map is determined by a *diffeomorphism*  $\varphi$ .

Though the usual mathematical model for the ‘time set’ is the real line, there are many situations where a representation of the time by integers is more to the point. An important class of examples is that of periodically driven systems: One chooses a fixed phase of the forcing. Then the state of the system is only recorded at the times where the forcing is in that particular phase. The map describing the dynamics assigns to each state the state which is reached after one period of the forcing (starting at the selected phase). Maps which are obtained from dynamical systems with periodic forcing in the above way are called *period maps* or *stroboscopic maps*. A closely related notion is that of a *Poincaré map*, see Section 1.4.2.

Also for this time set one sometimes deals with dynamical systems with an evolution map which is not defined on all of  $X \times \mathbb{Z}$ . This happens in particular in the case of Poincaré maps of periodic orbits, see Section 2.3.1. Also in this case we speak of *local dynamical systems*.

**1.2.4. Time set  $\mathbb{Z}_+$ .** These systems are just defined by any possibly non-invertible map, or *endomorphism*,  $\varphi : X \rightarrow X$ . The corresponding evolution map is again given by  $\Phi(x, n) = \varphi^n(x)$ , so  $\Phi$  ‘consists of’ the iterations of  $\varphi$ . The main difference between these systems and the previous case is that one cannot reconstruct the past from the present (compare the case with time set  $\mathbb{R}_+$ ). An important class of examples of such dynamical systems is formed by algorithms which are based on iterative procedures to obtain better approximations. A particular example is the Newton algorithm to obtain solutions of an equation of the form  $f(x) = 0$ . The iteration step in question is  $x \mapsto \varphi(x) = x - f(x)/f'(x)$  (we omit the details concerning the proper choice of the state space and the problem that  $f'(x)$  may be zero for some values of  $x$ ). The algorithm consists of choosing an initial value  $x$ , then iterating  $\varphi$  on  $x$ , i.e. calculating  $\varphi^n(x)$  for  $n = 1, \dots, N$ . If the algorithm is successful, these iterations converges to a solution of the equation  $f(x) = 0$ ;

then  $\varphi^N(x)$  will be a good approximation of that solution, provided  $N$  is big enough. For more information on the Newton algorithm as dynamical system, see [51].

**1.2.5. General conclusion concerning the various time sets.** If we restrict ourselves to finite dimensional state spaces, the only important time sets are  $\mathbf{R}$ ,  $\mathbf{Z}$ , and  $\mathbf{Z}_+$ . In these cases the evolution map is determined by a *differential equation* (or *vector field*), an *invertible map* (or *automorphism*), respectively, a *general map* (or *endomorphism*). Often one does not mention explicitly the notion of an evolution map at all but discusses differential equations, and maps (invertible or not).

### 1.3. The evolution map

In the cases where the state space is a differentiable manifold (or a vector space) it is common to require the evolution map to be smooth, or piecewise smooth. Apart from this there are situations where the evolution maps respect some extra structure on the state space. We mention some of them.

**1.3.1. Symplectic (and Hamiltonian) systems.** For symplectic systems the state space  $X$  is a differentiable manifold with a *symplectic form*, i.e. a 2-form  $\omega$  such that  $d\omega = 0$  and such that the map  $\Omega : T(X) \rightarrow T^*(X)$ , defined by  $\Omega(v) = \omega(v, \cdot)$ , induces an isomorphism from  $T_x(X)$  to  $T_x^*(X)$  for each  $x \in X$ . We note that these conditions imply that the state space  $X$  is a manifold of even dimension, say  $2n$ , and that the  $n$ -fold exterior product of  $\omega$  is a nowhere zero *volume form*. A manifold, equipped with such a 2-form is called a *symplectic manifold*.

For these systems one requires the evolution map to respect the symplectic structure in the sense that  $(\Phi^t)^*\omega = \omega$  for each  $t \in T$ , where  $\Phi^t$  is the time  $t$  map as defined in Section 1.2.1.

In the case that the time set is  $\mathbb{R}$ , these systems are often called Hamiltonian. The reason for this is the following. If a vector field  $V$  on a symplectic manifold generates a symplectic flow, then  $\iota_V\omega = \omega(V, \cdot)$  is a closed 1-form. Often this 1-form  $\iota_V\omega$  is even exact so that there is a function  $H : X \rightarrow \mathbb{R}$  such that  $dH = \iota_V\omega$ . Such a function  $H$  is called a *Hamiltonian* of the system; it is constant along evolution curves of the dynamical system. Classical mechanical systems, without friction, are in general described by Hamiltonian systems. The function  $H$ , is then interpreted as the energy. Classical references for these systems are [61,1,3]. It is important to note that this restriction on the evolution map has severe consequences for the possible types of dynamics, e.g. Hamiltonian systems can have no attractors.

As in classical mechanics, other special classes of dynamical systems have also been introduced, like non-holonomic systems and Poisson systems.

**1.3.2. Volume preserving systems.** Volume preserving systems are very much like the symplectic ones that we defined above: instead of a symplectic form we have a volume form  $\eta$ . For an  $n$ -dimensional state space  $X$  this is an  $n$ -form which is nowhere zero. Also, here, the requirement is that the evolution map respects the volume form in the sense that

$(\Phi^t)^*\eta = \eta$  for all  $t$ . Since, as mentioned above, the  $n$ -fold product of a symplectic form on a  $2n$ -dimensional manifold is a volume form, each symplectic system is also volume preserving. In the dimension 2 there is no difference between a symplectic and a volume preserving system. In the case where the time set is  $\mathbb{R}$ , and the evolution map is generated by a vector field  $V$ , the condition that the volume is preserved is equivalent to the condition that  $\iota_V \eta$ , the *flux form* of  $V$ , is a closed  $(n - 1)$ -form.

The main interest of volume preserving systems lies in the possibility of analysing those properties of symplectic systems which are a consequence of the volume conservation only. In this context we mention that there is an extensive theory (ergodic theory) concerning dynamical systems whose state space is a measure space and whose time evolution respects this measure, though the measures, studied in ergodic theory, are often of a very different nature: they are often singular with respect to the Lebesgue measure (class). For the ergodic theory of differentiable dynamical systems we refer to [20,40,34]. We note however that Hamiltonian systems are very exceptional in the class of all volume preserving systems: in terms of the notion of genericity, to be discussed in Section 2.1, volume preserving systems in dimension greater than 2 generically are not Hamiltonian.

Another reason for studying these volume preserving systems is the fact that flows of incompressible fluids are volume preserving. So these systems can give, in principle, information on how e.g. a well localized amount of pollution is transported due to a steady flow in such an incompressible medium (like the sea). In the case where such a flow is time dependent, but periodic in time (like the tidal motion), this leads to periodically driven volume preserving systems, and by restricting the period map, to a volume preserving system with time set  $\mathbb{Z}$ .

**1.3.3. Reversible systems.** Reversible systems form another class derived from classical mechanics. Reversible refers to the situation where, in some sense, reversing the time direction transforms evolutions to evolutions as is the case in the Newtonian gravitation theory. In this gravitation theory the equation of motion for a number of point masses is a second-order differential equation of the form  $q'' = f(q)$  where  $f$  is a vector field on the *configuration space*, which is an open subset in a vector space which has three dimensions for each point mass; a configuration specifies the positions of all the point masses. Indeed, if  $t \mapsto q(t)$  is a solution of such an equation, then also  $t \mapsto q(-t)$  is a solution of the same equation.

However note: the configuration space is *not* a state space for such a system: in order to predict the future, we need the positions, i.e. the configuration, *and* the velocities. When reversing the direction of time, the velocities change sign. So in terms of the evolution in the state space we have: if  $t \mapsto (q(t), v(t) = q'(t))$  is an evolution then  $t \mapsto (q(-t), -v(-t))$  is also an evolution.

In the general definition of a reversible system one assumes that the time set  $T$  is symmetric, i.e.  $t \in T$  implies that  $-t \in T$ , and that the state space  $X$  is a manifold with an *involution*  $I : X \rightarrow X$ , i.e.  $I^2 = \text{id}_X$ . The condition on the evolution map  $\Phi$  is then that if  $t \mapsto x(t) = \Phi(x(0), t)$  is an evolution, then also  $t \mapsto I(x(-t))$  is an evolution, and hence that  $I(\Phi(x(0), -t)) = \Phi(I(x(0)), t)$ , or  $I\Phi^{-t} = \Phi^t I$ .

For general information on this subject we refer to [9,22,60].

**1.3.4. Systems with symmetry.** Reversible systems can in some sense be considered as systems with symmetry, namely invariant under time reversal and the involution  $I$  when applied at the same time. In the standard definition of systems with symmetry however, one considers symmetries in the state space only. So let  $X$  be a state space of a dynamical system on which a group  $G$  acts, i.e. for which a map  $\alpha : G \times X \rightarrow X$  is defined such that  $\alpha(e, x) = x$  and  $\alpha(g_1, \alpha(g_2, x)) = \alpha(g_1 g_2, x)$  for all  $g_1, g_2 \in G$  and  $x \in X$ ,  $e$  being the unit element in  $G$ . In this situation  $\alpha(g, x)$  is also denoted by  $gx$ . We say that the dynamical system defined by  $\Phi : X \times T \rightarrow X$  has symmetry group  $G$  if  $\Phi(gx, t) = g\Phi(x, t)$  for all  $x \in X$ ,  $g \in G$  and  $t \in T$ , i.e. if each  $\Phi^t$  commutes with each  $g \in G$ .

Note that the diffeomorphisms which have such a symmetry group  $G$ , i.e. which commute with each element  $g \in G$ , themselves form a group. This is also the case for the diffeomorphisms respecting a given symplectic or volume form. However, for a given involution  $I$  on a state space  $X$ , the diffeomorphisms defining reversible systems (with respect to that involution) do *not* form a group.

Some classical references in this area are [24,59,26].

**1.3.5. Systems depending on parameters.** Dynamical systems which depend on parameters are the subject of bifurcation theory. These systems can formally also be defined in terms of systems with extra structure. For this we define the extended phase space as the product of the usual phase space with the parameter space; we denote the extended phase space by  $X$  and the parameter space by  $P$ . Then we have a natural projection  $\pi : X \rightarrow P$ . This projection is considered as an extra structure which is preserved, in the sense that we impose on the evolution map that  $\pi\Phi^t = \pi$  for all  $t \in T$ . This means that for each  $p \in P$ , such an evolution map defines, by restriction, a dynamical system with state space  $X_p = \pi^{-1}(p)$ . These systems with parameters are used to describe dynamics which depend on external parameters, i.e. on the coordinates of the point  $p \in P$ . One may think of parameters which can be set externally, but which are kept constant during the time evolution as described by the evolution map. (Systems where such parameters are not kept constant are the subject of *system theory*.)

General references on this subject are [19,28,38]; in Section 4 of this chapter, we discuss some local bifurcations in systems depending on one parameter. This type of extra structure can be combined with the previous structures, e.g. one may consider symplectic systems depending on parameters. For the bifurcation theory of symplectic systems we refer to the Chapter on KAM theory in this volume.

**1.3.6. Shifts.** A very different class of dynamical systems is formed by the *shifts* – here one also speaks of *symbolic dynamics*. In this case the state space is no longer a manifold. In the simplest case it is the space of all sequences  $S = (s_i)$  whose elements belong to a given finite set  $A$ , called the alphabet, and whose indices have values in  $\mathbb{Z}$  or in  $\mathbb{Z}_+$  depending on whether the time set is  $\mathbb{Z}$  or  $\mathbb{Z}_+$ . The usual topology on such a state spaces can be given by the metric  $\rho$  which assigns to two sequences  $S = (s_i)$  and  $T = (t_i)$  a distance  $2^{-k}$  if  $k$  is the smallest absolute value of an index  $i$  for which the corresponding elements  $s_i$  and  $t_i$  are different. This topology on this space of sequences, which can be identified with the infinite product  $A^{\mathbb{Z}}$  or  $A^{\mathbb{Z}_+}$ , is the so called product topology (where  $A$  is assumed to have the discrete topology). The evolution map  $\Phi$  assigns to the sequence

$S = (s_i)$  and  $n \in T$  the sequence  $\Phi(S, n) = \sigma^n(S) = \tilde{S} = (\tilde{s}_i = s_{i+n})$ . One calls  $\sigma$  the *shift map*. These spaces of sequences are also called *shift spaces*.

A systems whose state space is a closed subset of  $A^{\mathbb{Z}}$  or  $A^{\mathbb{Z}_+}$  which is invariant under the shift map is called a *sub-shift*.

These dynamical systems are important as a link between probability theory and differentiable dynamics: (sub-)shifts, with invariant measures on the state space, are at the basis of Bernoulli and Markov processes. They are even at the basis of the general invariant measures constructed by thermodynamic formalism. On the other hand one can prove that many differentiable dynamical systems on manifolds have invariant sets (so called horseshoes, and more general so-called hyperbolic basic sets) which are homeomorphic, as a space *with* an evolution map, to such a shift. A classical reference for these shifts and their relation to smooth dynamics is [10]; for more information see also [34].

**1.3.7. Gradient systems and (non-)recurrence.** Gradient systems form a class of a somewhat different nature: these systems cannot be defined in terms of a structure which has to be preserved. For the definition of a gradient system we need a Riemannian manifold  $(X, g)$  as a state space ( $g$  denotes the Riemannian metric tensor), together with a (potential) function  $V : X \rightarrow \mathbb{R}$ . The gradient vector field  $v$  of  $V$  on  $X$  is then defined by the equation  $g(v, \cdot) = dV$ . Note that if  $X = \mathbb{R}^n$  with the usual Euclidean metric, then this definition reduces to the usual definition of the gradient:  $\text{grad}(V) = (\partial_1 V, \dots, \partial_n V)$ . The gradient system is then defined as the flow of the vector field  $-\text{grad}V$ .

These gradient systems are in some sense the opposite of Hamiltonian systems: if one thinks of the function  $V$  as the energy, then an evolution of the corresponding gradient system loses its energy ‘as fast as possible’. In order to make this difference between Hamiltonian and gradient clear we introduce the notion of recurrence.

### Recurrence

For a dynamical system with state space  $X$ , time set  $T$ , and evolution map  $\Phi$ , we say that a state  $x \in X$  is *recurrent* if for each neighbourhood  $U$  of  $x$  and each  $\bar{t} > 0$  there is a  $t \in T$  such that  $|t| > \bar{t}$  and  $\Phi(x, t) \in U$ . We say that  $x$  is *positively* (or *negatively*) *recurrent* if the  $t$  in the above definition can be chosen in  $T \cap \mathbb{R}_+$  (or in  $T \cap \mathbb{R}_-$ ). It is clear that stationary points and periodic points are recurrent, but the present notion is much more general. If  $x$  is a recurrent point of the state space, then also the evolution starting at that point, given by  $t \mapsto \Phi(x, t)$ , is called a *recurrent evolution*. Note that all its points are recurrent.

The main special property of gradient systems as opposed to Hamiltonian systems is that for gradient systems the only recurrent points are the stationary points. This follows easily from the property that if  $x(t)$  is an evolution of a gradient system with potential function  $V$  and if  $t_1 < t_2$ , then either  $V(x(t_1)) > V(x(t_2))$  or  $x(t)$  is a stationary evolution, i.e.  $x(t) = x(0)$  for all  $t$ . For Hamiltonian, and even for volume preserving systems, the situation is completely different. According to the Poincaré return theorem (e.g. see [3] or [1]), whenever the state space is compact, or, in the case of Hamiltonian systems, the energy level is compact, almost all states (in the sense of measure theory) are recurrent.

Besides gradient systems one also considers *gradient-like* systems. They have the (weaker) property that for any evolution  $x(t)$  and  $t_1 < t_2 \in T$  one has either



$V(x(t_1)) > V(x(t_2))$  or  $x(t)$  is a stationary solution and  $dV(x(t)) = 0$ . Note that also for these gradient-like systems only the stationary points are recurrent.

It was the description of dynamical systems in terms of such potential functions, and the assumption that the dynamics was gradient-like, which was at the basis of *catastrophe theory*. The qualitative theory and the bifurcation analysis of these potential functions was a precursor of and a model for the theory of non-linear dynamical systems and their bifurcations as we know it now. We will come back to this subject when discussing bifurcations in Section 4. Here we only mention some main references: Catastrophe theory was strongly advocated as a unifying theory for the description of many natural systems, especially biological systems, by R. Thom, [70], see also [4]. For the mathematical theory of singularities and their bifurcations, which is the theoretical basis of catastrophe theory, we refer to [72,11,8,5].

#### 1.4. Relations between the different classes of dynamical systems

In this section we discuss a number of constructions to make a new dynamical system from a given one. These constructions, and in particular the suspension and Poincaré map as discussed below, give an indication of how to translate results for dynamical systems with time set  $\mathbb{R}$  to those with time set  $\mathbb{Z}$ , and vice versa.

**1.4.1. Suspension.** Suspension is a general construction by which one assigns to a systems with time set  $\mathbb{Z}$  a system with time set  $\mathbb{R}$ . If we start with a state space  $X$ , and an evolution map, determined by an invertible map  $\varphi : X \rightarrow X$ , then the new system has state space  $\tilde{X} = X \times [0, 1] / \sim$ , where  $\sim$  stands for the identification of  $(x, 1)$  with  $(\varphi(x), 0)$ . An equivalent way of defining this new state space is  $\tilde{X} = X \times \mathbb{R} / \sim'$ , where  $\sim'$  stands for the identification of  $(\varphi(x), t)$  with  $(x, t + 1)$ . The new evolution map  $\Phi'$  is given, in the second representation, by  $\Phi'([(x, t)], \tau) = [(x, t + \tau)]$ , where we use the square brackets  $[ ]$  to refer to the equivalence classes under the equivalence relation  $\sim'$ .

If we start with a dynamical system, the state space of which is a manifold, and which is determined by a diffeomorphism  $\varphi$ , then the state space of the suspended system also has the structure of a manifold. However if we start with a dynamical system with a vector space as state space, then suspension leads to a state space which cannot be given the structure of a vector space because it is not simply connected. If, in the latter case the diffeomorphism defining the evolution map is orientation preserving, then the suspended system can be realized in a vector space in the sense that its state space can be considered as an open subset of a vector space. Even in that case and if the original diffeomorphism is given by simple equations (like e.g. the Hénon system) it is usually hard to give explicit equations for the suspended system.

**1.4.2. Poincaré map.** The construction of a Poincaré map is almost the inverse of the above construction. Here we start with a dynamical system whose state space is a manifold  $X$  and whose time set is  $\mathbb{R}$ . We call a sub-manifold  $Y \subset X$  a Poincaré section if:

- (1) it has co-dimension 1;



- (2) for each  $y \in Y$ , there is both a positive  $t_+(y)$  and a negative  $t_-(y)$  such that  $\Phi(y, t_{\pm}(y)) \in Y$  is the first return in  $Y$  of the evolution starting in  $y$  for positive, respectively, negative, time;
- (3) in each point of  $Y$ , the evolution through that point is transverse to, i.e. not tangent to,  $Y$ .

In this case, there is a well defined map  $\varphi : Y \rightarrow Y$  which maps each point  $y \in Y$  to the point  $\Phi(y, t_+(y))$ , where  $t_+(y)$ , as above, is the smallest positive real number for which  $\Phi(y, t_+(y)) \in Y$ . Clearly,  $\varphi$  is a diffeomorphism: its inverse is the map  $y \mapsto \Phi(y, t_-(y))$ .

So the newly constructed system has state space  $Y$  and evolution map generated by  $Y \ni y \mapsto \Phi(y, t_+(y)) \in Y$ .

This construction of making a Poincaré map is related to the construction of the period map (or the stroboscopic map) for periodically driven systems: If we consider a dynamical system defined by a differential equation  $x' = f(x, t)$  (which is not allowed according to our definition of a dynamical system), in which  $f(x, t)$  is periodic in  $t$ , say with period  $P$ , then we obtain a dynamical system, according to our terminology, by taking an extra state variable which we denote by  $\tau$ . This extended dynamical system is then defined by the equations

$$\begin{aligned} x' &= f(x, \tau) \\ \tau' &= 1. \end{aligned}$$

Since  $f$  has period  $P$  in its second variable,  $\tau$  should be interpreted as element of  $\mathbb{R} \bmod P \cdot \mathbb{Z}$ . This means that in the extended phase space we can define a codimension 1 sub-manifold  $Y = \{(x, \tau) | \tau = 0\}$ . This sub-manifold serves as the domain of a Poincaré map. Note that this Poincaré map assigns to each state the state which is reached after one period of the driving force. In this case the functions  $t_+$  and  $t_-$ , as introduced above, are both constant and equal to  $P$ , respectively,  $P^{-1}$ .

Consider a system, which is obtained by suspending a system with a manifold  $X$  as state space and defined by a diffeomorphism  $\varphi$ , and whose state space consequently is  $X \times \mathbb{R} / \sim'$ , with  $\sim'$  as above. We can return to the original system by taking the Poincaré map on the section consisting of the points  $\{(x, t) | t = 0\}$ . Note, however, that not all systems with time set  $\mathbb{R}$  admit the construction of a Poincaré map as we defined it: e.g. the occurrence of one stationary evolution is already an obstruction.

One often considers Poincaré maps in situations in which not all the above requirements are satisfied. The section  $Y$  has to have co-dimension 1 in any case. It may however happen that there are points  $y \in Y$  for which there is no positive  $t_+(y)$  or negative  $t_-(y)$  such that  $\Phi(y, t_+(y))$  or  $\Phi(y, t_-(y)) \in Y$ ; in this case the Poincaré map or its inverse is only defined on a part of  $Y$ . It also may happen that in some points  $y \in Y$  the evolution through  $y$  is tangent to  $Y$ ; this usually leads to discontinuities in the Poincaré map.

In these cases we do not obtain dynamical systems in the sense we defined, but rather *local dynamical systems* as introduced in the Sections 1.2.1 and 1.2.3. This is in particular the case for Poincaré maps of periodic orbits, to be discussed in Section 2.3.1.

**1.4.3. Taking the time  $t$  map.** If we start with a dynamical system with time set  $T = \mathbb{R}$ , then we obtain a new dynamical system by restricting the evolution map to the time set

$T' = \{nt | n \in \mathbb{Z} \text{ or } \mathbb{Z}_+\}$  for some  $t > 0$ . This procedure is much simpler than taking a Poincaré map, but the dynamical systems one obtains in this way are usually very atypical. This notion ‘atypical’ will be made precise in Section 2.1 when discussing the notion of genericity. This construction of taking the time  $t$  map has some theoretical interest because many numerical methods for approximating solutions of differential equations are based on (an approximation of) a time  $t$  map, usually for a small value of  $t$ .

**1.4.4. Taking restrictions.** If we have a dynamical system with state space  $X$ , time set  $T$ , and evolution map  $\Phi$ , then we call a subset  $Y \subset X$  *invariant* if  $\Phi(Y \times T) = Y$ . In that case one can define a new dynamical system with state space  $Y$  and evolution map  $\Phi' = \Phi|Y \times T$ . There are theorems which claim, under various hypotheses, the existence of invariant manifolds, often of low dimension and containing all the interesting dynamics, see [33,18]. This is of importance in the geometric study of dynamical systems, because the properties of low dimensional systems (say dimensions one and two) are much better known than those of high dimensional systems.

At this point we should mention that this construction of taking a restriction is also used in situations where one does not obtain a dynamical system in the strict sense of our definition but only a local dynamical system. Compare the partially defined Poincaré maps discussed above. We also think of the so-called centre manifolds which we will discuss in more detail in Section 4.2.1, dealing with generic local bifurcations. These centre manifolds are discussed in more detail in the Chapter on Local invariant manifolds and normal forms in this volume.

## 2. Transversality and generic properties

### 2.1. The notion of genericity

The notion of genericity is strongly related to the notion ‘almost all’ as it is used in measure and probability theory, but applies to a context where one has a topology instead of a (probability) measure. A general reference on the relation between these two notions of ‘almost all’, in the probabilistic and the topological context, is [47]. This notion of genericity, as a systematic way to separate the general, typical or generic, cases from the exceptional, atypical or non-generic, cases has been strongly advocated by R. Thom, e.g. see [70], in relation to catastrophe theory, see Section 4.1. Later it has played an important role in the theory of dynamical systems; we refer to [63] in which S. Smale formulated the foundations of the geometric theory of dynamical systems. The most important general result exploiting this notion of genericity is Thom’s transversality lemma which we will discuss in detail in Section 2.2.

We say that a topological space has the *Baire property* if any countable intersection of open and dense subsets is dense. According to the Baire category theorem, see [47] or [35], each complete metric space has the Baire property. Let  $F$  be a space with the Baire property; a subset of  $F$  is called *residual* if it contains a countable intersection of open and dense subsets. Such sets are considered big in the topological sense: a countable intersection of such sets is still dense (this notion is the analogue of ‘full measure’ in

probability theory). Related notions are: a *set of first category* which is a countable union of nowhere dense sets (this is the same as the complement of a residual set) and a  $G_\delta$  set which is the countable intersection of open sets (so that a set is residual if and only if it contains a dense  $G_\delta$  set).

We say that a property, which may or may not hold for points of  $F$ , is *generic* if the set of all points of  $F$  which possess this property is residual. Often such properties hold even for all elements in an open and dense subset, in which case we call the property *strongly generic*. Note that even in the case of a space  $F$  which does not have the Baire property, the notion of ‘strongly generic’ makes sense as a formalization of ‘almost all’. In this context also the notion of persistence is important. We say that a property is *persistent* for an element  $f \in F$  if the set of points, for which the property holds, contains a neighbourhood of  $f$ : then such a property persists under perturbations, provided they are sufficiently small.

We mentioned before that a generic property holds for almost all elements, in the topological sense. Unfortunately there are certain paradoxical results; e.g. there are residual subsets of the real line which have Lebesgue measure zero. For an extensive discussion of such facts we refer again to [47].

We shall use these notions in situations where  $F$  is a class of dynamical systems, e.g. systems defined by  $C^k$  vector fields (or differential equations) on a (compact) manifold  $X$ . In that case  $F$  is a Banach space with respect to the  $C^k$  topology. One can then prove that for the elements of  $F$  certain properties are (strongly) generic (below we shall give examples). This means that in this class of dynamical systems, though these properties don’t hold for all elements, by an arbitrarily small perturbation one can obtain a system for which they do hold. In the strongly generic case, one can even choose such a perturbation so that the property becomes *persistent*. So if the equation of motion of a system is only known approximately, one may assume that it has all generic properties. One has to be somewhat careful however: we can only impose *countably many* of such generic properties at the same time if we want the combination of these properties to be again generic.

Other classes of dynamical systems to be considered are those defined by  $C^k$  mappings, or diffeomorphisms, of a manifold  $X$  to itself. As long as  $X$  is compact, these spaces are Banach manifolds and hence have a topology that can be defined by a complete metric. In the case where  $X$  is not compact, the situation is somewhat more complicated: there one has for vector fields, maps, and diffeomorphisms the strong and the weak  $C^k$  topology. The weak topology can be defined by a complete metric, so that it defines a Baire space. The strong topologies cannot be defined by a complete metric, but still the Baire property can be shown to hold. This means that also in this case the notion of genericity makes sense; for details see [32].

## 2.2. Sard’s theorem and Thom’s transversality lemma

We already mentioned the central role of Thom’s transversality lemma in relation to genericity. We shall describe its proof because this explains the structure of many genericity results: it shows how to transform results on perturbations into genericity statements. The proof of the transversality lemma depends heavily on Sard’s theorem which we state first.

Let  $f : M \rightarrow N$  be a  $C^r$  map between differentiable manifolds (or vector spaces),  $r > 0$ . Then we define the set  $\Sigma_f$  of critical values of  $f$  as

$$\Sigma_f = \{y \in N \mid \exists x \in M \text{ such that } f(x) = y \text{ and } d_x f \text{ is not surjective}\}.$$

**THEOREM** (Sard, [11,32]). *For a  $C^r$  map  $f : M \rightarrow N$ , with  $r > 0$  and  $r > (\dim(M) - \dim(N))$ , the complement  $N - \Sigma_f$  is a residual subset of  $N$ . Moreover  $\Sigma_f$  has measure zero for each measure which belongs, on each domain of local coordinates, to the same measure class as the Lebesgue measure with respect to such local coordinates.*  $\square$

**DEFINITION** (Transversality, [32,69]). Let  $M$  be a smooth manifold with smooth sub-manifold  $N \subset M$  (without loss of generality we may take  $M$  and  $N$  to be  $C^\infty$ ). We say that a map  $f : V \rightarrow M$ , where  $V$  also is a smooth manifold and where  $f$  is at least  $C^1$ , is transversal with respect to  $N$  if for each  $v \in V$  with  $f(v) \in N$ , we have

$$T_{f(v)}(N) + d_v f(T_v(V)) = T_{f(v)}(M).$$

Two sub-manifolds  $N$  and  $V$  of  $M$  are called transversal if the inclusion map  $V \hookrightarrow M$  is  $N$ -transversal. This condition is symmetric in  $N$  and  $V$ : it is equivalent to the condition that for every  $x \in N \cap V$ , one has  $T_x(N) + T_x(V) = T_x(M)$ .  $\square$

We mention a few elementary properties of transversal maps. If  $f$  is transversal as in the above definition, then  $f^{-1}(N)$  is a sub-manifold of  $V$  the co-dimension of which is equal to the co-dimension of  $N$  in  $M$ . In particular, if  $\dim(V) + \dim(N) < \dim(M)$ , then transversality implies  $f(V) \cap N = \emptyset$ . In the space of  $C^r$  maps from  $V$  to  $M$ , with  $r > 0$ , the set of  $N$  transversal maps is open, at least if one assumes that  $N$  is a *closed* sub-manifold and if, in case  $V$  is not compact, one uses the strong topology.

**THEOREM** (Transversality Lemma [32,69]). *For  $M$ ,  $N$ , and  $V$  as in the above definition and  $r > 0$ , the  $N$ -transversal maps form a residual subset of the space of all  $C^r$  maps from  $V$  to  $M$  (both in the weak and the strong topology).*  $\square$

**PROOF.** We shall only discuss the proof for the simplest case, namely where  $M = \mathbb{R}^m$ ,  $N$  is a closed submanifold, and where  $V$  is compact. Since we know that transversality is an open property, we only have to prove that any (smooth) map  $f : V \rightarrow M$  can be approximated by an  $N$ -transversal map. Since any  $C^k$  map can be  $C^k$  approximated by a  $C^\infty$  map, e.g. see [32], we may assume  $f$  to be  $C^\infty$ .

The construction of the approximation goes in two steps: first we construct a *perturbing family* (or *unfolding*, or *deformation*), and then we prove that, *in this family*, there is, arbitrarily close to  $f$ , a map which is  $N$ -transversal.

A perturbing family in this context is a map  $F : V \times \mathbb{R}^p \rightarrow M$ , for some integer  $p$ , such that:

- (1)  $F(v, 0) = f(v)$ ,
- (2)  $F$  has the same class of differentiability as  $f$  and
- (3) for each  $(v, x) \in V \times \mathbb{R}^p$ , the derivative  $dF_{(v,x)}$  is a surjective map to  $T_{F(v,x)}(M)$ .

Such a family can easily be constructed: we may just take  $p = m$  and  $F(v, x) = f(v) + x$ . For such a perturbing family  $F$ , and  $\mu \in \mathbb{R}^p$ , we define  $f_\mu(v) = F(v, \mu)$ .

Now we come to the second step: we want to show that arbitrarily close to  $0 \in \mathbb{R}^p$  there are  $\mu$  such that  $f_\mu$  is  $N$ -transversal. For this we first observe that, since the derivative of  $F$  is surjective at any point, this map is transverse to any sub-manifold of  $V \times \mathbb{R}^p$  and in particular to  $N$ . This means that  $\tilde{N} = F^{-1}(N)$  is a sub-manifold of  $V \times \mathbb{R}^p$ . Next we define the projection  $\pi$  of  $\tilde{N}$  on  $\mathbb{R}^p$  as the restriction of the canonical projection  $V \times \mathbb{R}^p \rightarrow \mathbb{R}^p$  to  $\tilde{N}$ . Let  $\Sigma_\pi \subset \mathbb{R}^p$  be the set of critical values of  $\pi$ . Then for any  $\mu \notin \Sigma_\pi$ ,  $f_\mu$  is  $N$ -transversal. This is a consequence of the following arguments.

If  $f_\mu(v) \in N$ , then  $(v, \mu) \in \tilde{N}$ . Since  $\mu \notin \Sigma_\pi$ , the tangent space of  $V \times \mathbb{R}^p$  at  $(v, \mu)$  is spanned by the tangent spaces of  $V \times \{\mu\}$  and  $\tilde{N}$ ; these tangent spaces are mapped by  $(dF)_{(v, \mu)}$  to  $df_\mu(T_v(V))$  and  $T_{f_\mu(v)}(N)$ , respectively. The transversality of  $f_\mu$  (at  $v$ ) now follows from the fact that the derivative of  $F$  is everywhere surjective. And hence that  $df_\mu(T_v(V)) + T_{f_\mu(v)}(N) = T_{f_\mu(v)}(M)$ .

This, together with Sard's theorem, immediately implies that, arbitrarily close to  $0 \in \mathbb{R}^p$ , there are  $\mu$  such that  $f_\mu$  is  $N$ -transversal. This proves the theorem, at least in this special case.  $\square$

There are many adaptations of this theorem in which the class of mappings is of a more special nature. Then the construction of the perturbing family is the only new thing – the rest of the proof remains unchanged. In Section 2.2.2 we discuss an example of this which is of importance in dynamical systems and an example related with potential functions is considered in Section 1.3.7. First however in Section 2.2.1 we give some generalizations of Thom's transversality lemma.

**2.2.1. Jet extensions and transversality.** First we discuss a number of definitions. Let  $M$  and  $N$  be manifolds (or vector spaces). The  $k$ -jet of a smooth map  $f : M \rightarrow N$  at a point  $x \in M$  is the equivalence class of  $f$  under the equivalence relation  $\sim_{x,k}$  defined by:

The maps  $f, g : M \rightarrow N$  are equivalent in the sense of  $\sim_{x,k}$  if the derivatives of  $f$  and  $g$ , up to and including order  $k$  are equal in  $x$ . In this definition, the zeroth derivative is the value of the map itself; equal derivatives up to and including order  $k$  means that in some, and hence in any, local coordinate systems (on  $M$  and  $N$  near  $x$  and  $f(x)$ , respectively) all partial derivatives of  $f$  and  $g$  up to and including order  $k$  are equal at  $x$ .

The equivalence classes of  $\sim_{x,k}$  are called  $k$  jets at  $x$ ; the  $k$ -jet of  $f$  at  $x$  is denoted by  $[f]_x^k$ . In the case where  $M$  and  $N$  are vector spaces, the  $k$ -jets at  $x$  can be uniquely represented as polynomial mappings from  $M$  to  $N$  of degree at most  $k$ .

The set of all equivalence classes of  $\sim_{x,k}$  is denoted by  $J_x^k(M, N)$  and the union of these sets for all  $x \in M$  is denoted by  $J^k(M, N)$ . We note that  $J^k(M, N)$  has, in a natural way, the structure of a differentiable manifold with projections  $\pi_M$  and  $\pi_N$  on  $M$  and  $N$ , respectively: the  $k$ -jet  $[f]_x^k$  projects to  $x \in M$  and to  $f(x) \in N$ .

For a smooth (at least  $C^k$ ) map  $f : M \rightarrow N$  we define its  $k$ -jet extension as the map  $J^k(f) : M \rightarrow J^k(M, N)$  which assigns to each  $x \in M$  the  $k$ -jet  $[f]_x^k$ . Note that by definition  $\pi_M J^k(f)$  is the identity on  $M$  and  $\pi_N J^k(f) = f$ . If  $f$  is  $C^\ell$ , then  $J^k(f)$  is  $C^{\ell-k}$ .

**THEOREM (Transversality for  $k$ -jet Extensions).** *Let  $M$  and  $N$  be manifolds and let  $W$  be a sub-manifold of the jet manifold  $J^k(M, N)$  as defined above. Then, for each  $r > k$ , it is a*

generic property of  $C^r$  maps  $f$  from  $M$  to  $N$ , that its  $k$ -jet extension  $J^k(f)$  is transversal with respect to  $W$ .  $\square$

**SKETCH OF THE PROOF.** This version of the transversality theorem is not an immediate consequence of the first version. The reason is that not every map from  $M$  to  $J^k(M, N)$  is the  $k$ -jet extension of a map from  $M$  to  $N$ . So the main problem is to construct a perturbing family for  $f$ .

Such a perturbing family is a map  $F : M \times \mathbb{R}^p \rightarrow N$ , for some integer  $p$ , with the following properties:

1.  $F(x, 0) = f(x)$ ;
2.  $F$  is as differentiable as  $f$ ;
3. the map  $J_1^k(F) : M \times \mathbb{R}^p \rightarrow J^k(M, N)$  has a derivative which is everywhere surjective;  $J_1^k(F)$  assigns to each  $(x, \mu) \in M \times \mathbb{R}^p$  the  $k$ -jet at  $x$  of  $f_\mu$  (as before:  $f_\mu(x) = F(x, \mu)$ ; the subscript 1 indicates that we only take the  $k$ -jet ‘in the  $M$ -direction’, i.e. in the direction of the first variable).

It is not hard to see that once we have such a perturbing family, we can complete the proof in the same way as before.

We only describe the construction of the perturbing family in the simplest case where  $M = \mathbb{R}^m$  and  $N = \mathbb{R}^n$ . Let  $\mathcal{P}$  be the vector space consisting of all polynomial maps of degree at most  $k$  from  $\mathbb{R}^m$  to  $\mathbb{R}^n$ . Since  $\mathcal{P}$  is a finite dimensional vector space, it can be identified with  $\mathbb{R}^p$  for some  $p$  (depending on  $m, n$ , and  $k$ ). Using this identification we obtain a perturbing family in the following form:

$$F(x, Q) = f(x) + Q(x), \quad \text{for } x \in M \text{ and } Q \in \mathcal{P}.$$

Also for these jet extensions we refer the reader to [32] for more details.  $\square$

**2.2.2. Transversality with respect to (semi-)algebraic sets.** Up to now we discussed transversality with respect to a manifold. Often one has to deal with a slightly more complex situation where one wants transversality with respect to a (semi-)algebraic set, i.e. a set defined by algebraic equalities (and inequalities). This can be reduced to the case of manifolds due to the fact that such (semi-)algebraic sets can be stratified, and thus can be decomposed as a finite union of manifolds.

An elementary introduction to this theory can be found in [11, chapter 12] and on a more advanced level in [25, chapters I and II].

In order to provide examples of how all this is used, we indicate the proof that for generic vector fields on  $\mathbb{R}^n$ , and hence on any manifold, all singularities are hyperbolic (this is part of the Kupka-Smale theorem, which we will discuss further in Section 2.3), and, as a second example, the proof that for generic functions all singularities are non-degenerate (this is a special case of the theory of bifurcations of (potential) functions which we will discuss further in Section 4.1).

These are only two of the many genericity results which will be discussed in following sections. The proofs of these results are all based on arguments similar to the ones we use here.

We recall that for a vector field  $Z$  on  $\mathbb{R}^n$ , or on a manifold, we call the point  $x$  a *singular point* or a *singularity* of  $Z$  if  $Z(x) = 0$ , which is equivalent to  $x$  being a stationary point for the dynamical system. In such a singularity, the derivative of  $Z$  can be interpreted as a linear map from  $T_x$  to itself: in  $\mathbb{R}^n$  this is done by interpreting  $Z$  as a (smooth) map from  $\mathbb{R}^n$  to itself and then taking the derivative in the ordinary sense; on a general manifold one uses the same procedure in local coordinates and verifies that the result is independent of these coordinates. We call the singularity  $x$  *hyperbolic* if the derivative of  $Z$  at  $x$  has no eigenvalues on the imaginary axis. So now we want to prove:

**THEOREM (Genericity of Hyperbolic Singularities).** *For  $C^r$  vector fields,  $r > 0$ , on  $\mathbb{R}^n$ , or on a manifold, it is a generic property that all the singularities are hyperbolic.*  $\square$

We only discuss the proof for vector fields on  $\mathbb{R}^n$ . In the space  $L(n, n)$  of  $n \times n$ -matrices, the set  $A$  of those matrices which have at least one eigenvalue on the imaginary axis is a semi-algebraic subset. This can be seen as follows: The set  $A^* \subset L(n, n) \times \mathbb{R}$ , defined by  $A^* = \{(\alpha, t) \mid \det(\alpha - i t \text{Id}) = 0\}$ , where  $\text{Id}$  is the  $n \times n$  identity matrix, is by definition an algebraic, and hence a semi-algebraic, subset of  $L(n, n) \times \mathbb{R}$ . Since the image of a (semi-) algebraic set under an algebraic map is at least semi-algebraic (see [25]),  $A$ , the projection of  $A^*$  in  $L(n, n)$ , is also semi-algebraic. Since  $A$  has no interior points, it is the union of sub-manifolds of codimension  $\geq 1$ .

**PROOF.** Now we apply the transversality theorem for 1-jet extensions to vector fields as maps from  $\mathbb{R}^n$  to itself. We use the fact that the space  $J^1(\mathbb{R}^n, \mathbb{R}^n)$  can be canonically identified with  $\mathbb{R}^n \times \mathbb{R}^n \times L(n, n)$  so that the 1-jet of  $f$  at  $x$  is identified with  $(x, f(x), df_x)$ . We want transversality with respect to the subset  $\bar{A} = \mathbb{R}^n \times \{0\} \times A$ . Due to the previous observations we know that this set is the union of sub-manifolds of co-dimensions  $\geq n + 1$ . Hence the 1-jet extension of a map  $f$  from  $\mathbb{R}^n$  to itself is transversal with respect to  $\bar{A}$ , if and only if the image of its 1-jet extension is disjoint from  $\bar{A}$ .

The transversality theorem for 1-jet extensions implies that, for generic  $C^r$  vector fields  $Z$ ,  $r \geq 2$ , the image of the 1-jet extension  $J^1(Z)$  is disjoint from  $\bar{A}$ . This implies the announced statement for  $r \geq 2$ . The statement for the case that  $r = 1$  follows from the following observations:

1. the property of only having hyperbolic singularities is open (on compact manifolds and also in general with respect to the strong topology), also in the  $C^1$  topology;
2. any vector field can be  $C^1$  approximated by a  $C^2$  vector field, which can then be approximated by a vector field without non-hyperbolic singularities, implying that the vector fields without non-hyperbolic singularities are dense in the  $C^1$  vector fields.  $\square$

Also for our second example we have to recall some definitions. We consider real functions  $f$  on  $\mathbb{R}^n$ , or on a manifold, which are at least  $C^2$ . A *singularity* of such a function is a point  $x$  where  $df_x = 0$ . At such a singularity, the second derivative is a symmetric  $n \times n$  matrix. Also this second derivative at a singularity can be defined in an intrinsic way as a quadratic form on the tangent space  $T_x$ , e.g. see [11]. The singularity is called *non-degenerate* if this second derivative, as a matrix, has a non-zero determinant. We now want to prove the following statement:



**THEOREM (Genericity of Non-degenerate Singularities).** *For  $C^r$  functions,  $r \geq 2$  on  $\mathbb{R}^n$ , or on a manifold, it is a generic property that all the singularities are non-degenerate.*  $\square$

**PROOF.** We only give the proof for functions on  $\mathbb{R}^n$ , which we assume for the moment to be at least  $C^3$ . We consider the jet space  $J^2(\mathbb{R}^n, \mathbb{R})$  which can be identified with  $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times S(n)$ , where  $S(n)$  is the space of symmetric  $n \times n$  matrices. With this identification, the 2-jet of  $f$  at  $x$  is identified with  $(x, f(x), f'(x), f''(x))$ . The subset of  $S(n)$ , consisting of those symmetric  $n \times n$  matrices with determinant zero, is denoted by  $DS(n)$  ( $D$  referring to ‘degenerate’). This subset  $DS(n)$  is even algebraic, and since it has no interior points, it is the union of sub-manifolds of codimension at least one.

In  $J^2(\mathbb{R}^n, \mathbb{R})$  we now introduce the subset  $\tilde{DS}(n) = \mathbb{R}^n \times \mathbb{R} \times \{0\} \times DS(n)$ . This subset is the union of sub-manifolds of codimension at least  $n + 1$ . Now it follows from the transversality theorem for jet extensions that for generic  $f$ , the image of the 2-jet extension  $J^2(f)$  is disjoint from  $\tilde{DS}(n)$ . For such  $f$ , we have for each  $x \in \mathbb{R}^n$ , whenever  $f'(x) = 0$  (i.e. whenever  $x$  is a critical point of  $f$ ), that the second derivative, as a matrix, has non-zero determinant. This means that all the singularities are non-degenerate. This proves our statement for the case where  $r \geq 3$ . The case  $r = 2$  is now obtained by the same method as we used above to lower the differentiability requirement in the example of genericity of hyperbolicity of singularities of vector fields.  $\square$

### 2.3. Generic properties of dynamical systems based on transversality (Kupka-Smale theorems)

In the previous section we have already discussed, as a first example, a generic property of dynamical systems, namely the property for vector fields (or differential equations) of having only hyperbolic singularities. This is part of the so-called Kupka-Smale theorem, see [36,37,62], which is the main subject of this section. Most of the further results in this section are adaptations of this theorem but to symplectic, Hamiltonian, and volume preserving systems. There are analogous results for systems with parameters, but these will be considered (in the simplest cases) in Section 4 which is devoted to generic local bifurcations in one parameter families of dynamical systems.

**2.3.1. The Kupka-Smale theorem.** Apart from the original publications, the Kupka-Smale theorem, and related results, are also discussed in text books such as [49,55]. First we have to discuss various definitions concerning hyperbolicity.

#### Hyperbolicity

For a differentiable map  $\varphi : X \rightarrow X$ ,  $X$  a manifold, we say that  $p \in X$  is a *fixed point* if  $\varphi(p) = p$ . We call the fixed point *hyperbolic* if all the eigenvalues of  $d\varphi_p$  have norm different from one.

We say that  $p \in X$  is a *periodic point* (of period  $k$ ) if it is a fixed point of  $\varphi^k$ , and if  $\varphi^\ell(p) \neq p$  for  $\ell = 1, \dots, k - 1$ . As a periodic point of  $\varphi$ ,  $p$  is called *hyperbolic* if it is a hyperbolic fixed point of  $\varphi^k$ .

We recall that for a vector field  $Z$  on a manifold  $X$ , a *singularity*  $p$ , i.e. a point  $p \in X$  such that  $Z(p) = 0$ , is *hyperbolic* if all the eigenvalues of  $dZ_p$  have real parts different from 0.



Note the connection between these definitions: a vector field defines a dynamical system with time set  $\mathbb{R}$  (disregarding possible problems related to non-uniqueness or non-existence of global solutions). So if  $\Phi$  is the evolution map of the dynamical system defined by  $Z$ , then each singularity  $p$  of  $Z$  is a fixed point of  $\Phi^t$ . Then  $p$  is hyperbolic as a singularity of  $Z$  if and only if it is a hyperbolic fixed point of  $\Phi^t$  whenever  $t \neq 0$ .

For a vector field  $Z$  and corresponding evolution map  $\Phi$  we say that a non-singular point  $q \in X$  belongs to a *periodic orbit*, or *periodic evolution*, of period  $t > 0$  if  $\Phi^t(q) = q$  (and  $\Phi^{t'}(q) \neq q$  for  $0 < t' < t$ ). The periodic orbit as a subset of  $X$  is then  $\Phi(\{q\} \times [0, t))$ . In that case,  $(d\Phi^t)_q$  has one eigenvalue equal to one: it maps  $Z(q)$  to itself; if all its other eigenvalues have norm different from one, the periodic orbit is called *hyperbolic*.

The relation with the definition of hyperbolicity for fixed points is based on the construction of a local Poincaré map for the periodic orbit: Consider a small co-dimension one manifold  $S$ , containing the point  $q$  on the periodic orbit and nowhere tangent to the vector field  $Z$ . The corresponding Poincaré map  $P$  is just the map from  $S$  to itself (at least defined in a neighbourhood of  $q$ ) as defined in Section 1.4.2. Then  $q$  is a fixed point of  $P$ , and it is a hyperbolic fixed point if and only if the periodic orbit of  $Z$  is hyperbolic. We note that the eigenvalues of the derivative of the Poincaré map at the fixed point (here  $q$ ) are also called Floquet multipliers of the periodic orbit.

**THEOREM (Kupka-Smale, part I).** *Let  $X$  be a manifold. For  $C^k$  maps from  $X$  to itself,  $k > 0$ , it is a generic property to have all its fixed points and periodic points hyperbolic. For  $C^k$  vector fields on  $X$ ,  $k > 0$  it is a generic property to have all its singularities and periodic orbits hyperbolic.*  $\square$

Apart from some technical complications, the proof of this theorem follows the pattern outlined in Section 2.2.2, for details see [49] or [55].

For the second part of the theorem under consideration, we need the definitions of stable and unstable manifolds. We shall restrict ourselves here to  $C^k$  dynamical systems,  $k \geq 1$ , with time set either  $\mathbb{R}$  or  $\mathbb{Z}$ . We shall only define stable manifolds: by reversing the time, stable manifolds become unstable manifolds, and vice versa.

### *Stable (and unstable) manifolds*

First we consider a fixed point  $p$  of a diffeomorphism  $\varphi : X \rightarrow X$  (we don't assume it to be hyperbolic). The *local stable manifold*  $W_{loc}^s(p)$  of  $p$  (for some neighbourhood  $U$  of  $p$ ) consists of the points  $x \in U$  such that  $\varphi^n(x) \in U$  for all positive  $n$  and such that the distance between  $\varphi^n(x)$  and  $p$  decreases exponentially in  $n$ . One can prove that  $W_{loc}^s$  is a sub-manifold, the dimension of which equals the number of eigenvalues of  $d\varphi_p$  with norm smaller than one, and which is of the same class of differentiability as  $\varphi$ ; e.g. see [33].

The *(global) stable manifold* of  $p$  is then defined as  $W^s(p) = \bigcup_{n < 0} \varphi^n(W_{loc}^s(p))$ . Since  $\varphi$  is a diffeomorphism and  $W_{loc}^s$  is a sub-manifold,  $W^s$  is at least an injectively immersed sub-manifold.

For a periodic point  $p$  of a diffeomorphism  $\varphi$ , say with period  $k$ , its (local) stable manifold is defined as the (local) stable manifold of  $p$  as a fixed point of  $\varphi^k$ .

Next we consider a singularity  $p$  of a vector field  $Z$  on a manifold  $X$ . Let  $\Phi$  be the evolution map corresponding to  $Z$ . The *local stable manifold*  $W_{loc}^s(p)$  of  $p$  (for some neighbourhood  $U$  of  $p$ ) consists of those points  $x \in U$ , for which  $\Phi(x, t)$ , for  $t > 0$ ,

stays in  $U$  and converges to  $p$  with exponentially (in  $t$ ) decreasing distance between  $p$  and  $\Phi(x, t)$ . Also in this case we obtain the (global) stable manifold by taking  $W^s(p) = \bigcup_{t < 0} \Phi^t(W_{loc}^s(p))$ . As in the previous case, this stable manifold has its dimension equal to the number of eigenvalues of  $dZ_p$  with negative real part; it has the same degree of differentiability as  $Z$ , and is an injectively immersed sub-manifold.

Finally we consider a periodic orbit  $\gamma$  of a vector field  $Z$ . The local stable manifold  $W_{loc}^s(\gamma)$  (for some neighbourhood  $U$  of  $\gamma$ ) is defined as the set of those points  $x \in U$  for which  $\Phi(x, t)$ , for positive  $t$ , stays in  $U$  and approaches  $\gamma$  exponentially. The definition of  $W^s(\gamma)$  in terms of  $W_{loc}^s(\gamma)$  is as before. Also this is an injectively immersed sub-manifold whose class of differentiability is equal to that of  $Z$ . In order to give a formula for the dimension, we take a point  $q \in \gamma$  and denote the period of  $\gamma$  by  $\tau$ . Then  $q$  is a fixed point of  $\Phi^\tau$ . Let  $r$  be the number of eigenvalues of  $(d\Phi^\tau)_q$  with norm smaller than one. Then the dimension of  $W^s(\gamma) = r + 1$ .

This last definition is linked to that of the stable manifold of a fixed point through the Poincaré map  $P$  on a section transverse to the periodic orbit  $\gamma$  at a point  $q$ . In fact, the intersection of a local stable manifold of  $\gamma$  with such a section  $S$  is a local stable manifold of the Poincaré map, defined on that section.

Unstable manifolds are stable manifolds of the dynamical system in which the time has been reversed. This means that a (local) unstable manifold for  $\varphi$  is a (local) stable manifold for  $\varphi^{-1}$ . Also a (local) unstable manifold for the vector field  $Z$  is a (local) stable manifold for the vector field  $-Z$ . Unstable manifolds are denoted by  $W^u$ .

**THEOREM (Kupka-Smale, part II).** *For generic  $C^k$  dynamical systems,  $k > 0$ , (diffeomorphisms or vector fields) on a manifold  $X$ , all intersections of stable and unstable manifolds are transversal.*  $\square$

We conclude this section with a few remarks on hyperbolicity. After all: why would we like to know that usually all fixed point etc. are hyperbolic? A main reason is that near a hyperbolic orbit the dynamics is simple in the sense that by a proper choice of coordinates one can obtain so called linearizations. We formulate this result, due to Grobman and Hartman, see [27,30], only for fixed points and singularities, but a similar result holds for periodic orbits:

**THEOREM (Grobman-Hartman).** *Let  $\varphi$ , respectively  $Z$ , be a diffeomorphism, respectively a vector field, with evolution map  $\Phi$ , on  $X$  and let  $p \in X$  be a fixed point, respectively a singularity. Then there is a continuous coordinate system  $x_1, \dots, x_m$  in a neighbourhood of  $p$  such that  $\varphi$ , respectively  $\Phi^t$ , expressed in these coordinates, is linear in a neighbourhood of  $p$ .*  $\square$

In general it is not possible to make such linearizing coordinates differentiable. If one imposes, however, stronger but still generic, conditions on the eigenvalues of  $d\varphi_p$ , respectively of  $dZ_p$ , then it is possible to find linearizing coordinates which are differentiable, see [64,65].

#### *Genericity in the real analytic case*

In the usual treatment of transversality and the corresponding generic properties, one works in the context of  $C^k$  or at most  $C^\infty$  mappings. This is convenient because then one

can make localized perturbations. However, a general method to extend these results to the real analytic case was given in [17]; this also applies to the genericity results below for special classes of dynamical systems.

### 2.3.2. Generic properties of symplectic, Hamiltonian and volume preserving systems.

The main reference for the results which we discuss here is [56]. We treat a number of cases separately.

**Symplectic diffeomorphisms** We start with the periodic and fixed points of symplectic maps (time set  $\mathbb{Z}$ ). Our starting point is a state space which is a manifold  $X$  with symplectic form  $\omega$ , see Section 1.3.1. For a symplectic diffeomorphism  $\varphi : X \rightarrow X$ , the 1-jet extension satisfies some extra properties which reflect the fact that  $\varphi$  has to respect the symplectic form. This is the reason that we define the 1-jet space differently here:  $J_\omega^1(X)$  is the space consisting of triples  $(x_1, x_2, L)$ , with  $x_1, x_2 \in X$ , and  $L$  a linear map from  $T_{x_1}$  to  $T_{x_2}$  such that  $L^*(\omega_{x_2}) = \omega_{x_1}$ , i.e. we require  $L$  to be a linear symplectic map from the tangent space at  $x_1$  to the tangent space at  $x_2$ . Note that this is consistent with what we defined in Section 1.3.1, i.e. a diffeomorphism  $\varphi$  is symplectic if and only if for each  $x \in X$ , the triple  $(x, \varphi(x), d\varphi_x)$  is in  $J_\omega^1(X)$  as defined here. The important fact is that with this definition of the 1-jet space, we have, within the class of symplectic diffeomorphisms, the transversality theorem with respect to submanifolds of these jet spaces.

This does however not lead to the same Kupka-Smale theorem we had before: the reason is that, in the group of linear symplectic maps of a vector space to itself, the set of non-hyperbolic maps has interior points. This is related to the fact that for a symplectic automorphism, the eigenvalues have to satisfy some extra conditions: whenever  $\lambda$  is an eigenvalue of such an automorphism, then so are  $\bar{\lambda}$ ,  $\lambda^{-1}$ , and  $\bar{\lambda}^{-1}$ . This implies that a pair of non-real eigenvalues  $\lambda, \bar{\lambda}$  on the unit circle (of multiplicity one) cannot be pushed off the unit circle by a small perturbation. In order to formulate a generic property for fixed points, excluding 1-jets of a certain type (like the non-hyperbolic ones in the non-symplectic case) we need a subset of  $G\ell_\omega(2n)$ , the group of symplectic automorphisms in a vector space of dimension  $2n$ , which:

1. has no interior points (and is semi-algebraic);
2. is invariant under symplectic conjugations, i.e. independent of the choice of a particular basis;
3. the elements of which cause dynamic complexity.

For this we take the complement of the set of those  $L \in G\ell_\omega(2n)$  for which each eigenvalue is either hyperbolic, i.e. in norm different from one, or has norm one, but has only multiplicity one and is not a root of unity (eigenvalues which are a root of unity or have multiplicity greater than one correspond to resonance, and complicated dynamics). Linear symplectic automorphisms which satisfy this last condition are called *elementary*.<sup>1</sup>

---

<sup>1</sup>One may also require all the eigenvalues on the unit circle to be multiplicatively independent, i.e. their logarithms, divided by  $2\pi i$  together with 1, to be independent over the rationals. Even with this extra condition the following result remains true.

We say that a fixed point  $p$  of a symplectic diffeomorphism  $\varphi$  is *elementary* if  $d\varphi_p$  is elementary; if  $p$  is a periodic point of such a diffeomorphism with  $\varphi^k(p) = p$  and  $\varphi^i(p) \neq p$  for  $0 < i < k$ , then it is called *elementary* if  $d\varphi_p^k$  is elementary.

**THEOREM (Kupka-Smale for Symplectic Diffeomorphisms, I).** *For  $C^r$  symplectic diffeomorphisms,  $r > 0$ , it is a generic property that all its fixed points and periodic points are elementary.*  $\square$

In this situation the stable and unstable manifolds of the fixed and periodic points are defined as in the general case. They have however some additional properties. First, on a  $2n$ -dimensional symplectic manifold, stable and unstable manifolds always have dimension  $\leq n$ . This is a consequence of the above mentioned restriction on the eigenvalues of a symplectic automorphism. Second, stable and unstable manifolds are always isotropic, in the following sense:

A sub-manifold  $Y$  of a symplectic manifold  $X$  with symplectic form  $\omega$  is called *isotropic* if for each  $p \in Y$  and  $v, w \in T_p(Y)$ ,  $\omega(v, w) = 0$ .

Still in this situation the second part of the Kupka-Smale theorem remains unchanged:

**THEOREM (Kupka-Smale for Symplectic Diffeomorphisms, II).** *For  $C^r$  symplectic diffeomorphisms,  $r > 0$ , it is a generic property that all stable and unstable manifolds intersect transversally.*  $\square$

The condition of transversal intersections in the last theorem has to be interpreted correctly: if  $p$  is a periodic or a fixed point, then one has to disregard  $p$  itself as an intersection of  $W^s(p)$  and  $W^u(p)$ : if  $p$  is not hyperbolic this is indeed a non-transversal intersection. If the non-hyperbolic eigenvalues of such a periodic or fixed point all have multiplicity 1, then, within the symplectic context, it is persistent as a non-hyperbolic fixed or periodic point, and hence the non-transversal intersection cannot be perturbed away.

**Volume preserving diffeomorphisms** In the one-dimensional case volume preserving diffeomorphisms are just translations; this is too trivial a case for further consideration. In dimension two, volume preserving diffeomorphisms are just symplectic diffeomorphisms which we have already discussed. In dimensions greater than two, the generic properties of fixed and periodic points are the same as in the general case. The reason is that the only restriction on the derivative of a volume preserving diffeomorphism at a fixed point is that the product of its eigenvalues (the determinant) equals  $\pm 1$ . In dimension three, in the space of linear volume preserving automorphisms, the set of non-hyperbolic ones has no interior points. So all the considerations from the general case carry over.

**Hamiltonian flows: the singularities** Here again the state space is a manifold  $X$  with symplectic form  $\omega$ . For a dynamical system with time set  $\mathbb{R}$  there is a generating vector field  $Z$ , i.e.  $Z(p) = \partial_t \Phi(p, 0)$ . We recall, see Section 1.3.1, that the requirement that the maps  $\Phi^t$  preserve  $\omega$  is equivalent to the requirement that the 1-form  $\iota_Z \omega = \omega(Z, \cdot)$  is closed. In many examples, e.g. when  $X$  is a vector space, this 1-form is even exact so that there is a function, the *Hamiltonian function*,  $H : X \rightarrow \mathbb{R}$  such that  $dH = \omega(Z, \cdot)$ . From now on we consider only the case where there is such a Hamiltonian function.

We recall an important fact from Hamiltonian dynamics: for each evolution  $x(t)$  of such a system, the function  $H(x(t))$  is constant. This means that such a dynamical system decomposes to a 1-parameter family of dynamical systems with state spaces  $X_h = H^{-1}(h)$ . This will be especially of importance in the next section on periodic orbits.

We now formulate the first generic property concerning the singularities of a Hamiltonian function. For this we recall that a singularity of a function is a point where its derivative is zero; it is non-degenerate if its second derivative, as a quadratic form, has maximal rank.

**THEOREM (Genericity of Critical Points).** *For  $C^r$  Hamiltonian systems,  $r > 0$ , on a symplectic manifold  $X$ , it is a generic property that the corresponding Hamiltonian function has only non-degenerate critical points and that in any two different critical points, the values of the Hamiltonian function are different.*  $\square$

We observe that whenever the Hamiltonian vector field is  $C^r$ , a corresponding Hamiltonian function is  $C^{r+1}$ . The Hamiltonian function is not unique, but any two possible Hamiltonian functions for the same Hamiltonian vector field locally only differ by a constant. So that the statement on the critical points of the Hamiltonian function makes sense for a  $C^1$  Hamiltonian vector field and is independent of this choice of the Hamiltonian (at least if the symplectic manifold, i.e. the state space, is connected).

For a singularity  $p$  of a Hamiltonian vector field  $Z$ , one can show that whenever  $\lambda$  is an eigenvalue of  $dZ_p$ , then so are  $-\lambda$ ,  $\bar{\lambda}$  and  $-\bar{\lambda}$ . This means that also in this case it is not a generic property that all singularities are hyperbolic; the above generic conditions on the Hamiltonian only imply, in terms of the eigenvalues of the derivative of the vector field, that there is no eigenvalue zero. We now call a singularity of a Hamiltonian vector field *elementary* if it has no eigenvalue zero and if all its eigenvalues on the imaginary axis have multiplicity one.<sup>2</sup>

**THEOREM (Kupka-Smale for Hamiltonian Singularities).** *For  $C^r$  Hamiltonian vector fields,  $r > 0$ , it is a generic property that all singularities are elementary.*  $\square$

The stable and unstable manifold of a singularity  $p$  of a Hamiltonian vector field  $Z$  with Hamiltonian  $H$  are contained in the level set of  $H$  which contains the singularity (of the vector field). This can lead to intersections which are non-transversal in  $X$  (which can be seen easily in the case where  $X$  has dimension 2). Still we have the following theorem:

**THEOREM (Kupka-Smale for Stable and Unstable Manifolds of Singularities).** *For  $C^r$  Hamiltonian vector fields,  $r > 0$ , it is a generic property that stable and unstable manifolds of singularities intersect transversally as sub-manifolds of their level set of the Hamiltonian function.*  $\square$

Note that, assuming the generic property for the Hamiltonian formulated in the previous theorem, the critical level of a singularity contains no other critical points, so that, except for the singularity itself, it is a smooth submanifold. As usual, the singularity itself is disregarded as the intersection of stable and unstable manifolds, so that the notion of transversal intersection within the level of the Hamiltonian is well defined.

---

<sup>2</sup>One may also require all the eigenvalues on the imaginary axis to be independent over the rationals: then the following results are true.

**Hamiltonian flows: the periodic orbits** Here the results are rather complicated to formulate in a more or less complete way, and also, as we shall explain, they belong already to the theory of bifurcations, i.e. to the theory of parametrized dynamical systems. So we only indicate the general flavour and give references.

As we observed before, the dynamics near a periodic orbit can be analysed in terms of a Poincaré map on some section of co-dimension 1. As our state space is a symplectic manifold, say of dimension  $2n$ , such a section  $S$  has dimension  $2n - 1$  and is partitioned in  $2n - 2$  dimensional level sets  $S_h = S \cap H^{-1}(h)$  of the Hamiltonian function. It turns out that the symplectic form  $\omega$  on  $X$  defines, by restriction, a symplectic structure in each of these levels. Furthermore, the Poincaré return map is a symplectic diffeomorphism in each level. This means that this return map is a 1-parameter family of symplectic diffeomorphisms (only locally defined) in dimension  $2n - 2$ .

Generic properties of such 1-parameter families of symplectic maps were studied in [44]; in [66] it was proved that these genericity results are also valid for Poincaré maps of generic Hamiltonian flows.

**Volume preserving flows** As before, the 1-dimensional case is completely trivial, the 2-dimensional case is the same as for Hamiltonian systems. So from now on we assume that the dimension is at least three. The singularities can be dealt with as in the case of volume preserving diffeomorphisms: now the condition is that at a singularity of a vector field, whose flow is volume preserving, the sum of the eigenvalues of the derivative is zero. In dimensions three and larger, in the set of volume preserving linear vector fields, the non-hyperbolic ones (in the sense that they have at least one eigenvalue on the imaginary axis) have no interior points. So for the singularities the same Kupka-Smale theorem as in the general case holds. The same is true for the transversality of intersections of stable and unstable manifolds.

Now we come to the periodic orbits. In the dimensions four and larger the situation is analogous: if we construct a Poincaré map for a periodic orbit, then the Poincaré return map is a volume preserving local diffeomorphism in dimension at least three. This means that, by the same arguments as used in the case of volume preserving diffeomorphisms, we get the same Kupka-Smale theorem as in the general case.

In the dimension three the situation is different. The Poincaré return map is now a 2-dimensional volume preserving diffeomorphism, i.e. a symplectic diffeomorphism. For the fixed points of such maps there are generically two possibilities for the eigenvalues of the derivative:

1. They are  $\lambda$  and  $\lambda^{-1}$ , both real and in norm different from 1 – in this case the periodic orbit is hyperbolic.
2. They are  $\lambda$  and  $\bar{\lambda}$ , both on the unit circle and not equal to a root of unity – in this case we call the periodic orbit elliptic (due to the fact that the eigenvalues are on the unit circle) and elementary (due to the fact that roots of unity are avoided). These elliptic periodic orbits have ‘no stable and unstable manifolds’, more precisely, these manifolds coincide with the periodic orbit.

**THEOREM (Kupka-Smale for Volume Preserving Flows).** *For  $C^r$  volume preserving flows on a manifold  $X$  of dimension at least four,  $r > 0$ , it is a generic property that all*

*singularities and periodic orbits are hyperbolic and that stable and unstable manifolds intersect transversally.*

*In the dimension three the conclusion is the same except that we then have to allow also elementary elliptic periodic orbits as defined above.*  $\square$

REMARK. As we observed earlier, the dynamical systems which preserve a symplectic structure also preserve a volume structure. But within the class of volume preserving systems, the symplectic (or Hamiltonian) systems are exceptional. This fact is illustrated by the above theorems: there are many types of fixed or periodic orbits which are allowed in generic symplectic systems, but not in volume preserving systems. This is because such types of fixed or periodic points cannot be perturbed away within the class of symplectic systems, but can be perturbed away within the class of volume preserving systems. A similar remark applies often when we have a class of dynamical systems together with a subclass. Not only with the symplectic systems as a subclass of the volume preserving systems, but also with the volume preserving systems as a subclass of general differentiable dynamical systems. This can be illustrated by a simple and concrete example. If we consider on the 2-sphere  $S^2$  a vector field  $Z$  defining an area preserving flow, then there is a corresponding Hamiltonian function  $H : S^2 \rightarrow \mathbb{R}$ . (This follows from the fact that in this dimension area preserving implies symplectic, and the fact that the 2-sphere is simply connected.) So then  $\iota_Z \omega = dH$  and each orbit of this system is contained in a level curve of  $H$ . Assuming that the function  $H$  is generic, this implies that there are only three types of orbits: stationary points, orbits which tend, both for  $t \rightarrow \infty$  and for  $t \rightarrow -\infty$ , to a stationary point, while all the other orbits are periodic and non-hyperbolic. Generically there are only finitely many orbits of the first two types so that almost all points of the 2-sphere belong to such a non-hyperbolic periodic orbit. This in spite of the fact that in the class of all vector fields on  $S^2$ , the generic elements have only periodic orbits which are hyperbolic.

### 3. Generic properties which are not based on transversality: the Closing Lemma

The main result which we consider in this section is Pugh's Closing Lemma [52]. Consider a dynamical system with state space  $X$ , time set  $T$  and evolution map  $\Phi$ . We recall that a point  $x \in X$  is *recurrent* if  $\Phi(x, t)$  comes arbitrarily close to  $x$  for arbitrarily large values of  $t$ , see Section 1.3.7.

A related notion is that of non-wandering. For  $X$ ,  $T$  and  $\Phi$  as above, we say that  $x \in X$  is *non-wandering* if for any neighbourhood  $U$  of  $x$  and positive real  $\bar{t}$ , there is a  $t \in T$  such that  $|t| > \bar{t}$  and  $\Phi(U \times \{t\}) \cap U \neq \emptyset$ .

It is clear that each point on a periodic orbit is recurrent and that each recurrent point is non-wandering, but not vice versa.

A basic question then is, given a recurrent point, whether one can make this point periodic by an arbitrarily small perturbation of the dynamics. Below we make this question precise and give known positive and negative results. For this we restrict ourselves to dynamical systems whose state space is a differentiable manifold and whose time set is either  $\mathbb{R}$  or  $\mathbb{Z}$ , i.e. we only consider vector fields and diffeomorphisms. The  $C^r$  closing problem is then:



Given a  $C^r$  vector field or diffeomorphism with a recurrent point  $x$ , is it true that by an arbitrarily small (in the  $C^r$  sense) perturbation one can obtain a new diffeomorphism or vector field such that  $x$  is part of a periodic orbit?

For  $r = 0$ , in which case the question has to be formulated slightly differently, the answer is positive and the proof is simple. For  $r = 1$  the answer is also positive, but the proof is very delicate, see the above reference [52]. In fact, this solution of the closing problem gives slightly more: not only can one obtain a periodic orbit with an arbitrarily small perturbation, but one can even make such a perturbation with support contained in an arbitrarily small neighbourhood of  $x$ .

With the positive result for the  $C^1$  case, and some arguments involving semi-continuity, one obtains the following genericity theorem:

**THEOREM** (Pugh [53]). *For  $C^1$  diffeomorphisms and vector fields on a manifold  $X$  it is a generic property that in the set of non-wandering points, the periodic points are dense.*  $\square$

The closing problem with  $r > 1$  is even harder. In fact the only positive result on the  $C^r$  closing problem for  $r \geq 2$  is that for diffeomorphisms of the circle and for a large class of differential equations on the torus, the answer to the  $C^r$  closing problem is positive for all  $r \geq 2$ . More surprising is the fact that, according to Gutierrez, on a large class of non-compact manifolds of dimensions  $\geq 2$ , the answer to the closing problem is negative for differential equations for all  $r \geq 2$ , at least with respect to the *strong* Whitney topology, see [29] where also further references are given. In fact the result of Gutierrez shows that the answer to the closing problem in the stronger sense, namely requiring that the perturbation takes place in an arbitrarily small neighbourhood of the recurrent point  $x$ , is negative for all  $r \geq 2$ .

We should point out an important difference between this result and the results obtained by the application of the transversality theorem. It is best explained with an example. We consider 1-parameter families of diffeomorphisms on the circle  $f_\mu : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  such that the corresponding map  $F : \mathbb{S}^1 \times \mathbb{R} \rightarrow \mathbb{S}^1$ , defined by  $F(x, \mu) = f_\mu(x)$  is at least  $C^3$ . If we assume that the rotation numbers of  $f_0$  and  $f_1$  are different (we don't go into the meaning of this condition, but it is not a serious restriction for our argument), then the set  $A \subset [0, 1]$ , of those  $\mu$  for which  $f_\mu$  has no periodic points, has positive Lebesgue measure, see [31]. On the other hand, on a compact manifold the non-wandering set of any diffeomorphism is non-empty, so that generically there should be periodic points. So, in the topological sense there are almost always periodic points, though in *all* 1-parameter families of circle diffeomorphisms as above, the parameter set for which there are no periodic points is non-negligible in the measure theoretic sense. This discrepancy between the two notions of 'almost all' has been already mentioned before.

Such paradoxical situations (namely the set of parameter values, in generic  $k$ -parameter families, for which the generic property fails to hold, having positive measure), are known not to exist in the case of genericity results based on transversality. This is related to the fact that, according to Sard's theorem, not only the complement of the set of critical values is residual, but also the set of critical values has measure zero for the Lebesgue measure class. In fact it is not hard to prove the following extension of the transversality theorem.



**THEOREM (Extended Transversality).** *Let the smooth manifolds  $N \subset M$  and  $V$  and  $r > 0$  be as in the transversality lemma (see Section 2.2). Then, for generic  $p$ -parameter families  $F : V \times \mathbb{R}^p \rightarrow M$ , the parameter values  $\mu \in \mathbb{R}^p$  for which  $f_\mu : V \rightarrow M$ , defined by  $f_\mu(x) = F(x, \mu)$ , is  $N$ -transversal, form a residual subset of  $\mathbb{R}^p$ , the complement of which has measure zero.  $\square$*

The proof of this extended transversality theorem follows the same lines as the proof of the transversality lemma. One only should note that the generic  $p$ -parameter families  $F$  are here those  $F$ , who, as a map from the manifold  $V \times \mathbb{R}^p$  to  $M$ , are  $N$ -transversal. For these generic families one then applies Sard's theorem to the projection of  $F^{-1}(N)$  to  $\mathbb{R}^p$ .

#### 4. Generic local bifurcations

In this section we summarize a number of well-known results concerning local bifurcations, first for (potential) functions and then for differential equations and diffeomorphisms. We mentioned already in Section 1.3.7 how this bifurcation theory for (potential) functions, or catastrophe theory, was a precursor of the corresponding theory for dynamical systems. In the present sections we shall exploit this: we first discuss the case of functions and then show how the theory for dynamical systems initially develops along the same lines but soon becomes much more complex. For further information on the theory of bifurcations of functions, [Z,1967] [11] are introductory texts, while [8,43] are extensive expositions on a more advanced level. For (local) bifurcations of dynamical systems some main references are [5,19,28,46,38].

##### 4.1. Generic local bifurcations of functions (Catastrophe Theory)

We consider real functions on  $\mathbb{R}^n$  depending on parameters  $\mu = (\mu_1, \dots, \mu_p)$ , which we assume to be  $C^\infty$ , as functions on  $\mathbb{R}^n \times \mathbb{R}^p$ . The theory of the bifurcations of such functions has been developed by a number of mathematicians, mainly R. Thom, B. Malgrange, and J. Mather, see e.g. [39,43,67,72].

For a point  $\bar{x} \in \mathbb{R}^n$  and a parameter value  $\mu_0 \in \mathbb{R}^p$  such that  $df_{\mu_0}(\bar{x})$  is non-zero, it follows from the implicit function theorem that there is a  $\mu$ -dependent  $C^\infty$  coordinate such that, for  $\mu$  near  $\mu_0$  and  $x$  near  $\bar{x}$  the function  $f_\mu$  equals the first coordinate. This means that for  $\mu$  near  $\mu_0$  and  $x$  near  $\bar{x}$  no abrupt changes take place: there are no (local) bifurcations.

Generically, the set of points  $(x, \mu)$  such that the derivative of  $f_\mu$  at  $x$  is zero, form a sub-manifold of dimension  $p$  in  $\mathbb{R}^n \times \mathbb{R}^p$ . This follows easily from the Thom transversality theorem for jet extensions, Section 2.2.1. Consequently we expect that for each value of  $\mu$  the corresponding points  $x$ , where  $df_\mu(x) = 0$ , will be locally isolated. In these *critical points* or *singularities*, one can consider the second derivative. We already saw in Section 2.2.2 that this second derivative is a quadratic form on the tangent space at the singularity. Another application of the transversality theorem (for 2-jet extensions) implies that generically the set of points  $(x, \mu)$ , such that  $x$  is a singularity of  $f_\mu$  with degenerated second derivative, is a union of manifolds of dimension less than  $n$ . This is due to the fact that in  $S(n)$ , the set of symmetric  $n \times n$  matrices, the matrices with determinant zero form an algebraic subset without interior points, see Section 2.2.2.

Now consider a point  $\bar{x}$  which is critical for  $f_{\mu_0}$  but such that the second derivative is non-degenerate (we call such points *non-degenerate critical points*). Due to the Morse lemma, after a suitable  $\mu$ -dependent  $C^\infty$  coordinate change, for  $x$  near  $\bar{x}$  and  $\mu$  near  $\mu_0$  we have

$$f_\mu(x_1, \dots, x_n) = \sum_i \pm x_i^2 + c(\mu).$$

So, apart from an additive constant, depending only on  $\mu$ , there are no changes for  $\mu$  near  $\mu_0$  and  $x$  near  $\bar{x}$ : there are no (local) bifurcations.

In order to prepare for the statement of further results, we mention a general result for degenerate critical points of functions depending on parameters. It is an easy consequence of the Morse lemma. Suppose  $f_{\mu_0}$  has a degenerate critical point in  $\bar{x}$ . Let  $k$  be the rank of the second derivative of  $f_{\mu_0}$  at  $\bar{x}$ .

**SPLITTING LEMMA, OR GENERALIZED MORSE LEMMA ([11]).** For  $\mu_0$  and  $\bar{x}$  as above, after a suitable  $\mu$  dependent  $C^\infty$  coordinate change, we have, for  $x$  near  $\bar{x}$  and  $\mu$  near  $\mu_0$ :

$$f_\mu(x_1, \dots, x_n) = \tilde{f}_\mu(x_1, \dots, x_{n-k}) + \sum_{i=n-k+1}^n \pm x_i^2.$$

We call  $\tilde{f}_\mu(x_1, \dots, x_{n-k})$  the essential part of  $f_\mu(x_1, \dots, x_n)$ . □

This means that, in order to describe what happens near a degenerate critical point, we are confined to the essential part. This enables us to formulate the main theorem of ‘elementary catastrophe theory’ which gives a full description of the local structure near any degenerate critical point in generic functions depending on a parameter of dimension at most 4.

*Main theorem of catastrophe theory (Malgrange, Thom, Mather).* A generic function depending on a parameter  $\mu$  of dimension at most 4, has, near each of its critical points an essential part which has, after a suitable coordinate transformation and up to an additive constant  $c(\mu)$  depending only on  $\mu$ , one of the following seven forms:

- $f_\mu(x) = x_1^3 + \mu_1 x_1$  (fold catastrophe);
- $f_\mu(x) = \pm x_1^4 + \mu_1 x_1 + \mu_2 x_1^2$  (cusp catastrophe);
- $f_\mu(x) = x_1^5 + \mu_1 x_1 + \mu_2 x_1^2 + \mu_3 x_1^3$  (dove tail catastrophe);
- $f_\mu(x) = \pm x_1^6 + \mu_1 x_1 + \mu_2 x_1^2 + \mu_3 x_1^3 + \mu_4 x_1^4$  (butterfly catastrophe);
- $f_\mu(x) = x_1^3 + x_2^3 + \mu_1 x_1 x_2 + \mu_2 x_1 + \mu_3 x_2$  (hyperbolic umbilic catastrophe);
- $f_\mu(x) = x_1^3 - x_1 x_2^2 + \mu_1 (x_1^2 + x_2^2) + \mu_2 x_1 + \mu_3 x_2$  (elliptic umbilic catastrophe).

These coordinate transformations are of the form  $(\tilde{x}, \tilde{\mu}) = (\tilde{x}(x, \mu), \tilde{\mu}(\mu))$  and transform the ‘central singularity’ to the origin, both in the  $x$ -space and in the  $\mu$ -space. □

### Remarks

1. In these above cases we speak of bifurcations, because abrupt changes occur: e.g. the number of critical points of  $f_\mu$  near  $x = 0$  changes discontinuously as a function of  $\mu$  for  $\mu$  near 0.

2. In the above representations we see that the number of  $\mu$ -coordinates in the explicit expressions is not always the same. The number of  $\mu$ -coordinates in these expressions is called the co-dimension; it is the smallest number of parameters needed to have such bifurcations in generic families. (If the actual dimension of  $\mu$  is larger, the dependence on the remaining parameters can be transformed away.)

#### 4.2. Centre manifolds and generic local bifurcations of differential equations

We consider differential equations  $x' = F_\mu(x)$ , where  $F_\mu$  can be considered as a vector field depending on parameters  $\mu = (\mu_1, \dots, \mu_p)$ , and their bifurcations at singular, or stationary, points, i.e. points where the vector field  $F_\mu$  is zero. These bifurcations have a strong relation with the bifurcations of functions as described in the previous section: If  $f_\mu$  is a function, depending on parameters, then  $F_\mu = -\text{grad} f_\mu$  is a vector field depending on parameters and the bifurcations of functions carry over to bifurcations of gradient vector fields (or of the corresponding differential equations); the critical points of the potential functions corresponding to stationary points of vector fields. Initially, the present theory for differential equations develops almost completely parallel to the corresponding theory for functions, but soon it diverges and the situation for vector fields becomes much more complicated: there are many vector fields which are not of gradient type. Since all considerations are local, we may assume our vector fields to be defined on  $\mathbb{R}^n$  and to depend on parameters  $\mu = (\mu_1, \dots, \mu_p) \in \mathbb{R}^p$ .

##### *Flow box lemma*

Let  $F_\mu$  be a smooth vector field on  $\mathbb{R}^n$  with  $F_{\mu_0}(x_0) \neq 0$ . Then, up to a smooth  $\mu$ -dependent change of coordinates and restricted to a neighbourhood of  $(x_0, \mu_0)$ , the vector field  $F_\mu$  equals the first coordinate vector field, so that the corresponding evolution map locally has the form  $\Phi_\mu((x_1, x_2, \dots, x_n), t) = (x_1 + t, x_2, \dots, x_n)$ ; the corresponding system of differential equations is then  $x'_1 = 1, x'_2 = \dots = x'_n = 0$ .  $\square$

This lemma is a direct consequence of the well known ‘existence, uniqueness, and smooth dependence on initial states’ theorem for differential equations, e.g. see [19]. This lemma illustrates the analogy between the non-critical points for functions and the non-stationary points for vector fields. Also here, the pairs  $(x, \mu)$ , such that  $x$  is a singularity of  $F_\mu$ , form generically a  $p$ -dimensional sub-manifold of  $\mathbb{R}^n \times \mathbb{R}^p$ . As pointed out already in Section 2.2.2, the analogue of non-degenerate critical points here is the *hyperbolic singularities*: we recall that a singularity  $x$  of a vector field  $F$  is hyperbolic if none of the eigenvalues of  $dF(x)$ , as a linear map of the tangent space  $T_x$  to itself, is on the imaginary axis.

The linearization theorem for flows, as formulated in Section 2.3.1 can be seen as an analogue of the Morse lemma for functions.

*Centre manifold theorem (see also the chapter ‘Local invariant manifolds and normal forms’ in this volume)*

Let  $F_\mu$  be a vector field and assume  $x_0$  is a non-hyperbolic singularity of  $F_{\mu_0}$  where  $dF_{\mu_0}(x_0)$  has  $k$  hyperbolic eigenvalues and  $(n - k)$  eigenvalues on the imaginary axis.

Then there is a sub-manifold  $W_{(x_0, \mu_0)}^c$  in  $\mathbb{R}^n \times \mathbb{R}^p$  of dimension  $n - k + p$ , called the *centre manifold*, such that:

1.  $(x_0, \mu_0) \in W_{(x_0, \mu_0)}^c$ ;
2. for each  $(x, \mu) \in W_{(x_0, \mu_0)}^c$ , the vector  $F_\mu(x)$  is tangent to  $W_{(x_0, \mu_0)}^c(\mu) = W_{(x_0, \mu_0)}^c \cap \mathbb{R}^n \times \{\mu\}$ ;
3. the derivative of  $F_{\mu_0}|_{W_{(x_0, \mu_0)}^c}(\mu_0)$  at  $x_0$  has only eigenvalues on the imaginary axis;

It can be shown that the centre manifold is in general not unique; if  $F_\mu$  is  $C^k$ ,  $0 < k < \infty$ , then  $W_{(x_0, \mu_0)}^c$  can be made  $C^k$ .

For the proof of this theorem, and its extensions, one can consult the classical reference [33] or the more recent version in [19]; the first appendix in [50] is intended as an introduction to the main ideas of [33] while some more recent results are also mentioned.

A few remarks concerning this centre manifold are in order. From the specification it follows that for  $(\bar{x}, \mu) \in W_{(x_0, \mu_0)}^c$ , the solution of the differential equation  $x' = F_\mu(x)$ , starting in  $\bar{x}$  will stay in the centre manifold, at least for some time. It will stay because of the above condition 2, but, since the centre manifold is usually not a closed sub-manifold, it may ‘disappear’ in finite time. Disappearing means, in terms of the intrinsic topology of the centre manifold, that it goes to infinity, i.e. is leaving every compact subset; in terms of the geometry of  $\mathbb{R}^n \times \mathbb{R}^p$  it means that it crosses ‘the boundary’ of the centre manifold.

A related observation is that  $F_{\mu_0}|_{W_{(x_0, \mu_0)}^c}(\mu_0)$  defines a local dynamical system: the evolution map is not defined on all of  $W_{(x_0, \mu_0)}^c(\mu_0) \times \mathbb{R}$  but only on a neighbourhood of  $\{(x_0, \mu_0)\} \times \mathbb{R}$  in  $W_{(x_0, \mu_0)}^c(\mu_0) \times \mathbb{R}$ . This is a consequence of the fact that, as we noted above, evolutions starting in the centre manifold may escape in finite time from it. These local dynamical systems have already been mentioned when discussing the various time sets, see Sections 1.2.1 and 1.2.3. So by being restricted to a centre manifold, see Section 1.4.4, we obtain a localized form of a parametrized family of dynamical systems, parametrized by  $\mu$  in a neighbourhood of  $\mu_0$  in  $\mathbb{R}^p$ . The state spaces of the new dynamical systems have dimension  $n - k$ . So we have lowered the dimension by an amount which equals the number of ‘hyperbolic’ eigenvalues, i.e. eigenvalues which are not on the imaginary axis.

As in the case of bifurcations of functions, compare the splitting lemma in Section 4.1, we want to be sure that, when restricted to a centre manifold we don’t lose ‘essential’ information about the dynamics. The precise statement of the results is rather technical, see [33], but the main point is that the only information we lose, when restricted to a centre manifold, is given by the number of contracting and expanding hyperbolic eigenvalues at the central singularity  $(x_0, \mu_0)$ , i.e. the number of eigenvalues of  $dF_{\mu_0}(x_0)$  which have negative (or positive) real part.

After the present discussions it may be clear that, when dealing with bifurcations of singularities of vector fields or differential equations, we can restrict ourselves to those situations where there are no hyperbolic eigenvalues at the central singularity in  $(x_0, \mu_0)$ , since this corresponds to the situation after restriction to a centre manifold. Also this is completely analogous to what we did in the case of bifurcations of functions, i.e. restriction to the essential part.

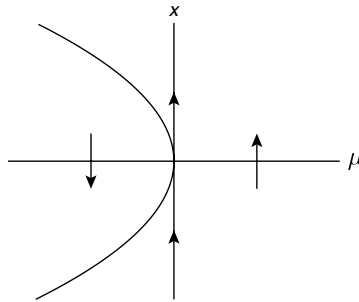


Fig. 1. Dynamics near a generic saddle-node (positive case).

**4.2.1. Generic bifurcations in centre manifolds.** Here we only consider 1-parameter families of vector fields. For more general cases see the concluding remarks in Section 4.4. Generically there are only two types of non-hyperbolic singularities which can occur, namely where there is one eigenvalue zero (saddle-node bifurcation) and where there is a pair of complex-conjugate non-zero eigenvalues on the imaginary axis (Hopf bifurcation).

**Saddle-node bifurcation** After restriction to a centre manifold, in this case we have a (local) 1-dimensional dynamical system depending on one parameter  $\mu$ . Without loss of generality we may assume that the saddle-node occurs at  $x = 0$  for  $\mu = 0$ . If we write our system as  $x' = F_\mu(x)$ , then we have  $F_0(0) = F'_0(0) = 0$ . Next we impose two conditions:  $F''_0(0) \neq 0$  and  $\partial_\mu F_0(0) \neq 0$ . If these are satisfied, the saddle-node is called *generic*. It is not hard to show that it is indeed a generic property for 1-parameter families of differential equations that all their saddle-nodes are generic in this sense. Also these conditions are independent of the coordinates used. Depending on the signs of the two quantities which we require to be non-zero, there are four cases. However they can be transformed into each other by inverting the direction of the  $x$ -axis and/or the  $\mu$ -axis. So we only have to consider the case where both  $F''_0(0)$  and  $\partial_\mu F_0(0)$  are positive; we call this the positive case. For the dynamics in a neighbourhood of  $x = 0$  and  $\mu = 0$ , see Figure 1, and the analysis in [19]. One can prove, e.g. see [46], that any two generic saddle-node bifurcations are *topologically conjugate* in the sense of the following theorem.

**SADDLE-NODE THEOREM.** Let  $F_\mu$  and  $\tilde{F}_\mu$  both have a generic saddle-node at  $\mu = 0$ ,  $x = 0$ . (Both are assumed to be positive in the above sense.) Then there is a homeomorphism  $H$  of the form  $H(x, \mu) = (H_1(x, \mu), h(\mu))$ , from a neighbourhood  $U$  of  $x = 0$ ,  $\mu = 0$  to another neighbourhood  $\tilde{U}$  of  $x = 0$ ,  $\mu = 0$ , which maps evolutions of  $F_\mu$  to evolutions of  $\tilde{F}_{h(\mu)}$  in the following sense. Denoting the evolution maps corresponding to  $F_\mu$ , and  $\tilde{F}_\mu$ , by  $\Phi_\mu$ , respectively  $\tilde{\Phi}_\mu$ , whenever  $\Phi_\mu(\{x\} \times [0, t]) \subset U$ , the homeomorphism  $H$ , or the homeomorphisms  $H_1$  and  $h$ , satisfy:

$$\tilde{\Phi}_{h(\mu)}(H_1(x, \mu), t) = H_1(\Phi_\mu(x, t), \mu).$$

□

A homeomorphism like  $H$  in the above theorem is called a local *topological conjugacy* between the (parametrized) flows defined by  $F_\mu$  and  $\tilde{F}_\mu$ .

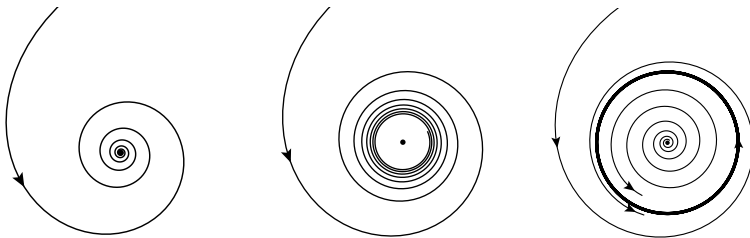


Fig. 2. Hopf bifurcation in the positive case: dynamics for  $\mu < 0$ ,  $\mu = 0$ , respectively  $\mu > 0$ .

The fact that we take here  $H$  to be a *homeomorphism* and not a diffeomorphism, which would seem more natural since we generally assume our objects to be ‘sufficiently differentiable’, is explained in Section 5 on structural stability and moduli.

**Hopf bifurcation** In this case we consider a 1-parameter family of 2-dimensional differential equations  $F_\mu$  such that for  $\mu = 0$  at  $x = 0$  there is a singularity, and such that the eigenvalues of  $dF_0(0)$  form a complex-conjugate non-zero pair on the imaginary axis. Also, here one has to impose further generic conditions before one can describe the dynamics near  $x = 0$ ,  $\mu = 0$ . In order to describe these conditions we first observe that, for  $\mu$  near 0 there is, by the implicit function theorem, a one parameter family of singularities  $x(\mu)$  of  $F_\mu$ , where  $x(\mu)$  depends smoothly on  $\mu$  and  $x(0) = 0$ . We denote the eigenvalues of the derivative of  $F_\mu$  at  $x(\mu)$  by  $\lambda(\mu)$  and  $\bar{\lambda}(\mu)$ . The first generic condition is then that the real part of  $\lambda'(0)$  is different from 0, i.e. that the eigenvalues cross the imaginary axis with positive speed. Before formulating the second condition we observe that if we approximate  $F_0$  by its linear part, then all solutions of the corresponding linear differential equation are periodic (except the zero solution). The second generic condition is that the terms up to order three of  $F_0$  at 0 force all the solutions, starting near 0, to converge to 0 either for positive time or for negative time. The explicit condition, in terms of the partial derivatives up to and including order three of  $F_0$ , can be found in [42] – it occupies about one full page.

Also, here there are four cases, depending on whether the real part of the derivative of the eigenvalues is positive or negative and on whether 0 is an attracting or repelling point for the evolutions of the dynamical system defined by  $F_0$ . In all these cases there is a periodic orbit which is splitting off a singularity. We call the bifurcation *supercritical* if the periodic orbit is attracting and *subcritical* if the periodic orbit is repelling. All these cases can be transformed into each other by reversing the time and/or the  $\mu$ -axis. So we only have to consider the case where the real part of the derivative of the eigenvalues is positive and where 0 is attracting; we call this the positive case. In this situation, a periodic orbit is splitting off the singularity when  $\mu$  moves from negative to positive values and the periodic orbit is attracting (so the bifurcation is supercritical). See Figure 2.

Even under the generic extra conditions, there is no analogue of the above saddle-node theorem: a conjugating homeomorphism has to map each periodic orbit to a periodic orbit of the other system, and corresponding periodic orbits must have the *same period*. It is not hard to make examples where these periods cannot be matched. For this reason one defines a weaker notion, namely *topological equivalence*. A topological equivalence is a

homeomorphism which maps evolutions of one dynamical system to evolutions of another dynamical system, preserving the direction of time but not necessarily preserving the exact time parametrization. Further motivation for the introduction of topological equivalences, in relation to structural stability, will be given in Section 5. In the sense of topological equivalence there is an analogue of the above saddle-node theorem. We can add to this topological description the fact that the diameter of the periodic solution grows like  $\sqrt{\mu}$ , in the sense that if, for  $\mu > 0$ ,  $d(\mu)$  denotes the diameter of the periodic solution of  $F_\mu$ , then  $d(\mu)/\sqrt{\mu}$  has a well defined and non-zero limit for  $\mu \rightarrow 0$ .

### 4.3. Centre manifolds and local bifurcations of fixed points of diffeomorphisms

One can obtain examples of bifurcations of fixed points of diffeomorphisms, by starting with a parametrized family of differential equations, with evolution maps  $\Phi_\mu$ , and then taking the parametrized family of diffeomorphisms given by  $\Phi_\mu^t$  for some positive  $t$ . In this way singularities of the vector fields, defining the differential equations, become fixed points of the diffeomorphisms. (If there are periodic solutions of the differential equations, there may be extra fixed points of these diffeomorphisms for exceptional  $t$ -values which are multiples of periods.) With this construction hyperbolic stationary points of vector fields are transformed to hyperbolic fixed points of diffeomorphisms, as we have already observed in Section 2.3.1. The linearization theorem, see also Section 2.3.1, can be seen as an analogue of the Morse lemma for non-degenerated critical points of functions. Also the centre manifold theorem in the previous section carries over to diffeomorphisms. Only one detail has to be formulated differently, namely the condition that the vector field has to be tangent to the centre manifold (condition 2 in the centre manifold theorem): in the case of a parametrized family  $\varphi_\mu$  of diffeomorphisms near a fixed point  $x_0$  of  $\varphi_{\mu_0}$ , we require that the centre manifold  $W_{(x_0, \mu_0)}^c$  is mapped locally to itself by  $\varphi(x, \mu) = (\varphi_\mu(x), \mu)$  in the sense that  $W_{(x_0, \mu_0)}^c \cap \varphi(W_{(x_0, \mu_0)}^c)$  is an open subset of  $W_{(x_0, \mu_0)}^c$  containing  $(x_0, \mu_0)$ . Also in this case we obtain, by restriction, a local dynamical system, depending on the parameter  $\mu$ , in the sense that the evolution map is only defined in a neighbourhood of  $\{(x_0, \mu_0)\} \times \mathbb{Z}$  in  $W_{(x_0, \mu_0)}^c \times \mathbb{Z}$ .

This means that also here we need only consider bifurcations *after restriction to a centre manifold*.

**4.3.1. Generic bifurcations in centre manifolds.** Again we are restricted to the bifurcations in generic 1-parameter families. For an extensive treatment of these bifurcations we refer to [46]. There are three possible types of bifurcations which are distinguished by the non-hyperbolic eigenvalue(s). The three cases are: one eigenvalue equal to one (saddle-node), one eigenvalue equal to minus one (period doubling), and a pair of non-zero complex conjugate eigenvalues on the unit circle (Hopf Neïmark Sacker bifurcation). As before, one can prove that these bifurcations, with some generic conditions on the higher order terms, are the only local bifurcations of fixed points which occur in generic 1-parameter families of diffeomorphisms.

**Saddle-node for diffeomorphisms** Here we consider a 1-parameter family of 1-dimensional diffeomorphisms  $\varphi_\mu(x)$ , and we assume that the saddle-node bifurcation

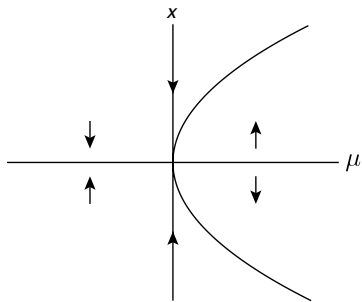


Fig. 3. The dynamics near a period doubling bifurcation (negative case); the arrows indicate the direction in which points move after iterating the diffeomorphism *twice*.

occurs for  $\mu = 0$  at  $x = 0$ . This implies that we have  $\varphi_0(0) = 0$  and  $\varphi'_0(0) = 1$ . We impose the following generic conditions:  $\varphi''_0(0) \neq 0$  and  $\partial_\mu \varphi_0(0) \neq 0$ . As in the previous case of the saddle-node for differential equations we have four cases, depending on the signs of these non-zero quantities. Without loss of generality we may assume these quantities to be positive (to be called the positive case), because we can obtain the other cases by reversing the  $x$ -axis and/or the  $\mu$ -axis. For a detailed discussion of this bifurcation we refer to [46]. The dynamics near this bifurcation is just like the dynamics of the time  $t$  map of the evolution maps corresponding to a saddle-node bifurcation for differential equations, so here again we can refer to Figure 1. In this case the analogue of the saddle-node theorem is still true (but much harder to prove), see [46]. The only adaptation one has to make is to adjust slightly the definition of a topological conjugacy between (families of) diffeomorphisms. For parametrized families of diffeomorphisms  $\varphi_\mu$  and  $\tilde{\varphi}_\mu$  a conjugacy is a homeomorphism of the form  $H(x, \mu) = (H_1(x, \mu), h(\mu))$  such that  $\tilde{\varphi}_{h(\mu)}(H_1(x, \mu)) = H_1(\varphi_\mu(x), \mu)$ .

**Period doubling bifurcations** Here we deal with a 1-parameter family  $\varphi_\mu$  of 1-dimensional diffeomorphisms satisfying  $\varphi_0(0) = 0$  and  $\varphi'_0(0) = -1$ . In order to formulate the generic conditions we impose on the higher order terms, we first observe that, by the implicit function theorem there is a smooth 1-parameter family  $x(\mu)$  of fixed points of  $\varphi_\mu$  (with  $x(0) = 0$ );  $\lambda(\mu)$  denotes the derivative  $\varphi'_\mu(x(\mu))$ . The first generic condition is  $\lambda'(0) \neq 0$ . For the second generic condition we consider the Taylor series  $\varphi_0^2(x) = \alpha x + \beta x^2 + \gamma x^3 + \text{higher order terms}$ . From the fact that the derivative  $\varphi'_0(0) = -1$ , we conclude that in this Taylor series  $\alpha = 1$  and  $\beta = 0$ . Our generic condition is that  $\gamma \neq 0$ . Again there are four cases which can be transformed into each other by reversing the time (i.e. by replacing  $\varphi_\mu$  by  $\varphi_\mu^{-1}$ ) and/or by reversing the  $\mu$ -axis. So we may restrict ourselves to the case where  $\lambda'(0) < 0$  and  $\gamma < 0$  (the negative case). In this situation, if we move  $\mu$  from negative values to positive values, the fixed point changes from attracting to repelling (the fixed point being attracting for  $\mu = 0$ ) while for  $\mu = 0$  a pair of periodic points (one periodic orbit) of period 2 splits off. The dynamics for this negative case is indicated in Figure 3. It is interesting to compare this bifurcation with the Hopf bifurcation for differential equations in Section 4.2.2. This relation can be made explicit in the following



way: in the Hopf bifurcation for differential equations we have (locally) one singularity at  $x(\mu)$  for each value of the parameter  $\mu$ . Consider a line  $\ell_\mu$ , depending smoothly on  $\mu$ , through  $x(\mu)$ . Then we obtain a (locally defined) 1-parameter family of diffeomorphisms in these lines  $\ell_\mu$  by mapping  $x(\mu)$  to itself and mapping each other point in  $\ell_\mu$  to the first intersection of the positive evolution, starting at that point, with  $\ell_\mu$ . (It is not completely trivial to prove that in this way one indeed obtains local diffeomorphisms in  $\ell_\mu$ .) This 1-parameter family of 1-dimensional diffeomorphisms has a period doubling bifurcation. This means that also for period doubling bifurcations we distinguish between supercritical (period 2 orbit attracting) and sub-critical (period 2 orbit repelling).

Finally, in order to explain the name of this bifurcation, we observe also that periodic points of period  $k$  can undergo such a bifurcation, as fixed points of the  $k$ -th power of the original diffeomorphism. In that case a periodic orbit with period  $2k$  (*double* the period) splits off a periodic orbit with period  $k$ .

**Hopf Neïmark Sacker bifurcation** Here we deal with a 1-parameter family of 2-dimensional diffeomorphisms  $\varphi_\mu$ . We assume that for  $\mu = 0$  we have a fixed point at  $x = 0$  and that the eigenvalues of  $d\varphi_0(0)$  form a complex conjugate pair on the unit circle, different from  $\pm 1$ ; we even assume that these eigenvalues are no  $k$ -th root of unity for  $k \leq 4$ . Also this bifurcation is related to the Hopf bifurcation in section 4.2.2: Taking the time  $t$  maps of the evolution maps corresponding to a Hopf bifurcation for differential equations, one obtains a 1-parameter family of diffeomorphisms with a Hopf Neïmark Sacker bifurcation (at least if one avoids exceptional  $t$ -values for which the eigenvalues of the derivative at the fixed point belong to the set of excluded roots of unity). Also here we have to impose generic conditions on the higher order terms. These are rather analogous to the case of the Hopf bifurcation for differential equations. So we first observe that, due to the implicit function theorem, there is a smooth family of fixed points  $x(\mu)$  of  $\varphi_\mu$  (with  $x(0) = 0$ ). We denote the eigenvalues of  $d\varphi_\mu(x(\mu))$  by  $\lambda(\mu)$  and  $\bar{\lambda}(\mu)$ . So  $|\lambda(0)| = 1$ . The first condition we impose on the higher order terms is that  $\partial_\mu |\lambda(0)| \neq 0$ . For the next condition we note that, since the eigenvalue  $\lambda(0)$  is not a  $k$ -th root of unity for  $k \leq 4$ , it can be shown that under generic conditions (and they form the second condition we impose) on the higher order terms (up to order 3) of  $\varphi_0$  at 0, the origin becomes an attractor or repeller for  $\varphi_0$ . This means that  $\varphi^n(x) \rightarrow 0$  for all  $x$  which are sufficiently close to the origin, either as  $n \rightarrow +\infty$  (attracting case) or as  $n \rightarrow -\infty$  (repelling case).

As before there are four cases that can be transformed into each other by reversing the time and/or the  $\mu$ -axis. So we may be restricted to the case where 0 is an attracting fixed point for  $\varphi_0$  and that the derivative of the norm of the eigenvalue  $\lambda(\mu)$  at the bifurcation is positive. Then the fixed point changes from attracting to repelling when  $\mu$  moves from negative to positive values. For  $\mu = 0$ , an attracting invariant closed curve is splitting off (if  $\mu$  moves to positive values). (This is the *supercritical* case as opposed to the *subcritical* case where the invariant closed curve is repelling.) By this we mean that for small positive values of  $\mu$  there is a closed invariant curve  $C_\mu$  around the fixed point  $x(\mu)$ , i.e.  $\varphi_\mu(C_\mu) = C_\mu$  (meaning that the curve is invariant), which is attracting in the sense that for points  $x$ , sufficiently close to  $C_\mu$ , the distance from  $\varphi_\mu^n(x)$  to  $C_\mu$  converges to 0 as  $n \rightarrow \infty$ . The diameter of the invariant curve grows like  $\sqrt{\mu}$ , just as the periodic orbit in the case of the Hopf bifurcation for differential equations. The dynamics on the

closed invariant curve  $C_\mu$  in general depends in a very complicated way on the parameter  $\mu$ . Consequently no general description, up to topological conjugacy or equivalence, of the dynamics is possible in a full neighbourhood of  $(0, 0)$  in the  $(x, \mu)$ -space. Still one can prove that the only real complications of the dynamics are due to the dynamics in the closed curves: there is a neighbourhood  $U$  of 0 in  $\mathbb{R}^2$  and a neighbourhood  $V$  of 0 in  $\mathbb{R}$  such that for each  $x \in U$  and  $\mu \in V$  there are only the following five possibilities for the evolution  $x_n = \varphi_\mu^n(x)$ :

1.  $x_n = x = x(\mu)$ , and we have the fixed point evolution;
2.  $x_n \in C_\mu$ , and we have an evolution in  $C_\mu$  (this happens only for  $\mu > 0$ );
3.  $x_n \rightarrow x(\mu)$  as  $n \rightarrow -\infty$  and for some  $\tilde{x} \in C_\mu$ , the distance between  $x_n$  and  $\varphi_\mu^n(\tilde{x})$  converges to 0 as  $n \rightarrow \infty$  (this happens only for  $\mu > 0$ );
4.  $x_n$  leaves  $U$  for negative  $n$  and there is some  $\tilde{x} \in C_\mu$  such that the distance between  $x_n$  and  $\varphi_\mu^n(\tilde{x})$  converges to 0 as  $n \rightarrow \infty$  (this happens only for  $\mu > 0$ );
5.  $x_n$  leaves  $U$  for negative  $n$  and converges to  $x(\mu)$  as  $n \rightarrow \infty$  (this happens only for  $\mu < 0$ ).

The analysis of this bifurcation was carried out first by Neĭmark [45] and Sacker [58]; see also [38].

The dynamics in the invariant closed curves is related with quasi-periodic motions and resonances; the historical reference is here [2], see also [14]. In order to study the dynamics on these invariant closed curves in relation to resonance phenomena, one has to consider families depending on at least two parameters, see also the Chapter on KAM theory in this volume.

#### 4.4. Concluding remarks

In this section on generic local bifurcations, we have only given explicit descriptions of those local bifurcations of singularities and fixed points which appear in generic 1-parameter families. the situation becomes very different if we consider the generic bifurcations in 1-parameter families of dynamical systems which preserve some extra structure, symplectic or volume preserving systems, see [44], respectively [12]. At this moment there exists a good understanding of those bifurcations which appear in generic 2-parameter families. The number of different cases, however, is too big for a complete exposition in this chapter. A good general reference is [38], in which many computational aspects related with these bifurcations are also discussed. Many of these bifurcations do not admit a complete description up to conjugacy or equivalence; we have seen this phenomenon already in the case of the Hopf Neĭmark Sacker bifurcation, where the dynamics in the closed invariant curve provoked this problem. With two and more parameters these problems become much more complex, and even seem to exhibit new aspects each time one more parameter is added.

All results concerning local bifurcations at fixed points of diffeomorphisms carry over to corresponding bifurcations of periodic evolutions of systems with time set  $\mathbb{R}$  through the use of (local) Poincaré maps. These bifurcations are not strictly local any more, since they are not concentrated in a small neighbourhood of one point. Bifurcations of

periodic point of diffeomorphisms, which are also not completely local, can be treated as bifurcations of fixed points, since these periodic points are fixed points of an iteration of the diffeomorphism.

Considering non-local bifurcations, we meet the homoclinic and heteroclinic bifurcations, i.e. bifurcations involving non-transversal intersections of stable and unstable manifolds of singularities, fixed points or periodic orbits, see [50,38]. These will be discussed in subsequent chapters and are still the subject of ongoing research; see however also the example of moduli in the next section. Other examples of non-local bifurcations are quasi-periodic Hopf bifurcations, where an (attracting)  $(n + 1)$ -torus splits off an  $n$ -torus, and other quasi-periodic bifurcations, see [15]. These bifurcations are closely related to resonance bifurcations and are discussed in the Chapter on KAM theory in this volume.

## 5. Structural stability and moduli

The work on moduli, or moduli of stability, in dynamical systems should be understood in a historical context. For this we first have to discuss the notion of structural stability. We give the relevant definitions below, but first explain the general idea. The main goal was to find a good equivalence relation in the set of dynamical systems. Equivalent systems should be equal from the *qualitative* or topological point of view (they should have the same ‘structure’). On the other hand ‘generic systems’, or at least ‘many’ systems should be interior points in their equivalence class (so that their structure does not change if the system is perturbed slightly). Systems which are in the interior of their equivalence class are then called *structurally stable* (with respect to that equivalence relation). In order to avoid confusion, we here already mention that no (useful) equivalence relation has been found for which structural stability is generic.

**Structural stability** There are several equivalence relations in the space of dynamical systems:  $C^\ell$  conjugacies and equivalences for  $\ell \geq 0$  (we recall the definitions below). For most of these equivalence relations however, equivalence classes will not, or only very exceptionally, have interior points, so that practically no system will be structurally stable with respect to that equivalence relation. This is what we will discuss first, since it gives us motivation for considering certain equivalence relations as more important than others. Here we deal with dynamical systems whose state space is a (compact) manifold, though for the main part one may think also of  $\mathbb{R}^n$ . As time sets we only consider  $\mathbb{R}$  and  $\mathbb{Z}$  and we assume that the evolution maps are generated by a smooth vector field or a smooth diffeomorphism. The dynamical system is said to be  $C^k$  if the generating vector field or diffeomorphism is  $C^k$ . We only consider  $C^k$  dynamical systems with  $k > 0$ . If we have two dynamical systems of the same type, i.e. with the same state space and the same time set, with evolution maps  $\Phi$  and  $\tilde{\Phi}$  then we say that these dynamical systems are  $C^\ell$  conjugate if there is a  $C^\ell$  diffeomorphism (for  $\ell = 0$  a homeomorphism)  $h$  from the state space to itself such that  $h(\Phi(x, t)) = \tilde{\Phi}(h(x), t)$  for all  $x$  in the state space and  $t$  in the time set. This defines an equivalence relation. However the equivalence classes (usually) contain no interior points, at least if  $\ell > 0$ . This can be seen as follows: if  $x_0$  is a stationary point for  $\Phi$ , and if  $h$  is a  $C^\ell$  conjugacy,  $\ell > 0$ , from  $\Phi$  to  $\tilde{\Phi}$ , then  $h(x_0)$  is a stationary point for  $\tilde{\Phi}$

and the eigenvalues of the derivatives of the generating vector fields or diffeomorphisms at  $x_0$  and  $h(x_0)$  have to be the same. This can be shown by the chain rule. Since it is possible to perturb these eigenvalues in an arbitrary direction by a perturbation which is arbitrarily small in the  $C^k$  sense, the equivalence class cannot have interior points (if there is at least one stationary point). The way to avoid this is to use  $C^0$  conjugacies, i.e. homeomorphisms. In this case we have indeed something like ‘local structural stability’ at hyperbolic stationary points. This is a consequence of the Hartman Grobman linearization theorem [27,30], as mentioned in Section 2.3.1. This is the reason that we mainly consider  $C^0$  conjugacies, also called *topological conjugacies*.

As already mentioned before (see the discussion on the Hopf bifurcation), a conjugacy maps periodic evolutions to periodic evolutions with exactly the same period. In the case when the time set is  $\mathbb{R}$ , the period can be any (positive) real number. Such a period can be changed by a perturbation which is arbitrarily small in the  $C^k$  sense. So for dynamical systems with time set  $\mathbb{R}$  the equivalence class of a system containing at least one periodic orbit has no interior points. This is the reason why, whenever the time set is  $\mathbb{R}$ , we only consider *topological equivalences*: a topological equivalence between  $\Phi$  and  $\tilde{\Phi}$  is a homeomorphism  $h$  such that, for any  $x$  and  $x'$  in the state space,  $\Phi(x, t) = x'$  holds for some positive  $t$ , if and only if  $\tilde{\Phi}(h(x), \tilde{t}) = h(x')$  holds for some positive  $\tilde{t}$ . With this equivalence relation, the Hartman Grobman linearization theorem also implies ‘local structural stability’ near hyperbolic periodic evolutions.

With the above definitions, we have a good geometric description of the  $C^k$  dynamical systems,  $k > 0$ , which are structurally stable: they are systems which satisfy strong hyperbolicity and transversality conditions. The central papers in this area, which rather belong to volume 1 of this series, are [54,57,41]. It was already noted a long time ago that, even with these equivalence relations, structural stability is not a generic property. The first example (as far as we know) of an open set of dynamical systems, none of which is structurally stable, is due to Smale, see appendix 24 in [7], where the authors refer to unpublished notes by Smale.

**Bifurcations and moduli** In the section on bifurcations we have already met non-structurally stable dynamical systems. We recall the saddle node which occurs in the 1-parameter family of differential equations  $x' = f_\mu(x) = \mu + x^2$ . Though, for  $\mu = 0$ , this system is not structurally stable, the whole family is (locally) structurally stable: if we perturb  $f_\mu$  slightly, we can even conjugate the original family with the perturbed family (in a neighbourhood of  $\mu = 0$  and  $x = 0$ ). Of course, given the fact that structural stability is not generic, it would not be realistic to hope that for generic bifurcations in  $k$  parameter families of dynamical systems we always have this form of structural stability. However in some cases this type of structural stability is violated in such a way that still a very good description, up to topological conjugacy or equivalence, can be given. This is the case where we have moduli of stability. We shall describe one such example (without proof) and then give references to some key papers and later developments, without even trying to be complete.

We consider a 1-parameter family of 2-dimensional diffeomorphisms  $\varphi_\mu$  with fixed points  $p_\mu$  and  $q_\mu$  of saddle type and such that the derivatives of  $\varphi_\mu$  at these fixed points have positive eigenvalues;  $\ell_\mu$  and  $m_\mu$  are finite arcs in  $W^u(p_\mu)$ , respectively,

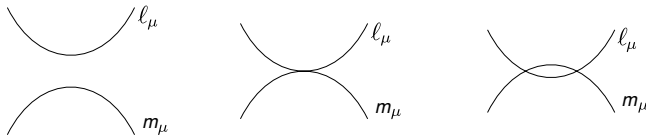


Fig. 4. Before, at, and after the tangency of separatrices.

$W^s(q_\mu)$ , containing  $p_\mu$ , respectively,  $q_\mu$ . We assume that  $p_\mu$ ,  $q_\mu$ ,  $\ell_\mu$ , and  $m_\mu$  all depend differentiably on the parameter  $\mu$ . We also assume that for  $\mu = 0$  there is a bifurcation in the sense that  $\ell_0$  and  $m_0$  have a point of tangency (and no other intersections). The arcs, in  $\ell_0$  and  $m_0$  from  $p_0$ , respectively  $q_0$ , to the point of tangency are denoted by  $\tilde{\ell}$ , respectively  $\tilde{m}$ . Such bifurcations occur in generic 1-parameter families of diffeomorphisms; generically (possibly up to reversing the direction of the  $\mu$ -axis), one then has for  $\mu < 0$  two intersections of  $\ell_\mu$  and  $m_\mu$ , and for  $\mu > 0$  no such intersections, at least for small  $|\mu|$ . We assume the tangency of  $\ell_0$  and  $m_0$  to be parabolic, i.e. locally diffeomorphic, to the pair of curves  $\{y = 0\}$  and  $\{y = x^2\}$  in the plane; also this is a generic condition for such bifurcations in the plane. The situation is indicated in Figure 4. Let now  $\lambda_\mu$  be the contracting eigenvalue of  $d\varphi_\mu$  at  $p_\mu$ , and  $\sigma_\mu$  be the expanding eigenvalue of  $d\varphi_\mu$  at  $q_\mu$ . Then  $M = \log(\lambda_0)/\log(\sigma_0)$  is called a modulus of stability in the following sense.

Let  $\varphi'_\mu$  be another 1-parameter family of 2-dimensional diffeomorphisms satisfying the same specifications. We denote the objects, similar to those for the first family  $\varphi_\mu$ , by the same symbol with an accent. So for this second family we have the curves  $\tilde{\ell}'$  and  $\tilde{m}'$ , the (positive) eigenvalues  $\lambda'_\mu$  and  $\sigma'_\mu$  and the modulus of stability  $M' = \log(\lambda'_0)/\log(\sigma'_0)$ . The reason that we call these quantities  $M$  and  $M'$  moduli of stability is that there is a local conjugacy from a neighbourhood of  $\tilde{\ell} \cup \tilde{m}$  to a neighbourhood of  $\tilde{\ell}' \cup \tilde{m}'$ , conjugating  $\varphi_0$  locally to  $\varphi'_0$ , if and only if  $M = M'$ . This was actually the first example of a modulus of stability for diffeomorphisms, see [48,21].

For a general discussion of moduli, due to tangencies of stable and unstable manifolds in generic 1-parameter families, we refer to [71]. Moduli of stability have also been found in relation to saddle-node bifurcations of diffeomorphisms (in combination with other separatrices), see [46] and even for singularities of gradient vector fields [68].

## References

- [1] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, 2nd edn, Benjamin (1978).
- [2] V.I. Arnold, *Small denominators I on the mapping of the circle to itself*, Transl. Amer. Math. Soc., Serie 2 **46** (1965), 213–284.
- [3] V.I. Arnold, *Méthodes mathématiques de la mécanique classique*, Mir (1976); English translation: *Mathematical methods of classical mechanics* (translation by K. Vogtmann, A. Weinstein), Springer-Verlag, 1989.
- [4] V.I. Arnold, *Catastrophe Theory*, Springer-Verlag (1986).
- [5] V.I. Arnold, ed., *Bifurcation theory and catastrophe theory*, Dynamical Systems V, Encyclopaedia of Mathematical Sciences, Vol. 5, Springer-Verlag (1986).
- [6] V.I. Arnold, ed., *Singularity theory I and II*, Dynamical Systems VI and VIII, Encyclopaedia of Mathematical Sciences, Vol. 6 respectively Vol. 39, Springer-Verlag (1993).

- [7] V.I. Arnold and A. Avez, *Problèmes ergodiques de la mécanique classique*, Gautier-Villars (1967); English translation: *Ergodic problems in classical mechanics* (translation by A. Avez), Benjamin, 1968.
- [8] V.I. Arnold, S.M. Gusein-Zade and A.N. Varchenko, *Singularities of Differentiable Maps* (2 vol.), Birkhäuser (1985–1988).
- [9] V.I. Arnold and M.B. Sevryuk, *Oscillations and bifurcations in reversible systems*, Nonlinear Phenomena in Plasma Physics and Hydrodynamics, R.Z. Sagdeev, ed., Mir, Moscow (1986), 31–64.
- [10] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer-Verlag (1975).
- [11] Th. Bröcker, *Differentiable Germs and Catastrophes*, Cambridge University Press (1975).
- [12] H.W. Broer, *Formal normal form theorems for vector fields and some consequences for bifurcations in the volume preserving case*, Lecture Notes in Math., Vol. 898, Springer (1981), 54–89.
- [13] H.W. Broer, F. Dumortier, S.J. van Strien and F. Takens, *Structures in Dynamics*, North-Holland (1991).
- [14] H.W. Broer, G.B. Huitema and M.B. Sevryuk, *Quasi-periodic motions in families of dynamical systems*, Lecture Notes in Math. **1645** (1996).
- [15] H.W. Broer, G.B. Huitema, F. Takens and B.L.J. Braaksma, *Unfoldings and bifurcations of quasi-periodic tori*, Mem. AMS **83** (421) (1990).
- [16] H.W. Broer and F. Takens, *Dynamical Systems and Chaos*, Epsilon Uitgaven, Vol. 64 (2009); Springer Appl. Math. Sci., Vol. 172 (2010).
- [17] H.W. Broer and F.M. Tangerman, *From a differentiable to a real analytic perturbation theory, applications to the Kupka-Smale theorems*, Erg. Th. & Dyn. Syst. **6** (1986), 345–362.
- [18] J. Carr, *Applications of Centre Manifold Theory*, Springer-Verlag (1981).
- [19] S.-N. Chow and J.K. Hale, *Methods of Bifurcation Theory*, Springer-Verlag (1982).
- [20] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic Theory*, Springer-Verlag (1982).
- [21] W.C. de Melo, *Moduli of stability of two-dimensional diffeomorphisms*, Topology **19** (1980), 9–21.
- [22] R.L. Devaney, *Reversible diffeomorphisms and flows*, Trans. Amer. Math. Soc. **218** (1976), 89–113.
- [23] E. Fabes, M. Luskin and G.R. Sell, *Construction of inertial manifolds by elliptic regularization*, J. Differential Equations **89** (1991), 355–387.
- [24] M. Field, *Equivariant dynamical systems*, Trans. Amer. Math. Soc. **259** (1980), 185–205.
- [25] C.G. Gibson, K. Wirthmüller, A.A. du Plessis and E.J.N. Looijenga, *Topological Stability of Smooth Mappings*, Lecture Notes in Mathematics, Vol. 552, Springer-Verlag (1976).
- [26] M. Golubitsky, I. Stewart and D.G. Schaeffer, *Singularities and Groups in Bifurcation Theory* (2-vols), Springer-Verlag (1985–1988).
- [27] D.M. Grobman, *Homeomorphisms of systems of differential equations*, Dokl. Akad. Nauk SSSR **128** (1959), 880–881.
- [28] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag (1983).
- [29] C. Gutierrez, *A counterexample to a  $C^2$  closing lemma*, Erg. Th. & Dyn. Syst. **7** (1987), 509–530.
- [30] P. Hartman, *On local homeomorphisms of Euclidean spaces*, Bol. Soc. Mat. Mexicana **5** (1960), 220–241.
- [31] M. Herman, *Mesure de Lebesgue et nombre de rotation*, Geometry and Topology, Lecture Notes in Mathematics, Vol. 597, Springer-Verlag (1977), 271–293.
- [32] M.W. Hirsch, *Differential Topology*, Springer-Verlag (1976).
- [33] M.W. Hirsch, C.C. Pugh and M. Shub, *Invariant Manifolds*, Lecture Notes in Math., Vol. 583, Springer-Verlag (1977).
- [34] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press (1995).
- [35] J.L. Kelley, *General Topology*, Van Nostrand (1955).
- [36] I. Kupka, *Contribution à la théorie des champs génériques*, Contr. Diff. Equ. **2** (1963), 457–484.
- [37] I. Kupka, *Contribution à la théorie des champs génériques II*, Contr. Diff. Equ. **3** (1964), 411–420.
- [38] Y.A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag (1995).
- [39] B. Malgrange, *Ideals of Differentiable Functions*, Oxford University Press (1966).
- [40] R. Mañé, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag (1987).
- [41] R. Mañé, *A proof of the  $C^1$  stability conjecture*, Publ. Math. IHES **66** (1988), 161–210.
- [42] J.E. Marsden and M.F. McCracken, *The Hopf Bifurcation and its Applications*, Springer-Verlag (1976).
- [43] J. Mather, *Stability of  $C^\infty$  mappings I – VI*, Ann. Math. **87** (1968), 89–104, **89** (1969) 254–291; Publ. Sci. IHES **35** (1968), 127–156, **37** (1969), 223–248; Adv. in Math. **4** (1970), 301–335; Lecture Notes in Math. **192** (1971), 207–253.

- [44] K.R. Meyer, *Generic bifurcations of periodic points*, Trans. Amer. Math. Soc. **149** (1970), 95–107.
- [45] Yu.I. Neimark, *On some cases of dependence of periodic motions on parameters*, Dokl. Akad. Nauk SSSR **129** (1959), 736–739.
- [46] S.E. Newhouse, J. Palis and F. Takens, *Bifurcations and stability of families of diffeomorphisms*, Publ. Math. IHES **57** (1983), 5–71.
- [47] J.C. Oxtoby, *Maß und Kategorie*, Springer-Verlag (1971).
- [48] J. Palis, *A differentiable invariant of topological conjugacies and moduli of stability*, Astérisque **51** (1978), 335–346.
- [49] J. Palis and W. de Melo, *Geometric Theory of Dynamical Systems*, Springer-Verlag (1982).
- [50] J. Palis and F. Takens, *Hyperbolicity & Sensitive Chaotic Dynamics at Homoclinic Bifurcations*, Camb. Univ. Press (1993).
- [51] H.-O. Peitgen, ed., *Newton's Method and Dynamical Systems*, Kluwer (1989).
- [52] C. Pugh, *The closing lemma*, Amer. J. Math. **89** (1967), 956–1009.
- [53] C. Pugh, *An improved closing lemma and a general density theorem*, Amer. J. Math. **89** (1967), 1010–1021.
- [54] J.W. Robbin, *A structural stability theorem*, Ann. of Math. **94** (1971), 447–493.
- [55] C. Robinson, *Dynamical Systems*, CRC Press (1995).
- [56] R.C. Robinson, *Generic properties of conservative systems, I and II*, Amer. J. Math. **92** (1970), 562–603, 879–906.
- [57] R.C. Robinson, *Structural stability of  $C^1$  flows*, Dynamical Systems – Warwick 1974, Lecture Notes in Mathematics, Vol. 468, Springer-Verlag (1975), 262–277.
- [58] R. Sacker, *A new approach to the perturbation theory of invariant surfaces*, Comm. Pure Appl. Math. **18** (1965), 717–732.
- [59] D.H. Sattinger, *Branching in the presence of symmetry*, CBMS-NSF Reg. Conf. Ser. Appl. Math., Vol. 40, SIAM (1983).
- [60] M.B. Sevryuk, *Reversible Systems*, Lecture Notes in Mathematics, Vol. 1211, Springer-Verlag (1986).
- [61] C.L. Siegel and J.K. Moser, *Lectures on Celestial Mechanics*, Springer-Verlag (1971).
- [62] S. Smale, *Stable manifolds for differential equations and diffeomorphisms*, Ann. Scuola Sup. Pisa **18** (1963), 97–116.
- [63] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.
- [64] S. Sternberg, *Local contractions and a theorem of Poincaré*, Amer. J. Math. **79** (1957), 809–824.
- [65] S. Sternberg, *On the structure of local homeomorphisms of Euclidean  $n$ -space, II*, Amer. J. Math. **80** (1958), 623–631.
- [66] F. Takens, *Hamiltonian systems, generic properties of closed orbits and local perturbations*, Math. Ann. **188** (1970), 304–312.
- [67] F. Takens, *Homoclinic points of conservative systems*, Inv. Math. **18** (1972), 267–292.
- [68] F. Takens, *Singularities of gradient vector fields and moduli*, Singularities and Dynamical Systems, North-Holland (1985), 81–88.
- [69] R. Thom, *Un lemme sur les applications différentiables*, Bol. Soc. Mat. Mex. **2** (1956), 59–71.
- [70] R. Thom, *Stabilité structurelle et morphogénèse*, Benjamin (1972).
- [71] S.J. van Strien, *Saddle connections of arcs of diffeomorphisms: moduli of stability*, Dynamical Systems and Turbulence, Warwick 1980, Lecture Notes in Mathematics, Vol. 898, Springer-Verlag (1981), 352–365.
- [72] E.C. Zeeman, *The classification of elementary catastrophes of codimension  $\leq 5$* , Structural Stability, the Theory of Catastrophes, and Applications in the Sciences, Springer Lecture Notes in Mathematics, Vol. 525, Springer-Verlag (1976), 263–327.

## CHAPTER 2

# Prevalence

Brian R. Hunt<sup>1</sup>

*University of Maryland, College Park, MD 20742, United States*

Vadim Yu. Kaloshin<sup>2</sup>

*Department of Mathematics, University of Maryland, College Park, MD, 20740, United States*

*Department of Mathematics, Penn State University, University Park, PA, 16801, United States*

### Contents

|                                                                                   |    |
|-----------------------------------------------------------------------------------|----|
| 1. Introduction . . . . .                                                         | 45 |
| 1.1. Examples where generic properties do not hold almost everywhere . . . . .    | 46 |
| 1.2. Examples of properties that hold generically and almost everywhere . . . . . | 49 |
| 2. Linear prevalence . . . . .                                                    | 51 |
| 2.1. Properties of prevalence . . . . .                                           | 55 |
| 2.2. Proving prevalent results . . . . .                                          | 59 |
| 2.3. Examples of linear prevalence . . . . .                                      | 60 |
| 2.4. Open problems: Generic results on linear spaces . . . . .                    | 66 |
| 3. Nonlinear prevalence . . . . .                                                 | 68 |
| 3.1. Nonlinear prevalence without an underlying linear structure . . . . .        | 69 |
| 3.2. Nonlinear prevalence with an underlying linear structure . . . . .           | 70 |
| 3.3. Examples of nonlinear prevalence . . . . .                                   | 72 |
| 3.4. Open problems: generic results in nonlinear spaces . . . . .                 | 74 |
| 4. Other notions of genericity . . . . .                                          | 78 |
| 4.1. Non-Abelian prevalence . . . . .                                             | 78 |
| 4.2. Relative prevalence and applications in mathematical economics . . . . .     | 79 |
| 4.3. Stronger forms of genericity and prevalence . . . . .                        | 80 |
| References . . . . .                                                              | 81 |

---

<sup>1</sup>The first author was partially supported by NSF Grant No. DMS0616585.

<sup>2</sup>The second author was partially supported by AIM and Sloan fellowships and NSF Grant No. DMS-0300229.





## 1. Introduction

This article surveys results and conjectures in dynamical systems and other areas that describe properties of ‘almost every’ function in some space, using a probabilistic (or measure-theoretic) notion called ‘prevalence’, which we define for complete metric linear spaces in Section 2. In many cases these properties coincide with the properties of ‘generic’ functions in the space, using the more classical notion that a property on a complete metric space (or more generally, a Baire space) is called *generic* if it is true on a countable intersection of open dense subsets. However, in many cases properties that are generic in the topological sense are not prevalent in the measure-theoretic sense, and vice-versa. Indeed, open dense subsets of  $\mathbb{R}^d$  can have arbitrarily small Lebesgue measure; see Section 1.1 for other examples.

While there have been many results on generic properties in function spaces, which we do not attempt to survey here, mathematicians have desired an analogue of the finite-dimensional notions of ‘Lebesgue almost every’ and ‘Lebesgue measure zero’ in infinite-dimensional function spaces. In his contribution to the 1954 International Congress of Mathematicians [97], Kolmogorov wrote:

In order to obtain negative results concerning insignificant or exceptional character of some phenomenon, we shall apply the following somewhat haphazard technique: if in a class  $K$  of functions  $f(x)$  one can introduce a finite number of functionals

$$F_1(f), F_2(f), \dots, F_r(f)$$

that in some sense can naturally be considered as taking ‘generally arbitrary’ values

$$F_1(f) = C_1, F_2(f) = C_2, \dots, F_r(f) = C_r$$

from some region in the  $r$ -dimensional space of points  $C = (C_1, \dots, C_r)$ , then any phenomenon that can take place only if  $C$  belongs to a set of zero  $r$ -dimensional measure will be regarded as exceptional and subject to ‘neglect’.

Arnol’d and Anosov also wrote of the need for a probabilistic perspective in studying dynamical systems; see Example 2 of Section 1.1 and the beginning of Section 2.

Prevalence generalizes and formalizes this notion, which is based on a finite-dimensional parameterization of the function space in which the property of interest occurs for a ‘Lebesgue almost every’ or a ‘Lebesgue measure zero’ set of parameters. In Section 1.1, we give some examples of the difference between the notions of ‘Lebesgue almost every’ and topological genericity. In Section 2, we define prevalence and related notions. In Section 2.1, we discuss fundamental properties of prevalence, including its equivalence to ‘Lebesgue almost every’ in  $\mathbb{R}^d$ . In Section 2.2, we give some simple examples of proving results with prevalence and discuss a prevalent transversality theorem from which many results follow. In Section 3, we discuss extensions of prevalence to nonlinear spaces. Sections 2 and beyond discuss results and conjectures from various areas that use prevalence or closely related notions. See also [74] and [135] for surveys of results involving prevalence.

### 1.1. Examples where generic properties do not hold almost everywhere

In  $\mathbb{R}^d$ , many properties are both topologically generic and hold Lebesgue almost everywhere. For example, Lebesgue almost every real number is irrational, and the set of irrational numbers is the intersection over all rational numbers  $x$  of the complement of  $\{x\}$ ; thus, it is a countable intersection of open dense sets. In Section 1.2 we give several examples where these notions of typicality coincide. However, there are many important cases in which properties that hold generically differ from those that hold Lebesgue almost everywhere.

As we have mentioned, in  $\mathbb{R}^d$  there are open dense sets with arbitrarily small Lebesgue measure. To see this, consider a countable dense set  $\{x_1, x_2, \dots\} \in \mathbb{R}^d$ , and for  $k = 1, 2, \dots$ , put an open ball with Lebesgue measure  $\varepsilon/2^k$  around  $x_k$  for some  $\varepsilon > 0$ . The union of these balls is open and dense, but has measure less than  $\varepsilon$ . Furthermore, taking the intersection of such sets over a sequence of  $\varepsilon$  values tending to zero, we get a countable intersection of open dense sets with Lebesgue measure zero. In other words, a property can be topologically generic but hold only on a set with Lebesgue measure zero.

While the construction above is artificial, there are many examples of naturally arising properties that are generic but occur with probability zero in the sense of Lebesgue measure. Some of the examples in this section are taken from [74]. Our first example, concerning how well a typical real number can be approximated by rational numbers, is similar to the construction above, with the rational numbers being the countable dense set.

**EXAMPLE 1 (Diophantine Numbers).** A real number  $x$  is said to have *Diophantine exponent*  $\gamma$  if the inequality  $|qx - p| < |q|^{-\beta}$  has infinitely many integer solutions  $(p, q)$  for  $\beta < \gamma$  but only finitely many solutions for  $\beta > \gamma$ . Thus, the larger a number's Diophantine exponent, the better it can be approximated by rational numbers. Irrational numbers with infinite Diophantine exponent are called *Liouville numbers*. Liouville numbers are generic in the real numbers but have Lebesgue measure zero [137]. To see this, let

$$S_{k,\beta} = \left\{ x \in \mathbb{R} : \left| x - \frac{p}{q} \right| < \frac{1}{q^{1+\beta}} \text{ for some } p, q \in \mathbb{Z} \text{ with } q > k \right\}.$$

Then the Liouville numbers are the intersections over all positive integers  $k$  and  $\beta$  of  $S_{k,\beta}$  (minus the rational numbers), and each  $S_{k,\beta}$  is open and dense. On the other hand, within an interval of length  $L$ , the Lebesgue measure of  $S_{k,\beta}$  is at most

$$\sum_{q=k+1}^{\infty} \frac{2qL}{q^{1+\beta}} < \frac{2L}{(\beta-1)k^{\beta-1}}$$

for  $\beta > 1$ . Thus, the Liouville numbers have Lebesgue measure zero. In fact, this estimate implies that Lebesgue almost every real number has Diophantine exponent 1.

**EXAMPLE 2 (Circle Diffeomorphisms).** In the space of orientation-preserving circle diffeomorphisms, those that are structurally stable (their dynamics is essentially unchanged by small perturbations) form an open dense set, consisting of those diffeomorphisms with a rational rotation number and no nonhyperbolic periodic orbits; see, for example, the book

by Arnol'd [12]. After presenting this result, Arnol'd comments:

The preceding theorems give the impression that a generic diffeomorphism of the circle has rational rotation number, and diffeomorphisms with irrational rotation number are exceptional. Nevertheless, numerical experiments usually lead to (at least apparently) everywhere dense orbits. To explain this phenomenon, we consider, for example, the family of the diffeomorphisms

$$A_{\alpha,\varepsilon} : y \mapsto y + \alpha + \varepsilon \sin y, \quad \alpha \in [0, 2\pi], \quad \varepsilon \in [0, 1).$$

We shall represent every diffeomorphism by a point in the  $(\alpha, \varepsilon)$ -plane. As is easily seen, the set of diffeomorphisms with rotation number  $\mu = p/q$  is bounded by a pair of smooth curves and approaches the axis  $\varepsilon = 0$  with increasingly narrow tongues as  $q$  increases. The union of these sets is dense. Nevertheless, it turns out that the measure of the set of points of the parameter plane for which the rotation number is rational, is small in the domain  $0 \leq \varepsilon \leq \varepsilon_0$ ,  $0 \leq \alpha \leq 2\pi$ , compared to the measure of the whole domain.

Consequently, a diffeomorphism chosen randomly from our family with small  $\varepsilon$  has irrational rotation number with great probability.

Moreover, an analogous result holds for any analytic or sufficiently smooth family of diffeomorphisms that are close to rotations; for example, for the family  $y \mapsto y + \alpha + \varepsilon a(y)$  with an arbitrary analytic function  $a$ : for small  $\varepsilon$ , the orbits are everywhere dense on the circle and the rotation number is irrational with a preponderant probability.

Consequently, the idea of structural stability is not the only approach to the notion of a generic system. The metric approach indicated above is more appropriate for the description of the actually observable behaviour of the system in some cases.

See Figure 1 for a picture of some of the tongues described by Arnol'd. In the setting above, the fact that the Lebesgue measure of  $\alpha$  for which the rotation number is rational approaches zero as  $\varepsilon \rightarrow 0$  was proved by Arnol'd in [11]. This result was generalized to diffeomorphisms with finite smoothness ( $C^3$ ) by Herman [67], who also proved that for an arbitrary 1-parameter family of  $C^3$  diffeomorphisms with  $C^1$  parameter dependence, if the rotation number is not identically constant then the parameters with irrational rotation number have positive Lebesgue measure. Thus for such families, the dynamical behaviour that is topologically generic (stable periodicity) does not have full Lebesgue measure. (On the other hand, Świątek [173] proved that for certain families of  $C^3$  homeomorphisms with a critical point, including the family  $A_{\alpha,\varepsilon}$  with  $\varepsilon = 1$ , the set of parameters with irrational rotation number has Lebesgue measure zero.)

**EXAMPLE 3 (Linearizability of Neutral Fixed Points).** A complex analytic map with a neutral fixed point at the origin can be written

$$z_{n+1} = e^{2\pi i \alpha} z + z^2 f(z)$$

with  $\alpha$  real and  $f$  analytic. Cremer [36] proved that if  $f$  is a nonzero polynomial, then for a generic set of  $\alpha$ , the map above is not conjugate to a rotation in any neighbourhood

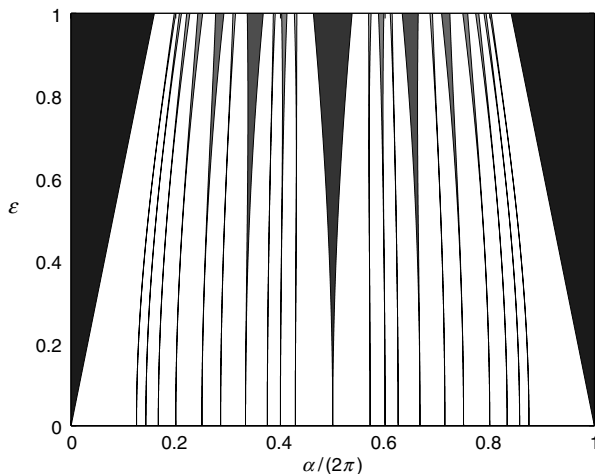


Fig. 1. Arnol'd tongues with rotation numbers  $p/q$  for  $q \leq 8$ . The shaded set that approaches  $\alpha/(2\pi) = p/q$  as  $\varepsilon \rightarrow 0$  is the set of parameters  $(\alpha, \varepsilon)$  for which  $A_{\alpha, \varepsilon}$  has rotation number  $p/q$ .

of the origin. On the other hand, Siegel [162] proved that for Lebesgue almost every  $\alpha$  (namely,  $\alpha$  not a Liouville number), the map above is analytically conjugate to a rotation in a neighbourhood of the origin. See [23, pp. 98–105] for further discussion of these results.

**EXAMPLE 4 (Transitivity of the Complex Exponential Map).** Misiurewicz [124] proved that the map  $z_{n+1} = \exp(z_n)$  on the complex plane is topologically transitive, which implies that a topologically generic initial condition has a dense orbit. On the other hand, Lyubich [102] and Rees [148] proved that Lebesgue almost every initial condition has a trajectory whose limit set lies on the real axis. See [103] for a discussion of both results.

**EXAMPLE 5 (Law of Large Numbers).** Consider an infinite sequence  $X_1, X_2, \dots$  of random variables that are uniformly distributed in the interval  $[0, 1]$ . By the strong law of large numbers, with probability one, the averages  $A_n = (X_1 + X_2 + \dots + X_n)/n$  converge to  $1/2$  as  $n \rightarrow \infty$ . (The underlying measure here is not, strictly speaking, Lebesgue measure, but rather the infinite product measure on  $[0, 1]^\omega$  derived from Lebesgue measure on  $[0, 1]$ .) However, using the product topology on  $[0, 1]^\omega$ , the set of sequences  $\{X_n\}$  for which the sequence of averages  $A_n$  diverges is generic. This follows from the fact that the set  $S_k$  of sequences for which  $A_m < 1/4$  and  $A_n > 3/4$  for some  $m, n \geq k$  is open and dense. In a similar manner, one can show that generically  $\limsup A_n = 1$  and  $\liminf A_n = 0$  as  $n \rightarrow \infty$ . We learned this example from [127].

**EXAMPLE 6 (Diophantine Properties of  $SO(3)$ ).** Denote the angle of rotation of an element  $A$  of  $SO(3)$  by  $\angle A$ . For  $A, B \in SO(3)$ , let  $W_n(A, B)$  be the set of reduced words of length  $n$  in  $A, A^{-1}, B$ , and  $B^{-1}$ . Define

$$e_n(A, B) = \min \angle W_n(A, B).$$

We say that a pair  $(A, B)$  is Diophantine if for some positive  $C$  and  $\tau$  we have  $e_n(A, B) > C\tau^n$  for all  $n \geq 1$ . One can show that a topologically generic pair  $(A, B) \in SO(3) \times SO(3)$  is not Diophantine even for  $C\tau^n$  replaced by  $C\tau^{n^2}$ . Kaloshin and Rodnianski [85] showed that for almost every pair we have  $e_n(A, B) > C\tau^{n^2}$  for some  $C, \tau > 0$ .

**EXAMPLE 7** (Dynamics of Unimodal Maps). A map of an interval  $f : I \rightarrow I$  is called *unimodal* if it has a unique critical point, which we will assume here to be quadratic. The simplest and most famous example is the real quadratic family  $f_\lambda : [0, 1] \rightarrow [0, 1]$  given by

$$f_\lambda(x) = \lambda x(1 - x),$$

for  $0 \leq \lambda \leq 4$ . Milnor and Thurston [123] showed that the quadratic family is qualitatively universal, in the sense that every unimodal map has essentially the same dynamics as some  $f_\lambda$ . Furthermore, Guckenheimer [61] proved that every unimodal map with negative Schwarzian derivative is topologically conjugate to some  $f_\lambda$ .

A unimodal map is called *regular* if its critical point belongs to the basin of a hyperbolic periodic attractor (that is, its forward orbit approaches an attracting periodic point), and all of its periodic points are hyperbolic. It is called *stochastic* if it has an invariant measure that is absolutely continuous with respect to Lebesgue measure. A unimodal map cannot be both regular and stochastic (see e.g. [77]).

Jakobson [77] proved that the quadratic map  $f_\lambda$  is stochastic for a positive Lebesgue measure set of  $\lambda$ , and that the same is true for families  $C^2$  close to  $f_\lambda$ . On the other hand, Graczyk and Świątek [60] and Lyubich [104] proved that  $f_\lambda$  is regular for an open dense set of  $\lambda$ . In the complement of this set, Lyubich [105] showed that  $f_\lambda$  is stochastic for Lebesgue almost every  $\lambda$ .

Extension of these results to any non-trivial family of real analytic unimodal maps was done for regular parameters by Kozlovski [98] and for stochastic parameters by Avila, Lyubich, and de Melo [18]. Note that these results fit nicely to the general program of studying attractors in finite parameter families of dynamical systems (in all dimensions) formulated by Palis [140]. See also Section 3.4.4.

## 1.2. Examples of properties that hold generically and almost everywhere

Here are several examples where similar phenomena are known (or appear) to hold for topologically generic parameters and for Lebesgue almost every parameter.

**EXAMPLE 8** (Projections of Finite-dimensional Sets). Mañé ([106, Lemma 1.1]; see also [49, p. 627], or [75, Section 2.2]) proved that if  $S$  is a compact subset of a Banach space  $B$  with box-counting dimension  $d$ , then if  $n > 2d + 1$ , a generic linear projection from  $B$  onto a subspace of dimension  $n$  is one-to-one. (This result is also true with  $2d$  replaced by the Hausdorff dimension of  $S \times S$ . However,  $2d$  cannot always be replaced by twice the Hausdorff dimension of  $S$ , as demonstrated by an example of Kan in an appendix to [154].)

In the case that  $B$  is finite-dimensional, the space of projections onto an  $n$ -dimensional subspace can be identified with an open dense subset of a finite-dimensional space, namely the full rank linear transformations from  $B$  to  $\mathbb{R}^n$  (see Section 2.3.2). Thus a generic linear transformation from  $B$  to  $\mathbb{R}^n$  is one-to-one on  $S$ . Sauer, Yorke, and Casdagli [154] proved that almost every (in the sense of Lebesgue measure for the corresponding matrices) linear transformation from  $B$  to  $\mathbb{R}^n$  is one-to-one on  $S$  for  $n > 2d$ . (This result is a special case of their result for a prevalent smooth function from  $B$  to  $\mathbb{R}^n$ ; see Section 2.3.2.)

EXAMPLE 9 (Kupka–Smale Theorem). Consider a smooth map on a compact manifold. It is called *Kupka–Smale* if all periodic points are hyperbolic and the stable and unstable manifolds of any two saddles intersect transversally. The classical Kupka–Smale theorem (see e.g. [142]) says that a generic  $C^r$  diffeomorphism is Kupka–Smale for  $r \geq 1$ . The Kupka–Smale theorem has been proven in many other infinite-dimensional settings, including smooth endomorphisms [160], real-analytic diffeomorphisms [28,101] and holomorphic automorphisms of class  $C^r$  [29]. In Section 3.3.2 we will discuss a prevalent Kupka–Smale theorem.

Buzzard, Hruska, and Ilyashenko [31] considered a finite-dimensional space of invertible polynomials of given degree in  $\mathbb{C}^2$ , and proved that the Kupka–Smale property is both generic and of full measure. More exactly, define the set  $\mathcal{P}_d$  of all the ‘normalized’ polynomial automorphisms of dynamical degree  $d \geq 2$ . For the precise definition see [31, Section 2.1]. An important irreducible component of  $\mathcal{P}_d$  is the set  $\mathcal{H}_d$  of generalized Henon maps

$$F : (x, y) \rightarrow (y, P(y) - ax),$$

where  $P$  is a degree  $d$  monic polynomial and  $a \neq 0$ . The parameter space for  $\mathcal{H}_d$  is  $\mathbb{C}^d \times (\mathbb{C} - \{0\})$ , while  $\mathcal{P}_d$  is formed by taking the composition of generalized Hénon maps to obtain total degree  $d$ . With this notation, Theorem 1.1 of [31] says that *Kupka–Smale maps in  $\mathcal{P}_d$  form a generic set of full measure*. The same statement is true for the corresponding set of polynomial maps on  $\mathbb{R}^2$  with real coefficients.

EXAMPLE 10 (Polygonal Billiards). For a compact planar domain  $D$  with piecewise smooth boundary, the billiard flow is defined as follows. A particle flows along a straight line with unit speed until it reaches the boundary, where it reflects as light does from a mirror, flowing along an outgoing line that makes an equal angle with the boundary as the incoming line. The state space for this flow consists of ordered pairs  $(x, \theta) \in D \times [0, 2\pi)$ , representing the position and velocity direction of the particle, minus the (Lebesgue measure zero) subset of initial conditions whose trajectories eventually reach a nonsmooth boundary point. It is not hard to show that Lebesgue measure on this state space is invariant under the billiard flow.

Kerckhoff, Masur, and Smillie [92] proved that for a generic polygonal domain, the billiard flow is ergodic. The space of polygons can be thought of as an open subset of a finite-dimensional space, and it is an open problem whether the set of polygons with ergodic billiard flow has full (or even positive) Lebesgue measure [62]. However, the generic result of Kerckhoff–Masur–Smillie is a corollary to the following ‘Lebesgue almost every’ result in the same paper.

Rational polygons (those whose angles are rational multiples of  $\pi$ ) do not have ergodic billiard flow because, for a given initial flow angle  $\theta$ , the future flow directions are confined to a finite set of angles that are rationally related to  $\theta$ . Thus, the state space is foliated by invariant subsets parameterized by  $\theta$ . Kerckhoff–Masur–Smillie [92] showed that for Lebesgue almost every  $\theta$ , the billiard flow restricted to the corresponding invariant subset is uniquely ergodic.

EXAMPLE 11 (Interval Exchange Transformations). Recently ergodic properties of the following beautiful dynamical system have attracted a lot of attention. Let  $d \geq 2$  be a natural number, and let  $\pi$  be an irreducible permutation of  $\{1, \dots, d\}$ , in the sense that  $\pi\{1, \dots, k\} \neq \{1, \dots, k\}$ ,  $1 \leq k \leq d$ . Given  $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}_+^d$  such that  $\sum_j \lambda_j = 1$  we define an *interval exchange transformation* (IET) in the usual way (see e.g. [35]): subdivide the unit interval according to the  $\lambda_j$ :  $I_k = [\sum_{j < k} \lambda_j, \sum_{j \leq k} \lambda_j)$  and permute them according to  $\pi$ . Such a map is clearly Lebesgue measure preserving. Mazur [116] and Veech [182] proved that for every irreducible  $\pi$  and Lebesgue almost every  $\lambda$ , the corresponding IET is uniquely ergodic. Katok [86] showed that an IET is never mixing. The question of weak mixing is a lot more subtle. First, Katok and Stepin [88] proved weak mixing for almost every IET on three intervals. Veech [183] established almost sure weak mixing for infinitely many irreducible permutations. Then Nogueira and Rudolph [132] proved that a.e. irreducible IET is topological weak mixing. Finally, Avila and Forni [16] proved that for any irreducible permutation that is not a rotation, almost every IET is weakly mixing. It seems that a generic IET is ergodic, but we could not find an appropriate reference.

## 2. Linear prevalence

Here are several statements about Banach spaces that can be made precise using prevalence (see [74]):

- Almost every function  $f \in L^1[0, 1]$  satisfies  $\int_0^1 f(x)dx \neq 0$ .
- Almost every sequence  $\{a_j\}_{j=1}^\infty$  in  $\ell^2$  has the property that  $\sum_{j=1}^\infty a_j$  diverges.
- Almost every  $f \in C^1(\mathbb{R})$  has no point  $x$  where  $f(x) = f'(x) = 0$ .
- Almost every  $C^1$  map on  $\mathbb{R}^d$  has the property that all of its periodic points are hyperbolic.

These statements are true in terms of topological genericity; for example, the set of  $f \in L^1[0, 1]$  for which  $\int_0^1 f(x)dx \neq 0$  is easily seen to be open and dense. The generic version of the third statement is a consequence of the Kupka–Smale theorem (see Example 9 in Section 1.2). However, in light of the examples from the previous section, we would like a probabilistic framework in which to prove such statements, one that coincides with the notion of ‘Lebesgue almost every’ on  $\mathbb{R}^d$ .

One may wonder whether Lebesgue measure itself can be extended to some or all infinite-dimensional Banach spaces. However, as we now show, a translation-invariant measure  $\mu$  on an infinite-dimensional Banach space  $B$  must either have the property that some open sets have measure zero or that all open sets have infinite measure. This follows



from the fact that an open ball in  $B$  with radius  $R$  contains infinitely many disjoint balls of radius  $R/4$ . One proves this by constructing the centres of these balls as a sequence  $\{x_n\}$  in the open ball with radius  $3R/4$  such that  $\|x_m - x_n\| > R/2$  for  $m \neq n$ . Given  $x_1, x_2, \dots, x_n$ , one chooses  $x_{n+1}$  by observing that since the span  $V_n$  of  $x_1, x_2, \dots, x_n$  is a proper closed subspace of  $B$ , there must be a point  $y$  that has a positive distance from  $V_n$ , and by scaling  $y$  we can make that distance lie between  $R/2$  and  $3R/4$ . Then  $\|y - v\| < 3R/4$  for some  $v \in V_n$ , and  $y - v$  lies at least  $R/2$  from  $V_n$ , so let  $x_{n+1} = y - v$ .) Since  $\mu$  is translation invariant, each of the balls of radius  $R/4$  has the same measure, so either they have measure zero or the ball with radius  $R$  has infinite measure. Thus if all balls have positive measure, all balls have infinite measure, and hence so do all open sets.

In order to regard measure zero sets as ‘negligible’, we do not want open sets to have measure zero. On the other hand, if all open sets have infinite measure, then the class of sets with measure zero may be unreasonably small. In  $\mathbb{R}^d$ , Lebesgue measure is not the only translation-invariant measure, but it is the only one (up to a constant) that is also  $\sigma$ -finite. Hausdorff  $s$ -dimensional measure with  $s < d$  is an example of a translation-invariant measure that is not  $\sigma$ -finite; for  $s = 0$  this is just a counting measure, which assigns to each set its cardinality. Furthermore, the Hausdorff measures are well-defined on infinite-dimensional Banach spaces. One could consider a set to be negligible if it has measure zero with respect to such a measure, but then being negligible would be a much stronger condition than having Lebesgue measure zero (for counting measure, only the empty set has measure zero).

Not only is there no  $\sigma$ -finite translation-invariant measure on an infinite-dimensional Banach space, there is in fact no nonzero  $\sigma$ -finite measure for which the sets of measure zero are translation-invariant [57, 171, 172, 180]. In light of these considerations, we abandon the idea of defining measure zero on infinite-dimensional Banach spaces with respect to a single measure.

Before proceeding to the general definition of prevalence, we consider a preliminary version based on considering  $d$ -parameter families of functions for finite  $d$ . Many results in dynamical systems describe behaviour that holds for Lebesgue almost every parameter, or a positive Lebesgue measure set of parameters, in a specific family of systems or in all families of a given type. This is the nature of the measure-theoretic results in Examples 2, 3, 7, 9 and 11 of the previous section, for example. Discussing genericity in dynamical systems, Anosov wrote [9] (using ‘second category’ to mean topologically generic):

When properties of points in  $\mathbb{R}^n$  are being discussed, one can base the notion of genericity on Lebesgue measure instead of sets of second category. The metric and category points of view are similar in many respects, but do not coincide; in fact a first-category set may have total measure. In infinite-dimensional function space there is no particular natural measure, and so the metric variant drops out. But, the metric point of view can be adopted for a family of dynamical systems depending on a finite number of parameters  $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ . There are interesting problems in which the distinction between this point of view and the category point of view manifests itself (Anosov here refers to Arnol’d discussion that is reproduced in

Example 2 above). Moreover, smooth (of some class or other) families of smooth dynamical systems themselves form a function space, so in this space the category approach may be applied: then the metric approach can be applied to a generic (in the sense of Baire category) family.

The last sentence forms the basis for the definition of prevalence we give for nonlinear spaces in Section 3. However, in Banach spaces we do not require consideration of generic  $d$ -parameter families, but rather a single  $d$ -parameter family that, by translation, foliates the space into finite-dimensional leaves.

We say that a set  $S$  in a Banach space  $B$  is *finite-dimensionally prevalent* if we can define a continuous family of perturbations  $v_x \in B$  for  $x$  in some bounded neighbourhood of 0 in  $\mathbb{R}^d$  (with  $v_0 = 0$ ), such that for every fixed  $v \in B$ , we have that  $v + v_x \in S$  for Lebesgue almost every  $x$ . In other words, we require that for some finite-dimensional family of perturbations, if we start at any point in  $B$ , then by adding a randomly chosen perturbation, we end up in  $S$  with probability one (where by randomly we mean uniformly with respect to normalized Lebesgue measure).

REMARK 1. Finite-dimensional prevalence has been formulated by several authors with the restriction that  $v_x$  is required to be a linear function of  $x$ , so that the perturbations span a finite-dimensional subspace of  $B$ . Quinn and Sard [147] called the complement of such a finite-dimensionally prevalent set ‘0-preconull’. Vishik and Kuksin formulated essentially the same notion in [184]. Sauer, Yorke, and Casdagli [154] used ‘prevalence’ to mean this finite-dimensional notion.

A set  $S \subset \mathbb{R}^d$  whose complement has Lebesgue measure zero is finite-dimensionally prevalent; one need simply use the perturbations  $v_x = x$  for  $x$  in the unit ball in  $\mathbb{R}^d$ . As we will see below in Proposition 2.5, the converse of this statement is also true. This follows from a Fubini argument; if, for example,  $S$  has probability one with respect to perturbations in a proper subspace of  $\mathbb{R}^d$ , then adding perturbations in the remaining directions to span  $\mathbb{R}^d$  does not change this fact.

We call the complement of a prevalent set *shy*. Finite-dimensional shyness satisfies the following conditions that are true also of ‘Lebesgue measure zero’:

- (1) Every subset of a shy set is shy.
- (2) Every translation of a shy set is shy.
- (3) No nonempty open set is shy.

The first two conditions are easily verified, while the third follows from the fact that for a nonempty open set  $S$  in a Banach space  $B$  and a continuous family  $v_x$  of perturbations, if  $v \in S$  then the inverse image of  $x \mapsto v + v_x$  is open and nonempty, and hence has positive Lebesgue measure.

REMARK 2. For topological genericity, the negligible sets are called *meagre* or said to be of *first (Baire) category*: countable unions of nowhere dense sets. Conditions (1) and (2) above are easily seen to be true with ‘shy’ replaced by ‘meagre’. The fact that meagre sets cannot contain an open set in a complete metric space is the Baire Category Theorem (see e.g. [137]). More generally, any space in which a meagre set must have no interior is called

a *Baire space*; in addition to complete metric spaces, locally compact Hausdorff spaces are also known to be Baire spaces. The complement of a meagre set is also called *residual*.

As mentioned above, we will also show that (in contrast to meagreness):

(4) In  $\mathbb{R}^d$ , a set is shy if and only if it has Lebesgue measure zero.

However, we do not know whether finite-dimensional shyness satisfies the following condition that measure zero sets with respect to any given measure must satisfy:

(5) Every countable union of shy sets is shy.

For example, the proof that almost every map has all of its points of period  $p$  hyperbolic involves the use of a  $2p$ -parameter family of perturbations. The set  $S$  of maps for which all periodic points are hyperbolic is then the intersection over  $p$  of finite-dimensionally prevalent sets, but this method of proof does not establish that  $S$  itself is finite-dimensionally prevalent. We could, of course, say that a countable intersection of finite-dimensionally prevalent sets is prevalent (or equivalently, that a countable union of finite-dimensionally shy sets is shy), but it is not immediately clear (though it turns out to be true) that condition (3) is still satisfied with this definition.

In order to satisfy condition (5) in a more natural way, we make a modest generalization of finite-dimensional prevalence and shyness. One can think of the family  $v_x$  of perturbations above, with  $x$  distributed according to normalized Lebesgue measure, as defining a probability measure on the Banach space  $B$ . Finite-dimensional prevalence requires a property to have probability one with respect to all translations of such a measure. In the definitions of prevalence and shyness below, we allow more general probability measures. (For ease of notation, we translate sets rather than measures below.)

Though we have thus far discussed only Banach spaces, we do not require scalar multiplication for the definitions below. We define prevalence more generally for *complete metric Abelian groups*, by which we mean topological Abelian groups (with operation denoted  $+$ ) whose topology is generated by a complete metric.

**DEFINITION 1.** Let  $B$  be a complete metric Abelian group. A Borel measure  $\mu$  on  $B$  is called *transverse* to a Borel set  $S \subset B$  if for all  $x \in B$ , we have  $\mu(S+x) = 0$ . A Borel set  $S \subset B$  is called *shy* if there exists a compactly supported Borel probability measure that is transverse to  $S$ . More generally, a subset of  $B$  is called shy if it is contained in a shy Borel set. A subset of  $B$  is called *prevalent* if its complement is shy.

Of course, the normalization of the measure is not important, and one can replace ‘probability measure’ with ‘nonzero finite measure’ in the definition above. The reason that we place such a restriction on the measure (and that we place restrictions on the topological group  $B$ ) is to allow us to prove that every countable union of shy sets is shy (see [Proposition 2.7](#)). However, in many cases it is convenient to use transverse measures that are not finite, such as Lebesgue measure supported on a finite-dimensional subspace of a Banach space  $B$ . In order to allow the use of such measures, we say that a Borel measure  $\mu$  on a topological space  $B$  is *admissible* if there is a compact set  $K \subset B$  such that  $0 < \mu(K) < \infty$ . The following lemma shows that ‘compactly supported Borel probability measure’ can be replaced by ‘admissible measure’ in the definition above.

LEMMA 2.1. *If a Borel set  $S \subset B$  has a transverse admissible measure, then it is shy.*

PROOF. Let  $\mu$  be an admissible measure on  $B$  that is transverse to  $S$ , and let  $K \subset B$  be a compact set for which  $0 < \mu(K) < \infty$ . For every Borel set  $T \subset B$ , let  $\tilde{\mu}(T) = \mu(T \cap K)/\mu(K)$ . Then  $\tilde{\mu}$  is a Borel probability measure that is transverse to  $S$ , and whose support is contained in  $K$ .  $\square$

REMARK 3. In the case where  $B$  is separable, Christensen [32–34] defined the notion of a ‘Haar zero set’, which is equivalent to the definition of ‘shy’ above. (Christensen requires only a transverse Borel probability measure, but as we will see below, for separable spaces all Borel probability measures are admissible.) We use the terminology of [74], which gave essentially the definition above, but built admissibility into the definition of ‘transverse’. Borwein and Moors [27] gave an equivalent definition of shyness; they and other authors call a shy set a ‘Haar null set’.

## 2.1. Properties of prevalence

In this section, we verify conditions (1)–(5) from the previous section and discuss other properties of shyness and prevalence. Throughout this section we assume that  $B$  is a complete metric Abelian group, and we largely follow the arguments in [74]. Conditions (1) and (2) are easy to check:

PROPOSITION 2.2. *If  $S$  is shy, then so is every subset of  $S$  and every translation of  $S$ .*

Next we verify condition (3).

PROPOSITION 2.3. *Every shy set has no interior. Equivalently, every prevalent set is dense.*

PROOF. Let  $S$  be a set containing an open set  $U$ , and let  $\mu$  be a Borel probability measure with compact support  $K \subset B$ . The collection of open sets  $U + x$  for  $x \in B$  cover  $K$ , so by compactness there is a finite subcover. Since  $\mu(K) = 1$ , we must have  $\mu(U + x) > 0$  for some  $x$ . Then  $\mu(S + x) > 0$  as well, so  $\mu$  is not transverse to  $S$ . Since  $\mu$  was arbitrary,  $S$  cannot be shy.  $\square$

Conditions (4) and (5) rely on the fact that the convolution of a  $\sigma$ -finite transverse measure with an arbitrary  $\sigma$ -finite Borel measure is also transverse. This, and several of the arguments below, rely on an earlier version of the Fubini theorem due to Tonelli (see, for example, [48]):

THEOREM (Tonelli). *Let  $\mu$  and  $\nu$  be  $\sigma$ -finite measures on spaces  $X$  and  $Y$ , respectively. If  $f : X \times Y \rightarrow \mathbb{R}$  is nonnegative and measurable with respect to the product measure  $\mu \times \nu$ , then*

$$\int_{X \times Y} f(x, y) d\mu \times \nu(x, y) = \int_Y \int_X f(x, y) d\mu(x) d\nu(y).$$

(In particular,  $f(x, y)$  is a measurable function of  $x$  for almost every  $y$ , and  $\int_X f(x, y) d\mu(x)$  is a measurable function of  $y$ , so that the right side makes sense.)

The advantage of this theorem over the Fubini theorem is that one need not verify *a priori* that the integral on the left side above is finite. (The Fubini theorem, on the other hand, does not require that  $f$  be nonnegative.) We will apply the Tonelli theorem below only to characteristic functions of Borel sets, so that measurability is not an issue.

Given Borel measures  $\mu$  and  $\nu$  on a group  $B$ , we define their convolution  $\mu * \nu$  as follows. Let  $S \subset B$  be a Borel set, and let  $\mu \times \nu$  be the product measure on  $B \times B$ . Let  $\mu * \nu(S) = \mu \times \nu(\{(x, y) : x + y \in S\})$ . The lemma below immediately implies that the union of two shy sets is shy.

**LEMMA 2.4.** *Let  $\mu$  be a  $\sigma$ -finite measure on a group  $B$  that is transverse to a Borel set  $S \subset B$ . For every  $\sigma$ -finite Borel measure  $\nu$  on  $B$ , the convolution  $\mu * \nu$  is transverse to  $S$ .*

**PROOF.** By the Tonelli theorem, for all  $z \in B$ ,

$$\begin{aligned} \mu * \nu(S + z) &= \int_{B \times B} \chi_{\{(x, y) : x + y \in S + z\}}(x, y) d\mu \times \nu(x, y) \\ &= \int_B \int_B \chi_{S + z - y}(x) d\mu(x) d\nu(y) \\ &= \int_B \mu(S + z - y) d\nu(y) = 0. \end{aligned} \quad \square$$

Now condition (4) follows by convolution with Lebesgue measure.

**PROPOSITION 2.5.** *A set  $S \subset \mathbb{R}^d$  is shy if and only if it has Lebesgue measure zero.*

**PROOF.** Let  $\nu$  be Lebesgue measure on  $\mathbb{R}^d$ . If  $\nu(S) = 0$ , then  $S$  is contained in a Borel set  $\tilde{S}$  with  $\nu(\tilde{S}) = 0$ . Since  $\nu$  is translation-invariant, this implies that  $\nu$  is transverse to  $\tilde{S}$ , and hence  $\tilde{S}$  and  $S$  are shy. If, on the other hand,  $S$  is shy, then it is contained in a shy Borel set  $\tilde{S}$ . Let  $\mu$  be a finite (hence  $\sigma$ -finite) Borel measure that is transverse to  $\tilde{S}$ . Then  $\mu * \nu$  is also transverse to  $\tilde{S}$ . But reversing the roles of  $\mu$  and  $\nu$  in the proof of [Lemma 2.4](#), we have that for every Borel set  $T \subset \mathbb{R}^d$ ,

$$\begin{aligned} \mu * \nu(T) &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \chi_{\{(y, x) : y + x \in T\}}(y, x) d\nu \times \mu(y, x) \\ &= \int_{\mathbb{R}^d} \nu(T - x) d\mu(x) = \nu(T) \mu(\mathbb{R}^d), \end{aligned}$$

again using translation-invariance of  $\nu$ . In other words,  $\mu * \nu$  is a multiple of  $\nu$ , and thus  $\nu$  is transverse to  $\tilde{S}$ . In particular,  $\nu(\tilde{S}) = 0$  and hence  $\nu(S) = 0$ .  $\square$

**REMARK 4.** More generally, the same argument applies to other topological Abelian groups with Lebesgue measure replaced by any (nonzero)  $\sigma$ -finite translation-invariant Borel measure (completed so that all subsets of Borel sets with measure zero also are measurable with measure zero). When  $B$  is a locally compact Hausdorff space, such a measure exists and is (up to a constant multiple) unique, and is called *Haar measure*.

Finally, we verify condition (5). This is the first result in which we use that the topology of  $B$  is generated by a complete metric.

REMARK 5. For simplicity, we assume below in several places that the metric on  $B$  is translation-invariant. If it were not, we could replace it with a translation-invariant metric that generates the same topology [94].

We start with a lemma that says that a shy set has a transverse probability measure with arbitrarily small support.

LEMMA 2.6. *If  $S \subset B$  is shy, then for every  $\varepsilon > 0$ , there is a compactly supported Borel probability measure that is transverse to  $S$  with its support contained in the closed ball with radius  $\varepsilon$  centred at 0.*

PROOF. Let  $\mu$  be a Borel probability measure with compact support  $K \subset B$  that is transverse to  $S$ . By compactness,  $K$  can be covered with finitely many closed balls of radius  $\varepsilon$ . At least one of these balls, call it  $C$ , has positive measure. Let  $x$  be the centre of  $C$ . For every Borel set  $T \subset B$ , let  $\tilde{\mu}(T) = \mu((T + x) \cap C)/\mu(C)$ . Then  $\tilde{\mu}$  is a Borel probability measure that is transverse to  $S$ , and its support is the compact set  $(K \cap C) - x$ . The support is contained in  $C - x$ , which is the closed ball with radius  $\varepsilon$  centred at 0.  $\square$

PROPOSITION 2.7. *Every countable union of shy sets is shy.*

PROOF. Every countable collection of shy subsets of  $B$  is contained in a collection  $S_1, S_2, \dots \subset B$  of Borel sets with transverse probability measures  $\mu_1, \mu_2, \dots$  that have compact supports  $K_1, K_2, \dots \subset B$ , respectively. By Lemma 2.6, we can assume, without loss of generality, that  $K_n$  is contained in the closed ball with radius  $2^{-n}$  centred at 0. This will allow us to construct a compactly supported Borel probability measure that is the infinite convolution of all  $\mu_n$ , and is thus transverse to  $S$ .

Let  $K = K_1 \times K_2 \times \dots$  be the Cartesian product of all  $K_n$ . Then in the product topology,  $K$  is compact by the Tychonoff theorem, and supports a product probability measure  $\tilde{\mu} = \mu_1 \times \mu_2 \times \dots$  defined on the Borel subsets of  $K$  [48]. For  $(x_1, x_2, \dots) \in K$ , let

$$\pi(x_1, x_2, \dots) = x_1 + x_2 + \dots$$

Since  $x_n$  lies at most  $2^{-n}$  from 0, the infinite sum above is a Cauchy sequence, and thus converges since  $B$  is complete. Furthermore, as we now show,  $\pi$  is continuous.

For every  $\varepsilon > 0$  and positive integer  $N$ ,  $(x_1, x_2, \dots)$  has a neighbourhood consisting of sequences  $(y_1, y_2, \dots)$  such that  $y_n$  is within  $\varepsilon$  of  $x_n$  for  $1 \leq n \leq N$ . Then  $\pi(y_1, y_2, \dots)$  is within  $N\varepsilon + 2^{1-N}$  of  $\pi(x_1, x_2, \dots)$ . By choosing  $N$  and  $\varepsilon$  appropriately, this distance can be made arbitrarily small. Thus  $\pi$  is continuous as claimed.

It follows that the image  $\pi(K)$  of  $K$  under  $\pi$  is compact. Define the Borel probability measure  $\mu = \pi(\tilde{\mu})$  supported on  $\pi(K)$  by  $\mu(T) = \tilde{\mu}(\pi^{-1}(T))$  for Borel sets  $T \in B$ . We think of  $\mu$  as the infinite convolution of all  $\mu_n$ . Next, for each  $n \geq 1$  we define the compactly supported Borel probability measure  $\tilde{\nu}_n$  to be the product of all  $\mu_m$  except for  $\mu_n$ , and let  $\nu_n = \pi(\tilde{\nu}_n)$ . Since (with appropriate rearrangement of indices)  $\tilde{\mu} = \mu_n \times \tilde{\nu}_n$ , and summation is commutative, it follows that  $\mu = \mu_n * \nu_n$ . Since  $\mu_n$  is transverse to  $S_n$ , by Lemma 2.4 so is  $\mu$ . Finally, since  $\mu$  is transverse to all  $S_n$ , it is transverse to their union.  $\square$

The fact that a transverse measure can be taken to have arbitrarily small support suggests that prevalence and shyness are essentially local properties. At least in the case when  $B$  is separable, this idea can be made precise.

**DEFINITION 2.** We say that a set  $S \subset B$  is *locally shy* if every point in  $B$  has a neighbourhood whose intersection with  $S$  is shy, and that a set is *locally prevalent* if its complement is locally shy.

Of course, all shy sets are locally shy, and likewise for prevalence. The converse is true in separable spaces.

**PROPOSITION 2.8.** *If  $B$  is separable, then all locally shy sets are shy.*

**PROOF.** If  $S \subset B$  is locally shy, then  $B$  can be covered by open sets each of whose intersection with  $S$  is shy. By the Lindelöf theorem [48], since  $B$  is separable, this cover has a countable subcover. Intersecting each set in this subcover with  $S$ , we express  $S$  as a countable union of shy sets, whence  $S$  is shy.  $\square$

If  $B$  is not separable, then local prevalence and local shyness may be weaker conditions than their global counterparts, but we think that it is still reasonable to call a set measure-theoretically generic if it is locally prevalent.

When  $B$  is separable, the condition of compact support can be removed from the definition of prevalence, because all probability measures are admissible in the sense of Lemma 2.1. The following result and proof are due to Ulam; see footnote 3 in [138].

**PROPOSITION 2.9.** *If  $B$  is separable, then all Borel probability measures are admissible.*

**PROOF.** Let  $\mu$  be a Borel probability measure on  $B$ , let  $\{x_n\}$  be a dense sequence of points in  $B$ , and let  $S_{k,N}$  be the union of the closed balls of radius  $2^{-k}$  centred at  $x_1, x_2, \dots, x_N$ . For each  $k \geq 1$ , define  $N_k$  inductively to be the smallest integer such that

$$\mu \left( \bigcap_{j=1}^k S_{j,N_j} \right) > 1/2.$$

Such  $N_k$  always exist because the (nested) union over  $N \geq 1$  of  $S_{k,N}$  is all of  $B$ , and in particular covers  $\bigcap_{j=1}^{k-1} S_{j,N_j}$ . Then  $\bigcap_{j=1}^{\infty} S_{j,N_j}$  then has measure at least  $1/2$ , and this intersection is compact because it is closed and totally bounded.  $\square$

Non-shy Borel sets in Euclidean spaces have positive Lebesgue measure; in infinite-dimensional spaces, they share some properties of positive Lebesgue measure sets but have some important differences. In the case that  $B$  is a separable Banach space, Matoušková [113] proved that non-shy Borel sets are uniformly non-shy in the following sense. If  $S \subset B$  is not shy, then there exist  $r > 0$  and  $\delta > 0$  such that for every compactly probability measure  $\mu$  whose support is contained in a ball of radius  $r$ , there exists  $x \in B$  such that  $\mu(S+x) \geq \delta$ . On the other hand, Dougherty [45] showed that if  $B$  is a separable infinite-dimensional Banach space, it contains uncountably many disjoint non-shy Borel sets (see also [169]).

In Section 2, we argued that in an infinite-dimensional Banach space, a ball with radius  $R$  contains infinitely many disjoint balls with radius  $R/4$ . This is a standard argument that also shows that a closed ball with positive radius is not compact. In particular, a compact set has no interior. The same is true for an infinite-dimensional vector space with a complete metric [155, p. 23]. As we will show in the next section (see Proposition 2.11), this implies that all compact sets are shy.

In Euclidean space, sets with Lebesgue measure zero are invariant under diffeomorphisms, but we do not know if shyness is invariant under sufficiently smooth homeomorphisms of Banach spaces. However, Matoušková [114] showed that in any separable infinite-dimensional Banach space, there is a bi-Lipschitz homeomorphism that maps a shy set to a non-shy set.

## 2.2. Proving prevalent results

Given a property that one thinks is prevalent or shy, in order to construct a proof one must find a transverse measure. In vector spaces, a convenient choice is often Lebesgue measure supported on a finite-dimensional subspace. (Though not compactly supported, this measure is admissible in the sense above that it can be restricted to a compactly supported probability measure.) Here we give some basic examples of this technique. First we consider statements 1 from the beginning of Section 2.

**PROPOSITION 2.10.** *The set of functions  $f \in L^1[0, 1]$  for which  $\int_0^1 f(x)dx \neq 0$  is prevalent.*

**PROOF.** Let  $S$  be the set of functions in  $L^1[0, 1]$  whose integral is nonzero, and let  $\mu$  be Lebesgue measure on the subspace of constant functions in  $L^1[0, 1]$ . For each  $g \in L^1[0, 1]$ , the set  $S + g$  is the set of functions whose integral is not equal to  $\int_0^1 g(x)dx$ . There is only one constant function in  $S + g$ , and hence  $\mu(S + g) = 0$ . This shows that  $\mu$  is transverse to  $S$ .  $\square$

**REMARK 6.** A more colloquial, but perhaps more intuitive, way to write the proof above is as follows. ‘For every  $f \in L^1[0, 1]$ , we have  $\int_0^1 (f(x) + c)dx = \int_0^1 f(x)dx + c$ , which is nonzero for Lebesgue almost every  $c$ . It follows that Lebesgue measure on the space of constant functions in  $L^1[0, 1]$  is transverse to the set of functions with integral zero.’ Here we have argued that the set  $S$  has measure zero with respect to the translation of the measure  $\mu$  by an arbitrary function  $f$ , which is equivalent to saying that  $\mu(S - f) = 0$ . While we used the latter notion of translating the set by an arbitrary function in our formal definition of prevalence, and in the proof above, the equivalent notion of translating the measure by an arbitrary function allows one to think of the measure as defining a class of perturbations of that function (as in our definition of finite-dimensional prevalence).

Statement 2 of Section 2, that a prevalent element of  $\ell^2$  is not summable, can also be proved with a one-dimensional space of perturbations. For example, consider the subspace spanned by  $(1, 1/2, 1/3, \dots) \in \ell^2$ . Every translation of this subspace intersects the set of summable elements of  $\ell^2$  in at most one point, and hence Lebesgue measure on this subspace is transverse to the set of summable elements.



Next we show that compact subsets of infinite-dimensional spaces are shy using a one-dimensional space of perturbations.

**PROPOSITION 2.11.** *If  $B$  is an infinite-dimensional vector space, then every compact set is shy.*

**PROOF.** Let  $S$  be a compact subset of an infinite-dimensional Banach space  $B$ . As above, it suffices to show that there is a one-dimensional subspace  $V$  of  $B$  such that every translation of  $V$  intersects  $S$  (or equivalently, every translation of  $S$  intersects  $V$ ) in at most one point. Consider the function  $f : \mathbb{R} \times S \times S \rightarrow B$  given by

$$f(c, x, y) = c(x - y).$$

A vector  $v$  that is not in the range of  $f$  has the property that no nonzero multiple of  $v$  is the difference between two elements of  $S$ ; hence,  $v$  spans a subspace  $V$  with the property above. Thus we need only show that the range of  $f$  is not all of  $B$ .

Since  $S$  is compact, the image  $I_n$  of  $[-n, n] \times S \times S$  under  $f$  is compact for each positive integer  $n$ . Since  $B$  is infinite-dimensional, a compact subset of  $B$  has no interior (a ball of radius  $R$  cannot be covered by finitely many balls of radius  $R/4$  by an argument similar to before), and hence is nowhere dense. Therefore the range of  $f$  is the union of countably many nowhere dense sets  $I_n$ ; that is, it is meagre. In particular, the range is not all of  $B$  by the Baire category theorem.  $\square$

### 2.3. Examples of linear prevalence

In this section we present some results that have been proved in terms of linear prevalence, and what analogous results are known for topological genericity.

**2.3.1. Regularity of functions.** A common type of generic or prevalent result on a function space says that a function is typically as ‘irregular’ as the function space allows. For example, a classical result is that a generic continuous function from  $\mathbb{R}$  (or an interval therein) to  $\mathbb{R}$  is nowhere differentiable [20,117] (see also [137]). Furthermore, continuous functions are generically nowhere Hölder continuous, and Hölder continuous functions with a given exponent are generically nowhere Hölder continuous with a higher exponent [15]. These results are also true for prevalent continuous functions in these spaces [72], and more strongly for HP-typical (see Section 4.3) continuous functions [96].

On the other hand, if we extend the usual definition of differentiability to include cases where the limit defining the derivative is  $\pm\infty$  (as opposed to cases where the  $\liminf$  and  $\limsup$  differ), then a generic continuous function is still nowhere differentiable, but Zajíček [188] proved that neither the nowhere differentiable functions nor their complements are prevalent in the continuous functions.

Shi [158] proved several prevalent results of this type previously known to be true generically – for example, almost every function in  $C^\infty[0, 1]$  is nowhere real analytic, and almost every function that is Riemann integrable on a bounded interval is discontinuous on a set of points with cardinality of the continuum within any subinterval.

Frayse and Jaffard [55] considered the Hölder regularity of functions in various Sobolev and Besov spaces. They showed that a prevalent function in one of these spaces is as irregular as possible in terms of the multifractal spectrum of Hölder exponents. Again this coincides with an earlier generic result [76].

Frayse [54] (see also [56]) also proved that in the space of complex analytic functions on the open unit disk (with the topology of uniform convergence on compact subsets), a prevalent function cannot be extended analytically through an arc of the unit circle; this is also true for a generic function from this space [93].

The graph of a generic [71] or prevalent [118] function in  $C[0, 1]$  has dimension 2 for various notions of dimension, including the upper box-counting dimension.

**2.3.2. Projections and embeddings.** In this section we discuss results of the following type. Given a space of functions from a set  $S$  of a Euclidean space, what properties of  $S$  or measure  $\mu$  on  $S$  are preserved in its image under a generic or prevalent function?

If  $S$  is a  $C^r$  manifold of dimension  $d$  for  $r \geq 1$ , the Whitney Embedding Theorem [185] says that  $S$  can be embedded in  $\mathbb{R}^{2d+1}$ . Whitney's proof is perturbative and implies that a generic  $C^r$  function from  $S$  to  $\mathbb{R}^n$  is an embedding (that is, a diffeomorphism between  $S$  and its image) for  $n > 2d$ . In this case all diffeomorphism-invariant properties of  $S$  are preserved in its image. The proof of Mañé's result [106] for linear projections (see Example 8 in Section 1.2) can be modified to show that if  $S$  is a compact subset of a Banach space with box-counting dimension  $d$ , then if  $n > 2d + 1$ , a generic  $C^r$  function in  $\mathbb{R}^n$  is one-to-one. Sauer, Yorke, and Casdagli [154] proved Whitney's result for a prevalent  $C^r$  function, and showed that if  $S$  is a compact subset of a Euclidean space with box-counting dimension  $d$ , a prevalent  $C^r$  function (defined at least on a neighbourhood of  $S$ ) in  $\mathbb{R}^n$  is an embedding on  $S$ , again provided that  $n > 2d$ . (Here and below, by 'embedding' on a set  $S$  that is not necessarily a manifold, we mean that the function is one-to-one and is an embedding on every compact manifold-with-boundary in  $S$ .)

Takens [174] proved a dynamical Embedding Theorem, which states that for a  $C^2$  flow  $\{\Phi_t\}_{t \in \mathbb{R}}$  on a  $d$ -dimensional  $C^2$  manifold  $S$  subject to certain nondegeneracy conditions, a generic  $C^2$  function  $h$  from  $S$  to  $\mathbb{R}$  generates an embedding into  $\mathbb{R}^{2d+1}$  when composed with the flow as follows:

$$x \mapsto (h(x), h(\Phi_{-1}(x)), h(\Phi_{-2}(x)), \dots, h(\Phi_{-2d}(x))).$$

Such a map is called a 'delay-coordinate embedding'. Sauer, Yorke, and Casdagli [154] proved a version of this result for prevalent  $h$ , and showed that if  $S$  is not a manifold but instead a compact invariant subset of  $\mathbb{R}^m$  with box-counting dimension  $d$ , then the delay-coordinate map into  $\mathbb{R}^n$  defined by a  $C^2$  function  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  is prevalently an embedding if  $n > 2d$ . Their non-degeneracy condition, somewhat weaker than Takens', is that the flow has finitely many fixed points, finitely many periodic orbits with integer period at most  $n$ , and no periodic orbits with period 1 or 2.

In the analysis of experimental data believed to arise from a chaotic attractor, a common technique is to try to embed the attractor into  $\mathbb{R}^n$ , either by making  $n$  simultaneous measurements of the state of the system at a series of times, or by using the delay-coordinate technique with a single measurement made at each time. In such cases the attractor

dimension is not known *a priori*, so one cannot know whether the embedding results above actually apply. Ott and Yorke [136] formulated versions of the prevalent results above in which the hypotheses can be checked on the image of  $S$  rather than on  $S$ . For example, with an appropriate definition of ‘tangent dimension’, they showed that a prevalent  $C^1$  function from a compact subset  $S$  of a Euclidean space into  $\mathbb{R}^n$  has the property that if the tangent dimension of the image of  $S$  is less than  $n/2$ , then the function is an embedding on  $S$ .

In cases where  $S$  is mapped to  $\mathbb{R}^n$  where  $n$  is not sufficiently large to guarantee that the mapping is generically or prevalently one-to-one, one may still expect certain features of  $S$  to be preserved in its image, for example its dimension. A classical result [110,91,115] is that if  $S$  is a compact subset of  $\mathbb{R}^m$  with Hausdorff dimension at most  $n$ , then almost every orthogonal projection onto an  $n$ -dimensional subspace preserves the Hausdorff dimension of  $S$ . Here ‘almost every’ is with respect to the measure on the Grassmanian manifold of  $n$ -dimensional subspaces of  $\mathbb{R}^m$  induced by Haar measure on the orthogonal group  $O(m)$ . We can restate this result in terms of Lebesgue measure by considering linear transformations from  $\mathbb{R}^m$  to  $\mathbb{R}^n$ . Since full-rank matrices have full Lebesgue measure in the space of  $n \times m$  matrices, and every full-rank linear transformation from  $\mathbb{R}^m$  to  $\mathbb{R}^n$  can be expressed in a unique way as a composition of an orthogonal projection onto an  $n$ -dimensional subspace and a linear isomorphism from that subspace to  $\mathbb{R}^n$ , it follows that Lebesgue almost every linear transformation from  $\mathbb{R}^m$  to  $\mathbb{R}^n$  preserves the Hausdorff dimension of  $S$  (again provided that this dimension is at most  $n$ ).

Using prevalence, one can extend these results to spaces of nonlinear functions and/or to the case that  $S$  lies in an infinite-dimensional space. Sauer and Yorke [153] proved that for compact  $S \subset \mathbb{R}^m$  with Hausdorff dimension at most  $n$ , a prevalent  $C^1$  function from  $\mathbb{R}^m$  to  $\mathbb{R}^n$  preserves the Hausdorff dimension of  $S$ , and that the same is true for the correlation dimension of a compactly supported measure on  $\mathbb{R}^m$ . (They also gave an example showing that no such result is possible for a box-counting dimension.) They also showed that for a compactly supported invariant measure  $\mu$  of a diffeomorphism  $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$ , the correlation dimension of  $\mu$  is preserved by the delay-coordinate map

$$x \mapsto (h(x), h(F^{-1}(x)), h(F^{-2}(x)), \dots, h(F^{-(n-1)}(x))).$$

for a prevalent  $C^1$  function  $h : \mathbb{R}^m \rightarrow \mathbb{R}$ , provided that  $n$  is at least the correlation dimension of  $\mu$  and that the set of periodic points of  $F$  with period at most  $n$  has Lebesgue measure zero.

In the case that  $S$  is a compact subset of a Banach space  $B$ , Hunt and Kaloshin [73] proved that if  $n$  is greater than twice the box-counting dimension of  $S$  and  $V$  is a space of functions from  $B$  to  $\mathbb{R}^n$  that contains the bounded linear functions and is contained in the locally Lipschitz functions, then a prevalent function in  $V$  is one-to-one on  $S$ . Thus, the conditions under which a prevalent function is one-to-one on  $S$  are the same as in the case where  $S$  lies in a Euclidean space. On the other hand, for each real  $d$  they gave, in the same paper, an example of a compact set  $S_0$  with Hausdorff dimension  $d$  in a Hilbert space, such that the image of  $S_0$  under every bounded linear function into  $\mathbb{R}^n$  has Hausdorff dimension less than  $d$ . With an additional assumption on the compact subset  $S$  of Banach space  $B$ , Ott, Hunt, and Kaloshin [134] proved that the Hausdorff dimension of  $S$  is preserved by a prevalent function from any of the function spaces  $V$  described above.

Kahane [79] (see also [56]) proved some results about images of sets and measures under prevalent continuous (rather than smooth) functions that differ strikingly from their generic counterparts. Specifically, he showed that given a Cantor set (a homeomorphic image of the canonical middle-third Cantor set)  $S$  in  $\mathbb{R}$ , the image of  $S$  under a prevalent continuous function  $f$  from  $\mathbb{R}$  to itself is the closure of its interior, i.e.  $f(S) = \overline{\text{Int } f(S)}$ , despite the fact that  $S$  itself has no interior. By contrast, Kaufmann [90] showed that the image of  $S$  under a generic continuous function is a Cantor set with Lebesgue measure zero. Kahane also proved that given a compactly supported non-atomic measure  $\mu$ , the image measure  $\mu \circ f^{-1}$  under a prevalent continuous function  $f$  is absolutely continuous (with respect to Lebesgue measure) with  $C^\infty$  density. By contrast, the image of  $\mu$  under a generic continuous function is singular with respect to Lebesgue measure.

**2.3.3. Symmetric attractors for generic or prevalent smooth  $\Gamma$ -invariant dynamical systems.** Let  $\Gamma$  be a compact Lie group acting on  $\mathbb{R}^n$ . Identify  $\Gamma$  with a closed subgroup of the group  $O(n)$  of the  $n \times n$  orthogonal matrices acting by matrix multiplication of vectors in  $\mathbb{R}^n$ . A mapping  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is called  $\Gamma$ -equivariant if  $f(\gamma x) = \gamma f(x)$  for every  $\gamma \in \Gamma$  and  $x \in \mathbb{R}^n$ .

Suppose that  $A$  is an  $\omega$ -limit set for the  $\Gamma$ -equivariant map  $f$ . We define the *symmetry group* of  $A$  to be the subgroup

$$\Sigma_A = \{\gamma \in \Gamma : \gamma A = A\}.$$

Since  $A$  is a closed subset,  $\Sigma_A$  is a closed subgroup of  $\Gamma$ . This subgroup has a physical interpretation as *symmetry on average* (see [40,120]).

A natural question is which closed subgroups of  $\Gamma$  can be realized as the symmetry group of such an  $\omega$ -limit set for a typical/prevalent map  $f$ .

The first elementary example is when  $A$  is a fixed point  $x$ , then  $\Sigma_A$  is the isotropy subgroup  $\Sigma_x = \{\gamma \in \Gamma : \gamma x = x\}$  of the point  $x$ . If  $A$  is a periodic point, then  $\Sigma_A$  contains the isotropy group of the points in  $A$  and is a cyclic extension of this isotropy group. Results from [14,51,120] give a good understanding of the case when  $\Gamma$  is finite. In particular, for ‘most’ actions of  $\Gamma$  the symmetry  $\Sigma_A$  can be any subgroup of  $\Gamma$ . Furthermore, each subgroup can be realized as an Axiom A attractor [51] and, therefore, such a symmetry is structurally stable.

An important remark is that for a more complicated  $\omega$ -limit set  $A$ , the subgroup  $\Sigma_A$  may be much larger.

In the case when  $\Gamma$  is not finite the situation is quite different. Here are a few examples from [121].

- Let  $\Gamma = SO(2)$ , then generically  $\Sigma_A = SO(2)$ ;
- Let  $\Gamma = O(2)$  and the dynamics in  $A$  is ‘mildly irregular’, then generically  $\Sigma_A$  is either  $SO(2)$  or  $O(2)$ ;
- Let  $\Gamma = O(3)$  and the dynamics in  $A$  is ‘very irregular’, then generically  $\Sigma_A$  is either  $SO(3)$  or  $O(3)$ ;

Define  $\Gamma^0$  to be the connected component of the identity in  $\Gamma$ . The following result of Melbourne and Stewart [121] generalizes these examples.

Suppose that  $\Gamma \subset O(n)$  in an Abelian compact Lie group and that  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a  $\Gamma$ -equivariant mapping with  $\omega$ -limit set  $A = \omega(x_0)$ . A  $\Gamma$ -cocycle is a map  $\phi : \mathbb{R}^n \rightarrow \Gamma^0$ . We form a perturbation of  $f$  by defining  $f_\phi(x) = \phi(x)f(x)$ . Note that  $f_\phi(x)$  is automatically  $\Gamma$ -equivariant (since  $\Gamma$  is Abelian). Let  $A_\phi$  be the  $\omega$ -limit set of  $x_0$  under  $f_\phi$ . Denote by  $\mathcal{Z}_k$  the space of compactly supported  $C^k$   $\Gamma$ -cocycles. Melbourne–Stewart [121] showed that for each nonnegative  $k$ , there is a generic and prevalent subset  $Z \subset \mathcal{Z}_k$  such that  $\Sigma_{A_\phi}$  contains  $\Gamma^0$  for each  $\phi \in Z$ .

**2.3.4. Prevalence of Nekhoroshev stability.** A well known result by Nekhoroshev [128] states that for an analytic nearly integrable Hamiltonian system, the action variables of the unperturbed Hamiltonian remain nearly constant over an exponentially long time with respect to the size of the perturbation, provided that the unperturbed Hamiltonian satisfies some generic transversality condition known as *steepness*. Niederman [131] proved that the Nekhoroshev estimates still hold for a prevalent unperturbed Hamiltonian.

**2.3.5. Convergence to equilibria in strongly order-preserving semiflows.** Enciso, Hirsch, and Smith [50] considered the dynamics of a prevalent initial condition for a strongly order-preserving semiflow on a separable Banach space. The space is assumed to have a nontrivial cone  $K$  and a partial order such that  $x \leq y$  if  $y - x \in K$ , and a continuous semiflow  $\Phi$  is ‘strongly order-preserving’ if for all  $x < y$ , there are neighbourhoods  $U \ni x$  and  $V \ni y$  such that  $\Phi_t(U) \leq \Phi_t(V)$  for  $t$  sufficiently large. For such semiflows, Hirsch [69] proved that of the initial conditions with compact orbit closure, a residual subset is quasiconvergent, meaning that their  $\omega$ -limit sets are contained in the set of equilibria. Under additional hypotheses, Smith and Thieme [168] proved convergence to a single equilibrium for a residual subset. Enciso, Hirsch, and Smith [50] proved the analogous results for prevalence, namely that the sets of exceptional initial conditions, proved meagre by the results above, are also shy. Many PDEs are known to be strongly order preserving, including a class of reaction-diffusion systems discussed in [50].

**2.3.6. Non-uniform hyperbolicity and reducibility of one-dimensional quasiperiodic cocycles in  $SL(2, \mathbb{R})$ .** A one-dimensional quasiperiodic  $C^r$ -cocycle in  $SL(2, \mathbb{R})$  is a pair  $(\alpha, A)$  where  $\alpha \in \mathbb{R}$  and  $A \in C^r(\mathbb{R}/\mathbb{Z}, SL(2, \mathbb{R}))$ . A cocycle could be viewed as a skew-product:

$$(\alpha, A) : \mathbb{R}/2\mathbb{Z} \times \mathbb{R}^2 \rightarrow \mathbb{R}/2\mathbb{Z} \times \mathbb{R}^2 \quad (x, w) \rightarrow (x + \alpha, A(x) \cdot w).$$

The Lyapunov exponent of  $(\alpha, A)$  is defined as

$$L(\alpha, A) = \lim_{n \rightarrow \infty} \frac{1}{n} \int_{\mathbb{R}/\mathbb{Z}} \ln \|A_n(x)\| dx \geq 0,$$

where  $A_n(x) = \prod_{j=n-1}^0 A(x + j\alpha) = A(x + (n-1)\alpha) \cdot \dots \cdot A(x)$ . A cocycle  $(\alpha, A)$  is called *uniformly hyperbolic* if there is a continuous splitting  $E^s(x) \oplus E^u(x) = \mathbb{R}^2$  and constants  $C > 0, \lambda < 1$  such that

$$\|A_n(x)w\| \leq C\lambda^n \|w\| \quad w \in E^s(x)$$

$$\|A_n(x - n\alpha)^{-1}w\| \leq C\lambda^n \|w\| \quad w \in E^u(x).$$

The set of uniformly hyperbolic cocycles is open in the  $C^0$ -topology allowing perturbation of both  $\alpha$  and  $A$ . If  $(\alpha, A)$  has positive Lyapunov exponents but is not uniformly hyperbolic we call it *nonuniformly hyperbolic*.

We say that a  $C^r$ -cocycle  $(\alpha, A)$  is  $C^r$ -reducible if there exists  $B \in C^r(\mathbb{R}/2\mathbb{Z}, SL(2, \mathbb{R}))$  and  $A_0 \in SL(2, \mathbb{R})$  such that

$$B(x + \alpha)A(x)B^{-1}(x) = A_0, \quad x \in \mathbb{R}$$

and it is  $C^r$ -reducible modulo  $\mathbb{Z}$  if one can take  $B \in C^r(\mathbb{R}/\mathbb{Z}, SL(2, \mathbb{R}))$ .

Given a function  $v \in C^r(\mathbb{R}/\mathbb{Z}, \mathbb{R})$ , consider the Schrödinger cocycle

$$S_{v,E}(x) = \begin{pmatrix} v(x) - E & -1 \\ 1 & 0 \end{pmatrix} \in C^r(\mathbb{R}/\mathbb{Z}, SL(2, \mathbb{R})),$$

where  $v$  is usually called the potential and  $E$  is the energy.

Let  $G(x) = \{x^{-1}\}$  be the Gauss map. We say that a real number  $\alpha$  is  $(\kappa, \tau)$ -Diophantine if  $|q\alpha - p| > \kappa|q|^{-\tau}$  for all integers  $p$  and  $q \neq 0$ . We say that  $\alpha$  is *recurrent Diophantine* if for some positive  $\kappa$  and  $\tau$ , we have that  $G^n(\alpha)$  is  $(\kappa, \tau)$ -Diophantine for some infinitely many  $n \in \mathbb{N}$ .

Avila and Krikorian [17] showed that for recurrent Diophantine  $\alpha$  and a  $C^r$  potential  $v : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}$  with  $r = \omega, \infty$  for almost every energy  $E$  the cocycle  $(\alpha, S_{v,E})$  is either nonuniformly hyperbolic or  $C^r$ -reducible.

This result resembles the Palis conjecture discussed in Section 3.4.4.

**2.3.7. Existence and uniqueness for differential equations.** Orlicz [133] proved that for a residual set of continuous functions  $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ , the ordinary differential equation  $dy/dx = f(x, y)$  has uniqueness of all solutions (that is, it does not have two distinct maximal solutions with the same initial condition). Lasota and Yorke [100] proved that the property of local existence for all initial conditions also holds for a residual set of  $f$ . They also extend these results to the case that  $y$  lies in a Banach space, in the following sense: for each  $\sigma$ -compact subset  $S$  of the Banach space, there is a residual set of continuous  $f$  for which every initial condition in  $S$  has a unique maximal solution. Many topologically generic existence and uniqueness results have also been proved for various spaces of initial value problems in partial differential equations (e.g. [4]) and stochastic differential equations (e.g. [68]).

**2.3.8. Nondegeneracy of double-periodic solutions of nonlinear Cauchy–Riemann equations.** Consider double-periodic solutions of quasilinear Cauchy–Riemann equations for maps with values in a compact complex manifold  $\mathcal{M}$

$$u_{\bar{z}} + f(z, u) = 0, \quad z = x + iy = \mathbb{C}/(\mathbb{Z} + i\mathbb{Z}), \quad u \in \mathcal{M},$$

where  $f$  is a vector field on  $\mathcal{M}$ , sufficiently smooth in  $z$  and  $u$ , and

$$u_{\bar{z}} = (\partial/\partial\bar{z})u = \frac{1}{2}(u_x + iu_y) \in T_y\mathcal{M}.$$

Consider such solutions where  $u : \mathbb{T}^2 \rightarrow \mathcal{M}$  is contractible and the manifold  $\mathcal{M}$  is Kählerian. Suppose also that  $\|f(z, u)\|_u \leq \sqrt{V}$  for each  $z \in \mathbb{T}^2$ ,  $u \in \mathcal{M}$ , where  $\|\cdot\|_u$  is normalized Kählerian metric and  $V = \inf\{\text{Area } g(\mathbb{S}^2) : g : \mathbb{S}^2 \rightarrow \mathcal{M} \text{ is a nontrivial holomorphic map}\}$ . Kuksin [99] proved that for an open dense (generic) set of  $f$  with the above properties, all solutions of the above equation are nondegenerate, i.e. for  $f'$  close to  $f$  there is only one solution  $u'$  close to  $u$  and  $u'$  smoothly depends on  $f'$ .

## 2.4. Open problems: Generic results on linear spaces

In this section we present several examples of properties on linear spaces that are known to be topologically generic, but for which we know no prevalent analogue. It would be interesting to know whether these properties are prevalent, shy, or neither.

**2.4.1. Generic invariant measures of hyperbolic invariant sets.** Let  $f : M \rightarrow M$  be a diffeomorphism of a compact manifold. Recall that a point is called non-wandering if for every neighbourhood  $U$  of  $x$  there exists  $n > 0$  with  $f^n(U) \cap U \neq \emptyset$ . Denote by  $\Omega$  the set of non-wandering points of  $f$ . Recall that  $f$  satisfies Axiom A if periodic points of  $f$  are dense in  $\Omega$  and the tangent bundle to  $\Omega$ , denoted  $T_\Omega M$ , has continuous splitting  $T_\Omega M = E^s \oplus E^u$ , invariant under the derivative  $df$ , such that  $df$  is contracting on  $E^s$  and  $df^{-1}$  is contracting on  $E^u$ . Smale's decomposition theorem says that  $\Omega$  is the disjoint union of finitely many closed invariant subsets on each of which  $f$  is topologically transitive.

Sigmund studied the space  $\mathcal{M}$  of invariant measures on such a component of  $\Omega$ . In [163] he showed that a generic element of  $\mathcal{M}$  with respect to the weak topology is non-atomic, positive on open sets, ergodic, and has zero entropy, but the set of strongly mixing measures is meagre. In [164] he continued the study and proved that the strongly mixing measures and those for which the system is a Bernoulli shift are dense, and the weakly mixing ones are generic. He also shows that the measures with positive entropy are dense in  $\mathcal{M}$ .

**2.4.2. Operators with singular continuous spectrum.** Let  $C_\infty(\mathbb{R}^n)$  be the space of infinitely differentiable functions vanishing at infinity in  $\|\cdot\|_\infty$ . For a continuous function  $V \in C_\infty(\mathbb{R}^n)$ , let  $\mathcal{S}(V)$  be the Schrödinger operator  $-\Delta + V$  on  $L^2(\mathbb{R}^n)$ . Simon [165] proved that for a Baire generic  $V$  we have that  $\mathcal{S}(V)$  has singular continuous spectrum on  $(0, \infty)$ .

In [166], Simon found a similar phenomenon for Laplace–Beltrami operators on  $\mathbb{S}^1 \times \mathbb{R}$ . In coordinates  $(x, \theta) \in \mathbb{R} \times \mathbb{S}^1 = M$ , consider metrics of the form  $dx^2 + f(x)^2 d\theta^2$ , where  $1 \leq f(x) \leq 2$  and  $f(x)$  is  $C^\infty$  smooth with the metric topology associated with  $\{\sup_{|x| \leq N} \|d^m f/dx^m\|_\infty\}_{m=0,1,\dots,N=1,2,\dots}$ . Let  $X$  be the set of such  $f$ . Then Simon proved that for a generic  $f$  in  $X$ , the Laplacian operator has spectrum  $[0, \infty)$  and is purely singular continuous.

Consider spectral results of prevalent flavour. Let  $H$  be the following operator on  $l_2(\mathbb{Z})$  depending on three parameters  $\lambda, \alpha, \theta$ , defined by

$$[H(\lambda, \alpha, \theta)u](n) = u(n+1) + u(n-1) + \lambda \cos(2\pi\alpha n + \theta)u(n).$$



Avron–Simon [19] showed that for an irrational number with the property  $|\alpha - p_k/q_k| < k^{-q_k}$  for a sequence  $q_k \rightarrow \infty$  and fixed  $\lambda > 2$  for a.e.  $\theta$ ,  $H(\lambda, \alpha, \theta)$  has purely singular continuous spectrum. Jitomirskay–Simon [78] proved the same property of the spectrum for generic  $\theta$ s, an irrational  $\alpha$ , and  $\lambda > 2$ ; their paper also contains further references of the occurrence of singular continuous spectra in the theory of Schrödinger operators.

Damanik, Killip, and Lenz [38] looked at the spectrum of  $H(\lambda, \alpha, \theta)$  with potential  $\cos(2\pi\alpha n + \theta)u(\theta)$  replaced by  $\chi_{[1-\alpha, 1)}(\alpha n + \theta \bmod 1)$ . For  $\lambda \neq 0$ , irrational  $\alpha$ , and all  $\theta$  the spectrum is again purely singular continuous.

Presence of a singular continuous spectrum for a Schrödinger type operator seems accidental and undesirable; however, it is generic.

**2.4.3. Prevalence and genericity of rapid mixing.** Let  $(\Sigma, \sigma)$  be a one-sided subshift of a finite type with transition matrix  $Q$  which is topologically mixing. Consider the distance  $d_\theta(\omega, \bar{\omega}) = \theta^k$ , where  $k = \max\{j : \omega_i = \bar{\omega}_i \text{ for } i \leq j\}$  and  $0 < \theta < 1$ . Let  $C_\theta(\Sigma)$  be the space of Lipschitz functions with respect to  $d_\theta$  endowed with the norm  $\|h\|_\theta = \max\{\|h\|_{C^0}, L(h)\}$ , where  $L(h)$  is the Lipschitz constant of  $h$ . If  $\tau \in C_\theta(\Sigma)$  is positive, we define the suspension flow  $S_\tau$  with the roof function  $\tau$  as follows. Let

$$\Sigma^\tau = \{(\omega, t) : 0 \leq t \leq \tau(\omega)\} / (\omega, \tau(\omega)) \sim (\sigma(\omega), 0),$$

and  $S^t(\omega, s) = (\omega, t + s)$  subject to the identification in  $\Sigma^\tau$ . Denote elements of  $\Sigma^\tau$  by  $q$ . To a  $d_\theta$ -Lipschitz function  $F$  on  $\Sigma^\tau$  one can naturally associate its Gibbs measure  $\mu_F$ . Recall that a function on  $\mathbb{R}_+ \ni t$  is called a *Schwartz function* if it goes to zero as  $t \rightarrow +\infty$  faster than any inverse power of  $t$  and so do all its derivatives. We say that the flow  $\{S^t\}$  is *rapidly mixing with respect to a function  $F$* , if for any pair of continuous functions  $A$  and  $B$  we have that the function

$$R_{A,B}(\mu_F, t) = \int A(q) B(S^t(q)) d\mu_F(q) - \int A(q) d\mu_F(q) \int B(q) d\mu_F(q),$$

which measures correlations, is a Schwartz function of  $t$ . The flow  $\{S^t\}$  is called *rapidly mixing* if it is rapidly mixing with respect to every  $d_\theta$ -Lipschitz function  $F$ . Denote the set of rapidly mixing flows by  $\mathcal{RM}$ .

Dolgopyat [42, Cor. 2] proved that for a generic 1-parameter family of roof functions  $\{\tau_s\}$ , for the corresponding flow denoted  $\{S_s^t\}$ , we have

$$\text{mes} \{ \varepsilon : \{S_s^t\} \notin \mathcal{RM} \} = 0.$$

Next, let  $\{S^t\}$  be a smooth flow on a manifold  $M$  preserving a measure  $\mu$ . The flow  $\{S^t\}$  is rapidly mixing with respect to  $\mu$  if  $R_{A,B}(\mu, t)$  is a Schwartz function of  $t$  for all  $A, B \in C^\infty(M)$ , and the corresponding map from  $C^\infty(M) \times C^\infty(M)$  to the space of Schwartz functions is continuous. If, in addition,  $R_{A,B}(\mu, t)$  decays exponentially with  $t$ , we say that  $\{S^t\}$  is exponentially mixing. In the above notations for  $\tau \in C_\theta(\Sigma)$  define  $\tau_n(\omega) = \sum_{j=0}^{n-1} \tau(\sigma^j \omega)$  and say that  $\tau$  is *eventually positive* if there exists  $n$  such that  $\tau_n > 0$ . The set of eventually positive roof functions is clearly open so we can talk about a generic property inside this open subset of  $C_\theta(\Sigma)$ .



Enlarging the space of roof functions to those which are eventually positive, Dolgopyat [43, Thm. 1.1] showed that *the set of roof functions such that  $S^t(\tau)$  is exponentially mixing contains an open dense subset of the eventually positive elements of  $C_\theta(\Sigma)$ .*

In this direction there is the following important conjecture.

**CONJECTURE.** *For each  $r > 1$ , the set of exponentially mixing Axiom A flows contains a  $C^r$ -open and dense subset of all Axiom A flows.*

Field, Melbourne, and Török [52] extended the latter result of Dolgopyat to show that *amongst  $C^r$  Axiom A flows,  $r \geq 2$ , there is a  $C^2$ -open  $C^r$ -dense set of flows, for which each nontrivial hyperbolic set is rapidly mixing.*

**2.4.4. Minimal Mather measures for convex Lagrangian systems.** Let  $TM$  be a tangent bundle of a closed compact manifold  $M$ , and  $L : TM \times \mathbb{T} \rightarrow \mathbb{R}$  be a  $C^r$  smooth time-periodic Lagrangian that is convex and superlinear in the fibre and has complete Euler–Lagrange flow.

For Euler–Lagrange flow satisfying a Lagrangian  $L$  as above, Mather [112] defined a rich family of invariant measures, which he called *minimal (or action-minimizing) measures*. Roughly, for each ‘rotation vector’ (each cohomology class on  $M$  to be more exact) there is at least one such a measure. Supports of these invariant measures provide invariant sets for the Euler–Lagrange flow which extend to both KAM tori and cantori.

We say that a property is generic in the sense of Mañé if, for each Lagrangian  $L$  satisfying the above hypothesis, there is a generic set of functions  $\mathcal{O} \subset C^\infty(M \times \mathbb{T}, \mathbb{R})$  such that the property holds for each  $L - u$ ,  $u \in \mathcal{O}$ .

Bernard and Contreras [22], extending Mañé’s result [107], proved that for a generic in the sense of the Mañé Lagrangian system and each ‘rotation vector’ there are only finitely many minimal measures.

### 3. Nonlinear prevalence

In this section we shall discuss a way to define prevalence in infinite dimensional spaces without natural linear structure. Examples of such spaces include the spaces of maps from one (compact) manifold into another, or of selfmaps of a (compact) manifold. For definiteness, consider the space  $C^r(M, N)$  of  $C^r$  smooth maps of a compact manifold  $M$  into a manifold  $N$  with the uniform  $C^r$  topology. In the case of noncompact  $M$  one can consider the Whitney  $C^r$  topology.

One of the first ideas that comes up in defining prevalence in such nonlinear spaces is to foliate this space by finite-dimensional submanifolds and rely on Fubini-type arguments. Here one has to be careful in view of the following paradoxical example due to Katok. We shall use Milnor’s version [122]:

*There exist a Lebesgue measurable set  $E$  with measure one in the unit square  $(0, 1) \times [0, 1]$ , together with a family of disjoint smooth real analytic curves  $\Gamma_\beta$  that fills out this square, so that each curve intersects the set  $E$  in at most a single point.*

The curves in question depend continuously on the parameter  $\beta \in [0, 1]$  and form a topological foliation of the square. There cannot be smooth dependence on  $\beta$ , which easily follows from the Fubini theorem.

There is a dynamical realization of this example. Let  $A_2$  be the automorphism of the 2-torus  $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ , given by  $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ . Let  $A_3$  be the automorphism of the 3-torus  $\mathbb{T}^3 = \mathbb{R}^3/\mathbb{Z}^3$ , given by  $\begin{pmatrix} A_2 & 0 \\ 0 & 1 \end{pmatrix}$ . Ruelle and Wilkinson [151], using results of Shub and Wilkinson [161], proved that there is a  $C^1$ -open set  $U$  of  $C^2$ -smooth volume-preserving diffeomorphisms of the 3-torus close to  $A_3$ , such that for each  $g \in U$  the central foliation  $W_g^c$  is *absolutely singular*. Moreover, there is a set  $S \subset \mathbb{T}^3$  of full measure and an integer  $k$  such that  $S$  meets every leaf of  $W_g^c$  in exactly  $k$  points.

### 3.1. Nonlinear prevalence without an underlying linear structure

Now we turn to a definition of prevalence in the space of maps, due to Kaloshin [80], which does not use any linear structure. In the next section we shall discuss another way to define prevalence in the space of maps, due to Tsujii [176], which relies on a linear structure given by an exponential map.

Let  $B^n \subset \mathbb{R}^n$  denote the unit ball. Consider the space  $C^r(M \times B^n, N)$  of  $n$ -parameter families  $\{f_\varepsilon\}_{\varepsilon \in B^n}$  of maps from  $C^r(M, N)$  that depend smoothly on the parameter  $\varepsilon$ .

**DEFINITION 3.** A set  $P \subset C^r(M, N)$  is called strictly  $n$ -prevalent if for a positive integer  $n$  there is an open dense set of  $n$ -parameter families  $Fm(P) \subset C^r(M \times B^n, N)$  such that the following conditions are satisfied:

- (1) for every  $n$ -parameter family  $\{f_\varepsilon\}_{\varepsilon \in B^n} \in Fm(P)$  we have

$$mes\{\varepsilon \in B^n : f_\varepsilon \notin P\} = 0, \quad (1)$$

where  $mes$  is the  $n$ -dimensional Lebesgue measure;

- (2) for every  $f \in C^r(M, N)$  there is a family  $\{f_\varepsilon\}_{\varepsilon \in B^n} \in Fm(P)$  such that  $f = f_0$ .

A set  $P \subset C^r(M, N)$  is called  $n$ -prevalent if  $P$  can be represented as the intersection of a countable number of strictly  $n$ -prevalent sets.

**REMARK 7.** Another possibility to define ‘nonlinear’ prevalence is to say that a set is prevalent if in every local coordinate chart it is prevalent with respect to a local linear structure in such a chart. See Quinn–Sard [147], where they call a set in the linear space *0-preconull* if it is finite-dimensionally shy with linear perturbations, and a set in the manifold is *0-conull* if it is a countable union of sets that are each *0-preconull* in every local coordinate chart. However, this definition is quite strong and may be difficult to apply. See also Tsujii’s definition in the next section.

Straightforward application of the definition shows that if  $P$  is  $n$ -prevalent for some  $n$ , then it is dense, and that a countable intersection of  $n$ -prevalent sets is  $n$ -prevalent; these are the analogues of conditions (3) and (4) for negligible sets in Section 2. In addition, we have the following analogue of condition (5).

**LEMMA 3.1.** *Let  $P \subset \mathbb{R}^k$  be  $n$ -prevalent for some  $n < k$ . Then  $P$  has full  $k$ -dimensional Lebesgue measure.*

PROOF. Suppose  $P$  does not have full measure. Then there is a density point  $f$  of the complement  $S = \mathbb{R}^n \setminus P$ . By conditions (2) and (1) of Definition 3, there is a family  $\{f_\varepsilon\}_{\varepsilon \in B^n}$  in  $Fm(P)$  such that  $f = f_0$  and (1) holds. Since  $Fm(P)$  is open, sufficiently close families  $\{\tilde{f}_\varepsilon\}_{\varepsilon \in B^n}$  also satisfy (1). In particular, there is a nondegenerate family  $\{f_\varepsilon^*\}_{\varepsilon \in B^n}$  passing through  $f = f_0^*$ , where nondegeneracy means that  $df_0^*$  has full rank. Now by the Implicit Function Theorem, there is a local coordinate system in  $\mathbb{R}^k$  near  $f$  such that  $\bigcup_{\varepsilon \in B^n} f_\varepsilon^*$  near  $f = f_0^*$  is a plane of dimension  $n$ . Consider nearby planes. They are also given by nearby  $n$ -parameter families. Application of the Fubini Theorem shows that in a small neighbourhood of  $f$  the set,  $P$  has full measure. Therefore, its complement  $S$  has measure zero. This contradiction proves the lemma.  $\square$

The finite-dimensional case is one of the natural tests of the definition of prevalence. Consider a set  $P \subset \mathbb{R}^k$  that is  $n$ -prevalent for some  $n < k$ . It turns out that without condition (2),  $P$  might have an arbitrarily small  $k$ -dimensional Lebesgue measure. For simplicity assume  $k = 2$  and  $n = 1$ . In this context, a  $C^r$  smooth 1-parameter family  $\{f_\varepsilon\}_{-1 \leq \varepsilon \leq 1}$  is the range of a  $C^r$  smooth map from  $[-1, 1]$  to  $\mathbb{R}^2$ . Then the countable set  $S = \{Q_i\}_{i \in \mathbb{N}}$  of all polynomial maps  $Q_i : [-1, 1] \rightarrow \mathbb{R}^2$  with rational coefficients, is a  $C^r$  dense set of 1-parameter families. For each sequence  $\varepsilon = \{\varepsilon_i\}_{i \in \mathbb{N}}$  of positive real numbers, let  $S_\varepsilon = \bigcup_{i \in \mathbb{N}} \mathcal{B}_{\varepsilon_i}$ , where  $\mathcal{B}_{\varepsilon_i}$  is the open  $\varepsilon_i$ -ball around  $Q_i$ ; then  $S_\varepsilon$  is open and dense in  $C^r([-1, 1], \mathbb{R}^2)$ . Next, let  $P_\varepsilon$  be the union of the ranges of all maps in  $S_\varepsilon$ . Then  $P_\varepsilon$  satisfies condition (1) of Definition 3 with  $Fm(P_\varepsilon) = S_\varepsilon$ . The union of the ranges of all maps in  $\mathcal{B}_{\varepsilon_i}$  is a neighbourhood of the curve  $Q_i([-1, 1])$ , and its Lebesgue measure can be made arbitrarily small by making  $\varepsilon_i$  arbitrarily small. Therefore, the measure of  $P_\varepsilon$  can be arbitrarily small.

### 3.2. Nonlinear prevalence with an underlying linear structure

Now we turn to a definition of prevalence in the space of maps, due to Tsujii [176], which relies on a linear structure given by an exponential map.

Let  $M$  be a compact manifold of dimension  $m$  and let  $\pi : V \rightarrow M$  be a  $C^\infty$  vector bundle of dimension  $p$  over  $M$ . We denote the set of  $C^r$  sections of the vector bundle  $V$  by  $\Gamma^r(V)$ , which is endowed with the  $C^r$  norm and  $C^r$  topology. Then there are natural inclusions of Banach spaces:

$$\Gamma^0(V) \supset \Gamma^1(V) \supset \Gamma^2(V) \supset \dots$$

Notice that in this sequence each space is dense in the bigger spaces and the Borel  $\sigma$ -algebra on it coincides with the restrictions of those in the bigger spaces.

Let  $\tau_\varphi : \Gamma^0(V) \rightarrow \Gamma^0(V)$  be the translation by  $\varphi \in \Gamma^0(V)$ , i.e.  $\tau_\varphi(f) = f + \varphi$ . We say a Borel probability measure  $\mu$  on  $\Gamma^0(V)$  is *quasi-invariant along the subspace*  $\Gamma^r(V)$  if  $\tau_\varphi(\mu)$  is equivalent to  $\mu$  for any element  $\varphi \in \Gamma^r(V)$ , and we denote the set of such measures by  $\mathcal{M}_r$ . Put  $\mathcal{M}_\infty = \bigcup_{r < +\infty} \mathcal{M}_r$ . One can show that the set  $\mathcal{M}_r$  is not empty for sufficiently large  $r$  (see [176, Lemma A, Sect. 4]).

In Euclidean spaces, a  $\sigma$ -finite measure is equivalent to the Lebesgue measure if and only if it is quasi-invariant along the whole space. However, it is known that there is no

such measure for infinite dimensional space [57,171,172,180] (see also the discussion in the beginning of Section 2). Thus it is reasonable to consider the measures that are quasi-invariant along dense subspaces and is a motivation for Tsujii's definition.

Denote by  $B(\Gamma^r(V))$  the Borel  $\sigma$ -algebra on  $\Gamma^r(V)$  and define a collection of sets

$$\tilde{Z}(\Gamma^r(V)) = \{E \in B(\Gamma^r(V)) : \mu(E) = 0 \text{ for every } \mu \in \mathcal{M}_\infty\}.$$

Let  $D(V)$  be the group of  $C^\infty$  diffeomorphisms  $\psi : V \rightarrow V$  satisfying  $\pi \circ \psi = \pi$  and let  $\psi_*$  be the action of the element  $\psi \in D(V)$  on  $\Gamma^r(V)$  such that

$$\text{graph}(\psi_*(\phi)) = \psi(\text{graph}(\phi)), \quad \phi \in \Gamma^r(V).$$

Introduce

$$Z(\Gamma^r(V)) = \bigcap_{\psi \in D(V)} \psi_*(\tilde{Z}(\Gamma^r(V))).$$

Now consider the space  $C^r(M, N)$  as above. Choose a  $C^\infty$  metric on  $N$  and for any  $f \in C^\infty(M, N)$  on a neighbourhood  $U_f$  of its zero section define a homeomorphism

$$\Phi_f : \Gamma^r(f^*TN) \rightarrow C^r(M, N)$$

by  $\Phi_f(h)(x) = \exp_{f(x)}(h(x))$ . Then the coordinate system  $\{(\Phi_f, U_f), f \in C^\infty(M, N)\}$  shows that  $C^r(M, N)$  is a Banach manifold (see e.g. [1]).

For the space  $C^r(M, N)$ , let  $Z(C^r(M, N))$  be the family of Borel subsets  $E \subset C^r(M, N)$  such that the set  $\Phi_f^{-1}(E \cap \Phi_f(U_f))$  belongs to  $Z(\Gamma^r(f^*TN))$  for every  $f \in C^\infty(M, N)$ . Since  $Z(\Gamma^r(V))$  is invariant under the action of  $D(V)$ , the definition of  $Z(C^r(M, N))$  does not depend on the choice of  $C^\infty$  metric on  $N$  nor on the choice of the neighbourhoods  $U_f$ .

**DEFINITION 4.** Call a Borel set  $E \subset C^r(M, N)$  *T-shy* if  $E$  belongs to  $Z(C^r(M, N))$ . Call a non-Borel set shy if it is contained in a shy Borel set. Call a set *T-prevalent* if its complement is shy.

**LEMMA 3.2** ([176, Lemma A, Section 1]). *A countable union of T-shy sets is T-shy. If a Borel set  $E$  is contained in a T-shy set then  $E$  is T-shy. A T-shy subset  $E$  has no interior with respect to the  $C^r$  topology.*

**THEOREM 3.3** ([176, Theorem B]). *If  $E$  is T-shy, then for each probability measure  $\lambda$  on  $[0, 1]^n$ , the set  $S_{E,\lambda} \subset C^r(M \times [0, 1]^n, N)$  is T-shy, where*

$$S_{E,\lambda} = \{F(x, \varepsilon) \in C^r(M \times [0, 1]^n, N) : \lambda(\{\varepsilon \in [0, 1]^n : F(\cdot, \varepsilon) \in E\}) > 0\}.$$

**REMARK 8.** In other words, if property  $E$  is *T-shy*, then for an  $n$ -parameter family  $F(x, \varepsilon)$  from a *T-prevalent* set of  $n$ -parameter families, the measure of the set of parameters corresponding to  $F(\cdot, \varepsilon) \in E$  is zero. This is an analogue of (1).

### 3.3. Examples of nonlinear prevalence

Probably, the key property for proving the application of either of these definitions of nonlinear prevalence to dynamical systems is Thom's transversality theorem (see [80, Theorem 1.2] or its analogue [176, Theorem C]), which asserts the following: Let  $M$  and  $N$  be smooth manifolds and let  $K$  be a smooth submanifold of the space of  $m$ -jet extensions  $J^m(M, N)$  for some  $m < r$ . Then the set of maps  $f$  in  $C^r(M, N)$  such that the  $m$ -jet extension  $j^m f$  is transversal to  $K$  is  $n$ -prevalent for any  $n > \dim J^m(M, N)$ .

Application of this theorem leads to a proof that many properties are  $n$ -prevalent for some  $n$ . This includes a multi-jet transversality theorem, Whitney's embedding theorem, and Mather's stability theorem (see [81, 80]). Also, Tsujii [177] proved that for a dense set of  $n$ -parameter families of  $C^r$ -smooth orientation-preserving circle diffeomorphisms ( $n, r \geq 1$ ), for almost every parameter value, the corresponding diffeomorphism has Diophantine rotation number.

**3.3.1. Prevalence of fat solenoidal attractors.** Consider a so-called generalized baker transform

$$B : [-1, 1] \times [-1, 1] \rightarrow [-1, 1] \times [-1, 1],$$

$$B(x, y) = \begin{cases} (2x - 1, \beta y + (1 - \beta)) & \text{for } x \geq 0 \\ (2x + 1, \beta y - (1 - \beta)) & \text{for } x < 0. \end{cases}$$

For  $\beta = 1/2$  this is the standard baker transform. Alexander and Yorke [3] studied the case  $1/2 < \beta \leq 1$  and discovered that  $B$  admits an absolutely continuous invariant measure (acim) provided  $\beta$  satisfies a certain numerical condition. Later Solomyak [170] proved the existence of an acim for almost every  $\beta$  in  $(1/2, 1]$ .

Consider now a class of dynamical systems generated by maps

$$T : \mathbb{S}^1 \times \mathbb{R} \rightarrow \mathbb{S}^1 \times \mathbb{R}, \quad T(x, y) = (\ell x, \lambda y + f(x)),$$

where  $\ell$  is an integer,  $0 < \lambda < 1$  is a real number and  $f$  is a  $C^2$  function on  $\mathbb{S}^1$ . This map is a skew product on the expanding map  $\tau : x \rightarrow \ell x$ . One can show the existence of an ergodic probability measure  $\mu$  on  $\mathbb{S}^1 \times \mathbb{R}$ , for which Lebesgue almost every point  $x \in \mathbb{S}^1 \times \mathbb{R}$  is generic, i.e.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i(x)} = \mu \quad \text{weakly.}$$

This measure  $\mu$  is usually called the *SBR measure* for  $T$ . A natural question is the smoothness of the SBR measure  $\mu$  with respect to the Lebesgue measure on  $\mathbb{S}^1 \times \mathbb{R}$ . In the case  $\lambda \ell < 1$ , the SBR measure is totally singular with respect to the Lebesgue measure, because  $T$  contracts the area. The case  $\lambda \ell > 1$  corresponds to  $\beta > 1/2$  for the generalized baker's transform and is definitely more interesting. Fix  $\ell$  and denote by  $\mathcal{D} \subset (0, 1) \times C^2(\mathbb{S}^1, \mathbb{R})$  the set of pairs  $(\lambda, f)$  for which the SBR measure is absolutely

continuous with respect to Lebesgue measure. Denote the interior of  $\mathcal{D}$  with respect to the natural product topology by  $\mathcal{D}^\circ$ .

Tsujii [178] proved that for  $\ell^{-1} < \lambda < 1$  there exists a finite collection of  $C^\infty$  functions  $\phi_i : \mathbb{S}^1 \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$ , such that, for every  $C^2$  function  $g \in C^2(\mathbb{S}^1, \mathbb{R})$ , the subset of  $\mathbb{R}^m$

$$\left\{ (t_1, \dots, t_m) \in \mathbb{R}^m : \left( \lambda, g(x) + \sum_{i=1}^m t_i \phi_i(x) \right) \notin \mathcal{D}^\circ \right\}$$

has Lebesgue measure zero.

Since  $\mathbb{R}$  has natural linear structure, by considering only perturbations to the  $\mathbb{R}$ -component of  $T$  we can use Definition 1 of linear prevalence. The above result, in particular, says that the set  $\mathcal{D}^\circ$  is prevalent. Notice also that this result implies that  $\mathcal{D}$  is open and dense. So this is another example of a topologically generic and prevalent set.

Later, Tsujii [179] extended the aforementioned result to a more general class of maps. Let  $\mathbb{T}^2$  be a two dimensional torus equipped with a metric  $\|\cdot\|$  and the Lebesgue measure. Call a  $C^1$ -mapping  $F : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  a *partially hyperbolic endomorphism* if there are positive constants  $\lambda$  and  $c$  and a continuous decomposition of the tangent bundle  $T\mathbb{T}^2 = E^c \oplus E^u$  with  $\dim E^c = \dim E^u = 1$ , such that

- (1)  $\|DF^n|_{E^u(z)}\| > \exp(\lambda n - c)$ ;
- (2)  $\|DF^n|_{E^c(z)}\| < \exp(-\lambda n + c)$

for all  $z \in \mathbb{T}^2$  and  $n \geq 0$ . The subbundles  $E^c$  and  $E^u$  are called the central and unstable subbundles, respectively. Notice that partially hyperbolic  $C^r$ -endomorphisms on  $\mathbb{T}^2$  form an open subset in the space  $C^r(\mathbb{T}^2, \mathbb{T}^2)$ , provided  $r \geq 1$ .

An invariant Borel probability measure  $\mu$  for  $F : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  is called a *physical measure* if its basin of attraction

$$\mathcal{B}(\mu) = \left\{ z \in \mathbb{T}^2 : \frac{1}{n} \sum_{i=0}^{n-1} \delta_{F^i(z)} \rightarrow \mu \text{ weakly as } n \rightarrow \infty \right\}$$

has positive Lebesgue measure. Tsujii [179] shows that a Baire generic partially hyperbolic  $C^r$ -endomorphism on  $\mathbb{T}^2$  admits finitely many ergodic physical invariant measures whose union of basins of attraction has full Lebesgue measure, provided that  $r \geq 19$ .

In this paper Tsujii also shows that partially hyperbolic  $C^r$ -endomorphisms on  $\mathbb{T}^2$  also form a prevalent set in a certain sense. See [179, Theorem 2.2] for the precise statement.

**3.3.2. Kupka–Smale theorem for diffeomorphisms and vector fields.** In Example 9 of Section 1.2, we discussed the Kupka–Smale property for maps on a compact manifold  $M$ . It can also be stated for the space of  $C^r$  vector fields on  $M$ . Kaloshin [80] proved a prevalent Kupka–Smale theorem for  $C^r$  diffeomorphisms and  $C^r$  vector fields, using a slight variation of Definition 3 above. Here is the precise statement.

In the notation of Definition 3, call a set  $P$  strictly  $(\delta, n)$ -prevalent for a positive integer  $n$  if for any  $\delta > 0$  there is an open dense set of  $n$ -parameter families  $Fm(P, \delta)$  satisfying

(1) with the measure bounded above by  $\delta$ , instead of being zero, and (2) with  $Fm(P)$  replaced by  $Fm(P, \delta)$ . Call a set  $P$  a  $(\delta, n)$ -prevalent if it can be represented as the intersection of a countable number of strictly  $(\delta, n)$ -prevalent sets.

Similarly to Lemma 3.1, one can show that a  $(\delta, n)$ -prevalent subset of a finite dimensional space has full Lebesgue measure.

This definition can be given for both the space of  $C^r$  diffeomorphisms of a compact manifold  $M$  and the space of  $C^r$  vector fields on  $M$ .

Kaloshin [80] proved that the Kupka–Smale property is  $(\delta, n)$ -prevalent both for  $C^r$ -diffeomorphisms and for  $C^r$ -vector fields.

**3.3.3. Growth of the number of periodic points.** Let  $M$  be a compact manifold of dimension at least 2 and  $\text{Diff}^r(M)$  be the space of  $C^r$ -smooth diffeomorphisms of  $M$ . Associate to each diffeomorphism the sequence  $P_n(f)$  of the number of isolated periodic points for  $f$  of period  $n$

$$P_n(f) = \#\{\text{isolated } x \in M : f^n(x) = x\}.$$

Surprisingly Kaloshin [82], using Gonchenko, Shilnikov, and Turaev [58], exhibits for each sequence  $\{a_n\}_{n \in \mathbb{Z}_+}$  an open set  $\mathcal{N}$  in the space of diffeomorphisms  $\text{Diff}^r(M)$  such that for a Baire generic diffeomorphism  $f \in \mathcal{N}$  the number of periodic points  $P_n(f)$  grows with the period  $n$  faster than  $\{a_n\}_{n \in \mathbb{Z}_+}$  along a subsequence, i.e.  $P_{n_i}(f) > a_{n_i}$  for some  $n_i \rightarrow \infty$  with  $i \rightarrow \infty$ . This result gives a negative answer to a long standing question of Smale [167].

In the case of surface diffeomorphisms, i.e.  $\dim M = 2$ , an open set  $\mathcal{N}$  with supergrowth of the number of periodic points turns out to be a Newhouse domain, which is defined in Section 3.4.4.

However, it turns out that for each  $r > 1$  and  $\delta > 0$ , for a prevalent diffeomorphism  $f \in \text{Diff}^r(M)$ , the growth of the number of periodic points is bounded as  $P_n(f) < C \exp(n^{1+\delta})$ . Here a different notion of a prevalent diffeomorphism is involved, as described in [83, 84].

### 3.4. Open problems: generic results in nonlinear spaces

In this section we present topologically generic results in nonlinear spaces, for which it would be interesting to prove analogues using some form of nonlinear prevalence as described in the preceding sections.

**3.4.1. Genericity of ergodicity, mixing, weak mixing.** Let  $(M, \mu)$  be a topological space  $M$  with a probability measure  $\mu$  on it. Let  $\mathcal{G}(M, \mu)$  (respectively,  $\mathcal{M}(M, \mu)$ ) be the space of  $\mu$ -preserving automorphisms (i.e. measure preserving bijections), endowed with the weak topology (respectively, uniform topology). An automorphism  $f \in \mathcal{M}(M, \mu)$  is called *ergodic* if for every pair of measurable sets  $A$  and  $B$  we have that  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(f^k(A) \cap B) = \mu(A)\mu(B)$ ; *weakly mixing* if  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(f^k(A) \cap B) - \mu(A)\mu(B)| = 0$ ; and *strongly mixing* if  $\lim_{n \rightarrow \infty} \mu(f^k(A) \cap B) = \mu(A)\mu(B)$ ;

Oxtoby and Ulam [139] showed that ergodicity is generic in  $\mathcal{M}(M, \mu)$ . Later Halmos proved that ergodicity [63] as well as weak mixing [64] are generic in  $\mathcal{G}(M, \mu)$ . However, the complement to strong mixing is generic in  $\mathcal{G}(M, \mu)$ , as was shown by Rokhlin [150]. It took many years before Katok and Stepin [89] proved that weak mixing is generic in  $\mathcal{M}(M, \mu)$ . A modern treatment of this topic can be found in [5].

Going back to von Neumann [35] to study finer properties of measure preserving automorphisms, it is fruitful to study the spectral theory of the unitary operators associated with those automorphisms. The general problem of studying spectral theory for Schrödinger operators is to distinguish pure point (discrete) spectrum ( $\sigma_{pp}$ ), singular continuous spectrum ( $\sigma_{sc}$ ), and absolutely continuous spectrum ( $\sigma_{ac}$ ). In general the spectral analysis of Schrödinger operators of a transformation is not easy, even though genericity results are relatively easy to find. For example, a singular continuous spectrum is often generic (see [41] and references therein). Concrete examples with purely singular continuous spectra can be found in [19,70,78].

**3.4.2. Genericity of non-Lipschitz Anosov foliations.** Let  $f$  be a diffeomorphism of a compact Riemannian manifold  $M$ . It is called *Anosov* if the tangent bundle splits (necessarily uniquely) into the sum of two  $Df$ -invariant subbundles  $TM = E^u \oplus E^s$ , and there are constants  $C$  and  $a < 1$  such that

$$\|Df^n(5)\| \leq Ca^n \|v\| \quad \|Df^{-n}(u)\| \leq Ca^n \|u\| \quad \text{for } v \in E^s, u \in E^u, n \in \mathbb{N}.$$

This implies that for nearby points either positive or negative iterates move apart exponentially fast. Any hyperbolic matrix  $A \in SL(m, \mathbb{Z})$  induces an Anosov diffeomorphism of the  $m$ -torus  $\mathbb{T}^m$ : the bundles  $E^u$  and  $E^s$  are defined by the expanding and contracting eigenspaces of  $A$ . Fix some  $m \geq 2$ .

Anosov [7] showed that Anosov diffeomorphisms are *structurally stable*: for any two sufficiently  $C^1$ -close Anosov diffeomorphisms  $f$  and  $g$  of  $\mathbb{T}^m$  there is a (unique) homeomorphism  $h : \mathbb{T}^m \rightarrow \mathbb{T}^m$  close to the identity such that  $h \circ f = g \circ h$ . One could ask the question is there is more regularity to the conjugacy  $h$ ? We say  $h$  is uniformly bi- $C^\beta$  for some  $\beta > 0$  if there are positive constants  $C$  and  $d$  with

$$\frac{\|x - x'\|^{1/\beta}}{C} \leq \|h(x) - h(x')\| \leq C \|x - x'\|^\beta$$

for  $\|x - x'\| \leq d$ . It is well known (see e.g. Katok and Hasselblatt [87, Thm. 19.1.2]) that the conjugacy is uniformly bi- $C^\beta$ . However, Hasselblatt and Wilkinson [65, Thm. 3] showed for any  $0 < \alpha < 1$  there is a linear symplectic Anosov diffeomorphism  $A$  and a  $C^k$  neighbourhood  $U$  of  $A$  of symplectic diffeomorphisms, such that for an open dense set of  $f \in U$ , the conjugacy to  $A$  is almost nowhere bi- $C^\alpha$ .

Let

$$\begin{aligned} B^u(f) &:= \inf_{p \in M} \sup\{(\log \mu_s - \log \nu_s) / \log \mu_f : \mu_f < \mu_s < 1 < \nu_s, \\ &\quad \mu_f^n \|v\| / C \leq \|Df^n(v)\| \leq C \mu_s^n \|v\|, \|Df^{-n}(u)\| \leq C \nu_s^{-n} \|u\| \\ &\quad \text{for } v \in E^s(p), u \in E^u(p), n \in \mathbb{N}\}. \end{aligned}$$



Denote by  $C^r$  the space of  $C^{[r]}$  maps whose  $[r]$ -derivatives have modulus of continuity  $O(x^{r-[r]})$ .

With these notations Hasselblatt and Wilkinson [65, Thm. 6] showed that

- If  $B^u(f) \notin \mathbb{N}$  then  $E^u \in C^{B^u(f)}$ , while if  $B^u(f) \in \mathbb{N}$  then  $E^u \in C^{B^u(f)-1, O(x|\log x|)}$ ;
- For an open dense set of symplectic diffeomorphisms the regularity of  $E^u$  is at most as indicated above;
- For an open dense set of diffeomorphisms with  $\text{codim}(E^s) = 1$  the regularity of  $E^u$  is at most as above.

**3.4.3. Genericity of zero Lyapunov exponents outside of hyperbolicity.** Bochi and Viana [25] showed that for a generic volume-preserving diffeomorphism on a compact manifold, almost every orbit is either projectively hyperbolic or has all Lyapunov exponents equal to zero. This result extends the result of Mañé and Bochi [108, 24] for surface diffeomorphisms.

**3.4.4. Newhouse phenomena.** The following results relate to the Newhouse phenomenon of maps that have infinitely many sinks [129, 130].

*Hénon family.* In the 1970s, M. Hénon [66] made an extensive numerical study of the behaviour under iteration of maps  $P_{a,b}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  of the form  $(x, y) \mapsto (1 - ax^2 + by, x)$ , where  $a, b \in \mathbb{R}$ . In particular, Hénon found numerical evidence supporting the existence of a strange attractor for  $P_{a,b}$  when  $a = 1.4$  and  $b = 0.3$ . Van Strien [181] proved that in the parameter plane  $(a, b)$  arbitrarily near  $(a, b) = (2, 0)$  there is an open set  $\mathcal{N}$  such that for a Baire generic parameter in it, the corresponding  $P_{a,b}$  has infinitely many coexisting sinks.

*Polynomial automorphisms of  $\mathbb{C}^2$ .* Buzzard [30], using results of Fornaess and Gavosto [53], showed that for a large enough  $d$  in the space of holomorphic self-maps  $H_d(\mathbb{C}^2)$  of  $\mathbb{C}^2$  of degree  $d$ , there exists an open set  $\mathcal{N} \subseteq H_d(\mathbb{C}^2)$  with the same property as in the previous example.

*Smooth diffeomorphisms.* Let a diffeomorphism  $f \in \text{Diff}^r(M)$  have a saddle periodic orbit  $p$ , where  $M$  is a smooth compact surface and  $r \geq 2$ . Suppose stable  $W^s(p)$  and unstable  $W^u(p)$  manifolds of  $p$  have a quadratic tangency. Such a diffeomorphism  $f$  is called a *diffeomorphism exhibiting a homoclinic tangency*.

Denote by  $\mathcal{HT} \subset \text{Diff}^r(M)$  the set of diffeomorphisms exhibiting a homoclinic tangency. Denote the  $C^r$ -closure of  $\mathcal{HT}$  by  $\mathcal{N}$ . A fundamental result of Newhouse [129] is that  $\mathcal{N}$  has nonempty interior for surface diffeomorphisms. Call  $\mathcal{N}$  a *Newhouse domain*. Its interior is analogous to the open sets from the previous two examples.

On this account we would also like to mention the following conjecture, which is due to Palis [143], about the space of diffeomorphisms of 2-dimensional manifolds:

**CONJECTURE.** *Every surface diffeomorphism  $f \in \text{Diff}^r(M)$  can be approximated by a diffeomorphism that either is hyperbolic or exhibits a homoclinic tangency.*

This conjecture was proved for approximations in the  $C^1$  topology by Pujals and Sambarino [146]. If this conjecture is true, then in the complement of the set of hyperbolic

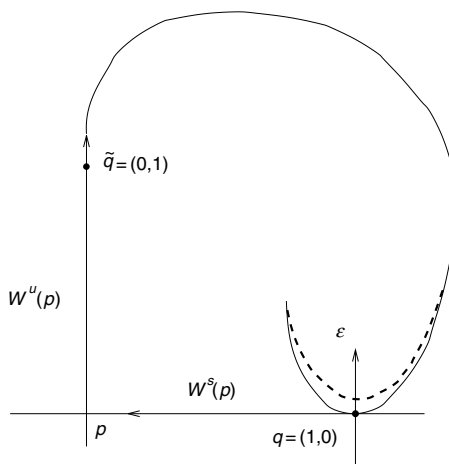


Fig. 2. Homoclinic tangency.

diffeomorphisms, those diffeomorphisms with arbitrarily quick growth of number of periodic orbits form a Baire generic set (see Section 3.3.3).

Unfolding of homoclinic tangencies is far from being understood. A striking phenomenon was discovered by Gonchenko, Shilnikov, and Turaev [58]. They show that *there does not exist a finite number of parameters to describe all bifurcations occurring near a homoclinic tangency*. This suggests that the complete description of bifurcations of diffeomorphisms with a homoclinic tangency is impossible.

Consider a generic one parameter family of perturbations  $\{f_\varepsilon\}_{\varepsilon \in I}$ ,  $I = [-\varepsilon_0, \varepsilon_0]$  of a 2-dimensional diffeomorphism  $f = f_0 \in \text{Diff}^r(M)$  with homoclinic tangency and small  $\varepsilon_0 > 0$  (see Figure 2). Roughly speaking,  $\varepsilon$  parameterizes the oriented distance from the tip of the unstable manifold to the stable manifold. Such a family is called an *unfolding of an HT*.

Palis [140,143] makes the following conjecture.

**CONJECTURE.** *For almost every  $\varepsilon$ , the corresponding  $f_\varepsilon$  has only finitely many sinks.*

Gorodetski and Kaloshin [59] obtained a partial proof of this conjecture, showing that there are only finitely many ‘localized’ sinks of bounded ‘complexity’. It can also be shown that for a Baire generic diffeomorphism in a Newhouse domain, there are infinitely many ‘localized’ sinks of bounded ‘complexity’.

**3.4.5. Stable intersection of two Cantor sets on a line.** Palis [141] conjectured that for a generic pair of Cantor sets  $(K_1, K_2)$  of the real line,  $K_1 - K_2$  either has measure zero or contains an interval. Moreira [125] introduced the concept of stable intersection: two Cantor sets  $K_1$  and  $K_2$  have *stable intersection* if there is a neighbourhood  $V$  of  $(K_1, K_2)$  in the set of pairs of  $C^k$ -regular Cantor sets (see [126, Section 2] for a definition of the topology) such that for  $(\tilde{K}_1, \tilde{K}_2) \in V$  we have  $\tilde{K}_1 \cap \tilde{K}_2 \neq \emptyset$ .

Moreira and Yoccoz [126] showed that there is an open dense set  $U$  in the set of pairs of  $C^k$ -regular Cantor sets such that, if  $(K_1, K_2) \subset U$  with  $HD(K_1) + HD(K_2) > 1$ , then there exists  $t \in \mathbb{R}$  such that  $(K_1, K_2 + t)$  has stable intersection. Moreover, generic  $n$ -parameter families of  $(K_1, K_2)$  are actually contained in  $U$ .

The aforementioned Palis conjecture is motivated by the Newhouse phenomenon discussed in the previous example. It turns out that for some one-parameter families of 2-dimensional diffeomorphisms, there exist invariant hyperbolic sets that are homeomorphic to the product of two Cantor sets. Moreover, the leaves of the unstable and stable manifolds of this invariant set form two families of laminations. Typically as the parameter changes, these two laminations meet and form a persistent tangency. The Newhouse phenomenon arises from this important observation (see e.g. [143, Chapter 6] for more).

**3.4.6. Bumpy metrics.** Let  $M$  be a closed compact manifold and  $g$  be a smooth Riemannian metric on  $M$ . A metric is called *bumpy* if all closed geodesics are non-degenerate. The bumpy metric theorem asserts that the set of  $C^r$  bumpy metrics is a residual subset of the set of all  $C^r$  metrics endowed with the  $C^r$  topology for all  $2 \leq r \leq \infty$ . The bumpy metric theorem is traditionally attributed to Abraham [2], but see also Anosov [8] and Klingenberg and Takens [95].

The set of smooth Riemannian metrics on a given  $M$  has a partial linear structure, namely that a linear combination of such metrics with positive coefficients is also in the set. One could thus formulate a prevalent bumpy metric theorem using the notion of relative prevalence [6] on a convex subset of a vector space, which we discuss in Section 4.2.

**3.4.7. Eigenvalues of the Laplacian for Riemannian metrics with negative curvature.** Sarnak [152] makes the following conjecture, which he attributes to [26]. (Here  $X$  is a compact 2-dimensional Riemannian manifold.)

CONJECTURE 3.4. *Fix  $r \geq 2$  and let  $\mathcal{R}_r(X)$  denote the space of all Riemannian metrics on  $X$  in the  $C^r$  topology. For the generic (in the sense of measure as defined in [74] for example)  $g \in \mathcal{R}_r(X)$  of negative curvature, the local spacing distributions between the eigenvalues of  $\Delta_g$  in the large  $\lambda$  limit, follow the laws of the eigenvalue spacings of random matrices from the Gaussian Orthogonal Ensemble (see [119] for a description of the latter).*

As in the previous example, relative prevalence (see Section 4.2) could be used for this conjecture.

## 4. Other notions of genericity

In this section we survey briefly extensions of prevalence to non-Abelian topological groups and to convex subsets of vector spaces, as well as some notions of genericity that are stronger than topological genericity or prevalence.

### 4.1. Non-Abelian prevalence

Recall that we originally defined prevalence for a complete metric Abelian group  $B$ . In the case where  $B$  is not Abelian, Christensen [34], Topsøe and Hoffmann–Jørgensen [175],

and Mycielski [127] have suggested that the definition of shyness can be generalized by defining transversality of a measure  $\mu$  to a subset  $S \subset B$  to be the property that  $\mu(x + S + y) = 0$  for all  $x$  and  $y$  in  $B$ . In this section we call this generalization ‘strong shyness’; one can also define in the natural way the notions of left transverse and right transverse, and hence left shyness and right shyness.

Conditions (1)–(5) of Section 2 hold for strong shyness, left shyness, and right shyness, but in general left shyness and right shyness are weaker than strong shyness. For example, Shi and Thompson [159] considered the space of orientation-preserving homeomorphisms on  $[0, 1]$  (with composition being the group operation, and the uniform topology), and showed that in this space a set can be both left shy and right shy but not shy in the stronger sense above. Specifically, they proved that for  $0 < \alpha < \infty$ , the set  $S_\alpha$  of homeomorphisms whose derivative at zero exists and equals  $\alpha$  is not shy but is both left shy and right shy. (In fact, the union over all such  $\alpha$  of  $S_\alpha$  is left shy and right shy, yet it contains uncountably many non-shy sets.)

On the other hand, for a set that is invariant under conjugation, strong shyness, left shyness, and right shyness are all equivalent. Thus, for automorphism groups, dynamical properties that are preserved under conjugation can be proved prevalent in the strong sense using perturbations by composition either only on the left or only on the right. For diffeomorphisms, this approach is an alternative to the notions of nonlinear prevalence discussed in Sections 3.1 and 3.2.

Mycielski [127] proposed, among several other problems, to determine whether ergodicity is prevalent in this sense on the space of homeomorphisms from the unit cube in  $\mathbb{R}^n$  to itself, as was proved generically in [139]. Dougherty and Mycielski [46] considered the space of permutations of the natural numbers (again a group under composition) with the topology that two permutations are close if they and their inverses agree for all natural numbers up to some large number  $n$ . They proved the following properties of the dynamics of permutations in this space: a prevalent permutation has only finitely many periodic orbits and infinitely many distinct nonperiodic orbits, whereas for a topologically generic permutation, all numbers are periodic.

#### 4.2. Relative prevalence and applications in mathematical economics

Anderson and Zame [6] developed a notion of prevalence *relative* to a completely metrizable convex subset  $C$  of a topological vector space  $X$ . Their motivation was that the parameter set for many problems in mathematical economics is a shy subset of the natural ambient infinite-dimensional space. For example, many economic models require functions that are nonnegative or concave, and in some function spaces these properties are shy. However, the sets of nonnegative functions and concave functions are convex.

Given the assumptions above on  $C$ , Anderson and Zame [6] define shyness relative to  $C$  in such a way that conditions (1)–(5) of Section 2 hold, and condition (3) is strengthened to say that no *relatively* open subset of  $C$  is shy relative to  $C$ . Using this notion, they proved several prevalent results.

One is that for a prevalent subset of a certain class of games, there are only finitely many pure strategy Nash equilibria; a similar topologically generic result was proved in [47]. In this case the space of games is a Banach space, but the existence of a Nash equilibrium

is guaranteed only on a convex subset of this space, and they also establish prevalence relative to this subset.

Another result in [6] is that in a certain class of commodity exchange models, a prevalent subset has only finitely many equilibria; a similar topologically generic result was proved in [111]. They also proved prevalent existence of equilibria in a different class of continuous-time commodity exchange models for which, by contrast, a topologically generic model was proved [10] to have no equilibria.

Shannon and Zame [157] also proved prevalent finiteness of equilibria and, in addition, continuous dependence of the equilibria on parameters, for a class of commodity exchange models related to those mentioned above, but with infinitely many commodities. This generalizes an earlier result [39] for Lebesgue almost every member of a finite-commodity class. Riedel [149] proved a similar prevalent result for a related class of continuous-time models.

Motivated by such applications, Shannon [156] formulated and proved a transversality theorem for Lipschitz functions in terms of relative prevalence.

### 4.3. Stronger forms of genericity and prevalence

Here we discuss both topological and measure-theoretic notions of genericity that are stronger than topological genericity and/or prevalence, in the sense that their class of negligible sets is a proper subclass of the meagre sets or the shy sets.

Several of these notions are based on the idea of ‘porosity’, which was used by Denjoy, and rediscovered and named by Dolžhenko [44]; see [186,187]. A subset of a metric space is considered to be negligible if it is  $\sigma$ -porous, meaning a countable union of ‘porous’ sets. The definition of porosity varies among different authors; the weakest is as follows. A subset  $S$  of a metric space  $X$  is called *porous* if for each  $x \in S$ , there is a constant  $0 < c < 1$  and a sequence  $\{R_k\}$  of positive numbers tending to zero, such that the ball of radius  $R_k$  centred at  $x$  contains a ball of radius  $cR_k$  that does not intersect  $S$ . This definition is sometimes strengthened to include all  $x \in X$  and/or to require that  $c$  be independent of  $x$  and/or to replace the sequence  $\{R_k\}$  with all  $R > 0$ , or at least all sufficiently small  $R > 0$ . Some of these variations, while affecting the class of porous sets, do not change the class of  $\sigma$ -porous sets. However, requiring the porosity condition for all sufficiently small  $R > 0$  strengthens the notion considerably; such sets are often called ‘very porous’ and countable unions of these sets are then called ‘ $\sigma$ -very porous’.

Porous sets are clearly nowhere dense, and hence  $\sigma$ -porous sets are meagre. It also follows from the Lebesgue density theorem that porous subsets of  $\mathbb{R}^n$  have Lebesgue measure zero, and hence the same is true for  $\sigma$ -porous subsets of  $\mathbb{R}^n$ . Thus,  $\sigma$ -porous subsets of  $\mathbb{R}^n$  are small in both the senses of category and measure. However, Preiss and Tišer [145] proved that every separable infinite-dimensional Banach space has a subset that is both  $\sigma$ -porous and prevalent. Thus, in such spaces, properties that are generic in the sense of porosity need not be prevalent, and vice-versa. See [187, Section 5] for a discussion of the relationship between these and several other notions of genericity on infinite-dimensional spaces.

Kolář [96] introduced an even stronger notion of negligibility of a subset of a Banach space, called ‘HP-small’, that implies both shyness and  $\sigma$ -porosity (and hence meagreness too). A subset  $S$  of a Banach space  $B$  is called *HP-small* if for some  $c > 0$  it can be expressed as a countable union of sets  $S_m$  with the following property. For every  $c' \in (0, c)$  and  $r > 0$  there exists  $K > 0$  and an infinite sequence of balls whose centres all have norm at most  $r$  and whose radii are all  $c'r$ , such that for all  $x \in X$ , the translate  $x + S$  intersects at most  $K$  of the balls. Kolář called a property *HP-typical* if it is true except on an HP-small set, and proved, for example, that an HP-typical continuous function from  $[0, 1]$  to  $\mathbb{R}$  is nowhere Hölder continuous. (See Section 2.3.1 for more prevalent and generic results of this type.)

For Borel subsets of a separable Banach space  $B$ , several measure-theoretic definitions of negligibility that are stronger than shyness have also been used in the literature. Mankiewicz [109] introduced a notion later called ‘cube null’, whereby a Borel subset of  $B$  is negligible if it has measure zero for every ‘cube measure’; a cube measure is a probability measure describing the distribution of a random variable  $v + \sum_{k=1}^{\infty} C_k v_k$ , where  $v, v_1, v_2, \dots \in B$  are fixed vectors for which the span of  $v_1, v_2, \dots$  is dense in  $B$ , and  $C_1, C_2, \dots$  are independent and uniformly distributed in  $[0, 1]$ . Aronszajn [13] defined a Borel subset  $S \subset B$  to be negligible, if for every sequence  $v_1, v_2, \dots \in B$  of a vector whose span is dense,  $S$  can be written as a union of sets  $S_k$  such that the Lebesgue measure on the line spanned by  $x_k$  is transverse to  $S_k$ ; such subsets are called ‘Aronszajn null’. Phelps [144] called a Borel subset of  $B$  ‘Gaussian null’ if it has measure zero for every nondegenerate Gaussian measure on  $B$ ; a probability measure is (nondegenerate) Gaussian if its pushforward to  $\mathbb{R}$  under each element of the dual space of  $B$  is a (nondegenerate) Gaussian measure. These three notions were proved equivalent by Csörnyei [37], and are all stronger than shyness; see [21, Chapter 6] for a discussion.

## References

- [1] R. Abraham, *Lectures of Smale on Differential Topology*, Columbia University, New York (1962).
- [2] R. Abraham, *Bumpy metrics*, Global Analysis, S.S. Chern and S. Smale, eds, Proc. Sympos. Pura Math., Vol. XIV (1970), 1–3.
- [3] J. Alexander and J. Yorke, *Fat Baker's transformations*, Ergodic Theory Dynam. Systems **4** (1) (1984), 1–23.
- [4] A. Alexiewicz and W. Orlicz, *Some remarks on the existence and uniqueness of solutions of the hyperbolic equation  $\partial^2 z / \partial x \partial y = f(x, y, z, \partial z / \partial x, \partial z / \partial y)$* , Stud. Math. **15** (1956), 201–215.
- [5] S. Alpern and V. Prasad, *Properties generic for Lebesgue space automorphisms are generic for measure-preserving manifold homeomorphisms*, Cambridge Tracts in Mathematics, Vol. 139, Cambridge University Press (2000).
- [6] R.M. Anderson and W.R. Zame, *Genericity with infinitely many parameters*, Adv. Theor. Econ. **1** (2001), 1–62.
- [7] D.V. Anosov, *Geodesic flows on Riemann manifolds with negative curvature*, Proc. Steklov Inst. **90** (1967).
- [8] D.V. Anosov, *On generic properties of closed geodesics*, Math. USSR, Izv. **21** (1983), 1–29.
- [9] D.V. Anosov, *Elementary theory*, Chapter 2 of Smooth Dynamical Systems by D.V. Anosov, I.U. Bronshtein, S.Kh. Aranson, V.Z. Grines, Dynamical System, Vol. 1, D.V. Anosov and V.I. Arnol'd, eds, Springer-Verlag, New York (1998).
- [10] A. Araujo and P.K. Monteiro, *Generic nonexistence of equilibria in finance models*, J. Math. Econ. **14** (1991), 61–66.

- [11] V.I. Arnol'd, *Small denominators I: on mappings of the circle onto itself*, Izv. Akad. Nauk Ser. Mat. **25** (1961), 21–86 (Russian); Transl. Amer. Math. Soc. Ser. 2, **46** (1965), 213–284.
- [12] V.I. Arnol'd, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, New York (1983).
- [13] N. Aronszajn, *Differentiability of Lipschitzian mappings between Banach spaces*, Studia Math. **57** (1976), 147–190.
- [14] P. Ashwin and I. Melbourne, *Symmetry groups of attractors*, Arch. Ration. Mech. Anal. **126** (1994), 59–78.
- [15] H. Auerbach and S. Banach, *Über die Höldersche Bedingung*, Studia Math. **3** (1931), 180–184.
- [16] A. Avila and G. Forni, *Weak mixing for interval exchange transformations and translation flows*, Ann. Math. (2) **165** (2) (2007), 637–664, preprint.
- [17] A. Avila and R. Krikorian, *Reducibility or non-uniform hyperbolicity for quasiperiodic Schrödinger cocycles*, Ann. Math. (2) **164** (2006), 911–940.
- [18] A. Avila, M. Lyubich and W. de Melo, *Regular and stochastic dynamics in real analytic families of unimodal maps*, Inv. Mathemat. **154** (2003), 451–550.
- [19] Y. Avron and B. Simon, *Singular continuous spectrum for a class of almost periodic Jacobi matrices*, Bull. AMS **6** (1982), 81–85.
- [20] S. Banach, *Über die Baire'sche Kategorie gewisser Funktionenmengen*, Studia Math. **3** (1931), 174–179.
- [21] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis, Vol. 1*, Amer. Math. Soc. Colloquium Publications, Vol. 48 (2000).
- [22] P. Bernard and G. Contreras, *A generic property of families of Lagrangian systems*, Ann. Math. (2) **167** (3) (2008), 1–10.
- [23] P. Blanchard, *Complex analytic dynamics on the Riemann sphere*, Bull. Amer. Math. Soc. **11** (1984), 85–141.
- [24] J. Bochi, *Genericity of Lyapunov exponents*, Ergodic Theory Dynam. Systems **22** (2002), 1667–1696.
- [25] J. Bochi and M. Viana, *The Lyapunov exponents of generic volume-preserving and symplectic maps*, Ann. Math. **161** (3) (2005), 1423–1485.
- [26] O. Bohigas, M.J. Giannoni and C. Schmit, *Spectral properties of the Laplacian and random matrix theories*, J. Physique Lett. **45** (1984), 1015–1022.
- [27] J.M. Borwein and W.B. Moors, *Null sets and essentially smooth Lipschitz functions*, SIAM J. Optim. **8** (1998), 309–323.
- [28] H.W. Broer and F.M. Tangerman, *From a differentiable to a real analytic perturbation theory, applications to the Kupka–Smale theorem*, Ergodic Theory Dynam. Systems **6** (3) (1986), 345–362.
- [29] G. Buzzard, *Kupka–Smale theorem for automorphisms of  $C^n$* , Duke Math. J. **93** (3) (1998), 487–503.
- [30] G. Buzzard, *Infinitely many periodic attractors for holomorphic maps of 2 variables*, Ann. Math. (2) **145** (2) (1997), 389–417.
- [31] G. Buzzard, S. Lynch Hruska and Y. Ilyashenko, *Kupka–Smale theorem for polynomial automorphisms of  $C^2$  and persistence of heteroclinic intersections*, Invent. Math. **161** (1) (2005), 45–89.
- [32] J.P.R. Christensen, *On sets of Haar measure zero in abelian Polish groups*, Israel J. Math. **13** (1972), 255–260.
- [33] J.P.R. Christensen, *Measure theoretic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings*, Publ. Dept. Math. Lyon t. 10-2 (1973), 29–39.
- [34] J.P.R. Christensen, *Topology and Borel Structure*, North-Holland, Amsterdam (1974).
- [35] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic theory*, Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Vol. 115, Springer-Verlag (1982).
- [36] H. Cremer, *Zum zentrumproblem*, Math. Ann. **98** (1927), 151–163.
- [37] M. Csörnyei, *Aronszajn null sets and Gaussian null sets coincide*, Israel J. Math. **111** (1999), 191–202.
- [38] D. Damanik, R. Killip and D. Lenz, *Uniform spectral properties of one-dimensional quasicrystals. III.  $\alpha$ -continuity*, Comm. Math. Phys. **212** (1) (2000), 191–204.
- [39] G. Debreu, *Economies with a finite set of equilibria*, Econometrica **38** (1970), 387–392.
- [40] M. Dellnitz, M. Golubitsky and I. Melbourne, *Mechanism of symmetry creation*, Bifurcation and Symmetry ISNM 104, E. Allgower, K. Böhmer and M. Golubitsky, eds, Birkhäuser, Basel (1992), 99–109.
- [41] R. del Rio, S. Jitomirskaya, N. Makarov and B. Simon, *Singular continuous spectrum is generic*, Bull. AMS **31** (2) (1994), 208–212.
- [42] D. Dolgopyat, *Prevalence of rapid mixing I*, Ergodic Theory Dynam. Systems **18** (1998), 1097–1114.
- [43] D. Dolgopyat, *Prevalence of rapid mixing II*, Ergodic Theory Dynam. Systems **20** (2000), 1045–1059.

- [44] E.P. Dolženko, *Boundary properties of arbitrary functions*, Izv. Akad. Nauk SSSR Ser. Mat. **31** (1967), 3–14.
- [45] R. Dougherty, *Examples of non-shy sets*, Fund. Math. **144** (1994), 73–88.
- [46] R. Dougherty and J. Mycielski, *The prevalence of permutations with infinite cycles*, Fund. Math. **144** (1994), 89–94.
- [47] P. Dubey, *Inefficiency of Nash equilibria*, Math. Oper. Res. **11** (1986), 1–8.
- [48] N. Dunford and J.T. Schwartz, *Linear Operators, Part I*, Interscience, New York (1958).
- [49] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Rev. Mod. Phys. **57** (1985), 617–656.
- [50] G.A. Enciso, M.W. Hirsch and H.L. Smith, *Prevalent behaviour of strongly order preserving semiflows*, J. Dynam. Differential Equations **20** (1) (2008), 115–132.
- [51] M. Field, I. Melbourne and M. Nicol, *Symmetric attractors for diffeomorphisms and flows*, Proc. London Math. Soc. **72** (1996), 657–696.
- [52] M. Field, I. Melbourne and A. Török, *Stability of mixing and rapid mixing for hyperbolic flows*, Ergodic Theory Dynam. Systems **25** (5) (2005), 517–551.
- [53] J. Fornæss and E. Gavosto, *Existence of generic homoclinic tangencies for Henon mappings*, J. Geom. Anal. **2** (5) (1992), 429–444.
- [54] A. Fraysse, *Résultats de généricité en analyse multifractale*, Ph.D. thesis, Université Paris XII (2005).
- [55] A. Fraysse and S. Jaffard, *How smooth is almost every function in a Sobolev space?* Rev. Mat. Iberoamericana **22** (2) (2006), 663–682.
- [56] A. Fraysse, S. Jaffard and J.-P. Kahane, *Quelques propriétés génériques en analyse*, C.R. Acad. Sci. Paris, Ser. I **340** (2005), 645–651.
- [57] I.V. Girsanov and B.S. Mityagin, *Quasi-invariant measures and linear topological spaces*, Nauchn. Dokl. Vys. Skol. **2** (1959), 5–10 (Russian).
- [58] S. Gonchenko, L. Shilnikov and D. Turaev, *Homoclinic tangency of an arbitrary order in Newhouse regions (Russian)*, Dynamical Systems, Itogi Nauki Tekh. Ser. Sovrem. Mat. Prilozh. Temet. Obz. Vseross. Nauchn. i Tekhn. Inform. (VINITI), 67, Moscow (1999), 69–128.
- [59] A. Gorodetski and V. Kaloshin, *How often surface diffeomorphisms have infinitely many sinks and hyperbolicity of periodic points near a homoclinic tangency*, Adv. Math. **208** (2) (2007), 710–797.
- [60] J. Graczyk and G. Świątek, *Generic hyperbolicity in the logistic family*, Ann. Math. **146** (1997), 1–52.
- [61] J. Guckenheimer, *Sensitive dependence on initial conditions for one-dimensional maps*, Comm. Math. Phys. **70** (1979), 133–160.
- [62] E. Gutkin, *Billiard dynamics: a survey with the emphasis on open problems*, Regul. Chaotic Dynam. **8** (2003), 1–13.
- [63] P. Halmos, *Approximation theories for measure-preserving transformation*, Trans. AMS **55** (1944), 1–18.
- [64] P. Halmos, *In general a measure preserving transformation is mixing*, Ann. Math. (2) **45** (1944), 786–792.
- [65] B. Hasselblatt and A. Wilkinson, *Prevalence of non-Lipschitz Anosov foliations*, Ergodic Theory Dynam. Systems **19** (1999), 643–656.
- [66] M. Hénon, *A two-dimensional mapping with a strange attractor*, Comm. Math. Phys. **50** (1) (1976), 69–77.
- [67] M.R. Herman, *Mesure de Lebesgue et nombre de rotation*, Geometry and Topology, J. Palis and M. do Carmo, eds, Springer Lecture Notes in Mathematics, Vol. 597 (1977), 271–293.
- [68] A.J. Heunis, *On the prevalence of stochastic differential equations with unique strong solutions*, Ann. Prob. **14** (1986), 653–662.
- [69] M.W. Hirsch, *Systems of differential equation which are competitive or cooperative II: convergence almost everywhere*, SIAM J. Math. Anal. **16** (1985), 423–439.
- [70] A. Hof, O. Knill and B. Simon, *Singular continuous spectrum for palindromic Schrödinger operators*, Comm. Math. Phys. **174** (1) (1995), 149–159.
- [71] P.D. Humke and G. Petruska, *The packing dimension of a typical continuous function is 2*, Real Anal. Exch. **14** (1988–1989), 345–358.
- [72] B.R. Hunt, *The prevalence of continuous nowhere differentiable functions*, Proc. Amer. Math. Soc. **122** (1994), 711–717.
- [73] B.R. Hunt and V.Yu. Kaloshin, *Regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces*, Nonlinearity **12** (1999), 1263–1275.
- [74] B.R. Hunt, T. Sauer and J.A. Yorke, *Prevalence: a translation-invariant almost every for infinite dimensional spaces*, Bull. Amer. Math. Soc. **27** (1992), 217–238; *Prevalence: an addendum*, Bull. Amer. Math. Soc. **28** (1993), 306–307.



- [75] Yu. Ilyashenko and Weigu Li, *Nonlocal bifurcations*, Mathematical Surveys and Monographs, Vol. 66, American Mathematical Society, Providence, RI (1999).
- [76] S. Jaffard, *On the Frisch-Parisi conjecture*, J. Math. Pures Appl. **79** (2000), 525–552.
- [77] M.V. Jakobson, *Absolutely continuous invariant measures for one-parameter families of one-dimensional maps*, Comm. Math. Phys. **81** (1981), 39–88.
- [78] S. Jitomirskaya and B. Simon, *Operators with singular continuous spectrum: III. Almost periodic Schrödinger operators*, Comm. Math. Phys. **165** (1994), 201–205.
- [79] J.-P. Kahane, *Propriétés prévalentes, versus génériques, des images continues*, Bull. Sci. Math. **130** (2006), 97–109.
- [80] V.Yu. Kaloshin, *Some prevalent properties of smooth dynamical systems*, Proc. Steklov Inst. Math. **213** (1997), 123–151.
- [81] V.Yu. Kaloshin, *Prevalence in spaces of smooth mappings and a prevalent Kupka–Smale theorem*, Dokl. Akad. Nauk **355** (1997), 306–307.
- [82] V. Kaloshin, *Generic diffeomorphisms with superexponential growth of number of periodic orbits*, Comm. Math. Phys. **211** (1) (2000), 253–271.
- [83] V. Kaloshin and B. Hunt, *A stretched exponential bound on the rate of growth of the number of periodic points for prevalent diffeomorphisms. I*, Electron. Res. Announc. Amer. Math. Soc. **7** (2001), 17–27; 27–36.
- [84] V. Kaloshin and B. Hunt, *Stretched exponential estimates on growth of the number of periodic points for prevalent diffeomorphisms. I*, Ann. Math. (2) **165** (1) (2007), 89–170.
- [85] V. Kaloshin and I. Rodnianski, *Diophantine properties of elements of  $SO(3)$* , Geom. Funct. Anal. **11** (5) (2001), 953–970.
- [86] A.B. Katok, *Interval exchange transformations and some special flows are not mixing*, Israel J. Math. **35** (4) (1980), 301–310.
- [87] A.B. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press (1995).
- [88] A.B. Katok and A.M. Stepin, *Approximations in ergodic theory*, Uspehi Mat. Nauk **22** (5 (137)) (1967), 81–106.
- [89] A.B. Katok and A.M. Stepin, *Metric properties of homeomorphisms that preserve measure*, Uspehi Mat. Nauk **25** (2 (152)) (1970), 193–220 (Russian); English translation in Russian Math. Surveys, **25** (1970), 191–220.
- [90] R. Kaufmann, *A functional method for linear sets*, Israel J. Math. **5** (1967), 185–187.
- [91] R. Kaufmann, *On Hausdorff dimension of projections*, Mathematika **15** (1968), 153–155.
- [92] S. Kerckhoff, H. Masur and J. Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. Math. **124** (1986), 293–311.
- [93] S. Kierst and E. Szpilrajn, *Sur certaines singularités des fonctions analytiques uniformes*, Fund. Math. **21** (1933), 267–291.
- [94] V.L. Klee Jr., *Invariant metrics in groups (solution of a problem of Banach)*, Proc. Amer. Math. Soc. **3** (1952), 484–487.
- [95] W. Klingenberg and F. Takens, *Generic properties of geodesic flows*, Math. Ann. **197** (1972), 323–334.
- [96] J. Kolář, *Porous sets that are Haar null, and nowhere approximately differentiable functions*, Proc. Amer. Math. Soc. **129** (2001), 1403–1408.
- [97] A.N. Kolmogorov, *Proc. Intern. Congr. Math.* (1954), pp. 315–333.
- [98] O. Kozlovski, *Axiom A maps are dense in the space of unimodal maps in the  $C^k$ -topology*, Ann. Math. **157** (2003), 1–43.
- [99] S. Kuksin, *Double-periodic solutions of Nonlinear Cauchy–Riemann equations*, Comm. Pure Appl. Math. **49** (1996), 639–676.
- [100] A. Lasota and J.A. Yorke, *The generic property of existence of solutions of ordinary differential equations in Banach space*, J. Differential Equations **13** (1973), 1–12.
- [101] J.A. Leslie, *A ‘generic’ property of real analytic differential equations on a compact manifold*, Boll. Un. Mat. Ital. B (6) **2** (3) (1983), 669–679.
- [102] M.Yu. Lyubich, *On the typical behaviour of the trajectories of the exponential*, Russ. Math. Surv. **41** (2) (1986), 207–208.
- [103] M.Yu. Lyubich, *Measurable dynamics of the exponential*, Siberian Math. J. **28** (1987), 780–793.
- [104] M.Yu. Lyubich, *Dynamics of quadratic polynomials, I–II*, Acta Math. **178** (1997), 185–297.

- [105] M. Lyubich, *Almost every real quadratic map is either regular or stochastic*, Ann. Math. **156** (2002), 1–78.
- [106] R. Mañé, *On the dimension of the compact invariant sets of certain non-linear maps*, Dynamical Systems and Turbulence, Springer Lect. Notes Math., Vol. 898 (1981), 230–242.
- [107] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (2) (1996), 273–310.
- [108] R. Mañé, *The Lyapunov exponents of area-preserving diffeomorphism*, Int. Conference on Dynamical Systems (Montevideo, 1995), London, Harlow, Putnam Res. Notes Math. Ser., Vol. 362 (1996), 100–119.
- [109] P. Mankiewicz, *On the differentiability of Lipschitz mappings in Fréchet spaces*, Studia Math. **45** (1973), 15–29.
- [110] J.M. Marstrand, *Some fundamental geometrical properties of plane sets of fractional dimension*, Proc. London Math. Soc. **4** (1954), 257–302.
- [111] A. Mas-Colell, *The Theory of General Economic Equilibrium: A Differentiable Approach*, Cambridge University Press (1985).
- [112] J. Mather, *Action minimizing measures for positive Lagrangian systems*, Math. Z. **207** (1993), 169–207.
- [113] E. Matoušková, *The Banach–Saks property and Haar null sets*, Comment. Math. Univ. Carolinae **39** (1998), 71–80.
- [114] E. Matoušková, *Lipschitz images of Haar null sets*, Bull. London Math. Soc. **32** (2000), 235–244.
- [115] P. Mattila, *Hausdorff dimension, orthogonal projections and intersections with planes*, Ann. Acad. Sci. Fennicae **A1** (1975), 227–244.
- [116] H. Mazur, *Interval exchange transformations and measures foliations*, Ann. Math. (2) **115** (1) (1982), 169–200.
- [117] S. Mazurkiewicz, *Sur les fonctions non dérivables*, Studia Math. **3** (1931), 92–94.
- [118] M. McClure, *The prevalent dimension of graphs*, Real Anal. Exch. **23** (1997–1998), 241–246.
- [119] M. Mehta, *Random Matrices and the Statistical Theory of Energy Levels*, 2nd edn, Academic Press, Boston (1990).
- [120] I. Melbourne, M. Delnitz and M. Golubitsky, *The structure of symmetric attractors*, Arch. Rational Mech. Anal. **123** (1993), 75–98.
- [121] I. Melbourne and I. Stewart, *Symmetric  $\omega$ -limit sets for smooth  $\Gamma$ -equivariant dynamical systems with  $\Gamma^0$  Abelian*, Nonlinearity **10** (1997), 1551–1567.
- [122] J. Milnor, *Fubini Foiled: Katok’s paradoxical example in measure theory*, The Mathematical Intelligencer, 1997, 30–32.
- [123] J. Milnor and W. Thurston, *On iterated maps of the interval*, Dynamical Systems: Proc., Univ. Md. 1986–87, J. Alexander, ed., Springer Lect. Notes Math., Vol. 1342 (1988), 465–563 (this paper was circulated in preprint form starting in 1977).
- [124] M. Misiurewicz, *On iterates of  $e^z$* , Ergodic Theory Dynam. Systems **1** (1981), 103–106.
- [125] C.G. Moreira, *Stable intersections of Cantor sets and homoclinic bifurcations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (6) (1996), 741–781.
- [126] C.G. Moreira and J.-C. Yoccoz, *Stable intersection of regular Cantor sets with large Hausdorff dimension*, Ann. Math. **154** (2001), 45–96.
- [127] J. Mycielski, *Some unsolved problems on the prevalence of ergodicity, instability, and algebraic independence*, Ulam Quarterly **1** (3) (1992), 30–37.
- [128] N.N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable hamiltonian systems*, Russ. Math. Surveys **32** (6) (1977), 1–65.
- [129] S.E. Newhouse, *Diffeomorphisms with infinitely many sinks*, Topology **13** (1974), 9–18.
- [130] S.E. Newhouse, *On the abundance of wild hyperbolic sets*, Publ. Math. Inst. Hautes Etud. Sci. **50** (1979), 101–151.
- [131] L. Niederman, *Prevalence of exponential stability among nearly-integrable Hamiltonian systems*, Ergodic Theory Dynam. Systems **27** (2007), 905–928.
- [132] A. Nogueira and D. Rudolph, *Topological mixing of interval exchange transformations*, Ergodic Theory Dynam. Systems **17** (5) (1997), 1183–1209.
- [133] W. Orlicz, *Zur theorie der differentialgleichung  $y' = f(x, y)$* , Bull. Acad. Polon. Sci. Cracovie, Ser. A (1932), 221–228 (German).
- [134] W. Ott, B.R. Hunt and V.Yu. Kaloshin, *The effect of projections on fractal sets and measures in Banach spaces*, Ergodic Theory Dynam. Systems **26** (3) (2006), 869–891.
- [135] W. Ott and J. Yorke, *Prevalence*, Bull. Amer. Math. Soc. **42** (2005), 263–290.

- [136] W. Ott and J. Yorke, *Learning about reality from observation*, SIAM J. Appl. Dynam. Syst. **2** (2003), 297–322.
- [137] J.C. Oxtoby, *Measure and Category*, 2nd edn, Springer-Verlag, New York (1980).
- [138] J.C. Oxtoby and S.M. Ulam, *On the existence of a measure invariant under a transformation*, Ann. Math. **40** (1939), 560–566.
- [139] J.C. Oxtoby and S.M. Ulam, *Measure-preserving homeomorphisms and metrical transitivity*, Ann. Math. **42** (1941), 874–920.
- [140] J. Palis, *A global view of dynamics and a conjecture on the denseness of finitude of attractors*, Géométrie complexe et système dynamiques, Orsay, 1995, Astérisque 261 (2000) 335–347.
- [141] J. Palis, *Homoclinic orbits, hyperbolic dynamics and dimension of Cantor sets*, Contemp. Math. **58** (1987), 203–216.
- [142] J. Palis and W. de Melo, *Geometric Theory of Dynamical Systems. An Introduction*, Springer-Verlag, New-York, Berlin (1982) (Translated from the Portuguese by A.K. Manning).
- [143] J. Palis and F. Takens, *Hyperbolicity and Sensitive-chaotic Dependence at Homoclinic Bifurcations, Fractal Dimensions and Infinitely Many Attractors*, Cambridge University Press (1990).
- [144] R.R. Phelps, *Gaussian null sets and differentiability of Lipschitz maps on Banach spaces*, Pacific J. Math. **77** (1978), 523–531.
- [145] D. Preiss and J. Tišer, *Two unexpected examples concerning differentiability of Lipschitz functions on Banach spaces*, Oper. Theory Adv. Appl. **77** (1995), 219–238.
- [146] E. Pujals and M. Sambarino, *Homoclinic tangencies and hyperbolicity for surface diffeomorphisms*, Ann. Math. (2) **151** (3) (2000), 961–1023.
- [147] F. Quinn and A. Sard, *Hausdorff conullity of critical images of Fredholm maps*, Amer. J. Math. **94** (1972), 1101–1110.
- [148] M. Rees, *The exponential map is not recurrent*, Math. Z. **191** (1986), 593–598.
- [149] F. Riedel, *Generic determinacy of equilibria with local substitution*, J. Math. Econ. **41** (2005), 603–616.
- [150] V.A. Rokhlin, *‘General’ measure-preserving transformation is not mixing*, Dokl. Acad. Nauk SSSR (N.S.) **60** (1948), 349–351.
- [151] D. Ruelle and A. Wilkinson, *Absolutely singular dynamical foliations*, Comm. Math. Phys. **219** (2001), 481–487.
- [152] P. Sarnak, *Some problems in number theory, analysis and mathematical physics*, Mathematics: Frontiers and Perspectives, V. Arnold, M. Atiyah, P. Lax and B. Mazur, eds (2000).
- [153] T. Sauer and J.A. Yorke, *Are the dimensions of a set and its image equal under typical smooth functions?* Ergodic Theory Dynam. Systems **17** (1997), 941–956.
- [154] T. Sauer, J.A. Yorke and M. Casdagli, *Embedology*, J. Stat. Phys. **65** (1991), 579–616.
- [155] H.H. Schaefer, *Topological Vector Spaces*, Macmillan, New York (1966).
- [156] C. Shannon, *A prevalent transversality theorem for Lipschitz functions*, Proc. Amer. Math. Soc. **134** (2006), 2755–2765.
- [157] C. Shannon and W.R. Zame, *Quadratic concavity and determinacy of equilibrium*, Econometrica **70** (2002), 631–662.
- [158] H. Shi, *Prevalence of some known typical properties*, Acta Math. Univ. Comenianae **70** (2001), 185–192.
- [159] H. Shi and B. Thompson, *Haar null sets in the space of automorphisms on  $[0, 1]$* , Real Anal. Exch. **24** (1998–1999), 337–350.
- [160] M. Shub, *Endomorphisms of compact differential manifolds*, Amer. J. Math. **91** (1969), 175–199.
- [161] M. Shub and A. Wilkinson, *Pathological foliations and removable zero exponents*, Inv. Math. **139** (2000), 495–508.
- [162] C.L. Siegel, *Iteration of analytic functions*, Ann. Math. **43** (1942), 607–612.
- [163] K. Sigmund, *Generic properties of invariant measures for Axiom A diffeomorphisms*, Invent. Math. **11** (1970), 99–109.
- [164] K. Sigmund, *On mixing measures for Axiom A diffeomorphisms*, Proc. Amer. Math. Soc. **36** (1972), 497–504.
- [165] B. Simon, *Operators with singular continuous spectrum: I. General operators*, Ann. Math. **141** (1995), 131–145.
- [166] B. Simon, *Operators with singular continuous spectrum. VI. Graph Laplacians and Laplace–Beltrami operators*, Proc. Amer. Math. Soc. **124** (4) (1996), 1177–1182.
- [167] S. Smale, *Differentiable Dynamical Systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.

- [168] H.L. Smith and H.R. Thieme, *Convergence for strongly order preserving semiflows*, SIAM J. Math. Anal. **22** (1991), 1081–1101.
- [169] S. Solecki, *On Haar null sets*, Fund. Math. **149** (1996), 205–210.
- [170] B. Solomyak, *On the random series  $\sum \pm \lambda^n$  (an Erdos problem)*, Ann. Math. **142** (1995), 611–625.
- [171] V.N. Sudakov, *Linear sets with quasi-invariant measure*, Dokl. Akad. Nauk SSSR **127** (1959), 524–525 (Russian).
- [172] V.N. Sudakov, *On quasi-invariant measures in linear spaces*, Vestnik Leningrad Univ. **15** (19) (1960), 5–8 (Russian, English summary).
- [173] G. Świątek, *Rational rotation numbers for maps of the circle*, Comm. Math. Phys. **119** (1988), 109–128.
- [174] F. Takens, *Detecting Strange Attractors in Turbulence*, Lect. Notes Math., Vol. 898, Springer-Verlag (1981), 366–381.
- [175] F. Topsøe and J. Hoffmann-Jørgensen, *Analytic spaces and their application*, Analytic Sets, C.A. Rogers, et al. eds, Academic Press (1980), 317–401.
- [176] M. Tsujii, *A measure on the space of smooth mappings and dynamical system theory*, J. Math. Soc. Japan **44** (3) (1992), 415–425.
- [177] M. Tsujii, *Rotation number and one-parameter families of circle diffeomorphisms*, Ergodic Theory Dynam. Systems **12** (1992), 359–363.
- [178] M. Tsujii, *Fat solenoidal attractors*, Nonlinearity **14** (2001), 1011–1027.
- [179] M. Tsujii, *Physical measures for partially hyperbolic surface endomorphisms*, Acta Math. **194** (2005), 37–132.
- [180] Y. Umemura, *Measures on infinite dimensional vector spaces*, Publ. Res. Inst. Math. Sci. **1** (1965), 1–47.
- [181] S. van Strien, *On the bifurcations creating horseshoes*, Lecture Notes in Mathematics, Vol. 898, D. Rand and L.-S. Young, eds, Springer, Berlin, Heidelberg, New York (1981), 316–351.
- [182] W.A. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. Math. (2) **115** (1) (1982), 201–242.
- [183] W.A. Veech, *The metric theory of interval exchange transformations. I. Generic spectral properties*, Amer. J. Math. **106** (6) (1984), 1331–1356.
- [184] M.I. Vishik and S.B. Kuksin, *Perturbations of quasilinear elliptic equations and Fredholm manifolds*, Math. USSR Sbornik **58** (1987), 223–243.
- [185] H. Whitney, *Differentiable manifolds*, Ann. Math. **37** (1936), 645–680.
- [186] L. Zajíček, *Porosity and  $\sigma$ -porosity*, Real Anal. Exch. **13** (1987–1988), 314–350; Errata, Real Anal. Exch. **14** (1988–1989), 5.
- [187] L. Zajíček, *On  $\sigma$ -porous sets in abstract spaces*, Abstr. Appl. Anal. **2005** (2005), 509–534.
- [188] L. Zajíček, *On differentiability properties of typical continuous functions and Haar null sets*, Proc. Amer. Math. Soc. **134** (2006), 1143–1151.



## CHAPTER 3

# Local Invariant Manifolds and Normal Forms

Floris Takens

*Johann Bernoulli Institute for Mathematics and Computer Science,  
P.O. Box 407 9700 AK Groningen, The Netherlands  
Bernoulliborg, Building 5161, Nijenborgh 9, 9747 AG Groningen, The Netherlands*

André Vanderbauwhede

*Department of Mathematics, Ghent University, Krijgslaan 281, B-9000 Gent, Belgium*

### *Contents*

|                                                                       |     |
|-----------------------------------------------------------------------|-----|
| 1. Introduction . . . . .                                             | 91  |
| 2. Construction of invariant manifolds: the graph transform . . . . . | 91  |
| 2.1. (Locally) invariant manifolds of fixed points . . . . .          | 102 |
| 2.2. Further generalizations . . . . .                                | 104 |
| 3. Invariant foliations . . . . .                                     | 104 |
| 4. Linearizations and partial linearizations . . . . .                | 108 |
| 5. Normal forms . . . . .                                             | 112 |
| 5.1. Normal forms for vector fields . . . . .                         | 117 |
| 5.2. Normal forms in bifurcation theory . . . . .                     | 117 |
| 6. Liapunov-Schmidt reduction . . . . .                               | 120 |
| References . . . . .                                                  | 123 |



## 1. Introduction

In this chapter we review results and constructions which are of importance for the local bifurcation theory of diffeomorphisms and vector fields. In particular we describe the construction of invariant manifolds and foliations and also discuss the construction of normal forms.

We start with invariant manifolds and foliations. It is not our purpose to give complete proofs of all the results; most of these can be found in [12]. It is rather our intention to describe the principles of the various constructions, so that the reader can conclude which results must be true, and to indicate how these constructions can be applied to bifurcation analysis. We refer the reader to the first appendix of [23] where a related but different treatment was given of results concerning invariant manifolds (and foliations) associated with hyperbolic basic sets. Also [1] (in particular the second section on hyperbolic sets and stable manifolds) may give some further insights.

In our review we are particularly interested in centre manifolds, or more generally in normally hyperbolic invariant manifolds, and the reduction of the dynamics to such manifolds. This also involves the construction of invariant foliations and (partial) linearizations.

In the second part we discuss normal forms for vector fields and diffeomorphisms at singularities and fixed points, respectively. These are mainly of importance for the investigation of local bifurcations, but the same methods also provide Taylor expansions for invariant manifolds which are sometimes useful for the numerical approximation of these manifolds.

We treat the cases of both diffeomorphisms and vector fields. Usually we consider only one of these cases in detail and indicate, where necessary, how to adapt the statement of the results and proofs for the other case.

Historically there have been two different approaches to the construction of invariant manifolds: the Hadamard or graph transform method [10] which we will use here, and the more analytical Perron method [24–26] which is used for example in [36]. The first general proofs of the existence of centre manifolds were given by Kelley [14] and by Hirsch, Pugh and Shub [12]. A modern treatment of invariant manifolds, both for vectorfields and for diffeomorphisms, is given in [28]. For a recent general reference to the topics treated in this chapter we refer the reader to [6]; many examples and applications can be found in [15].

## 2. Construction of invariant manifolds: the graph transform

We describe the method of constructing invariant manifolds by graph transforms for diffeomorphisms, and then discuss how to extend it to vector fields and further generalisations. First we consider a very trivial case. Let  $L : E \rightarrow E$  be a linear map on a vector space  $E$ . We assume that  $E$  has two complementary subspaces  $E^+$  and  $E^-$ , which are invariant under  $L$ , and such that the norms of the proper values of  $L|_{E^+}$  are all bigger than those of  $L|_{E^-}$ , and such that  $L|_{E^-}$  has all its proper values of norm smaller than 1. We denote  $L|_{E^\pm}$  by  $L^\pm$ . We can choose a Euclidean structure on  $E$  which makes  $E^+$  and  $E^-$  orthogonal and which is such that, for some  $0 < a < 1$ , we have, for each  $0 \neq v \in E^+$  and  $0 \neq w \in E^-$ :  $\|L(v)\| > a\|v\|$  and  $\|L(w)\| < a\|w\|$ . In this situation we can define



a graph transform  $\Gamma_L$  for maps from  $E^+$  to  $E^-$  as follows. For a map  $h : E^+ \rightarrow E^-$  we define its graph as  $G_h = \{x + h(x) \mid x \in E^+\}$ . Then we define  $G_{\Gamma_L(h)} = L(G_h)$ . It is not hard to see that this is indeed a valid definition of a map  $\Gamma_L(h) : E^+ \rightarrow E^-$ . In fact  $\Gamma_L(h)(v) = L^-(h((L^+)^{-1}(v)))$ . Also it follows from  $a < 1$ , that for every uniformly bounded map  $h$ , repeated application of the graph transform leads to uniform convergence to zero.

The main question that we investigate in this section is: what happens with the definition of the graph transform (and the convergence under repeated application of the graph transform) if we replace  $L$  by a mapping  $L'$  which is only close to  $L$  in the  $C^1$ -norm?

A first observation is that we not only have to choose  $L'$  near  $L$  in the  $C^1$ -norm, but also have to restrict the mappings to which we want to apply the graph transform in order to obtain a well defined result: even if  $L'$  is close to  $L$  in the  $C^1$ -sense, there is no reason for the graph of a discontinuous map  $h : E^+ \rightarrow E^-$  to be mapped by  $L'$  to the graph of some map from  $E^+$  to  $E^-$ . Even restricting this to continuous maps  $h$  is not enough. In order to get a graph transform which is defined for a suitably general class of maps from  $E^+$  to  $E^-$ , we put a restriction on  $L'$  which we describe now.

In each tangent space  $T_x(E)$  we define the (unstable) cone by  $T_x(E) \supset C_x = \{(v+w) \mid v \in E^+, w \in E^-, \|v\| \geq \|w\|\}$ , where we use an identification of  $T_x(E)$  with  $E$ . It is clear that under (the derivative of)  $L$  we have that  $C_x$  is mapped into the interior of  $C_{L(x)}$  (at least outside the origin). Our first restriction on the  $C^1$ -distance between  $L$  and  $L'$  is that, also under the derivative of  $L'$ , for each  $x$ ,  $C_x$  is mapped into the interior of  $C_{L'(x)}$  (at least outside the origin). We put one more restriction: the distance between  $L(x)$  and  $L'(x)$  should be uniformly bounded. With these restrictions, we have the case that the graph transform is defined for each uniformly bounded map  $h : E^+ \rightarrow E^-$  with Lipschitz constant 1. We recall that a map  $h$  has Lipschitz constant 1 if for all  $x$  and  $y$  we have  $\|h(x) - h(y)\| \leq 1 \cdot \|x - y\|$ . It is indeed clear that the image under  $L'$  of the graph of such an  $h$  is locally the graph of a map with Lipschitz constant 1. This follows from the fact that a map  $h : E^+ \rightarrow E^-$  has Lipschitz constant 1 if and only if for each  $x \in E^+$  the graph of  $h$  is (locally) contained in the cone

$$\tilde{C}_{x+h(x)} = \{x + h(x) + v + w \mid v \in E^+, w \in E^-, \|v\| \geq \|w\|\}.$$

Globally,  $E^+ \ni v \mapsto \pi_+ L'(v + h(v))$ , where  $\pi_+$  is the orthogonal projection on  $E^+$ , is a map from  $E^+$  to itself which is proper (because  $h$  and  $L - L'$  are uniformly bounded) and which is a local homeomorphism; this implies that it is a global homeomorphism. Hence  $L'(G_h)$  is the graph of the map (with Lipschitz constant 1) which we denote by  $\Gamma_{L'}(h)$ .

Next we want to prove that repeated application of the graph transform leads to a limit (which is of course, in general, not identically zero). Such a limit is a fixed element for the graph transform and hence has a graph which is an invariant manifold for  $L'$ . For this limit to exist we have to restrict the  $C^1$ -distance between  $L$  and  $L'$  even further.

We observe that, from the fact that  $L \mid E^-$  is a contraction and that  $L - L'$  is uniformly bounded, it follows that there is some constant  $A$  such that the set  $U_A^+ = \{v + w \mid v \in E^+, w \in E^-, \|w\| < A\}$  is mapped by  $L'$  into itself, and such that for any  $v \in E^+$  and  $w \in E^-$  there is an  $i$ , depending on  $\|w\|$ , such that  $(L')^i(v + w) \in U_A^+$ . This means that if  $h$  is

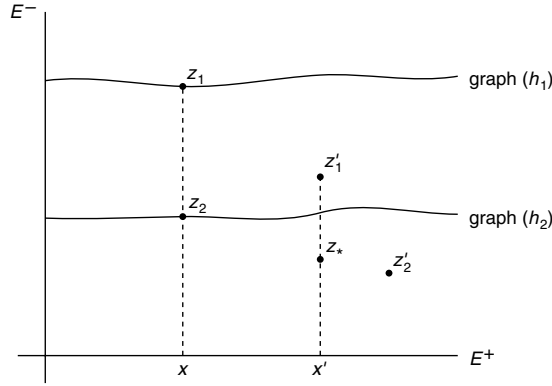


Fig. 1. The action of the graph transform:  $L'$  maps  $z_1, z_2$  to  $z'_1, z'_2$ .

uniformly bounded and with Lipschitz constant 1, repeated application of the graph transform to  $h$  leads eventually to maps with  $C^0$ -norm smaller than  $A$ , i.e. with graph in  $U_A^+$ .

Now we impose a sharper condition on the  $C^1$ -distance between  $L$  and  $L'$  in order to prove that repeated application of  $\Gamma_{L'}$  to a uniformly bounded map with Lipschitz constant 1 leads to a limit. In fact our condition guarantees that  $\Gamma_{L'}$  is a contraction (with respect to the uniform  $C^0$  norm) on uniformly bounded maps from  $E^+$  to  $E^-$  with Lipschitz constant 1. This condition is such that, for each  $x \in E$  and each vector  $w \in E^- \subset T_x(E)$ , the  $E^+$  and  $E^-$  components of  $dL'_x(w)$ , denoted by  $\tilde{w}_+$  and  $\tilde{w}_-$ , satisfy:

$$\begin{aligned}\|\tilde{w}_-\| &\leq \frac{1+2a}{3}\|w\|; \\ \|\tilde{w}_+\| &\leq \frac{1-a}{3}\|w\|.\end{aligned}$$

Indeed this implies that  $\Gamma_{L'}$  is a contraction on uniformly bounded maps with Lipschitz constant 1: let  $h_1$  and  $h_2$  be such maps and let  $x \in E^+$ ; we prove that  $\|\Gamma_{L'}(h_1)(x') - \Gamma_{L'}(h_2)(x')\| \leq \frac{2+a}{3}\|h_1(x) - h_2(x)\|$ , where  $x'$  is the projection of  $L'(x + h_1(x))$  on  $E^+$ . We introduce the following points, see Figure 1.

$z_1$  and  $z_2$  are defined by  $z_i = x + h_i(x)$ ;  $z'_1$  and  $z'_2$  are the images of  $z_1$  and  $z_2$  under  $L'$ ;  $z_* = \Gamma_{L'}(h_2)(x')$ . Then it follows from the last two conditions on the derivative of  $L'$  that  $z'_1 - z'_2$  has  $E^-$  and  $E^+$  components which are, in norm, at most  $\frac{1+2a}{3}$  and  $\frac{1-a}{3}$  times the norm of  $z_1 - z_2$ . Then, using the fact that  $\Gamma_{L'}(h_2)$  has Lipschitz constant 1, it follows that the norm of  $z'_1 - z_* = \Gamma_{L'}(h_1)(x') - \Gamma_{L'}(h_2)(x')$  is at most  $\frac{2+a}{3} < 1$  times the norm of  $z_1 - z_2 = h_1(x) - h_2(x)$ . This proves that  $\Gamma_{L'}$  is indeed a contraction and hence has a unique fixed point. This fixed point is automatically a map with Lipschitz constant 1, whose graph is invariant under the action of  $L'$ .

We summarize the result of the above arguments in the following:

**THEOREM.** *Let  $E$  be a Euclidean vector space with orthogonal splitting  $E = E^+ + E^-$  and let  $L : E \rightarrow E$  be a linear map leaving invariant this splitting and such that for some*

$0 < a < 1$  we have for  $0 \neq v \in E^+$ ,  $0 \neq w \in E^-$ :  $\|L(v)\| > a\|v\|$  and  $\|L(w)\| < a\|w\|$ . Then, for any  $L'$ , sufficiently near  $L$  in the uniform  $C^1$ -norm, there is a unique  $L'$ -invariant manifold which is the graph of a uniformly bounded map from  $E^+$  to  $E^-$  with Lipschitz constant 1.

Moreover, for any bounded map from  $E^+$  to  $E^-$  with Lipschitz constant 1, repeated application of the map  $L'$  to its graph results in convergence to this invariant manifold.  $\square$

### Remarks

The following remarks contain several refinements and generalizations of the above theorem. In these remarks we will often use the notations introduced in the above proof of the theorem.

#### 1. Vector fields

The above result has an immediate analogue for vector fields (or differential equations). We recall that for a vector field  $X$  on a vector space  $E$  (or, more generally, on a manifold) there is a ‘general solution’  $\varphi : E \times \mathbf{R} \rightarrow E$ , in the sense that  $\varphi(x, 0) = x$  and  $\partial_t \varphi(x, t) = X(\varphi(x, t))$  for each  $x \in E$  and  $t \in \mathbf{R}$ . This is the general solution of  $X$ , considered as a differential equation. (We ignore the problem that solutions may not be defined for all time because this does not occur for linear vector fields or for vector fields which are sufficiently close to a linear vector field.) We also write  $X^t$  for the map  $E \ni x \mapsto \varphi(x, t) \in E$ .

Now we consider a linear vector field  $X$  on the vector space  $E$  with orthogonal splitting  $E = E^+ + E^-$ , such that the linear map  $X^t$ , for any  $t > 0$ , satisfies the assumptions which we imposed on  $L$  in the above theorem. Note that the value of  $a$  will depend on  $t$ ; in general this dependence will be exponential in the sense that we can choose  $a = e^{At}$  for some  $A < 0$ . We observe that if  $X'$  is  $C^1$ -close to  $X$ , then also  $X''$  is  $C^1$ -close to  $X^t$ . However for this to be true *uniformly*, we need to impose a further condition, e.g. we may require that the support of  $X' - X$  is compact. We will show later (see remark 5 below) why this restriction of the generality is relatively harmless. So in this case we obtain from the above theorem that for some  $t > 0$  there is an  $X''$ -invariant manifold which is the graph of a function from  $E^+$  to  $E^-$  with Lipschitz constant 1. But more is true: if  $X'$  is sufficiently close to  $X$  in the  $C^1$ -sense then for each  $t > 0$  such an invariant manifold exists and it is independent of  $t$ . For this we impose the following condition on the smallness of the  $C^1$ -difference of  $X$  and  $X'$ : we require that for each  $t > 0$  and each  $x \in E$  the unstable cone  $C_x$  is mapped by the derivative of  $X''$  into the interior of the unstable cone  $C_{X''(x)}$  (at least outside the origin). Then, for each  $t > 0$  the graph transform  $\Gamma_{X''}$  transforms maps from  $E^+$  to  $E^-$  with Lipschitz constant 1 to such maps with Lipschitz constant 1. As we assumed, there is a value  $\bar{t} > 0$  for which the graph transform  $\Gamma_{X''}$  is a contraction on maps from  $E^+$  to  $E^-$  with Lipschitz constant 1, and hence has a unique fixed map which we denote by  $h_0$ . For any  $t$ , the maps  $X^{\bar{t}}$  and  $X''$  commute. So also  $\Gamma_{X''}(h_0)$  is a fixed element for  $\Gamma_{X^{\bar{t}}}$ . From the uniqueness theorem it follows that  $\Gamma_{X''}(h_0) = h_0$ .

#### 2. Differentiability

As stated, the invariant manifold in our theorem is only Lipschitz (in the sense that the invariant manifold is the graph of a map with Lipschitz constant 1). Actually more is true:

it is at least  $C^1$ . A complete proof of this fact, based on the fibre contraction theorem can be found in [12].

Here we show how this can be proved after possibly restricting the  $C^1$ -distance between  $L$  and  $L'$  even further.

Let the invariant manifold in the conclusion of the theorem be denoted by  $W$  and the function of which it is a graph by  $h_0$ . We construct a fibre bundle  $\mathcal{F}$  with base  $W$ : for  $x \in W$ , the fibre  $F_x$  consists of the linear subspaces of  $T_x(E)$  which are graphs of linear maps from  $E^+$  to  $E^-$  (using again the identification of  $T_x(E)$  with  $E$ ) with norm at most equal to 2. If  $L'$  is sufficiently  $C^1$ -close to  $L$ , the derivative of  $L'$  maps each fibre  $F_x$  into the fibre  $F_{L'(x)}$  in a contracting way.

This implies that there is a unique continuous section  $\sigma$  of  $\mathcal{F}$  which is invariant under the derivative of  $L'$ . We will show that for each  $x \in W$ ,  $\sigma(x)$  is the tangent plane of  $W$  at  $x$ . Since these tangent planes represent the derivative of  $h_0$ , it then follows that  $h_0$  is  $C^1$ .

In order to show that  $\sigma(x)$  is the tangent plane of  $W$  at  $x$  we first construct a cone field  $\mathcal{C}_0$  over  $W$ : for each  $x \in W$ ,  $\mathcal{C}_0(x)$  is the cone in  $T_x(E)$  consisting of those vectors  $(v_+, v_-)$ ,  $v_+ \in E^+$  and  $v_- \in E^-$ , such that  $\|v_-\| \leq 2\|v_+\|$ . By induction we define the cone fields  $\mathcal{C}_i$  for  $i = 1, 2, \dots$ :  $\mathcal{C}_i(x)$  is the image under the derivative of  $L'$  of  $\mathcal{C}_{i-1}(L'^{-1}(x))$ . The fact that the derivative of  $L'$  contracts fibres of  $\mathcal{F}$  implies that the cones  $\mathcal{C}_i(x)$ , for  $i \rightarrow \infty$ , converge to the linear subspace  $\sigma_x$ .

For each  $x \in W$  and  $i \geq 0$  we define  $C_i(x) = \{x + v | v \in \mathcal{C}_i(x)\}$ , using again the identification of  $E$  and  $T(E)$ . We will show that for each  $x \in W$  and  $i \geq 0$  there is a neighbourhood  $U_{x,i}$  of  $\pi_+(x)$  in  $E^+$  such that  $\pi_+^{-1}(U_{x,i}) \cap W$  is 'strictly contained in'  $C_i(x)$ . By 'strictly contained in' we mean that there is a cone  $\tilde{C}_i(x)$ , contained in the interior of  $C_i(x)$  (except for the origin) such that  $\pi_+^{-1}(U_{x,i}) \cap W$  is contained in  $\tilde{C}_i(x)$ , where  $\tilde{C}_i(x) = \{x + v | v \in \tilde{C}_i(x)\}$ .

For  $i = 0$  the above statement follows from the fact that  $h_0$  has Lipschitz constant 1. For  $i > 0$  it follows inductively from the definition of  $\mathcal{C}_i$  and the fact that  $L'$  is differentiable.

It now follows easily that  $\sigma_x$  indeed is the tangent plane of  $W$  at  $x$ . This completes the proof that  $h_0$  is  $C^1$ .

### 3. More differentiability and the centre unstable manifold

For our vector space with splitting  $E = E^+ + E^-$  we introduce the Grassmannian bundle  $G(E)$  over  $E$  whose fibre over a point  $x \in E$  consists of the Grassmannian manifold of all linear subspaces of  $T_x(E)$  whose dimensions equal the dimension of  $E^+$ . In this bundle we consider the sub-bundle  $G_0(E)$  which consists of those linear subspaces which project isomorphically onto  $E^+$ . So the elements of  $G_0(E)$  can be represented in a canonical way by pairs  $(x, A)$  with  $x \in E$  and  $A$  a linear map from  $E^+$  to  $E^-$ ; the corresponding linear subspace of  $T_x(E)$  is, up to the identification of  $E$  and  $T_x(E)$ , the graph of  $A$ . This representation of the elements of  $G_0(E)$  shows that  $G_0(E)$  can be given the structure of a vector space. With respect to this structure of a vector space the derivative of  $L$  (in our main theorem) induces a linear map in  $G_0(E)$  which we denote by  $DL$ ; the element represented by  $(x, A)$  is mapped by  $DL$  to the element represented by  $(L(x), L^-A(L^+)^{-1})$ , where  $L^+$  and  $L^-$  are again the restrictions of  $L$  to  $E^+$  and  $E^-$  respectively. The proper values of  $DL$  can be obtained from the proper values of  $L$ :

because of the action on the first component of  $(x, A)$ , each proper value of  $L$  is a proper value of  $DL$ . Then for each pair of proper values  $\mu$  of  $L^+$  and  $\lambda$  of  $L^-$ , the transformation  $A \mapsto L^-A(L^+)^{-1}$  has a proper value  $\mu^{-1}\lambda$ .

Next we consider an even smaller sub-bundle  $G_1(E) \subset G_0(E)$  which consists of the linear subspaces which are represented by pairs  $(x, A)$  with  $\|A\| \leq 1$ . Note that the subspaces in  $G_1(E)$  are just the possible tangent spaces of graphs of smooth functions from  $E^+$  to  $E^-$  with Lipschitz constant 1. By the action of the derivative of  $L$  on  $G_0(E)$ , as described above,  $G_1(E)$  is mapped into its interior. Due to the restrictions imposed on  $L'$ , also the derivative of  $L'$  maps  $G_1(E)$  into its interior. Observe that, in general, the derivative of  $L'$  does *not* define a transformation in  $G_0(E)$ . We denote these maps in  $G_1(E)$ , induced by the derivatives of  $L$  and  $L'$  also by  $DL$  and  $DL'$  respectively.

The idea here is to show that, under certain conditions, our theorem is also applicable to the linear map  $DL$  and its approximation  $DL'$ .

First, of course, we need the map  $L'$  to be  $C^2$  so that  $DL'$  is at least  $C^1$ . Then we need to know the proper values of the derivative of  $DL$ . These were obtained above: we have the proper values of  $L$  together with quotients of a proper value of  $L^-$  and a proper value of  $L^+$ . The relevant splitting in  $G_0(E)$  is given by  $DE^+ = \{(x, A) | x \in L^+, A = 0\}$  and  $DE^- = \{(x, A) | x \in L^-\}$  so that all the 'new' proper values belong to  $DL^- = DL|DE^-$ . In order to formulate the condition on  $DL$ , for our main theorem to be valid for  $DL$  and  $DL'$ , in terms of the proper values of  $L$ , we denote the smallest norm of a proper value of  $L^+$  by  $m^+$  and the biggest norm of a proper value of  $L^-$  by  $m^-$ . The condition imposed on  $L$  in our main theorem can then be formulated as:

$$\begin{aligned} m^- &< m^+ \quad \text{and} \\ m^- &< 1. \end{aligned}$$

Here we have to impose the extra condition

$$m^- < (m^+)^2.$$

Indeed, the biggest norm of a proper value of  $DL^-$  is  $\max\{m^-(m^+)^{-1}, m^-\}$  which is smaller than 1 and which is, by the extra condition, also smaller than  $m^+$ , the smallest norm of a proper value of  $L^+$  and hence of  $DL^+$ .

If this condition is satisfied, and if  $DL$  and  $DL'$  are sufficiently  $C^1$ -close, i.e. if  $L$  and  $L'$  are sufficiently  $C^2$ -close, then we find an invariant manifold for  $DL'$  which is  $C^1$ . (In order to be complete we should add here that, though  $DL'$  is not defined on all of  $G_0(E)$ , it is defined on  $G_1(E)$ , which is mapped by  $DL'$  into itself, so that the graph transform can be iterated on maps from  $DE^+$  to  $DE^- \cap G_1(E)$  with Lipschitz constant 1.) The points of the invariant manifold are represented by pairs  $(x, A)$ , where  $x$  is in the invariant manifold of  $L'$  as constructed before as the graph of the function  $h_0 : E^+ \rightarrow E^-$ ; we claim that  $A$  is the derivative of  $h_0$  in  $\pi_+(x)$ .

Indeed for a suitable and differentiable  $h : E^+ \rightarrow E^-$  we have a corresponding  $Dh : DE^+ \rightarrow DE^-$  which is defined by  $Dh(x, 0) = (h(x), dh(x))$ . Then the graph transforms of  $L'$  and  $DL'$  'agree' in the sense that  $D\Gamma_{L'}(h) = \Gamma_{DL'}(Dh)$ . This means that  $\Gamma_{DL'}^i(Dh) = D(\Gamma_{L'}^i h)$  converges to  $Dh_0$ .

Since  $Dh_0$  is  $C^1$ , by remark 2,  $h_0$  is  $C^2$  in this case.

By repeating this construction we find that if  $L$  and  $L'$  are sufficiently close in the  $C^k$ -topology, and if, in the above notation, also  $m^- < (m^+)^k$ , then the invariant manifold in our theorem is even  $C^k$ . We will consider later (in remark 5) the problem of removing the condition of being  $C^k$ -close.

An important special case is when the constant  $a$  in the theorem can be chosen to be arbitrarily close to 1, i.e. if  $L^+$  has no contracting proper values, or if  $m^+$ , as introduced above, is  $\geq 1$ . Then the resulting invariant manifold is called the *centre unstable manifold*. It can be made  $C^k$ , for any finite  $k$ , if  $L'$  is  $C^k$  and sufficiently close to  $L$  in the  $C^k$ -sense.

#### 4. Unstable and strong unstable manifolds—the condition ‘ $a < 1$ ’ removed

The restriction ‘ $a < 1$ ’ in our main theorem can be removed, provided that we require that the perturbed map still has a fixed point which we may assume, without loss of generality, to be located at the origin. The main modification in the proof consists of the following: instead of considering maps  $h : E^+ \rightarrow E^-$  which are uniformly bounded and have Lipschitz constant 1, we consider maps  $h$  which have Lipschitz constant 1 and which are zero at the origin, so that  $\|h(x)\| \leq \|x\|$  for all  $x \in E^+$ . These maps need not be uniformly bounded and also the limit which we will finally obtain may not be uniformly bounded. Graphs of such maps are contained in  $C = \{v + w | v \in E^+, w \in E^-, \|w\| \leq \|v\|\}$ . Instead of the ordinary  $C^0$ -norm for these maps  $h$  we use the adapted norm  $\|h\|' = \sup_{0 \neq x \in E^+} \frac{\|h(x)\|}{\|x\|} \leq 1$ .

In this situation we have to put different conditions on the smallness of the  $C^1$ -norm of  $L - L'$ . In order to formulate them we ‘split’ the constant  $a$  and choose  $a_1 < a < a_2$  so that for all  $0 \neq v \in E^+$  and  $0 \neq w \in E^-$  we have  $\|L(v)\| > a_2\|v\|$  and  $\|L(w)\| < a_1\|w\|$ . Then the conditions which  $L'$  has to satisfy are:

- i for each  $x \in E$ , the derivative of  $L'$  in  $x$  maps the unstable cone  $C_x$  in the interior (except for the origin) of the unstable cone  $C_{L'(x)}$ ;
- ii for each  $x \in C$ , with  $E^+$ -component  $x_+$ , the  $E^+$ -component  $\tilde{x}_+$  of  $L'(x)$  satisfies  $\|\tilde{x}_+\| \geq a_2\|x_+\|$ ;
- iii for each  $x \in C$  and  $w \in E^- \subset T_x(M)$ , the  $E^+$ - and  $E^-$ -components of  $dL'_x(w)$ , denoted by  $\tilde{w}_+$  and  $\tilde{w}_-$  satisfy:

$$\|\tilde{w}_-\| \leq \frac{2a_1 + a_2}{3} \|w\|$$

and

$$\|\tilde{w}_+\| \leq \frac{a_2 - a_1}{3} \|w\|.$$

With these adaptations the convergence under repeated application of the graph transform remains valid.

If all the proper values of  $L|E^+$  are bigger than 1 and all proper values of  $L|E^-$  are smaller than or equal to 1, the resulting manifold is called the *unstable manifold*. If also  $L|E^-$  has proper values of norm bigger than 1, the resulting manifold is called a *strong unstable manifold*.

These (strong) unstable manifolds are as differentiable as  $L'$ . This is in part based on the localization construction below.

### 5. Localization—locally invariant manifolds—general uniqueness of global (strong) unstable manifolds

Here we relax the restrictions on the map  $L'$  in our theorem. So we assume here to have some  $C^k$ -map  $f : E \rightarrow E$  with  $f(0) = 0$ ; we denote the derivative of  $f$  in 0, as a linear map on  $E$ , by  $L_f$ , the linear part of  $f$ . We assume that there is some  $0 < a < 1$  (due to the preceding remark 4 we may even drop the restriction  $a < 1$ ), and a splitting  $E = E^+ + E^-$  which is  $L_f$ -invariant and such that the proper values of  $L_f$  restricted to  $E^+$  and  $E^-$ , have norms bigger than  $a$  and smaller than  $a$ , respectively. Then we take a Euclidean metric on  $E$  which makes  $E^+$  and  $E^-$  orthogonal and for which we also have that  $\|L_f(v)\| > a\|v\|$  and  $\|L_f(w)\| < a\|w\|$  for  $0 \neq v \in E^+$  and  $0 \neq w \in E^-$ . The only obstruction for our theorem to be valid for  $L_f$  and  $f$  is now that  $f$  and  $L_f$  may not be close enough (in the  $C^1$ - or in the  $C^k$ -sense). For this we introduce the following two procedures.

First we write  $f$  as the sum of  $L_f$  and its nonlinear part  $\tilde{f} = f - L_f$ . Next we take a  $C^\infty$ -function  $\psi : E \rightarrow \mathbf{R}$  which is identically equal to 1 in a neighbourhood of the origin  $0 \in E$  and which has a compact support. We define, for  $\varepsilon > 0$ ,  $\psi_\varepsilon : E \rightarrow \mathbf{R}$  by  $\psi_\varepsilon(x) = \psi(\frac{x}{\varepsilon})$ . Since  $\tilde{f}$  and its first derivative are zero at the origin,  $\tilde{f}_\varepsilon = \psi_\varepsilon \tilde{f}$  converges to zero in the  $C^1$ -sense for  $\varepsilon \rightarrow 0$ . So for  $\varepsilon$  sufficiently small,  $L_f$  and  $f_\varepsilon = L_f + \tilde{f}_\varepsilon$  are sufficiently close in the  $C^1$ -sense to apply our theorem. In this way we find an invariant  $C^1$ -manifold  $W$  for  $f_\varepsilon$ . Since, however,  $f$  and  $f_\varepsilon$  are equal in some neighbourhood of the fixed point 0, such an invariant manifold is still *locally invariant* for  $f$  in the sense that for some neighbourhood  $U$  of 0 we have  $f(W) \cap U = W \cap U$ .

In order to get  $C^k$ -results, we need to make the  $C^k$ -size of the nonlinear part of  $f$  small. This can simply be done by rescaling: if we replace  $f$  by  $f^\lambda$ , defined by  $f^\lambda(x) = \lambda^{-1}f(\lambda x)$ , then, for  $\lambda \rightarrow 0$ , the  $l$ th order derivatives of  $f^\lambda$  go to 0 as  $\lambda^{l-1}$  for  $l \geq 2$ , while the  $C^1$ -size remains the same. In other words, if we apply this rescaling to the linear and the nonlinear parts  $L_f$  and  $\tilde{f}$  of  $f$ , then  $L_f$  does not change while  $\tilde{f}$  goes to zero in the  $C^k$ -sense with  $\lambda$ . Note that  $f$  and  $f^\lambda$  are conjugated as maps on  $E$  (by the scalar multiplication by  $\lambda$  on  $E$ ), so they define the same ‘dynamics’. Hence, by this rescaling we can reduce the higher order derivatives without modifying the ‘dynamics’.

So, in order to apply our theorem (even in the  $C^k$ -version), we only have to modify  $f$  outside some neighbourhood of the fixed point 0 (and do some rescaling). In general, the resulting invariant manifold depends on such a modification, *even arbitrarily close to the fixed point*, i.e. the locally invariant manifold is not even *locally unique*. There are however exceptions: if all the proper values of  $L_f|_{E^+}$  are in norm bigger than 1, i.e. if we are dealing with a (strong) unstable manifold, then the resulting manifold is *locally unique*. The reason for this uniqueness is the following. Suppose  $W_{loc}^1$  and  $W_{loc}^2$  are two locally invariant manifolds such that there is no neighbourhood of the origin  $0 \in E$  in which the two manifolds are equal. By intersecting these local manifolds, if necessary, with a small neighbourhood of the origin we can obtain that  $f(W_{loc}^i) \supset \overline{W_{loc}^i}$  (for this to be true one needs the proper values of  $L_f|_{E^+}$  all to have norm bigger than 1); in this case we say that  $f$  is overflowing on  $W_{loc}^i$ . We may of course assume that both these locally invariant manifolds are graphs of (locally defined) maps from  $E^+$  to  $E^-$  with Lipschitz constant 1.

Next we make these locally invariant manifolds into globally invariant manifolds for one of the maps, say  $f_\varepsilon$ , used to construct one of the locally invariant manifolds, by taking  $W^i = \cup_{j=0}^\infty f_\varepsilon^j(W_{loc}^i)$ . This leads to a contradiction: on the one hand these globally  $f_\varepsilon$ -invariant manifolds are different, but on the other hand we know that the globally invariant manifold for  $f_\varepsilon$  is unique. This is the contradiction which implies that these invariant manifolds are locally unique. We note that a locally invariant (strong) unstable manifold  $W_{loc}$  on which  $f$  is overflowing can be made into a *global (strong) unstable manifold* by taking  $W = \cup_{j=0}^\infty f^j(W_{loc})$ . If we have a map  $f$  which is  $C^\infty$ , then, for each  $k$  we can make global (strong) unstable manifolds which are  $C^k$ . Due to the uniqueness of (strong) unstable manifolds, the manifolds with the various classes of differentiability (and same tangent space in 0) must coincide. Hence they are all  $C^\infty$ .

It is this localization procedure by which we can reduce, also for vector fields, the construction of locally invariant manifolds to the case where the non-linear part of the vector field or the map has compact support.

### 6. The inclination lemma

It is suggested by the previous remarks, and proved in [12], that if we know that the graph transform has a limit which is  $C^k$ , then we have also convergence in the  $C^k$ -sense. To be more precise: if we start with a  $C^k$ -function  $h : E^+ \rightarrow E^-$  and if we know, by the above arguments, that the graph transform  $\Gamma_{L'}$  has a limit which is  $C^k$ , then  $(\Gamma_{L'})^i(h)$  will converge in the  $C^k$ -sense.

This has an important consequence for the case that we started with a linear map  $L$  having no proper values of norm 1. In that case we have for  $L'$ , which is again sufficiently near  $L$  in the  $C^1$ -norm, a stable and an unstable manifold (the stable manifold is the unstable manifold of the inverse  $L'^{-1}$ —see also remark 8 below), which we denote by  $W^s$  and  $W^u$ , respectively. If  $U$  is a (small) manifold which intersects  $W^s$  transversally in one point, then the forward images  $(L')^i(U)$  of  $U$  converge to  $W^u$ , and this convergence is in the  $C^k$  sense whenever  $L'$  is  $C^k$ . This last observation is the content of the inclination lemma or  $\lambda$ -lemma, see [19].

In order to make this statement precise, we have to take iterations of  $U$  under  $L'$  ‘in a restricted sense’, i.e. we claim that there exists a neighbourhood  $V$  of the origin, containing  $U \cap W^s$ , such that the manifolds  $U_i$ , defined by  $U_0 = U \cap V$  and  $U_{i+1} = L'(U_i) \cap V$ , are nonempty and converge in the  $C^k$ -sense to  $V \cap W^u$ .

### 7. The centre stable and the centre manifold

Let  $f : E \rightarrow E$  be a differentiable map with fixed point 0. Assuming that  $df_d(0)$  has proper values of norm 1, we say that  $W^c$  is a *centre manifold* for the fixed point 0 if:

- $0 \in W^c$ ;
- $W^c$  is locally invariant under  $f$  in the sense that for some neighbourhood  $U$  of 0,  $W^c \cap U = f(W^c) \cap U$ ;
- the  $df_0$ -invariant subspace  $T_0(W^c)$  is the maximal invariant subspace of  $T_0(E)$  such that all the proper values of  $df_0|_{T_0(W^c)}$  have norm one.

The construction of such centre manifolds for *invertible*  $f$  is, based on the above observations, straightforward: we first construct a centre-unstable manifold for  $f$ , then a centre unstable manifold for  $f^{-1}$ , which is, by definition, a *centre stable manifold* for  $f$ ,



and then take the intersection of the two. The centre manifold can be made  $C^k$  for finite  $k$  whenever  $f$  is  $C^k$ ; it is in general not locally unique. If  $f$  is  $C^\infty$ , it is not always possible to construct a centre manifold which is  $C^\infty$ , see [40].

In the last two remarks we restricted ourselves to invertible maps (diffeomorphisms). In the next remark we show how to eliminate this restriction.

### 8. Inverse graph transform

Though the term ‘inverse graph transform’ is not commonly used, the idea is very close to the way in which e.g. Moser analysed the dynamics of the horseshoe in [16]. The main idea is to construct a graph transform which can replace the above construction of first taking the inverse of the map and then using the graph transform of that inverse, in particular for cases where we deal with maps  $L$  and  $L'$  which are not invertible.

So we assume that the spaces  $E = E^+ + E^-$  and the maps  $L$  and  $L'$  are as in the beginning of this section, except that we now assume that for some  $a > 1$  and all  $0 \neq v \in E^+$  and  $0 \neq w \in E^-$  we have  $\|L(v)\| > a\|v\|$  and  $\|L(w)\| < a\|w\|$ . (Using the arguments in remark 4, we can even weaken the condition on  $a$  to  $a > 0$ .) We consider maps  $h : E^- \rightarrow E^+$  and their graphs  $G_h$  as defined before. Though the inverse of the map  $L'$  may not be defined, the inverse image of  $G_h$  under  $L'$  is still defined. As for the ordinary graph transform, we should take  $h$  with Lipschitz constant 1, and  $L'$  sufficiently close to  $L$  in the  $C^1$ -sense in order to expect  $(L')^{-1}(G_h)$  to be again the graph of a function with Lipschitz constant 1. In the present case we assume even that  $h$  is  $C^1$  (we show later that this restriction is harmless). The assumption that  $h$  has Lipschitz constant 1 then means that the derivative of  $h$  has norm at most 1 in each point.

We made the assumption that  $h$  is  $C^1$  in order to be able to apply transversality. As in the proof of our main theorem, we assume that  $L - L'$  is uniformly bounded and that for each  $x \in E$  the derivative of  $L'$  maps the unstable cone  $C_x$  into the interior (except for the origin) of  $C_{L'(x)}$ . Then it is easy to verify that the map  $L'$  is transversal with respect to  $G_h$ . This implies that  $(L')^{-1}(G_h)$  is a smooth manifold whose dimension equals the dimension of  $E^-$ . It also follows that for each point  $x \in (L')^{-1}(G_h)$ , the tangent space of  $(L')^{-1}(G_h)$  at  $x$  is ‘transverse’ to the unstable cone  $C_x$ , i.e. the intersection of that tangent space and the unstable cone consists only of  $0 \in T_x(E)$ . Hence the projection of  $(L')^{-1}(G_h)$  on  $E^-$  is locally invertible in every point and  $(L')^{-1}(G_h)$  is locally the graph of a differentiable map from  $E^-$  to  $E^+$ . Due to the transversality with respect to the unstable cones  $C_x$  the derivative of such a local map has norm smaller than 1 everywhere.

Next, because  $L^+$  is expanding and since we assume that  $L - L'$  is uniformly bounded, there is an  $A > 0$  such that the  $A$ -neighbourhood  $U_A^- = \{v + w \mid v \in E^+, w \in E^-, \|v\| < A\}$  of  $E^-$  has the property that  $L'^{-1}(\bar{U}_A^-) \subset U_A^-$ . If we now assume that  $h$  is uniformly bounded by  $A$ , then it is clear that  $(L')^{-1}(G_h)$  is also contained in  $U_A^-$  so that the orthogonal projection of  $(L')^{-1}(G_h)$  on  $E^-$  is a proper map. Since we already know that it is a local diffeomorphism, it now follows that this projection of  $(L')^{-1}(G_h)$  on  $E^-$  is a global diffeomorphism. Hence  $(L')^{-1}(G_h)$  is the graph of a map from  $E^-$  to  $E^+$ , which we denote by  $\Gamma_{L'}^{-1}(h)$ ;  $\Gamma_{L'}^{-1}$  is called the inverse graph transform for  $L'$ .

The proof that, for  $L'$  sufficiently  $C^1$ -close to  $L$ , this inverse graph transform, acting on uniformly bounded  $C^1$  maps from  $E^-$  to  $E^+$  with norm of their derivative at most 1, is a contraction with respect to the uniform  $C^0$ -norm, can be given in the same way as

for the ordinary graph transform. Since the  $C^1$ -maps (whose derivative has everywhere norm at most one) are dense, with respect to the  $C^0$  norm, in the maps with Lipschitz constant 1, the inverse graph transform has a unique continuous extension to the uniformly bounded Lipschitz maps with Lipschitz constant 1. This means that the domain of the graph transform can be extended to all uniformly bounded maps with Lipschitz constant 1 and that it is a contraction. As before the contraction has a unique fixed element, which is a map whose graph is an  $L'$ -invariant manifold.

This means in fact that, even in the case where  $L'$  is not invertible, the inverse graph transform just works as if the inverse of  $L'$  did exist.

### 9. Invariant manifolds for vector fields and differentiable structures

We have seen various methods to construct invariant manifolds for maps and vector fields; the differentiability of these invariant manifolds is often less than the differentiability of the map or vector field with which we started. In the case of vector fields there is an extra problem which arises when we want to restrict the vector field, or the dynamics, of our system to an invariant manifold which is, say,  $C^k$ . Then in principle the restriction of the vector field to the invariant manifold can be at most  $C^{k-1}$ . The reason for this is that the tangent bundle of a  $C^k$ -manifold is only a  $C^{k-1}$  manifold and a vector field on such a manifold, being a cross section in the tangent bundle, can be at most  $C^{k-1}$ .

This extra loss of differentiability can be avoided by a construction which appeared in [22]. The idea of this construction can be best explained for the special case where we have a  $C^k$ -vector field  $X$  on a vector space  $E$  with splitting  $E = E^+ + E^-$ , with an invariant manifold  $N$ , which is the graph of a  $C^k$ -map  $h : E^+ \rightarrow E^-$ . On  $N$  we take the  $C^{k+1}$ -structure for which the  $C^{k+1}$  functions are the compositions  $f\pi_+$ , for  $C^{k+1}$ -functions  $f : E^+ \rightarrow \mathbf{R}$ ; as before  $\pi_+$  is the projection of  $E$  on  $E^+$  with kernel  $E^-$ . Then clearly,  $\pi_+ | N : N \rightarrow E^+$  is a  $C^{k+1}$ -diffeomorphism and  $(\pi_+)_*(X | N)$  is a  $C^k$ -vector field on  $E^+$ , as can be seen from the following composition:

$$E^+ \ni x \mapsto (x, h(x)) \mapsto X(x, h(x)) \mapsto \pi_+(X(x, h(x))) = ((\pi_+)_*(X | N))(x)$$

of three maps which are all at least  $C^k$ .

For the general case of a  $C^k$  submanifold  $N$  of a manifold  $M$  which we assume to be at least  $C^{k+1}$ , we proceed as follows. First we choose any  $C^{k+1}$ -structure on  $N$  which is compatible with the given  $C^k$  structure; for the possibility of doing this whenever  $k \geq 1$  see the section on smoothing of maps and manifolds in [17]. We denote  $N$ , equipped with this  $C^{k+1}$  structure, by  $N_1$ . Next we choose some local  $C^k$ -projection  $\pi$  of a neighbourhood  $U$  of  $N$  in  $M$  on  $N$  (so that  $\pi | N$  is the identity on  $N$ ). Then we take  $\pi_1$  to be a  $C^{k+1}$  (with respect to the  $C^{k+1}$ -structure on  $U$  induced from  $M$  and the  $C^{k+1}$ -structure of  $N_1$ ) approximation of  $\pi$  which is so close to  $\pi$  in the  $C^1$ -sense that for each  $x \in N$  we have that  $\pi_1^{-1}(x)$  is transversal to  $N$  and that  $\pi_1^{-1}(x) \cap N$  consists of only one point. With this we define the map  $\Pi : N \rightarrow N$  by  $\Pi(x) = \pi_1^{-1}(x) \cap N$ . Clearly  $\Pi$  is a  $C^k$ -diffeomorphism on  $N$ . Finally we define the ultimate  $C^{k+1}$ -structure on  $N$ , and we denote  $N$  equipped with that structure by  $N_2$ , so that  $\Pi$  is a  $C^{k+1}$ -diffeomorphism from  $N_1$  to  $N_2$ . With this differentiable structure the restriction to  $N$  of any  $C^k$  vector field  $X$  on  $M$  for which  $N$  is an invariant manifold is still  $C^k$ . This follows essentially by the same arguments as we

used above in the special situation: one obtains  $\Pi_*^{-1}X$  from the following composition of  $C^k$  maps:

$$x \in N_1 \mapsto \Pi(x) \mapsto X(\Pi(x)) \in T(M) \mapsto d\Pi_{\Pi(x)}(X(\Pi(x))) = ((\Pi^{-1})_*X)(x)$$

### 2.1. (Locally) invariant manifolds of fixed points

We consider now a, not necessarily invertible, map  $\varphi$  with a fixed point and analyse what  $\varphi$ -invariant manifolds we can associate with this fixed point by the constructions which we have introduced so far. Since we are primarily interested in locally invariant manifolds, it is no restriction of generality to assume that  $\varphi$  is defined on a vector space  $E$  and that the fixed point is at the origin. We assume  $\varphi$  to be  $C^k$  for some  $1 \leq k \leq \infty$ . For the various manifolds we want to decide whether they are unique and how differentiable they are.

We denote the derivative of  $\varphi$  in 0, as a linear map on  $E$ , by  $L_\varphi$ . We define a splitting  $E = E^s + E^c + E^u$  so that these subspaces are invariant under  $L_\varphi$  and such that the proper values of  $L_\varphi$ , restricted to  $E^s$ ,  $E^c$ , and  $E^u$ , have proper values with norms smaller than 1, equal to 1, and bigger than 1, respectively.

In order to apply the previous constructions, we need  $L_\varphi$  and  $\varphi$  to be close in the  $C^1$  sense, or even in the  $C^l$  sense for some  $1 \leq l \leq k$ . We have seen how to achieve this by modifying  $\varphi$ , leaving it unchanged, up to conjugation by a linear dilatation, in some (small) neighbourhood of 0. The price which we pay for this is that the resulting manifolds are for the original map  $\varphi$  only *locally invariant* in the sense that for some neighbourhood  $U$  of the fixed point 0 the manifold  $V$  satisfies  $V \cap U = \varphi(V) \cap U$ . As we have seen in remark 5, in some cases we can conclude that there is uniqueness and that we can construct ‘global objects’. This is the case for locally (strong) unstable manifolds, and, due to remark 8, for locally (strong) stable manifolds

#### Global (strong) stable and unstable sets

We have seen how to construct locally (strong) stable and unstable manifolds so that they satisfy  $\overline{W_{loc}^u} \subset \varphi(W_{loc}^u)$  for the (strong) unstable manifolds and  $\overline{\varphi(W_{loc}^s)} \subset W_{loc}^s$  for the (strong) stable manifolds. These local manifolds are unique due to the property of  $\varphi$  or ‘ $\varphi^{-1}$ ’ being overflowing on them.

As we observed before, one can define the corresponding non-local (strong) unstable set as  $W^u = \cup_{i \geq 0} \varphi^i(W_{loc}^u)$ . However, if  $\varphi$  is not a diffeomorphism,  $W^u$  need not be a manifold; if  $\varphi$  is a diffeomorphism, then  $W^u$  is at least an injectively immersed submanifold: it need not be an embedded submanifold because it may accumulate on itself. Such accumulation of an invariant manifold on itself is rather frequent in (strongly) nonlinear systems, e.g. see Smale’s horseshoe example in [29].

Also we can define the (strong) stable set as  $W^s = \cup_{i \geq 0} \varphi^{-i}(W_{loc}^s)$ . All the above remarks concerning the (strong) unstable sets also apply here.

For other locally invariant manifolds these constructions do not give meaningful results.

#### Differentiability

As we observed before, the (strong) stable and unstable (local) manifolds are as differentiable as  $\varphi$ . The centre stable and the centre unstable manifold can be made  $C^k$ , for  $k$  finite, if  $\varphi$  is  $C^k$ .

Then there are invariant manifolds, the differentiability of which is restricted by the various proper values of  $L_\varphi$ . These are the manifolds which are constructed by applying the graph transform method to a splitting  $E = E^+ + E^-$  where  $E^+$  contains  $E^c + E^u$  as a proper subspace. Let  $m^-$ ,  $m^+$  be the maximal norm of a proper value of  $L_\varphi|E^-$ , and the minimal norm of a proper value of  $L_\varphi|E^+$ , respectively. So they are a measure of the weakest contraction in  $E^-$ , and the strongest contraction in  $E^+$ , respectively. We say that the resulting submanifold (tangent to  $E^+$ ) is  $r$  normally attracting if  $r$  is the biggest integer such  $(m^+)^r > m^-$ . From our methods of improving the differentiability it follows that such (locally) invariant manifolds can be made of differentiability class  $\min\{r, k\}$ .

By applying similar arguments, using the inverse graph transform, we obtain locally invariant manifolds containing a centre stable manifold.

Finally, intersections of (locally) invariant manifolds are again locally invariant. So putting everything together we obtain the following:

**THEOREM.** *Let  $\varphi$  be a  $C^k$ -map on a vector space  $E$  with a fixed point at the origin and let  $\{\alpha_1 < \alpha_2 < \dots < \alpha_s\}$  be the set of norms of proper values of the derivative  $L_\varphi$  of  $\varphi$  at the origin. Then, for every interval  $[\alpha_i, \alpha_j]$ , with  $i \leq j$  (so the 'point interval' is not excluded), there is a locally invariant manifold  $V_{ij}$  which is at least  $C^1$  and whose tangent space at the origin  $T_0(V_{ij})$  is the maximal  $L_\varphi$ -invariant subspace such that the proper values of  $L_\varphi$  restricted to this subspace have norms  $\alpha_i, \dots, \alpha_j$ . We say that this manifold is  $r^s$  normally contracting if  $i > 1$ ,  $\alpha_{i-1} < 1$ , and if  $r^s$  is the biggest integer such that  $\alpha_i^{r^s} > \alpha_{i-1}$  (if  $\alpha_i \geq 1$  then the manifold is infinitely normally attracting). We say that this manifold is  $r^u$  normally expanding if  $j < s$ ,  $\alpha_{j+1} > 1$  and if  $r^u$  is the biggest integer such that  $\alpha_j^{r^u} < \alpha_{j+1}$  (if  $\alpha_j \leq 1$  then the manifold is infinitely normally attracting).*

*If*

–  $\alpha_i > 1$  and  $j = s$

*or*

–  $\alpha_j < 1$  and  $i = 1$

*then  $V_{ij}$  is an invariant manifold which is  $C^k$  and which is unique; these are the only cases in which the invariant manifold is unique and as differentiable as  $\varphi$ ;*

*if*

–  $\alpha_i \geq 1$  and  $j = s$

*or*

–  $\alpha_j \leq 1$  and  $i = 1$

*or*

–  $i = j$  and  $\alpha_i = \alpha_j = 1$

*then  $V_{ij}$  can be made  $C^l$  for any finite  $l \leq k$ —in general such a manifold is not necessarily unique; in all other cases the maximal differentiability of the invariant manifold which can be concluded from the linear part  $L_\varphi$  of  $\varphi$  is bounded by a finite number, which depends on the proper values of  $L_\varphi$ ;*

*if  $i > 1$  and  $\alpha_j \leq 1$  then  $V_{ij}$  can be made so that its class of differentiability is  $\min\{k, r^s\}$ ;*

if  $j < s$  and  $\alpha_i \geq 1$  then  $V_{ij}$  can be made so that its class of differentiability is  $\min\{k, r^u\}$ ;

if  $i > 1$ ,  $j < s$ , and  $\alpha_i < 1 < \alpha_j$  then  $V_{ij}$  can be made so that its class of differentiability is  $\min\{k, r^s, r^u\}$ .  $\square$

## 2.2. Further generalizations

In this subsection we briefly mention a number of other instances where invariant manifolds can be constructed with the method of graph transforms. The simplest generalization is to singularities of vector fields. Here we get exactly the analogue of the above theorem. In order to formulate the correct conditions on the proper values of the derivative of the vector field at the singularity, we have to exponentiate these proper values and then subject them to the conditions in the case of diffeomorphisms.

The next case which we consider is that of periodic orbits of vector fields. The dynamics near such a periodic orbit is described by the Poincaré map, i.e. by a return map in a local codimension with one section transversal to the periodic orbit. It is a local diffeomorphism so the last theorem is valid for the Poincaré map. Invariant manifolds for the flow are obtained by saturating the invariant manifolds of the Poincaré map with integral curves of the vector field.

A less straightforward generalization is concerned with the so-called normally hyperbolic invariant manifold. For this we refer to [27].

Finally there are the (families of) invariant manifolds associated with hyperbolic basic sets. For a review on these matters we refer to [23].

## 3. Invariant foliations

Let  $M$  be a  $m$ -dimensional manifold with open subset  $U \subset M$ . A  $k$ -dimensional foliation of  $U$  consists of a partition of  $U$  into equivalence classes, called *leaves*, which are injectively immersed connected  $k$ -dimensional sub-manifolds of  $U$ , so that near each point  $x \in U$  there is a local coordinate system  $H : V \rightarrow \mathbf{R}^m$ , where  $V$  is a neighbourhood of  $x$  in  $U$ , such that each connected component of the intersection of  $V$  with a leaf of the foliation has, in these coordinates, the form  $\{(x_1, \dots, x_m) | x_{k+1} = c_{k+1}, \dots, x_m = c_m\}$  for constants  $c_{k+1}, \dots, c_m$ .

The equivalence class of  $x$  is denoted by  $\mathcal{F}_x$ . If  $U' \subset U$  is an open subset, then the restriction  $\mathcal{F}_{U'}$  of  $\mathcal{F}$  to  $U'$  is the foliation, the leaves of which are the connected components of the intersections of leaves of  $\mathcal{F}$  with  $U'$ . For a diffeomorphism  $\varphi : M \rightarrow M$ ,  $\varphi_*\mathcal{F}$  is the foliation on  $\varphi(U)$  whose leaves are the images under  $\varphi$  of the leaves of  $\mathcal{F}$ . We say that the foliation  $\mathcal{F}$  on  $U$  is invariant under a diffeomorphism  $\varphi : M \rightarrow M$  if the restrictions of the foliations  $\mathcal{F}$ ,  $\varphi_*\mathcal{F}$ , and  $\varphi_*^{-1}\mathcal{F}$  to  $U \cap \varphi(U) \cap \varphi^{-1}(U)$  are equal. Finally we say that a foliation  $\mathcal{F}$  on  $U$  is invariant under (the flow of) a vector field  $X$  if it is invariant under  $X^t$  for all sufficiently small  $|t|$ .

We say that a foliation  $\mathcal{F}$  is  $C^k$  if the map  $F$ , which assigns to each  $x \in U$  the tangent plane  $F_x$  of  $\mathcal{F}_x$  at  $x$ , is  $C^k$ . In that case the coordinates  $H$ , as above, can be made  $C^k$ ; this is

implied by Frobenius' theorem, see [18]. Apart from these smooth foliations we shall also encounter foliations which are only continuous but whose leaves are smooth manifolds.

### Integrability

We recall that for a given smooth map  $F$ , which assigns to each  $x \in M$  a  $k$ -dimensional subspace  $F(x) \subset T_x(M)$ , it is not necessarily so that there exists a corresponding foliation (whose leaf through  $x$  has tangent space  $F(x)$ ). For this  $F$  has to satisfy a so-called integrability condition. This integrability condition can be formulated in various ways; an intrinsic form is: for any two  $C^1$ -vector fields  $X, Y$  on  $M$  which are in  $F$ , in the sense that for each  $x \in M$ ,  $X(x), Y(x) \in F(x)$ , the Lie product  $[X, Y]$  is also in  $F$ . This is essentially the content of Frobenius' theorem which can be found in many texts on differential geometry, e.g. see [18]. For us an important aspect of this integrability condition is that it is closed in the  $C^1$ -topology, in the sense that if a sequence  $\{F_i\}$  of  $k$ -plane fields converges in the  $C^1$ -sense to  $F$  then, and if each of the  $F_i$  satisfies the integrability condition, this also holds for  $F$ .

### Example 1: The unstable foliation in a centre unstable manifold

For a diffeomorphism  $\varphi$  with fixed point  $p$  we have seen how to construct a locally invariant centre unstable manifold, which we denote by  $W^{cu}$ . Here we want to construct a foliation  $\mathcal{F}$  of a neighbourhood of  $p$  in  $W^{cu}$  which is  $\varphi$ -invariant and which has the unstable manifold  $W^u$ , at least restricted to a neighbourhood of  $p$  in  $W^{cs}$ , as a leaf. Such a foliation is called an *unstable foliation of a centre unstable manifold*. Without loss of generality we may start with  $\varphi|W^{cu}$ , i.e. we may assume that  $d\varphi(0)$  has no proper values of norm smaller than 1. Since all constructions are local, we also may assume that we work in a vector space and have the fixed point at the origin.

This means that it is enough to consider, changing to the notation of the subsection on graph transforms, a linear map  $L$  in a vector space  $E$  with a  $C^1$ -near perturbation  $L'$  such that:

- $E$  has a splitting  $E = E^c + E^u$  which is invariant under  $L$  and such that the proper values of  $L|E^c$  have norm 1 and the proper values of  $L|E^u$  have norm bigger than 1;
- $L'$  has a fixed point in  $0 \in E$  and its derivative in 0 equals  $L$ .

It is clear that the foliation of  $E$  by affine subspaces, all parallel to  $E^u$ , is an invariant foliation for  $L$ . We denote this foliation by  $\mathcal{F}_0$ . We shall prove that if  $L'$  is sufficiently close to  $L$  in the  $C^2$ -sense, the foliations  $\mathcal{F}_i = (L')^i_*(\mathcal{F}_0)$  have a limit which is differentiable. We note that this restriction of  $C^2$ -closeness is not serious if we only want a *locally* invariant foliation (see remark 5 in the previous section); we need however  $L'$  to be at least  $C^2$ .

We introduce a Euclidean metric in  $E$  so that  $E^c$  and  $E^u$  are perpendicular and so that for two constants  $1 < a < b$ , we have that, for all non-zero vectors  $v \in E^c$  and  $w \in E^u$ ,  $\|L(v)\| < a\|v\|$  and  $\|L(w)\| > b\|w\|$ .

We use again the Grassmannian bundle  $G_1(E)$  whose fibre over  $x \in E$  consists of the linear subspaces of  $T_x(E)$ , whose dimensions equal the dimension of  $E^u$ , and which are graphs of linear functions from  $E^u$  to  $E^c$  of norm at most 1 (using again the identification of  $T_x(E)$  with  $E$ ). The elements of  $G_1(E)$  are represented again by pairs  $(x, A)$  with  $x \in E$  and  $A : E^u \rightarrow E^c$  with  $\|A\| \leq 1$ . Then the element represented by  $(0, 0)$  is a fixed point

of the (linear) map  $DL$  induced by the derivative of  $L$  in  $G_1(E)$  and  $DL$  maps  $G_1(E)$  into its interior. The proper values of this linear mapping are:

- the proper values of  $L$ ;
- for each proper value  $\alpha$  of  $L|E^c$  and proper value  $\beta$  of  $L|E^u$ , the proper value  $\alpha/\beta$  — from the assumptions it follows that this latter collection of proper values consists of contracting proper values only.

This means that the field of tangent spaces of  $\mathcal{F}_0$ , interpreted as a subset of  $G_1(E)$  is a centre-unstable manifold of  $(0, 0)$  as a fixed point of  $DL$ . If we now recall the construction of the centre-unstable manifold, the inclination lemma, and the conditions imposed on  $L$  and  $L'$ , we see that the fields of tangent spaces of  $\mathcal{F}_i = L_*^i \mathcal{F}_0$  converge in the  $C^1$ -sense. So the limit is again an integrable field of subspaces, which is now invariant under the derivative of  $L'$ . Hence this limit defines the  $L'$ -invariant foliation which we claimed to exist.  $\square$

*Example 2: the unstable foliation near a hyperbolic fixed point*

Here we consider an invertible linear map  $L : E \rightarrow E$ , with a perturbation  $L'$  which is  $C^1$ -close to  $L$ , for which there is a splitting  $E = E^s + E^u$ , such that both subspaces are invariant under  $L$ , and such that  $L|E^u$  has only proper values of norm greater than 1 and  $L|E^s$  has only proper values of norm smaller than 1 (hyperbolic for fixed points of diffeomorphisms means: ‘no proper values with norm 1’). We assume again that we have a Euclidean structure on  $E$  such that the subspaces  $E^u$  and  $E^s$  are orthogonal and such that  $L|E^s$ , respectively  $L|E^u$ , is a contraction, respectively an expansion. For this linear map an unstable foliation can be defined as the foliation whose leaves are affine subspaces of  $E$  which are parallel to  $E^u$ . We want to show that for  $L'$  sufficiently close to  $L$  in the  $C^1$ -sense there is also a (perturbed) unstable foliation, i.e. a locally invariant foliation  $\mathcal{F}^u$ , defined on a neighbourhood of the fixed point 0 and with the unstable manifold  $W^u$ , restricted to this neighbourhood of 0, as one of its leaves.

In general we cannot proceed as in the above case (example 1) because in the Grassmannian bundle  $G_1(E)$ , such a foliation should define an invariant manifold, which however may not be one of those constructed by our graph transform methods and which may even not exist as a  $C^1$ -manifold. The reason is that the proper values of the (linear) map in  $G_1(E)$  induced by  $DL$  have, apart from the proper values of  $L$ , new proper values of the form  $\alpha/\beta$  where  $\alpha$  is a proper value of  $L|E^s$  and  $\beta$  is a proper value of  $L|E^u$ . All these new proper values are contracting, but it need not be the case that they are all, in norm, smaller than those of  $L|E^s$ , and this is what would be needed for the application of the graph transform limit procedure. One can even construct examples of hyperbolic fixed points such that there is no unstable foliation which is  $C^1$ . So we have to proceed differently.

In order to construct an unstable foliation, we will use a *fundamental domain* in the stable manifold  $W^s$  for  $L'$ . We say that a subset  $D \subset W^s$ , with boundary components  $\partial_+ D$  and  $\partial_- D$  is a fundamental domain in  $W^s$  if:

- $L'(\partial_+ D) = \partial_- D$ ;
- each orbit  $\{(L')^i(x)\}_{i \in \mathbb{Z}}$ , with  $x \in W^s$  and different from the fixed point, has either one point in  $D$  or two points in  $D$  one of which is in  $\partial_+ D$  and one is in  $\partial_- D$ .

The construction of such a fundamental domain can be done as follows. Assuming  $L'$  to be sufficiently close to  $L$ , the set  $U_1^u = \{v + w | v \in E^s, w \in E^u, \|v\| \leq 1\}$  is mapped, by  $L'$  into its interior. Then we can take as the fundamental domain  $D = U_1^u \cap W^s \setminus (\text{interior}(L'(U_1^u)) \cap W^s)$ ; the boundary components are given by  $\partial_+ D = \partial U_1^s \cap W^s$  and  $\partial_- D = L'(\partial_+ D)$ . Next we choose a foliation  $\mathcal{F}_+$  in  $\tilde{\partial}_+ D = \{v + w | v \in E^s, w \in E^u, \|v\| = 1, \|w\| < 1\}$  (which contains  $\partial_+ D$  since  $L$  and  $L'$  are sufficiently close) such that each leaf intersects  $\partial_+ D$  transversally in one point. We can take, for example, all leaves to be parallel translations of  $E^u$ , of course restricted to  $\tilde{\partial}_+ D$ . By applying  $L'$  we obtain the foliation  $\mathcal{F}_-$  of  $L'(\tilde{\partial}_+ D) = \tilde{\partial}_- D$ . The next step is to ‘interpolate’ these two foliations to obtain a foliation  $\mathcal{F}_D$ , which is defined in the region between  $\tilde{\partial}_+ D$  and  $\tilde{\partial}_- D$ , such that each of its leaves intersects  $D$  transversally in exactly one point. If  $L'$  and  $L$  are sufficiently close such an interpolation is not hard to construct; for the details see [19] or [20]. We denote the domain of definition of  $\mathcal{F}_D$  by  $\tilde{D}$ .

Further extension of the foliation  $\mathcal{F}_D$  is fixed by the fact that we want to obtain a (local) foliation which is  $L'$ -invariant. The first step is to apply  $L'$  to it and so obtain a foliation on  $L'(\tilde{D})$ . Since the two foliations agree by construction on  $\tilde{\partial}_- D$ , they, together, can be considered as one foliation on  $\tilde{D} \cup L'(\tilde{D})$ . This extended foliation is only continuous (not differentiable) at the points of  $\tilde{\partial}_- D$ . If one takes somewhat more care in the construction of the extension to  $\tilde{D}$  the extension to  $\tilde{D} \cup L'(\tilde{D})$  will even be differentiable, also at the points of  $\tilde{\partial}_- D$ . After this we can continue the extensions to  $L'^2(\tilde{D})$ ,  $L'^3(\tilde{D})$  etcetera. In this way we obtain a foliation which is defined on the complement of  $W^u$ , the unstable manifold of  $L'$ , in a neighbourhood of the fixed point. From the inclination lemma however, we know that the leaves which are close to  $W^u$  are even close to  $W^u$  as differentiable submanifolds. So the foliation is defined on a complete neighbourhood if we add  $W^u$  as a leaf. This is the unstable foliation which we denote by  $\mathcal{F}^u$ . It is the last extension, adding  $W^u$  as a leaf, which generally makes the foliation only continuous. Still much differentiability is left: if  $L'$  is  $C^k$  then the leaves of  $\mathcal{F}$ , as we constructed them, are  $C^k$ ; the foliation can be made so that it is  $C^{k-1}$  outside  $W^u$ . Moreover it follows from the inclination lemma that the map, which assigns to each point  $x \in E$  in the domain of  $\mathcal{F}^u$  the leaf  $\mathcal{F}_x^u$  through  $x$ , is a continuous map to the space of  $C^k$ -submanifolds.

It is important to note that this foliation is far from unique: many choices were made, namely the foliation  $\mathcal{F}_+$  and the extension  $\mathcal{F}_D$ . In some constructions this freedom in the choice of the unstable foliation is important.  $\square$

The construction of the corresponding *stable foliations* proceeds in the usual way: first take the inverse of the map, and then construct the unstable foliation of this inverse.

#### *Foliations associated with a general fixed point*

We now consider an invertible linear map  $L : E \rightarrow E$  with an invariant splitting  $E = E^s + E^c + E^u$  such that the proper values of  $L$  restricted to  $E^s$ ,  $E^c$ ,  $E^u$  are in norm smaller than, equal to, or bigger than 1, respectively. Let  $L'$  be a map which is sufficiently close to  $L$  in the  $C^1$  sense and which has a fixed point 0. Then we have seen how to construct the invariant manifolds  $W^s$ ,  $W^{cs}$ ,  $W^c$ ,  $W^{cu}$ , and  $W^u$  which are tangent to  $E^s$ ,  $E^c + E^s$ ,  $E^c$ ,  $E^s + E^u$ ,  $E^u$ , respectively. We have also seen how to construct *inside*  $W^{cs}$  and  $W^{cu}$  the stable and the unstable foliations  $\mathcal{F}^s$  and  $\mathcal{F}^u$ , which are differentiable. It turns out that these last two foliations can be extended to  $L'$ -invariant foliations which are defined



in a full neighbourhood of the fixed point. These extended foliations cannot in general be made differentiable, but can be so constructed that they are differentiable outside  $W^{cs}$ , or, respectively,  $W^{cu}$ . The construction is based on the construction in the above example 2 and the inclination lemma. For the details we refer to [12] and [27].

### *Invariant foliations for vector fields*

The construction of foliations which are invariant under the flow of a vector field follows along the same lines as in the case of diffeomorphisms. The main difference is the notion of a fundamental domain. In the case of vector fields, a fundamental domain is not a ‘ring’ with two boundary components of codimension one, but a codimension one submanifold. In order to make this more explicit, we consider again the construction of the unstable foliation, but now for a hyperbolic singularity of a vector field. We have discussed already the stable manifold in that case. We denote the stable manifold again by  $W^s$ . A *fundamental domain* in this case is a codimension one submanifold  $D$  of  $W^s$  such that each integral curve in  $W^s$ , different from the singularity, has exactly one point in  $D$  where it intersects transversally. Assuming again that we are working in a vector space  $E = E^+ + E^-$  with splitting and Euclidean metric as before, so that  $E^-$  is the tangent space of  $W^s$  in the origin, one can take as a fundamental domain set  $S_\varepsilon = \{v + w | v \in E^-, \|v\| = \varepsilon, w \in E^u, v + w \in W^s\}$ , at least for  $\varepsilon$  sufficiently small. We assume that  $S_\varepsilon$  is contained in  $\tilde{S}_\varepsilon = \{v + w | v \in E^s, w \in E^u, \|v\| = \varepsilon, \|w\| \leq 1\}$  and that our vector field  $X$  is everywhere transversal to  $\tilde{S}_\varepsilon$ . In order to construct an (invariant) unstable foliation we first foliate  $\tilde{S}_\varepsilon$  with leaves parallel to  $E^u$ ; we denote this foliation by  $\mathcal{F}_S$  (it is the analogue of the foliation  $\mathcal{F}_D$  in the case of diffeomorphisms). In order to extend the foliation we proceed as follows: for each leaf  $\mathcal{F}_x$  of  $\mathcal{F}_S$  and each  $t \in \mathbf{R}$ ,  $X^t(\mathcal{F}_x)$  is a leaf of the extended foliation—finally, as in the case of diffeomorphisms, one has to add  $W^u$  as the last leaf.

## **4. Linearizations and partial linearizations**

In the present section we discuss applications of the constructions in the previous sections which are of importance for the local analysis of dynamical systems near fixed points. First we consider *hyperbolic fixed points*, i.e. fixed points such that the derivative of the map at that fixed point has no proper values of norm 1 or 0; after that we also consider more general fixed points. The theorem of Grobman and Hartman, see [9] and [11] deals with this hyperbolic case. We introduce first some notation. Let  $\varphi_i : X_i \rightarrow X_i$ ,  $i = 1, 2$  be two (continuous) maps on topological spaces  $X_1$  and  $X_2$ . We say that these maps are *conjugated* if there is a homeomorphism  $h : X_1 \rightarrow X_2$  such that  $h\varphi_1 = \varphi_2h$ . Then we call  $h$  a *conjugacy*.

If  $x_1$  and  $x_2$  are fixed points of  $\varphi_1$  and  $\varphi_2$ , then these maps are *locally conjugated* near these fixed points if there is a homeomorphism  $h$  from a neighbourhood of  $x_1$  to a neighbourhood of  $x_2$  such that  $h(x_1) = x_2$  and such that  $h\varphi_1 = \varphi_2h$  wherever defined.

**THEOREM** (Grobman [9] and Hartman [11]). *Let  $\varphi : E \rightarrow E$  be a diffeomorphism with  $\varphi(0) = 0$  and such that the derivative  $d\varphi(0)$ , which we also denote by  $L_\varphi$ , has no proper value of norm 0 or 1. Then  $\varphi$  is locally conjugated with  $L_\varphi$ .*

SKETCH OF THE PROOF. Let  $\varphi$  be as in the theorem. Since we only want to conclude the existence of a *local* conjugacy, we may assume that the support of  $\varphi - L_\varphi$  is compact and that the  $C^1$ -distance between  $\varphi$  and  $L_\varphi$  is sufficiently small. First we consider the simpler case where all the proper values of  $d\varphi(0) = L_\varphi$  have norm smaller than 1. Then we may assume that we have a Euclidean metric on  $E$  such that  $L_\varphi(0)$  is a contraction in the sense that for each  $0 \neq v \in E$  we have  $\|L_\varphi(v)\| < \|v\|$ . Then  $L_\varphi$  maps each  $\varepsilon$ -disc  $D_\varepsilon$  into its interior, so that  $D_\varepsilon \setminus L_\varphi(D_\varepsilon)$  is a fundamental domain for  $L_\varphi$ . For  $\varepsilon$  sufficiently small the same will then hold for the map  $\varphi$  itself, i.e.  $D_\varepsilon \setminus \varphi(D_\varepsilon)$  is a fundamental domain for  $\varphi$ . In order to construct a local conjugacy  $h$  (satisfying  $hL_\varphi = \varphi h$ ) we start with a fundamental domain (like with the construction of the (un)stable foliation near a hyperbolic fixed point). Here one may take  $h|_{\partial D_\varepsilon}$  to be the identity. The fact that we are constructing a conjugacy fixes  $h$  on  $L_\varphi(\partial D_\varepsilon)$  as  $\varphi h(L_\varphi)^{-1}$ . The interpolation of  $h$  between these two boundaries of the fundamental domain is free, except that it has to be a homeomorphism; for the details of the construction see [19] and [20]. After we have defined  $h$  in such a fundamental domain, we can extend the domain of definition further inside  $D_\varepsilon$  by using the conjugacy equation as follows. For  $0 \neq x \in D_\varepsilon$  there is some non-negative  $i$  such that  $(L_\varphi)^{-i}(x)$  is in the fundamental domain. Then  $h(x)$  should be equal to  $\varphi^i(h((L_\varphi)^{-i}(x)))$ . Finally  $h(0) = 0$ . In this way we obtain the conjugacy as a homeomorphism on  $D_\varepsilon$ .

The above arguments, applied to  $\varphi^{-1}$ , prove a corresponding statement for expanding maps.

Next we consider the case where we have a diffeomorphism  $\varphi : E \rightarrow E$  for which 0 is a fixed point at which the derivative has no eigenvalues of norm 0 or 1. We denote the derivative of  $\varphi$  at 0 again by  $L_\varphi$ . Using the above results, we see that  $\varphi$  and  $L_\varphi$  are locally conjugated when restricted to their stable manifolds and also when restricted to their unstable manifolds. We denote these conjugacies by  $h_s$  and  $h_u$ .

In order to extend the domain of these partially defined local conjugacies we use the stable and unstable foliations as constructed in the previous section. Any point  $x$  can be written as  $x = x_u + x_s$  with  $x_u$ , respectively  $x_s$  in the unstable, respectively stable, manifold of  $L_\varphi$ . Then, for  $x$  sufficiently near 0,  $h(x)$  can be defined as the intersection of the leaf of the stable foliation through  $h_u(x_u)$  and the leaf of the unstable foliation through  $h_s(x_s)$ .

This completes our sketch of the proof of the theorem of Grobman and Hartman.  $\square$

We observe, and this is not hard to prove, that on a given vector space any two invertible linear contractions are conjugated if and only if they are both orientation preserving or both orientation reversing. So up to local conjugacy a diffeomorphism near a hyperbolic fixed point is completely determined by the dimensions of stable and unstable manifolds, together with the fact whether or not the diffeomorphism, restricted to stable and unstable manifold, is orientation preserving.

The corresponding theorem for vector fields can be proved in a similar way. The main difference is in the fundamental domains, as we discussed before. The adaptations here are very similar to those discussed in the case of invariant foliations.

### Partial linearizations

In the case that we have a fixed point at which some of the proper values have norm 1, there is in general not a local conjugacy with the derivative (or linearization). In order to formulate the corresponding result in this case, we need some more notation. Let

$\varphi : M \rightarrow M$  be a diffeomorphism with invariant manifold  $V \subset M$ , i.e. such that  $\varphi(V) = V$ . We define the normal bundle  $N(V)$  of  $V$  in  $M$  as the vector bundle over  $V$  with fibres  $T_x(M)/T_x(V)$ . Since, for each  $x \in V$ ,  $d\varphi(x)$  maps  $T_x(M)$  to  $T_{\varphi(x)}(M)$  and  $T_x(V)$  to  $T_{\varphi(x)}(V)$ , we get an induced map, from  $N_x(V)$  to  $N_{\varphi(x)}(V)$ ; all these maps together define a map from  $N(V)$  to itself, which we denote by  $D\varphi$ ; this map is also called the *normal linearization* of  $\varphi$  along  $V$ . Also for such invariant manifolds one can define the notion as being normally hyperbolic, and the Hartman Grobman theorem was generalized in [27] to such normally hyperbolic invariant manifolds. We don't go into the formal definition of normal hyperbolicity but note that it means that the weakest contractions, respectively expansions, in the normal directions are stronger than the strongest contractions, respectively expansions in the submanifold  $V$ . So the centre manifolds, as we constructed them, are, at least in a sufficiently small neighbourhood of the fixed point, normally hyperbolic. The result for compact normally hyperbolic invariant manifolds is that the diffeomorphism near such a manifold is locally (in a neighbourhood of that manifold) conjugated to the normal linearization  $D\varphi$  in  $N(V)$ . With some modifications this result also applies to our centre manifolds. Without going into the details of the proof, the result is:

**THEOREM.** *Let  $\varphi : E \rightarrow E$  be a diffeomorphism with a fixed point in the origin. We assume that  $E$  has a splitting  $E = E^c + E^h$  which is invariant under the derivative  $L_\varphi$  of  $\varphi$  in the origin and such that all the proper values of  $L_\varphi|E^c$  have norm 1 and all proper values of  $L_\varphi|E^h$  have norm different from 1. Then  $\varphi$  is locally conjugated to the product of  $\varphi$  restricted to a centre manifold and  $L_\varphi|E^h$ .*  $\square$

This means that the local dynamics at a fixed point is determined, up to local conjugacy, by the dynamics in the centre manifold, together with the dimensions of stable and unstable manifolds and the fact whether or not the map is orientation preserving in the stable and the unstable manifold. Putting it in a more formal way:

**COROLLARY.** *If  $\varphi_1, \varphi_2 : E \rightarrow E$  are two diffeomorphisms as in the theorem with splittings  $E = E_1^c + E_1^h$  and  $E = E_2^c + E_2^h$  such that the dimensions of  $E_1^c$  and  $E_2^c$  are equal and if moreover the maps  $L_{\varphi_1}|E_1^h$  and  $L_{\varphi_2}|E_2^h$  are conjugated, then  $\varphi_1$  and  $\varphi_2$  are locally conjugated if and only if the restrictions to their centre manifolds are locally conjugated.*  $\square$

### Bifurcations

An important special case of the last corollary deals with bifurcations of diffeomorphisms. In bifurcation theory, one studies diffeomorphisms which depend on one or more parameters. In this case we denote the diffeomorphism by  $\varphi_\mu : \mathbf{R}^n \rightarrow \mathbf{R}^n$ , where  $\mu \in \mathbf{R}^p$  stands for the  $p$  parameters on which  $\varphi$  depends. Such a parametrized diffeomorphism can also be considered as one diffeomorphism in dimension  $n + p$ , namely  $\Phi : \mathbf{R}^{n+p} \rightarrow \mathbf{R}^{n+p}$ , defined by  $\Phi(x, \mu) = (\varphi_\mu(x), \mu)$ . If  $\varphi_0$  has a fixed point in 0, then  $\Phi$  also has a fixed point 0, but now the derivative of  $\Phi$  in 0 has a proper value 1 with multiplicity at least  $p$ . So  $\Phi$  has a centre manifold of dimension at least  $p$ . It turns out that in this case the local conjugacy with the normal linearization, as in the above theorem, can be made so that it does not change the  $\mu$ -coordinates, i.e. the conjugacy can be written

as  $H(x, \mu) = (h_\mu(x), \mu)$ . This means that it is in fact a  $p$ -parameter family of local conjugacies. So if we study bifurcation problems, and are only interested in aspects which are preserved by parametrized local conjugacies, we may restrict ourselves to bifurcations in centre manifolds.

### *Partial linearizations of vector fields*

As in the hyperbolic case, the results for vector fields are completely similar. For vector fields however one considers, besides (local) conjugacies also (local) equivalences which are defined as follows. Let  $X_1$  and  $X_2$  be vector fields on manifolds  $M_1$  and  $M_2$ . A homeomorphism  $h : M_1 \rightarrow M_2$  is an equivalence between  $X_1$  and  $X_2$  if  $h$  maps integral curves of  $X_1$  to integral curves of  $X_2$  in such a way that the ‘direction’ is preserved, i.e. such that, whenever  $x \in M_1$  and  $t_1 > 0$ , there is a  $t_2 > 0$ , depending on  $x$  and  $t_1$ , such that  $h(X_1^{t_1}(x)) = X_2^{t_2}(h(x))$ . Also with conjugacy replaced by equivalence, the above theorem remains true. The proof is somewhat more complicated, see [21].

### *Differentiable linearizations*

The local conjugacies in the above results are  $C^0$  and can in general not be made  $C^1$ . If however extra conditions are imposed on the proper values, then one can find smooth local conjugacies. For the linearizations of hyperbolic fixed points, this was done by Sternberg, [30] and [31]. The main results can be formulated, for  $C^\infty$ -diffeomorphisms, as follows.

**THEOREM (Formal Linearization, Sternberg [31]).** *Let  $\varphi : \mathbf{R}^n \rightarrow \mathbf{R}^n$  be a  $C^\infty$ -diffeomorphism with a fixed point in the origin with proper values  $\alpha_1, \dots, \alpha_n$ . Suppose these proper values satisfy the following condition:*

*For non-negative integers  $m_1, \dots, m_n$  with  $\sum_j m_j \geq 2$  and  $i = 1, \dots, n$  we have  $\alpha_i \neq \alpha_1^{m_1} \cdot \dots \cdot \alpha_n^{m_n}$  (note that this condition implies that none of the proper values can have norm 1).*

*Then there is a  $C^\infty$ -diffeomorphism  $h : \mathbf{R}^n \rightarrow \mathbf{R}^n$  with  $dh(0) = \text{Id}$  and such that the Taylor series of  $h\varphi h^{-1}$  in the origin contains only linear terms.*  $\square$

We say that  $h$ , as in the theorem, defines a formal linearization of  $\varphi$ . The proof of the theorem is based on formal manipulations with power series (or Taylor series) together with a theorem of Borel which states that for any formal power series there is a  $C^\infty$ -function which has that formal power series as Taylor series, see [2]. In a later section on normal forms we shall deal with these formal manipulations.

**THEOREM.** *Let  $\varphi_1, \varphi_2 : \mathbf{R}^n \rightarrow \mathbf{R}^n$  be  $C^\infty$ -diffeomorphisms which have a hyperbolic fixed point in the origin. If the infinite Taylor series of  $\varphi_1$  and  $\varphi_2$  in the origin are the same, then there is a  $C^\infty$ -diffeomorphism  $h : \mathbf{R}^n \rightarrow \mathbf{R}^n$  such that in some neighbourhood of the origin  $h\varphi_1 = \varphi_2h$ . In other words,  $h$  is a local  $C^\infty$ -conjugacy between  $\varphi_1$  and  $\varphi_2$ . Moreover this map  $h$  can be made so that its Taylor series at the origin is the Taylor series of the identity.*  $\square$

Also in the non hyperbolic case there are results concerning the construction of differentiable local conjugacies with normal linearizations, see [32].

## 5. Normal forms

Since, according to the previous results on partial linearizations, the local dynamics near a non-hyperbolic fixed point or equilibrium is essentially determined by the dynamics on a centre manifold, it is important to have techniques which allow us to calculate (or approximate) such centre manifolds and the dynamics on them. The main one of such techniques is given by normal form theory which uses coordinate transformations to bring the Taylor expansion of the diffeomorphism or vectorfield under consideration in a form which is ‘as simple as possible’. Such normal forms also play an important role in bifurcation theory. For a general survey on normal form theory we refer to [3].

To explain the basic idea (which is rather simple) we need some notation and (later on) a little bit of algebra. Let  $\varphi : E \rightarrow E$  be a smooth diffeomorphism with a fixed point in the origin, and let  $L_\varphi$  be the derivative of  $\varphi$  at this fixed point. For each integer  $k \geq 1$  we denote by  $[\varphi]_k : E \rightarrow E$  the Taylor expansion of  $\varphi$  at the origin up to order  $k$ , and by  $T_k \varphi : E \rightarrow E$  the homogeneous  $k$ -th order term in this Taylor expansion. So  $[\varphi]_1 = L_\varphi$  and  $T_k \varphi = [\varphi]_k - [\varphi]_{k-1}$  for  $k \geq 2$ . By  $H_k(E)$  we denote the vector space of homogeneous polynomial mappings  $\tilde{h}_k : E \rightarrow E$  of degree  $k$ . For each linear mapping  $L : E \rightarrow E$  we define the linear operator  $\text{ad}_k(L) : H_k(E) \rightarrow H_k(E)$  as the commutator with  $L$ :  $\text{ad}_k(L)\tilde{h}_k := L\tilde{h}_k - \tilde{h}_k L$  for each  $\tilde{h}_k \in H_k(E)$ .

Fix some  $k \geq 2$  and let  $h_k : E \rightarrow E$  be a  $C^\infty$ -diffeomorphism such that  $[h_k]_{k-1} = I$ ; further on we will then denote  $T_k h_k$  by  $\tilde{h}_k \in H_k(E)$ , such that  $[h_k]_k = I + \tilde{h}_k$ . It is then an elementary exercise to show that

$$[h_k \varphi h_k^{-1}]_k = [\varphi]_k - L_\varphi \tilde{h}_k + \tilde{h}_k L_\varphi = [\varphi]_k - \text{ad}_k(L_\varphi) \tilde{h}_k.$$

This shows that the conjugation with  $h_k$  only affects terms of order  $k$  or higher in the Taylor series, which allows us to use a step by step approach: first we handle the second order terms, then the third order terms, and so on. At the  $k$ -th step the aim is to choose  $\tilde{h}_k \in H_k(E)$  such that  $T_k(h_k \varphi h_k^{-1})$  becomes as simple as possible, preferably zero (here  $\varphi$  is not the original diffeomorphism, but the diffeomorphism obtained after the conjugation of the foregoing step). We can only achieve setting  $T_k(h_k \varphi h_k^{-1}) = 0$  if we can choose  $\tilde{h}_k \in H_k(E)$  such that  $\text{ad}_k(L_\varphi) \tilde{h}_k = T_k \varphi$ , i.e. if  $T_k \varphi$  belongs to the range of  $\text{ad}_k(L_\varphi)$ ; the easiest way to satisfy this condition is to assume that  $\text{ad}_k(L_\varphi)$  is invertible. As we will show further on, the proper values of  $\text{ad}_k(L_\varphi)$  have the form

$$\beta_{i,m_1,\dots,m_n} = \alpha_i - \prod_{j=1}^n \alpha_j^{m_j},$$

with  $i = 1, \dots, n$  and  $m_1, \dots, m_n$  non-negative integers such that  $\sum_j m_j = k$ , and where  $\alpha_1, \dots, \alpha_n$  are the proper values of  $L_\varphi$ . When  $\beta_{i,m_1,\dots,m_n} \neq 0$  for all  $i = 1, \dots, n$  and for all non-negative integers  $m_1, \dots, m_n$  with  $\sum_j m_j \geq 2$  then  $\text{ad}_k(L_\varphi)$  is invertible for all  $k \geq 2$ , and at each step in the normal form reduction we can choose  $\tilde{h}_k$  such that  $T_k(h_k \varphi h_k^{-1}) = 0$ . In the end we obtain a conjugation of the original  $\varphi$  with a diffeomorphism whose Taylor series is that of the linear map  $L_\varphi$ , thus proving the formal linearization theorem of Sternberg (see Section 4). The conjugating homeomorphism

$h$  should be such that  $[h]_1 = I$  and  $T_k h = T_k(h_k h_{k-1} \cdots h_2)$  for  $k \geq 2$ ; such a diffeomorphism exists by the theorem of Borel mentioned earlier.

In case where the conditions of Sternberg's theorem are not satisfied (in particular when the fixed point is non-hyperbolic), we can proceed as follows. For each  $k \geq 2$  we choose a complement  $N_k$  of the range  $\text{Rg}(\text{ad}_k(L_\varphi))$  of  $\text{ad}_k(L_\varphi)$  in  $H_k(E)$ , we write (at the  $k$ -th step in the normalization process)  $T_k \varphi = \psi_k + v_k$ , with  $\psi_k \in \text{Rg}(\text{ad}_k(L_\varphi))$  and  $v_k \in N_k$ , and we choose  $\tilde{h}_k \in H_k(E)$  such that  $\text{ad}_k(L_\varphi)\tilde{h}_k = \psi_k$ . Then  $T_k(h_k \varphi h_k^{-1}) = v_k \in N_k$ , and by the same argument as above we obtain the following result.

**THEOREM (Formal Normalization).** *Let  $\varphi : E \rightarrow E$  be a  $C^\infty$ -diffeomorphism with a fixed point at the origin and with derivative  $L_\varphi$  at this fixed point. For each  $k \geq 2$  let  $N_k$  be a subspace of  $H_k(E)$  such that  $H_k(E) = \text{Rg}(\text{ad}_k(L_\varphi)) + N_k$ . Then there exists a  $C^\infty$ -diffeomorphism  $h : E \rightarrow E$  with  $h(0) = 0$  and  $dh(0) = I$  such that  $T_k(h\varphi h^{-1}) \in N_k$  for each  $k \geq 2$ .*

Under the conditions of the theorem we call the Taylor series of  $h\varphi h^{-1}$  the *normal form* of  $\varphi$ , and  $[h\varphi h^{-1}]_k$  (with  $k \geq 2$ ) its *truncated normal form up to order  $k$* . When the diffeomorphism  $\varphi$  is such that  $T_k \varphi \in N_k$  for all  $k \geq 2$  then we say that  $\varphi$  is *in normal form*; when  $T_m(\varphi) \in N_m$  for  $2 \leq m \leq k$  we say that  $\varphi$  is *in normal form up to order  $k$* .

The formulation of the foregoing theorem leaves much freedom in the choice of the 'normal form subspaces'  $N_k$ , and as a consequence there are many types of normal forms. An obvious question is whether it is possible to choose these normal form subspaces in a more systematic way; for the answer some algebra will be useful. By the Jordan-Chevalley decomposition theorem (see e.g. [13]) every linear operator  $L$  on  $E$  can be split in a unique way as  $L = S + N$ , where  $S$  is semi-simple (i.e. complex diagonalizable),  $N$  is nilpotent, and  $SN = NS$ . The proper values of  $L$  and  $S$  coincide, the space  $E$  can be split as  $E = \text{Ker}(S) + \text{Rg}(S)$ , and the range of  $L$  can be written as  $\text{Rg}(L) = \text{Rg}(S) + (\text{Ker}(S) \cap \text{Rg}(N))$ . If  $V$  is a complement of  $\text{Ker}(S) \cap \text{Rg}(N)$  in  $\text{Ker}(S)$  then  $V$  is also a complement of  $\text{Rg}(L)$  in  $E$ :  $E = \text{Rg}(L) + V$ . We will show now that for each  $k \geq 1$  the Jordan-Chevalley decomposition of  $\text{ad}_k(L)$  is given by  $\text{ad}_k(L) = \text{ad}_k(S) + \text{ad}_k(N)$ ; at the same time we will also obtain the proper values of  $\text{ad}_k(L)$ .

It follows immediately from the definitions that  $\text{ad}_k(L) = \text{ad}_k(S) + \text{ad}_k(N)$  and that  $\text{ad}_k(S)\text{ad}_k(N) = \text{ad}_k(N)\text{ad}_k(S)$ . For each integer  $r \geq 1$  and each  $\tilde{h}_k \in H_k(E)$  the expression for  $(\text{ad}_k(N))^r \tilde{h}_k$  consists of a finite number of terms of the form  $\pm N^p \tilde{h}_k N^q$ , with  $p + q = r$ ; since  $N$  is nilpotent the same is true for  $\text{ad}_k(N)$ . It remains to show that  $\text{ad}_k(S)$  is semi-simple. Let  $\alpha_1, \dots, \alpha_n$  be the proper values of  $S$  (and of  $L$ ), and let  $\zeta_1, \dots, \zeta_n \in E^c$  (the complexification of  $E$ ) be a corresponding set of proper vectors:  $S\zeta_j = \alpha_j \zeta_j$  for  $1 \leq j \leq n$ . For each  $i = 1, \dots, n$  and for each set of non-negative integers  $m_1, \dots, m_n$  with  $\sum_j m_j = k$  we define an element  $h_{i,m_1,\dots,m_n}$  of  $H_k(E^c)$  by

$$h_{i,m_1,\dots,m_n} \left( \sum_j z_j \zeta_j \right) := \left( \prod_j z_j^{m_j} \right) \zeta_i, \quad \forall z_1, \dots, z_n \in \mathbb{C}.$$

These polynomial mappings form a basis of  $H_k(E^c)$ , and a direct calculation shows that

$$\text{ad}_k(S)h_{i,m_1,\dots,m_n} = (\alpha_i - \prod_j \alpha_j^{m_j})h_{i,m_1,\dots,m_n},$$

i.e.  $h_{i,m_1,\dots,m_n}$  is a proper vector of  $\text{ad}_k(S)$  with proper value  $\beta_{i,m_1,\dots,m_n} = \alpha_i - \Pi_j \alpha_j^{m_j}$ . We conclude that  $\text{ad}_k(S)$  is semi-simple, that  $\text{ad}_k(L) = \text{ad}_k(S) + \text{ad}_k(N)$  is a Jordan-Chevalley decomposition, and that  $\text{ad}_k(L)$  has the same proper values as  $\text{ad}_k(S)$ , namely  $\beta_{i,m_1,\dots,m_n}$  with  $1 \leq i \leq n$ ,  $m_j \geq 0$  and  $\sum_j m_j = k$ .

We now apply these results on Jordan-Chevalley decompositions to our normal form reduction: writing the Jordan-Chevalley decomposition of  $L_\varphi$  as  $L_\varphi = S_\varphi + N_\varphi$  we conclude from these results that we can choose a complement  $N_k$  of  $\text{Rg}(\text{ad}_k(L_\varphi))$  in  $H_k(E)$  by taking for  $N_k$  a complement of  $\text{Ker}(\text{ad}_k(S_\varphi)) \cap \text{Rg}(\text{ad}_k(N_\varphi))$  in  $\text{Ker}(\text{ad}_k(S_\varphi))$ . Before saying more on how to choose this complement let us already point out an important consequence: using this approach always results in normal form subspaces  $N_k$  which are subspaces of  $\text{Ker}(\text{ad}_k(S_\varphi))$ , which means that the corresponding normal form commutes with  $S_\varphi$ . In combination with the formal normal form theorem, this gives the following.

**THEOREM.** *Let  $\varphi : E \rightarrow E$  be a  $C^\infty$ -diffeomorphism with a fixed point at the origin, and let  $S_\varphi$  be the semi-simple part of the derivative  $L_\varphi = d\varphi(0)$ . Then there exists a  $C^\infty$ -diffeomorphism  $h : E \rightarrow E$  with  $h(0) = 0$  and  $dh(0) = I$  such that the Taylor series of  $h\varphi h^{-1}$  at the origin commutes with  $S_\varphi$ .*

In the case where all proper values of  $L_\varphi$  have norm equal to 1 (or when we are restricted to a centre manifold) the foregoing result can be sharpened. Then, not only does the Taylor series of the transformed diffeomorphism commute with  $S_\varphi$ , there actually exists a diffeomorphism which commutes with  $S_\varphi$  and has the same Taylor series; this is stated in the following theorem, a proof of which one can find in [33].

**THEOREM.** *Let  $\varphi : E \rightarrow E$  be a  $C^\infty$ -diffeomorphism with a fixed point at the origin, and let  $S_\varphi$  be the semi-simple part of the derivative  $L_\varphi = d\varphi(0)$ . Suppose that all proper values of  $L_\varphi$  have norm equal to 1. Then there exists a  $C^\infty$ -diffeomorphism  $h : E \rightarrow E$  and a  $C^\infty$  vector field  $X$  on  $E$  such that*

- $h(0) = 0$  and  $dh(0) = I$ ;
- $X$  commutes with  $S_\varphi$ ;
- the Taylor series at the origin of  $h\varphi h^{-1}$  and  $X_1 S_\varphi$  coincide.

Here  $X_1 : E \rightarrow E$  is the time one map of the flow on  $E$  defined by the vector field  $X$ ; clearly  $X_1$  and hence also  $X_1 S_\varphi$  commute with  $S_\varphi$ .

Under the conditions of this theorem we say that  $X_1 S_\varphi$  is an *integrable approximation* of  $h\varphi h^{-1}$ , and that  $\varphi^{(i)} := h^{-1} X_1 S_\varphi h$  is an integrable approximation of  $\varphi$ . It is not hard to prove (see e.g. [5]) that every neighbourhood of  $\varphi$  in the  $C^\infty$ -topology contains a diffeomorphism  $\tilde{\varphi}$  such that  $\tilde{\varphi} - \varphi^{(i)} \equiv 0$  on some neighbourhood of the origin in  $E$ . In general one does not have that  $\varphi - \varphi^{(i)} \equiv 0$  on some neighbourhood of the origin. Also, although integrable approximations always exist, they are in general not unique.

In order to understand some of the consequences of the foregoing theorems let us consider for a moment the case of an integrable diffeomorphism  $\varphi$ , that is, we assume that  $\varphi$  commutes with  $S_\varphi$ ; also let  $E = E^s + E^c + E^u$  be the  $L_\varphi$ -invariant splitting of  $E$  such that the proper values of  $L_\varphi$  restricted to  $E^s$ ,  $E^c$  and  $E^u$  are in norm respectively



smaller, equal to and bigger than one. Then these subspaces are also invariant under  $S_\varphi$ , and the proper values of  $S_\varphi$  restricted to  $E^s$ ,  $E^c$  and  $E^u$  are the same as for the corresponding restrictions of  $L_\varphi$ . We claim that due to our assumption that  $\varphi S_\varphi = S_\varphi \varphi$  the subspaces  $E^s$ ,  $E^c$  and  $E^u$  are also invariant under the diffeomorphism  $\varphi$  itself. To see this observe that for  $u, v \in E$  we have  $\lim_{m \rightarrow \infty} S_\varphi^m u = 0$  if and only if  $u \in E^s$ , and  $\lim_{m \rightarrow \infty} S_\varphi^{-m} v = 0$  if and only if  $v \in E^u$ . For each  $u \in E^s$  we have

$$\lim_{m \rightarrow \infty} S_\varphi^m \varphi(u) = \lim_{m \rightarrow \infty} \varphi(S_\varphi^m u) = \varphi(0) = 0,$$

and hence  $\varphi(u) \in E^s$ ; in a similar way one shows that  $\varphi(v) \in E^u$  for all  $v \in E^u$ . For each  $w \in E^c$  we can find a strictly increasing sequence of integers  $m_1 < m_2 < \dots$  such that  $\lim_{p \rightarrow \infty} S_\varphi^{m_p} w = \lim_{p \rightarrow \infty} S_\varphi^{-m_p} w = w$  (this follows from the fact that  $S_\varphi$  is semi-simple and that the restriction of  $S_\varphi$  to  $E^c$  has only proper values with norm 1). Consequently

$$\varphi(w) = \lim_{p \rightarrow \infty} \varphi(S_\varphi^{\pm m_p} w) = \lim_{p \rightarrow \infty} S_\varphi^{\pm m_p} \varphi(w),$$

which is only possible when  $\varphi(w) \in E^c$ . From this  $\varphi$ -invariance of  $E^s$ ,  $E^u$  and  $E^c$  it follows that the (unique) stable and unstable sets of the fixed point at the origin are given by  $W^s = E^s$  and  $W^u = E^u$ , while  $W^c = E^c$  forms a centre manifold.

Combining these considerations with the last theorem we conclude that for a general diffeomorphism  $\varphi : E \rightarrow E$  with a fixed point at the origin there exists a smooth conjugacy  $h : E \rightarrow E$  which formally flattens the stable and unstable sets of the fixed point and which also gives a flat formal centre manifold. More precisely: up to terms which are  $C^\infty$ -flat at the origin we have  $W^s = h(E^s)$  and  $W^u = h(E^u)$ , and for each  $k \geq 2$  the origin has a local centre manifold of class  $C^k$  which, up to terms of order  $k + 1$ , can be approximated by  $h(E^c)$ . The local dynamics on this centre manifold is, again up to terms of order  $k + 1$ , conjugated to the dynamics of the truncated normal form  $[h\varphi h^{-1}]_k$  restricted to the centre subspace  $E^c$ . This way the normal form reduction gives us, up to any desired order, approximations for a local centre manifold and the dynamics on this centre manifold. This is particularly important in bifurcation theory, but we will come back to that later on.

We still have to describe a way to choose the normal form subspace  $N_k$  as a complement of  $\text{Ker}(\text{ad}_k(S_\varphi)) \cap \text{Rg}(\text{ad}_k(N_\varphi))$  in  $\text{Ker}(\text{ad}_k(S_\varphi))$ . There are several ways to do this; probably the simplest one is based on the fact that for a linear operator on a (finite-dimensional) vectorspace the range of the operator and the kernel of its transpose (adjoint) are complementary, where the transpose can be taken with respect to any convenient Euclidian structure. We explain now how this can be used to make a choice for  $N_k$ ; our approach is a modification of the one in [8]. We start with a technical lemma.

**LEMMA.** *Let  $(\cdot, \cdot)$  be a scalar product on  $E$ . Then there exists for each  $k \geq 1$  an associated scalar product  $\langle \cdot, \cdot \rangle_k$  on  $H_k(E)$  such that for each linear operator  $L : E \rightarrow E$  the transpose  $(\text{ad}_k(L))^T$  of  $\text{ad}_k(L)$  with respect to  $\langle \cdot, \cdot \rangle_k$  is given by  $\text{ad}_k(L^T)$ , where  $L^T$  is the transpose of  $L$  with respect to  $(\cdot, \cdot)$ .*

**PROOF.** Using an orthonormal basis of  $E$  we can identify  $E$  with  $\mathbf{R}^n$ ,  $H_k(E)$  with  $H_k(\mathbf{R}^n)$ , and  $(\cdot, \cdot)$  with the standard scalar product on  $\mathbf{R}^n$ . For each  $n$ -vector  $m = (m_1, \dots, m_n) \in$



$\mathbf{N}^n$  of non-negative integers we define  $|m| := \sum_i m_i$  and  $m! := \prod_i m_i!$ , and a homogeneous polynomial  $p_m : \mathbf{R}^n \rightarrow \mathbf{R}$  given by  $p_m(x) := \prod_i x_i^{m_i}$ . For two polynomials  $p, q : \mathbf{R}^n \rightarrow \mathbf{R}$  given by respectively  $p(x) = \sum' a_m p_m(x)$  and  $q(x) = \sum' b_m p_m(x)$  (where  $a_m, b_m \in \mathbf{R}$  and  $\sum'$  indicates that only a finite number of terms are different from zero and that  $|m| \geq 1$ ) we define the polynomial  $p(\partial) \cdot q$  by

$$(p(\partial) \cdot q)(x) = \sum' a_m \partial_1^{m_1} \partial_2^{m_2} \dots \partial_n^{m_n} q(x).$$

For each linear operator  $L$  on  $\mathbf{R}^n$  (with  $(Lx)_i = \sum_j L_{ij}x_j$  and  $(L^T x)_i = \sum_j L_{ji}x_j$ ) we have

$$p(\partial) \cdot (qL) = ((pL^T)(\partial) \cdot q)L;$$

this follows easily from the chain rule when  $p$  has degree one, and can then be extended to general polynomials  $p$  by induction on the degree. Elements  $g, h$  of  $H_k(\mathbf{R}^n)$  can be written as  $g = \sum_{|m|=k} \alpha_m p_m$  and  $h = \sum_{|m|=k} \beta_m p_m$ , with  $\alpha_m, \beta_m \in \mathbf{R}^n$ ; we define the scalar product  $\langle g, h \rangle_k$  of these two elements as

$$\langle g, h \rangle_k := \sum_{|m|=k} m! (\alpha_m, \beta_m) = \sum_{|m|=k, |m'|=k} (\alpha_m, \beta_{m'}) p_m(\partial) \cdot p_{m'}(0).$$

We have then for each linear operator  $L$  on  $\mathbf{R}^n$  that

$$\begin{aligned} \langle g, \text{ad}_k(L)h \rangle_k &= \langle g, Lh \rangle_k - \langle g, hL \rangle_k \\ &= \sum_{|m|=k, |m'|=k} [(\alpha_m, L\beta_{m'}) p_m(\partial) \cdot p_{m'}(0) \\ &\quad - (\alpha_m, \beta_{m'}) p_m(\partial) \cdot (p_{m'}L)(0)] \\ &= \sum_{|m|=k, |m'|=k} \left[ (L^T \alpha_m, \beta_{m'}) p_m(\partial) \cdot p_{m'}(0) \right. \\ &\quad \left. - (\alpha_m, \beta_{m'}) (p_m L^T)(\partial) \cdot (p_{m'})(0) \right] \\ &= \langle \text{ad}_k(L^T)g, h \rangle_k. \end{aligned}$$

This shows that  $\text{ad}_k(L^T)$  is indeed the transpose of  $\text{ad}_k(L)$  on  $H_k(\mathbf{R}^n)$ .  $\square$

In order to apply this to our normal forms we will take  $L = L_\varphi$  and choose the scalar product  $(\cdot, \cdot)$  in an appropriate way which we explain now. Observe that since  $\text{ad}_k(N_\varphi)$  and  $\text{ad}_k(S_\varphi)$  commute,  $\text{ad}_k(N_\varphi)$  leaves  $\text{Ker}(\text{ad}_k(S_\varphi))$  invariant. If we choose the scalar product  $(\cdot, \cdot)$  on  $E$  such that not only  $S_\varphi$  and  $N_\varphi$  commute, but also  $S_\varphi$  and  $N_\varphi^T$ , then  $\text{ad}_k(N_\varphi^T)$  and  $\text{ad}_k(S_\varphi)$  commute, and  $\text{ad}_k(N_\varphi^T)$  leaves  $\text{Ker}(\text{ad}_k(S_\varphi))$  invariant. In combination with the Lemma this proves that

$$N_k = \text{Ker}(\text{ad}_k(S_\varphi)) \cap \text{Ker}(\text{ad}_k(N_\varphi^T))$$

is then a complement of  $\text{Ker}(\text{ad}_k(S_\varphi)) \cap \text{Rg}(\text{ad}_k(N_\varphi))$  in  $\text{Ker}(\text{ad}_k(S_\varphi))$ . The existence of the required scalar product  $(\cdot, \cdot)$  can easily be shown by considering the standard real Jordan normal form of  $L_\varphi$ .

The normal forms which result from the above choice of  $N_k$  commute with  $S_\varphi$  and, except for the single linear term  $N_\varphi$ , also with  $N_\varphi^T$ . Observe that this approach requires the choice of a scalar product which is tied to the structure of the linearization  $L_\varphi$ .

In most cases the actual calculation of these normal forms for a particular homeomorphism  $\varphi$  remains a formidable task, certainly when the dimension of  $E$  increases and when one wants a normal form up to orders higher than 2 or 3. The bottleneck lies in the fact that for each  $k \geq 2$  and for some  $\psi_k \in \text{Rg}(\text{ad}_k(L_\varphi))$  obtained from the foregoing steps, one has to determine some  $\tilde{h}_k \in H_k(E)$  such that  $\text{ad}_k(L_\varphi)\tilde{h}_k = \psi_k$ . The dimension of this linear problem grows very fast with  $k$  and the dimension of  $E$ , thus making it difficult to write a sufficiently general algorithm which can handle all reasonable cases in an efficient way. A different approach which seems to considerably alleviate this problem is based on the representation theory of  $\text{sl}(2, \mathbf{R})$ -algebras (see e.g. [13]). In this approach the transpose  $N_\varphi^T$ , which we used in the foregoing, is replaced by some other nilpotent linear operator  $M_\varphi$  which also commutes with  $S_\varphi$  and which is such that  $N_\varphi$  and  $M_\varphi$ , together with their commutator  $H_\varphi := [N_\varphi, M_\varphi]$ , generate an  $\text{sl}(2, \mathbf{R})$ -algebra; this allows us to use the extensive representation theory which exists for such algebras. It would lead us too far to go here into further details about this approach, we just refer to [7] for more information.

### 5.1. Normal forms for vector fields

The normal form theory for diffeomorphisms as outlined above can easily be adapted to the case of vector fields; we just briefly indicate a few of the changes which have to be made. Let  $X$  be a  $C^\infty$ -smooth vector field on  $E$ , with an equilibrium at the origin, and with linearization  $A_0 = dX(0)$  at that equilibrium. A diffeomorphism  $h : E \rightarrow E$  transforms this vector field in a new vector field  $h_*X$  (the *push-forward* of  $X$  under  $h$ ) given by  $(h_*X)(h(u)) = dh(u) \cdot X(u)$  for each  $u \in E$ . When  $h$  is such that  $[h]_k = I + \tilde{h}_k$  for some  $k \geq 2$  and  $\tilde{h}_k \in H_k(E)$ , then

$$[h_*X]_k = [X]_k - \text{ad}_k(A_0) \cdot \tilde{h}_k,$$

where for each linear operator  $A$  on  $E$  and for each  $k \geq 1$  the linear operator  $\text{ad}_k(A)$  on  $H_k(E)$  is defined by

$$(\text{ad}_k(A) \cdot \tilde{h}_k)(u) := A(\tilde{h}_k(u)) - d\tilde{h}_k(u) \cdot A(u), \quad \forall \tilde{h}_k \in H_k(E), \quad \forall u \in E.$$

Note that  $\text{ad}_k$  here is different from the  $\text{ad}_k$  as used before. With these changes in mind the foregoing results on normal forms and integrable approximations for diffeomorphisms can be easily adapted to vector fields, see e.g. [36,34] and [5] for details.

### 5.2. Normal forms in bifurcation theory

When instead of one single diffeomorphism  $\varphi$  we have a family of diffeomorphisms  $\varphi_\mu : E \rightarrow E$  depending smoothly on a (multi)-parameter  $\mu \in \mathbf{R}^p$  as encountered typically in bifurcation problems, several approaches are possible. One can directly apply the foregoing theory to the augmented diffeomorphism  $\Phi : E \times \mathbf{R}^p \rightarrow E \times \mathbf{R}^p$  given by

$\Phi(u, \mu) := (\varphi_\mu(u), \mu)$ ; the only requirement is that for  $\mu = 0$  the diffeomorphism  $\varphi_0$  has a fixed point at the origin:  $\varphi_0(0) = 0$ , and hence  $\Phi(0, 0) = (0, 0)$ . The special triangular structure of  $\Phi$  (which distinguishes between phase space variables  $u$  and parameters  $\mu$ ) can be retained during the normal form transformation. Formally, the diffeomorphism preserves the projection  $\pi$  of  $E \times \mathbf{R}^p$  onto  $\mathbf{R}^p$ , in the sense that  $\pi \circ \Phi = \pi$ . By allowing only transformations  $h : E \times \mathbf{R}^p \rightarrow E \times \mathbf{R}^p$  which have the same property  $\pi \circ h = \pi$ —this means that for  $k \geq 2$  one considers only those polynomial mappings in  $H_k(E \times \mathbf{R}^p)$  for which the  $\mathbf{R}^p$ -components are identically zero—the final normal form  $h\Phi h^{-1}$  will also preserve  $\pi$ . A basic aspect of this approach is that one takes Taylor expansions with respect to both the state variable  $u$  and the parameter  $\mu$ ; from a bifurcation point of view (‘How does the system change with the parameter?’) this may sometimes be less desirable. However, in those cases where the fixed point of  $\varphi_0$  does not persist for all parameter values in a full neighbourhood of  $\mu = 0$  (for example, at a saddle node of fixed points), this approach seems to be the only one available. The other approach which we will describe next requires that (possibly after some preliminary transformations) we have  $\varphi_\mu(0) = 0$  for all  $\mu$ .

Before going on let us remark that the idea described in the previous paragraph (a normal form reduction which preserves the projection  $\pi$ ) can be extended to several other special structures: symplecticity, preservation of volumes, equivariance, reversibility, and so on. For some examples on how this works see e.g. [4].

Returning to our parametrized family of diffeomorphisms it is not a good idea to treat each of the  $\varphi_\mu$  as a separate diffeomorphism and bring it into normal form: although  $\text{ad}_k(L_\mu)$  (with  $L_\mu := d\varphi_\mu(0)$ ) depends smoothly on  $\mu$ , the mapping  $\mu \mapsto \dim \text{Rg}(\text{ad}_k(L_\mu))$  is only lower-semi-continuous, and hence also the normal form subspaces  $N_k$  may show sudden jumps when the parameter is changed. To avoid this we will use the linearization  $L_0$  at  $\mu = 0$  and the corresponding operators  $\text{ad}_k(L_0)$  to define the normal form subspaces  $N_k$ , even for  $\mu \neq 0$ . As we will see the drawback of this is that the validity range in parameter space may shrink at each step in the normal form reduction, and therefore only normal forms up to a finite (but arbitrary) order can be obtained. We now make this more precise.

We assume that  $\varphi_\mu(0) = 0$  and set  $L_\mu := d\varphi_\mu(0)$ . Let  $L_0 = S_0 + N_0$  be the Jordan-Chevalley decomposition of  $L_0$ , and choose a Euclidean structure on  $E$  such that both  $N_0$  and  $N_0^T$  commute with  $S_0$ . Then the following result holds.

**THEOREM.** *For each  $k \geq 1$  there exists a neighbourhood  $\omega_k$  of the origin in the parameter space  $\mathbf{R}^p$  and a smooth mapping  $h^{(k)} : E \times \omega_k \rightarrow E$  such that the following holds:*

- for each  $\mu \in \omega_k$  the mapping  $h_\mu^{(k)} = h^{(k)}(\cdot, \mu) : E \rightarrow E$  is a diffeomorphism;
- $h_\mu^{(k)}(0) = 0$  for all  $\mu \in \omega_k$ , and  $dh_0^{(k)}(0) = I$ ;
- $T_1 \left( h_\mu^{(k)} \varphi_\mu(h_\mu^{(k)})^{-1} \right) - L_0 \in \text{Ker}(\text{ad}_1(S_0)) \cap \text{Ker}(\text{ad}_1(N_0^T))$  for all  $\mu \in \omega_k$ ;
- $T_m \left( h_\mu^{(k)} \varphi_\mu(h_\mu^{(k)})^{-1} \right) \in \text{Ker}(\text{ad}_m(S_0)) \cap \text{Ker}(\text{ad}_m(N_0^T))$  for all  $\mu \in \omega_k$  and for all  $m$  with  $2 \leq m \leq k$ .

As noted above the neighbourhoods  $\omega_k$  may shrink to just the origin as  $k$  increases to infinity, and therefore the theorem is in general not valid for  $k = \infty$ .

PROOF. For each  $k \geq 1$  we have the splitting

$$H_k(E) = \text{Rg}(\text{ad}_k(L_0)) + \left[ \text{Ker}(\text{ad}_k(S_0)) \cap \text{Ker}(\text{ad}_k(N_0^T)) \right];$$

we denote by  $\pi_k$  the projection of  $H_k(E)$  onto  $\text{Rg}(\text{ad}_k(L_0))$  associated with this splitting. We prove the theorem using induction in  $k$ , starting with  $k = 1$ . Define  $\psi_1 : H_1(E) \times \mathbf{R}^p \rightarrow \text{Rg}(\text{ad}_1(L_0))$  by setting

$$\psi_1(A, \mu) := \pi_1 \left[ T_1(A\varphi_\mu A^{-1}) - A_0 \right] = \pi_1 \left[ AL_0 A^{-1} - L_0 \right]$$

for all  $A \in H_1(E)$  (the space of linear operators on  $E$ ) and for all  $\mu \in \mathbf{R}^p$ . (Formally  $\psi_1$  is only defined on the open subset of  $H_1(E)$  consisting of *invertible* linear operators). Then  $\psi_1(I, 0) = 0$  and  $d_A \psi_1(I, 0) = -\pi_1 \text{ad}_1(L_0)$  is surjective as a linear operator from  $H_1(E)$  into  $\text{Rg}(\text{ad}_1(L_0))$ ; this shows that  $\psi_1$  is a submersion at  $(A, \mu) = (I, 0)$ , and by the implicit function theorem we can find a smooth mapping  $\mu \mapsto A_\mu$  such that  $A_0 = I$  and  $\psi_1(A_\mu, \mu) = 0$  for all  $\mu$  in an open neighbourhood  $\omega_1$  of the origin in  $\mathbf{R}^p$ . Setting  $h_\mu^{(1)} = A_\mu$  then proves the theorem for the case  $k = 1$ .

Next fix some  $k \geq 2$ ; by the induction hypothesis, and by considering  $h_\mu^{(k-1)} \varphi_\mu (h_\mu^{(k-1)})^{-1}$  as a new  $\varphi_\mu$  we may then assume that for all  $\mu$  in a neighbourhood  $\omega_{k-1}$  of the origin in  $\mathbf{R}^p$  we have  $L_\mu - L_0 \in \text{Ker}(\text{ad}_1(S_0)) \cap \text{Ker}(\text{ad}_1(N_0^T))$  and  $T_m(\varphi_\mu) \in \text{Ker}(\text{ad}_m(S_0)) \cap \text{Ker}(\text{ad}_m(N_0^T))$  for  $2 \leq m \leq k-1$ . Consider the family of linear operators  $\pi_k \text{ad}_k(L_\mu)$  ( $\mu \in \omega_{k-1}$ ) from  $H_k(E)$  into  $\text{Rg}(\text{ad}_k(L_0))$ ; for  $\mu = 0$  this operator is surjective (by definition of  $\pi_k$ ), and therefore this remains so for all  $\mu$  in a neighbourhood  $\omega_k \subset \omega_{k-1}$  of the origin in  $\mathbf{R}^p$ . It follows that we can find a smooth mapping  $\mu \in \omega_k \mapsto \tilde{h}_k(\mu) \in H_k(E)$  such that

$$\pi_k \left[ T_k(\varphi_\mu) - \text{ad}_k(L_\mu) \tilde{h}_k(\mu) \right] = 0, \quad \forall \mu \in \omega_k.$$

Finally, let  $h_\mu^{(k)}$  ( $\mu \in \omega_k$ ) be a family of diffeomorphisms on  $E$  such that  $[h_\mu^{(k)}]_k = I + \tilde{h}_k(\mu)$ ; since

$$\left[ h_\mu^{(k)} \varphi_\mu (h_\mu^{(k)})^{-1} \right]_k = [\varphi_\mu]_k - \text{ad}_k(L_\mu) \tilde{h}_k(\mu)$$

such a family satisfies all the requirements of the theorem.  $\square$

A somewhat different version of the foregoing theorem can be found in [37]; a similar result also holds for vectorfields depending smoothly on parameters.

We finish this subsection with a simple example which nicely illustrates the fact that in general it is impossible to bring the full family  $\varphi_\mu$  in normal form up to infinite order. We take  $E = \mathbf{R}^2$ ,  $p = 1$ , and we consider a family  $\varphi_\mu$  of diffeomorphisms on  $E$  with  $\varphi_\mu(0) = 0$  and such that

$$L_\mu = \begin{pmatrix} e^{\Omega+\mu} & 0 \\ 0 & e^{-1} \end{pmatrix};$$

here  $\Omega$  is a positive irrational number. One can easily verify that  $L_0$  satisfies the conditions of the Sternberg theorem, and hence the diffeomorphism  $\varphi_0$  can be formally linearized. This also implies that in the proof above the projection  $\pi_k$  will be the identity on  $H_k(E)$ , and the  $k$ -th step in the normalization procedure for the family  $\varphi_\mu$  will work for those parameter values  $\mu$  for which  $\text{ad}_k(L_\mu)$  is invertible on  $H_k(E)$  (the family  $\varphi_\mu$  is already in normal form up to first order, by the assumed form of  $L_\mu$ ). This means, in particular, that  $\omega_k$  (for  $k \geq 2$ ) will be an open interval around the origin such that  $\Omega + \omega_k$  does not contain any rational number  $p/q$  with  $p, q > 0$ ,  $\gcd(p, q) = 1$  and  $p + q \leq k - 1$ ; indeed, when  $\Omega + \mu = p/q$  for such  $p, q$ , then one can verify, using our earlier results on the proper values of  $\text{ad}_k(L_\varphi)$ , that  $\text{ad}_k(L_\mu)$  is not invertible. As  $k$  increases to infinity this excludes *all* rational numbers as possible values for  $\Omega + \mu$ , and since  $\Omega$  can be approximated arbitrarily close by rationals this means that  $\omega_k$  will shrink to a point (the origin) when  $k$  goes to infinity.

## 6. Liapunov-Schmidt reduction

In order to study bifurcations near the fixed point at the origin for a parametrized family  $\varphi_\mu$  of diffeomorphisms (with  $\mu \in \mathbf{R}^p$ ) we can restrict our study to a centre manifold, and on such a centre manifold the family can, up to any finite order, be approximated by an integrable family  $\varphi_\mu^{(i)}$ . When this integrable family shows some interesting behaviour (such as e.g. periodic or quasi-periodic orbits), we would like to find out whether or not this behaviour persists for the original family. To answer such question we have to take the remainder terms into account, thus destroying the  $S_0$ -equivariance of the integrable family. As explained e.g. in [5] this destruction of the symmetry may lead to quite different behaviours for the original family and for the integrable approximation.

There is one exception to the foregoing, namely for the bifurcation of periodic orbits: as we will explain in this section, this problem can be reformulated in such a way that we have the same symmetry for both the original family and the integrable approximation. The approach is based on a rather elementary version of the Liapunov-Schmidt reduction technique. A general survey of this technique can be found in [38], and an extensive discussion in the case of vector fields in e.g. [35]. The presentation given here is from [37].

Consider a family of  $C^\infty$ -diffeomorphisms  $\varphi_\mu : E \rightarrow E$  ( $\mu \in \mathbf{R}^p$ ) with a fixed point at the origin, and fix some integer  $q \geq 1$ . We want to determine, for all sufficiently small values of the parameter  $\mu$ , all small  $q$ -periodic orbits of  $\varphi_\mu$ . When  $\varphi_\mu$  is linear, say  $\varphi_\mu = d\varphi_\mu(0) = L_\mu$ , this is a simple algebra problem: we just have to determine the proper space corresponding to the proper value one of the linear operator  $L_\mu^q$ . For the general case we will reformulate the problem in the *orbit space*  $\mathcal{O}_q$  consisting of all bi-infinite sequences  $v = (v_j)_{j \in \mathbf{Z}}$  in  $E$  which are  $q$ -periodic:  $v_j \in E$  and  $v_{j+q} = v_j$  for all  $j \in \mathbf{Z}$ . The space  $\mathcal{O}_q$  is finite-dimensional and isomorphic to  $E^q$ . For any mapping  $\psi : E \rightarrow E$  we define the *lifting*  $\hat{\psi} : \mathcal{O}_q \rightarrow \mathcal{O}_q$  by setting  $(\hat{\psi}(v))_j := \psi(v_j)$  for all  $v \in \mathcal{O}_q$  and all  $j \in \mathbf{Z}$ . We also define the *shift operator*  $\sigma : \mathcal{O}_q \rightarrow \mathcal{O}_q$  by  $(\sigma(v))_j := v_{j+1}$  for all  $v \in \mathcal{O}_q$  and all  $j \in \mathbf{Z}$ . Our problem is then to find, for all small  $\mu \in \mathbf{R}^p$ , all small  $v \in \mathcal{O}_q$  which satisfy the equation

$$\Phi_\mu(v) := \hat{\varphi}_\mu(v) - \sigma(v) = 0.$$

It is important to observe that  $\sigma$  generates a cyclic  $\mathbf{Z}_q = \mathbf{Z}/q\mathbf{Z}$ -action on  $\mathcal{O}_q$  (since  $\sigma^q$  equals the identity), and that  $\Phi_\mu$  commutes with  $\sigma$ . We also have  $\Phi_\mu(0) = 0$  for all  $\mu$ .

The linearization of  $\Phi_\mu$  at the origin gives  $d\Phi_\mu(0) = \hat{L}_\mu - \sigma$ ; for  $\mu = 0$  this takes the form  $d\Phi_0(0) = \hat{L}_0 - \sigma = (\hat{S}_0 - \sigma) + \hat{N}_0$ , where  $L_0 = S_0 + N_0$  is the Jordan-Chevalley decomposition of  $L_0 = d\varphi_0(0)$ . It is easy to verify that  $\hat{S}_0 - \sigma$  is semi-simple, that  $\hat{N}_0$  is nilpotent, and that both commute; hence  $d\Phi_0(0) = (\hat{S}_0 - \sigma) + \hat{N}_0$  is the Jordan-Chevalley decomposition of  $d\Phi_0(0)$ . Also, we can decompose  $\mathcal{O}_q$  as  $\mathcal{O}_q = \text{Ker}(\hat{S}_0 - \sigma) + \text{Rg}(\hat{S}_0 - \sigma)$ ; this decomposition is invariant under  $\sigma$  and under  $\hat{L}_0 - \sigma$ , and the restriction of  $d\Phi_0(0) = \hat{L}_0 - \sigma$  to  $W := \text{Rg}(\hat{S}_0 - \sigma)$  is invertible. Moreover, the elements of  $\text{Ker}(\hat{S}_0 - \sigma)$  correspond to  $q$ -periodic orbits under  $S_0$ ; hence, if we set  $U_q := \text{Ker}(S_0^q - I) \subset E$  and define  $\zeta : U_q \rightarrow \mathcal{O}_q$  by  $(\zeta(u))_j := S_0^j(u)$  for all  $u \in U_q$  and for all  $j \in \mathbf{Z}$ , then  $\zeta$  is an isomorphism between  $U_q$  and  $\text{Ker}(\hat{S}_0 - \sigma)$ , and  $\zeta(S_0(u)) = \sigma(\zeta(u))$  for all  $u \in U_q$ . Every element  $v \in \mathcal{O}_q$  can be written in a unique way as  $v = \zeta(u) + w$ , with  $u \in U_q$  and  $w \in W$ ; in particular:

$$\hat{\varphi}_\mu(\zeta(u) + w) = \zeta(\psi_\mu(u, w)) + \Theta_\mu(u, w)$$

for some smooth mappings  $\psi_\mu : U_q \times W \rightarrow U_q$  and  $\Theta_\mu : U_q \times W \rightarrow W$ . The equation  $\Phi_\mu(v) = 0$  (with  $v = \zeta(u) + w$ ) decomposes as a system of two equations, namely

$$\psi_\mu(u, w) = S_0(u) \quad \text{and} \quad \Theta_\mu(u, w) - \sigma(w) = 0.$$

For  $(u, w, \mu) \in U_q \times W \times \mathbf{R}^p$  sufficiently small the second of these equations can be solved by the implicit function theorem for  $w = w_\mu^*(u)$ ; bringing this solution in the first equation we arrive at our final equation

$$\varphi_\mu^{(r)}(u) = S_0(u), \quad \text{with } \varphi_\mu^{(r)}(u) := \psi_\mu(u, w_\mu^*(u)).$$

This gives us the following.

**LEMMA.** *Let  $\varphi_\mu$  be a family of  $C^\infty$ -diffeomorphisms on  $E$ , depending smoothly on the parameter  $\mu \in \mathbf{R}^p$  and such that  $\varphi_\mu(0) = 0$  for all  $\mu$ . Let  $S_0$  be the semi-simple part of  $L_0 = d\varphi_0(0)$ . Fix some integer  $q \geq 1$  and define  $\mathcal{O}_q$ ,  $\hat{S}_0$  and  $\sigma$  as above. Let  $U_q := \text{Ker}(S_0^q - I) \subset E$  and  $W := \text{Rg}(\hat{S}_0 - \sigma) \subset \mathcal{O}_q$ ; define  $\zeta : U_q \rightarrow \text{Ker}(\hat{S}_0 - \sigma)$  by  $\zeta(u) := (S_0^j(u))_{j \in \mathbf{Z}}$ . Then there exist smooth mappings  $\varphi_\mu^{(r)} : U_q \rightarrow U_q$  and  $w_\mu^* : U_q \rightarrow W$  such that for all sufficiently small  $(\mu, v) \in \mathbf{R}^p \times \mathcal{O}_q$  the following holds:  $v$  is a  $q$ -periodic orbit of  $\varphi_\mu$  if and only if there exists some small  $u \in U_q$  such that  $v = \zeta(u) + w_\mu^*(u)$  and  $\varphi_\mu^{(r)}(u) = S_0(u)$ . Moreover*

$$\varphi_\mu^{(r)}(0) = 0, \quad d\varphi_0^{(r)}(0) = L_0|_{U_q} \quad \text{and} \quad \varphi_\mu^{(r)}(S_0(u)) = S_0(\varphi_\mu^{(r)}(u)),$$

while also

$$w_\mu^*(0) = 0, \quad dw_0^*(0) = 0 \quad \text{and} \quad w_\mu^*(S_0(u)) = \sigma(w_\mu^*(u)).$$

PROOF. The first part follows from the reduction above, the other properties of  $\varphi_\mu^{(r)}$  and  $w_\mu^*$  come from the uniqueness given by the implicit function theorem, from  $\varphi_\mu(0) = 0$ , from the  $\mathbf{Z}_q$ -equivariance of the problem, and by differentiating the identity  $\Phi_\mu(\zeta(u) + w_\mu^*(u)) = \zeta(\varphi_\mu^{(r)}(u) - S_0(u))$ .  $\square$

REMARK. When  $L_0$  restricted to  $U_q$  has some nilpotency (i.e.  $N_0|_{U_q} \neq 0$ ) then a further reduction is possible. Therefore the equation  $\varphi_\mu^{(r)}(u) = S_0(u)$  is, in general, not what in bifurcation theory is usually called the ‘bifurcation equation’.

Next we explore the relation between the mapping  $\varphi_\mu^{(r)}$  on  $U_q$  and the normal form of the original family  $\varphi_\mu$ . Suppose for a moment that  $\varphi_\mu$  commutes with  $S_0$ :  $\varphi_\mu(S_0(u)) = S_0(\varphi_\mu(u))$ . Then  $\hat{\varphi}_\mu(\zeta(u)) = \zeta(\varphi_\mu(u))$  for all  $u \in U_q$ ; it follows that  $\psi_\mu(u, 0) = \varphi_\mu(u)$  and  $\Theta_\mu(u, 0) = 0$ , which in turn implies that  $w_\mu^*(u) = 0$  and  $\varphi_\mu^{(r)}(u) = \varphi_\mu(u)$  for all sufficiently small  $(u, \mu) \in U_q \times \mathbf{R}^p$ . We conclude that in this case all small  $q$ -periodic orbits of  $\varphi_\mu$  have the form  $v = \zeta(u)$  for some  $u \in U_q$  satisfying the equation  $\varphi_\mu(u) = S_0(u)$ . Although the condition  $\varphi_\mu(S_0(u)) = S_0(\varphi_\mu(u))$  will in general not be satisfied, we observe that both the mapping  $\varphi_\mu^{(r)}$  and integrable approximations  $\varphi_\mu^{(i)}$  actually do satisfy this condition. For  $\varphi_\mu^{(r)}$  this means that the solutions of the equation  $\varphi_\mu^{(r)}(u) = S_0(u)$  which appears in the Lemma generate all small  $q$ -periodic orbits for the family  $\varphi_\mu^{(r)}$ . As for integrable approximations, a careful analysis of the reduction shows that if the family  $\varphi_\mu$  is in normal form with respect to  $S_0$  up to some finite order  $k \geq 1$  (by this we mean that  $[\varphi_\mu]_k$  commutes with  $S_0$ ), then  $[\varphi_\mu^{(r)}]_k = [\varphi_\mu]_k|_{U_q}$ . So we can approximate  $\varphi_\mu^{(r)}$  by bringing the original family  $\varphi_\mu$  in normal form. The following theorem summarizes our results.

THEOREM. Under the conditions of the previous Lemma there exists for all sufficiently small parameter values  $\mu \in \mathbf{R}^p$  a one-to-one relation between the small  $q$ -periodic orbits of  $\varphi_\mu$  and the small  $q$ -periodic orbits of the ‘reduced mapping’  $\varphi_\mu^{(r)}$  on  $U_q$ . This reduced mapping is equivariant with respect to the  $\mathbf{Z}_q$ -action generated by  $S_0$  on  $U_q$ , and its small  $q$ -periodic orbits are also  $\mathbf{Z}_q$ -orbits. Finally, if for some  $k \geq 1$  the truncated Taylor expansion  $[\varphi_\mu]_k$  commutes with  $S_0$ , then  $[\varphi_\mu^{(r)}]_k = [\varphi_\mu]_k|_{U_q}$ .

In order to find the actual  $q$ -periodic orbits we must of course further study the equation  $\varphi_\mu^{(r)}(u) = S_0(u)$  on  $U_q$ ; this ‘reduced phase space’  $U_q$  is the proper space corresponding to those proper values of  $S_0$  which are  $q$ -th roots of unity. Under generic assumptions on those proper values one can show (see e.g. [37]) that the  $\mathbf{Z}_q$ -equivariance of  $\varphi_\mu^{(r)}$  leads to so-called *resonance horns* or *Arnold tongues* in parameter space, consisting of those parameter values for which there exist (non-trivial)  $q$ -periodic orbits.

Similar reduction results can also be obtained for the bifurcation of periodic orbits from an equilibrium in families of vector fields (generalized Hopf bifurcation), although there the reduction technique is somewhat heavier since it requires the use of Banach spaces of periodic mappings. It is also possible to take special structures such as equivariance, symplecticity or reversibility into account. As an example we refer to [39] for the case of Hamiltonian vector fields.

## References

### Surveys in volume 1A

- [1] B. Hasselblatt, *Hyperbolic dynamical systems*, Handbook of Dynamical Systems, Vol. 1A, B. Hasselblatt and A. Katok, eds, North-Holland (2002), 249–319.

### Other sources

- [2] E. Borel, *Sur quelques points de la théorie des fonctions*, thèse, Ann. Sci. École Norm. Sup. **4** (1895), 9–35.
- [3] H.W. Broer, *Normal forms in perturbation theory*, Encyclopaedia of Complexity and System Science, R. Meyers, ed., Springer-Verlag (2009), 6310–6329.
- [4] H.W. Broer, *Formal normal form theorems for vector fields and some consequences for bifurcations in the volume preserving case*, Dynamical Systems and Turbulence, Warwick 1980, D.A. Rand and L.-S. Young, eds, Lecture Notes in Mathematics, Vol. 898, Springer-Verlag (1981), 54–74.
- [5] H.W. Broer and F. Takens, *Formally symmetric normal forms and genericity*, Dynamics Reported, Vol. 2, U. Kirchgraber and H.O. Walter, eds, John Wiley & Sons (1989), 39–59.
- [6] H.W. Broer and F. Takens, *Dynamical Systems and Chaos*, Epsilon Uitgaven, Vol. 64 (2009); Springer Appl. Math. Sci., Vol. 172 (2010).
- [7] R. Cushman and J. Sanders, *Nilpotent normal forms and representation theory of  $\mathfrak{sl}(2, \mathbf{R})$* , A.M.S. Contemp. Math. **56** (1986), 31–51.
- [8] C. Elphick, E. Tirapegui, M. Brachet, P. Coullet and G. Iooss, *A simple global characterization for normal forms of singular vector fields*, Physica **29D** (1987), 95–127.
- [9] D. Grobman, *Homeomorphisms of systems of differential equations*, Dokl. Akad. Nauk SSSR **129** (1959), 880–881.
- [10] J. Hadamard, *Sur l'itération et les solutions asymptotiques des équations différentielles*, Bull. Soc. Math. France **29** (1901), 224–228.
- [11] P. Hartman, *On the local linearization of differential equations*, Proc. Amer. Math. Soc. **14** (1963), 568–573.
- [12] M. Hirsch, C. Pugh and M. Shub, *Invariant Manifolds*, Lecture Notes in Mathematics, Vol. 583, Springer-Verlag (1977).
- [13] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag (1972).
- [14] A. Kelley, *The stable, center stable, center, center unstable, and unstable manifolds*, J. Differential Equations **3** (1967), 546–570.
- [15] Y.A. Kuznetsov, *Elements of applied bifurcation theory*, Applied Mathematics Series, Vol. 112, Springer-Verlag (1995).
- [16] J. Moser, *Stable and Random Motion in Dynamical Systems*, Princeton University Press (1973).
- [17] J.R. Munkres, *Elementary Differential Topology*, Princeton University Press (1963).
- [18] R. Narasimhan, *Analysis on Real and Complex Manifolds*, North-Holland Publishing (1968).
- [19] J. Palis, *On Morse Smale dynamical systems*, Topology **8** (1969), 385–405.
- [20] J. Palis and S. Smale, *Structural stability theorems*, Proc. Symp. Pure Math. **14** (1970).
- [21] J. Palis and F. Takens, *Topological equivalence of normally hyperbolic dynamical systems*, Topology **16** (1977), 335–345.
- [22] J. Palis and F. Takens, *Stability of parametrized families of gradient vector fields*, Ann. of Math. **118** (1983), 383–421.
- [23] J. Palis and F. Takens, *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations*, Camb. Univ. Press (1993).
- [24] O. Perron, *Über Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssystemen*, Math. Z. **29** (1928), 129–160.
- [25] O. Perron, *Über Stabilität und asymptotisches Verhalten der Lösungen eines Systems endlicher Differenzengleichungen*, J. Reine Angew. Math. **161** (1929), 41–64.
- [26] O. Perron, *Die Stabilitätsfrage bei Differentialgleichungen*, Math. Z. **32** (1930), 703–728.
- [27] C. Pugh and M. Shub, *Linearization of normally hyperbolic diffeomorphisms and flows*, Inv. Math. **10** (1970), 187–198.
- [28] C. Robinson, *Dynamical Systems*, CRC Press (1995).
- [29] S. Smale, *Differentiable dynamical systems*, Bull. A.M.S. **73** (1967), 747–817.



- [30] S. Sternberg, *Local contractions and a theorem of Poincaré*, Amer. J. Math. **79** (1957), 809–824.
- [31] S. Sternberg, *On the structure of local homeomorphisms of Euclidean  $n$ -space*, II, Amer. J. Math. **80** (1958), 623–631.
- [32] F. Takens, *Partially hyperbolic fixed points*, Topology **10** (1971), 133–147.
- [33] F. Takens, *Forced oscillations and bifurcations*, Applications of Global Analysis I, Comm. Math. Inst. Rijksuniversiteit Utrecht **3** (1974), 1–59.
- [34] F. Takens, *Singularities of vector fields*, Publ. Math. IHES **43** (1974), 47–100.
- [35] A. Vanderbauwhede, *Local Bifurcation and Symmetry*, Research Notes in Math., Vol. 75, Pitman (1982).
- [36] A. Vanderbauwhede, *Centre manifolds, normal forms and elementary bifurcations*, U. Kirchgraber and H.O. Walter, eds, Dynamics Reported, Vol. 2, John Wiley & Sons (1989), 86–169.
- [37] A. Vanderbauwhede, *Subharmonic bifurcation at multiple resonances*, Proceedings of the Mathematics Conference (Birzeit 1998), S. Elaydi, F. Allen, A. Ekhader, T. Mughrabi and M. Saleh, eds, World Scientific (2000), 254–276.
- [38] A. Vanderbauwhede, *Lyapunov-Schmidt method for dynamical systems*, Encyclopaedia of Complexity and System Science, R. Meyers, ed., Springer-Verlag (2009), 5299–5315.
- [39] A. Vanderbauwhede and J.-C. van der Meer, *A general reduction method for periodic solutions near equilibria in Hamiltonian systems*, Normal Forms and Homoclinic Chaos, W.F. Langford and W. Nagata, eds, A.M.S. Providence (1995), 273–294.
- [40] S. van Strien, *Centre manifolds are not  $C^\infty$* , Math. Z. **166** (1979), 143–145.

## CHAPTER 4

# Complex Exponential Dynamics

Robert L. Devaney

*Department of Mathematics & Statistics, Boston University, 111 Cummington St., Boston, MA 02215 USA,  
617-353-4560*

*E-mail: bob@bu.edu*

### Contents

|                                                  |     |
|--------------------------------------------------|-----|
| 1. Introduction                                  | 127 |
| 2. Basic notions                                 | 127 |
| 2.1. Preliminaries from dynamics                 | 127 |
| 2.2. Types of fixed or periodic points           | 129 |
| 2.3. Preliminaries from complex analysis         | 132 |
| 2.4. Role of the singular values                 | 134 |
| 2.5. The Julia set                               | 136 |
| 2.6. The Fatou set                               | 138 |
| 3. Quadratic dynamics                            | 140 |
| 3.1. The filled Julia set                        | 140 |
| 3.2. The fundamental dichotomy                   | 141 |
| 3.3. The Mandelbrot set                          | 141 |
| 3.4. External rays                               | 143 |
| 3.5. Some folk theorems                          | 144 |
| 3.6. Landing points of external rays             | 146 |
| 3.7. Rays landing on the $p/q$ bulb              | 148 |
| 3.8. The size of limbs and the Farey tree        | 151 |
| 3.9. Further remarks                             | 152 |
| 3.10. Appendix: The Farey tree                   | 153 |
| 3.11. Appendix: Angle doubling                   | 153 |
| 4. Exponential dynamics                          | 155 |
| 4.1. Computing the Julia set                     | 155 |
| 4.2. Explosions                                  | 156 |
| 4.3. Misiurewicz points                          | 157 |
| 4.4. Hyperbolic components of period 1–3         | 158 |
| 4.5. Hyperbolic components with higher periods   | 160 |
| 5. Cantor bouquets                               | 163 |
| 5.1. The idea of the construction                | 164 |
| 5.2. Cantor N-bouquets                           | 166 |
| 5.3. Straight brushes                            | 170 |
| 5.4. Connectedness properties of Cantor bouquets | 174 |
| 5.5. Uniformization of the attracting basin      | 175 |

|                                                            |     |
|------------------------------------------------------------|-----|
| 6. Indecomposable continua . . . . .                       | 176 |
| 6.1. Topological preliminaries . . . . .                   | 177 |
| 6.2. Construction of $\Lambda$ . . . . .                   | 178 |
| 6.3. Dynamics on $\Lambda$ . . . . .                       | 180 |
| 6.4. Final comments and questions . . . . .                | 182 |
| 7. The parameter plane . . . . .                           | 183 |
| 7.1. Structural instability . . . . .                      | 183 |
| 7.2. Other near-real parameters . . . . .                  | 187 |
| 7.3. Hairs in the parameter plane . . . . .                | 188 |
| 7.4. Questions and problems . . . . .                      | 189 |
| 8. Untangling hairs . . . . .                              | 190 |
| 8.1. The period doubling bifurcation . . . . .             | 191 |
| 8.2. Fingers . . . . .                                     | 193 |
| 8.3. The kneading sequence . . . . .                       | 195 |
| 8.4. Augmented itineraries . . . . .                       | 197 |
| 8.5. Untangling the hairs . . . . .                        | 199 |
| 8.6. Back to the parameter plane . . . . .                 | 202 |
| 9. Back to polynomials . . . . .                           | 203 |
| 9.1. The polynomial family . . . . .                       | 204 |
| 9.2. External rays . . . . .                               | 207 |
| 10. Other families of maps . . . . .                       | 208 |
| 10.1. Maps with polynomial Schwarzian derivative . . . . . | 209 |
| 10.2. The tangent family . . . . .                         | 213 |
| 10.3. Asymptotic values that are poles . . . . .           | 215 |
| 10.4. Bifurcation to an entire function . . . . .          | 218 |
| 10.5. Cantor bouquets and cantor sets . . . . .            | 219 |
| References . . . . .                                       | 221 |

## 1. Introduction

Our goal in this paper is give a somewhat quirky introduction to the field of complex dynamics. Usually, in the many introductions to this field that have appeared in the past twenty years or so, the main emphasis is on polynomial (in fact, quadratic polynomial) or rational dynamics. In this paper, we will take a slightly different tack; while we will of course discuss the polynomial case (and, yes, the Mandelbrot set), our main emphasis will be on entire functions and, most often, the complex exponential function.

There are several reasons for taking this approach. First, as mentioned above, there are many other papers and books where the emphasis is on polynomials [9,13,17,59,60,67]. Second, the dynamics of entire functions has experienced a tremendous surge of interest in recent years. And third, there are many very interesting connections between transcendental dynamics and planar topology. Indeed, one of our main subthemes in this paper is to highlight many of the very interesting types of sets from a topological point of view that arise as the Julia sets of entire functions.

In this paper, we assume that the reader is familiar with the basic ideas of discrete dynamical systems, general topology, and complex analysis. We will then use these fields as the foundation for our discussion of complex dynamics. This paper should be readable by people who have a background in discrete dynamics. Readers familiar with the basic ideas of complex dynamics dealing with polynomials or rational functions should also get something out of this paper, as we shall emphasize many of the topics that are different when entire functions are considered.

We will not provide all of the proofs: these can be found in the books cited above or in the research literature. We will, however, provide at least one proof in each section (or, at least, a somewhat detailed sketch), just to give the reader a sampling of how things are worked out in complex dynamics as well as the interplay between the background fields.

It is a pleasure to thank Keller Devaney, who digested this entire article and returned many highly stylized comments; all errors that remain are due to her.

## 2. Basic notions

Our goal in this section is to introduce some of the basic definitions and tools used in complex dynamics. These are drawn from the fields of discrete dynamical systems theory and complex analysis. In later sections we will specialize the discussion to entire transcendental functions.

### 2.1. Preliminaries from dynamics

Suppose that  $F : \mathbb{C} \rightarrow \mathbb{C}$  is complex analytic. As in discrete dynamics, we are interested in the dynamics of  $F$ , so we are concerned with the iteration of  $F$ . Given  $z_0 \in \mathbb{C}$ , the *orbit* of  $z_0$  is the sequence

$$z_0, F(z_0), F(F(z_0)), \dots$$

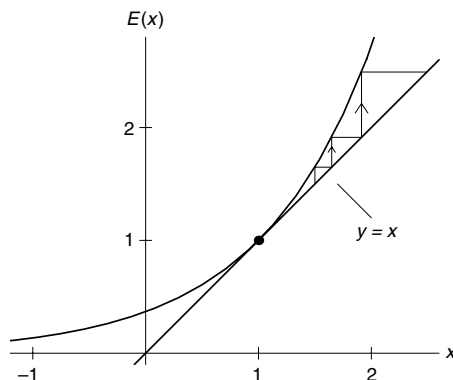


Fig. 1. The graph of  $E(x) = (1/e)e^x$ .

For simplicity, we write  $F^n$  for the  $n$ -fold composition of  $F$  with itself. Then the main question in dynamics is: can we predict the fate of orbits? That is, what happens to the sequence  $\{F^n(z_0)\}$  as  $n$  tends to  $\infty$ ?

There are many different types of orbit for a typical function. Perhaps the most important are the *fixed points*  $z_0$  for which  $F(z_0) = z_0$ . Next in importance are the *periodic points of period  $n$*  (also called cycles of period  $n$  or  $n$ -cycles): they are points that satisfy  $F^n(z_0) = z_0$ . The *period* of the cycle is the least  $n \geq 1$  for which  $F^n(z_0) = z_0$ . The point  $z_0$  is an *eventual fixed point* (or cycle) if  $z_0$  is not itself fixed (or periodic), but  $F^n(z_0)$  is fixed (or periodic) for some  $n \geq 1$ . Other important orbits are those that tend to a fixed point or periodic orbit, or those that tend to  $\infty$ .

EXAMPLE. Let  $D(z) = z^2$ . The orbits of 0 and 1 are fixed. The orbits of  $-1$  and  $\pm i$  are eventually fixed, since these orbits land on the fixed point at 1 after several iterations. If  $|z_0| > 1$ , then  $D^n(z_0) \rightarrow \infty$  as  $n \rightarrow \infty$ . If  $|z_0| < 1$  then  $D^n(z_0) \rightarrow 0$  as  $n \rightarrow \infty$ . The points  $z_1 = e^{2\pi i/3}$  and  $z_2 = e^{4\pi/3}$  lie on a cycle of period 2 for  $D$ . In fact,  $e^{2\pi i(p/q)}$  is periodic if both  $p$  and  $q$  are integers with  $q$  odd. If  $z_0$  lies on the unit circle, then the entire orbit of  $z_0$  remains on this circle.

EXAMPLE. Let  $E(z) = (1/e)e^z$ . We have  $E(1) = 1$  and  $E'(1) = 1$ . If  $x \in \mathbb{R}$  and  $x < 1$ , then  $E^n(x_0)$  tends to the fixed point at 1. If  $x_0 > 1$ , then  $E^n(x_0) \rightarrow \infty$  as  $n \rightarrow \infty$ . This can be shown using the web diagram as shown in Figure 1.

EXAMPLE. Let  $S(z) = \sin z$ . Then 0 is a fixed point for  $S$ . If  $x_0 \in \mathbb{R}$ , then either  $S(x_0) = 0$  (so the orbit is eventually fixed), or  $S^n(x_0) \rightarrow 0$ . On the other hand, points on the imaginary axis have orbits that tend to  $\infty$  since  $\sin(iy) = i \sinh(y)$ .

EXAMPLE. Let  $C(z) = \cos z$ . Then  $C$  has a real fixed point at  $x_0 = 0.7390\dots$  and the orbit of any  $x \in \mathbb{R}$  either lands on or tends to  $x_0$ .

## 2.2. Types of fixed or periodic points

In this section we will assume that  $z_0$  is a fixed point for  $F$ . If  $z_0$  lies on an  $n$ -cycle, then everything below goes through using  $F^n$  instead of  $F$ .

DEFINITION 2.1. The fixed point  $z_0$  is:

1. attracting if  $0 < |F'(z_0)| < 1$ ;
2. super-attracting if  $F'(z_0) = 0$ ;
3. repelling if  $|F'(z_0)| > 1$ ;
4. neutral or indifferent if  $F'(z_0) = e^{2\pi i\theta_0}$ . If  $\theta_0$  is rational, then  $z_0$  is rationally indifferent or parabolic, otherwise  $z_0$  is irrationally indifferent.

Note that there are no fixed points of saddle type for complex functions as we have just one ‘eigenvalue’ for the derivative, namely  $F'(z_0)$ , not two distinct eigenvalues.

The dynamical behaviour of  $F$  near attracting, repelling or super-attracting fixed points is completely understood, as we discuss below. In the neutral case, the behaviour near certain irrationally indifferent fixed points is still not completely understood.

Suppose  $F$  and  $G$  are two analytic functions. We say that  $F$  is (analytically) conjugate to  $G$  if there is an (analytic) homeomorphism  $h : \mathbb{C} \rightarrow \mathbb{C}$  such that  $h \circ F = G \circ h$ . We also define local conjugacy on subsets of  $\mathbb{C}$  in the natural manner.

The reason why analytic functions are easy to understand near attracting fixed points is given by the following theorem.

THEOREM 2.2 (Linearization Theorem). *Suppose  $z_0$  is an attracting fixed point for  $F$  and  $F'(z_0) = \lambda$  with  $0 < |\lambda| < 1$ . Then there is a neighbourhood  $U$  of  $z_0$  and an analytic map  $h : U \rightarrow \{z \mid |z| < 1\}$  such that  $h \circ F(z) = \lambda \cdot h(z)$ . That is,  $F$  is analytically conjugate to the linear map  $z \rightarrow \lambda z$  on  $U$ .*

In the repelling case, we use the Inverse Function Theorem to conjugate the branch of the inverse of  $F$  that fixes  $z_0$  to a linear map of the form  $z \rightarrow \lambda^{-1}z$  with  $|\lambda| > 1$ . Then  $F$  is locally conjugate to  $z \rightarrow \lambda z$ .

Finally, if  $z_0$  is super-attracting, then there is a neighbourhood  $U$  of  $z_0$  and  $n > 1$  such that  $F$  is analytically conjugate to  $z \rightarrow z^n$  on  $U$ , where  $h$  takes values in some disk  $\{z \mid |z| < r < 1\}$ .

As we mentioned above, the dynamics near neutral fixed points is extremely complicated, and we will not go into details here. But there are two cases that are completely understood. The first is the case of a rationally indifferent fixed point. In this case the Fatou Flower Theorem [60] asserts that, for some  $k > 0$ , there are  $k$  attracting and  $k$  repelling petals meeting at  $z_0$ . Roughly speaking, an attracting petal is an open set  $U$  bounded by a simple closed curve  $\gamma$  passing through  $z_0$  that has the property that

1.  $F(\overline{U}) \subset U \cup \{z_0\}$
2.  $F(\gamma) \cap \gamma = \{z_0\}$ .

It can be shown that all orbits within  $U$  tend to  $z_0$  under iteration.

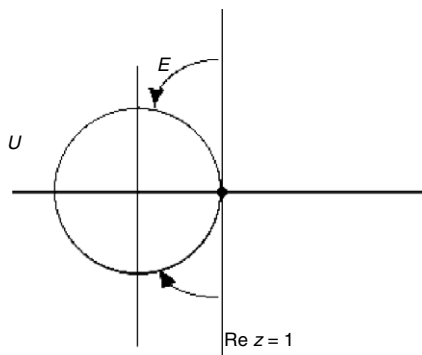


Fig. 2. The plane  $\operatorname{Re} z \leq 1$  is an attracting petal for  $E(z) = (1/e)e^z$ .

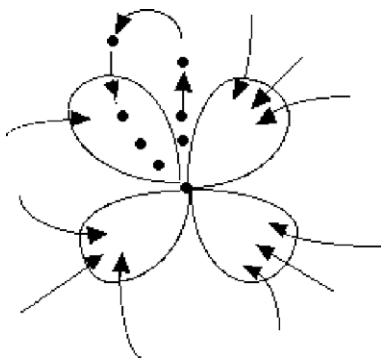


Fig. 3. Four Fatou flowers for a rationally indifferent fixed point.

EXAMPLE. Let  $E(z) = (1/e)e^z$ . Then  $z_0 = 1$  is a neutral fixed point with  $E'(z_0) = 1$ . Consider the vertical line segment  $\operatorname{Re} z = 1$ . Let  $U = \{z \mid \operatorname{Re} z < 1\}$ . Then  $U$  is an attracting petal for  $E$  (not worrying too much about the fact that the boundary of  $U$  is not a closed curve, though it is if we consider the Riemann sphere). Indeed,  $E$  maps  $U$  to the interior of the unit circle in  $\mathbb{C}$  and so  $E(\overline{U})$  meets  $\operatorname{Re} z = 1$  only at  $z_0 = 1$ . Thus it follows that all orbits of  $E$  in the half-plane  $\operatorname{Re} z < 1$  tend to 1 under iteration. See Figure 2.

A repelling petal is simply an attracting petal for  $F^{-1}$ , which exists locally near  $z_0$  since  $F'(z_0) \neq 0$ . In the above example, the preimage of the circle of radius 1 centred at  $x = 2$  on the real axis (under the branch of the inverse of  $E$  that fixes  $z_0 = 1$ ) and its interior form a repelling petal. Note that any orbit in the repelling petal must eventually leave the petal (except, of course, the neutral fixed point). See Figure 3.

EXAMPLE. Let  $S(z) = \sin z$ . Then 0 is a rationally indifferent fixed point. There are now two attracting petals, both straddling the real axis. Orbits on the imaginary axis leave a neighbourhood of 0. Indeed, on the imaginary axis, as we saw earlier, we have  $\sin(iy) = i \sinh(y)$ , so orbits tend to  $\infty$  along this axis. See Figure 4.

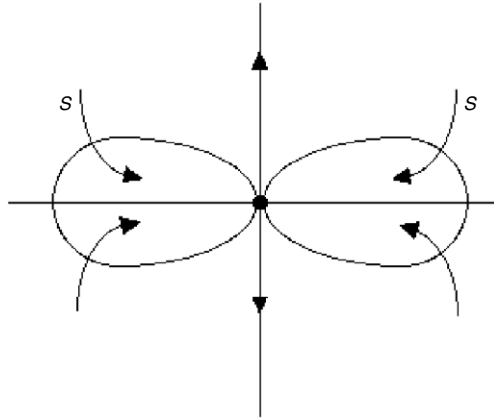


Fig. 4. There are two attracting petals  $S$  for  $\sin z$  and two repelling directions.

Now suppose that  $z_0$  is an irrationally indifferent fixed point with  $F'(z_0) = e^{2\pi i\theta_0}$ . We will deal only with the case that  $\theta_0$  is a Brjuno number. To define these irrationals, we make a brief digression to consider continued fractions.

Since  $\theta_0$  is irrational, it has an infinite continued fraction expansion of the form

$$\theta_0 = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}$$

where each  $a_j$  is a positive integer (and  $a_0$  may be 0). The convergents of  $\theta_0$  are defined to be the rational numbers

$$\frac{p_n}{q_n} = a_0 + \frac{1}{a_1 + \frac{1}{\ddots + a_n}}$$

It is known that the convergents of  $\theta_0$  are the closest rational approximants of  $\theta_0$  (in the sense that there are no rationals  $p/q$  between  $\theta_0$  and  $p_n/q_n$  with  $q < q_n$ ).

The rational number  $\theta_0$  is said to be a *Brjuno number* if the series

$$\sum_{n=1}^{\infty} \frac{\log(q_{n+1})}{q_n}$$

converges. For example, all diophantine numbers are Brjuno, so the Brjunos are dense in the reals. Recall that  $\theta$  is diophantine if  $\theta$  is 'far' from rationals in the sense that there are constants  $c > 0$ ,  $k \geq 2$  such that

$$\left| \theta - \frac{p}{q} \right| > \frac{c}{q^k}$$



for all rationals  $p/q$ . Unfortunately for dynamical systems, the complement of the Brjunos in the irrationals is also dense. The Brjuno numbers form a set of full measure however.

The natural question is when is a nonlinear function conjugate to an irrational rotation? As stated, this question clearly cannot depend on the rotation angle  $\theta_0$ , as we can always conjugate the irrational rotation  $z \rightarrow e^{2\pi i\theta_0}z$  by an analytic function to obtain a nonlinear function that acts like an irrational rotation. So the question is: are there any irrationals  $\theta_0$  for which every nonlinear function of the form

$$z \rightarrow e^{2\pi i\theta_0}z + a_2z^2 + \dots$$

for a given  $\theta$  is conjugate to an irrational rotation? The following theorem answers this question.

**THEOREM 2.3 (Yoccoz' Theorem).** *The function  $z \rightarrow e^{2\pi i\theta}z + a_2z^2 + \dots$  is locally conjugate to  $z \rightarrow e^{2\pi i\theta}z$  (independent of any higher order terms) if  $\theta$  is a Brjuno number. If  $\theta$  is not a Brjuno number, then the quadratic polynomial*

$$z \mapsto e^{2\pi i\theta}z + z^2$$

*has a fixed point at the origin which is not linearizable.*

Suppose  $F'(z_0) = e^{2\pi i\theta_0}$  where  $\theta_0$  is Brjuno. By Yoccoz' Theorem, there is a neighbourhood  $U$  of  $z_0$  such that each orbit in  $U - \{z_0\}$  is dense on an invariant simple closed curve surrounding  $z_0$ . The maximal simply connected open set about  $z_0$  with this property is called a Siegel disk after C.L. Siegel who first proved that maps of the form

$$z \rightarrow e^{2\pi i\theta}z + a_2z^2 + \dots$$

are linearizable if  $\theta$  is diophantine.

### 2.3. Preliminaries from complex analysis

The main reason that one dimensional complex analytic dynamics is so special is that all of the tools of complex analysis are available. In this section we review some of the most basic results from this field that we will use over and over again.

One of the most important tools is the:

**THEOREM 2.4 (Riemann Mapping Theorem).** *Let  $U$  be an open, simply connected subset of  $\mathbb{C}$  and suppose that  $U$  is not equal to  $\mathbb{C}$  itself. Let  $z_0 \in U$ . Then there is an analytic map  $\phi$  taking  $U$  onto the unit disk  $\mathbb{D}$  in one-to-one fashion and satisfying  $\phi(z_0) = 0$ . Moreover, if we normalize so that  $\phi'(z_0) > 0$ , then the map  $\phi$  is unique.*

The importance of this result is that, whenever we encounter some dynamical behaviour on an open simply connected set, we may transfer the dynamics to the unit disk  $\mathbb{D}$  via the Riemann map  $\phi$ . Many of the important theorems in complex analysis are stated in terms of maps of the disk. A principal example of this is the:

**THEOREM 2.5** (Schwarz Lemma). *Suppose  $F : \mathbb{D} \rightarrow \mathbb{D}$  is analytic and  $F(0) = 0$ . Then  $|F'(0)| \leq 1$  and, moreover, either  $|F(z)| < |z|$  for all  $z \neq 0$ , or else  $F(z) = e^{i\theta} \cdot z$  for some  $\theta \in \mathbb{R}$ .*

Note that, if  $F$  is not a rotation, the condition  $|F(z)| < |z|$  guarantees that all points in  $\mathbb{D}$  move closer to 0 at each application of  $F$  and so all orbits in  $\mathbb{D}$  tend to 0 under iteration of  $F$ . Therefore the importance of this result is the fact that either  $F : \mathbb{D} \rightarrow \mathbb{D}$  is a rotation, or else all orbits in  $\mathbb{D}$  tend to the fixed point at 0.

We will often invoke the following modification of this result.

**COROLLARY 2.6.** *Suppose  $U$  is a simply connected open set not equal to  $\mathbb{C}$  itself. Suppose further that  $F(\overline{U}) \subset U$ . Then  $F$  has a fixed point in  $U$  and, moreover, all orbits of  $F$  tend to this fixed point.*

This is true since there is a fixed point in  $U$  by the Brouwer Fixed Point Theorem. Then we can conjugate  $F$  to a map of the disk via the Riemann map  $\phi$ , and we may assume that the fixed point is sent to 0. Since the map  $\phi \circ F \circ \phi^{-1}$  is analytic on  $\mathbb{D}$  and fixes 0, the Schwarz Lemma applies. Since  $F(\overline{U}) \subset U$ , it follows that  $\phi \circ F \circ \phi^{-1}$  cannot be a rotation, and so  $F$  must have a globally attracting fixed point.

This is one of the principal differences between real and complex dynamics. In real dynamics we can have all sorts of dynamics inside a simply connected region. For example, the classical horseshoe map may be defined inside a simply connected region. For complex maps however, this is never the case. We may only have a globally attracting fixed point or a rotation domain in such a region. This is one of the reasons that complex dynamical systems are so well understood.

There is another way to view this. On  $\mathbb{D}$ , there is a very special metric called the *Poincaré metric*. If  $F : \mathbb{D} \rightarrow \mathbb{D}$  is analytic, then it turns out that either  $F$  is a strict contraction in the Poincaré metric, or else  $F$  is an isometry. This is often referred to as the Schwarz-Pick Lemma. Thus, by the Riemann Mapping Theorem, we can transfer the Poincaré metric to any simply connected domain in (but not equal to all of)  $\mathbb{C}$ .

Perhaps the most important types of orbits are those that lie in the *Julia set* which is named for the French mathematician Gaston Julia who first studied these orbits around 1918. To define the Julia set, we need to make a digression into the theory of normal families of functions. Let  $\{G_i\}$  be a family of analytic functions in  $\mathbb{C}$ . Often, but not always, the  $G_i$  will be the iterates of a given function.

**DEFINITION 2.7.** The family of functions  $\{G_i\}$  is a normal family on an open set  $U \subset \mathbb{C}$  if every sequence of the  $G_i$ 's has a subsequence that either

1. converges uniformly on compact subsets of  $U$ , or
2. converges uniformly to  $\infty$  on compact subsets of  $U$ .

Recall that, in case 1, the subsequence converges to a function on  $U$  that necessarily is analytic. By the Arzela-Ascoli theorem, the  $G_i$  form a normal family on  $U$  if the  $G_i$  are

equicontinuous (in the spherical metric) on compact subsets of  $U$ . Another main tool from complex analysis is:

**THEOREM 2.8** (Montel's Theorem). *Suppose the  $\{G_i\}$  are a family of analytic functions on  $\mathbb{C}$  and that there are two distinct values  $z_1$  and  $z_2$  in  $\mathbb{C}$  that are never assumed by the  $G_i$ . Then  $\{G_i\}$  is a normal family of functions.*

**EXAMPLE.** Let  $D(z) = z^2$ . Then the family of iterates of  $D$ ,  $\{D^n\}$  is

1. normal on any open subset of  $\{z \mid |z| < 1\}$ ;
2. normal on any open subset of  $\{z \mid |z| > 1\}$ ;
3. not normal on any open set that intersects the unit circle.

Indeed, in case 1, the  $D^n$  converge uniformly on compact subsets to 0. In case 2, the  $D^n$  converge uniformly on compact subsets to  $\infty$ . But in case 3, there are open sets in  $U$  on which the  $D^n$  converge either to 0 or to  $\infty$ , and hence this sequence does not converge to an analytic function.

**DEFINITION 2.9.** A function  $E : \mathbb{C} \rightarrow \mathbb{C}$  is entire if  $E(z) \neq \infty$  for any  $z \in \mathbb{C}$  and  $E$  is not a polynomial. Sometimes such functions are called entire transcendental functions.

Our main goal in this paper is to study the dynamics of entire functions, but we will occasionally deal with polynomials in order to contrast these two cases. With the exception of Section 10, we will not treat meromorphic functions, i.e. functions with poles.

One of the main differences between polynomials and entire functions is the behaviour at  $\infty$ . For a polynomial  $P$ , we may extend  $P$  to the entire Riemann sphere by defining  $P(\infty) = \infty$ . We have  $P'(\infty) = 0$ , so that  $\infty$  is a super-attracting fixed point. This can be seen by conjugating  $P$  via  $\phi(z) = 1/z$ . The fixed point at  $\infty$  is sent to 0 and then one calculates that  $(\phi \circ P \circ \phi^{-1})'(0) = 0$ .

For an entire function  $E$ , we cannot extend  $E$  to  $\infty$  continuously, never mind analytically. Indeed, the behaviour of  $E$  near  $\infty$  is quite complicated. This is illustrated by the:

**THEOREM 2.10** (Great Picard Theorem.). *In any neighbourhood of  $\infty$  an entire function assumes all values in  $\mathbb{C}$  infinitely often with the possible exception of one value.*

**EXAMPLE.** Let  $E(z) = e^z$ . A neighbourhood of  $\infty$  contains a region of the form  $\{z \mid |z| > r\}$  for some  $r$ . This region contains infinitely many strips of the form  $k\pi \leq \text{Im } z < (k+2)\pi$ , and each of these strips is mapped in one-to-one fashion onto  $\mathbb{C} - \{0\}$ .

For the record,  $\infty$  is called an *essential singularity* of the entire function.

## 2.4. Role of the singular values

For entire functions, there are two types of *singular values* that play an important role in determining the dynamics. These are the critical and asymptotic values. A *critical value* is an image of a critical point, i.e.  $F(z_0)$  where  $F'(z_0) = 0$ .

**DEFINITION 2.11.** The point  $z_0 \in \mathbb{C}$  is an asymptotic value of  $E$  if there is a curve  $\gamma(t)$  satisfying  $\lim_{t \rightarrow \infty} \gamma(t) = \infty$  and  $\lim_{t \rightarrow \infty} E(\gamma(t)) = z_0$ .

EXAMPLE. The omitted value 0 is an asymptotic value for  $E_\lambda(z) = \lambda e^z$  since any curve  $\gamma(t)$  that satisfies  $\lim_{t \rightarrow \infty} \operatorname{Re}(\gamma(t)) = -\infty$  also satisfies  $\lim_{t \rightarrow \infty} E(\gamma(t)) = 0$ . Also, the point at  $\infty$  is an asymptotic value.

One of the main reasons that singular values are important is the fact that any attracting fixed point must have a singular value in its immediate basin of attraction; indeed, this is one of the major features that distinguishes complex dynamics from other branches of discrete dynamical systems. The *basin of attraction* of  $z_0$  is the set of all points whose orbit tends to  $z_0$ . The *immediate* basin of attraction of  $z_0$  is the component of the basin that contains  $z_0$ . Basins of attraction for attracting  $n$ -cycles are defined using  $F^n$  instead of  $F$ .

**THEOREM 2.12** (Role of the Singular Values Theorem). *Suppose the analytic function  $F$  has an attracting fixed point or cycle. Then there is at least one singular value in the immediate basin of attraction of this point.*

PROOF. To see why the immediate basin of the fixed point  $z_0$  contains a singular value, we argue by contradiction. We may assume that  $F'(z_0) \neq 0$  for otherwise we are done. Let  $U_0$  be a bounded neighbourhood of  $z_0$  contained in the immediate basin of  $z_0$  and satisfying  $F(U_0) \subset U_0$ . We can find such a  $U_0$  using the linearization result for attracting fixed points.

We may assume also that  $F$  is one-to-one on  $U_0$ . Now we pull back by a branch of  $F^{-1}$ : let  $U_1$  be the preimage of  $U_0$  that contains  $U_0$ . Now if  $U_1$  is unbounded, we may find a curve  $\gamma(t) \subset U_1$  with  $\gamma(t) \rightarrow \infty$  as  $t \rightarrow \infty$  and  $F(\gamma(t))$  approaching a limit in  $U_0$ . This yields an asymptotic value in  $U_0$ , so it follows that  $U_1$  must be bounded. Also,  $F|_{U_1}$  must be one-to-one for otherwise there would be a critical point in  $U_1$  and therefore a critical value in  $U_0$ .

Continuing in this fashion we construct  $U_{n+1} \supset U_n$  for each  $n \geq 1$  with  $U_{n+1}$  bounded and  $F : U_{n+1} \rightarrow U_n$ . Let  $W$  be the union of the  $U_n$ .  $W$  is an open set in  $\mathbb{C}$  and we have  $F : W \rightarrow W$  is one-to-one. Hence  $F^{-1} : W \rightarrow W$  is well-defined and analytic.

Now  $W \neq \mathbb{C}$  for otherwise  $F$  would be a Möbius transformation and hence not entire. Also, there are (many more than) two points not in  $W$ , since the other preimages of  $U_0$  do not lie in  $W$ . As a consequence, the family of functions  $\{F^{-n}\}$  is a normal family on  $W$ . Therefore the  $F^{-n}$  has a subsequence that converges to an analytic function on  $W$ . But each  $F^{-n}$  fixes  $z_0$  whereas  $F^{-n}(z)$  tends to the boundary of  $W$  as  $n \rightarrow \infty$ . As a consequence, the limit function is not even continuous. This contradiction establishes the result.  $\square$

**COROLLARY 2.13.** *Suppose  $F$  has at most  $n$  singular values. Then  $F$  can have at most  $n$  attracting cycles.*

EXAMPLE. The exponential function  $E_\lambda(z) = \lambda e^z$  has no critical values and only one asymptotic value, 0. Hence  $E_\lambda$  can have at most one attracting cycle. If  $0 < \lambda < 1/e$ , the graph of  $E_\lambda$  lies below that of  $E_{1/e}$ , and so  $E_\lambda$  has a real attracting fixed point. Figure 5 shows that 0 is attracted to this attracting fixed point.

EXAMPLE. The sine family  $S_\lambda(z) = \lambda \sin z$  has no asymptotic values and infinitely many critical points but only 2 critical values,  $\pm\lambda$ . When  $|\lambda| < 1$  the origin is an attracting fixed

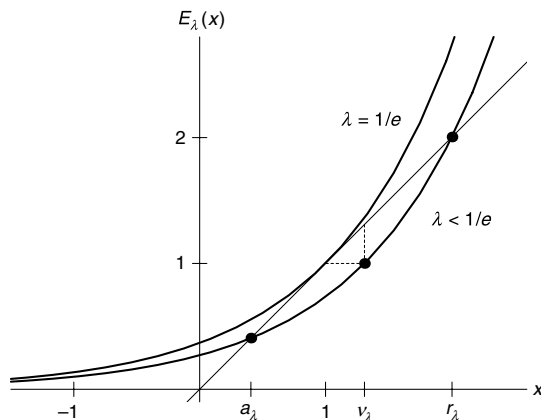


Fig. 5. The graphs of  $E_\lambda$  for  $\lambda = 1/e$  and  $\lambda < 1/e$ .

point which must attract one critical value. Since  $S_\lambda$  is odd, in fact 0 attracts both critical values, and so for these  $\lambda$ -values, 0 is the only attracting cycle.

## 2.5. The Julia set

The most important orbits in a complex dynamical system lie in the *Julia set*, named for Gaston Julia, who initiated the study of these sets in 1918.

**DEFINITION 2.14.** Let  $F : \mathbb{C} \rightarrow \mathbb{C}$  be complex analytic. The Julia set of  $F$ , denoted  $J(F)$ , is the set of points at which the family of iterates  $\{F^n\}$  fails to be a normal family.

Points in  $J(F)$  have orbits that are sensitive to initial conditions. If  $z \in J(F)$  and  $U$  is any neighbourhood of  $z$ , then by Montel's Theorem,  $\cup F^n(U)$  contains all points in  $\mathbb{C}$ , with at most one exception. We think of this as meaning that the Julia set is the chaotic regime for  $F$ , because orbits here are extremely sensitive to initial conditions.

**EXAMPLE.** For  $D(z) = z^2$ ,  $J(D)$  is the unit circle. If  $U$  is a neighbourhood of  $z$  with  $|z| = 1$ , then we may always find a smaller neighbourhood  $U'$  of  $U$  satisfying

$$U' = \{z \mid \theta_1 < \text{Arg } z < \theta_2, 0 < r_1 < |z| < r_2\}$$

It is then easy to see that  $\bigcup D^n(U')$  covers  $\mathbb{C} - \{0\}$ .

Note that repelling cycles always lie in the Julia set. To see this, suppose that  $F^n(z_0) = z_0$  and  $|(F^n)'(z_0)| = \lambda > 1$ . Then the sequence of analytic functions  $F^{kn}$  cannot converge to  $\infty$  in any neighbourhood of  $z_0$  since each  $F^{kn}$  fixes  $z_0$ . Also,  $F^{kn}$  cannot converge to an analytic function in any neighbourhood of  $z_0$  since  $|(F^{kn})'(z_0)| = \lambda^{kn} \rightarrow \infty$  as  $k \rightarrow \infty$ . So any limit function would be nondifferentiable at  $z_0$ .

Rationally indifferent periodic points also lie in the Julia set. Inside the attracting petals, iterates of the map tend toward the cycle, but in the repelling petals, orbits move

away. Hence we cannot have uniform convergence of iterates in any neighbourhood of the periodic point.

EXAMPLE. The fixed point  $z_0 = 1$  for  $E(z) = (1/e)e^z$  is neutral with  $E'(z_0) = 1$ . The web diagram shows that if  $x < 1$ , the orbit of  $x$  tends to 1, but if  $x > 1$ , the orbit tends to  $\infty$ . Hence  $1 \in J(E)$ . See Figure 1.

EXAMPLE. The fixed point 0 for  $S(z) = \sin z$  also satisfies  $S'(0) = 1$ . Orbits on the real axis tend to 0 (or are eventually fixed at 0), but nonzero orbits on the imaginary axis tend to  $\infty$ . Again, we cannot have uniform convergence in any neighbourhood of 0.

THEOREM 2.15 (Properties of the Julia Set). *The Julia set is*

1. *closed*;
2. *nonempty*;
3. *forward invariant (i.e. if  $z \in J(F)$ , then  $F(z) \in J(F)$ )*;
4. *backward invariant*;
5. *equal to the closure of the set of repelling cycles of  $F$ .*

The fact that  $J(F)$  is nonempty is surprisingly not so easy to show, at least as far as I am aware. Properties 1, 3, and 4 are straightforward exercises. Property 5 gives a more dynamical definition of the Julia set, so let's prove this part. We will prove this under the simplifying assumption that  $F$  has a repelling periodic point. Every example considered in this paper has such a point, and indeed every analytic function has infinitely many such points.

PROOF. By definition, the family of functions  $\{F^n\}$  is not normal at any point in  $J(F)$ . So we will show that there is a repelling periodic point in any neighbourhood of a point where  $\{F^n\}$  fails to be normal. Toward that end, suppose  $\{F^n\}$  is not normal at  $p$  and let  $W$  be a neighbourhood of  $p$ . We will produce a repelling periodic point in  $W$ .

Under our assumption, there exists a repelling periodic point  $z_0$  somewhere. We may assume that  $z_0$  is a fixed point for  $F$ . Using the linearization around this fixed point, there is a neighbourhood  $U_0$  of  $z_0$  such that  $F : U_0 \rightarrow \mathbb{C}$  is a diffeomorphism onto its image which we may assume contains  $U_0$ . Hence  $F^{-1}$  is well-defined on  $U_0$  and maps  $U_0$  inside itself. Let  $U_i = F^{-i}(U_0)$  and note that  $U_{i+1} \subset U_i$  and  $\cap U_i = \{z_0\}$ .

Since  $\{F^n\}$  is not normal at  $p$ , there is a point  $z_1 \in W$  and an integer  $n$  such that  $F^n(z_1) = z_0$ <sup>1</sup>. Similarly, since  $\{F^n\}$  is not normal at  $z_0$ , there is a point  $z_2 \in U_0$  and an integer  $m$  such that  $F^m(z_2) = z_1$ . Hence  $F^{m+n}(z_2) = z_0$ .

We now make the simplifying assumption that  $(F^{m+n})'(z_2) \neq 0$ . If  $z_2$  is a critical point for  $F^{m+n}$ , then some modifications to the following argument are necessary. We leave these details to the reader. Since  $(F^{m+n})'(z_2) \neq 0$ , there is a neighbourhood  $V$  of  $z_2$  which is contained in  $U_0$  and that is mapped diffeomorphically onto a neighbourhood of  $z_0$  by  $F^{m+n}$ . By adjusting  $V$ , we may assume that  $F^m(V) \subset W$  and that  $F^{m+n}$  maps

---

<sup>1</sup> $z_0$  cannot be the 'exceptional point' (that is, the point not hit by  $F$ ) since it is easy to see that such a point has no preimages and so is a super-attracting fixed point.

$V$  diffeomorphically onto  $U_j$  for some  $j$ . It follows that  $F^{m+n+j}$  is a diffeomorphism mapping  $V$  onto  $U_0$ . Consequently, this map has an inverse which contracts  $U_0$  onto  $V$ . There is a fixed point for  $F^{m+n+j}$  in  $V$  which, by the Schwarz lemma, must be repelling. The orbit of this repelling periodic point enters  $W$ , since  $F^m(V) \subset W$ . This completes the proof.  $\square$

One of the most interesting topological facts about Julia sets is the following:

**PROPOSITION 2.16.** *Either the Julia set of  $F$  is nowhere dense in the plane, or else  $J(F) = \mathbb{C}$ .*

Indeed, if  $J(F)$  contains an open set  $U$ , then  $\cup F^n(U)$  covers the whole plane (except at most one point) and, by forward invariance, this set is contained in  $J$ . We can exclude that nasty ‘except one point’ too since  $J(F)$  is closed.

As a remark, if  $F$  is a polynomial, then  $J(F)$  can never be the entire plane since all orbits in a neighbourhood of  $\infty$  tend to  $\infty$  (so  $\{F^n\}$  is normal there). But we will see that there are many examples of entire functions for which  $J = \mathbb{C}$ . Indeed, one of the most interesting aspects of entire dynamics is the abrupt transition from nowhere dense Julia sets to Julia sets that are the whole plane.

## 2.6. The Fatou set

The complement of the Julia set is called the *Fatou set* (or *stable set*). Attracting cycles and their basins of attraction always lie in the Fatou set since iterates here tend to the cycle and thus form a normal family.

If  $z_0$  is a fixed point (or cycle) for which we have a Siegel disk, then  $z_0$  (and any point in the rotation domain) also lies in the Fatou set. To see this, recall that  $F^n$  is conjugate in a neighbourhood  $U$  of  $z_0$  to  $z \rightarrow e^{2\pi i n \theta_0}$  for some irrational  $\theta_0$ . A subsequence  $F^{n_i}$  is then conjugate in  $U$  to irrational rotation by  $2\pi n_i \theta_0$ . We think of  $2\pi n_i \theta_0$  as a sequence of points on the unit circle. As such, this sequence must have a limit point. Then the corresponding  $F^n$  tend to a map that is conjugate to rotation by this limiting angle.

Another type of region that lies in the Fatou set is a wandering domain. This set is a component of the Fatou set that is never periodic or eventually periodic, that is, this domain wanders forever.

**EXAMPLE.** Let  $G_\lambda(z) = z + \lambda \sin z$  where  $\lambda$  is chosen as follows. The real graph of  $G_\lambda(x)$  has infinitely many critical values, and we may choose  $\lambda$  so that there is an orbit that consists entirely of critical points and tends to  $\infty$  as shown in Figure 6. If  $x_0$  lies on this orbit of critical points, then we may choose a disk  $U$  about  $x_0$  so that  $G_\lambda(U) \subset U + 2\pi$ . It follows that the orbit of any point in  $U$  tends to  $\infty$ . Thus  $U$  is a wandering domain provided  $G_\lambda^n(U)$  lies in a different component of the Julia set for each  $n$ . This follows from the following exercise.

**EXERCISE.** The vertical lines  $\operatorname{Re} z = 2k\pi$  for  $z \in \mathbb{Z}$  lie in the Julia set of  $G_\lambda$ .

Another type or region in the Fatou set is a *Baker domain* or *domain at  $\infty$* . This is a set that is forward invariant and in which all orbits tend to  $\infty$ .

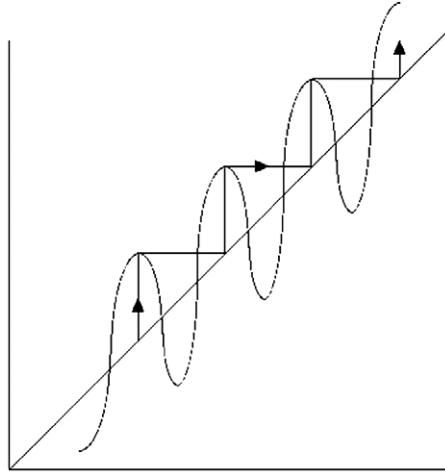


Fig. 6. The graph of  $G_\lambda(z) = z + \lambda \sin z$ .

EXAMPLE. Let  $H(z) = z + 2 + e^{-z}$ . Then any point in the half-plane  $\operatorname{Re} z > 0$  has orbit tending to  $\infty$ . This is true since  $z + 2$  moves  $z$  two units to the right, but adding  $e^{-z}$  moves  $z + 2$  at most one unit in any direction. Hence  $\operatorname{Re} H(z) > \operatorname{Re} z + 1$  for  $\operatorname{Re} z > 0$ . Therefore  $\operatorname{Re} z > 0$  is a domain at  $\infty$ .

Note that each example of a wandering domain and Baker domain features a function with infinitely many singular values.

Finally, basins of attraction of parabolic periodic points also lie in the Fatou set. These basins consist of all points whose orbits eventually enter an attracting petal.

EXAMPLE. For  $E(z) = (1/e)e^z$ , we have seen that the half-plane  $\operatorname{Re} z < 1$  is mapped inside itself and that 1 is a neutral fixed point. Consequently, this region (and all its preimages) lies in the basin of attraction of 1.

We now describe one of the fundamental theorems in complex dynamics, the ‘No Wandering Domains’ Theorem, due to Dennis Sullivan [68]. The version we will use is an extension to the entire case due independently to Goldberg and Keen [44] and Eremenko and Lyubich [39]. This theorem really should be called the:

**THEOREM 2.17** (Classification of Stable Domains). *Suppose  $E$  is an entire function that has only finitely many singular values. Then every component of the Fatou set is eventually periodic. Moreover, if  $U$  is a component of the Fatou set that is periodic, then  $U$  is one of the following:*

1. *the basin of attraction of an attracting or super-attracting cycle;*
2. *the basin of attraction of a rationally indifferent cycle;*
3. *a Siegel disk.*



And that's it! There are no wandering domains, domains at infinity, or any other type of Fatou set for entire functions with finitely many singular values. Clearly, these kinds of maps deserve a name.

**DEFINITION 2.18.** An entire function that has only finitely many singular values is called critically finite.

Luckily, everyone's favourite entire functions are critically finite. This includes the exponential ( $\lambda e^z$ ), sine ( $\lambda \sin z$ ), and cosine ( $\lambda \cos z$ ) families, where  $\lambda \in \mathbb{C}$ .

### 3. Quadratic dynamics

In this section we present a review of some well known (and a few not-so-well-known) results involving the family of quadratic polynomials  $Q_c(z) = z^2 + c$ . Our primary goal is to introduce the external rays in both the dynamical and parameter planes and to show how these can be used to understand some of the complicated structures that live in these planes. We will see the 'remnants' of these rays, the so-called hairs, when we look at Julia sets and parameter planes for entire functions [35,36].

#### 3.1. The filled Julia set

For the quadratic family, the only singular value is the critical value  $c = Q_c(0)$  since 0 is the only critical point. Polynomials never have (finite) asymptotic values.

**DEFINITION 3.1.** The filled Julia set of  $Q_c$ , denoted  $K_c$ , is the set of points whose orbits are bounded.

It is easy to check that the boundary of  $K_c$  in  $\mathbb{C}$  is the Julia set  $J(Q_c)$ . Furthermore,  $\infty$  is a super-attracting fixed point for  $Q_c$ , so  $K_c$  is compact and we can use the linearization theorem to show that  $Q_c$  is conjugate to  $Q_0(z) = z^2$  in a neighbourhood of  $\infty$ . That is, we can find a closed disk  $U$  about  $\infty$  in  $\overline{\mathbb{C}}$  and an  $r > 1$  together with an analytic homeomorphism

$$\phi_c : U \rightarrow \{z \mid |z| \geq r\}$$

such that  $\phi_c \circ Q_c = (\phi_c(z))^2$ .

Now suppose that the orbit of 0 under  $Q_c$  is bounded. Then we can extend the conjugacy  $\phi_c$  to the entire exterior of  $K_c$ . Here is how to do this. Each point in  $U$  has exactly two preimages in  $Q_c^{-1}(U)$ . Similarly, each point in  $\{z \mid |z| > r\}$  has exactly two preimages under  $z^2$  (in  $\{z \mid |z| \geq \sqrt{r}\}$ ). Using continuity, we can extend  $\phi_c$  analytically to  $Q_c^{-1}(U)$  in a unique fashion. The only impediment to this would be if  $c \in U$ . Then there would be only one preimage of the critical value  $c$  and therefore no way to extend  $\phi_c$  as a function.

So, as long as  $c$  lies in the filled Julia set we can continue this procedure infinitely often and thus we have a conjugacy with  $z^2$  defined on the exterior of  $K_c$  and taking values in  $\{z \mid |z| > 1\}$ .

### 3.2. The fundamental dichotomy

These facts allow us to prove the following important result for quadratic maps.

**THEOREM 3.2 (The Fundamental Dichotomy).** *Suppose  $0 \in K_c$ . Then  $K_c$  is connected. If  $0 \notin K_c$ , then  $K_c$  is a Cantor set.*

For the connectedness portion of this result, we have that, for each  $n$ ,  $Q_c^{-n}(U)$  is a closed disk in the Riemann sphere. Its complement is therefore an open, simply connected subset of  $\mathbb{C}$ . The closure of these subsets are nested as  $n$  increases and hence their intersection is a closed, connected set. Clearly, this intersection is  $K_c$ .

Now suppose  $Q_c^n(0) \rightarrow \infty$ . We may assume (by pulling our original  $U$  back enough times), that  $c$  lies on the boundary of  $U$ . Now we pull back once more. Every point in  $U$  (with the exception of  $c$ ) has two preimages symmetrically located about 0. The critical value has only one preimage, namely 0, so the boundary of the complement of  $Q_c^{-1}(U)$  is topologically a closed figure eight curve.

Now the complement of  $Q_c^{-1}(U)$  consists of two open, simply connected regions which we denote by  $I_0$  and  $I_1$ . Each  $I_j$  is mapped in one-to-one fashion by  $Q_c$  onto the complement of  $U$  which we call  $V$ . Hence  $Q_c(I_j) = V \supset (I_0 \cup I_1)$  in its interior.

It follows that if  $z \in K_c$ , then the entire orbit of  $z$  lies in  $I_0 \cup I_1$ . Thus we can assign a sequence of symbols 0 and 1 to each  $z$  in the usual way:

$$S(z) = s_0 s_1 s_2 \dots$$

where each  $s_j$  is 0 or 1 and  $s_j = 0$  if  $Q_c^j(z) \in I_0$ ,  $s_j = 1$  if  $Q_c^j(z) \in I_1$ .

Another way to say this is as follows. Let  $P_j$  be the inverse of  $Q_c$  on  $V$  that takes its values in  $I_j$ . So  $P_0$  maps  $V$  onto  $I_0$  and  $P_1$  maps  $V$  onto  $I_1$ . We may then say that

$$S(z) = s_0 s_1 s_2 \dots$$

if

$$z \in \bigcap_{n=0}^{\infty} P_{s_0} \circ P_{s_1} \circ \dots \circ P_{s_n}(V).$$

In fact,  $z = \bigcap_{n=0}^{\infty} P_{s_0} \circ \dots \circ P_{s_n}(V)$  since each  $P_j$  is a strict contraction in the Poincaré metric on  $V$ .

This then gives a one-to-one mapping from  $K_c$  onto the space of sequences of 0's and 1's. Standard arguments show that this map is continuous with continuous inverse. Since the space of sequences is homeomorphic to a Cantor set, we're done.

### 3.3. The Mandelbrot set

The Fundamental Dichotomy is really quite amazing. It tells us that, for quadratic maps, there are only two types of filled Julia sets: those that are connected and those that are

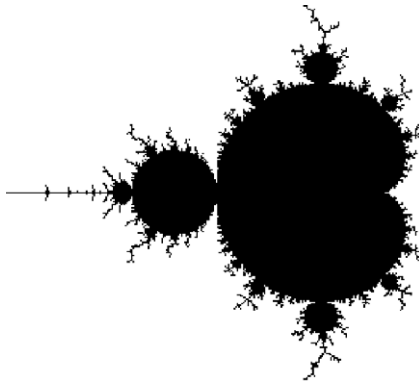


Fig. 7. The Mandelbrot set.

totally disconnected. There are no Julia sets that consist of 2 or 20 or 200 disjoint pieces. Moreover, it is the orbit of the critical value that determines which case we have. This leads to the definition of the well-known Mandelbrot set. See [16].

**DEFINITION 3.3.** The Mandelbrot set  $\mathcal{M}$  is the set of all  $c$ -values for which the orbit of 0 under  $Q_c$  does not tend to  $\infty$ . Equivalently,  $\mathcal{M}$  is the set of  $c$ -values for which the Julia set of  $Q_c$  is connected.

The image of the Mandelbrot set has become somewhat of an icon in complex dynamics. See Figure 7.

The visible bulbs in  $\mathcal{M}$  correspond to  $c$ -values for which  $Q_c$  has an attracting cycle of some given period. For example, the main central cardioid in  $\mathcal{M}$  consists of  $c$ -values for which  $Q_c$  has an attracting fixed point. This can be seen by solving for the fixed points

$$z^2 + c = z$$

that are attracting

$$|Q'_c(z)| = |2z| < 1.$$

Solving these two equations simultaneously, we see that the boundary of this region is given by

$$c = z - z^2$$

where  $z = \frac{1}{2}e^{2\pi i\theta}$ . That is, the function

$$c(\theta) = \frac{1}{2}e^{2\pi i\theta} - \frac{1}{4}e^{4\pi i\theta}$$

parametrizes the boundary of the cardioid. At  $c(\theta)$ ,  $Q_{c(\theta)}$  has a fixed point that is neutral; the derivative of  $Q_{c(\theta)}$  at this fixed point is  $e^{2\pi i\theta}$ .

For each rational value of  $\theta$ , there is a bulb tangent to the main cardioid at  $c(\theta)$ . For  $c$ -values in the bulb attached to the cardioid at  $c(p/q)$ ,  $Q_c$  has an attracting cycle of period  $q$ . We call this bulb the  $p/q$  bulb attached to the main cardioid and denote it by  $B_{p/q}$ .

It is known that, as  $c$  passes from the main cardioid, through  $c(p/q)$ , and into  $B_{p/q}$ ,  $Q_c$  undergoes a  $p/q$ -bifurcation. By this we mean: when  $c$  lies in the main cardioid near  $c(p/q)$ ,  $Q_c$  has an attracting fixed point with a nearby repelling cycle of period  $q$ . At  $c(p/q)$  the attracting fixed point and repelling cycle merge to produce the neutral fixed point with derivative  $e^{2\pi ip/q}$ . When  $c$  lies in  $B_{p/q}$ ,  $Q_c$  now has an attracting cycle of period  $q$  and a repelling fixed point. We will discuss these bulbs in more detail below, but first we pause to take up one of the most important subjects in complex dynamics.

### 3.4. External rays

Suppose  $K_c$  is connected. We now have a conjugacy  $\phi_c$  from the exterior of  $K_c$  to the exterior of the unit disk. In the language of the Riemann Mapping Theorem, this map is the *exterior Riemann map*. The map  $\phi_c$  conjugates  $Q_c$  in the exterior of  $K_c$  to  $Q_0(z) = z^2$  in  $\{z \mid |z| > 1\}$ . In particular, since  $z^2$  preserves the straight rays  $\theta = \text{constant}$ , it follows that  $Q_c$  preserves the preimages of these rays. See [3].

**DEFINITION 3.4.** The external ray of angle  $\theta$  is the preimage of the straight ray  $re^{2\pi i\theta}$  with  $r > 1$  under  $\phi_c$ .

Thus we see that the action of  $Q_c$  in the exterior of  $K_c$  is the same as the action of doubling on the straight rays outside the unit circle. Of course, we understand doubling completely (if not, see the Appendix to this section), so the conjugacy implies that we understand the dynamics of  $Q_c$  completely as well, at least outside  $K_c$ .

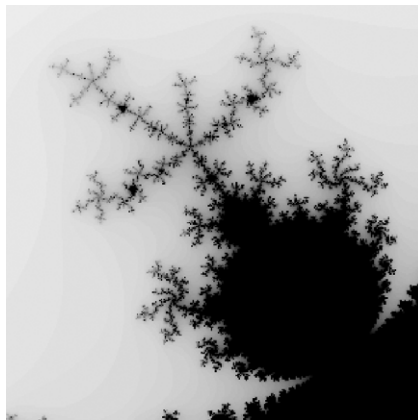
Certain of the external rays ‘land’ on the boundary of  $K_c$ . That is, for certain  $\theta$  values

$$\lim_{r \rightarrow \infty} \phi_c^{-1}(re^{2\pi i\theta})$$

exists. We call this point in  $J(Q_c)$  the *landing point* of the external ray with angle  $\theta$ .

It is known that if  $J(Q_c)$  is locally connected, then all rays actually land. (This is a consequence of Carathéodory theory, since  $Q_c$  actually gives the uniformization of the exterior of  $K_c$ .) In this case, the conjugacy with  $z^2$  shows that  $Q_c \mid J(Q_c)$  is effectively a quotient of the squaring map.

The major importance of the external rays, however, does not lie in the ‘dynamical plane’. Rather, the amazing results of Douady and Hubbard allow us to extend the exterior Riemann map to the exterior of the Mandelbrot set. Define  $\Phi(c) = \phi_c(c)$  for  $c \notin \mathcal{M}$ . Then it is known [37] that  $\Phi$  is an analytic map that takes the exterior of the Mandelbrot set onto the exterior of the unit disk in one-to-one fashion. In particular, the preimages of the straight rays under  $\Phi$  are the external rays for  $\mathcal{M}$ . We will discuss how these rays land in a subsequent section.

Fig. 8. The  $2/5$  bulb.

### 3.5. Some folk theorems

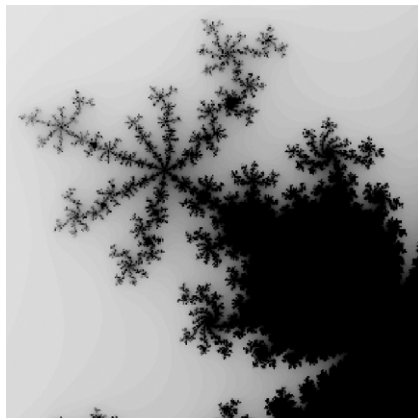
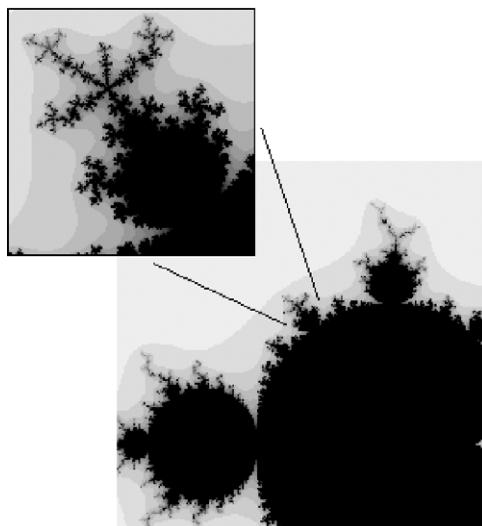
Now let's pause to have some fun. Strictly speaking, the following sections have nothing to do with our main topic, but they do at least illustrate the power of the external rays of  $\mathcal{M}$ . One of our major thrusts later will be to develop a similar theory for the more difficult exponential case.

As we described above, the Mandelbrot set consists of a basic cardioid shape (the attracting fixed point region) from which hang numerous 'bulbs' or 'decorations'. The bulbs directly attached to the cardioid are the  $p/q$  bulbs  $B(p/q)$ .

One of the surprising folk theorems we discuss below is that we can recognize the  $p/q$ -bulb from the geometry of the bulb itself. That is, we can read off dynamical information from the geometric information contained in the Mandelbrot set.

For example, the  $2/5$  bulb is displayed in Figure 8. For any  $c$ -value in the largest disk in this figure,  $Q_c$  has an attracting cycle with rotation number  $2/5$  about a central repelling fixed point. Note that the  $2/5$  bulb possesses an antenna-like structure that features a junction point from which five spokes emanate. One of these spokes is attached directly to the  $2/5$  bulb; we call this spoke the principal spoke. Now look at the 'smallest' of the non-principal spokes. Note that this spoke is located, roughly speaking,  $2/5$  of a turn in the counterclockwise direction from the principal spoke. This is how we geometrically identify this bulb as the  $2/5$ -bulb.

As another example, in Figure 9 we display the  $3/7$  bulb. Note that this bulb has 7 spokes emanating from the junction point, and the smallest is located  $3/7$  of a turn in the counterclockwise direction from the principal spoke. This then is the folk theorem: You can recognize the  $p/q$  bulb by locating the 'smallest' spoke in the antenna and determining its location relative to the principal spoke. Of course, the word 'smallest' needs some clarification here; later in this section we will make this notion precise. As an additional disclaimer, this folk theorem is only about 80% true using the Euclidean notion of 'smallness' or Lebesgue measure. Our goal is to provide a somewhat different framework in which this result is always true.

Fig. 9. The  $3/7$  bulb.Fig. 10.  $\frac{1}{2} \oplus \frac{1}{3} = \frac{2}{5}$ .

There is more to the story of interplay between the geometry of the Mandelbrot set and the corresponding dynamics. In Figure 10, we display the  $1/2$  and  $1/3$  bulbs. The  $1/2$  bulb is the large bulb to the left; the  $1/3$  bulb is the topmost bulb. In between these two bulbs are infinitely many smaller bulbs, but the largest we recognize as the  $2/5$  bulb. Now note that  $2/5$  can be obtained from  $1/2$  and  $1/3$  by ‘Farey addition’:

$$\frac{1}{2} \oplus \frac{1}{3} = \frac{2}{5}.$$

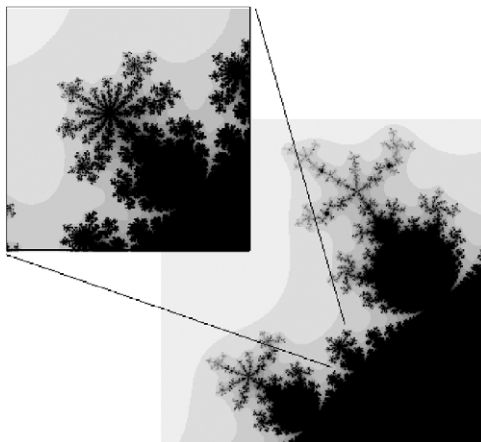


Fig. 11.  $\frac{2}{5} \oplus \frac{3}{7} = \frac{5}{12}$ .

That is, to obtain the largest bulb between two given bulbs (in a particular ordering), we simply add the corresponding fractions just the way we always wanted to add them, namely by adding the numerators and adding the denominators. This is the second of the folk theorems we discuss below. In particular it follows that the size of bulbs is determined by the Farey tree. For a discussion of the basic properties of the Farey tree, see the Appendix in this section.

As a second example, note that

$$\frac{2}{5} \oplus \frac{3}{7} = \frac{5}{12}$$

and that the  $5/12$  bulb is the largest between the  $2/5$  and  $3/7$  bulbs. See Figure 11.

While we will not give complete proofs of each of these folk theorems in this paper, we will indicate some of the combinatorial arguments involved in making the statements precise. For more folk theorems and complete proofs, we refer to [30].

### 3.6. Landing points of external rays

Now we can begin to make precise the folk theorems mentioned above. In order to do this, we will use some well known facts about both the Farey tree and angle doubling mod 1. See the Appendix to this section for more details about these prerequisites.

Let  $\mathcal{O}$  denote the exterior of the unit circle in the plane, i.e.

$$\mathcal{O} = \{z \mid |z| > 1\}.$$

As mentioned above, there is a unique analytic homeomorphism  $\Phi$  mapping the exterior of the Mandelbrot set to  $\mathcal{O}$ . The mapping  $\Phi$  takes positive reals to positive reals. This mapping is the uniformization of the exterior of the Mandelbrot set, or the exterior Riemann map.

The importance of  $\Phi$  stems from the fact that the image under  $\Phi^{-1}$  of the straight rays  $\theta = \text{constant}$  in  $\mathcal{O}$  has dynamical significance. In the Mandelbrot set, we define the external ray with external angle  $\theta_0$  to be the image of  $\theta = \theta_0$  under  $\Phi^{-1}$ . It is known that an external ray whose angle  $\theta_0$  is rational actually ‘lands’ on  $\mathcal{M}$ . That is

$$\lim_{r \rightarrow 1} \Phi^{-1}(re^{2\pi i \theta_0})$$

exists and is a unique point on the boundary of  $\mathcal{M}$ . This  $c$ -value is called the landing point of the ray with angle  $\theta_0$ .

For example, the ray with angle 0 lies on the real axis and lands on  $\mathcal{M}$  at the cusp of the main cardioid, namely the parameter  $c = 1/4$ . Also, the ray with angle  $1/2$  lies on the negative real axis and lands on  $\mathcal{M}$  at the tip of the ‘tail’ of  $\mathcal{M}$  which can be shown to be  $c = -2$ .

Consider now the interior of  $\mathcal{M}$ . The interior consists of infinitely many simply connected regions. A *bulb* of  $\mathcal{M}$  is a component of the interior of  $\mathcal{M}$  in which each  $c$ -value corresponds to a quadratic function which admits an attracting cycle. The period of this cycle is constant over each bulb. In many cases, a bulb is attached to a component of lower period at a unique point called the *root point* of the component.

The important result of Douady and Hubbard [37] is:

**THEOREM 3.5.** *Suppose a bulb  $B$  consists of  $c$ -values for which the quadratic map has an attracting  $q$ -cycle. Then the root point of this bulb is the landing point of exactly 2 rays, and the angles of each of these rays have period  $q$  under doubling.*

Thus, how the angles of the external rays of  $\mathcal{M}$  are arranged determines the ordering of the bulbs in  $\mathcal{M}$ . For example, the large bulb directly to the left of the main cardioid is the  $1/2$  bulb, so two rays with period 2 under doubling must land there. Now the only angles with period 2 under doubling are  $1/3$  and  $2/3$ , so these are the angles of the rays that land at the root point of  $B_{1/2}$ .

Now consider the  $1/3$  bulb atop the main cardioid. This bulb lies ‘between’ the rays 0 and  $1/3$ . There are only two angles between 0 and  $1/3$  that have period 3 under doubling, namely  $1/7$  and  $2/7$ , so these are the rays that land at the root point of  $B_{1/3}$ .

The  $2/5$  bulb lies between the  $1/3$  and  $1/2$  bulbs. Hence the rays that land at this root point must have period 5 under doubling and lie between  $2/7$  and  $1/3$ . The only angles that have this property are  $9/31$  and  $10/31$ , so these rays must land at the root point. See Figure 12.

These ideas allow us to measure the ‘largeness’ or ‘smallness’ of portions of the Mandelbrot set. Suppose we have two rays with angles  $\theta_-$  and  $\theta_+$  that both land at a point  $c_*$  in the boundary of  $\mathcal{M}$ .

Then, by the homeomorphism  $\Phi$ , all rays with angles between  $\theta_-$  and  $\theta_+$  must approach the component of  $\mathcal{M} - \{c_*\}$  cut off by  $\theta_-$  and  $\theta_+$ . (We remark that it is not known that all such rays actually land on  $\mathcal{M}$  – indeed, this is the major open conjecture about  $\mathcal{M}$ .) Thus it is natural to measure the size of this portion of  $\mathcal{M}$  by the length of the interval  $[\theta_-, \theta_+]$ .

The root point of the  $p/q$  bulb of  $\mathcal{M}$  divides  $\mathcal{M}$  into two sets. The component containing the  $p/q$  bulbs is called the  *$p/q$  limb*. We can then measure the size of the



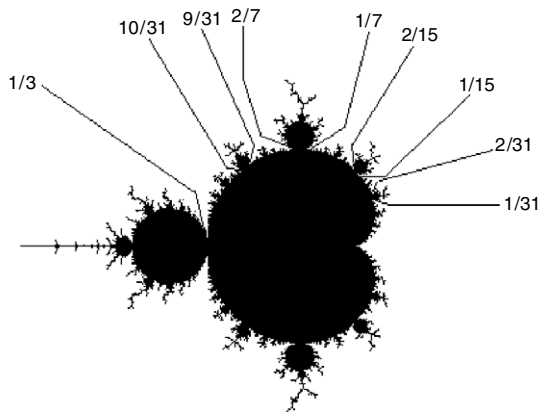


Fig. 12. Rays landing on the Mandelbrot set.

$p/q$  limb if we know the angles of the two external rays that land on the root point of the  $p/q$  bulb. This is the subject of the next section.

### 3.7. Rays landing on the $p/q$ bulb

In order to make the notion of ‘large’ or ‘small’ precise in the statement of the folk theorems, we need a way to determine the angles of the rays landing at the root point of  $B_{p/q}$ . We denote the angles of these two rays in binary by  $s_{\pm}(p/q)$ , where  $s_{-}(p/q) < s_{+}(p/q)$ . We call  $s_{-}(p/q)$  the *lower angle* of  $B_{p/q}$  and  $s_{+}(p/q)$  the *upper angle*.

As we will see,  $s_{\pm}(p/q)$  is a string of  $q$  digits (0 or 1) and so  $s_{\pm}(p/q)$  denotes the infinite repeating sequence whose basic block is  $s_{\pm}$ . Douady and Hubbard [37] have a geometric method involving Julia sets to determine these angles. Our method is more combinatorial and resembles algorithms due to Atela [4], LaVurs [53], and Lau and Schleicher [52].

To describe this algorithm, let  $R_{p/q}$  denote rotation of the unit circle through  $p/q$  turns, i.e.

$$R_{p/q}(\theta) = e^{2\pi i(\theta + p/q)}.$$

We will consider the itineraries of points in the unit circle under  $R$  using two different partitions of the circle.

The *lower partition* of the circle is defined as follows. Let  $I_0^- = \{\theta \mid 0 < \theta \leq 1 - p/q\}$  and  $I_1^- = \{\theta \mid 1 - p/q < \theta \leq 1\}$ . Note that the boundary point 0 belongs to  $I_1^-$  and  $-p/q = 1 - p/q$  belongs to  $I_0^-$ . We then define  $s_{-}(p/q)$  to be the itinerary of  $p/q$  under  $R_{p/q}$  relative to this partition. We call  $s_{-}(p/q)$  the *lower itinerary* of  $p/q$ . That is,  $s_{-}(p/q) = s_1 \dots s_q$  where  $s_j$  is either 0 or 1 and the digit  $s_j = 0$  if and only if  $R_{p/q}^{j-1}(p/q) \in I_0^-$ . Otherwise,  $s_j = 1$ .

For example,  $s_-(1/3) = 001$  since

$$I_0^- = (0, 2/3]$$

$$I_1^- = (2/3, 1]$$

and the orbit  $\frac{1}{3} \rightarrow \frac{2}{3} \rightarrow 1 \rightarrow \frac{1}{3} \dots$  lies in  $I_0^-$ ,  $I_0^-$ ,  $I_1^-$ , respectively.

Similarly,  $s_-(2/5) = 01001$  since

$$I_0^- = (0, 3/5]$$

$$I_1^- = (3/5, 1]$$

and the orbit is  $\frac{2}{5} \rightarrow \frac{4}{5} \rightarrow \frac{1}{5} \rightarrow \frac{3}{5} \rightarrow 0 \rightarrow \frac{2}{5} \dots$ .

We also define the *upper partition*  $I_0^+$  and  $I_1^+$  as follows

$$I_0^+ = [0, 1 - p/q)$$

$$I_1^+ = [1 - p/q, 1).$$

The *upper itinerary* of  $p/q$ ,  $s_+(p/q)$ , is then the itinerary of  $p/q$  relative to this partition. Note that  $I_0^+$  and  $I_1^+$  differ from  $I_0^-$  and  $I_1^-$  only at the endpoints.

For example,  $s_+(1/3) = 010$  since the orbit is  $\frac{1}{3} \rightarrow \frac{2}{3} \rightarrow 0 \dots$  and

$$I_0^+ = [0, 2/3)$$

$$I_1^+ = [2/3, 1).$$

This orbit starts in  $I_0^+$ , hops to  $I_1^+$ , and then returns to  $I_0^+$  before cycling. For  $2/5$ , we have

$$I_0^+ = [0, 3/5)$$

$$I_1^+ = [3/5, 1)$$

and  $s_+(2/5) = 01010$ .

The following theorem provides the algorithm for computing the angles of rays landing at the root point of the  $p/q$  bulb. For a proof, we refer to [37] or [30].

**THEOREM 3.6.** *The two rays landing at the root point of the  $p/q$  bulb are  $\overline{s_-(p/q)}$  and  $\overline{s_+(p/q)}$ .*

Note that  $s_{\pm}(p/q)$  differ only in their last two digits (provided  $q \geq 2$ ). Indeed we may write

$$s_-(p/q) = s_1 \dots s_{q-2} 0 1$$

$$s_+(p/q) = s_1 \dots s_{q-2} 1 0.$$

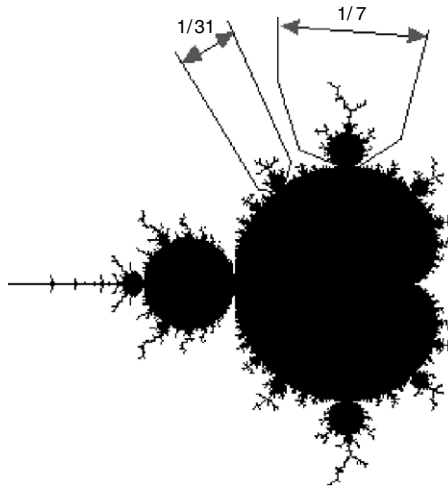


Fig. 13. Size of the  $2/5$  and  $1/3$  limbs of  $\mathcal{M}$ .

The reason for this is that the upper and lower itineraries are the same except at  $R_{p/q}^{q-2}(p/q) = -p/q$  and  $R_{p/q}^{q-1}(p/q) = 0$ , which form the endpoints of the two partitions of the circle.

We now define the *size of the  $p/q$  limb* to be the length of the interval  $[s_-(p/q), s_+(p/q)]$ . That is, the size of the  $p/q$  limb is given by the ‘number’ of external rays that approach this limb. We may compute size of these bulbs explicitly by using the fact that  $s_{\pm}(p/q)$  differ only in the last two digits.

**THEOREM 3.7.** *The size of the  $p/q$  limb is  $1/(2^q - 1)$ . That is*

$$\overline{s_+(p/q)} - \overline{s_-(p/q)} = \frac{1}{2^q - 1}.$$

To see this, just write the binary expansion of the difference in the form

$$\begin{aligned} \overline{s_+(p/q)} - \overline{s_-(p/q)} &= \frac{1}{2^{q-1}} + \frac{1}{2^{2q-1}} + \frac{1}{2^{3q-1}} \\ &\quad + \cdots - \left( \frac{1}{2^q} + \frac{1}{2^{2q}} + \frac{1}{2^{3q}} + \cdots \right) \\ &= \frac{1}{2^{q-1}} \cdot \frac{2^q}{2^q - 1} - \frac{1}{2^q} \cdot \frac{2^q}{2^q - 1} \\ &= \frac{1}{2^q - 1}. \end{aligned}$$

As we see in [Figure 13](#), the visual size of the bulbs does indeed correspond to the size as defined above.

### 3.8. The size of limbs and the Farey tree

In this section we relate the size of a  $p/q$  limb to the size of the limbs corresponding to the Farey parents of  $p/q$ . The following proposition relates the upper and lower itineraries of  $p/q$  and its Farey parents.

PROPOSITION 3.8. *Suppose*

$$0 < \frac{\alpha}{\beta} < \frac{\gamma}{\delta} < 1$$

*are the Farey parents of  $p/q$ . Then the lower itinerary  $s_-(p/q)$  consists of the first  $q$  digits of the upper angle  $\overline{s_+(\alpha/\beta)}$  of the smaller parent, and the upper itinerary  $s_+(p/q)$  consists of the first  $q$  digits of the lower angle  $\overline{s_-(\gamma/\delta)}$  of the larger parent.*

For a proof, we refer to [30].

In case one of the Farey parents are 0 or 1, we must modify the above proposition.

PROPOSITION 3.9. *Suppose that a Farey parent of  $p/q$  is 0. Then the  $q$  digits in the lower itinerary of  $p/q$  are given by*

$$s_-(p/q) = 0 \dots 01.$$

*If a Farey parent of  $p/q$  is 1, then we have*

$$s_+(p/q) = 1 \dots 10.$$

We now sketch the proof of one of the folk theorems mentioned earlier.

THEOREM 3.10. *Suppose  $\alpha/\beta < \gamma/\delta$  are the Farey parents of  $p/q$ . Then the size of the  $p/q$  limb is larger than the size of any other limb between the  $\alpha/\beta$  and  $\gamma/\delta$  limbs.*

PROOF. Assume first that neither of the parents are 0 or 1. By the previous proposition, we have that  $\overline{s_-(p/q)}$  and  $\overline{s_+(\alpha/\beta)}$  agree in their first  $q$  digits. Using these binary representations, we have

$$\overline{s_-(p/q)} - \overline{s_+(\alpha/\beta)} \leq \frac{1}{2^q}.$$

Similarly

$$\overline{s_-(\gamma/\delta)} - \overline{s_+(p/q)} \leq \frac{1}{2^q}.$$

This implies that the arc of rays between the  $p/q$  limb and either of its parents' limbs has length no larger than  $1/2^q$ . Thus any limb between them has size smaller than  $1/2^q$ .

From the previous section, we know that

$$\overline{s_+(p/q)} - \overline{s_-(p/q)} = \frac{1}{2^q - 1}.$$

As this quantity is larger than  $1/2^q$ , it follows that the  $p/q$  limb attracts the largest number of rays between its two parents.

In case one of the parents of  $1/q$  is 0, then we have that the size of the  $1/q$  bulb is  $1/(2^q - 1)$  as above, while the gap between 0 and  $\overline{s_-(p/q)} = 0. \dots 01$  is also  $1/(2^q - 1)$ . But then any limb between the  $1/q$  limb and the cusp of the cardioid must have size strictly smaller than  $1/(2^q - 1)$ , again showing that the  $1/q$  limb is the largest. The gap between the limbs of  $1/q$  and its other Farey parent  $1/(q - 1)$  is handled as above. The case of Farey parent 1 is handled similarly.

### 3.9. Further remarks

The technique of measuring the size of certain portions of the Mandelbrot set by the length of the interval of rays that land on that portion provides justification for other folk theorems involving the size of  $\mathcal{M}$ . For example, this is the same technique that is used to identify the  $p/q$  bulb using the ‘lengths’ of the spokes in its antenna. Once we know these rays, we can easily compute the lengths of the various spokes.

As an example of this, it can be shown that the two rays that land at the junction point of the antenna adjacent to the principal spoke are given by  $s_-\overline{s_+}$  and  $s_+\overline{s_-}$  where we have dropped the  $p/q$  for clarity. These two rays are therefore given by preperiodic binary sequences that begin to repeat only after the  $q$ th entry.

This fact shows that the vast majority of rays that land on the  $p/q$  limb actually approach the spokes of the antenna. For we have the following ordering of the rays landing on the  $p/q$  bulb:

$$\overline{s_-} < s_-\overline{s_+} < s_+\overline{s_-} < \overline{s_+}.$$

It is easy to check, using the above techniques, that the length of the arc of rays approaching the antenna between  $s_-\overline{s_+}$  and  $s_+\overline{s_-}$  is

$$\frac{1}{2^{q-1}} - \frac{2}{2^q(2^{q-1})}.$$

This number is much larger than the length of the arc between  $\overline{s_-}$  and  $s_-\overline{s_+}$  or between  $\overline{s_+}$  and  $s_+\overline{s_-}$ , each of which has length

$$\frac{1}{2^q(2^{q-1})}.$$

We can also use these two rays separating the principal spoke from the rest of the antenna to determine a list of the  $q$  rays that land on the junction point. Then using the techniques above we can determine that the shortest is located  $p/q$  turns in the counterclockwise direction from the principal spoke. See [30] for details.

So the moral of the stories in this section is: there is a way, using some simple geometry and dynamics, to understand large portions of the parameter plane for quadratic polynomials. We shall see similar stories unfold for the exponential map, though the dynamics will no longer be quite so simple.

### 3.10. Appendix: The Farey tree

We recall here a few facts about the *Farey tree*. The Farey tree is a tree containing all of the rationals between 0 and 1. At each stage of its construction, the Farey tree consists of a finite list of rationals. Adjacent rationals in this list are called *Farey neighbours*. The inductive step in the construction of the tree is: Each pair of Farey neighbours produces a *Farey child*, which is the rational between the two whose denominator is the smallest. Naturally, the rationals that produce a Farey child are called its *Farey parents*. See [54].

One of the most intriguing features of the Farey tree is that we obtain Farey children by Farey addition. That is, the fraction between the Farey neighbours  $\alpha/\beta$  and  $\gamma/\delta$  is given by

$$\frac{\alpha}{\beta} \oplus \frac{\gamma}{\delta} = \frac{\alpha + \gamma}{\beta + \delta}.$$

That is, to obtain the fraction between two Farey neighbours whose denominator is the smallest, we simply add the numerators and add the denominators of the parents to obtain the child.

We begin the construction of the tree with the pair of rationals 0 and 1 which we write as 0/1 and 1/1. Their child is 1/2, so the second stage of the construction gives the list

$$\frac{0}{1} \quad \frac{1}{2} \quad \frac{1}{1}.$$

At the next stage we obtain two new Farey children

$$\frac{0}{1} \quad \frac{1}{3} \quad \frac{1}{2} \quad \frac{2}{3} \quad \frac{1}{1}.$$

At generation four we find

$$\frac{0}{1} \quad \frac{1}{4} \quad \frac{1}{3} \quad \frac{2}{5} \quad \frac{1}{2} \quad \frac{3}{5} \quad \frac{2}{3} \quad \frac{3}{4} \quad \frac{1}{1}.$$

It is a fact that the Farey tree contains all rationals. See [45] or [43] for more details.

One other fact about the Farey tree is that  $\alpha/\beta$  and  $\gamma/\delta$  are Farey neighbours if and only if  $\alpha\delta - \gamma\beta = \pm 1$ . Consequently, we have

$$\left| \frac{\alpha}{\beta} - \frac{\gamma}{\delta} \right| = \frac{1}{\beta\delta}.$$

This is easily proved by induction.

### 3.11. Appendix: Angle doubling

In order to use the fundamental results of Douady and Hubbard [37] regarding the Mandelbrot set we recall in this appendix some facts about the doubling function. The

doubling function is defined on the circle considered as the reals modulo one and is given by  $D(\theta) = 2\theta \bmod 1$ .

We need two facts about  $D$ :

FACT 1. The angle  $\theta$  is periodic under  $D$  if and only if  $\theta$  is a rational of the form  $p/q$  (in lowest terms) with  $q$  odd.

For example, the  $D$ -orbit of  $1/3$  is

$$\frac{1}{3} \rightarrow \frac{2}{3} \rightarrow \frac{1}{3} \cdots$$

which has period 2. The rational  $1/7$  has period 3 under doubling:

$$\frac{1}{7} \rightarrow \frac{2}{7} \rightarrow \frac{4}{7} \rightarrow \frac{1}{7} \cdots$$

while  $1/5$  has period 4:

$$\frac{1}{5} \rightarrow \frac{2}{5} \rightarrow \frac{4}{5} \rightarrow \frac{3}{5} \rightarrow \frac{1}{5} \cdots$$

The rationals with even denominator are eventually periodic but not periodic. For example,  $1/6$  lies on an eventual 2-cycle

$$\frac{1}{6} \rightarrow \frac{1}{3} \rightarrow \frac{2}{3} \rightarrow \frac{1}{3} \cdots$$

and  $1/8$  is eventually fixed:

$$\frac{1}{8} \rightarrow \frac{1}{4} \rightarrow \frac{1}{2} \rightarrow 1 \rightarrow 1 \cdots$$

A second important fact about doubling is that we can read off the binary expansion of  $\theta$  by noting the *itinerary* of  $\theta$  in the circle relative to  $D$ . To define the itinerary, we denote the upper semicircle  $0 \leq \theta < 1/2$  by  $I_0$  and the lower semicircle  $1/2 \leq \theta < 1$  by  $I_1$ . Given  $\theta$ , we attach an infinite string of 0s and 1s to  $\theta$  as follows: The itinerary of  $\theta$  is  $B(\theta) = (s_0 s_1 s_2 \dots)$  where  $s_j$  is either 0 or 1, and  $s_j = 0$  if  $D^j(\theta) \in I_0$ ,  $s_j = 1$  if  $D^j(\theta) \in I_1$ . That is, we simply watch the orbit of  $\theta$  in the circle under doubling and assign 0 or 1 to the itinerary whenever  $D^j(\theta)$  lands in the arc  $I_0$  or  $I_1$ .

FACT 2. The itinerary  $B(\theta)$  is the binary expansion of  $\theta$ .

For example, if  $\theta = 1/3$ , then  $\theta \in I_0$ , while  $D(\theta) \in I_1$  and  $D^2(\theta) = \theta$ . Hence  $B(1/3)$  is the repeating sequence  $\overline{01}$ , which of course is the binary expansion of  $1/3$ . Similarly,  $B(1/7) = \overline{001}$  while  $B(1/5) = \overline{0011}$ .

## 4. Exponential dynamics

In this section we begin the study of the most important entire function, the complex exponential  $E_\lambda(z) = \lambda e^z$ . Our goal is to begin to paint the picture of the parameter plane or *bifurcation diagram* for this family, the analogue of the Mandelbrot set for the quadratic family. First seen around 1984 [6,7,20,39], this set and the Mandelbrot set are very different.

### 4.1. Computing the Julia set

One of the principal differences between quadratic and exponential dynamics deals with orbits that tend to  $\infty$ . For polynomials, any orbit that tends to  $\infty$  is necessarily in the Fatou set as the point at  $\infty$  is a super-attracting fixed point. This is not the case for  $E_\lambda$ :

**THEOREM 4.1.** *Suppose  $E_\lambda^n(z) \rightarrow \infty$ . Then  $z \in J(E_\lambda)$ . In fact,  $J(E_\lambda)$  is the closure of the set of points that escape to  $\infty$ .*

**PROOF.** If the orbit of  $z$  lies in the Fatou set, then the classification of stable domains says that the orbit of  $z$  must eventually enter an attracting basin or a Siegel disk. Thus, escaping orbits do not lie in the Fatou set.

Now there must be at least one escaping orbit for  $E_\lambda$ , since, roughly speaking, many points in the far right half-plane are mapped further to the right. Thus, given any  $z \in J(E_\lambda)$ , and any neighbourhood of  $z$ , by Montel's Theorem the iterates of  $E_\lambda$  map this neighbourhood over  $\mathbb{C} - \{0\}$ . In particular, there is a preimage of the escaping orbit in this neighbourhood, and so the escaping points are dense in the Julia set of  $E_\lambda$ .  $\square$

**REMARK.** In general, escaping points for entire functions do not lie in the Julia set, as our earlier examples of wandering domains and domains at infinity show. But critically finite entire functions do have this property.

This fact gives us a method to compute the Julia sets for complex exponentials. If the orbit of  $z$  tends to  $\infty$ , then it must do so with  $\operatorname{Re} E_\lambda^n(z) \rightarrow \infty$ . Hence we paint the picture of a Julia set using the following algorithm.

- ALGORITHM.**
1. Compute the orbit of  $z$  up to  $N$  iterations.
  2. If the orbit of  $z$  enters the region  $\operatorname{Re} z \geq 50$  at iteration  $j \leq N$ , then 'colour'  $z$  with a colour corresponding to  $j$ . Then  $z \in J(E_\lambda)$ .
  3. If the orbit never enters this half-plane, then colour  $z$  black and declare that  $z \notin J(E_\lambda)$ .

It is interesting to note that this very crude algorithm actually yields quite accurate pictures of the Julia set. See [38]. As we will see later, 'most' of the points in the Julia set actually have orbits that tend to  $\infty$ . Since  $e^{50}$  is quite large, we capture at least some of these escape orbits by tracking them into the half-plane  $\operatorname{Re} z > 50$ .



## 4.2. Explosions

The complex exponential family has only one singular value, namely the asymptotic value at 0. Just as in the quadratic case, we use the orbit of this point to paint the picture of the parameter plane for  $E_\lambda$ . We do not get as sharp a dichotomy in the exponential case, as the topology of the Julia sets for  $E_\lambda$  do not fall neatly into two distinct categories as in the quadratic case. However, the Classification of Stable Domains Theorem does allow us to say something depending upon the fate of the orbit of 0.

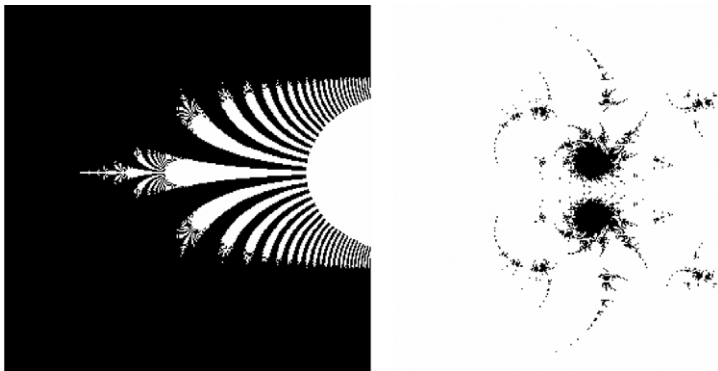
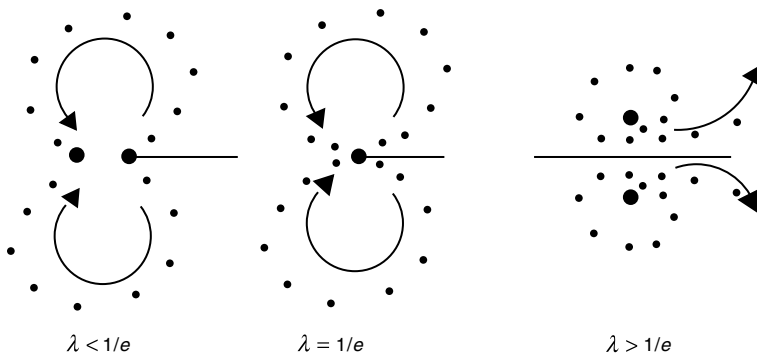
**THEOREM 4.2.**    1. Suppose the orbit of 0 tends to  $\infty$  under  $E_\lambda$ . Then  $J(E_\lambda) = \mathbb{C}$ .  
                           2. Suppose the orbit of 0 under  $E_\lambda$  is preperiodic. Then  $J(E_\lambda) = \mathbb{C}$ .  
                           3. Suppose that  $E_\lambda$  has an attracting or neutral periodic orbit. Then  $J(E_\lambda)$  is a nowhere dense subset of  $\mathbb{C}$ . In fact,  $J(E_\lambda)$  is confined to a half-plane of the form  $\operatorname{Re} z \geq v$  for some  $v \in \mathbb{R}$ .

**PROOF.** In the first two cases, the Classification of Stable Domains shows that the Fatou set must be empty, as any stable domain must absorb a singular value. In the final case, the orbit of 0 tends to the attracting or neutral periodic orbit. Hence there is a far left half-plane that is mapped into the basin of this cycle, so  $J(E_\lambda)$  must lie to the right of this plane.  $\square$

**EXAMPLE.** As we have seen,  $E_\lambda$  has an attracting fixed point if  $0 < \lambda < 1/e$ . When  $\lambda = 1/e$ , this fixed point turns rationally indifferent. When  $\lambda > 1/e$ , there are no fixed points on the real axis and the orbit of 0 tends to  $\infty$ . Hence  $J(E_\lambda)$  is nowhere dense if  $\lambda \leq 1/e$  while  $J(E_\lambda) = \mathbb{C}$  when  $\lambda > 1/e$ . In fact, as we will see shortly,  $J(E_\lambda)$  is contained in the half-plane  $\operatorname{Re} z \geq 1$  for  $0 < \lambda \leq 1/e$ . So we have an *explosion* when  $\lambda$  passes through  $1/e$ : Suddenly the Julia set changes from occupying a small portion of the plane to filling the entire plane! See Figure 14. In these images, the black region is the Fatou set. (In the case of  $\lambda = 0.4$  we see some black only because we iterated only 50 times to produce these images.) Also, it appears that the Julia set for  $\lambda < 1/e$  contains open sets. This, of course, is not the case. We will discuss this artifact of our computing algorithm in Section 5.

The explosion in the Julia set is really quite remarkable. We know that repelling periodic points are dense in  $J(E_\lambda)$ . As  $\lambda$  passes through  $1/e$ , all repelling periodic points move continuously, and only one new repelling periodic point is born. Yet, when  $\lambda \leq 1/e$ , all of these periodic points occupy a (rather small subset of) the right half-plane, but when  $\lambda > 1/e$ , they are dense in  $\mathbb{C}$ . Julia sets for  $E_\lambda$  may change in an interesting manner as  $\lambda$  varies.

The explosion in  $J(E_\lambda)$  is a global consequence of a relatively simple complex bifurcation, the *saddle-node bifurcation*. As  $\lambda$  approaches  $1/e$  from below, an attracting and a repelling fixed point come together and merge to form the neutral fixed point at  $\lambda = 1/e$ . When  $\lambda$  becomes larger than  $1/e$ , two new fixed points emerge, both of which are repelling with complex derivatives. The local picture of this bifurcation is shown in Figure 15. We will see that similar ‘global’ bifurcations almost always accompany simple bifurcations like the saddle-node or period-doubling.

Fig. 14. The Julia sets for  $\lambda = 1/e$  and  $\lambda = 0.4$ .Fig. 15. The local saddle-node bifurcation at  $\lambda = 1/e$ .

### 4.3. Misiurewicz points

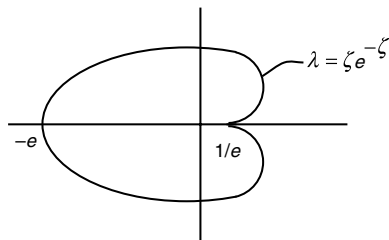
In case the orbit of 0 is preperiodic,  $J(E_\lambda) = \mathbb{C}$  and we say that  $\lambda$  is a Misiurewicz point. For example,  $\lambda = 2k\pi i$  is a Misiurewicz point if  $k \in \mathbb{Z}$  since

$$0 \mapsto 2k\pi i \mapsto 2k\pi i \cdots.$$

Similarly,  $\lambda = (2k + 1)\pi i$  is also a Misiurewicz point since

$$0 \mapsto (2k + 1)\pi i \mapsto -(2k + 1)\pi i \mapsto -(2k + 1)\pi i.$$

Misiurewicz points play the same role in the parameter plane for  $E_\lambda$  as the junction points in the Mandelbrot set. See [26].

Fig. 16. The attracting fixed point region  $C_1$ .

#### 4.4. Hyperbolic components of period 1–3

Clearly, when  $E_\lambda$  has an attracting cycle, there is an open set about  $\lambda$  in the bifurcation diagram in which all of the corresponding exponentials have an attracting cycle of the same period. The maximal such open connected region in the  $\lambda$ -plane is called a *hyperbolic component*.

The set of  $\lambda$ -values for which  $E_\lambda$  has an attracting fixed point is, as in the case of the quadratic family, cardioid-like. For  $E_\lambda$  has an attracting fixed point if

$$E_\lambda(z) = \lambda e^z = z$$

$$E'_\lambda(z) = \lambda e^z = \zeta$$

where  $|\zeta| < 1$ . These equations yield

$$\lambda = \zeta e^{-\zeta}, \quad |\zeta| < 1.$$

This region is a cardioid  $C_1$  in the plane as depicted in Figure 16.

Note that, as  $\lambda$  traverses the boundary of  $C_1$ , the multiplier  $\zeta$  wraps once around the unit circle. Hence we expect period- $n$  bifurcations to occur whenever  $\zeta$  is an  $n$ th root of unity just as in the case of the Mandelbrot set. Indeed, when  $\lambda = 1/e$ ,  $\zeta = 1$  and we have the saddle node bifurcation.

EXERCISE. Show that when  $\lambda = -e$ , the exponential family undergoes a period doubling bifurcation.

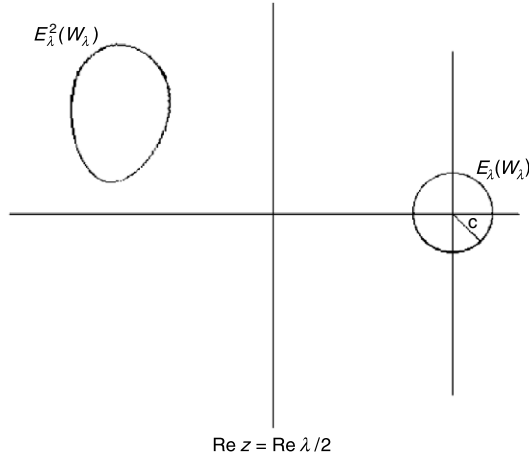
We now turn to the attracting period 2 region,  $C_2$ . This region occupies a large region in the left half-plane. For each  $\nu < 0$ ,  $\eta > 0$ , define

$$Z_{\nu,\eta} = \{\lambda \in \mathbb{C} \mid \operatorname{Re} \lambda \leq \nu, |\operatorname{Im} \lambda| \leq \eta\}.$$

$Z_{\nu,\eta}$  is a closed half strip in the left portion of the parameter plane.

PROPOSITION 4.3. For each  $\eta \in \mathbb{R}^+$ , there is  $\nu = \nu(\eta) < 0$  such that if  $\lambda \in Z_{\nu,\eta}$  then

1.  $E_\lambda$  has an attracting two-cycle, and
2.  $W_\lambda = \{z \mid \operatorname{Re} z \leq \operatorname{Re}(\lambda/2)\}$  is contained in the basin of attraction of the two-cycle.

Fig. 17.  $E_\lambda^2(W_\lambda) \subset W_\lambda$ .

PROOF. Fix  $\eta > 0$  and suppose that  $|\operatorname{Im} \lambda| \leq \eta$ . Define

$$c = c(\lambda) = |\lambda| \exp(\operatorname{Re}(\lambda/2)).$$

Note that  $|\lambda|c(\lambda) \rightarrow 0$  and  $|\operatorname{Re} \lambda|c(\lambda) \rightarrow 0$  as  $\operatorname{Re} \lambda \rightarrow -\infty$  since  $|\operatorname{Im} \lambda|$  is bounded.

We may choose  $\nu = \nu(\eta)$  such that, if  $\lambda \in Z_{\nu, \eta}$ , then

$$|\lambda|ce^c \leq (|\operatorname{Re} \lambda| + \eta)ce^c \leq |\operatorname{Re} \lambda|/4$$

since  $(|\operatorname{Re} \lambda| + \eta)ce^c$  tends to 0 as  $\operatorname{Re} \lambda \rightarrow -\infty$ .

We claim that  $E_\lambda^2$  maps  $W_\lambda$  inside itself. To see this, first note that  $E_\lambda(W_\lambda)$  is a punctured disk of radius  $c$  centred at 0. On this disk

$$|E'_\lambda(z)| = |E_\lambda(z)| = |\lambda| \exp(\operatorname{Re} z)$$

which is at most  $|\lambda|e^c$ . Clearly,  $E_\lambda(0) = \lambda$ , so points  $z \in E_\lambda(W_\lambda)$  are mapped a maximum distance of

$$\begin{aligned} |E_\lambda(z) - \lambda| &= |z - 0| \max_{z \in E_\lambda(W_\lambda)} |E'_\lambda(z)| \\ &\leq |\lambda|e^c \cdot c \\ &\leq |\operatorname{Re} \lambda|/4. \end{aligned}$$

away from  $\lambda$  under  $E_\lambda$ , and thus well to the left of the line  $\operatorname{Re} z = \operatorname{Re} \lambda/2$ , see Figure 17.

As a consequence,  $E_\lambda^2(W_\lambda)$  is contained in the interior of  $W_\lambda$ , provided  $\lambda \in Z_{\nu, \lambda}$ . By the Schwarz lemma,  $E_\lambda^2$  has an attracting fixed point in  $W_\lambda$  and, moreover, each point in  $W_\lambda$  tends to this point under iteration of  $E_\lambda^2$ . This fixed point gives an attracting 2-cycle for  $E_\lambda$ .  $\square$

Now we turn to the period three regions. Unlike  $C_1$  and  $C_2$ , the set  $C_3$  has infinitely many components.

Let  $\mu = a + i\pi$  with  $a$  sufficiently large. We claim that the map  $E_\mu$  has an attracting cycle of period 3. To see this, we first note that the real part of  $E_\mu^i(0)$  satisfies

$$\begin{aligned}\operatorname{Re} E_\mu(0) &= a \\ \operatorname{Re} E_\mu^2(0) &\approx -ae^a.\end{aligned}$$

Thus

$$|E_\mu^3(0)| \approx ae^{-ae^a}$$

which is very close to 0 when  $a$  is large.

Let  $B_\delta$  denote the ball of radius  $\delta$  centred at the origin. Then  $E_\mu^j(B_\delta)$  contains a ball whose radius is of the order of  $a\delta$  centred at  $\mu = E_\mu(0)$  if  $j = 1$ ; of the order of  $ae^a \cdot a\delta$  centred at  $E_\mu^2(0)$  if  $j = 2$ ; and of the order of  $ae^{-ae^a} \cdot ae^a \cdot a\delta$  centred at  $E_\mu^3(0)$  if  $j = 3$ . One checks easily that this latter radius is much smaller than  $\delta$  for  $\delta$  of the order of  $1/a$ . Moreover, the distance from  $E_\mu^3(0)$  to 0 is much smaller than  $\delta$ . Consequently,  $E_\mu^3$  maps  $B_\delta$  inside itself, and so  $E_\mu$  has an attracting cycle of period 3.

EXERCISE. Consider  $v = a + (2k + 1)\pi i$  for  $a$  large and  $k \in \mathbb{Z}$ . Show that  $E_v$  has an attracting cycle of period 3.

Consequently, there are open sets that contain the portions of the horizontal line  $\operatorname{Im} z = (2k+1)\pi$  in the far right half-plane that give  $\lambda$ -values for which  $E_\lambda$  has an attracting cycle of period 3. In fact all of these sets are distinct.

Thus the bifurcation diagram or parameter plane for the exponential family is quite different from that of the quadratic map. See [Figures 20](#) and [21](#). The small cardioid in the centre of [Figure 18](#) is  $C_1$ , the large region to the left is  $C_2$ ; and the large horizontal black regions tending to infinity in the right half-plane make up components of  $C_3$ . A portion of  $C_1$  is magnified in [Figure 19](#), and further zooms are in [Figures 20](#) and [21](#).

#### 4.5. Hyperbolic components with higher periods

Suppose  $z_0 = z_0(\lambda)$  is an attracting periodic point for  $E_\lambda$ . We write  $z_i = E_\lambda^i(z_0)$ . Let  $C_k$  denote the set of  $\lambda$ -values for which  $E_\lambda$  has an attracting cycle of period  $k \geq 3$ . We claim that each component of  $C_k$  is unbounded and simply connected. To prove this, we first need a lemma.

LEMMA 4.4. *Let  $U \subset C_k$  be open with  $\overline{U}$  compact and  $0 \notin \overline{U}$ . Then there exist constants  $c_1, c_2$  such that, if  $\lambda \in U$ , then  $c_1 \leq |z_j(\lambda)| \leq c_2$  for  $j = 1, \dots, k-1$ .*

PROOF. Suppose some  $|z_j(\lambda)| \rightarrow \infty$  as  $\lambda \rightarrow \lambda_0$ . Let  $D_r(\lambda_0)$  denote the ball of radius  $r$  about  $\lambda_0$  in the parameter plane. For each  $\lambda \in D_r(\lambda_0) \cap C_k$ , there is at least one  $z_i(\lambda)$  for

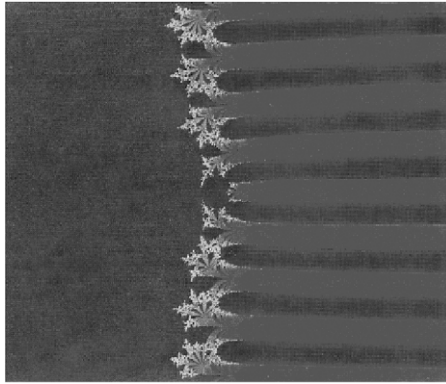


Fig. 18. The parameter plane for  $E_\lambda$ .

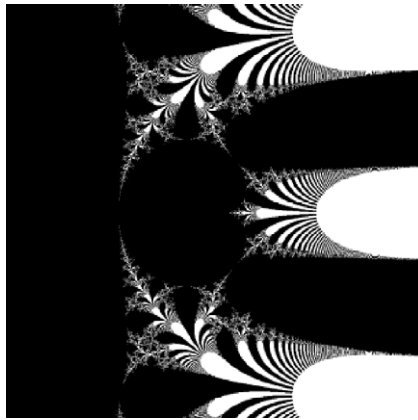


Fig. 19. The cardioid  $C_1$ .

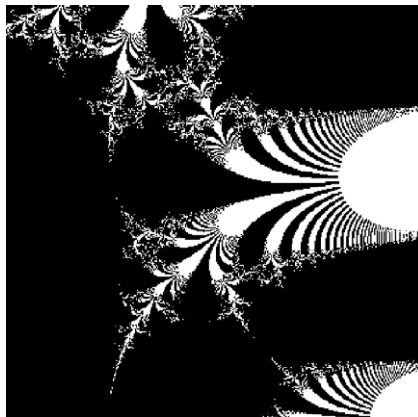


Fig. 20. A zoom into the previous figure just above  $C_1$ .

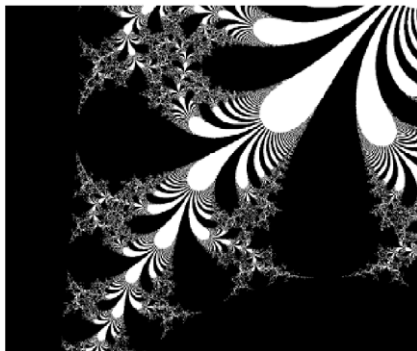


Fig. 21. A further zoom.

which  $|z_i(\lambda)| < 1$ . This follows from the Chain Rule and the fact that  $E'_\lambda(z_i) = z_{i+1}$  for all  $i$ . Then

$$c_2 = \max_{s=1, \dots, k-1} |E'_\lambda(z_i)|$$

is the required upper bound. Similar arguments give a lower bound.  $\square$

On  $C_k$  we have the ‘multiplier map’. This is the function  $\chi : C_k \rightarrow D$  where  $D$  is the open unit disk and

$$\chi(\lambda) = E'_\lambda(z_i(\lambda)).$$

**PROPOSITION 4.5.** *Each component of  $C_k$  is simply connected.*

**PROOF.** Let  $G_n(\lambda) = E^n_\lambda(0)$ ;  $\{G_n\}$  is a family of entire functions. Let  $C$  be a component of  $C_k$ ,  $\gamma \subset C$ , a simple closed curve bounding a region  $D$ , and  $U$ , a small neighbourhood of  $\gamma$  in  $C$ . We will show that  $D \subset C$ .

Since  $U \subset C$ , the functions  $\{G_{nk}\}$  converge to the periodic point  $z_0(\lambda)$  on  $U$ . By the Lemma,  $|z_0(\lambda)| < c_2 < \infty$  on  $U$ , so it follows that the  $\{G_{nk}\}$  converge on all of  $D$ , and the limit function determines a period  $k$  periodic point  $z_0(\lambda)$  for all  $\lambda$ . By the maximum principle,  $|\chi(\lambda)| < 1$  on  $D$ , where  $\chi(\lambda)$  is the multiplier.  $\square$

**PROPOSITION 4.6.** *If  $k \geq 2$ , each component of  $C_k$  is unbounded.*

**PROOF.** Let  $C$  be a relatively compact component of  $C_k$ . We know that  $\chi(\lambda)$  is bounded away from 0. Therefore, if  $\lambda_i \rightarrow \lambda \in \overline{C} - C$ , then  $|\chi(\lambda)| = 1$ .

In order to see that  $\chi(C)$  is closed, we take a sequence of points in this set and see that they have a limit point in the set. Consider the sequence  $\chi(\lambda_i)$ . If we look the sequence of  $\lambda_i$ , it will have limit  $\lambda$  either in  $\overline{C} - C$  or  $C$ . In the former case, we will have  $z_j(\lambda_i) \rightarrow z'_j(\lambda)$ . Then,  $E^k_\lambda(z'_j) = z'_j$ . The eigenvalue map at  $\lambda$  cannot lie inside the unit circle (otherwise,  $z'_j$  would be attracting, and hence  $\lambda \in C$ ). In the latter case,  $\chi(\lambda)$  lies inside the unit circle. Thus,  $\chi(C)$  is a relatively closed set in  $D$ .

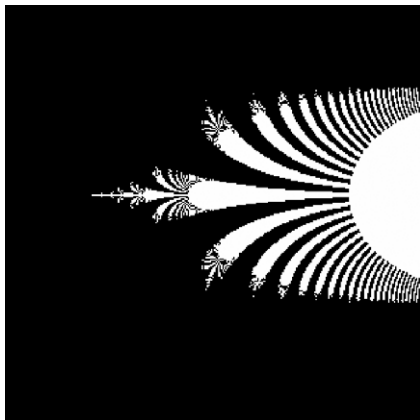


Fig. 22. The Julia set for  $\lambda = 1/e$ .

But since  $\chi$  is analytic on  $C$ , the image of  $C$  is also open. Hence  $\chi(C) = D$ , but  $\chi(\lambda) \neq 0$  if  $k \geq 2$ , since  $\chi(\lambda)$  is bounded away from zero.  $\square$

We remark that the period two region is unique in that it is the only period, other than one, for which there is a unique component in  $C_k$ . All larger periods have infinitely many components.

**PROBLEM.** Does each component of  $C_k$  have only one boundary component?

One may actually say quite a bit more about the eigenvalue map  $\chi: C_k \rightarrow D$  when  $k \geq 2$ . The proof of the following Theorem involves one of the main tools in complex dynamics, the measurable Riemann Mapping Theorem [2].

**THEOREM 4.7.** *Let  $C$  be a connected component of  $C_k$  with  $k \geq 2$ . Let  $D$  be the open unit disk in the plane. Then  $\chi: C \rightarrow D^* = D - \{0\}$  is a universal covering.*

## 5. Cantor bouquets

For quadratic polynomials, as we have seen, there are basically two types of Julia set: Cantor sets and Julia sets that are connected. For critically finite entire functions there is a similar dichotomy: either the Julia set is the entire plane or else it is nowhere dense and contains Cantor bouquets. In this section we will sketch the construction of a simple Cantor bouquet for the exponential map  $E_\lambda(z) = \lambda e^z$  where  $\lambda$  is real and satisfies  $0 < \lambda \leq 1/e$ . This construction is easily extended to the case of general  $\lambda$ . Much of the work in this section was done in collaboration with Clara Bodelon, Michael Hayes, Michal Krych, Gareth Roberts, Lisa Goldberg, and John Hubbard [14,15].

In Figure 22, we display the Julia set for  $E_{1/e}$ . The complement of the Julia set is displayed in black. It appears that this Julia set contains large open sets, but this in fact is





Fig. 23. Magnification of the Julia set for  $\lambda = 1/e$ .

not the case. The Julia set actually consists of uncountably many curves or ‘hairs’ extending to  $\infty$  in the right half-plane.

Each of the ‘fingers’ in this Figure seems to have many smaller fingers protruding from them.

As we zoom in to this image, we see more and more of the self-similar structure, as each finger generates more and more fingers. In fact, each of these fingers consists of a cluster of hairs that are packed together so tightly that the resulting set has Hausdorff dimension 2 [58].

### 5.1. The idea of the construction

Here is a rough idea of the construction of a Cantor bouquet. We will ‘tighten up’ some of these ideas in ensuing sections.

Let  $E(z) = (1/e)e^z$ . As we have seen,  $E$  has a neutral fixed point at 1 on the real axis, and  $E'(1) = 1$ . The vertical line  $\operatorname{Re} z = 1$  is mapped to the circle of radius 1 centred at the origin. In fact,  $E$  is a contraction in the half-plane  $H$  to the left of this line, since

$$|E'(z)| = \frac{1}{e} \exp(\operatorname{Re} z) < 1$$

if  $z \in H$ . Consequently, all points in  $H$  have orbits that tend to 1. Hence this half-plane lies in the Fatou set. We will try to paint the picture of the Julia set of  $E$  by painting instead its complement, the Fatou set.

Since the half-plane  $H$  is forward invariant under  $E$ , we can obtain the entire Fatou set by considering all preimages of this half-plane. Now the first preimage of  $H$  certainly contains the horizontal lines  $\operatorname{Im} z = (2k+1)\pi$ ,  $\operatorname{Re} z \geq 1$ , for each integer  $k$ , since  $E$  maps these lines to the negative real axis which lies in  $H$ . Hence there are open neighbourhoods of each of these lines that lie in the Fatou set. The first preimage of  $H$  is shown in Figure 24.

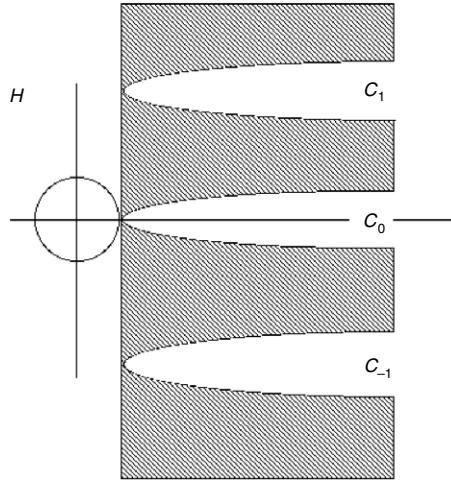


Fig. 24. The preimage of  $H$  consists of  $H$  and the shaded region.



Fig. 25. The second preimage of  $H$  in one of the fingers  $C_j$ .

The complement of  $E^{-1}(H)$  consists of infinitely many ‘fingers’. The fingers are  $2k\pi i$  translations of each other, and each is mapped onto the complementary half-plane  $\operatorname{Re} z \geq 1$ .

We denote the fingers in the complement of  $E^{-1}(H)$  by  $C_j$  with  $j \in \mathbb{Z}$ , where  $C_j$  contains the half line  $\operatorname{Im} z = 2j\pi$ ,  $\operatorname{Re} z \geq 1$ , which is mapped into the positive real axis. That is, the  $C_j$  are indexed by the integers in order of increasing imaginary part. Note that  $C_j$  is contained within the strip  $-\frac{\pi}{2} + 2j\pi \leq \operatorname{Im} z \leq \frac{\pi}{2} + 2j\pi$ .

Now each  $C_j$  is mapped in one-to-one fashion onto the entire half-plane  $\operatorname{Re} z \geq 1$ . Consequently each  $C_j$  contains a preimage of each other  $C_k$ . Each of these preimages forms a subfinger which extends to the right in the half-plane  $H$ . See Figure 25. The complement of these subfingers necessarily lies in the Fatou set.

Now we continue inductively. Each subfinger is mapped onto one of the original fingers by  $E$ . Consequently, there are infinitely many sub-subfingers which are mapped to the  $C_j$ ’s by  $E^2$ . So at each stage we remove the complement of infinitely many subfingers from each remaining finger.

This process is reminiscent of the construction of the Cantor set in Section 3.2. There we removed the complements of pairs of disks; here we remove the complement of infinitely many fingers. As a result, after performing this operation infinitely many times, we do not end up with points. Rather, the intersection of all of these fingers is a simple curve extending to  $\infty$ .

This collection of curves forms the Julia set.  $E$  permutes these curves and each curve consists of a well-defined endpoint together with a ‘hair’ that extends to  $\infty$ . It is tempting to think of this structure as a ‘Cantor set of curves’, i.e. a product of the set of endpoints and the half-line. However, this is not the case as the set of endpoints is not closed.

Note that we can assign symbolic sequences to each point on these curves. We simply see which of the  $C_j$ ’s these orbits of the point lie in after each iteration and assign the corresponding index  $j$ . That is, to each hair in the Julia set we attach an infinite sequence  $s_0s_1s_2\dots$  where  $s_j \in \mathbb{Z}$  and  $s_j = k$  if the  $j$ th iterate of the hair lies in  $C_j$ . The sequence  $s_0s_1s_2\dots$  is called the *itinerary* of the curve.

For example, the portion of the real line  $\{x \mid x \geq 1\}$  lies in the Julia set since all points (except 1) tend to  $\infty$  under iteration, not to the fixed point, and this hair has itinerary  $(000\dots)$ .

Another temptation is to say that there is a hair corresponding to every sequence  $s_0s_1s_2\dots$ . This, unfortunately, is not true either, as certain sequences simply grow too quickly to correspond to orbits of  $E$ . It is known [33] that a sequence  $(s_0s_1s_2\dots)$  is allowable in the sense that it corresponds to a point in  $J(E_\lambda)$  if and only if there exists  $x \in \mathbb{R}$  such that  $|2\pi s_j| < E^j(x)$  for each  $j$ . Here,  $E(z) = e^z$  [22,32].

So this is a rough picture of  $J(E)$ : a ‘hairy’ object extending toward  $\infty$  in the right half-plane. We call this object a *Cantor bouquet*. We will see that this bouquet has some rather interesting topological properties as we investigate further.

## 5.2. Cantor $N$ -bouquets

In this section we begin the construction of a Cantor bouquet. We first construct a Cantor set on which  $E_\lambda$  is conjugate to the shift map on  $2N + 1$  symbols.

The graph of  $E_\lambda$  (see Figure 26) shows that  $E_\lambda$  has two fixed points on the real axis, an attracting fixed point at  $q_\lambda$  and a repelling fixed point at  $p_\lambda$ , with  $0 < q_\lambda < p_\lambda$ . Note that  $q_\lambda < -\log \lambda < p_\lambda$  since

$$E'_\lambda(q_\lambda) < 1 = E'_\lambda(-\log \lambda) < E'_\lambda(p_\lambda).$$

Fix a real number  $\ell_\lambda$  satisfying  $-\log \lambda < \ell_\lambda < p_\lambda$  and observe that, if  $\operatorname{Re} z \geq \ell_\lambda$ , then

$$|E'_\lambda(z)| = \lambda e^{\operatorname{Re} z} \geq \lambda e^{\ell_\lambda} > \mu > 1$$

for some constant  $\mu > 1$ . Thus  $E_\lambda$  is ‘expanding’ in the half-plane  $\operatorname{Re} z \geq \ell_\lambda$ .

Note also that  $E_\lambda$  maps the half-plane  $\operatorname{Re} z < \ell_\lambda$  inside itself, in fact to the disk of radius  $E_\lambda(\ell_\lambda)$  centred at 0, since  $E_\lambda(\ell_\lambda) < \ell_\lambda$ . Now  $E_\lambda$  has a fixed point in this half-plane, namely  $q_\lambda$ . See Figure 27. It follows from the Corollary to the Schwarz lemma that

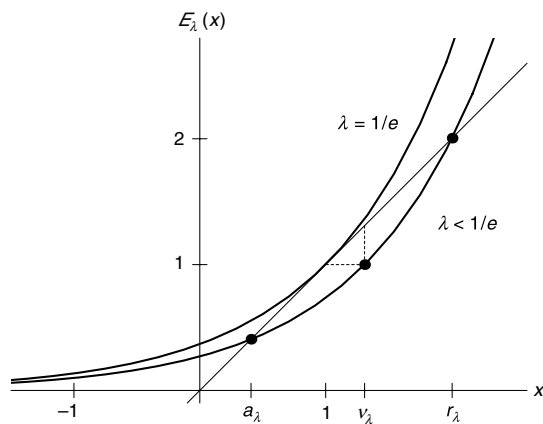


Fig. 26. The graphs of  $E_\lambda$  for  $\lambda = 1/e$  and  $\lambda < 1/e$ .

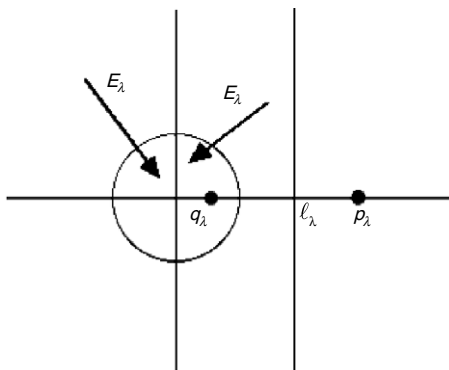


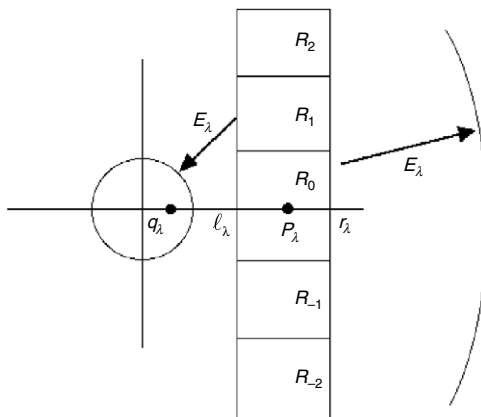
Fig. 27.  $E_\lambda$  maps the half-plane  $\operatorname{Re} z < \ell_\lambda$  inside the disk.

all orbits in this half-plane tend to  $q_\lambda$ , and so the Julia set of  $E_\lambda$  is contained in the right half-plane  $\operatorname{Re} z \geq \ell_\lambda$ .

We will now construct a collection of invariant Cantor sets for  $E_\lambda$  in the right half-plane  $\operatorname{Re} z \geq \ell_\lambda$  on which  $E_\lambda$  is conjugate to the one-sided shift map on  $2N + 1$  symbols. Fix an integer  $N \geq 1$ . Consider the rectangle  $B_N$  bounded as follows:

1. on the left by  $\operatorname{Re} z = \ell_\lambda$ ;
2. above and below by  $\operatorname{Im} z = \pm(2N + 1)\pi$ ;
3. on the right by  $\operatorname{Re} z = r_\lambda$  where  $r_\lambda$  satisfies  $\lambda e^{r_\lambda} > r_\lambda + (2N + 1)\pi$ .

The inequality in part 3 guarantees that  $E_\lambda$  maps the right-hand edge of the rectangle to a circle of radius  $\lambda e^{r_\lambda}$  centred at 0 that contains the entire rectangle  $B_N$  in its interior.

Fig. 28. Construction of the  $R_i$ .

Note also that  $E_\lambda$  maps  $B_N$  onto the annular region  $\{z \mid \lambda e^{\ell_\lambda} \leq |z| \leq \lambda e^{r_\lambda}\}$  and that  $B_N$  is contained in the interior of this annulus.

For each integer  $i$  with  $-N \leq i \leq N$ , consider the subrectangle  $R_i \subset B_N$  defined by  $\ell_\lambda \leq \operatorname{Re} z \leq r_\lambda$  and  $(2i - 1)\pi \leq \operatorname{Im} z \leq (2i + 1)\pi$ . Note that  $E_\lambda$  maps each  $R_i$  onto the annular region above, and that  $|E'_\lambda(z)| > \mu > 1$ . Moreover, if we restrict  $E_\lambda$  to the interior of  $R_i$  we obtain an expanding analytic homeomorphism that maps the interior of  $R_i$  onto a region that covers all of  $B_N$ . See Figure 28. As a consequence, we may define an analytic branch of the inverse of  $E_\lambda$ ,  $L_{\lambda,i}$ , that takes  $B_N$  into  $R_i$  for each  $i$ . Clearly,  $L_{\lambda,i}$  is a contraction for each  $i$ . In particular, there is a constant  $\nu < 1$  such that  $|L'_{\lambda,i}(z)| < \nu$  for all  $i$  and all  $z$  in  $B_N$ .

Now define

$$\Lambda_N = \{z \in B_N \mid E_\lambda^j(z) \in B_N \text{ for } j = 0, 1, 2, \dots\},$$

that is,  $\Lambda_N$  is the set of points whose orbits remain for all iterations in  $B_N$ . Let  $\Sigma_N$  denote the space of infinite sequences  $s = (s_0 s_1 s_2 \dots)$  where each  $s_j$  is an integer,  $-N \leq s_j \leq N$ . Endow  $\Sigma_N$  with the product topology. For each  $s \in \Sigma_N$ , we identify a unique point in  $\Lambda_N$  via

$$\phi(s) = \lim_{n \rightarrow \infty} L_{\lambda,s_0} \circ L_{\lambda,s_1} \circ \dots \circ L_{\lambda,s_n}(z)$$

where  $z$  is any point in  $B_N$ . The fact that  $\phi(s)$  is a unique point follows from the fact that each  $L_{\lambda,s_i}$  is a contraction in  $B_N$ . In fact  $\phi(s)$  is independent of  $z$ , and  $\phi$  defines a homeomorphism between  $\Sigma_N$  and  $\Lambda_N$ . Moreover,  $\phi$  gives a conjugacy between the shift map on the sequence space  $\Sigma_N$  and  $E_\lambda|_{\Lambda_N}$ .

Since  $|E'_\lambda(z)| > \mu > 1$  for all  $z \in \Lambda_N$ , it follows that  $\Lambda_N$  is a hyperbolic set. Moreover, we have an increasing sequence of these Cantor sets  $\Lambda_1 \subset \Lambda_2 \subset \dots$ . Since each point in  $\Lambda_N$  lies in the complement of the basin of attraction of  $q_\lambda$ , it follows that  $\Lambda_N \subset J(E_\lambda)$ . We have proved:

**THEOREM 5.1.** *Suppose  $0 < \lambda < 1/e$ . Then the set of points  $\Lambda_N$  whose orbit remains for all time in the rectangle  $B_N$  is a Cantor set in  $J(E_\lambda)$ . The action of  $E_\lambda$  on this Cantor set is conjugate to the shift map on  $2N + 1$  symbols.*

We next claim that each point  $z_s = \phi(s)$  in one of the  $\Lambda_i$  comes with a unique ‘hair’ attached. This hair is a curve associated with a natural parametrization

$$h_{\lambda,s}: [1, \infty) \rightarrow \mathbb{C}$$

that satisfies

1.  $h_{\lambda,s}(1) = z_s$ ;
2.  $h_{\lambda,s}$  is a homeomorphism;
3. If  $t \neq 1$ , then  $\operatorname{Re} E_\lambda^n(h_{\lambda,s}(t)) \rightarrow \infty$  as  $n \rightarrow \infty$ ;
4. For each  $t$ ,  $h_{\lambda,s}(t)$  lies in the horizontal strip

$$(2s_0 - 1)\pi < \operatorname{Im} z < (2s_0 + 1)\pi;$$

5.  $E_\lambda(h_{\lambda,s}(t)) = h_{\lambda,\sigma(s)}(E_{1/e}(t))$  where  $\sigma(s)$  denotes the shift applied to the sequence  $s$ .

By condition 3, the orbit of  $h_{\lambda,s}(t)$  tends to  $\infty$ . Hence this point does not lie in the basin of attraction of  $q_\lambda$  and as a consequence each point on a hair also lies in the Julia set.

To define  $h_{\lambda,s}$ , recall that  $E_{1/e}$  has a unique fixed point at 1 and that  $E_{1/e}: [1, \infty) \rightarrow [1, \infty)$ . If  $t > 1$ , then  $E_{1/e}^n(t) \rightarrow \infty$  as  $n \rightarrow \infty$ . Recall that  $L_{\lambda,s_i}$  is the branch of the inverse of  $E_\lambda$  that will now take values in the horizontal strip given by

$$(2s_i - 1)\pi \leq \operatorname{Im} z \leq (2s_i + 1)\pi.$$

Then define

$$h_{\lambda,s}(t) = \lim_{n \rightarrow \infty} L_{\lambda,s_0} \circ \cdots \circ L_{\lambda,s_{n-1}}(E_{1/e}^n(t)).$$

It can be shown that  $h_{\lambda,s}$  has all of the properties listed above. See [29].

Thus, for each  $N \geq 1$ , we have a map

$$H_\lambda: \Sigma_N \times [1, \infty) \rightarrow \mathbb{C}.$$

This map is also a homeomorphism. We call the image of  $H_\lambda$  a *Cantor  $N$ -bouquet* and denote it by  $\mathcal{B}_{\lambda,N}$ .

These Cantor  $N$ -bouquets form the skeleton of the Julia set. Indeed, every repelling periodic point has bounded itinerary, and hence lies in some  $\mathcal{B}_{\lambda,N}$  for some  $N$ . In particular, such a point lies at the endpoint of a hair in some Cantor  $N$ -bouquet (and hence any  $M$ -bouquet for  $M > N$ ). This means that the  $N$ -bouquets are dense in the Julia set. So

$$J(E_\lambda) = \text{Closure of } \bigcup_{N=1}^{\infty} \mathcal{B}_{\lambda,N}.$$

We call the closure of the union of the  $N$ -bouquets a *Cantor bouquet*.

Now there are points in the Cantor bouquet that do not lie in  $\mathcal{B}_{\lambda,N}$  for any  $N$ . Indeed, there are many points whose itineraries are unbounded. To understand these points, we need to introduce the notion of a straight brush. This is the topic of the following section.

### 5.3. Straight brushes

To describe more completely the structure of a Cantor bouquet, we introduce the notion of a *straight brush* due to Aarts and Oversteegen [1].

To each irrational number  $\zeta$ , we assign an infinite string of integers  $n_0 n_1 n_2 \dots$  as follows. We will break up the real line into open intervals  $I_{n_0 n_1 \dots n_k}$  which have the following properties

1.  $I_{n_0 \dots n_k} \supset I_{n_0 \dots n_{k+1}}$ ;
2. The endpoints of  $I_{n_0 \dots n_k}$  are rational;
3.  $\zeta = \bigcap_{k=1}^{\infty} I_{n_0 \dots n_k}$ .

Now there are many ways to do this. We choose the following method based on the Farey tree. Inductively, we first define  $I_k = (k, k+1)$ . Given  $I_{n_0 \dots n_k}$  we define  $I_{n_0 \dots n_k j}$  as follows. Let

$$I_{n_0 \dots n_k} = \left( \frac{\alpha}{\beta}, \frac{\gamma}{\delta} \right).$$

Let  $p_0/q_0 = (\alpha + \gamma)/(\beta + \delta)$  be the Farey child of  $\alpha/\beta$  and  $\gamma/\delta$ . Let  $p_n/q_n$  be the Farey child of  $p_{n-1}/q_{n-1}$  and  $\gamma/\delta$  for  $n > 0$ , and let  $p_{n-1}/q_{n-1}$  be the Farey child for  $\alpha/\beta$  and  $p_n/q_n$  for  $n \leq 0$ . We then set  $I_{n_0 \dots n_k j}$  to be the open interval  $(p_j/q_j, p_{j+1}/q_{j+1})$ .

EXAMPLE.  $I_0 = (0, 1)$ . The Farey child of  $0/1$  and  $1/1$  is  $1/2$ , so  $p_0/q_0 = 1/2$ . Then  $p_1/q_1 = \frac{1}{2} \oplus \frac{1}{1} = 2/3$ ,  $p_2/q_2 = \frac{2}{3} \oplus \frac{1}{1} = \frac{3}{4}$ , and  $p_n/q_n = n+1/n+2$  for  $n > 0$ .

For the remaining  $n$  we have

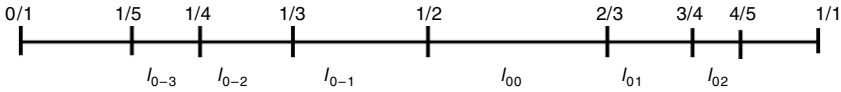
$$\begin{aligned} p_{-1}/q_{-1} &= \frac{0}{1} \oplus \frac{1}{2} = \frac{1}{3} \\ p_{-2}/q_{-2} &= \frac{0}{1} \oplus \frac{1}{3} = \frac{1}{4} \\ p_{-n}/q_{-n} &= \frac{1}{n+2}. \end{aligned}$$

Therefore, if  $n \geq 0$ ,

$$I_{0n} = \left( \frac{n+1}{n+2}, \frac{n+2}{n+3} \right)$$

and if  $n < 0$ ,

$$I_{0n} = \left( \frac{1}{-n+2}, \frac{1}{-n+1} \right).$$

Fig. 29. Construction of  $I_{0n}$ .

See Figure 29. Note that we exhaust all of the rationals via this procedure, so each irrational is contained in a unique  $I_{n_0 n_1 \dots}$ .

We now define a straight brush.

**DEFINITION 5.2.** A straight brush  $B$  is a subset of  $[0, \infty) \times \mathcal{N}$ , where  $\mathcal{N}$  is a dense subset of irrationals.  $B$  has the following three properties.

1.  $B$  is ‘hairy’ in the following sense. If  $(y, \alpha) \in B$ , then there exists a  $y_\alpha \leq y$  such that  $(t, \alpha) \in B$  if and only if  $t \geq y_\alpha$ . That is, the ‘hair’  $(t, \alpha)$  is contained in  $B$  where  $t \geq y_\alpha$ .  $y_\alpha$  is called the endpoint of the hair corresponding to  $\alpha$ .
2. Given an endpoint  $(y_\alpha, \alpha) \in B$  there are sequences  $\beta_n \uparrow \alpha$  and  $\gamma_n \downarrow \alpha$  in  $\mathcal{N}$  such that  $(y_{\beta_n}, \beta_n) \rightarrow (y_\alpha, \alpha)$  and  $(y_{\gamma_n}, \gamma_n) \rightarrow (y_\alpha, \alpha)$ . That is, any endpoint of a hair in  $B$  is the limit of endpoints of other hairs from both above and below.
3.  $B$  is a closed subset of  $\mathbb{R}^2$ .

**EXERCISE.** For any rational number  $v$  and any sequence of irrationals  $\alpha_n \in \mathcal{N}$  with  $\alpha_n \rightarrow v$ , show that the hairs  $[y_{\alpha_n}, \alpha_n]$  must tend to  $[\infty, v]$  in  $[0, \infty) \times \mathbb{R}$ .

**EXERCISE.** Show that condition 2 in the definition of a straight brush may be changed to: if  $(y, \alpha)$  is any point in  $B$  ( $y$  need not be the endpoint of the  $\alpha$ -hair), then there are sequences  $\beta_n \uparrow \alpha$ ,  $\gamma_n \downarrow \alpha$  so that the sequences  $(y_{\beta_n}, \beta_n) \rightarrow (y, \alpha)$  and  $(y_{\gamma_n}, \gamma_n) \rightarrow (y, \alpha)$  in  $B$ .

**EXERCISE.** Let  $(y, \alpha) \in B$  and suppose  $y$  is not the endpoint  $y_\alpha$ . Prove that  $(y, \alpha)$  is inaccessible in  $\mathbb{R}^2$  in the sense that there is no continuous curve  $\gamma : [0, 1] \rightarrow \mathbb{R}^2$  such that  $\gamma(t) \notin B$  for  $0 \leq t < 1$  and  $\gamma(1) = (y, \alpha)$ .

**EXERCISE.** Prove that  $(y_\alpha, \alpha)$  is accessible in  $\mathbb{R}^2$ .

These exercises show that a straight brush is a remarkable object from the topological point of view. Let’s view a straight brush as a subset of the Riemann sphere and set  $B^* = B \cup \infty$ , i.e. the straight brush with the point at infinity added. Let  $\mathcal{E}$  denote the set of endpoints of  $B$ , and let  $\mathcal{E}^* = \mathcal{E} \cup \infty$ . Then we have the following result, due to Mayer [56]:

**THEOREM 5.3.** *The set  $\mathcal{E}^*$  is a connected set, but  $\mathcal{E}$  is totally disconnected.*

That is, the set  $\mathcal{E}^*$  is a connected set, but if we remove just one point from this set, the resulting set is totally disconnected. Topology really is a weird subject!

The reason for this is that, if we draw the straight line in the plane  $(\gamma, t)$  where  $\gamma$  is a fixed rational, and then we adjoin the point at infinity, we find a disconnection of  $\mathcal{E}$ . This, however, is not a disconnection of  $\mathcal{E}^*$ . Moreover, the fact that any non-endpoint in  $B$  is inaccessible shows that we cannot disconnect  $\mathcal{E}^*$  by any other curve.



REMARK. Aarts and Oversteegen have shown that any two straight brushes are ambiently homeomorphic, i.e. there is a homeomorphism of  $\mathbb{R}^2$  taking one brush onto the other. This leads to a formal definition of a Cantor bouquet.

DEFINITION 5.4. A Cantor bouquet is a subset of  $\overline{\mathbb{C}}$  that is homeomorphic to a straight brush (with  $\infty$  mapped to  $\infty$ ).

Our main goal in this section is to sketch a proof of the following result:

THEOREM 5.5. Suppose  $0 < \lambda < 1/e$ . Then  $J(E_\lambda)$  is a Cantor bouquet.

PROOF. To construct the homeomorphism between the brush and  $J(E_\lambda)$  we first introduce symbolic dynamics. Recall that  $E_\lambda$  has a repelling fixed point  $p_\lambda > 0$  in  $\mathbb{R}$  and that the half-plane  $\operatorname{Re} z < p_\lambda$  lies in the Fatou set. Similarly the horizontal strips

$$\frac{\pi}{2} + 2k\pi < \operatorname{Im} z < \frac{\pi}{2} + (2k+1)\pi$$

are contained in the Fatou set since  $E_\lambda$  maps these strips to  $\operatorname{Re} z < 0$  which is contained in  $\operatorname{Re} z < p_\lambda$ .

We denote by  $S_k$  the closed halfstrip given by

$$\operatorname{Re} z \geq p_\lambda \quad \text{and} \quad -\frac{\pi}{2} + 2k\pi \leq \operatorname{Im} z \leq \frac{\pi}{2} + 2k\pi.$$

Note that these strips contain the Julia set since the complement of the strips lies in the Fatou set.

Given  $z \in J(E_\lambda)$ , we define the itinerary of  $z$ ,  $S(z)$ , as usual by

$$S(z) = s_0 s_1 s_2 \dots$$

where  $s_j \in \mathbb{Z}$  and  $s_j = k$  if and only if  $E_\lambda^j(z) \in S_k$ . Note that  $S(z)$  is an infinite string of integers that indicates the order in which the orbit of  $z$  visits the  $S_k$ . We will associate to  $z$  the irrational number given by the itinerary of  $z$  (and the decomposition of the irrationals described above). This will determine the hair in the straight brush to which  $z$  is mapped. See Figure 30. Thus we need only define the  $y$ -value along this hair. This takes a little work.

We will construct a sequence of rectangles  $R_k(j)$  for each  $j, k \geq 0$ . The point  $E_\lambda^j(z)$  will be contained in  $R_k(j)$  for each  $k \geq 0$ . And we will have  $R_{k+1}(j) \subset R_k(j)$  for each  $j$  and  $k$ . Each  $R_k(j)$  will have sides parallel to the axes and be contained in a strip  $S_\alpha$ . Finally each  $R_k(j)$  will have height  $\pi$ . Since the  $R_k(j)$  are nested with respect to  $k$ , the intersection  $\bigcap_{k=0}^\infty R_k(j)$  will be a nonempty rectangle of height  $\pi$  that contains  $E_\lambda^j(z)$ . We then define  $h(z)$  to be the real part of the left-hand edge of this limiting rectangle.

To begin the construction, we set  $R_0(j)$  to be the square centred at  $E_\lambda^j(z)$  with side of length  $\pi$  and contained in the appropriate strip  $S_\alpha$ . Observe that  $E_\lambda(R_0(j)) \supset R_0(j+1)$ . Indeed, the image of  $R_0(j)$  is an annulus whose inner radius is  $e^{-\pi/2}|E_\lambda^{j+1}(z)|$  and outer radius  $e^{\pi/2}|E_\lambda^{j+1}(z)|$ . Now  $e^{\pi/2} > 4$  and  $e^{-\pi/2} < 1/4$  so the image annulus is much larger than  $R_0(j+1)$ . See Figure 31.

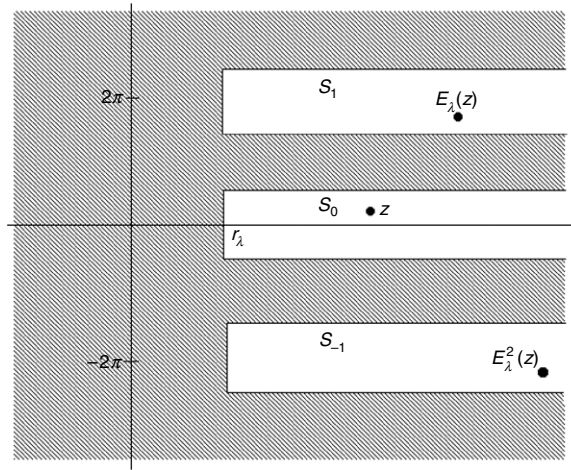


Fig. 30. The itinerary of  $z$  is  $0, 1, -1, \dots$

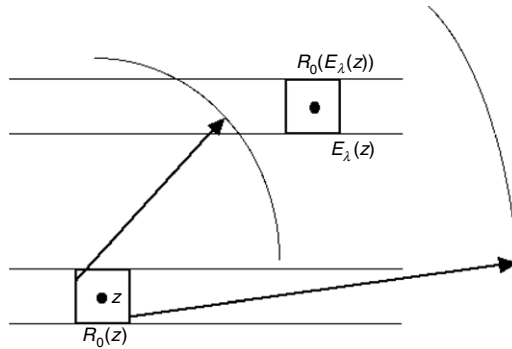
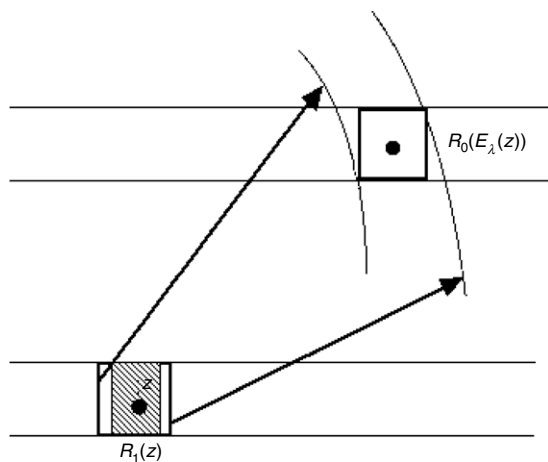
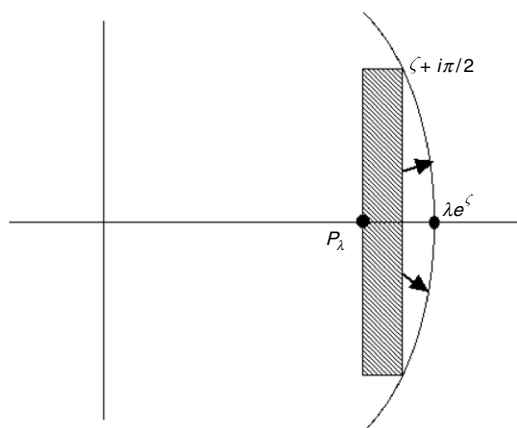


Fig. 31. Construction of  $R_0(0)$  and  $R_0(1)$ .

It follows that we may find a narrower rectangle  $R_1(j)$  strictly contained in  $R_0(j)$  having the property that the height of  $R_1(j)$  is  $\pi$  and the image  $E_\lambda(R_1(j))$  just covers  $R_0(j+1)$ . That is,  $R_1(j)$  is the smallest rectangle in  $R_0(j)$  whose image annulus is just wide enough so that  $R_0(j+1)$  fits inside. See Figure 32. Note that  $E_\lambda^j(z) \in R_1(j)$ .

Continue inductively by setting  $R_k(j)$  to be the subrectangle of  $R_{k-1}(j)$  whose image just covers  $R_{k-1}(j+1)$ . The  $R_k(j)$  are clearly nested for each fixed  $j$ .

EXAMPLE. Suppose  $z = p_\lambda$ . We have that  $R_0(j)$  is the square bounded by  $\operatorname{Re} z = p_\lambda \pm \pi/2$  and  $\operatorname{Im} z = \pm\pi/2$  for each  $j$ . One may check that, for each  $j$ ,  $\bigcap_{k=0}^{\infty} R_k(j)$  is the strip bounded by  $\operatorname{Re} z = p_\lambda$  and  $\operatorname{Re} z = \zeta$  where the circle of radius  $\lambda e^\zeta$  passes through  $\zeta \pm i\pi/2$ . See Figure 33.

Fig. 32. Construction of  $R_1(0)$ .Fig. 33. The intersection of  $R_j(0)$ .

Suppose  $z$  has itinerary  $S(z) = s_0 s_1 s_2 \dots$ . Let  $I(S(z))$  denote the irrational number determined by the sequence  $S(z)$  as above. Then set  $\phi(z) = (h(z), I(S(z)))$ . Then, as shown in [1],  $\phi$  is a homeomorphism onto a straight brush.

#### 5.4. Connectedness properties of Cantor bouquets

We call the set of endpoints of a Cantor bouquet the *crown*. Since a Cantor bouquet is homeomorphic to a straight brush with the points at  $\infty$  coinciding, it follows that any Cantor bouquet has the amazing connectedness property that the crown together with  $\infty$

is connected, but the crown alone is totally disconnected. Note also that the bouquet is nowhere locally connected (except at  $\infty$ ).

It can be shown that the construction above works for any exponential for which there exists an attracting or neutral periodic point. However, in the general case, some of the hairs in the Cantor bouquet may be attached to the same point in the crown. We will discuss this phenomenon in detail in Section 8.

McMullen [59] has shown that the Hausdorff dimension of the Cantor bouquet constructed above is 2 but its Lebesgue measure is zero. This accounts for why Figures 22 and 23 seem to have open regions in the Julia set. Also, Viana [69] has shown that each of the hairs in the Cantor bouquet is actually a  $C^\infty$  curve.

Cantor bouquets arise in many critically finite entire maps. In order to see this, suppose all singular points of  $F$  lie in some disk  $B_r$  of radius  $r$  centred at the origin. Consider the preimage  $F^{-1}(C - B_r)$  and let  $U$  be any component of this set. Now  $F$  maps  $U$  analytically onto  $C - B_r$  without singular values, so  $F$  must be a universal covering. As such,  $F$  acts like an exponential map. If, in fact,  $U$  is disjoint from  $\mathbb{C} - B_r$  and  $F$  has sufficient growth in  $U$ , it can be shown (see [31]) that there is an invariant Cantor bouquet for  $F$  in  $U$ . For a specific example dealing with the complex standard map  $z \mapsto z + \omega + \epsilon \sin z$  we refer to [40].

**EXERCISE.** Construct a similar Cantor bouquet for the map  $S_\lambda(z) = \lambda \sin z$  when  $0 < \lambda < 1$ . *Hint:* The rectangles will now be arranged vertically and there will be two bouquets: one in the upper half-plane and one in the lower.

### 5.5. Uniformization of the attracting basin

The basin of attraction  $\Omega_\lambda$  of  $E_\lambda$  is an open, dense, and simply connected subset of the Riemann sphere. Hence the Riemann Mapping Theorem guarantees the existence of a uniformization  $\phi_\lambda: D \rightarrow \Omega_\lambda$ . Given such a uniformization, it is natural to ask if the uniformizing map extends to the boundary of  $D$ .

In order to extend  $\phi_\lambda$  to the boundary, we need to show that the image of a straight ray  $re^{i\theta}$ , where  $\theta$  is constant under  $\phi_\lambda$ , converges to a single point as  $r \rightarrow 1$ . It is known that if the boundary of the uniformizing region is locally connected, then in fact  $\phi_\lambda$  does extend continuously to  $D$ . On the other hand, if the boundary of the region is not locally connected, then not all rays need converge (though a full measure set of them must converge). In our case, the boundary of  $\Omega_\lambda$  is nowhere locally connected (except at  $\infty$ ). However, it is a fact that all rays do converge. In fact, they land at precisely the endpoints of the Cantor bouquet. See [24]. This means that we can induce a map on the set of endpoints, but that map is necessarily nowhere continuous [64].

**EXERCISE.** Show that if we normalize the Riemann map  $\phi_\lambda$  so that 0 is mapped to 0, then the induced map  $\phi_\lambda^{-1} \circ E_\lambda \circ \phi_\lambda$  on the unit disk is given by

$$T_\mu(z) = \exp\left(i \left( \frac{\mu + \bar{\mu}}{1 + z} \right)\right).$$

Here  $\mu$  is a parameter that lies in the upper half-plane and depends upon  $\lambda$ .

EXERCISE. Discuss the dynamics of  $T_\mu$  on the boundary of the unit disk.

REMARK. While the Cantor bouquets constructed in this exercise are homeomorphic to those for  $E_\lambda$ , McMullen has shown [58] that the Lebesgue measure of the Cantor bouquet in the exponential case is zero whereas it is infinite in the case of the sine function. The rough reason for this is, in the case of sine, we have two Cantor bouquets, so most orbits tend to infinity by jumping back and forth between these two bouquets. So there are lots of additional itineraries. In the exponential case, orbits that tend to infinity just do so by heading to the right.

We have seen that a Cantor bouquet is a very interesting object from the topological point of view. But there is more to this story. For Karpinska has shown the following remarkable result [47]:

THEOREM 5.6. *The Julia set of  $E_\lambda$  with  $0 < \lambda < 1/e$  divides into two disjoint subsets: the ‘small’ set consisting of the endpoints and the ‘large’ set consisting of all the other points, i.e. the stems without the endpoints. The Hausdorff dimension of the set of stems is 1, but the Hausdorff dimension of the ‘smaller’ set of endpoints is 2!*

## 6. Indecomposable continua

As we have seen, the Julia set of the exponential function  $E_\lambda(z) = \lambda e^z$  is a rich structure from the topological point of view. For many values of  $\lambda$ , the Julia set contains Cantor bouquets. For other values of  $\lambda$ , the Julia set is the entire plane. Now it might appear that this type of Julia set, while quite chaotic dynamically, is ‘tame’ topologically. As we will see in this section, this is far from the truth: There are invariant sets for  $E_\lambda$  that are homeomorphic to complicated sets known as indecomposable continua. Specifically, we will investigate in this section the dynamics of  $E_\lambda$  when  $\lambda$  is real and  $\lambda > 1/e$ . As we have seen, the Julia set of  $E_\lambda$  is the entire plane in this case.

Consider the horizontal strip

$$S = \{z \mid 0 \leq \operatorname{Im} z \leq \pi\}$$

(or its symmetric image under  $z \rightarrow \bar{z}$ ). The exponential map  $E_\lambda$  takes the boundary of  $S$  to the real axis and the interior of  $S$  to the upper half-plane. Thus,  $E_\lambda$  maps certain points outside of  $S$  while other points remain in  $S$  after one application of  $E_\lambda$ . Our goal is to investigate the set of points whose entire orbit lies in  $S$ . Call this set  $\Lambda$ . The set  $\Lambda$  is clearly invariant under  $E_\lambda$ . There is a natural way to compactify this set in the plane to obtain a new set  $\Gamma$ . Moreover, the exponential map extends to  $\Gamma$  in a natural way. Our main results in this section include:

THEOREM 6.1.  *$\Gamma$  is an indecomposable continuum.*

Moreover, we will see that  $\Lambda$  is constructed in a similar fashion to a family of indecomposable continua known as *Knaster continua*. See [57].

As we will show in Section 6.2, the topology of  $\Lambda$  is quite intricate. Despite this, we will show that the dynamics of  $E_\lambda$  on  $\Lambda$  is quite tame. Specifically, we will prove:

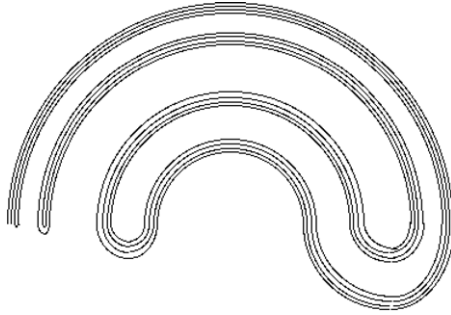


Fig. 34. The Knaster continuum.

**THEOREM 6.2.**  $E_\lambda$  has a unique repelling fixed point  $w_\lambda \in \Lambda$ , and the  $\alpha$ -limit set of all points in  $\Lambda$  is  $w_\lambda$ . On the other hand, if  $z \in \Lambda$ ,  $z \neq w_\lambda$ , then the  $\omega$ -limit set of  $z$  is either

1. the point at  $\infty$ , or
2. the orbit of 0 under  $E_\lambda$  together with the point at  $\infty$ .

Thus we see that  $E_\lambda$  possesses an interesting mixture of topology and dynamics in the case where the Julia set is the whole plane. When  $J = \mathbb{C}$ , the dynamics of  $E_\lambda$  are quite chaotic, but the overall topology is tame. On our invariant set  $\Lambda$ , however, it is the topology that is rich, but the dynamics are tame.

### 6.1. Topological preliminaries

In this section we review some of the basic topological ideas associated with indecomposable continua. See [51] for a more extensive introduction to these concepts.

Recall that a continuum is a compact, connected space. A continuum is decomposable if it is the union of two proper subcontinua (we emphasize the fact that these subcontinua are not disjoint but rather overlap). Otherwise, the continuum is indecomposable. One famous example of an indecomposable continuum is the Knaster continuum,  $K$ . One way to construct this set is to begin with the Cantor middle-thirds set sitting on the real axis in the plane between  $x = 0$  and  $x = 1$ . Then draw the semi-circles lying in the upper half-plane with centre at  $(1/2, 0)$  that connect each pair of points in the Cantor set that are equidistant from  $1/2$ . Next draw all the semicircles in the lower half-plane that have, for each  $n \geq 1$ , centres at  $(5/(2 \cdot 3^n), 0)$  and pass through each point in the Cantor set lying in the interval

$$2/3^n \leq x \leq 1/3^{n-1}.$$

The resulting set is partially depicted in Figure 34.

For a proof that this set is indecomposable, we refer to [51]. Dynamically, this set appears as the closure of the unstable manifold of Smale's horseshoe map (see [8,66]).

Note that the curve passing through the origin in this set is dense, since it passes through each of the endpoints of the Cantor set. It also accumulates everywhere upon itself. Such

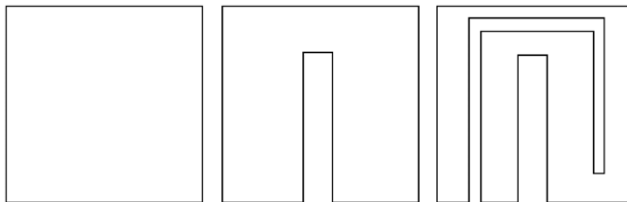


Fig. 35. A different construction of the Knaster continuum.

a phenomenon gives a criterion for a continuum to be indecomposable, as was shown by S. Curry. See his paper [18] for a proof.

**THEOREM 6.3.** *Suppose  $X$  is a one-dimensional non-separating plane continuum which is the closure of a ray that limits upon itself. Then  $X$  is indecomposable.*

Another view of the Knaster continuum which is intimately related to our own construction is as follows. Begin with the unit square  $S_0$  in the plane. Next remove a ‘canal’  $C_1$  from  $S_0$  whose boundary lies within a distance  $1/3$  of each boundary point of  $S_0$  as depicted in Figure 2. Call this set  $S_1$ . Next remove a new canal  $C_2$  from  $S_1$ . This time the boundary of  $C_2$  should be within  $1/9$  of the boundary of  $S_1$  as depicted in Figure 35. It is possible to continue this construction inductively in such a way that the resulting set is homeomorphic to the Knaster continuum.

## 6.2. Construction of $\Lambda$

Recall that the strip  $S$  is given by  $\{z \mid 0 \leq \text{Im}(z) \leq \pi\}$ . Note that  $E_\lambda$  maps  $S$  in one-to-one fashion onto  $\{z \mid \text{Im } z \geq 0\} - \{0\}$ . Hence  $E_\lambda^{-1}$  is defined on  $S - \{0\}$  and, in fact,  $E_\lambda^{-n}$  is defined for all  $n$  on  $S - \{\text{orbit of } 0\}$ . We will always assume that  $E_\lambda^{-n}$  means  $E_\lambda^{-n}$  restricted to this subset of  $S$ .

Define

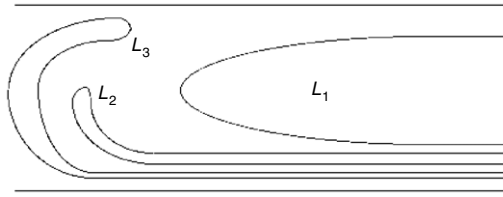
$$\Lambda = \{z \mid E_\lambda^n(z) \in S \text{ for all } n \geq 0\}.$$

If  $z \in \Lambda$  it follows immediately that  $E_\lambda^n(z) \in S$  for all  $n \in \mathbb{Z}$  provided  $z$  does not lie on the forward orbit of 0. Our goal is to understand the structure of  $\Lambda$ . Toward that end we define  $L_n$  to be the set of points in  $S$  that leave  $S$  at precisely the  $n$ th iteration of  $E_\lambda$ . That is,

$$L_n = \{z \in S \mid E_\lambda^i(z) \in S \text{ for } 0 \leq i < n \text{ but } E_\lambda^n(z) \notin S\}.$$

Let  $B_n$  be the boundary of  $L_n$ .

Recall that  $E_\lambda$  maps a vertical segment in  $S$  to a semi-circle in the upper half-plane centred at 0 with endpoints in  $\mathbb{R}$ . Either this semi-circle is completely contained in  $S$  or else an open arc lies outside  $S$ . As a consequence,  $L_1$  is an open, simply connected region which extends to  $\infty$  toward the right in  $S$  as shown in Figure 36. There is a natural

Fig. 36. Construction of the  $L_n$ .

parametrization  $\gamma_1: \mathbb{R} \rightarrow B_1$  defined by

$$E_\lambda(\gamma_1(t)) = t + i\pi.$$

As a consequence,

$$\lim_{t \rightarrow \pm\infty} \operatorname{Re} \gamma_1(t) = \infty.$$

If  $c > 0$  is large, the segment  $\operatorname{Re} z = c$  in  $S$  meets  $S - L_1$  in two vertical segments  $v_+$  and  $v_-$  with  $\operatorname{Im} v_- > \operatorname{Im} v_+$ .  $E_\lambda$  maps  $v_-$  to an arc of a circle in  $S \cap \{z \mid \operatorname{Re} z < 0\}$  while  $E_\lambda$  maps  $v_+$  to an arc of a circle in  $S \cap \{z \mid \operatorname{Re} z > 0\}$ . As a consequence, if  $c$  is large,  $v_+$  meets  $L_2$  in an open interval. Since  $L_2 = E_\lambda^{-1}(L_1)$ , it follows that  $L_2$  is an open simply connected subset of  $S$  that extends to  $\infty$  in the right half-plane *below*  $L_1$ , at least in the far right half-plane.

Continuing inductively, we see that  $L_n$  is an open, simply connected subset of  $S$  that extends to  $\infty$  toward the right in  $S$ . We may also parametrize the boundary  $B_n$  of  $L_n$  by  $\gamma_n: \mathbb{R} \rightarrow B_n$  where

$$E_\lambda^n(\gamma_n(t)) = t + i\pi$$

as before. Again

$$\lim_{t \rightarrow \pm\infty} \operatorname{Re} \gamma_n(t) = \infty.$$

Since each  $L_n$  is open, it follows that  $\Lambda$  is a closed subset of  $S$ .

**PROPOSITION 6.4.** *Let  $J_n = \bigcup_{i=n}^{\infty} B_i$ . Then  $J_n$  is dense in  $\Lambda$  for each  $n > 0$ .*

**PROOF.** Let  $z \in \Lambda$  and suppose  $z \notin B_i$  for any  $i$ . Let  $U$  be an open connected neighbourhood of  $z$ . Fix  $n > 0$ . Since  $E_\lambda^i(z) \in S$  for all  $i$ , we may choose a connected neighbourhood  $V \subset U$  of  $z$  such that  $E_\lambda^i(V) \subset S$  for  $i = 0, \dots, n$ .

Now the family of functions  $\{E_\lambda^i\}$  is not normal on  $V$ , since  $z$  belongs to the Julia set of  $E_\lambda$ . Consequently,  $\bigcup_{i=0}^{\infty} E_\lambda^i(V)$  covers  $\mathbb{C} - \{0\}$ . In particular, there is  $m > n$  such that  $E_\lambda^m(V)$  meets the exterior of  $S$ . Since  $E_\lambda^m(z) \in S$ , it follows that  $E_\lambda^m(V)$  meets the boundary of  $S$ . Applying  $E_\lambda^{-m}$ , we see that  $B_m$  meets  $V$ .



In fact, it follows that for any  $z \in \Lambda$  and any neighbourhood  $U$  of  $z$ , all but finitely many of the  $B_m$  meet  $V$ . This follows from the fact that  $E_\lambda$  has fixed points outside of  $S$  (in fact one such point is in each horizontal strip of width  $2\pi$  – see [29]), so we may assume that  $E_\lambda^m(V)$  contains this fixed point for all sufficiently large  $m$ . In particular, we have shown:

**PROPOSITION 6.5.** *Let  $z \in \Lambda$  and suppose that  $V$  is any connected neighbourhood of  $z$ . Then  $E_\lambda^m(V)$  meets the boundary of  $S$  for all sufficiently large  $m$ .*

**PROPOSITION 6.6.**  *$\Lambda$  is a connected subset of  $S$ .*

**PROOF.** Let  $G$  be the union of the boundaries of the  $L_i$  for all  $i$ . Since  $\Lambda$  is the closure of  $G$ , it suffices to show that  $G$  is connected. Suppose that this is not true. Then we can write  $G$  as the union of two closed and disjoint sets  $A$  and  $B$ . One of  $A$  or  $B$  must contain infinitely many of the boundaries of the  $L_i$ . Say  $A$  does. But then, if  $b \in B$ , the previous proposition guarantees that infinitely many of these boundaries meet any neighbourhood of  $b$ . Hence  $b$  belongs to the closure of  $A$ . This contradiction establishes the result.

We can now prove:

**THEOREM 6.7.** *There is a natural compactification  $\Gamma$  of  $\Lambda$  that makes  $\Gamma$  into an indecomposable continuum.*

**PROOF.** We first compactify  $\Lambda$  by adjoining the backward orbit of 0. To do this we identify the ‘points’  $(-\infty, 0)$  and  $(-\infty, \pi)$  in  $S$ : this gives  $E_\lambda^{-1}(0)$ . We then identify the points  $(\infty, \pi)$  and  $\lim_{t \rightarrow -\infty} \gamma_1(t)$ . This gives  $E_\lambda^{-2}(0)$ . For each  $n > 1$  we identify

$$\lim_{t \rightarrow \infty} \gamma_n(t)$$

and

$$\lim_{t \rightarrow -\infty} \gamma_{n+1}(t)$$

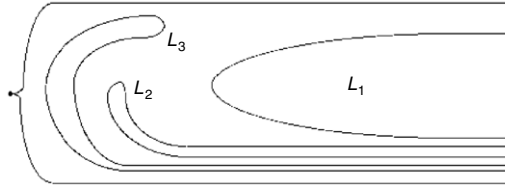
to yield  $E_\lambda^{-n-1}(0)$ . This augmented space  $\Gamma$  may easily be embedded in the plane. See Figure 37. Moreover, if we extend the  $B_i$  and the lines  $y = 0$  and  $y = \pi$  in the natural way to include these new points, then this yields a curve which accumulates everywhere on itself but does not separate the plane. See the proposition above. By the theorem of Curry [18], it follows that  $\Gamma$  is indecomposable.  $\square$

As a consequence of this theorem,  $\Lambda$  must contain uncountably many components (see [51, p. 213]). In fact, in [29] it is shown that  $\Lambda$  contains uncountably many curves.

### 6.3. Dynamics on $\Lambda$

In this section we describe completely the dynamics of  $E_\lambda$  on  $\Lambda$ .

**PROPOSITION 6.8.** *There exists a unique fixed point  $w_\lambda$  in  $S$  if  $\lambda > 1/e$ . Moreover,  $w_\lambda$  is repelling and, if  $z \in S$  – orbit of 0,  $E_\lambda^{-n}(z) \rightarrow w_\lambda$  as  $n \rightarrow \infty$ .*

Fig. 37. Embedding  $\Gamma$  in the plane.

PROOF. First consider the equation

$$\lambda e^{y \cot y} \sin y = y.$$

Since  $y \cot y \rightarrow 1$  as  $y \rightarrow 0$  and  $\lambda e > 1$ , we have  $\lambda e^{y \cot y} \sin y > y$  for  $y$  small and positive. Since the left-hand side of this equation vanishes when  $y = \pi/2$ , it follows that this equation has at least one solution  $y_\lambda$  in the interval  $0 < y < \pi/2$ .

Let  $x_\lambda = y_\lambda \cot y_\lambda$ . Then one may easily check that  $w_\lambda = x_\lambda + iy_\lambda$  is a fixed point for  $E_\lambda$  in the interior of  $S$ . Since the interior of  $S$  is conformally equivalent to a disk and  $E_\lambda^{-1}$  is holomorphic, it follows from the Schwarz Lemma that  $w_\lambda$  is an attracting fixed point for the restriction of  $E_\lambda^{-1}$  to  $S$  and that  $E_\lambda^{-n}(z) \rightarrow w_\lambda$  for all  $z \in S$ .

REMARKS. 1. Thus the  $\alpha$ -limit set of any point in  $\Lambda$  is  $w_\lambda$ .

2. The bound  $\lambda > 1/e$  is necessary for this result, as we have seen that  $E_\lambda$  has two fixed points on the real axis for any positive  $\lambda < 1/e$ . These fixed points coalesce at 1 as  $\lambda \rightarrow 1/e$  and then separate into a pair of conjugate fixed points, one of which lies in  $S$ .

We now describe the  $\omega$ -limit set of any point in  $\Lambda$ . Clearly, if  $z \in B_n$  then  $E_\lambda^{n+1}(z) \in \mathbb{R}$  and so the  $\omega$ -limit set for  $z$  is infinity.

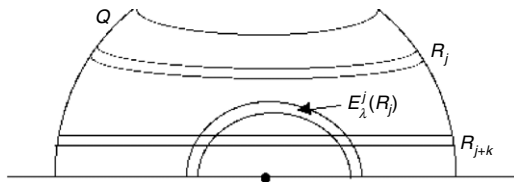
**THEOREM 6.9.** *Suppose  $z \in \Lambda$  and  $z \neq w_\lambda$ ,  $z \notin B_n$  for any  $n$ . Then the  $\omega$ -limit set of  $z$  is the orbit of 0 under  $E_\lambda$  together with the point at infinity.*

We first need a lemma.

**LEMMA 6.10.** *Suppose  $z \in \Lambda$ ,  $z \neq w_\lambda$ . Then  $E_\lambda^n(z)$  approaches the boundary of  $S$  as  $n \rightarrow \infty$ .*

PROOF. Let  $h$  be the uniformization of the interior of  $S$  taking  $S$  to the open unit disk and  $w_\lambda$  to 0. Recall that  $E_\lambda^{-1}$  is well defined on  $S$  and takes  $S$  inside itself. Then  $g = h \circ E_\lambda^{-1} \circ h^{-1}$  is an analytic map of the open disk strictly inside itself with a fixed point at 0. This fixed point is therefore attracting by the Schwarz Lemma. Moreover, if  $|z| > 0$  we have  $|g(z)| < |z|$ . As a consequence, if  $\{z_n\}$  is an orbit in  $\Lambda$ , we have  $|h(z_{n+1})| > |h(z_n)|$ , and so  $|h(z_n)| \rightarrow 1$  as  $n \rightarrow \infty$ . This completes the proof of the lemma.  $\square$

The remainder of the proof is essentially contained in [29] (see pp. 45–49). In that paper it is shown that there is a ‘quadrilateral’  $Q$  containing a neighbourhood of 0 in  $\mathbb{R}$  as

Fig. 38. The return map on  $Q$ .

depicted in Figure 38. The set  $Q$  has the following properties:

1. If  $z \in \Lambda - \bigcup_n B_n$  and  $z \neq w_\lambda$ , then the forward orbit of  $z$  meets  $Q$  infinitely often.
2.  $Q$  contains infinitely many closed ‘rectangles’  $R_k, R_{k+1}, R_{k+2}, \dots$  for some  $k > 1$  having the property that if  $z \in R_j$ , then  $E_\lambda^j(z) \in Q$  but  $E_\lambda^i(z) \notin Q$  for  $0 < i < j$ .
3. If  $z \notin \bigcup_{j=k}^\infty R_j$ , then  $z \in L_n$  for some  $n$ .
4.  $E_\lambda^j(R_j)$  is a ‘horseshoe’ shaped region lying below  $R_j$  in  $Q$  as depicted in Figure 5.
5.  $\lim_{j \rightarrow \infty} E_\lambda^j(R_j) = \{0\}$ .

As a consequence of these facts, any point in  $\Lambda$  has an orbit that meets the  $\bigcup R_j$  infinitely often. We may thus define a return map

$$\Phi: \Lambda \cap \left( \bigcup_j R_j \right) \rightarrow \bigcup_j R_j \cap \Lambda$$

by

$$\Phi(z) = E_\lambda^j(z)$$

if  $z \in R_j$ . By item 4,  $\Phi(z)$  lies in some  $R_k$  with  $k > j$ . By item 5, it follows that

$$\Phi^n(z) \rightarrow 0$$

for any  $z \in \Lambda \cap Q$ . Consequently, the  $\omega$ -limit set of  $z$  contains the orbit of 0 and infinity.

For the opposite containment, suppose that the forward orbit of  $z$  accumulates on a point  $q$ . By the lemma,  $q$  lies in the boundary of  $S$ . Now the orbit of  $q$  must also accumulate on the preimages of  $q$ . If  $q$  does not lie on the orbit of 0, then these preimages form an infinite set, and some points in this set lie on the boundaries of the  $L_n$ . But these points lie in the interior of  $S$ , and this contradicts the lemma. Thus the orbit of  $z$  can only accumulate in the finite plane on points on the orbit of 0. Since the ‘preimage’ of 0 is infinity, the orbit also accumulates at infinity. This completes the proof.  $\square$

#### 6.4. Final comments and questions

A very interesting result of Lyubich [55] asserts that the exponential map  $e^z$ , though quite chaotic, is not ergodic. Indeed, a full measure set of points have orbits that accumulate only on the orbit of 0 together with the point at  $\infty$ .

Douady and Goldberg [34] have shown that if  $\lambda, \mu > 1/e$ , then  $E_\lambda$  and  $E_\mu$  are not topologically conjugate. Each such map possesses invariant indecomposable continua  $\Lambda_\lambda$  and  $\Lambda_\mu$  in  $S$ , and the dynamics on each are similar, as shown above. However, we conjecture:

CONJECTURE. *If  $\lambda, \mu > 1/e$ ,  $\lambda \neq \mu$ , then  $\Lambda_\lambda$  and  $\Lambda_\mu$  are not homeomorphic.*

We expect that each  $\Lambda_\lambda$  yields a different Knaster-like continuum (when suitably compactified).

There are many other ways that indecomposable continua arise in the dynamics of the exponential family. For example, in [25], it is shown that, for  $0 < \lambda < 1/e$ , the set of points that correspond to uncountably many other itineraries are also indecomposable continua. These itineraries consist on infinitely many blocks of all zeroes whose length grows quickly. These sets are no longer invariant and all orbits in them accumulate on the orbit of 0 together with  $\infty$ .

In [27] other kinds of indecomposable continua were shown to exist in the case of a Misiurewicz parameter value, for example, when  $\lambda = 2\pi i$ . In the Julia set for this map, there are indecomposable continua for which a pair of ‘hairs’ accumulate on one another. There are also other instances where the set of points that share the same itinerary is an indecomposable continuum together with a separate hair that accumulates on this continuum.

## 7. The parameter plane

In this section, we continue to paint the picture of the parameter plane for the exponential family. Thus far we have concentrated on the hyperbolic components  $C_k$  in this picture, the  $\lambda$ -values for which  $E_\lambda$  has an attracting cycle of period  $k$ . Here we concentrate on the structure of the set of  $\lambda$ -values for which  $E_\lambda^n(0) \rightarrow \infty$ , so that  $J(E_\lambda) = \mathbb{C}$ . First recall the following facts about the parameter plane from Section 4:

1. The attracting fixed point region  $C_1$  is a cardioid-shaped region surrounding the origin.
2.  $C_2$  is a simply connected region filling a large portion of the left half-plane.
3. When  $k \geq 3$ ,  $C_k$  consists of infinitely many components, each of which is simply connected and extends to  $\infty$  in the right-half-plane.
4. The portion of the real axis  $(1/e, \infty)$  consists of  $\lambda$ -values for which  $E_\lambda^n(0) \rightarrow \infty$ , so that  $J(E_\lambda) = \mathbb{C}$ .

Thus we have two different types of region in the parameter plane: the hyperbolic components where the Julia set is a nowhere dense collection of tangled hairs (usually), and the line  $(1/e, \infty) \subset \mathbb{R}^+$ . Our goal in this section is to show that there are uncountably many other ‘hairs’ in the parameter plane like  $(1/e, \infty)$  on which the Julia set is the entire plane, and that, moreover, these hairs and the hyperbolic components are arranged in a rather interesting manner.

### 7.1. Structural instability

In order to underscore the complexity of the parameter plane of  $\lambda e^z$ , we first describe the dynamics of  $E_\lambda$  when  $\lambda$  lies off  $\mathbb{R}$  but very close to 1. That is, we consider perturbations

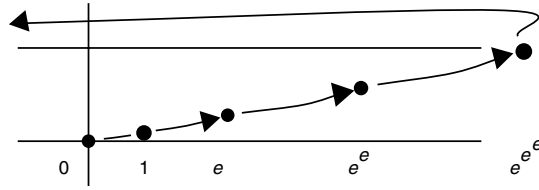


Fig. 39. The orbit of 0 under  $E_\lambda$ ,  $\lambda = e^{i\theta}$ .

of  $e^z$  within the exponential family. We shall show that, in every neighbourhood of  $\lambda = 1$ , there are parameters for which  $E_\lambda$  has an attracting cycle of period  $n$  for infinitely many different values of  $n$ , and there are also infinitely many other parameters, for which the orbit of 0 is preperiodic, of different periods (so  $J(E_\lambda) = \mathbb{C}$ ). As a consequence, the parameter plane for  $E_\lambda$  is extremely complicated near this parameter, and, in particular,  $e^z$  is (highly) structurally unstable.

As we have seen, it is the orbit of 0 that determines much of the dynamics of  $E_\lambda$ . If  $E_\lambda$  has an attracting periodic orbit, 0 is attracted to this orbit. If  $E_\lambda^n(0) \rightarrow \infty$  or if 0 is preperiodic, then  $J(E_\lambda) = \mathbb{C}$ . So let us consider the orbit of 0 for each  $\lambda = e^{i\theta}$  with  $\theta > 0$  and small. Intuitively, multiplication by  $e^{i\theta}$  gives successive points on the orbit of 0 a slight counterclockwise twist after each application of the exponential function. So we expect the orbit of 0 to 'climb' in the imaginary direction when  $\theta \neq 0$ . That being the case, it is entirely possible that, for certain  $\theta$ s, the orbit of 0 will land on the line  $\text{Im } z = \pi$ . If this occurs, then the next iterate lands in the far left-hand half-plane, and one more iterate carries 0 back extremely close to itself. See Figure 39.

We emphasize that 0 comes back extremely close to itself, for the real parts of the orbit of 0 form a sequence which is approximately given by

$$0, 1, e, e^e, e^{e^e}, \dots, E_1^{n-1}(e), -E_1^n(e), e^{-E_1^n(e)}$$

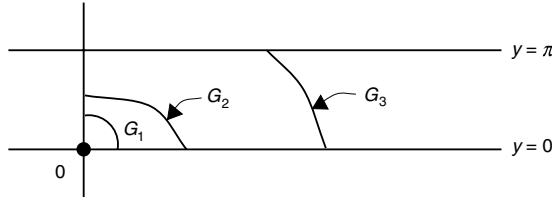
and, of course,  $e^{-E_1^n(e)}$  is extremely small. We also observe that it should be possible to select values  $\theta_n$  of the parameter with  $n$  sufficiently large so that 0 hops to the far left half-plane precisely at iteration  $n$ .

Let us make all of this a little more precise. We will write  $E_\lambda = E_\theta$ , with  $0 \leq \theta < \pi/2$ . Let  $S$  be the strip determined by  $\text{Re } z \geq 2$ ,  $0 \leq \text{Im } z \leq \pi$ . Let  $z_0 = x_0 + iy_0$ . We write

$$z_1 = z_1(\theta) = x_1 + iy_1 = E_\theta(z_0).$$

**LEMMA 7.1.** *There exists  $\theta_1$  with  $0 < \theta_1 < \pi/2$  such that, if  $0 \leq \theta \leq \theta_1$  and both  $z_0, z_1 \in S$ , then*

1.  $x_1(\theta) > 2x_0 + 1$ ,
2.  $y_1(\theta) > 2y_0$ .

Fig. 40. The curves  $G_n(\theta)$  for  $n = 1, 2, 3$ .

PROOF. First let  $\theta = 0$ . Then  $e^{x_0} \sin y_0 \leq \pi$ , so that  $\sin y_0 \leq \pi e^{-2}$ . Hence  $\cos y_0 > 0.8$ . Therefore,

$$x_1 = e^{x_0} \cos y_0 > 0.8e^{x_0} > 2x_0 + 1.5.$$

Also,

$$y_1 = e^{x_0} \sin y_0 > e^{x_0} y_0 / 2 > y_0 e^2 / 2 > 3y_0.$$

So (1) and (2) certainly hold for  $E_0$ . Now, for  $\theta \neq 0$ , we note that  $E_\theta(z) = e^{i\theta} E_0(z)$ . If  $z, E_\theta(z) \in S$ , then it follows that  $E_0(z) \in S$  as well. Hence it suffices to find  $\theta_1$  such that, if  $\theta < \theta_1$ , then

$$0 < \operatorname{Re}(w - e^{i\theta} w) < 1/2, \quad (1)$$

$$\operatorname{Im}(e^{i\theta} w) > \operatorname{Im}(w) \quad (2)$$

for all  $w \in S$  such that  $e^{i\theta} w \in S$ . (2) clearly holds if  $\theta_1 < \pi/2$ . For (1), we observe that if  $e^{i\theta} w \in S$ , then  $\operatorname{Re} w < \pi / \sin \theta$ . Therefore,

$$\begin{aligned} 0 < \operatorname{Re}(w - e^{i\theta} w) &< (\operatorname{Re} w)(1 - \cos \theta) + \pi \sin \theta \\ &< \pi(1 - \cos \theta) / \sin \theta + \pi \sin \theta. \end{aligned}$$

Since the right side approaches 0 as  $\theta \rightarrow 0$ , we may choose  $\theta_1$  small enough so that (1) holds for  $\theta < \theta_1$ .  $\square$

We remark that there exists  $\theta_2 > 0$  such that  $E_\theta^2(0) \in S$  for  $0 \leq \theta \leq \theta_2$ . From now on we assume that  $\theta < \min(\theta_1, \theta_2)$ .

Define  $G_n(\theta) = E_\theta^n(0)$ .  $G_1(\theta)$  traces out the unit circle, while  $G_2(\theta)$  gives a cardioid-like curve, part of which meets  $S$ . See Figure 40. Let  $\gamma_2(\theta)$  denote the piece of  $G_2(\theta)$  in  $S$ . Note that  $\gamma_2(\theta)$  meets  $y = 0$  at  $E_0^2(0) = e$ . Let  $\gamma_n(\theta)$  denote the connected component of  $S \cap G_n(\theta)$  that contains  $E_0^n(0)$ . For  $n$  sufficiently large (numerically,  $n \geq 3$ ),  $\gamma_n(\theta)$  connects  $y = 0$  to  $y = \pi$  in  $S$ . More precisely, for  $n \geq 3$ , there exists  $\theta_n$  such that  $\operatorname{Im}(\gamma_n(\theta_n)) = \pi$ ,  $\operatorname{Re}(\gamma_n(\theta_n)) \geq 2$ , and for all  $\theta$  with  $0 \leq \theta < \theta_n$ ,  $\gamma_n(\theta) \in S$ . See Figure 10. This can be seen by applying the Lemma repeatedly to  $E_\theta(\gamma_2(\theta))$ . If  $\theta > 0$ , there exists  $n = n(\theta)$  such that  $E_\theta^n(\gamma_2(\theta)) \notin S$ .

Clearly,  $\exp(\gamma_n(\theta))$  is a curve in the upper half-plane that connects the positive real axis to the negative real axis. Since  $0 < \theta < \pi/2$ , the curve  $E_\theta(\gamma_n(\theta)) = e^{i\theta} \exp(\gamma_n(\theta))$  also meets the negative real axis. When  $\theta = \theta_n$ , the image  $E_\theta(\gamma_n(\theta))$  is negative and real.

Denote this point by  $z_{n+1}$ , so that  $z_{n+1} = E_\theta^{n+1}(0)$ . Also let  $z_j = E_\theta^j(0) = x_j + iy_j$  for  $0 \leq j \leq n+1$ . Note that  $z_0 = 0$ ,  $|z_1| = 1$ , and  $z_j \in S$  for  $2 \leq j \leq n$ .

LEMMA 7.2. *We have  $\exp(x_n) \geq 2 + \sum_{j=1}^n (x_j + 1)$ .*

PROOF. We have  $x_2 \geq x_1 + 1$ . Moreover, by the previous Lemma, for  $2 \leq j \leq n-1$ , we have  $x_{j+1} \geq 2x_j + 1$ . Hence,

$$2 + \sum_{j=1}^n (x_j + 1) \leq \sum_{j=2}^{n-1} (x_{j+1} - x_j) + x_2 + x_n + 3 = 2x_n + 3 < e^{x_n},$$

since  $x_n \geq 2$ . □

We now construct a disk about 0 which is contracted inside itself by  $E_\theta^{n+2}$ . Let  $r_{n+1} = 1$  and define  $r_k = r_{k+1}/e^{x_k}e$  for  $0 \leq k \leq n$ . Note that  $r_j < 1$  for  $j \leq n$  and  $r_0 = (e^{n+1} \prod_{j=0}^n e^{x_j})^{-1}$ . Let  $B_j$  be the disk of radius  $r_j$  about  $z_j$ .

PROPOSITION 7.3. *If  $z \in B_0$ , then  $E_\theta^j(z) \in B_j$  for  $j \leq n+1$ . Moreover,*

$$\left| (E_\theta^{n+1})'(z) \right| \leq e^{n+1} \prod_{j=0}^n e^{x_j}.$$

PROOF. Suppose  $|z - z_j| < r_j$ . Let  $M_j = \sup |E'_\theta(z)|$ . Then, for  $j \leq n$ ,

$$|E_\theta(z) - z_{j+1}| \leq M_j r_j \leq |\lambda e^{x_j+r_j}| r_j \leq e^{x_j+r_j} \cdot r_j < e^{x_j} e \cdot r_j = r_{j+1},$$

since  $r_j < 1$ . Consequently,  $E_\theta$  maps  $B_j$  strictly inside  $B_{j+1}$ , and we have  $|E'_\theta(z)| \leq e^{x_j+1}$ . The result follows immediately. □

THEOREM 7.4.  *$E_\theta$  has an attracting periodic point of period  $n+2$  in  $B_0$ .*

PROOF. By the preceding proposition,  $E_\theta^{n+1}(B_0) \subset B_{n+1}$ . We now show that  $E_\theta(B_{n+1}) \subset B_0$ . Let  $z \in B_{n+1}$ . Then  $\operatorname{Re} z \leq x_{n+1} + 1 = -e^{x_n} + 1$ . Applying the previous Lemma, we have

$$\begin{aligned} |E_\theta(z)| &\leq \exp(-e^{x_n} + 1) \leq e \exp\left(-\sum_{j=1}^n (x_j + 1) - 2\right) \\ &= e^{-1} \left(\prod_{j=0}^n e^{x_j}\right)^{-1} e^{-n} = r_0, \end{aligned}$$

as required. Hence, there exists a periodic point  $w$  such that  $E_\theta^j(w) \in B_j$  for  $0 \leq j \leq n+1$  and  $E_\theta^{n+2}(w) = w$ .

Finally, observe that  $|E'_\theta(z)| < r_0$  if  $z \in B_{n+1}$ . Combined with the results of Proposition 7.3, this yields  $|(E_\theta^{n+2})'(w)| < 1$ . □

COROLLARY 7.5. *There is a sequence  $\theta_n \rightarrow 0$  such that, if  $n$  is sufficiently large,  $E_{\theta_n}$  has an attracting periodic orbit of period  $n$ .*

## 7.2. Other near-real parameters

There is nothing sacred about our using the initial point  $\lambda = 1$ ; one may check easily that the above arguments go through for any  $\lambda > 1/e$  and parameter values  $\lambda_\theta = \lambda e^{i\theta}$ . The result is a sequence of open sets in the  $\lambda$  plane converging to the interval  $(1/e, \infty)$ . For  $\lambda$ -values in these open sets,  $E_\lambda$  admits an attracting periodic orbit of period  $n$ .

As a corollary, we have

**COROLLARY 7.6.**  *$E_\lambda$  is not structurally stable if  $\lambda > 1/e$ .*

See [21] for more details on the above results. Douady and Goldberg [34] have proved that if  $\lambda, \mu > 1/e$ , then  $E_\lambda$  and  $E_\mu$  are not topologically conjugate. This is the case despite the fact that  $J(E_\lambda) = J(E_\mu) = \mathbb{C}$ .

There is much more to the parameter plane picture for  $E_\lambda$  when  $\lambda$  is near 1. Let us consider the functions

$$H_n(\lambda) = E_\lambda^n(0).$$

$H_n$  is a function of the parameter  $\lambda$  and tells where the  $n$ th iterate of 0 lands for each  $\lambda$ . As a consequence of our previous work, the family of functions  $\{H_n\}$  is not normal at any parameter value  $\lambda > 1/e$ . Fix  $\lambda_0 > 1/e$  and let  $U$  be any neighbourhood of  $\lambda_0$  in the  $\lambda$ -plane. Then Montel's theorem guarantees that the  $H_n$  assume all but two values in  $U$  infinitely often. These two omitted values are 0 and  $\infty$ , so, for example, there is  $\lambda \in U$  such that  $H_n(\lambda) = 2\pi i$ . But then

$$E_\lambda^{n+1}(0) = E_\lambda(2\pi i) = \lambda = E_\lambda(0).$$

Consequently, for this  $\lambda$ -value, 0 is preperiodic, and so  $J(E_\lambda) = \mathbb{C}$ . It follows that every  $\lambda > 1/e$  is a limit of points for which 0 is preperiodic.

Thus we have

**THEOREM 7.7.** *If  $\lambda > 1/e$ , there is a sequence  $\lambda_n \rightarrow \lambda$  such that 0 is preperiodic for  $E_{\lambda_n}$ , so that  $J(E_{\lambda_n}) = \mathbb{C}$ .*

We will discuss in the next section the fact that each of these  $\lambda_n$ -values also has a 'hair' attached in the  $\lambda$ -plane. This hair consists of  $\lambda$ -values for which  $E_\lambda^n(0) \rightarrow \infty$ . Consequently,  $J(E_\lambda) = \mathbb{C}$  in this case too.

This means that the picture of the  $\lambda$ -plane near  $\lambda > 1/e$  is quite complicated. There are open regions in which  $E_\lambda$  has attracting periodic points, as well as preperiodic  $\lambda$ -values with hairs attached for which  $J = \mathbb{C}$ . Thus we may continue to draw the parameter plane picture for the exponential family that we started in Section 4. In Figure 41 we see the attracting fixed point cardioid to the left. The black regions in this image represent regions where we have an attracting cycle. Note how they accumulate on the real axis, which we showed must be the case.



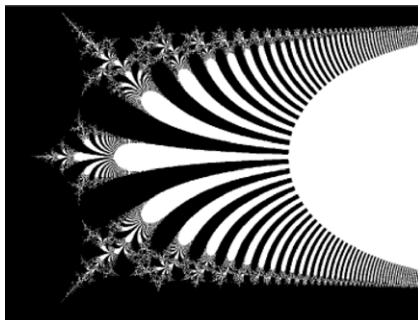


Fig. 41. Detail of the parameter plane near  $\lambda = 1/e$ . The white regions are not open sets; rather, they are collections of hairs in the parameter plane.

### 7.3. Hairs in the parameter plane

We now turn our attention to showing that there are infinitely many hairs in the parameter plane on which  $E_\lambda^n(0) \rightarrow \infty$  so that  $J(E_\lambda) = \mathbb{C}$ . We recall first the construction of the hairs in the dynamical plane. Given  $\lambda$  and an itinerary  $s = (s_0 s_1 s_2 \dots)$ , we considered the family of functions

$$G_s^n(\lambda, t) = L_{\lambda, s}^n \circ E^n(t)$$

where  $E(t) = E_{1/e}(t)$  and  $t \in [1, \infty)$ . That is,  $G_s^n(\lambda, t)$  is determined by iterating forward the model map  $E$  along the real axis, then pulling back by appropriate branches of the inverse of  $E_\lambda$ , as determined by the itinerary  $s$ . We know that this family of functions converges to a continuous function in  $t$  which parametrizes a hair in the dynamical plane, at least if  $t$  is large enough. Moreover, this function depends analytically on  $\lambda$ .

The hairs in the parameter plane consist of  $\lambda$ -values for which the orbit of 0 under  $E_\lambda$  tends to  $\infty$  with a specified itinerary. For simplicity, we will restrict attention to sequences in  $\Sigma_K$ , the set of bounded, regular itineraries, i.e. sequences of integers  $s_0 s_1 s_2 \dots$  where  $0 < |s_j| \leq K$ . So we specifically exclude  $s_j = 0$  in this section. The hairs have an endpoint that is given by a  $\lambda$ -value for which the orbit of 0 is bounded. All other  $\lambda$ -values on the hair have the property that the orbit of 0 tends to  $\infty$  with the specified itinerary. As a consequence,  $J(E_\lambda) = \mathbb{C}$  for these  $\lambda$ -values.

**DEFINITION 7.8.** Let  $s = s_0 s_1 s_2 \dots$ . A continuous curve  $H_s: [1, \infty) \rightarrow \mathbb{C}$  is called a *hair* with itinerary  $s$  if  $H_s$  satisfies:

1. If  $\lambda = H_s(t)$  and  $t > 1$ , then  $\text{Re } E_\lambda^n(0) \rightarrow \infty$  and the itinerary of  $\lambda$  under  $E_\lambda$  is  $s$ .
2. If  $\lambda = H_s(1)$ , then  $E_\lambda(0) = \lambda = z_\lambda(s)$  where  $z_\lambda(s)$  is the endpoint of the hair with itinerary  $s$  in the dynamical plane. Hence, the orbit of  $\lambda$  under  $E_\lambda$  is bounded and has itinerary  $s$ .
3.  $\lim_{t \rightarrow \infty} \text{Re } H_s(t) = \infty$ .

**REMARK.** We use the term ‘hair’ for curves in both the dynamical plane and parameter plane. When necessary, we use the terms dynamical hair and parameter hair to distinguish

between them. Our goal for the remainder of this section is to sketch the proof of the following result.

**THEOREM 7.9.** *Suppose  $s$  is a bounded sequence with no zeroes. Then there exists a hair in parameter space with itinerary  $s$ . Moreover, if  $s$  is periodic or preperiodic, then 0 is preperiodic under  $E_\lambda$  for  $\lambda = H_s(1)$ .*

The proof of the theorem depends upon several technical lemmas so, for simplicity, we merely sketch the main idea of the proof.

Given the itinerary  $s$ , we will work in a simply connected region  $Q_s$  in parameter space.  $Q_s$  will be the union of the horizontal strips  $R_\lambda(s_0)$  for  $\lambda \in \mathbb{C} - \mathbb{R}^+$ .  $Q_s$  is an open horizontal strip in  $\mathbb{C}$  with height  $4\pi$  bounded by horizontal lines  $\text{Im } z = (2s_0 - 2)\pi$  and  $\text{Im } z = (2s_0 + 2)\pi$ , so  $Q_s$  does not include  $\mathbb{R}$  since  $s_0 \neq 0$ . Given the sequence  $s$ , for each  $t$ , the dynamical hair  $h_{\lambda,s}(t)$  lies in  $R_\lambda(s_0) \subset Q_s$ . Consider the map  $F_t(\lambda) = h_{\lambda,s}(t)$ . Note that  $F_t$  is a function of  $\lambda$  and assigns to  $\lambda$  the point on the dynamical hair with itinerary  $s$  and time parameter  $t$ . It can be shown that  $F_t$  is an analytic function of  $\lambda$ . Furthermore,  $F_t$  maps the closure of  $Q_s$  strictly into its interior so therefore  $F_t$  has a unique fixed point in  $Q_s$ . This fixed point is a  $\lambda$ -value that satisfies  $\lambda = h_{\lambda,s}(t)$ , so  $\lambda = E_\lambda(0)$  lies on the hair in the dynamical plane that is attached to  $z_\lambda(s)$ . We therefore define the point  $H_s(t)$  on the hair in the parameter plane as the unique fixed point of  $F_t$  for each  $t \geq 1$ . If  $t > 1$ , it follows that  $E_\lambda^n(0) \rightarrow \infty$ , whereas, if  $t = 1$ , 0 maps after one iteration of  $E_\lambda$  onto  $z_\lambda(s)$  and so this orbit is bounded. As we vary  $t$ , the fixed point of  $F_t$  varies, and this curve of fixed points produces the hair in parameter plane.

For further details of the proof, we refer to [14].

#### 7.4. Questions and problems

We have shown that there is a unique hair in the parameter plane corresponding to any bounded, regular sequence and that this hair is attached to a  $\lambda$ -value for which the orbit of 0 is bounded. Note that if the itinerary  $s$  is periodic, the the orbit of 0 is *preperiodic*. This is true since the itinerary corresponds to the orbit of  $E_\lambda(0) = \lambda$ , not to the orbit of 0. These are the Misiurewicz points discussed in [25].

When the itinerary contains 0s, the situation is much more complicated. We conjecture that some of these are hairs land at bifurcation points in the parameter plane. We also expect that there are many non-regular hairs that land at the same point in the parameter plane, much as happens in the dynamical plane (see Section 8).

**PROBLEM.** Describe where the hairs with non-regular itineraries land in the parameter plane.

Recall that there is a hair in the parameter plane that reaches the saddle-node bifurcation point at  $\lambda = 1/e$ . This of course is the portion of the real-axis  $(1/e, \infty)$ . In every other component of  $C_k$  with  $k > 1$ , the multiplier map is a universal covering, so there are infinitely many points on the boundary of a component with multiplier 1. These are the saddle-node points. They are the visible cusps along the boundary of these components.

**PROBLEM.** Determine an algorithm that describes the set of all hairs that tend to the saddle-node points on the boundary of a component of  $C_k$  with  $k > 1$ .

**PROBLEM.** Find an algorithm that determines when two different hairs in the parameter plane land on the same parameter value.

In a sense, the hairs in the parameter plane play the same role as the external rays in the exterior of the Mandelbrot set for the quadratic family. Indeed, in each case the singular orbit tends to  $\infty$  with a specific itinerary when the parameter lies on one of these curves. Of course, we have no uniformization of a neighbourhood of  $\infty$  for  $E_\lambda$  as we do in the quadratic case, since  $\infty$  is an essential singularity rather than a superstable fixed point. Thus we expect no uniformization near  $\infty$  in parameter space as well. Of course, as we have seen in Section 4, all of the hyperbolic components (except  $C_1$ ) tend to  $\infty$  in the exponential case, whereas they are all bounded for the quadratic family.

## 8. Untangling hairs

In this section we discuss the topology of the Julia set for exponential maps for which  $E_\lambda$  has an attracting periodic orbit. As we have seen, there can be at most one attracting periodic orbit. When  $E_\lambda$  has an attracting fixed point, there is a single component of the basin of attraction and the Julia set is a Cantor bouquet. When the attracting orbit has period larger than 1, the topology of the Julia set changes. There are now infinitely many different components in the basin of the cycle. There still are invariant Cantor sets and hairs in the Julia set. However, several of these hairs may actually be attached to the same point in the Cantor set. This is what separates the various components of the basin of the cycle. When this happens, we say that the hairs are *attached* or *tied together*. We present a method in this section for ‘untangling’ these hairs. That is, we provide an algorithm that enables us to read off, using symbolic dynamics, which hairs are attached to the same point, given the itinerary of 0.

For example, in Figure 42, we display the Julia set when  $\lambda = 5 + i\pi$ . As we saw in Section 4, this exponential admits an attracting 3-cycle. In this figure we also display the Julia set when  $\lambda = 10 + 3\pi i$ . This map also has an attracting cycle of period 3. Note that different hairs now seem to be attached to one another. The symbolic dynamics will allow us to understand the differences between these two cases. In contrast, the Julia set for  $\lambda = 3.14i$  (Figure 43) shows that the structure of the attached hairs can be extremely complicated.

Our algorithm will depend on the *kneading sequence* associated with  $E_\lambda$ . The kneading sequence is a sequence of  $n - 2$  integers that specifies the topology of the basin of attraction of the attracting  $n$ -cycle (we assume that  $n > 2$  since the period 1 and 2 cases are trivial). It also allows us to codify which hairs land on which points in the Julia set. As an illustration, we will prove that, if the last integer in the kneading sequence is nonzero, then the corresponding exponential must have infinitely many distinct periodic points that have multiple hairs attached.

A similar procedure has been carried out for exponentials for which the orbit of 0 is preperiodic in collaboration with Xavier Jarque. We refer to [25] for details. Much of the

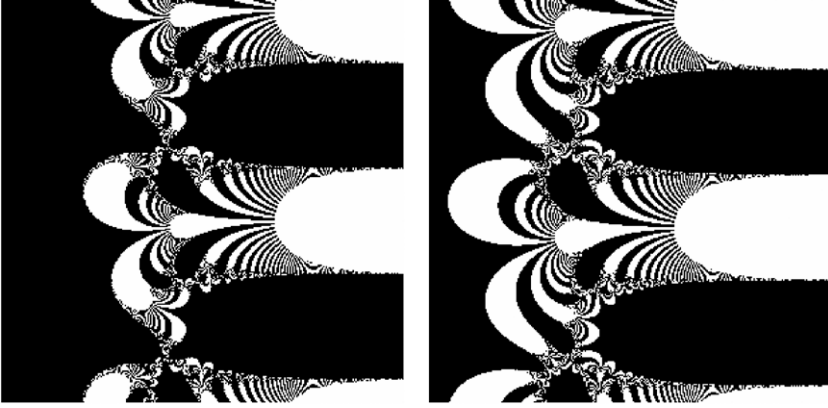


Fig. 42. The Julia sets for  $\lambda = 5 + \pi i$  and  $10 + 3\pi i$ .

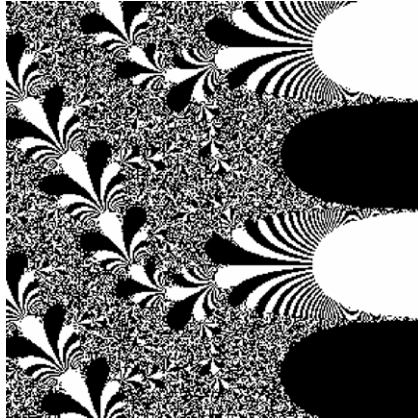


Fig. 43. The Julia set for  $\lambda = 3.14i$ .

work in the present chapter has been done jointly with Ranjit Bhattacharjee [12]. We refer to this paper for many of the details of this construction.

### 8.1. The period doubling bifurcation

We begin with a simple example of how two hairs become attached to the same endpoint. As we have seen, the exponential family undergoes a period doubling bifurcation as  $\lambda$  decreases through  $-e$ . An attracting fixed point becomes neutral when  $\lambda = -e$  and then becomes repelling. This is the story on the real axis. In the complex plane, the picture is somewhat different. Before the bifurcation, there is a repelling cycle of period 2 on either side of the real axis, and each of these points has a single hair attached. As  $\lambda$  approaches  $-e$ , these points approach the attracting fixed point, taking their hairs along with them.

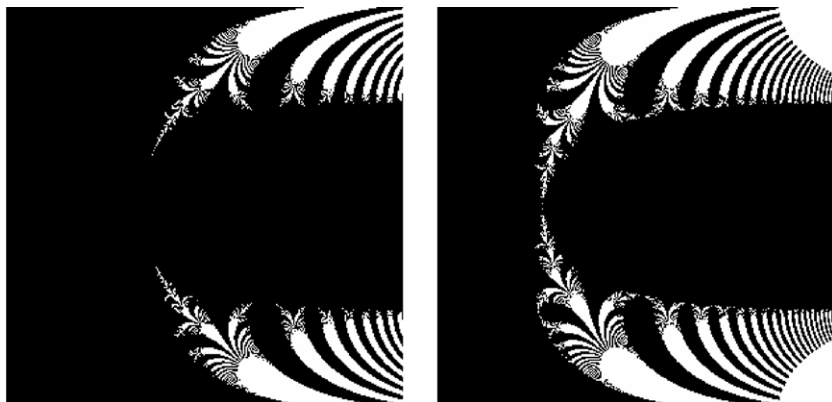


Fig. 44. The Julia sets for  $\lambda = -2.5$  and  $\lambda = -3.5$ .

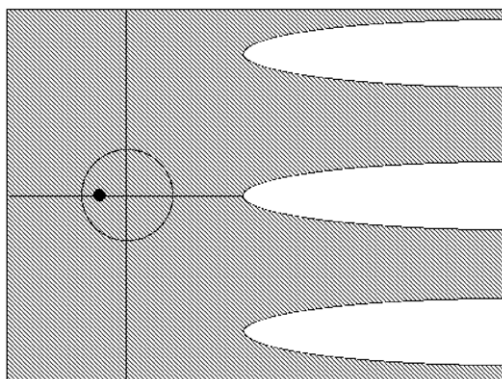


Fig. 45. The shaded region is the second preimage of  $B$  when  $\lambda > -e$ .

When  $\lambda = -e$ , the repelling 2-cycle collides with the attracting fixed point and produces the neutral fixed point. At this point, the neutral fixed point inherits the two hairs. The hairs lie inside the repelling petal at this fixed point, while portions of the real axis on either side of the fixed point lie in the attracting petals. When  $\lambda < -e$ , the new attracting 2-cycle splits away, leaving its former hairs attached to the repelling fixed point. This is a *hair transplant*. See Figure 44.

Here's the idea behind how this happens. Suppose first that  $-e < \lambda < 0$ . Consider a ball  $B$  centred at 0 that contains the attracting fixed point  $z_\lambda$ . Then the preimage of  $B$  is a half-plane that contains  $B$  and the second preimage of  $B$  is a 'glove', as depicted in Figure 45. This is precisely the picture we obtained in the case where  $0 < \lambda < 1/e$  (see Section 5), and the arguments there show that the Julia set is a Cantor bouquet.

When  $\lambda < -e$  the topology of these preimages changes. The first preimage of  $B$  is still a half-plane, but this half-plane no longer contains  $B$ . The second preimage of this half-plane is now infinitely many 'fingers', one of which necessarily contains  $B$ . Call this

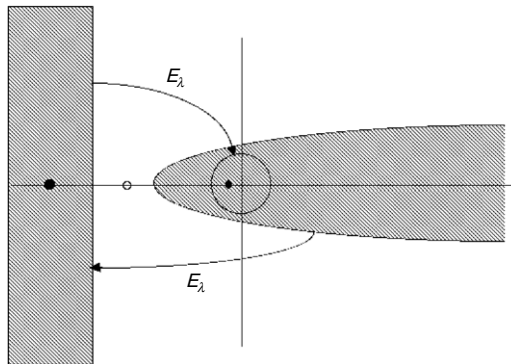


Fig. 46. The second preimage of  $B$  when  $\lambda < -e$ . There are actually infinitely many fingers that make up this preimage.

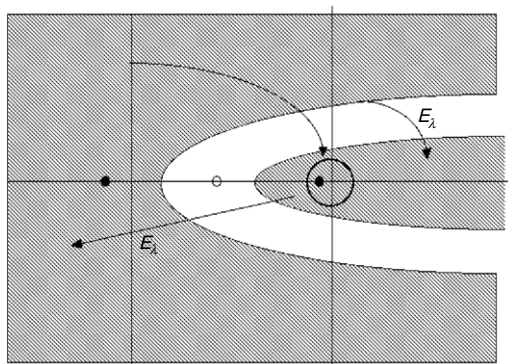


Fig. 47. The third preimage of  $B$  when  $\lambda < -e$  is a glove.

finger  $F$ . See Figure 46. If we then take the preimage of  $F$ , we see that this preimage is a glove as depicted in Figure 47. Call this glove  $G$ . Then we have  $E_\lambda^3(G) = B$ .

Now the Julia set of  $E_\lambda$  does not meet either  $F$  or  $G$ . A portion of  $J(E_\lambda)$  must be contained between  $F$  and  $G$ , as we will see below. Actually, this portion contains the pair of hairs that meet at the repelling fixed point.

## 8.2. Fingers

In this section we generalize the construction of the gloves for the period 2 case. We assume that  $E_\lambda$  has an attracting periodic cycle  $z_0, \dots, z_n = z_0$  of period  $n$ . Throughout we assume that  $n \geq 3$ .

As we have seen, the asymptotic value 0 lies in the immediate basin of attraction of some point on the cycle. Without loss of generality we will assume that  $0 \in A^*(z_1)$  where  $A^*(z)$  is the immediate basin of attraction of  $z$ . The reason for assuming  $0 \in A^*(z_1)$  rather than  $0 \in A^*(z_0)$  will become apparent soon. We will define a collection of open sets  $B_i$

about each of the  $z_i$ . Starting with the point  $z_1$  we first define a set  $B_{n+1}$  with the following properties:

1.  $B_{n+1}$  is an open and simply connected subset of  $A^*(z_1)$ ;
2.  $0, z_1 \in B_{n+1}$ ;
3.  $B_{n+1}$  has compact closure and is also a fundamental domain, i.e.  $E_\lambda^n(B_{n+1}) \subset B_{n+1}$ .

Next we will obtain a neighbourhood of  $z_0$  by considering the preimage of  $B_{n+1}$ . Define

$$B_n = E_\lambda^{-1}(B_{n+1}).$$

One checks easily that  $B_n$  is simply connected neighbourhood of  $z_0$  and  $B_n$  contains a half-plane  $\operatorname{Re} z \leq \xi_1$  and is contained in a half-plane  $\operatorname{Re} z \leq \xi_2$  for some  $\xi_1, \xi_2 \in \mathbb{R}$ .

Now we can extend this construction to all the points on the cycle. For  $j = 1, \dots, n$ , let  $B_{n-j}$  be the connected component of  $E_\lambda^{-1}(B_{n-j+1})$  that contains  $z_{n-j}$ . Note that  $B_1$  is contained in the immediate basin of  $z_1$  and  $B_1 \supset B_{n+1}$ . Indeed,  $E_\lambda^n(B_1) = B_{n+1} - \{0\}$ . We also have  $B_0 \supset B_n$  and  $E_\lambda^n(B_0) = B_n$ , and, for  $j = 1, \dots, n-1$ ,  $B_j$  is a simply connected set which is mapped univalently onto  $B_{j+1}$  by  $E_\lambda$ .

Note that  $E_\lambda : B_0 \rightarrow B_1 - \{0\}$  is a universal covering and hence this map is not univalent.

**DEFINITION 8.1.** An unbounded, simply connected  $F \subset \mathbb{C}$  is called a finger of width  $c$  if  $F$  is bounded by a simple curve  $\gamma \subset \mathbb{C}$  and there exists a  $\nu > 0$  such that  $F \cap \{z \mid \operatorname{Re} z > \nu\}$  is simply connected, extends to infinity, and satisfies

$$\{F \cap \{z \mid \operatorname{Re} z > \nu\}\} \subset \left\{z \mid \operatorname{Im} z \in \left[\xi - \frac{c}{2}, \xi + \frac{c}{2}\right]\right\}.$$

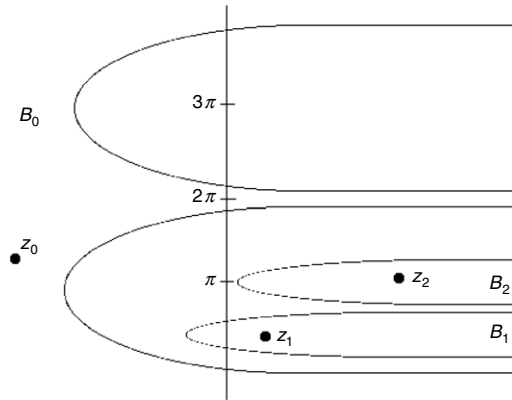
Suppose  $F$  is a finger of width  $c$  and suppose further that  $0 \notin F$ . Then one checks easily that  $E_\lambda^{-1}(F)$  consists of infinitely many disjoint fingers, each of width  $d \leq 2\pi$ . As a consequence, we have that, for  $j = 1, \dots, n-1$ ,  $B_j$  is a finger of width  $b_j \leq 2\pi$ . This construction stops at  $B_0$ , since  $B_0$  is not a finger due to the fact that  $0 \in B_1$ . Indeed, we have that the complement of  $B_0$  consists of infinitely many fingers of width  $\leq 2\pi$ . In this sense  $B_0$  resembles a ‘glove’, since it contains a left half-plane and has infinitely many fingers extending to the right. To summarize:

**THEOREM 8.2.** Suppose  $z_0, \dots, z_{n-1}$  is an attracting periodic orbit for  $E_\lambda$  with  $n \geq 3$ . Suppose  $0 \in A^*(z_1)$ . Then there exist disjoint, open, simply connected sets  $B_0, \dots, B_{n-1}$  such that

1.  $z_j \in B_j, B_j \subset A^*(z_j)$ ;
2.  $E_\lambda(B_j) = B_{j+1}, j = 0, \dots, n-2$  and  $E_\lambda(B_{n-1}) \subset B_0$ ;
3.  $B_1, \dots, B_{n-1}$  are fingers of width  $b_j \leq 2\pi$ ;
4. The complement of  $B_0$  consists of infinitely many disjoint fingers.

We say that a collection of open subsets  $B_0, \dots, B_{n-1}$  satisfying the conditions in [Theorem 8.2](#) is a *fundamental set* of attracting domains for the cycle  $z_0, \dots, z_{n-1}$ . The



Fig. 48. Fingers for  $E_\mu$ .

fingers  $B_1, \dots, B_{n-1}$  are called *stable fingers*. We remark that we may always assume that, in the far right half-plane, each of the stable fingers is bounded by a curve above and below that is nearly horizontal. See [12] for the details on this modification.

**EXAMPLE A.** Let  $\mu = a + i\pi$  where  $a$  is sufficiently large. As we saw in Section 4, the map  $E_\mu$  has an attracting cycle of period 3. Let  $D(\delta, 0)$  be a disk of radius  $\delta$  centred at the origin. One can choose  $\delta$  so that  $D(\delta, 0)$  lies in the basin of attraction of a point on the cycle. According to the above construction, we set  $B_4 = D(\delta, 0)$ . Then the  $B_j$  for  $j = 0, 1, 2$  form a fundamental set of attracting regions and are as displayed in Figure 48. Note that this picture is again a caricature of the  $B_j$ .

**EXAMPLE B.** Now let  $\nu = a + 3\pi i$  where  $a$  is sufficiently large. As we saw in Section 4,  $E_\nu$  also has an attracting cycle of period 3. In Figure 49 we sketch the location of the various  $B_j$  for  $E_\nu$ . Note that the only difference is the placement of  $B_2$  relative to the fingers in the complement of  $B_0$ .

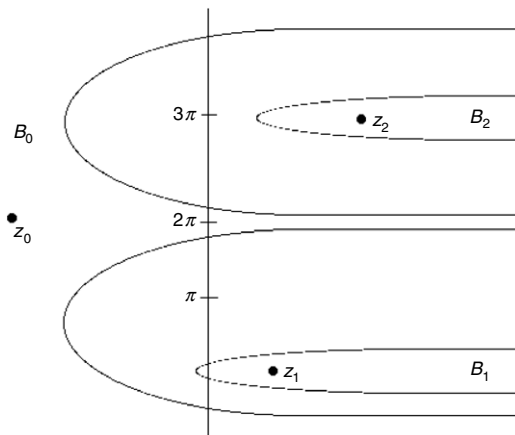
### 8.3. The kneading sequence

Using the fundamental set of attracting domains, we can now assign itineraries to each point in the Julia set. Recall that the complement of  $B_0$  consists of infinitely many closed fingers, unbounded in the right half-plane. We denote these fingers by  $\mathcal{H}_k$  where  $k \in \mathbb{Z}$ . We index the  $\mathcal{H}_k$  so that  $0 \in \mathcal{H}_0$  and so that  $k$  increases with increasing imaginary parts. Clearly,  $J(E_\lambda)$  is contained in the union of the  $\mathcal{H}_k$ .

Let  $\Sigma = \{(s) = (s_0 s_1 s_2 \dots) \mid s_j \in \mathbb{Z} \text{ for each } j\}$  be the sequence space on infinitely many symbols. As always, the *shift map*  $\sigma$  on  $\Sigma$  is given by

$$\sigma(s_0 s_1 s_2 \dots) = (s_1 s_2 s_3 \dots).$$



Fig. 49. Fingers for  $E_v$ .

We then define the *itinerary*  $S(z)$  of  $z \in J(E_\lambda)$  in the natural manner by setting by

$$S(z) = (s_0 s_1 s_2 \dots) \quad \text{where } s_j = k \text{ if and only if } E_\lambda^j(z) \in \mathcal{H}_k.$$

Note that  $S(E_\lambda(z)) = \sigma(S(z))$ .

We will be primarily concerned with itineraries whose entries are bounded. Therefore we set

$$\Sigma_N = \{s \in \Sigma \mid |s_j| \leq N \text{ for each } j\}.$$

Note that these itineraries are defined slightly differently from those encountered in Section 5.

Then, arguing exactly as in Section 5, we have:

**THEOREM 8.3.** *For each  $N > 0$  there is an invariant subset  $\Gamma_N$  of  $J(E_\lambda)$  that is homeomorphic to  $\Sigma_N$  and on which  $E_\lambda$  is conjugate to the shift map.*

So this gives a collection of Cantor sets of bounded orbits for  $E_\lambda$ ; now we need to describe how the hairs are attached to these points.

For each  $B_j$  with  $1 \leq j \leq n-1$ , there exists  $\mathcal{H}_k$  such that  $B_j \subset \mathcal{H}_k$ . We define the kneading sequence for  $\lambda$  as follows.

**DEFINITION 8.4.** Let  $E_\lambda$  have a attracting cycle of period  $n \geq 3$ . The kneading sequence is the string of  $n-2$  integers

$$K(\lambda) = k_1 k_2 \dots k_{n-2}$$

where  $k_i = j$  if and only if  $E_\lambda^i(0) \in \mathcal{H}_j$ .

Note that the kneading sequence gives the location of  $E_\lambda(0), \dots, E_\lambda^{n-2}(0)$  relative to the  $\mathcal{H}_k$ . We do not include the location of 0 since 0 always lies in  $\mathcal{H}_0$ . Similarly,  $E_\lambda^{n-1}(0)$

lies in  $B_0$ , which is the complement of the  $\mathcal{H}_k$ , and so this index is not included in  $K(\lambda)$  as well. However, we remark that, in some papers in the literature, these indices are included, so the kneading sequence begins with a 0 and ends with an \*. Equivalently, the kneading sequence indicates which  $\mathcal{H}_k$  contain the points  $z_2, \dots, z_{n-1}$  on the orbit of the cycle.

For  $\tau \gg 0$  as defined above, the set

$$\Lambda_\tau = \{z \in \mathbb{C} \mid \operatorname{Re} z \geq \tau\} - \bigcup_{j=0}^{n-1} B_j$$

consists of infinitely many closed fingers. Each finger in  $\Lambda_\tau$  is included in precisely one  $\mathcal{H}_j$  since all of the fingers in the glove  $B_0$  which bounds the  $\mathcal{H}_k$  are removed with the other  $B_j$ . If  $j$  is not one of the entries in the kneading sequence, then there is only one finger in  $\Lambda_\tau$  that lies in  $\mathcal{H}_j$  (namely the far right portion of  $\mathcal{H}_j$  itself). We denote this finger in  $\Lambda_\tau$  by  $H_j$ . However, for  $j$  in the kneading sequence, there is more than one finger in  $\Lambda_\tau$  that meets  $\mathcal{H}_j$  since the  $B_i$  separate  $\Lambda_\tau \cap \mathcal{H}_j$  into at least two fingers. The fingers that lie in such an  $\mathcal{H}_j \cap \Lambda_\tau$  will be denoted  $H_{j_k}$ , where  $j_k$  orders them with ascending imaginary part beginning with  $j_0$ . Note that all of these fingers lie in the half-plane  $\operatorname{Re} z \geq \tau$ .

EXAMPLE A. Recall the example  $E_\mu$  where  $\mu = a + i\pi$  as described in the previous section. In this case both  $B_1$  and  $B_2$  lie in  $\mathcal{H}_0$ . Since the kneading sequence only involves the location of  $B_2$  in this case, we have  $K(\mu) = 0$ . Furthermore, the fingers  $B_1$  and  $B_2$  subdivide  $\{\operatorname{Re} z \geq \tau\} \cap \mathcal{H}_0$  into three fingers which we denote by  $H_{0_0}$ ,  $H_{0_1}$ , and  $H_{0_2}$ .

EXAMPLE B. In Example B of the previous section, the kneading sequence is now  $K(\nu) = 1$ , since  $B_2$  lies in  $\mathcal{H}_1$ . Thus  $B_1$  and  $B_2$  subdivide both  $\{\operatorname{Re} z \geq \tau\} \cap \mathcal{H}_0$  and  $\{\operatorname{Re} z \geq \tau\} \cap \mathcal{H}_1$  into two subfingers, denoted by  $H_{0_0}$ ,  $H_{0_1}$ ,  $H_{1_0}$ , and  $H_{1_1}$ .

#### 8.4. Augmented itineraries

We can describe the itinerary of certain points in the Julia set even more precisely by defining an augmented itinerary for  $z \in J(E_\lambda) \cap \{z \in \mathbb{C} \mid \operatorname{Re} z \geq \tau\}$ . In an augmented itinerary, we specify which of the  $H_{j_k}$  the orbit of  $z$  visits. More precisely, let  $\mathbb{Z}'$  denote the set whose elements are either integers not contained in the kneading sequence, or subscripted integers  $j_k$  corresponding to an  $H_{j_k}$  if  $j$  is an entry in the kneading sequence. Then the *augmented itinerary* of  $z$  is

$$S'(z) = (s_0 s_1 s_2 \dots)$$

where each  $s_j \in \mathbb{Z}'$  and  $s_j$  specifies the finger in  $\Lambda_\tau$  containing  $E_\lambda(z)$ . Let  $\Sigma'$  denote the set of augmented itineraries. Of course, the augmented itinerary is defined only for points whose orbits remain for all time in  $\Lambda_\tau$ .

DEFINITION 8.5. The deaugmentation map is a map  $D : \Sigma' \rightarrow \Sigma$  such that if  $s_n = j_k$  then  $(D(s))_n = j$ . If  $s_n = j$ , then  $(D(s))_n = j$ .

That is,  $D$  simply removes the subscript from each subscripted entry in a sequence in  $\Sigma'$ , and leaves other entries alone.

It turns out that not all augmented itineraries actually correspond to orbits in the far right half-plane. In order to describe which augmented itineraries do correspond to points in  $J(E_\lambda)$ , we introduce the concept of allowable transitions.

**DEFINITION 8.6.** Let  $s = (s_0 s_1 s_2 \dots) \in \Sigma'$ . A transition is defined as any two adjacent entries  $(s_i, s_{i+1})$  in  $s$ . The transition is called allowable if

$$E_\lambda(H_{s_i}) \cap H_{s_{i+1}} \neq \emptyset.$$

In this case we say  $E_\lambda(H_{s_i})$  meets  $H_{s_{i+1}}$ . An allowable transition will be denoted as  $s_i \rightarrow s_{i+1}$ . An itinerary  $s' \in \Sigma'$  will be called allowable if for all  $s_j$  it follows that  $s_j \rightarrow s_{j+1}$ . The set of allowable itineraries will be denoted  $\Sigma^*$ .

For the remainder of this section we assume that  $N$  satisfies  $|k_j| \leq N$  for all entries  $k_j$  in the kneading sequence. Let  $\Sigma_N^*$  denote the set of sequences in  $\Sigma^*$  whose deaugmentation is a sequence in  $\Sigma_N$ . Then it can be shown that:

**PROPOSITION 8.7.** *Let  $s \in \Sigma_N^*$ . There is a unique tail of a hair in  $\Lambda_\tau \cap J(E_\lambda)$  that has augmented itinerary  $s$ .*

Thus, for each allowable sequence  $s'$  in  $\Sigma_N^*$ , we have a well defined hair in the portion of the Julia set to the right of  $\operatorname{Re} z = \tau$  that has itinerary  $s'$ . Given the hair  $h_{\lambda, \sigma(s)}(t)$ , we may pull this curve back into the region  $\operatorname{Re} z < \tau$  by applying the appropriate branch of the inverse,  $L_{\lambda, s_0}$ . The result is a curve that extends the hair  $h_{\lambda, s}(t)$  into the region  $\operatorname{Re} z < \tau$ . This follows since  $E_\lambda \circ h_{\lambda, s}(t)$  is properly contained in the hair  $h_{\lambda, \sigma(s)}(t)$  in the far right half-plane. We continue this process by applying

$$L_{\lambda, s_0} \circ \dots \circ L_{\lambda, s_n}$$

to the hair  $h_{\lambda, \sigma^{n+1}(s)}(t)$ . Each time we take an inverse, we extend the original hair. The *full hair* corresponding to the sequence  $s \in \Sigma_N^*$  is given by

$$\lim_{n \rightarrow \infty} L_{\lambda, s_0} \circ \dots \circ L_{\lambda, s_n} h_{\lambda, \sigma^{n+1}(s)}(t).$$

Then, as in the proof that  $\Gamma_N$  is a Cantor set, these full hairs each tend to a unique point in  $\Gamma_N$ . Now there is only one point in  $\Gamma_N$  that has the same non-augmented itinerary as the hair, namely the point whose deaugmented itinerary is given by  $D(s)$ . Therefore the full hair with itinerary  $s$  must terminate at this point. So we have:

**THEOREM 8.8.** *Let  $s \in \Sigma_N^*$ . The full hair corresponding to  $s$  is a curve in the Julia set that tends to  $\infty$  in the right half-plane and limits on  $\gamma_{D(s)} \in \Gamma_N$ .*

It follows from [Theorem 8.8](#) that hairs that correspond to different sequences in  $\Sigma_N^*$  that have the same deaugmentation must limit on the same point in  $\Gamma_N$ . To go back to our two examples:

EXAMPLE A. Recall that for  $E_\mu$ , the kneading sequence is  $K(\mu) = 0$  and that the region  $H_0$  contained the only two fingers  $B_1$  and  $B_2$ . These fingers subdivide  $\Lambda_\tau$  into the three fingers which we denoted by  $H_{00}$ ,  $H_{01}$ , and  $H_{02}$ .

Hence there are three full hairs in  $H_0$ , one tending to  $\infty$  in each of these three fingers. As we will see in the next section, all of these hairs have deaugmented sequence  $(000\dots)$ . Hence, by [Theorem 8.3](#), each of these hairs must be attached to  $\gamma_s$  with  $s = (000\dots)$ , which is a fixed point for  $E_\lambda$ . Furthermore, any preimage of  $\gamma_s$  must have three hairs attached, by invariance of the Julia set. These triple attachments are clear in [Figure 42](#), which shows  $J(E_\mu)$ .

EXAMPLE B. For the map  $E_\nu$ , the kneading sequence is  $K(\nu) = 1$  and we have two fingers,  $B_1 \subset H_0$  and  $B_2 \subset H_1$ . In  $H_0$  we have two fingers  $H_{00}$  and  $H_{01}$ , and there are two in  $H_1$  with indices  $1_0$  and  $1_1$ . Each of these fingers contains a hair, and we will see that the pair in  $H_0$  is attached to a point of period 2 with itinerary  $(010101\dots)$ , while the pair in  $H_1$  is attached to the point with itinerary  $(101010\dots)$ . These, as well as many other attachments, are visible in [Figure 42](#). Note the rather obvious difference between  $J(E_\nu)$  shown in this figure compared to  $J(E_\mu)$ .

### 8.5. Untangling the hairs

In this section, we show how to determine when two hairs are attached to the same point in the Julia set. By [Proposition 8.7](#), if we have an allowable itinerary in  $s' \in \Sigma_N^*$ , then there is a unique tail of a hair in  $J(E_\lambda)$  with that itinerary. If an augmented sequence is not allowable, then there is no such tail of a hair. Then, using [Theorem 8.3](#), we can pull each of these hairs back until it lands at a point in  $\Gamma_N$ . The landing point is then given by the point whose deaugmented itinerary is  $D(s')$ . Therefore, to determine whether we have more than one hair attached to a given point, all we need to do is to determine when we have multiple allowable augmented sequences, each of which has the same deaugmentation. This reduces the geometry of the hairs to a combinatorial problem, as we show below.

Our main tool is the following lemma.

LEMMA 8.9. *Let  $s_0, s_1, \dots, s_j \in \mathbb{Z}$ . Let  $s'_j \in \mathbb{Z}'$  with  $D(s'_j) = s_j$ . Then there is a unique sequence  $s'_0, s'_1, \dots, s'_{j-1}$  such that*

1.  $D(s'_i) = s_i$  for  $i = 0, 1, \dots, j-1$ .
2. The transitions

$$s'_0 \rightarrow s'_1 \rightarrow \dots \rightarrow s'_j$$

*are all allowable.*

PROOF. Suppose that  $i_j \rightarrow k_\ell$ . Recall that this means that  $E_\lambda(H_{i_j})$  meets  $H_{k_\ell}$  in the far right half-plane. Equivalently, we must have

$$L_{\lambda,i}(H_{k_\ell}) \cap \Lambda_\tau \subset H_{i_j}.$$

Now if  $i_m \rightarrow k_\ell$  also, we must have  $E_\lambda(H_{i_m})$  meets  $H_{k_\ell}$  in the far right half-plane as well. But both  $H_{i_j}$  and  $H_{i_m}$  are contained in  $H_i$  and  $E_\lambda$  is injective on  $H_i$ . Hence there can be at most one allowable transition of the form  $i_* \rightarrow k_\ell$ . This shows that the sequence above is unique, if it exists.

To see that there is a transition  $i_j \rightarrow k_\ell$ , recall that  $E_\lambda(H_i)$  covers  $\mathbb{C} - B_1$ . Hence  $E_\lambda(H_i)$  meets all of the fingers in  $\Lambda_\tau$ . In particular, there is a subfinger in  $\Lambda_\tau \cap H_i$  that maps over  $H_{k_\ell}$  in the far right half-plane. This proves existence.  $\square$

Thus, according to this lemma, given any  $s_j \in \mathbb{Z}'$ , we can find one and only one initial portion of an allowable sequence whose  $j$ th entry is  $s_j$ . Therefore we have:

**COROLLARY 8.10.** *Suppose  $s \in \Sigma'_N$  contains infinitely many entries that are nonsubscripted. Then there is at most one hair corresponding to this sequence.*

**COROLLARY 8.11.** *The only points in  $\Gamma_N$  that can have multiple hairs attached are those*

1. *whose itineraries consist only of subscripted entries in  $\mathbb{Z}'$ , or*
2. *are preimages of such points.*

Therefore, to determine which hairs are attached to which points in  $\Gamma_N$ , we need only consider allowable sequences that consist entirely of subscripted entries. These allowable sequences together with their preimages are the only sequences that may have multiple hairs attached. So we have reduced the question to: Which sequences  $s' \in \Sigma_N^*$  with only subscripted entries have the property that there is a second sequence  $t'$  with  $D(s') = D(t')$ ? We will describe the algorithm for determining this after returning to our examples.

**EXAMPLE A.** In this case we consider  $E_\lambda(z) = \mu e^z$  where  $\mu = 5 + i\pi$ . We have  $K(\mu) = 0$ . By the previous corollary, the only points in  $\Gamma_N$  that may have multiple hairs attached are those whose itineraries end  $(s_0 \dots s_n \bar{0} \dots)$ . That is, only the single (repelling) fixed point in  $H_0$  (and its preimages) can have multiple hairs attached.

We claim that there are exactly three hairs attached to each such point. To determine this, we need to ask which sequences in  $\Sigma_N^*$  have deaugmentation  $(000 \dots)$ . This in turn is determined by the allowable transitions among the  $0_j$ .

For  $E_\mu$ , the allowable entries in a sequence in  $\Sigma_N^*$  are  $0_0, 0_1, 0_2$  and all nonzero integers. The way the corresponding fingers are mapped show that the transition rules among these entries are:

1.  $0_0 \rightarrow 0_1$ ;
2.  $0_1 \rightarrow 0_2, k \geq 1$ ;
3.  $0_2 \rightarrow 0_0, k \leq -1$ ;
4.  $j \rightarrow k, 0_0, 0_1, 0_2$ , for any two nonzero integers  $j$  and  $k$ .

As a consequence, the only three allowable sequences consisting of only the  $0_j$  are

1.  $(\overline{0_0 0_1 0_2})$
2.  $(\overline{0_1 0_2 0_0})$
3.  $(\overline{0_2 0_0 0_1})$ .

Hence we have:

**THEOREM 8.12.** *For  $\lambda = \mu$ , the only points in  $\Gamma_N$  with multiple hairs attached are the fixed point with itinerary  $(000\dots)$  and all of its preimages. Each of these points has exactly three hairs attached. All other points have a single hair attached.*

Notice that we can capture the information about these hairs in matrix form using a *transition matrix*. In this matrix, the  $(i, j)$  entry is either 0 or 1 depending on whether  $i \rightarrow j$  is either not allowed or allowed. Here the rows and columns of the matrix are specified by the subscripted entries in  $\mathbb{Z}'$ . In this case, the transition matrix involves the entries  $0_0, 0_1$ , and  $0_2$  and is given by

$$T_\mu = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

**EXAMPLE B.** Now recall the function  $E_\lambda(z) = ve^z$  where  $v = a + 3\pi i$  where  $a$  is sufficiently large. In this case  $B_1$  lies in  $H_0$  but  $B_2$  now lies in  $H_1$ . So  $K(v) = 1$ . Therefore the relevant entries in  $\Sigma_N^*$  are  $0_0, 0_1, 1_0$ , and  $1_1$  and we need only consider sequences involving just 0s and 1s. One again checks easily that the transition rules among these entries are:

1.  $0_0 \rightarrow 0_1, 1_0$ ;
2.  $0_1 \rightarrow$  all others;
3.  $1_0 \rightarrow 0_1, 1_0, 1_2, k > 0$ ;
4.  $1_1 \rightarrow$  all others.

Thus the transition matrix now involves the four subscripted entries in  $\Sigma_N^*$  and is given by:

$$T_v = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The hair structure for  $E_v$  is much different from that of  $E_\mu$ . For example, the period 2 transitions

$$\begin{aligned} 0_0 &\rightarrow 0_1 \rightarrow 0_0 \\ 0_1 &\rightarrow 0_0 \rightarrow 0_1 \end{aligned}$$

are both allowable. Also, the transitions

$$\begin{aligned} 0_0 &\rightarrow 1_0 \rightarrow 0_1 \\ 0_1 &\rightarrow 1_1 \rightarrow 0_0 \end{aligned}$$

are allowable. Let  $\alpha$  denote the pair  $0_00_1$  and  $\beta$  the opposite pair  $0_10_0$ . Then we can string together any number of  $\alpha$ 's, say  $k$ , follow it with a  $1_1$  and then repeat periodically and we

obtain an allowable sequence in  $\Sigma_N^*$ . Similarly, the same number of  $\beta$ 's followed by a  $1_0$  and then repeated periodically is also allowable. But both of these sequences have the same deaugmentation, namely

$$(\overline{0 \dots 01})$$

with  $2k$  0s in each repeating block. Hence the hairs corresponding to each of these sequences are attached to a periodic point of period  $2k + 1$ .

Now none of these periodic points are preimages of each other. So, unlike the case of  $E_\mu$ , we have infinitely many distinct periodic points with multiple hairs attached. Of course, each of their infinitely many preimages also has a pair of hairs attached.

REMARK. Multiple hairs can be attached to nonperiodic points as well. For example, let  $\alpha = 0_00_1$  and  $\beta = 0_10_0$ . Then we have the following allowable sequences

$$\begin{aligned} \alpha 1_1 \alpha \alpha 1_1 \alpha \alpha 1_1 \dots \\ \beta 1_0 \beta \beta 1_0 \beta \beta 1_0 \dots \end{aligned}$$

Note that each of these sequences has the same nonperiodic deaugmentation.

By analogy with the previous example, one can prove more generally that:

**THEOREM 8.13.** *Suppose that  $K(\lambda) = k_1 k_2 \dots k_{n-2}$  where  $k_{n-2} \neq 0$ . Then the corresponding exponential has the property that there are infinitely many distinct periodic points that have multiple hairs attached.*

For the proof, we refer to [12]. This theorem is by no means optimal. A natural problem is to determine exactly which kneading sequences lead to infinitely many attachments.

In the case of the quadratic family, hair attachments are the same as external rays that meet at a common landing point. There are infinitely many such distinct attachments whenever the parameter  $c$  is drawn from any portion of the Mandelbrot set that is not connected to the main cardioid by a finite sequence of bifurcations. Perhaps the same is true for the exponential parameter space.

## 8.6. Back to the parameter plane

We can use the kneading sequences described in this section to begin to describe the structure of the parameter plane for  $E_\lambda$ . For it is known that all parameters in a given hyperbolic component have the same kneading sequence. Thus we may associate a string of  $n - 2$  integers to any hyperbolic component of period  $n$ . For technical reasons, in the parameter plane, it is customary to precede this sequence with a 0 (if the period is greater than 1) and to follow it with an asterisk. So the fixed point component has kneading sequence  $*$ , the period 2 component has kneading sequence  $0*$ , and the period three components have kneading sequences  $0k*$  where  $k \in \mathbb{Z}$ . Some of these components are displayed in Figure 50.

One of the main results regarding the structure of the parameter plane is the following [23].

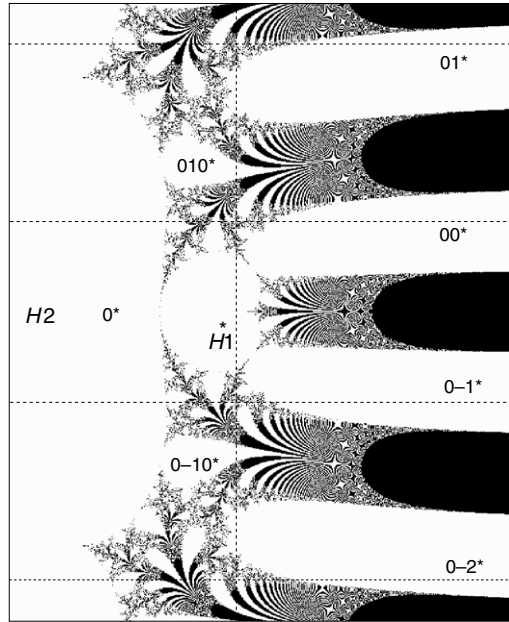


Fig. 50. The parameter plane indicating some of the kneading sequences of period 3 and 4 hyperbolic components.  $H1$  represents the period 1 region and  $H2$  the period 2 component.

**THEOREM 8.14.** *Fix  $n \geq 3$  and let  $s_1, \dots, s_{n-2} \in \mathbb{Z}$ . There exists a hyperbolic component  $\Omega_{0s_1 \dots s_{n-2}^*}$  that extends to  $\infty$  in the right half-plane and such that if  $\lambda \in \Omega_{0s_1 \dots s_{n-2}^*}$ , the map  $E_\lambda$  has an attracting cycle of period  $n$  with parameter plane kneading sequence  $s = 0s_1 \dots s_{n-2}^*$ . Moreover, the components  $\Omega_{0s_1 \dots s_{n-2}^*}$  are ordered lexicographically in the far right half-plane.*

In particular, it follows that there are infinitely many hyperbolic components of period  $n$  for each  $n \geq 3$ . From the proof of this theorem one obtains the following corollary (see Figure 51).

**COROLLARY 8.15.** *Let  $\Omega_{0s_1 \dots s_{n-2}^*}$  be as in Theorem A. Then between this hyperbolic component and the hyperbolic component  $\Omega_{0s_1 \dots (s_{n-2}+1)^*}$  there exist hyperbolic components  $\Omega_{0s_1 \dots (s_{n-2}+1)^* k^*}$  for each  $k \in \mathbb{Z}$ .*

In this statement the word ‘between’ refers to the ordering given by the imaginary part, since all hyperbolic components of period 3 or higher extend to infinity in the right half-plane.

## 9. Back to polynomials

At this juncture, there appears to be little similarity between the parameter plane for  $E_\lambda$  and the Mandelbrot set, the parameter plane for the maps  $z \rightarrow z^2 + c$ . It is true that each



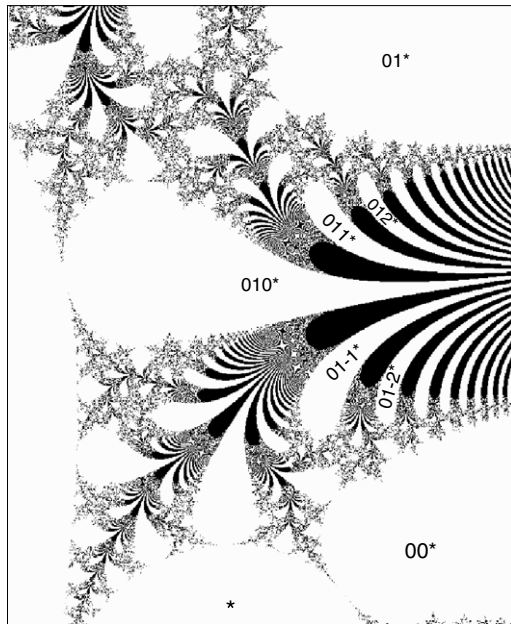


Fig. 51. A magnification of the parameter plane showing infinitely many hyperbolic components between two period 3 components.

family is a ‘natural’ one parameter family of maps each of which has a one singular value, and both parameter planes feature a central cardioid-like region in which the associated maps have an attracting fixed point. But there the similarity seems to end. In this section, however, we will show that there is a natural connection between the two sets.

### 9.1. The polynomial family

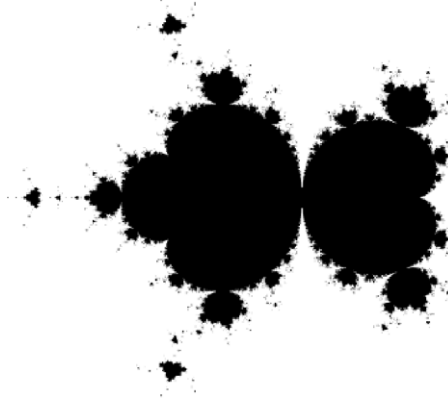
The connection is given by the family of maps

$$P_{d,\lambda}(z) = \lambda \left(1 + \frac{z}{d}\right)^d.$$

$P_{d,\lambda}$  is a polynomial of degree  $d$  which has a single critical point, at  $-d$ , and a single critical value at 0. Of course,  $P_{d,\lambda}$  converges uniformly on compact subsets to  $E_\lambda$ . But the convergence is dynamical as well.

Let  $Q_{d,c}(z) = z^d + c$ .  $Q_{d,c}$  has a single critical point at 0 and critical value  $c$ . Hence we may construct the parameter plane for  $Q_{d,c}$  just as in the case of the quadratic family (as in [37]). Define

$$v_c(z) = d \left( \frac{z}{c} - 1 \right).$$

Fig. 52.  $d = 4$ .

Then one may check easily that the affine map  $v_c$  conjugates  $Q_{d,c}$  to  $P_{d,\lambda}$ , where  $c$  is any of the  $d - 1$  choices for which  $\lambda = dc^{d-1}$ , provided  $c \neq 0$ . Thus  $v_c$  gives a ramified covering map from the parameter plane for  $Q_{d,c}$  to the parameter plane of  $P_{d,\lambda}$ . This is a  $d - 1$ -fold covering, ramified only at 0. Let  $B_d$  denote the parameter plane for the  $P_{d,\lambda}$ , i.e. the analogue of the Mandelbrot set for these maps. Then, arguing as in [37] as extended to this case in [63], we have

**THEOREM 9.1.**  *$B_d$  is connected and the complement of  $B_d$  in  $\overline{\mathbb{C}}$  is an open disk.*

The Fundamental Dichotomy for quadratic maps holds for the  $P_{d,\lambda}$  as well. That is, if  $\lambda \in B_d$ , then the filled Julia set of  $P_{d,\lambda}$  is connected, whereas, if  $\lambda \in \mathbb{C} - B_d$ , then

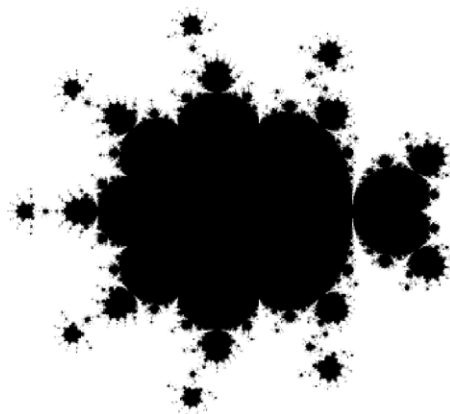
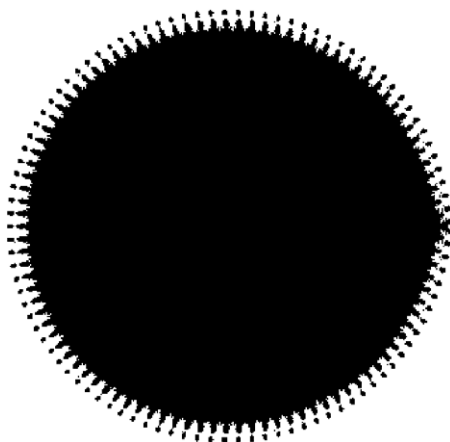
$$P_{d,\lambda}^n(0) \rightarrow \infty$$

and so  $J(P_{d,\lambda})$  is a Cantor set. See [13].

In Figures 52–54, we have displayed  $B_d$  for  $d = 4$  and  $d = 8$  and  $d = 100$ . The size of these images in the parameter plane varies. For example, when  $d = 4$ , we display a box of side length 6, but when  $d = 100$ , the image is a box of side length 200. Each of these images features a main cardioid of roughly the same size (it is hardly visible on the right in Figure 54, so we have magnified this region in Figure 55). Note that in this magnification, the large protuberances heading to the right are the period three regions that tend to the period three regions straddling the lines  $\text{Im } z = \pm\pi$  in the exponential parameter plane.

Note also how the period 2 region just to the left of the main cardioid also grows with  $d$ . Similarly, the number of cusps on the ‘decorations’ attached to the basic cardioid grows with  $d$ . Indeed, as  $d \rightarrow \infty$ , these figures ‘converge’ to the parameter plane of  $E_\lambda$ . This is what provides the link between the quadratic and exponential families.

**EXERCISE.** Suppose  $\lambda \in C_k$  for the exponential family. Show that there is  $d_0$  such that, if  $d > d_0$ , then  $P_{d,\lambda}$  has an attracting cycle of period  $k$ .

Fig. 53.  $d = 8$ .Fig. 54.  $d = 100$ .

Note that there is a fundamental difference between the  $B_d$  and the parameter plane for  $E_\lambda$  which we now write as  $B_\infty$ . For each  $d$ ,  $B_d$  is compact and its exterior is isomorphic to the unit disk. It is easier to work with the family  $z^d + c$  to prove this, but then the parameter plane for this family of maps is a ramified cover (ramified  $d - 1$  times over 0) of that of  $P_{d,\lambda}$ . On the other hand, as we showed in 4, each component of  $C_k$ ,  $k \geq 2$ , in  $B_\infty$  is non-compact, and the ‘exterior’ of  $B_\infty$  contains the hairs described above. (Actually, the exterior of  $B_\infty$  is not defined since any hair in the  $\lambda$ -plane is a limit of components of the  $C_k$ .) For more information about the convergence of hyperbolic components, we refer to [50].

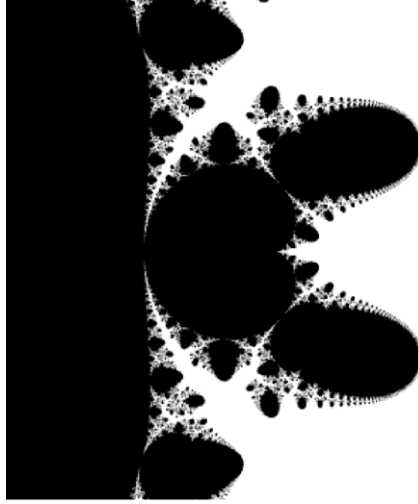


Fig. 55. A magnification of the main cardioid in  $B_\lambda$  when  $d = 100$ .

## 9.2. External rays

One natural question is what happens to the external rays of Douady and Hubbard which foliate the disk in complement of  $B_d$ ? In [15], it is shown that certain of these rays have a limit as  $d \rightarrow \infty$ , and that this limit is precisely one of the hairs in  $B_\infty$ .

How do we identify the hairs which are limit curves of the external rays? The precise answer is spelled out in [15], but we will give some special cases here. The external arguments of Douady and Hubbard are given by rational numbers between 0 and 1. For example, for each  $d$ , the external argument of angle  $1/d$  converges to a preperiodic  $\lambda$ -value in  $B_d$  which is attached to the first ‘bulb’ of the large period 2 region. These angles are depicted in Figure 56.

Note that, in base  $d$ , we may write

$$\frac{1}{d} = (d-1) \sum_{i=1}^{\infty} \frac{1}{d^{i+1}}.$$

The result in [15] is that this ray converges as  $\lambda \rightarrow \infty$  to the hair on which  $\lambda$  has itinerary (111...) (This is the hair which terminates at  $2\pi$ .)

In general, if for each  $d$ ,  $p(d)$  is the angle given by in base  $d$

$$p(d) = (d-1) \sum_{i=1}^{\infty} \frac{s_i}{d^{i+1}}$$

where the  $s_i$  form a repeating, regular sequence, then the rays with angle  $p(d)$  in  $B_d$  converge as  $d \rightarrow \infty$  to the hair on which all  $\lambda$ 's have itinerary  $(s_1 s_2 s_3 \dots)$ .

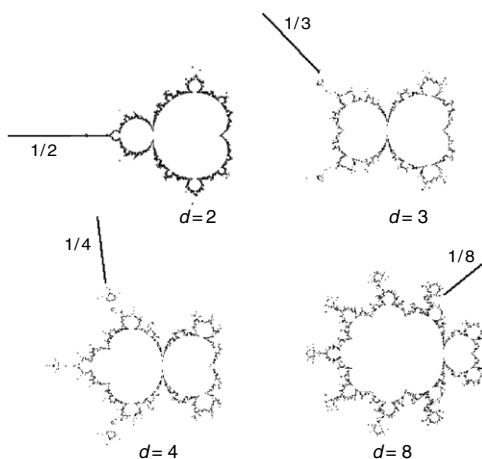


Fig. 56. The external ray corresponding to angle  $1/d$  for  $d = 2, 3, 4, 8$ .

It is an open question as to just how generally this result holds. For example, we conjecture that if  $s = (s_0 s_1 s_2 \dots)$  is a repeating sequence with  $s_0 = 0$ , then the corresponding hair converges to a bifurcation point in the parameter plane with external angle

$$p(d) = (d-1) \sum_{i=0}^{\infty} \frac{s_i}{d^{i+1}}.$$

In particular, the hairs corresponding to itineraries of the form  $(0j0j0j\dots)$  should have external angle  $j/(d+1)$ . See Figure 57.

**PROBLEM.** Determine the rays in  $B_{\infty}$  that land at bifurcation points on the boundaries of the  $C_k$ . How are they related to the external rays in  $B_n$ ?

## 10. Other families of maps

Thus far, except for some brief excursions into the land of polynomials, we have concentrated mostly on the exponential family. Of course, there are many other complex analytic maps whose dynamics are rather intriguing. Among entire functions that have been studied, we single out the trigonometric functions and the complex standard family  $z \mapsto z + \omega + \beta \sin z$  [40] as having received the most attention. More recently, the class of meromorphic functions has received considerable attention. In this section, we describe a few families of such families, paying special attention to how their dynamics differ from that of the exponential. For an excellent survey of these maps, we refer to [10].

One of the principal differences arising in the iteration of meromorphic (non-rational) functions is the fact that, strictly speaking, iteration of these maps does not lead to a

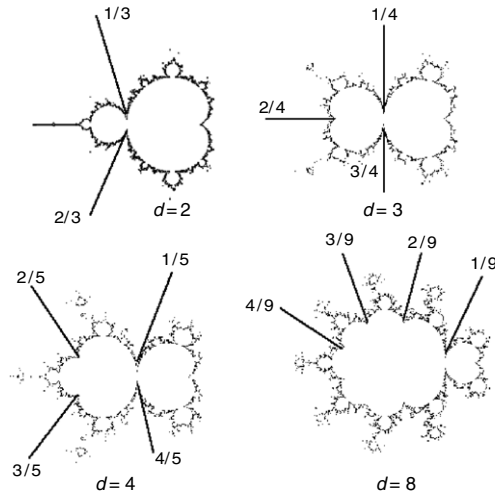


Fig. 57. The external ray corresponding to angle  $1/d$  for  $d = 2, 3, 4, 8$ .

dynamical system. Infinity is an essential singularity for such a map, and so the map cannot be extended continuously to infinity. Hence the forward orbit of any pole terminates, and, moreover, any preimage of a pole also has a finite orbit. All other points have well defined forward orbits.

Despite the fact that certain orbits of a meromorphic map are finite, the iteration of such maps is important. For example, the iterative processes associated with Newton's method applied to entire functions often yields a meromorphic function as the root-finder. See [19,11]. See also [42,41,49].

### 10.1. Maps with polynomial Schwarzian derivative

In this section we will deal exclusively with a very special and interesting class of meromorphic functions, namely, those whose Schwarzian derivative is a polynomial. This class of maps includes a number of dynamically important families of maps, including  $\lambda \tan z$  and  $\lambda \exp z$ .

The Schwarzian derivative has played a role in the analysis of dynamical systems in other settings. For example, if the Schwarzian derivative of two  $C^3$  maps of the interval is negative, the same is true for their composition. Singer [65] has used this to show that the class of functions satisfying these conditions share many of the special properties of complex analytic maps.

The main property of maps with polynomial Schwarzian derivative that makes this class special was noted first by Nevanlinna [62]. These maps are precisely the maps that have only finitely many asymptotic values and no critical values. As we have seen, the fate of the asymptotic values and critical values under iteration plays a crucial role in determining the dynamics. A further important fact about these maps concerns the covering properties

of the map itself. Hille [46] has shown that the plane may be decomposed into exactly  $p$  sectors of equal angle (when the Schwarzian derivative has degree  $p - 2$ ) each of which is associated with one of the asymptotic values. This structure theorem allows us to say much about the Julia set of the map.

DEFINITION 10.1. If  $F(z)$  is a meromorphic function, its Schwarzian derivative is defined by

$$S(F(z)) = \frac{F'''(z)}{F'(z)} - \frac{3}{2} \left( \frac{F''(z)}{F'(z)} \right)^2.$$

Associated to the Schwarzian differential equation

$$S(F(z)) = Q(z) \tag{*}$$

is a linear differential equation obtained by setting

$$g(z) = (F'(z))^{-\frac{1}{2}}.$$

The resulting equation is

$$g'' + \frac{1}{2}Q(z)g = 0. \tag{**}$$

If  $g_1, g_2$  are linearly independent (locally defined) solutions of (\*\*), their Wronskian is a non-zero constant  $k$ . Since

$$\left( \frac{g_1}{g_2} \right)' = \frac{k}{g_2^2},$$

it follows that  $F(z) = g_1(z)/g_2(z)$  is a solution of (\*). Conversely, each solution of (\*) may be written locally as a quotient of independent solutions of (\*\*).

There is a wide class of maps whose Schwarzian derivatives are polynomials. These include such maps as  $\lambda \tan z$  and  $\lambda \exp z$ , for which the Schwarzian derivative is a constant, and  $\int^z \exp(R(u))du$  where  $R$  is a polynomial.

The results of Nevanlinna [62] and Hille [46] allow us to describe the asymptotic properties of the solutions to (\*\*) when  $Q$  is a polynomial of degree  $p - 2$ . There are exactly  $p$  special solutions,  $G_0 \dots G_{p-1}$ , called *truncated* solutions, which have the following property: in any sector of the form

$$\left| \arg z - \frac{2\pi v}{p} \right| < \frac{3\pi}{p} - \epsilon$$

with  $\epsilon > 0$ ,  $G_v(z)$  has the asymptotic development

$$\log G_v(z) \sim (-1)^{v+1} z^{p/2}.$$

Each  $G_\nu$  is an entire function of order  $p/2$ . It follows that each  $G_\nu$  tends to zero as  $z \rightarrow \infty$  along each ray in a sector  $W_\nu$  of the form

$$\left| \arg z - \frac{2\pi\nu}{p} \right| < \frac{\pi}{p}.$$

Moreover,  $G_\nu(z) \rightarrow \infty$  in the adjacent sectors  $W_{\nu+1}$  and  $W_{\nu-1}$ . Note that  $G_\nu$  and  $G_{\nu+1}$  are necessarily linearly independent. However,  $G_\nu$  and  $G_{\nu+k}$  for  $|k| \geq 2$  need not be independent.

Any solution of the associated Schwarzian equation may therefore be written in the appropriate sector in the form

$$\frac{AG_\nu(z) + BG_{\nu+1}(z)}{CG_\nu(z) + DG_{\nu+1}(z)} = F(z) \quad (\dagger)$$

with  $AD - BC \neq 0$ . Note that  $F(z)$  tends to  $A/C$  along any ray in the interior of  $W_{\nu+1}$  and to  $B/D$  in  $W_\nu$ . Recall that an asymptotic (or critical) path for a function  $F(z)$  is a curve  $\alpha : [0, 1) \rightarrow \mathbb{C}$  such that

$$\lim_{t \rightarrow 1} \alpha(t) = \infty$$

and

$$\lim_{t \rightarrow 1} F(\alpha(t)) = \omega.$$

The point  $\omega$  is called an asymptotic value of  $F$ . Thus  $A/C$  and  $B/D$  are asymptotic values.

EXAMPLE. Let  $G_0(z) = e^{z/2}$  and  $G_1(z) = e^{-z/2}$ . Then we may write

$$e^z = \frac{g_0(z) + 0 \cdot G_1(z)}{0 \cdot G_0(z) + G_1(z)}.$$

So  $A/C = \infty$  and  $B/D = 0$ . If  $W_0$  is the left half-plane and  $W_1$  is the right half-plane, then  $e^z$  tends to 0 along rays in  $W_0$  and to  $\infty$  in  $W_1$ . So, as we have seen, both 0 and  $\infty$  are asymptotic values for the exponential.

Asymptotic values can be classified. Nevanlinna's results show that the assumption that  $Q$  is a polynomial implies that  $F$  has only finitely many asymptotic values. They are therefore all isolated. Let  $B$  be a neighbourhood of the asymptotic value  $\omega$  that contains no other asymptotic values. Consider the components of  $F^{-1}(B - \omega)$ . Since the only points at which  $F$  is not a covering of its image are the asymptotic values,  $F$  is a covering map on these components. Hence these components are either disks or punctured disks. If some component is a disk, then the asymptotic value  $\omega$  is called a *logarithmic singularity*.

EXAMPLE. Let  $F(z) = \tan z$ . Then  $i$  and  $-i$  are logarithmic singularities.  $F$  maps the half-plane  $\text{Im } z > y_0 > 0$  onto a punctured neighbourhood of  $i$ . The image of any path  $\alpha(t)$  such that  $\lim_{t \rightarrow 1} \text{Im } \alpha(t) = \infty$  is a path  $\beta(t)$  such that  $\lim_{t \rightarrow 1} \beta(t) = i$ . Similarly the



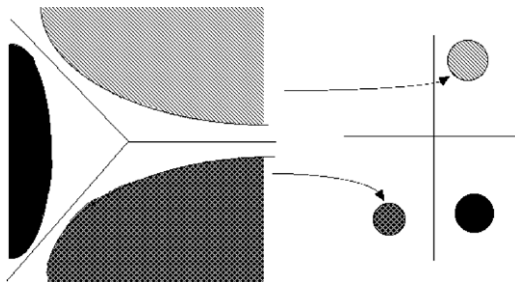


Fig. 58. Exponential tracts and Julia rays.

image of a lower half-plane  $\text{Im } z < v_1 < 0$  is mapped onto a punctured neighbourhood of  $-i$ . The point at  $\infty$  is an accumulation point of the poles; it is not an asymptotic value.

Since a map with a polynomial Schwarzian derivative assumes the special form (†) in each sector  $W_\nu$ , it follows that such a map has exactly  $p$  asymptotic values. Two or more of these values may coincide, but in this case, non-adjacent sectors of asymptotic paths correspond to this value. Also,  $F$  has no critical points since

$$F'(z) = \frac{k}{g_2^2(z)}$$

where  $k$  is a constant and  $g_2$  is entire. We summarize these facts in a theorem originally proved by Nevanlinna [62].

**THEOREM 10.2.** *Functions whose Schwarzian derivatives are degree  $p - 2$  polynomials are precisely the functions that have  $p$  logarithmic singularities,  $a_0, \dots, a_{p-1}$ . The  $a_i$  need not be distinct. There are exactly  $p$  disjoint sectors  $W_0, \dots, W_{p-1}$ , at  $\infty$ , each with angle  $2\pi/p$  in which  $F$  has the following behaviour: there is a collection of disks  $B_i$ , one around each of the  $a_i$ , satisfying  $F^{-1}(B_i - a_i)$  contains a unique unbounded component  $U_i \subset W_i$  and  $F : U_i \rightarrow B_i - a_i$  is a universal covering.*

The  $U_i$  are called *exponential tracts*. Since the truncated solutions tend to 0 or  $\infty$  in adjacent sectors, it follows that the boundary curve of each  $U_i$  has as asymptotic directions the pair of rays which bound the sectors  $W_i$ . A ray  $\beta$  is called a *Julia ray* for  $F$  if, in any angle about  $\beta$ ,  $F$  assumes all (but at most one) values infinitely often. See Figure 1. It is an immediate consequence of (†) that the ray  $\beta_i$  bounding a pair of adjacent sectors  $W_{i-1}$  and  $W_i$  for  $F$  is a Julia ray. See Figure 58.

**EXAMPLE.** The positive and negative imaginary axes are Julia rays for  $e^z$ . Similarly, the rays

$$\arg z = \frac{(2k+1)\pi}{2p}$$

for  $k \in \mathbb{Z}$  are Julia rays for  $\exp(z^p)$ .

When  $F$  is meromorphic, the arguments of the poles of  $F$  accumulate on the argument of the Julia rays. Therefore, except for finitely many poles, it is possible to associate to a given pole  $p$ , that particular Julia ray,  $\beta_{i(p)}$ , such that the pole lies in a small angle about  $\beta_{i(p)}$ . We also associate two asymptotic values to each pole;  $v_1(p)$  is the asymptotic value corresponding to  $U_{i(p)-1}$  and  $v_2(p)$  is the asymptotic value corresponding to  $U_{i(p)}$ .

EXAMPLE. The positive and negative axes are Julia rays for  $\tan z$ , the poles are contained in these rays and  $v_1(p) = i$ ,  $v_2(p) = -i$  for each of the positive poles while  $v_1(p) = -i$ ,  $v_2(p) = i$  for each of the negative poles.

## 10.2. The tangent family

Consider the equation

$$S(F(z)) = k \quad (**)$$

where  $k \in \mathbb{R} - \{0\}$ . The truncated solutions of (\*\*) are given by

$$e^{\pm\sqrt{-k/2}z}$$

and the general solution is

$$\frac{Ae^{\sqrt{-k/2}z} + Be^{-\sqrt{-k/2}z}}{Ce^{\sqrt{-k/2}z} + De^{-\sqrt{-k/2}z}}$$

with  $AD - BC \neq 0$ . Two of these parameters can be fixed by affine conjugation. We will consider one parameter subfamilies of this family in this and the next two sections.

Let

$$T_\lambda(z) = \lambda \tan z = \frac{\lambda e^{iz} - e^{-iz}}{i e^{iz} + e^{-iz}}$$

where  $\lambda > 0$ . We have  $S(T_\lambda(z)) = 2$ . As we have seen,  $T_\lambda$  has asymptotic values at  $\pm\lambda i$ , and  $T_\lambda$  preserves the real axis.

To define the Julia set of this map (and other maps in this class), we adopt the usual definition:  $J(T_\lambda)$  is the set of points at which the family of iterates of the map is not a normal family in the sense of Montel. As in the case of entire functions,  $J(T_\lambda)$  is also the closure of the set of repelling periodic points. But there is a new equivalent formulation of the Julia set:  $J(T_\lambda)$  is also the closure of the set which consists of the union of all of the preimages of the poles of  $T_\lambda$ .

EXERCISE. Prove that all the poles and their preimages are dense in the Julia set.

Usually, when the Julia set is not the entire plane, this set is a ‘fractal’. Some exceptions are the quadratic maps  $z \mapsto z^2$  whose Julia set is the unit circle, and  $z \mapsto z^2 - 2$ ,

whose Julia set is the interval  $[-2, 2]$ . For all other values of  $c$ , the Julia set of  $z^2 + c$  is a fractal. The tangent family provides another example of a map whose Julia set is a smooth submanifold of  $\mathbb{C}$ .

**PROPOSITION 10.3.** *If  $\lambda \in \mathbb{R}$ ,  $\lambda > 1$ , then  $J(T_\lambda)$  is the real line and all other points tend asymptotically to one of two fixed sinks located on the imaginary axis.*

**PROOF.** Write  $T_\lambda(z) = L_\lambda \circ E(z)$  where

$$E(z) = \exp(2iz)$$

$$L_\lambda(z) = -\lambda i \left( \frac{z-1}{z+1} \right).$$

$E$  maps the upper half-plane onto the unit disk minus 0 and  $L_\lambda$  maps the disk back to the upper half-plane. Both  $E$  and  $L_\lambda$  preserve boundaries, so  $T_\lambda$  maps the interior of the upper half-plane into itself. Now  $T_\lambda$  also preserves the imaginary axis and we have

$$T_\lambda(iy) = i\lambda \tanh(y).$$

The graph of  $\lambda \tanh y$  shows that  $T_\lambda$  has a pair of attracting fixed points located symmetrically about 0 if  $\lambda > 1$ . By the Schwarz lemma, all points in the upper (respectively, lower) half-plane tend, under iteration, to one of these points.

Hence neither the upper nor the lower half-plane is in  $J(T_\lambda)$ . The real line is in  $J(T_\lambda)$ . This follows from the facts that the real line satisfies  $T_\lambda^{-1}(\mathbb{R}) \subset \mathbb{R}$  and  $T_\lambda(\mathbb{R}) = \mathbb{R} \cup \infty$ , and that  $T'_\lambda(x) > 1$  for all  $x \in \mathbb{R}$  if  $\lambda > 1$  ( $T'_\lambda(x) \geq 1$  if  $\lambda = 1$ ). Each interval of the form

$$\left( \frac{2k-1}{2}\pi, \frac{2k+1}{2}\pi \right)$$

is expanded over all of  $\mathbb{R}$ . If  $U$  is any open interval in  $\mathbb{R}$ , then there is an integer  $k$  such that  $T_\lambda^k(U)$  covers one of these intervals of length  $\pi$ . Hence  $T_\lambda^{k+1}(U)$  covers  $U$ . It follows that there exist repelling fixed points and poles of  $T_\lambda^{k+1}$  in  $U$ .  $\square$

**REMARKS.** 1. If  $\lambda = 1$ , then  $J(T_\lambda) = \mathbb{R}$ , and all points with non-zero imaginary parts tend asymptotically to the neutral fixed point at 0.

2. When  $\lambda < -1$ , the dynamics of  $T_\lambda$  are similar to those for  $\lambda > 1$ , except that  $T_\lambda$  has an attracting periodic cycle of period two. Points in the upper and lower half-planes hop back and forth as they are attracted to the cycle. Since  $|T'_\lambda(x)| > 1$  for  $x \in \mathbb{R}$ , it follows, as above, that  $J(T_\lambda) = \mathbb{R}$  for  $\lambda < -1$ .

For  $0 < |\lambda| < 1$ , 0 is an attracting fixed point for  $T_\lambda$ . In this case, the Julia set of  $T_\lambda$  breaks up into a Cantor set, as we show below. We will as usual employ symbolic dynamics to describe the Julia set in this case. Let  $\Gamma$  denote the set of one-sided sequences whose entries are either integers or the symbol  $\infty$ . If  $\infty$  is an entry in a sequence, then we terminate the sequence at this entry, i.e.  $\Gamma$  consists of all infinite sequences  $(s_0, s_1, s_2, \dots)$  where  $s_j \in \mathbb{Z}$  and all finite sequences of the form  $(s_0, s_1, \dots, s_j, \infty)$  where  $s_i \in \mathbb{Z}$ .

The topology on  $\Gamma$  was described in [61]. For completeness, we will recall this topology here. If  $(s_0, s_1, s_2, \dots)$  is an infinite sequence, we choose as a neighbourhood basis of this

sequence the sets

$$U_k = \{(t_0, t_1, \dots) \mid t_i = s_i \text{ for } i \leq k\}.$$

If, on the other hand, the sequence is finite  $(s_0, \dots, s_j, \infty)$ , then we choose the  $U_k$  as above for  $k \leq j$  as well as sets of the form

$$V_\ell = \{(t_0, t_1, \dots) \mid t_i = s_i \text{ for } i \leq j \text{ and } |t_{j+1}| \geq \ell\}$$

for a neighbourhood basis.

There is a natural shift map  $\sigma : \Gamma \rightarrow \Gamma$  which is defined as usual by  $\sigma(s_0 s_1 s_2 \dots) = (s_1 s_2 \dots)$ . Note that  $\sigma(\infty)$  is not defined. In Moser's topology,  $\sigma$  is continuous and  $\Gamma$  is a Cantor set.  $\Gamma$  provides a model for many of the Julia sets of maps in our class, and  $\sigma|_\Gamma$  is conjugate to the action of  $F$  on  $J(F)$ . One such instance of this is shown in the following proposition.

**PROPOSITION 10.4.** *Suppose  $\lambda \in \mathbb{R}$  and  $0 < |\lambda| < 1$ . Then  $J(T_\lambda)$  is a Cantor set in  $\hat{\mathbb{C}}$  and  $T_\lambda|_{J(T_\lambda)}$  is topologically conjugate to  $\sigma|_\Gamma$ .*

**PROOF.** Since  $0 < |\lambda| < 1$ , 0 is an attracting fixed point for  $T_\lambda$ . Let  $B$  denote the immediate basin of attraction of 0 in  $\mathbb{R}$ .  $B$  is an open interval of the form  $(-p, p)$  where  $T_\lambda(\pm p) = \pm p$ . (The points  $\pm p$  lie on a periodic orbit of period two if  $-1 < \lambda < 0$ .) The preimages  $T_\lambda^{-1}(B)$  consist of infinitely many disjoint open intervals. Let  $I_j$ ,  $j \in \mathbb{Z}$ , denote the complementary intervals, enumerated left to right so that  $I_0$  abuts  $p$ . Then  $T_\lambda : I_j \rightarrow (\mathbb{R} \cup \infty) - B$  for each  $j$ , and  $|T'_\lambda(x)| > 1$  for each  $x \in I_j$ . Standard arguments [61] then show that

$$\Lambda = \{x \in \mathbb{R} \cup \{\infty\} \mid T_\lambda^j(x) \in \cup I_j \text{ for all } j\}$$

is a Cantor set and  $T_\lambda|_\Lambda$  is conjugate to  $\sigma|_\Gamma$ .

Now  $\Lambda$  is invariant under all branches of the inverse of  $T_\lambda$ . It therefore contains preimages of poles of all orders and is closed. Hence  $\Lambda$  is the Julia set of  $T_\lambda$ . The classification of stable regions tells us that all other points lie in the basin of 0.  $\square$

**REMARKS.** 1. The basin of 0 is therefore infinitely connected. This contrasts with the situation for polynomial or entire maps in which *finite* attracting fixed points always have a simply connected immediate basin of attraction.

2. In fact, the Julia set of  $T_\lambda$  is a similar Cantor set for all  $\lambda$  with  $|\lambda| < 1$ . See [28].

3. A full picture of the parameter plane for the tangent family may be found in [48].

### 10.3. Asymptotic values that are poles

As we have seen, entire transcendental functions of finite type often have Julia sets which contain analytic curves. Indeed, for a wide class of these maps (see [31]), all repelling periodic orbits lie at the endpoints of invariant curves which connect the orbit to the essential singularity at  $\infty$ .

In this section we give an example of a family of maps with constant Schwarzian derivatives for which certain of the repelling fixed points lie on analytic curves in the Julia set, but for which many of the other periodic points do not. This lack of homogeneity in the Julia set is caused by the fact that one of the asymptotic values is a pole.

Consider the family of maps

$$F_\lambda(z) = \frac{\lambda e^z}{e^z - e^{-z}}$$

with  $\lambda > 0$ . We have  $S(F_\lambda(z)) = -2$  and  $F_\lambda$  is periodic with period  $\pi i$ . These maps have asymptotic values at 0 and  $\lambda$ , and 0 is also a pole.

The graph of  $F_\lambda$  restricted to  $\mathbb{R}$  shows that  $F_\lambda$  has two fixed points in  $\mathbb{R}$  at  $p$  and  $q$  with  $p < 0 < q$ . We note that  $F_\lambda(z) = L_\lambda \circ E(z)$  where  $E(z) = \exp(-2z)$  and  $L_\lambda$  is the linear fractional transformation

$$L_\lambda(z) = \frac{\lambda}{1 - z}.$$

$F_\lambda$  has poles at  $k\pi i$  where  $k \in \mathbb{Z}$  as well as the following mapping properties:

1.  $F_\lambda$  preserves  $\mathbb{R}^+$  and  $\mathbb{R}^-$ .
2.  $F_\lambda$  maps the horizontal lines  $\text{Im } z = \frac{1}{2}(2k+1)\pi$  onto the interval  $(0, \lambda)$  in  $\mathbb{R}$ .
3.  $F_\lambda$  maps the imaginary axis onto the line  $\text{Re } z = \lambda/2$ , with the points  $k\pi i$  mapped to  $\infty$ .
4.  $F_\lambda$  maps horizontal lines onto circular arcs passing through both 0 and  $\lambda$ .
5.  $F_\lambda$  maps vertical lines with  $\text{Re } z > 0$  to a family of circles orthogonal to those in 4 which are contained in the plane  $\text{Re } z > \lambda/2$ .
6.  $F_\lambda$  maps vertical lines with  $\text{Re } z < 0$  to a family of circles orthogonal to those in 4 which are contained in the plane  $\text{Re } z < \lambda/2$ .

As a consequence of these properties, we have

**PROPOSITION 10.5.** *If  $\lambda > 0$ , then the fixed point  $q$  is attracting. Moreover, if  $\text{Re } z > 0$ , then  $F_\lambda^n(z) \rightarrow q$  as  $n \rightarrow \infty$ . Hence  $J(F_\lambda)$  is contained in the half-plane  $\text{Re } z \leq 0$ .*

To see this, just compute  $|F'_\lambda(q)| < 1$ . Then use property 5 above and the Schwarz lemma.

**PROPOSITION 10.6.**  *$J(F_\lambda)$  contains  $\mathbb{R}^- \cup \{0\}$ .*

**PROOF.** The fixed point  $p$  is repelling. This follows from the fact that  $F_\lambda$  has negative Schwarzian derivative: if  $|(F_\lambda)'(p)| \leq 1$ , then it follows that  $p$  would have to attract a critical point or asymptotic value of  $F_\lambda$  on  $\mathbb{R}$ . This does not occur since  $q$  attracts  $\lambda$  and 0 is a pole.

Let  $x \in (-\infty, p)$ . One may check easily that

$$|(F_\lambda^2)'(x)| > 1.$$

Moreover,

$$|(F_\lambda^{2n})'(x)| \rightarrow \infty$$

as  $n \rightarrow \infty$ . This again follows from the fact that  $F_\lambda^2$  has negative Schwarzian derivative on  $\mathbb{R}^-$ . Let  $U$  be a neighbourhood of  $x$  in  $\mathbb{C}$ . Note that  $F_\lambda^{2n}$  expands  $U$  until some image overlaps the horizontal lines  $y = \pm\pi/2$ . By the above properties, these points are in the basin of  $q$ . Hence the family  $\{F_\lambda^{2n}\}$  is not normal at  $x$ , and so  $(-\infty, p) \subset J(F_\lambda)$ . The image of this interval under  $F_\lambda$  is  $(p, 0)$ , so  $\mathbb{R}^- \subset J(F_\lambda)$ .  $\square$

Thus some points in the Julia set lie on analytic curves; for example,  $\mathbb{R}^-$  and all of its preimages. But not all points in the Julia set lie on smooth invariant curves:

**PROPOSITION 10.7.** *There is a unique repelling fixed point  $p_1$  in the half strip*

$$\pi/2 < \operatorname{Im} z < 3\pi/2$$

*and this point does not lie on any smooth invariant curve in  $J(F_\lambda)$ .*

**PROOF.** Let  $R$  be the rectangle  $\pi/2 < \operatorname{Im} z < 3\pi/2$ ,  $\nu < \operatorname{Re} z < 0$  where  $\nu$  is chosen far enough to the left in  $\mathbb{R}$  that

$$|F_\lambda(\nu + iy)| < \pi/4.$$

Then  $F_\lambda(R)$  is a ‘disk’ which covers  $R$  and  $F_\lambda|_R$  is 1-1. So,  $F_\lambda^{-1}$  has a unique attracting fixed point  $p_1$  in  $R$ . Since this argument is independent of  $\nu$  for  $\nu$  large enough negative, the first part of the proposition follows.

Now suppose that  $p_1$  lies on a smooth invariant curve  $\gamma$  in  $J(F_\lambda)$ . Since  $J(F_\lambda)$  is invariant under  $F_\lambda^{-1}$ , we may assume that  $\gamma$  accumulates on the boundary of the strip  $\pi/2 < \operatorname{Im} z < 3\pi/2$ ,  $\operatorname{Re} z < 0$  by taking iterates of  $F_\lambda^{-1}$ , as above. The upper and lower boundaries of the strip are stable by property 2; hence  $\gamma$  cannot meet  $y = \pi/2$  or  $y = 3\pi/2$ . Similarly,  $\gamma$  cannot meet the line  $x = 0$  (except possibly at  $i\pi$ ). So  $\gamma$  can only accumulate at  $\infty$  or  $i\pi$ . If  $\gamma$  accumulates at  $\infty$ , then  $\gamma$  must also accumulate at  $i\pi$ , since  $F_\lambda(i\pi) = \infty$ . Since all points on  $\gamma$  leave the strip under iteration, it follows that  $\gamma$  must contain  $i\pi$ . Now  $\gamma$  cannot have a tangent vector at  $i\pi$ , for if so,  $\gamma$  would enter the region  $\operatorname{Re} z \geq 0$ ,  $\operatorname{Im} z \neq i\pi$ , which lies in the Fatou set.  $\square$

**REMARK.** There is a continuous invariant curve which lies in the Julia set and accumulates on  $p_1$ . Indeed, the horizontal line  $\ell_0$  given by  $y = \pi$ ,  $x \leq 0$  lies in  $J(F_\lambda)$  since it is mapped onto  $\mathbb{R}^-$  by  $F_\lambda$ . Consider the successive preimages  $\ell_n = F_\lambda^{-n}(\ell_0)$ , where  $F_\lambda^{-1}$  is the branch of the inverse of  $F_\lambda$  whose image is  $\pi/2 < \operatorname{Im} z < 3\pi/2$ . Then  $\ell_1$  meets  $\ell_0$  at  $i\pi$ ,  $\ell_2$  meets  $\ell_1$  at  $F_\lambda^{-1}(i\pi)$ , and so forth. Since  $p_1$  is an attracting fixed point for  $F_\lambda^{-1}$ , the curve  $\ell$  formed by concatenating the  $\ell_i$  is invariant and accumulates on  $p_i$  as  $i \rightarrow \infty$ . Note that this curve is considerably different, from a dynamical point of view, from the invariant curve  $\mathbb{R}^-$  through  $p$ .

#### 10.4. Bifurcation to an entire function

Most maps with polynomial Schwarzian derivatives are bona fide meromorphic functions, but occasionally they are entire functions. In this section we give an example of an ‘explosion’ in the Julia set which occurs when a member of a meromorphic family suddenly becomes an entire function. An *explosion* occurs at a parameter value for a family of functions whenever the Julia sets of the functions in the family change suddenly, when the parameter is reached, from a nowhere dense subset of  $\mathbb{C}$  to all of  $\mathbb{C}$ .

Consider the family

$$F_\lambda(z) = \frac{e^z}{\lambda e^z + e^{-z}} = \frac{1}{\lambda + e^{-2z}}.$$

When  $\lambda = 0$ , the corresponding element of this family is the entire function  $F_0(z) = \exp(2z)$  whose dynamics are well understood. Indeed,  $e^{2z}$  is linearly conjugate to  $2e^z$ , and so  $J(F_0) = \mathbb{C}$ , since the orbit of the asymptotic value 0 tends to  $\infty$ .

When  $\lambda > 0$ ,  $J(F_\lambda) \neq \mathbb{C}$ . This follows since  $F_\lambda$  has a unique attracting fixed point  $p_\lambda$  on the real line. In fact, we can say much more about  $J(F_\lambda)$ .

**PROPOSITION 10.8.** *For all  $\lambda > 0$ , the Julia set of  $F_\lambda$  is a Cantor set in  $\mathbb{C}$  and  $F_\lambda \mid J(F_\lambda)$  is the shift map on infinitely many symbols.*

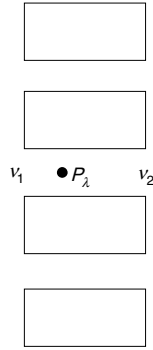
**PROOF.** First note that the entire real axis lies in the basin of attraction of  $p_\lambda$ . This follows since  $F_\lambda$  has negative Schwarzian derivative and maps  $\mathbb{R}$  diffeomorphically onto the interval bounded by the asymptotic values,  $(0, 1/\lambda)$ . In particular, both asymptotic values lie in the immediate basin of  $p_\lambda$  and so there are disks about these points which lie in the basin. Taking preimages of these disks, it follows that there are half-planes of the form  $\operatorname{Re} z < v_1$  and  $\operatorname{Re} z > v_2$  with  $v_1 < p_\lambda < v_2$  which lie in the immediate basin of  $p_\lambda$ .

We may find a strip  $S_\mu$  surrounding the interval  $[v_1, v_2]$  of the form

$$\{z \mid |\operatorname{Im} z| < \mu, v_1 \leq \operatorname{Re} z \leq v_2\}$$

which is mapped inside itself. Now let  $B$  denote the ‘ladder-shaped’ region consisting of the two half-planes together with  $S_\mu$  and all of its  $\pi i$  translates (see Figure 59). Clearly,  $F_\lambda$  maps  $B$  inside itself as long as the  $v_i$  are chosen large enough.

The complement of  $B$  consists of infinitely many congruent rectangles  $R_j$  where  $j \in \mathbb{Z}$  and the  $R_j$  are indexed according to the increasing imaginary part.  $F_\lambda$  maps each  $R_j$  diffeomorphically onto  $\mathbb{C} - F_\lambda(B)$ . In particular,  $F_\lambda(R_j)$  covers each  $R_k$  and  $\infty$ . It follows that there exists at least one point  $z$  corresponding to any sequence  $(s_0, s_1, \dots)$  in the sequence space which has the property that  $F_\lambda^n(z) \in R_{s_n}$  for each  $n$ . Then our usual arguments show that this point is unique, lies in the Julia set, and  $J(F_\lambda)$  therefore is a Cantor set modelled on the sequence space with infinitely many symbols, as described in Section 2.  $\square$

Fig. 59. The region  $B$ .

### 10.5. Cantor bouquets and cantor sets

Our goal in this section is to describe the topology of the Julia set for a general meromorphic map  $F$  with Schwarzian derivative that is a polynomial  $P(z)$ . By Nevanlinna's Theorem, each such map  $F$  has  $p$  asymptotic values  $a_0, \dots, a_{p-1}$ . To each  $a_i$  there corresponds a sector  $W_i$  with angle  $2\pi/p$  in which  $F$  has the following behaviour: We may choose a small disk  $B_i$  about  $a_i$  such that, if  $U_i$  is the component of  $F^{-1}(B_i)$  meeting  $W_i$ , then  $F : U_i \rightarrow B_i - a_i$  is a universal covering map. The sectors are separated by the Julia rays  $\beta_i$ . Almost all poles  $p$  have associated Julia rays,  $\beta_{i(p)}$ , and asymptotic values  $v_1(p)$  and  $v_2(p)$ .

Note that all of the  $a_i$  need not be distinct, but it follows from the linear-independence of the truncated solutions  $G_v$  and  $G_{v+1}$  that  $a_i$  corresponding to adjacent  $W_i$  are distinct. Thus the basic mapping properties of  $F$  are as depicted in Figure 60.

Let us assume that all of the  $a_i$  are finite. We may choose  $R$  sufficiently large so that  $D_R = \{z \mid |z| \leq R\}$  contains all of the  $B_i$  in its interior. Let  $\Gamma_R$  denote the disk in  $\hat{\mathbb{C}}$  which is the complement of  $D_R$ . Let

$$\tilde{\Gamma}_R = \Gamma_R - \bigcup_{i=0}^{p-1} U_i.$$

If  $R$  is large enough,  $\tilde{\Gamma}_R$  consists of exactly  $p$  'arms' which extend to  $\infty$  in  $\Gamma_R$  and which separate the  $U_i$ . Let  $A_i$  denote the arm between  $U_i$  and  $U_{i+1}$ .  $A_i$  contains the Julia ray  $\beta_i$  (see Figure 60).

Since  $F \mid A_i$  is a covering map that covers  $\hat{\mathbb{C}} - (B_i \cup B_{i+1})$  infinitely often, it follows that  $F^{-1}(\Gamma_R) \cap A_i$  consists of infinitely many disks, each of which is mapped by  $F$  in a one-to-one fashion over  $\Gamma_R$ . These disks accumulate only at  $\infty$ . This is similar to the situation in the previous section.

**PROPOSITION 10.9.** *Suppose all of the asymptotic values of  $F$  are finite. Let  $\Lambda_R = \{z \mid F^j(z) \in \Gamma_R \text{ for all } j\}$ . If  $R$  is chosen large enough, then  $\Lambda_R$  is a closed, forward invariant*



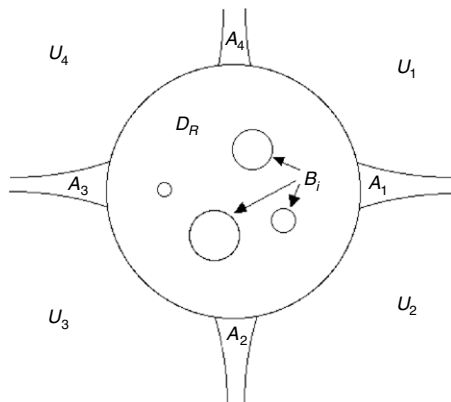


Fig. 60. Exponential tracts and Julia rays.

subset of  $J(F)$ . Moreover,  $\Lambda_R$  is homeomorphic to a Cantor set that is modelled on the shift space with infinitely many symbols.

PROOF. As there are only countably many disks in  $F^{-1}(\Gamma_R) \cap A_i$  for each arm  $A_i$ , we may choose an indexing of these disks by the natural numbers. Say  $F^{-1}(\Gamma_R) \cap (\cup A_i) = \bigcup_{j=0}^{\infty} D_j$ . Thus each  $D_j \subset \Gamma_R$  and  $F$  maps each  $D_j$  onto  $\Gamma_R$ . In particular,  $F|D_j$  covers each other  $D_k$  and  $\infty$ . Standard arguments as described in Section 2 then yield the result.  $\square$

Thus the set of points whose orbits remain in a neighbourhood of  $\infty$  form a closed forward invariant subset of the Julia set which is homeomorphic to a Cantor set. We may apply these ideas on a global level if we can guarantee that all of the asymptotic values lie in a single immediate attracting basin of a fixed point.

**COROLLARY 10.10.** *Suppose each of the  $a_i$  lie in the immediate attracting basin of an attracting fixed point. Then  $J(F)$  is a Cantor set and  $F|J(F)$  is conjugate to the shift map on infinitely many symbols.*

PROOF. Our assumption allows us to choose a simple closed curve in  $\mathbb{C}$  which bounds an open set in the immediate attractive basin, and which contains all of the  $B_i$ . The Julia set is contained in the complement of this set. Applying the above argument to this curve instead of  $\Gamma_R$  yields the result.  $\square$

Recall that the Julia set of a rational map is also a Cantor set under this hypothesis so that these meromorphic maps are dynamically similar to rational maps. By contrast, there are no entire transcendental functions whose Julia sets are Cantor sets [5]. On the other hand, if one or more of the asymptotic values is the point at  $\infty$ , then the Julia set contains Cantor bouquets.

**THEOREM 10.11.** *Suppose  $F(z)$  has polynomial Schwarzian derivative with degree  $p-2$ . Suppose that  $F$  has an asymptotic value  $a_i$  which is also a pole. Let  $W_i$  be the sector*

containing the exponential tract corresponding to  $a_i$ . Then for each  $N > 0$ ,  $J(F)$  contains a Cantor  $N$ -bouquet in  $W_i$  which is invariant under  $F^2$ .

The proof may be found in [28].

## References

- [1] J. Aarts and L. Oversteegen, *The geometry of Julia sets*, Trans. Amer. Math. Soc. **338** (1993), 897–918.
- [2] L. Ahlfors, *Lectures on Quasiconformal Mappings*, American Mathematical Society (2006).
- [3] L. Ahlfors and L. Bers, *The Riemann mappings theorem for variable metrics*, Ann. Math. **72** (1960), 385–404.
- [4] P. Atela, *Bifurcations of dynamic rays in complex polynomials of degree two*, Ergodic Theory Dynam. Systems **12** (1991), 401–423.
- [5] I.N. Baker, *Wandering domains in the iteration of entire functions*, Proc. London Math. Soc. **49** (1984), 563–576.
- [6] I.N. Baker, J. Kotus and Y. Lü, *Iteration of exponential functions*, Ann. Acad. Sci. Fenn. **9** (1984), 49–57.
- [7] I.N. Baker and P. Rippon, *Iteration of exponential functions*, Ann. Acad. Sci. Fenn., Series 1A Math. **9** (1984), 49–77.
- [8] M. Barge, *Horseshoe maps and inverse limits*, Pacific J. Math. **121** (1986), 29–39.
- [9] A. Beardon, *Iteration of Rational Functions*, Cambridge University Press (1994).
- [10] W. Berweiler, *Iteration of meromorphic functions*, Bull. Amer. Math. Soc. **29** (1993), 151–188.
- [11] W. Bergweiler, et al., *Newton's method for meromorphic functions*, Complex Analysis and its Applications, Pitman Res. Notes Math., Vol. 305 (1994), 147–158.
- [12] R. Bhattacharjee and R.L. Devaney, *Tying hairs for structurally stable exponentials*, preprint.
- [13] P. Blanchard, *Complex analytic dynamics on the Riemann sphere*, B.A.M.S. **2** (1984), 85–141.
- [14] C. Bodelón, R.L. Devaney, M. Hayes, L. Goldberg, J. Hubbard and G. Roberts, *Hairs for the complex exponential family*, Internat. J. Bifur. Chaos **9** (1999), 1517–1534.
- [15] C. Bodelón, R.L. Devaney, M. Hayes, L. Goldberg, J. Hubbard and G. Roberts, *Dynamical convergence of polynomials to the exponential*, J. Difference Equ. Appl. **6** (2000), 275–307.
- [16] B. Branner, *The Mandelbrot set*, Chaos and Fractals: The Mathematics Behind the Computer Graphics, American Mathematical Society (1989), 75–106.
- [17] L. Carleson and T. Gamelin, *Complex Dynamics*, Springer-Verlag (1994).
- [18] S. Curry, *One-dimensional nonseparating plane continua with disjoint  $\epsilon$ -dense subcontinua*, Topology Appl. **39** (1991), 145–151.
- [19] J. Curry, L. Garnett and D. Sullivan, *On the iteration of a rational function*, Comm. Math. Phys. **91** (1983), 267–277.
- [20] R.L. Devaney, *Julia sets and the bifurcation diagrams for exponential maps*, Bull. Amer. Math. Soc. **11** (1984), 167–171.
- [21] R.L. Devaney, *Structural instability of  $\text{Exp}(z)$* , Proc. A.M.S. **94** (1985), 545–548.
- [22] R.L. Devaney,  *$e^z$ : dynamics and bifurcations*, Internat. J. Bifur. Chaos **1** (1991), 287–308.
- [23] R.L. Devaney, N. Fagella and X. Jarque, *Hyperbolic components the complex exponential family*, Fund. Math. **174** (2002), 193–215.
- [24] R.L. Devaney and L. Goldberg, *Uniformization of attracting basins for exponential maps*, Duke Math. J. **55** (1987), 253–266.
- [25] R.L. Devaney and X. Jarque, *Misiurewicz points for complex exponentials*, Internat. J. Bifur. Chaos **7** (1997), 1599–1616.
- [26] R.L. Devaney and X. Jarque, *Indecomposable continua in exponential dynamics*, Conformal Geom. Dyn. **6** (2002), 1–12.
- [27] R.L. Devaney, X. Jarque and M. Moreno Rocha, *Indecomposable continua and misiurewicz points in exponential dynamics*, Internat. J. Bifur. Chaos **10** (2005), 3281–3293.
- [28] R.L. Devaney and L. Keen, *Dynamics of meromorphic maps: Maps with polynomial Schwarzian derivative (with Linda Keen)*, Ann. Sci. Éc. Norm. Sup. **22** (1989), 55–79.
- [29] R.L. Devaney and M. Krych, *Dynamics of  $\text{Exp}(z)$* , Ergodic Theory Dynam. Systems **4** (1984), 35–52.

- [30] R.L. Devaney and M. Moreno Rocha, *Geometry of the antennas in the Mandelbrot set*, Fractals **10** (2002), 39–46.
- [31] R.L. Devaney and F. Tangerman, *Dynamics of entire functions near the essential singularity*, Ergodic Theory Dynam. Systems **6** (1986), 489–503.
- [32] R.E.L. Deville, *Brjuno numbers and the symbolic dynamics of the complex exponential*, Qual. Theory Dyn. Sys. **1** (1999), 71–82.
- [33] R.E.L. Deville, *Itineraries of entire functions*, J. Difference Equ. Appl. **7** (2001), 193–214.
- [34] A. Douady and L. Goldberg, *The nonconjugacy of certain exponential functions*, Holomorphic Functions and Moduli, MSRI Publ., Springer Verlag (1988), 1–8.
- [35] A. Douady and J. Hubbard, *Itération des Polynômes quadratiques complexes*, C.R. Acad. Sci. Paris **t.29** (Serie I) (1982), 123–126.
- [36] A. Douady and J. Hubbard, *On the dynamics of polynomial-like mappings*, Ann. Sci. Éc. Norm. Sup. 4<sup>e</sup> Séries **t.18** (1985), 287.
- [37] A. Douady and J. Hubbard, *Étude Dynamique des Polynôme Complexes*, Publications Mathématiques d'Orsay, preprint.
- [38] M.B. Durkin, *The accuracy of computer algorithms in dynamical systems*, Internat. J. Bifur. Chaos **1** (1991), 625–639.
- [39] A. Eremenko and M.Yu. Lyubich, *Iterates of entire functions*, Dokl. Akad. Nauk SSSR **279** (1984), 25–27; English translation in Soviet Math. Dokl. **30** (1984), 592–594.
- [40] N. Fagella, *Limiting dynamics for the complex standard family*, Internat. J. Bifur. chaos Appl. Sci. Engrg. **5** (1995), 673–679.
- [41] N. Fagella and A. Garijo, *Capture zones of the family of functions  $\lambda z^m e^z$* , Internat. J. Bifur. Chaos **13** (2003), 2623–2640.
- [42] N. Fagella and C. Henriksen, *Deformation of entire functions with Baker domains*, Discrete Contin. Dyn. Syst. **15** (2006), 379–394.
- [43] J. Farey, *On a curious property of vulgar fractions*, Phil. Mag. J. London **47** (1816), 385–386.
- [44] L.R. Goldberg and L. Keen, *A finiteness theorem for a dynamical class of entire functions*, Ergodic Theory Dynam. Systems **6** (1986), 183–192.
- [45] L. Goldberg and C. Tresser, *Rotation orbits and the Farey tree*, Ergodic Theory Dynam. Systems **16** (1996), 1011–1029.
- [46] E. Hille, *On the zeroes of the functions of the parabolic cylinder*, Ark. Mat. Astron. Fys. **18** (26) (1924).
- [47] B. Karpinska, *On the accessible points in the Julia sets of some entire functions*, Fund. Math. **180** (2003), 89–98.
- [48] L. Keen and J. Kotus, *Dynamics of the family  $\lambda \tan z$* , Conform. Geom. Dyn. **1** (1997), 28–57.
- [49] J. Kotus and M. Urbanski, *Conformal, geometric, and expanding measures for transcendental expanding functions*, Math. Ann. **324** (2002), 619–656.
- [50] B. Krauskopf and H. Kriete, *Kernel convergence of hyperbolic components*, Ergodic Theory Dynam. Systems **17** (1997), 1137–1146.
- [51] K. Kuratowski, *Topology*, Vol. 2, Academic Press, New York (1968).
- [52] E. Lau and D. Schleicher, *Internal addresses in the Mandelbrot set and irreducibility of polynomials*, SUNY Stony Brook Institute for Mathematical Sciences, preprint #1994–19.
- [53] P. LaVauers, *Une Description Combinatoire de l'involution définie par  $M$  sur les rationnelles a denuminateur impair*, C.R. Acad. Sci. Paris Sér. I Math. **303** (1986), 143–146.
- [54] E. Lee, *The structure and topology of the Brjuno numbers*, Proc. 1999 Topology and Dynamics Conference **24** (1999), 189–201.
- [55] M. Lyubich, *Measurable dynamics of the exponential*, Soviet Math. Dokl. **35** (1987), 223–226.
- [56] J. Mayer, *An explosion point for the set of endpoints of the Julia set of  $\lambda \exp(z)$* , Ergodic Theory Dynam. Systems **10** (1990), 177–184.
- [57] J. Mayer and J.T. Rogers, *Indecomposable continua and the Julia sets of polynomials*, preprint.
- [58] C. McMullen, *Area and Hausdorff dimension of Julia sets of entire functions*, Trans. A.M.S. **300** (1987), 329–342.
- [59] C. McMullen, *Complex Dynamics and Renormalization*, Princeton University Press (1994).
- [60] J. Milnor, *Dynamics in One Complex Variable*, Princeton University Press (2006).
- [61] J. Moser, *Stable and Random Motions in Dynamical Systems*, Princeton University Press, Princeton (1973).
- [62] R. Nevanlinna, *Über Riemannsche Flächen mit endlich vielen Windungspunkten*, Acta Math. **58** (1932).

- [63] C. Petersen and G. Ryd, *Convergence of rational rays in parameter spaces*, The Mandelbrot Set: Theme and Variations, London Mathematical Society, Lecture Note Series, Vol. 274, Cambridge University Press (2000), 161–172.
- [64] G. Piranian, *The boundary of a simply connected domain*, Bull. Amer. Math. Soc. **64** (1958), 45–55.
- [65] D. Singer, *Stable orbits and bifurcations of maps of the interval*, SIAM J. Appl. Math. **35** (1978), 260–267.
- [66] S. Smale, *Diffeomorphisms with many periodic points*, Differential and Combinatorial Topology, Princeton University Press (1964), 63–80.
- [67] N. Steinmetz, *Rational Iteration: Complex Analytic Dynamical Systems*, De Gruyter, Berlin (1993).
- [68] D. Sullivan, *Quasiconformal maps and dynamical systems I, Solutions of the Fatou–Julia problem on wandering domains*, Ann. Math. **122** (1985), 401–418.
- [69] M. Viana, *The differentiability of the hairs of  $\exp(z)$* , Proc. Amer. Math. Soc. **103** (1988), 1179–1184.



## CHAPTER 5

# Some Applications of Moser's Twist Theorem

Mark Levi<sup>1</sup>

*Department of Mathematics, Pennsylvania State University, United States*

*E-mail: [levi@math.psu.edu](mailto:levi@math.psu.edu)*

### *Contents*

|                                                                                 |     |
|---------------------------------------------------------------------------------|-----|
| 1. Background: the action-angle variables, the generating functions . . . . .   | 227 |
| 1.1. Examples of action-angle variables . . . . .                               | 227 |
| 1.2. Action-angle variables for particles in $\mathbb{R}^n$ . . . . .           | 229 |
| 1.3. A generating function construction of the action-angle variables . . . . . | 230 |
| 1.4. A mechanical interpretation of generating functions . . . . .              | 231 |
| 1.5. The twist condition . . . . .                                              | 233 |
| 2. Basic statements of KAM theory . . . . .                                     | 235 |
| 2.1. Moser's twist theorem . . . . .                                            | 236 |
| 2.2. The Kolmogorov–Arnold theorem . . . . .                                    | 237 |
| 3. A variational approach to Moser's twist theorem . . . . .                    | 237 |
| 3.1. The generating function of an area-preserving map . . . . .                | 238 |
| 3.2. Reduction to a difference equation . . . . .                               | 238 |
| 4. Applications . . . . .                                                       | 240 |
| 5. Arnold diffusion . . . . .                                                   | 245 |
| References . . . . .                                                            | 245 |

---

<sup>1</sup>Partially supported by NSF grant DMS-9704554.



## 1. Background: the action-angle variables, the generating functions

In addition to a quick review of general definitions, this section contains the discussion of specific examples of action-angle variables.

In this article we discuss Hamiltonian systems

$$\begin{cases} \dot{x} = H_y, \\ \dot{y} = -H_x \end{cases} \quad (1)$$

defined by the Hamiltonian  $H : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ . One can write (1) more compactly:

$$\dot{z} = J \nabla H, \quad (2)$$

where

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad z = \begin{pmatrix} x \\ y \end{pmatrix},$$

where  $I$  is the  $n \times n$  identity matrix,  $\nabla H$  is the gradient treated as a column vector and where  $x, y$  are column vectors.

A Hamiltonian system (1) is said to be *completely integrable* if it possesses  $n$  independent integrals in involution – that is, if there exist  $n$  functions  $I_k : \mathbb{R}^{2n} \rightarrow \mathbb{R}$  satisfying

1.  $\{H, I_k\} \equiv \langle J \nabla H, \nabla I_k \rangle = 0$  (that is,  $\frac{d}{dt} I_k = 0$ , where the derivative is taken along the flow of (1)).
2.  $\{I_k, I_j\} = 0$ , and
3. the gradients  $\nabla I_1, \dots, \nabla I_n$  are linearly independent for all  $z$  in a neighbourhood of a level surface  $I_k = c_k$ .

**The Liouville–Arnold theorem ([4])** Assume that (1) is completely integrable and that, in addition to the conditions 1-3 above, all the level sets  $I_k = c_k$  are compact for any choice of constants  $c_k, k = 1, \dots, n$  from an open set in  $\mathbb{R}^n$ . Then the level set  $\{z \in \mathbb{R}^{2n} : I_k(z) = c_k, k = 1, \dots, n\}$  is an  $n$ -torus, and there exists a symplectic transformation  $\varphi = z \mapsto (I, \theta) \in \mathbb{R}^n \times \mathbb{T}^n$  in which the Hamiltonian becomes  $\theta$ -independent:  $H \circ \varphi^{-1}(I, \theta) = K(I)$ , and our Hamiltonian system therefore takes the form

$$\begin{cases} \dot{I} = 0 \\ \dot{\theta} = \omega(I), \end{cases} \quad (3)$$

where  $\omega(I) = K'(I)$ , where  $K'$  denotes the gradient.

### 1.1. Examples of action-angle variables

In this section we give several specific examples of the action-angle variables.



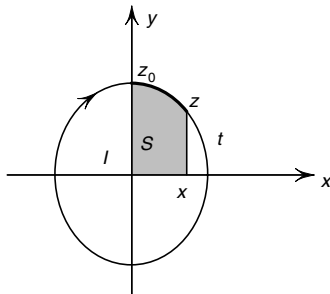


Fig. 1. Action-angle variables for the convex Hamiltonian in  $\mathbb{R}^2$ .

**1.1.1. Convex Hamiltonians with  $n = 1$  degree of freedom.** As our first example, let us consider the Hamiltonian system in  $\mathbb{R}^2$

$$\begin{cases} \dot{x} = H_y(x, y) \\ \dot{y} = -H_x(x, y) \end{cases}$$

where we assume that the level curves of  $H$  are convex curves enclosing the origin, Figure 1.

Now the action-angle map  $\varphi : z = (x, y) \mapsto (I, \theta)$  is constructed as follows. Given any point  $z \neq 0$  in the plane we define the action  $I = I(z)$  as the area enclosed by that level curve of  $H$  which passes through  $z$ :

$$I = \oint_{C_z} y \, dx,$$

where the level curve  $C_z = \{w : H(w) = H(z)\}$  is oriented clockwise. The angle variable  $\theta = \theta(z)$  is defined as the proportion of the period it takes for the solution to reach  $z$  from the positive  $y$ -axis:

$$\theta = t(z)/T(z); \tag{4}$$

here  $T(z)$  is the period of the orbit passing through  $z$  and  $t(z)$  is the time of travel from the positive  $y$ -axis to  $z$ . Since this time is defined modulo  $T(z)$ , the variable  $\theta$  is defined modulo 1.

The variable  $\theta$  has a purely geometrical meaning without a reference to time. Namely, consider an infinitesimally thin ring between the level curve of  $H$  passing through  $z$  and a nearby level curve of  $H$ . Consider also the segment of this ring lying between the  $y$ -axis and the vertical line through  $z$ . The angle  $\theta$  is then simply the ratio of the area of this segment to the area of the ring. This geometrical interpretation of  $\theta$  is equivalent to the definition (4) due to the fact that the Hamiltonian flow in the plane is incompressible, and thus, loosely speaking, the area swept by a moving short segment (which sweeps out the above-mentioned ring) increases in direct proportion to the time.

An equivalent construction of the action-angle variables via a generating function is given in Section 1.3.

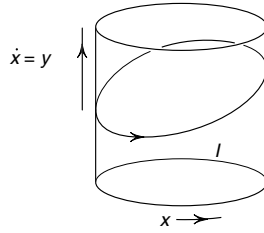


Fig. 2. The action variable for the case of periodic potential.

**1.1.2. Periodic Hamiltonians with one degree of freedom.** Let us now consider Hamiltonians satisfying the periodicity condition  $H(x, y) = H(x + 1, y)$ . The prime example is the pendulum, or, more generally, a particle in a periodic potential; the corresponding Hamiltonian is given by the kinetic plus potential energy, the latter being periodic:

$$H(x, y) = \frac{1}{2}y^2 + V(x), \quad V(x + 1) = V(x),$$

and we limit our attention to this example. Let us consider the ‘fast’ trajectories, corresponding to the ‘tumbling’ motions in the particular example of the pendulum. These trajectories correspond to the energy levels  $H > \max V$  chosen so that the particle has enough energy to climb the highest hill of the potential. Formally, from the conservation of energy  $y^2/2 + V(x) = E$  (here  $y = \dot{x}$ ), we conclude that  $\dot{x} = y = \pm\sqrt{2(E - V)} \neq 0$  for all time, and the corresponding trajectory wraps around the phase cylinder  $\mathbb{T} \times \mathbb{R} = \{(x \bmod 1, y)\}$ , as shown in Figure 2.

The action variable  $I$  corresponding to the point  $(x, y)$  is the area under the trajectory on the phase cylinder:

$$I = \int_0^1 y \, dx,$$

where the integration is taken over the energy curve passing through the point  $(x, y)$ , Figure 2. The angle variable  $\theta$  is the time normalized by the period of one full revolution:

$$\theta = \frac{1}{T} \int_0^x dx/y = \frac{\int_0^x dx/y}{\int_0^1 dx/y},$$

where, again, the integration is taken over the same energy curve. Alternatively, the angle variable can be defined via infinitesimal areas as described in the preceding example.

## 1.2. Action-angle variables for particles in $\mathbb{R}^n$

Consider a particle in  $\mathbb{R}^n$  moving under the influence of a potential force:

$$\ddot{x} = -\nabla V(x), \quad x \in \mathbb{R}^n. \quad (5)$$

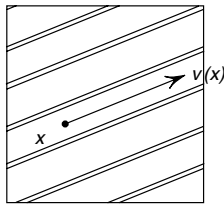


Fig. 3. An invariant 2-torus in  $\mathbb{R}^2 \times \mathbb{T}^2$  is an invariant vectorfield on the configuration torus.

Let us assume  $V$  to be periodic of period 1 in each of its variables, and let us treat each coordinate of  $x$  as an angle variable on the torus  $\mathbb{T}^n = \mathbb{R}^n(\text{mod } 1)$ . For our purposes it suffices to consider the cases of  $n = 2$  and  $n = 3$  only since these two cases illustrate all the phenomena discussed in this article.

In the simple special ‘unperturbed’ case of  $V = \text{const.}$  we have a free particle, and the associated Hamiltonian system (1) with  $y = \dot{x}$  and  $H = y^2/2$  is completely integrable. The action variables are the components of the velocity  $I = y = (y_1, \dots, y_n)$ , and the angle variables are the coordinates of  $x$ .

**The geometrical interpretation of the invariant tori** Without the loss of generality, let us consider the case  $n = 2$ . The invariant tori live in the phase space  $\mathbb{T}^2 \times \mathbb{R}^2$ . It is easier to visualize these tori as follows. Each invariant torus is given by

$$\{(x, y) : x \in \mathbb{T}^2, y = (v_1, v_2) = \text{const.}\}.$$

In other words, *an invariant torus in  $\mathbb{T}^2 \times \mathbb{R}^2$  corresponds precisely to a constant vectorfield on the torus  $\mathbb{T}^2$* . The same interpretation survives a perturbation of the potential  $V$ ; an invariant torus (if it exists) in the perturbed case is again a vectorfield  $v(x)$  on the configuration torus  $\mathbb{T}^2$  which is invariant under the dynamics of the system. The invariance means that if a solution  $x(t)$  of (5) satisfies  $\dot{x}(t) = v(x(t))$  for one value of  $t$ , then it satisfies the same relation for all  $t$ . In particular, an invariant torus with an irrational rotation number is manifested as a dense trajectory on  $\mathbb{T}^2$  without self-intersections (see Figure 3).

### 1.3. A generating function construction of the action-angle variables

We describe now an alternative construction of the action-angle variables, well suited to the non-autonomous case

$$\begin{cases} \dot{x} = H_y(x, y, t) \\ \dot{y} = -H_x(x, y, t) \end{cases} \quad (6)$$

referred to as the case of 1.5 degrees of freedom. As a side remark, for the higher degrees of freedom the action-angle variables are only defined in the exceptional case of complete

integrability (such as the preceding example), and the particular interpretation of these variables depends on the system.

Let  $S(x, I, t)$  be the area (Figure 1) bounded by the level curve  $H$  enclosing the area  $I$ , the  $y$ -axis and the line  $x = \text{const}$ . More precisely, we consider the level curve of  $H$  enclosing the area  $I$ , and let  $C_{x,I,t}$  be the arc of this curve lying above the segment  $[0, x]$ ; we define

$$S(x, I, t) = \int_{C_{x,I,t}} y \, dx. \quad (7)$$

We can now define the action-angle variables  $(\theta, I)$  implicitly via

$$S_x(x, I, t) = y, \quad S_I(x, I, t) = \theta. \quad (8)$$

This map is symplectic: indeed, using (8) gives

$$\begin{aligned} dx \wedge dy &= dx \wedge (S_{xx}dx + S_{xI}dI) = S_{xI}dx \wedge dI, \\ d\theta \wedge dI &= (S_{Ix}dx + S_{II}dI) \wedge dI = S_{Ix}dx \wedge dI \end{aligned}$$

so that  $dx \wedge dy = d\theta \wedge dI$ .

The action-angle variables satisfy the Hamiltonian system with the Hamiltonian

$$K(\theta, I, t) = H \circ \varphi^{-1}(\theta, I, t) + S_t(x, I, t),$$

where  $\varphi := (x, y; t) \mapsto (\theta, I; t)$  is the map defined by (8), and where  $x = x(\theta, I, t)$  is defined implicitly by (8).

As a side remark, in the autonomous case, we have  $K'(I) = 1/T$ , where  $T$  is the period of the trajectory enclosing the area  $I$ . This can be seen directly as follows. Let us consider two nearby orbits, i.e. two level curves  $\gamma : H = c$  and  $\gamma_+ : H = c + \varepsilon$ . The width of the thin annulus between  $\gamma$  and  $\gamma_+$ , measured along the perpendicular to  $\gamma$ , is  $\varepsilon/|\nabla H|$  (we drop the higher order terms in  $\varepsilon$  throughout this paragraph), and the area of the annulus is thus  $dI = \varepsilon \int ds/|\nabla H|$ , where  $ds$  is the length along  $\gamma$ . But the speed of the solution is  $|\nabla H|$ , and thus  $ds = |\nabla H|d\tau$ , where  $\tau$  is the time measured along the solution (with the time  $t$  appearing in the right-hand side of (6) frozen). Thus  $dI = \varepsilon \int d\tau = \varepsilon T$ , and thus  $K'(I) = dH/dI = 1/T$ , as claimed.

#### 1.4. A mechanical interpretation of generating functions

The generating function has an illuminating mechanical interpretation. Consider a seemingly very specific example shown in Figure 4. Two bars can slide without friction along the horizontal line. The bars are attached to each other and to the walls by springs, as shown in the figure.

The potential  $P(x, X)$  of the mechanical system is a function of the positions  $x, X$  of the sliding bars. Let us hold the bars in fixed positions  $x, X$  as shown; this requires that we

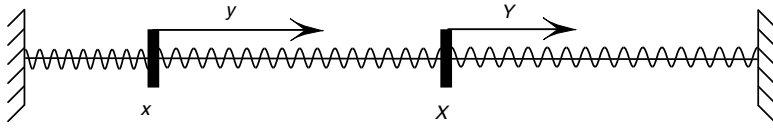


Fig. 4. The bars attached to springs can slide on the line. The forces required to keep the bars in respective positions  $x$ ,  $X$  are denoted by  $y$  and  $-Y$  correspondingly. The potential energy  $P(x, X)$  of the system is then a generating function for a symplectic map  $(x, y) \mapsto (X, Y)$ . This mapping is defined by (9).

apply some force to prevent the bars from moving. The bars are trying to move with forces  $f = -P_x(x, X)$  and  $F = -P_X(x, X)$ , where the subscripts denote partial derivatives. The last two relations follow from the definition of potential energy. These relations define the map  $\psi : (x, f) \mapsto (X, F)$ , provided  $P_{xX} \neq 0$ . Note that  $\psi$  preserves area, up to a change in sign: indeed, imagine holding the bar at  $x$  with the left hand, and the bar at  $X$  with the right hand. Then move both hands in some periodic fashion, returning to the original position. During the motion, both points:  $(x, f)$  and  $(X, F)$  describe closed paths  $c$  and  $C$  in the plane, with  $C = \psi(c)$ . But the net work we would do in such a motion is zero:

$$\int_c f dx + \int_C F dX = 0,$$

where the first term is the work done by the left hand and the second term is the work done by the right hand. But this mechanical statement expresses the fact that the two areas are equal, up to a sign. To avoid this sign reversal, we change the sign of one of the forces, by setting  $Y = -F$ . For the sake of uniformity of notation, let  $y = f$  (no sign change here). The resulting mapping  $\varphi : (x, y) \mapsto (X, Y)$  is defined by

$$y = -P_x(x, X), \quad Y = P_X(x, X). \quad (9)$$

We showed by a mechanical argument that this mapping preserves area:

$$\int_c y dx = \int_c Y dX = \int_{\varphi(c)} y dx.$$

The interpretation extends trivially to higher dimensions.

In most applications the generating function  $P$  arises not as the energy of a system of springs, but rather as the action of a Lagrangian (the difference of the kinetic and potential energies) of a moving mechanical system:

$$P(x, X) = \min_{x(\cdot)} \int_{t_0}^{t_1} L(x(t), \dot{x}(t)) dt, \quad x(t_0) = x, \quad x(t_1) = X.$$

As a side remark, any such action can be realized as the potential energy of a certain elastic spring in a potential field. In that case the momenta defined by  $p(t_0) = -P_x = L_{\dot{x}}|_{t=t_0}$ ,  $p(t_1) = P_X = L_{\dot{x}}|_{t=t_1}$  acquire the meaning of forces at the ends of this spring!

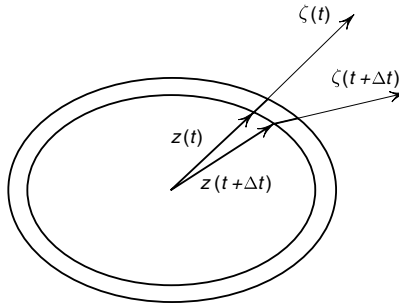


Fig. 5. Towards the proof of the twist condition (10).

### 1.5. The twist condition

Twist is a crucial property required for application of KAM theory. In this section we give a simple sufficient condition for twist for Hamiltonian systems with one degree of freedom. We consider the two cases: (1) the integral curves are closed in the plane, as in, say, a cubic potential, and (2) the integral curves encircle the cylinder, as in the case of the pendulum.

**Convexity of the potential and the twist** Let us first consider an important class of 'Newtonian' Hamiltonians  $H(x, y) = y^2/2 + V(x)$  with  $V$  either sub- or super-quadratic, in the sense made precise shortly. If  $V$  is a quadratic function, the system is linear and there is no twist. It is thus natural to expect that the non-quadratic growth of the potential should lead to twist; the following theorem makes this intuition precise.

Recall that the action  $I$  associated with a point  $(x, y)$  is the area enclosed by the level curve through  $(x, y)$ , while the angle  $\theta = t/T$  is the time, measured in the units of the period that it takes for the solution to travel from the positive  $y$ -axis to  $(x, y)$ . The system with the Hamiltonian  $H$ , written in the action-angle form (3) has  $\omega = \frac{1}{T}$ , and the twist condition  $\omega'(I) \neq 0$  amounts to  $T'(I) \neq 0$ , implying the monotone dependence of the period on the area, or on the amplitude of the periodic solution. Here is a convenient sufficient condition for the twist.

**THEOREM 1.** *If the potential  $V$  satisfies*

$$\frac{V'}{x} < V'', \quad (10)$$

*for all  $x \neq 0$  in some interval  $[-a, a]$ , then the period  $T$  of solutions of the Hamiltonian system*

$$\begin{cases} \dot{x} = y \\ \dot{y} = -V'(x) \end{cases} \quad (11)$$

*is a monotonic decreasing function of the amplitude, for those solutions which satisfy  $-a < x(t) < a$  for all  $t$ . If the opposite of the inequality (10) holds, then the period is a monotonic increasing function of the amplitude (see Figure 5).*

For the pendulum we have  $V' = \sin x$ , and the opposite of (10) holds:  $\sin x/x > \cos x$ , or  $|\tan x| > |x|$  (for  $x \neq 0$ ). We conclude that the period of the pendulum is a monotonic increasing function of the amplitude, for the class of oscillatory solutions. In the cases of superquadratic potentials  $V(x) = x^4$ , or  $V(x) = x^2 + x^4$ , or more generally,  $V(x) = \sum_{k=0}^n a_k x^{2k}$ ,  $a_k \geq 0$ , our criterion (10) shows that the period is a decreasing function of the energy, while for the subquadratic potential  $V(x) = x^\alpha$  with  $1 < \alpha < 2$  the period increases with the energy.

**PROOF OF THEOREM 1.** Let us write the linearization of (11) along a solution  $z = \text{col}(x, y)$ :

$$\begin{cases} \dot{\xi} = \eta \\ \dot{\eta} = -V''(x)\xi. \end{cases} \quad (12)$$

The linearized solution  $\zeta(t) = \text{col}(\xi, \eta)$  describes the infinitesimal difference between two nearby solutions of (11). Consider now any solution  $z(t)$  of (11) along with the associated solution  $\zeta(t)$  of the linearized system, with  $\zeta(0) \parallel z(0)$ . If the vector  $\zeta(t)$  turns clockwise faster than  $z(t)$  at  $t = 0$ , then the solutions on larger curves travel with higher angular velocity and thus the period of a solution  $z(t)$  is a decreasing function of the amplitude, or of the area enclosed by the orbit in the phase plane. The condition on the angular velocities of  $z$  and  $\zeta$  can be written as

$$\frac{d}{dt} \left( \frac{y}{x} \right) > \frac{d}{dt} \left( \frac{\eta}{\xi} \right), \quad \text{whenever} \quad \frac{y}{x} = \frac{\eta}{\xi}.$$

This condition reduces to the desired condition on  $V$  without reference to solutions, as follows. Carrying out the differentiations, we obtain

$$\frac{x\dot{y} - \dot{x}y}{x^2} > \frac{\dot{\eta}\xi - \dot{\xi}\eta}{\xi^2} \quad \text{when} \quad \frac{y}{x} = \frac{\eta}{\xi}.$$

Using (11) and (12) to eliminate the derivatives, and using the collinearity:  $\frac{y}{x} = \frac{\eta}{\xi}$ , we obtain (10).  $\square$

**Twist in planar Hamiltonian systems** The same idea leads to a twist condition for Hamiltonian vector fields in the plane (not necessarily of the kinetic+potential form):

$$\begin{cases} \dot{x} = H_y(x, y) \\ \dot{y} = -H_x(x, y), \end{cases} \quad (13)$$

or  $\dot{z} = J\nabla H(z)$ , where  $z = \text{col}(x, y)$ ,  $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  and  $\nabla \equiv \text{grad}$ . Assume that the level curves of  $H$  are closed, convex and enclose the origin, which is the rest point:  $\nabla H(0, 0) = 0$ .

Then all solutions of (13) are periodic. Denote by  $T(E)$  the period of the solution on the level curve  $H(z) = E$ .

**THEOREM 2.** *In the assumptions of the last paragraph, if  $\langle H''(z)z, z \rangle < \langle \nabla H(z), z \rangle$ , i.e.*

$$H_{xx}x^2 + 2H_{xy}xy + H_{yy}y^2 < H_x x + H_y y, \quad (14)$$

*then  $T(E)$  is a monotonic increasing function of  $E$ . If the opposite of (14) holds, then  $T(E)$  is a monotonic decreasing function.*

The proof is exactly the same as for the preceding theorem.

**Hamiltonians on a cylinder** Consider now the class of Newton-type Hamiltonians  $y^2/2 + V(x)$  with periodic potentials  $V(x+1) = V(x)$ .

**THEOREM 3.** *The time advance map  $\varphi^t : (x, y) \mapsto (X(t; x, y), Y(t; x, y))$  for the system (11) with a periodic potential  $V$  is a monotonic twist map in the sense that*

$$\frac{\partial}{\partial y} X(t; x, y) > 0 \quad (15)$$

*if  $t > 0$ , for all  $(x, y)$  satisfying*

$$y^2/2 + V(x) > \sup_{x \in \mathbb{R}} V. \quad (16)$$

**PROOF.** Let  $Z(t) = (X(t; x, y), Y(t; x, y))$  be a solution of (11) with initial conditions  $(x, y)$  satisfying (16). This means that the solution has enough energy to clear every hill of the potential and thus will travel always to the left or always to the right in the phase plane. Assume, without the loss of generality, that  $\frac{\partial}{\partial t} X(t; x, y) > 0$ . This means that the velocity  $\dot{Z} = (\dot{X}, \dot{Y})$  lies in the right-half plane for all  $t$ . Note also that  $\dot{Z}$  is a solution of the linearized system (12), and we thus established that *the linearized system possesses a solution that lies in the right half-plane for all time*. To finish the proof, note that the  $y$ -derivative

$$\zeta = (\xi, \eta) = \frac{\partial}{\partial y} Z(t; x, y) \quad (17)$$

(we suppress the initial conditions  $(x, y)$  from now on) satisfies the linearized system. Consider the time-dependent sector  $S(t)$  defined by the vectors  $\dot{Z}(t)$  and  $(0, 1)$ . Since the vector field (12) crosses the positive  $\eta$ -axis into the first quadrant, any solution of (12) starting in  $S(0)$  at  $t = 0$  lies in  $S(t)$  at time  $t > 0$ . Thus  $\zeta(t)$  lies in  $S(t)$ . But both rays forming  $S(t)$  lie in the right-half plane, and thus  $\zeta(t)$  lies in the right half-plane, which proves (15).  $\square$

## 2. Basic statements of KAM theory

In this section we state the two results: Moser's twist theorem and the Kolmogorov–Arnold theorem. These and related results form what is now referred to as the Kolmogorov–Arnold–Moser (KAM) theory. In the following section we describe a variational approach to



Moser's twist theorem, referring to the paper of Salamon and Zehnder [48] for a similar variational proof of the Kolmogorov–Arnold theorem.

### 2.1. Moser's twist theorem

Consider a perturbation of the twist mapping:

$$\varphi : (\theta, I) \mapsto (\theta + \alpha(I), I) + r(\theta, I) \quad (18)$$

where the perturbation  $r$  is small and periodic in  $\theta$  of period 1. It is assumed that all the functions are defined for all  $I \in (0, 1)$  and for all  $\theta \in \mathbb{R}$ , and, moreover, that the twist condition

$$\frac{d}{dI}\alpha(I) > 0$$

holds for all  $I \in (0, 1)$ . The mapping  $\varphi$  is a lift of an annulus map. In the unperturbed case where  $r = 0$  the annulus  $0 < I < 1$  is foliated by invariant circles  $I = \text{const}$ . It is easy to produce an arbitrarily small area-preserving perturbation that will destroy all the invariant circles – for example, a radial perturbation  $r = (0, \varepsilon)$ . We forbid such ‘non-exact’ perturbations by insisting that  $\varphi$  preserves area under any circle  $I = f(\theta)$ , where  $f$  is periodic of period 1. In the statement below,  $|\cdot|_s$  denotes the  $C^s$ -norm of a function, i.e. the supremum of the sum of absolute values of all derivatives up to order  $s$ .

**THEOREM 4 (Moser's Twist Theorem).** *Given any  $\varepsilon > 0$ ,  $\tau > 0$  and  $s \geq 1$ , let  $\omega$  satisfy the Diophantine conditions*

$$\left| \omega - \frac{p}{q} \right| \geq \varepsilon \frac{1}{q^{2+\tau}}, \quad (19)$$

*for all integers  $p, q$  with  $q > 0$ , and  $\alpha(0) + \varepsilon < \omega < \alpha(1) - \varepsilon$ . Let  $\alpha(I_0) = \omega$ . Then the mapping (18) has an invariant curve with the rotation number  $\omega$ :*

$$\gamma : u \mapsto (u, I_0) + \rho(u),$$

*where  $\rho$  is a periodic function of period one with  $|\rho|_s < \varepsilon$ , provided the following hypotheses are satisfied.*

1. *Every non-contractible closed curve on the annulus intersects its image under  $\varphi$ .*
2. *The twist is bounded from below and from above:*

$$c_0^{-1} \leq \alpha'(I) \leq c_0.$$

3. *The perturbation  $r$  is smaller than a certain constant  $\delta_0$  depending on  $\varepsilon, s, c_0$ :*

$$|r|_0 < \delta_0,$$

*with a bound on the higher derivatives up to order 4:*

$$|\alpha|_4 + |r|_4 < c_0.$$

Moreover, the mapping  $\varphi$  restricted to the invariant curve  $\gamma$  is conjugate to a rotation:

$$\varphi(\gamma(u)) = \gamma(u + \omega).$$

As stated, this theorem is actually a strengthening of Moser's original theorem [40]; see Herman [19] for further references and details.

## 2.2. The Kolmogorov–Arnold theorem

The Hamiltonian of a completely integrable Hamiltonian system written in the action-angle variables depends on the action alone. We will now consider perturbations of a completely integrable Hamiltonian system with the Hamiltonian

$$H(I, \varphi) = h(I) + R(I, \varphi), \quad (20)$$

where the remainder  $R$  is going to be small. We assume  $h$  and  $R$  to be real analytic functions of their arguments. To make the concept of smallness precise, we introduce a supremum norm as follows. Given  $\rho > 0$ , we define

$$\|R\|_\rho \stackrel{\text{def}}{=} \sup |R(I, \varphi)|,$$

where the supremum is taken over the complex neighbourhood  $|\text{Im } I| \leq \rho, |\text{Im } \varphi| \leq \rho$ .

We also need a vector version of the Diophantine condition. Given positive constants  $C$  and  $\sigma$ , we will say that the vector  $\omega \in \mathbb{R}^n$  is of  $(C, \sigma)$ -type if it satisfies

$$|\langle \omega, k \rangle| \geq \frac{C}{|k|^\sigma} \quad (21)$$

for all nonzero vectors  $k \in \mathbb{Z}^n$ .

**THEOREM 5 (Kolmogorov–Arnold).** *Assume that the Hamiltonian (20) is a real analytic function on the set  $|\text{Im } I| \leq \rho, |\text{Im } \varphi| \leq \rho$  for some  $\rho > 0$ . Assume that for some  $I_0$  the frequency vector  $h'(I_0) = \omega$  satisfies Diophantine conditions (21) for some positive  $C, \sigma$ . Finally, assume that the Hessian  $h''$  satisfies*

$$|\det h''(I)| \geq \mu > 0$$

*in a neighbourhood of  $I_0$ . Then, if  $\|R\|_\rho$  is small enough, the Hamiltonian system corresponding to the Hamiltonian (20) possesses an invariant  $n$ -torus  $\mathcal{T}_\omega$  given by  $I = I_0 + f(\varphi)$  where  $f$  is analytic and small, and any solution with initial conditions on  $\mathcal{T}_\omega$  is dense on  $\mathcal{T}_\omega$ . In fact, the flow on  $\mathcal{T}_\omega$  is analytically conjugate to the rigid translation  $\theta \mapsto \theta + \omega t$  on the  $n$ -torus.*

## 3. A variational approach to Moser's twist theorem

In the variational approach to KAM theory one reduces finding an invariant torus in  $\mathbb{R}^n \times \mathbb{T}^n$  to solving the Euler–Lagrange equation of Percival's variational functional. A physical

motivation for this approach is suggested by the Frenkel–Kontorova model. For the case of mappings, covered by Moser’s twist theorem, the existence of KAM tori is reduced to solving a simple-looking second order difference equation. In this section we outline this reduction for area-preserving maps. We do not discuss here the solution of the difference equation itself; the details can be found in [25] for maps and in [48] for flows.

### 3.1. The generating function of an area-preserving map

Consider an area-preserving twist mapping  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  of the covering plane of the cylinder  $\mathbb{R}(\text{mod } 1) \times \mathbb{R}$ . Assume that  $\varphi : (x_1, y_1) \mapsto (x_2, y_2)$  has a generating function  $h$ :

$$\begin{aligned} h_1(x_1, x_2) &= -y_1, \\ h_2(x_1, x_2) &= y_2. \end{aligned} \tag{22}$$

Here  $h_1, h_2$  are the derivatives with respect to the first or second argument of  $h$ . Similarly we will denote the second derivatives by  $h_{11}, h_{12}, h_{22}$ .

A large class of area-preserving mappings is given by (22), according to the following theorem. This class includes all area preserving maps that are covered by Moser’s twist theorem.

**THEOREM 6.** *Any smooth twist cylinder map  $\varphi$  satisfying the monotonic twist condition  $\partial x_2 / \partial y_1 > 0$  possesses a generating function  $h$  with  $h_{12} < 0$  such that the map is given by (22) implicitly. The map is exact:  $\int_{\varphi(\gamma)} y dx = \int_{\gamma} y dx$ , where  $\gamma$  is an arbitrary smooth noncontractible circle on the cylinder, if and only if  $h(x_1 + 1, x_2 + 1) = h(x_1, x_2)$ .*

### 3.2. Reduction to a difference equation

The problem of finding an invariant curve of a map  $\varphi$  given by (22) reduces to a difference equation, as follows. We seek an invariant curve in the parametric form  $w(\theta) = (u(\theta), v(\theta))$ , where  $u(\theta + 1) = u(\theta) + 1$  is monotonic in  $\theta$ , and  $v(\theta + 1) = v(\theta)$ . We seek  $w$  satisfying

$$\varphi(w(\theta)) = w(\theta + \omega), \tag{23}$$

with a prescribed rotation number  $\omega$ .

**THEOREM 7.** *The curve  $w(\theta) = (u(\theta), v(\theta))$  satisfies the invariance condition (23) under the map  $\varphi$  given by (22) if and only if the horizontal coordinate  $u(\theta)$  satisfies the second order difference equation*

$$E(u(\theta)) \equiv h_1(u(\theta), u(\theta + \omega)) + h_2(u(\theta - \omega), u(\theta)) = 0. \tag{24}$$

*The vertical coordinate  $v$  is then given by  $v(\theta) = -h_1(u(\theta), u(\theta + \omega))$ .*

PROOF. Since  $\varphi$  is given by (22), the condition (23) is equivalent to the system

$$\begin{cases} h_1(u, u^+) = -v \\ h_2(u, u^+) = v^+, \end{cases} \quad (25)$$

where  $u = u(\theta)$ ,  $u^+ = u(\theta + \omega)$ , etc. Replacing  $\theta$  by  $\theta - \omega$  in the second equation and adding it to the first, we obtain (24).  $\square$

REMARK 8. The Equation (24) is the Euler–Lagrange variational equation for the variational problem  $\delta \int_0^1 h(u, u^+) d\theta = 0$  (Percival's variational principle).

REMARK 9. The mean value of  $u_\theta E(u)$  is zero:

$$\int_0^1 u_\theta E(u) d\theta = 0 \quad (26)$$

as follows from the invariance of the Lagrangian  $\int_0^1 h(u, u^+) d\theta$  under  $\theta$ -translations  $u(\theta) \mapsto u(\theta + c)$ , or from the identity

$$u_\theta E(u) = \frac{\partial}{\partial \theta} h(u, u^+) - \nabla(u_\theta h_2(u^-, u))$$

where  $\nabla f = f(\theta + \omega) - f(\theta)$ . Integration gives the claim.

EXAMPLE. Consider the standard map

$$\begin{aligned} x_2 &= x_1 + y_1 + V'(x_1) \\ y_2 &= y_1 + V'(x_1); \end{aligned}$$

with a periodic function  $V$  of period 1. The generating function is

$$h(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2 + V(x_1),$$

and the Euler–Lagrange difference equation (24) takes a particularly simple form

$$u(\theta + \omega) - 2u(\theta) + u(\theta - \omega) = V'(u(\theta)).$$

Here is a quick sketch outlining how a statement on the difference equation (24) leads to a proof of Moser's twist theorem. One first proves that near an approximate solution  $u_0(\theta)$  of (24) (approximate in the sense that  $E(u(\theta))$  is sufficiently small in a certain norm) there exists an exact solution  $u$ . This statement requires some assumptions on  $E$ , one of which is the Diophantine nature of  $\omega$ . Now this theorem implies Moser's twist theorem as follows. Consider a generating function for Moser's twist map, fix a Diophantine  $\omega$  and consider the corresponding Euler–Lagrange operator  $E$ . The obvious fact that any unperturbed circle is 'approximately invariant' under the twist map implies that the function  $u_0$  corresponding to this circle is an 'approximate solution' of  $E(u) = 0$ . There exists therefore an exact solution near  $u_0$ , and this exact solution gives rise to an invariant curve, thus proving Moser's twist theorem. All the details of this can be found in [25] (or [48] in the Hamiltonian case).

## 4. Applications

This section lists some applications of KAM theory. Our goal here is to avoid generality and to rather focus on maximally simple representative examples. More general statements, as well as more detailed history, can be found in the references mentioned in the text.

EXAMPLE 1. Consider the particle in a periodic potential  $V$  in  $\mathbb{R}^n$ :

$$\ddot{x} = -\varepsilon \nabla V(x), \quad x \in \mathbb{R}^n, \quad (27)$$

where  $V$  is periodic of period one in each of its variables. To be specific, let us fix the energy  $\dot{x}^2/2 + V(x) = 1$ . We ask the same questions as in the previous example. (1) For  $\varepsilon = 0$  every action remains constant:  $\dot{x} = \text{const}$ . Is it true that for  $\varepsilon \neq 0$  every action  $\dot{x}$  remains close to its initial value for all time? (2) For  $\varepsilon = 0$   $x(t)$  is a quasiperiodic function of  $t$ . What can be said for  $\varepsilon \neq 0$ ?

The best known answer to (1) is this: For  $n = 2$ , the actions stay close to their initial values for *all* solutions, for  $\varepsilon$  sufficiently small. For  $n \geq 3$ , for most (in the sense of Lebesgue measure) initial conditions the actions remain nearly constant and the corresponding solutions are quasiperiodic, for  $\varepsilon$  sufficiently small. However, there exist potentials  $V$  such that some actions do not remain close to their original values, no matter how small  $\varepsilon$  is. This phenomenon of action drift is referred to as *Arnold diffusion*, and we mention a little more on it at the end of this chapter.

EXAMPLE 2. The system

$$\ddot{x} + (2 + \cos t) \sin x = 0$$

can be interpreted as a pendulum whose pivot undergoes periodic oscillations in the vertical direction. It doesn't seem unreasonable to conjecture that resonance will cause some solutions to gain speed without bound. Do such unbounded solutions exist? KAM theory gives a negative answer, as described later. This system is not close to completely integrable, and it may seem surprising that the KAM theory applies. It turns out, however, that this system is close to a completely integrable one in the range of large  $|\dot{x}|$ .

EXAMPLE 3. Consider the system

$$\ddot{x} + x^3 = \varepsilon \cos t.$$

For small  $\varepsilon$  this system is obviously close to a completely integrable one and KAM theory applies directly, with the resulting claim that there the Poincaré map  $\varphi_\varepsilon : (x, \dot{x})_{t=0} \mapsto (x, \dot{x})_{t=2\pi}$  has 'many' invariant circles. Any solution starting between two such circles stays between them, thus remaining bounded for all time. In addition, any solution starting on an invariant circle is quasiperiodic with two frequencies. But what if  $\varepsilon$  is not small, say,  $\varepsilon = 1$ , or  $\varepsilon = 10$ ? Surprisingly, even in that case all solutions stay bounded for all time. This result, due to Morris [39], follows from the fact that the Poincaré map is close to an integrable one near infinity.

EXAMPLE 4 (A Large Perturbation). The perturbation in the preceding example was of bounded magnitude. Let us consider a much more violently perturbed oscillator:

$$\ddot{x} + (2 + \cos t)x^3 = 0.$$

This perturbation is vastly stronger than the one in the previous example. Nevertheless, even here the solutions are all bounded for all time, as has been shown in [21,11,23]. It turns out that even in this case the system is close to a completely integrable one near infinity, after an appropriate (somewhat delicate) reduction.

**A more general theorem for potentials** The following theorem [23] includes all of the above examples and much more.

THEOREM 10. Assume that the function  $V(x, t) = V(x, t + 1)$  tends to  $\infty$  as  $|x| \rightarrow \infty$  and that it satisfies conditions (28)–(30) below. Then the equation

$$\ddot{x} + V_x(x, t) = 0$$

is near-integrable for large energies; more precisely, for any  $0 < \omega < 1$  satisfying the Diophantine conditions

$$\left| \omega - \frac{p}{q} \right| \geq K|q|^{-5/2} \quad \text{for all } 0 \neq q, p \in \mathbb{Z}$$

(where  $K > 0$  is independent of  $p, q$ ), the Poincaré map  $P : (x, \dot{x})_{t=0} \mapsto (x, \dot{x})_{t=1}$  possesses countably many invariant circles with rotation number  $\omega$ . These circles cluster at infinity in the  $(x, \dot{x})$ -plane, and their relative measure near infinity approaches full measure. Each of these circles is a section of an invariant torus in the extended phase space  $(x, \dot{x}, t \bmod 1)$ . The flow restricted to each torus is quasiperiodic with basic frequencies 1 and  $k + \omega$ , where  $k$  is an integer. All integers  $k \geq k_0(\omega)$  are represented by an invariant torus. In particular,

1. All solutions are bounded for all time:  $\sup_{\mathbb{R}} (|x| + |\dot{x}|) < \infty$ .
2. Most solutions with large amplitudes are quasiperiodic, i.e. most (in the sense of Lebesgue measure) initial conditions with large  $|x(0)| + |\dot{x}(0)|$  give rise to quasiperiodic solutions:  $x(t) = f(t, \omega t)$  where  $f$  is a function periodic in both variables.
3. Any sufficiently large number is a rotation number for some solution, i.e. there exists  $\rho_0$  such that for any  $\rho > \rho_0$  there exists a solution  $x_\rho(t)$  with that rotation number.<sup>1</sup>

**Conditions on the potential**  $V(x, t)$  For some positive constants  $c, c_1, a$  and  $0 \leq \mu < \frac{1}{100}(\frac{1}{2} - a)$  and for all  $x$  and  $t$  we have

$$|\partial_x^k \partial_t^\tau V| \leq c|x|^{-k} V^{1+\mu}, \quad k + \tau \leq 6, \quad (28)$$

$$-\frac{1}{2} + c_1 \leq \partial_x W \leq a < \frac{1}{2}, \quad \text{where } W = V/V_x, \quad (29)$$

<sup>1</sup>This statement is a consequence of the Aubry–Mather theory, [6,32].

and

$$|\partial_x^k \partial_t^\tau U|, \quad |\partial_x^k \partial_t^\tau W| \leq c|x|^{1-k}, \quad k + \tau \leq 5, \quad \text{where } U = V_t / V_x. \quad (30)$$

REMARK. Even a very special case of  $\mu = 0$  includes, together with the polynomial, potentials with exponential growth e.g.  $V(x, t) = (2 + \cos t) \cosh x$ . If  $\mu > 0$  then potentials with spatial oscillations are allowed as well; an example is  $V = p(t)(x + \cos \sqrt{x})^{2n}$ , where  $n$  is sufficiently large.

More details on this as well as the proof can be found in [23].

**Littlewood's counterexample** In the early 1960s Littlewood produced an example of a periodically forced oscillator with superquadratic potential:

$$\ddot{x} + V'(x) = p(t), \quad p(t+1) = p(t), \quad \lim_{x \rightarrow \infty} V'(x)/x \rightarrow \infty$$

possessing unbounded solutions. A geometrical consequence of the superquadratic growth of  $V$  is a twist property of the period map of the plane, in the sense that the angle by which an initial vector  $(x, y)$  rotates during one period tends to infinity as  $|(x, y)| \rightarrow \infty$ . It follows that the period map does not satisfy the assumptions of Moser's twist theorem. Littlewood's example has a discontinuous  $p$ . Zharnitsky [51] extended Littlewood's counterexample to the considerably more delicate case of continuous forcing  $p$ . The (non-fatal) errors in Littlewood's original paper [27] were corrected by Y. Long [30]. A much simplified analysis of Littlewood's counterexample with an optimal estimate on a possible  $V$  which makes such an example possible, is given in [24].

EXAMPLE 5 (A Quasiperiodic Perturbation). Consider the system in Example 4 with quasiperiodic time dependence, replacing  $2 + \cos t$  with  $f(t) = F(\omega_1 t, \dots, \omega_n t)$ , where  $F(\xi_1, \dots, \xi_n) > 0$  is a function periodic of period one in each of its  $n$  variables, and where  $\omega_k$  are rationally independent frequencies satisfying, moreover, an appropriate Diophantine condition:

$$\ddot{x} + f(t)x^3 = 0.$$

For example, one can take  $f(t) = 3 + \cos t + \cos \sqrt{2}t$ . Even in this case all solutions are bounded for all time, and most solutions with large initial conditions are quasiperiodic.

A much more general class of oscillators with quasiperiodic forcing has been analyzed in [26].

EXAMPLE 6 (Fermi-Ulam's Ping-pong). A particle bounces between two parallel walls, moving along a line perpendicular to the walls. The walls oscillate periodically, both with the same period. Between the collisions the particle moves with constant velocity. The collisions are perfectly elastic, which means that, at the moment of collision, the particle changes only the sign of its velocity when viewed from the reference frame of the wall:  $v_2 - w = -(v_1 - w)$ , where  $w$  is the speed of the wall and  $v_1, v_2$  are the speeds before and after collision, all measured in the frame attached to the origin. This gives  $v_2 = -v_1 + 2w$ :

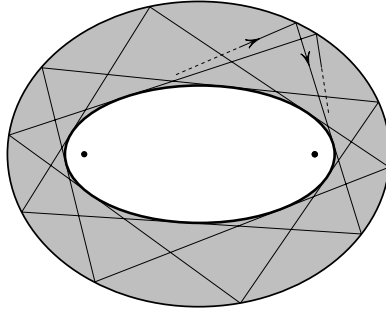


Fig. 6. An invariant circle for a billiard map corresponds to a caustic.

the ball gains twice the speed of the wall in collision. It is natural to ask whether there are initial conditions for which the particle can have unbounded speed, i.e. whether it is possible for a particle to gain an unbounded amount of energy from the moving walls. If the motion of the walls is sufficiently smooth, then the answer, according to KAM theory, is negative: for any initial condition the velocity is bounded for all time. The proof can be found in [16,47,21]. A similar result for quasiperiodically moving walls has been proven in [52].

EXAMPLE 7. Instead of a particle bouncing between two walls, consider a ball bouncing up and down on a periodically oscillating racket. If the oscillations of the racket are smooth, then any sufficiently small amplitude bounce will remain of small amplitude, bounded from both above and below, for all time. Most small amplitude oscillations are quasiperiodic [52].

EXAMPLE 8. Consider a billiard with a smooth boundary close to an ellipse, in the  $C^{k+1}$ -norm. The corresponding billiard map is then  $C^k$ -close to the integrable billiard map of the ellipse. Integrability of the latter map manifests itself in the fact that all segments of a trajectory are tangent to the same confocal ellipse (as shown in Figure 6) or hyperbola. If  $k \geq 4 + \varepsilon$ , Moser's invariant curve theorem, as strengthened by Herman [19], applies. The resulting invariant curves of the billiard mapping correspond to quasiperiodic billiard orbits. The infinite family of segments of such an orbit has an envelope, called the *caustic*, Figure 6.

The caustics cluster at the boundary, and their density, in the sense of measure, approaches one near the boundary. Moreover, one has stability in the following sense: if one chord of an orbit does not meet a caustic, then the entire orbit is forever trapped in the ring outside the caustic.

The first application of Moser's twist theorem to billiards is due to Lazutkin [22]. Zharnitsky gave a simplified proof; we refer to [53] for further details and references.

This result is intimately related to the previous one of the ball bouncing on a periodically oscillating racket, as follows. Consider a billiard particle travelling close to the boundary. To better keep track of this particle, let us consider its shadow on the boundary: the foot of the perpendicular from the particle to the boundary. Let us put ourselves into the reference frame of that foot, and imagine the inward normal direction as the direction vertically up.



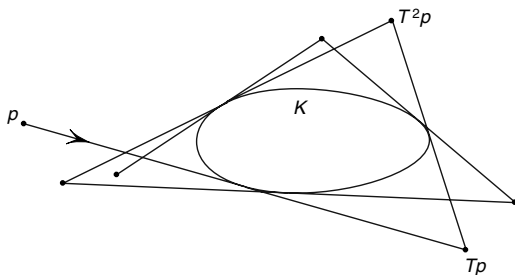


Fig. 7. The outer billiard.

With this choice, we are now in an accelerating frame, and the particle will no longer seem to be travelling with constant acceleration. Instead, it will seem to us as if the particle were under the influence of centrifugal gravity which forces the particle to bounce up and down repeatedly. In short, we are dealing with a particle bouncing on a plate under the influence of periodically varying gravity.

**EXAMPLE 9.** The outer billiard map, also referred to as the Neumann billiard map, is defined as follows. We are given a closed curve  $K$ , Figure 7. From a point  $p$  outside  $K$  we draw the tangent line  $l$  to  $K$  in the counterclockwise direction, and let  $Tp \neq p$  be the point on  $l$  equidistant with  $p$  from the point of tangency. Are there any points  $p$  whose iterates  $T^k p$  escape to infinity? If the boundary  $K$  is sufficiently smooth and its curvature is never zero, then the answer is negative, [17,21]. The proof relies on the non-obvious observation that near infinity  $T$  is a near-integrable map, provided  $K$  is a ‘nice’ curve.

**A brief background on adiabatic invariants** Before describing in the next paragraph a result due to Arnold, we give a brief sketch on adiabatic invariants. To focus on a simple case, consider a slowly changing Hamiltonian system

$$\begin{cases} \dot{x} = H_y(x, y, \varepsilon t) \\ \dot{y} = -H_x(x, y, \varepsilon t). \end{cases} \quad (31)$$

Since the derivative along solutions is small:  $\frac{d}{dt} H = \varepsilon \partial_\tau H(x, y, \tau)_{\tau=\varepsilon t} = O(\varepsilon)$ , we expect  $H$  to change by  $O(1)$  during time  $1/\varepsilon$ , along a typical solution. It turns out, however, that despite the ‘large’ change of  $H$ , the area enclosed by the energy curve  $H(x, y, \tau) = \text{const.}$  associated with the initial and the final points of the solution  $(x, y)(t)$  changes little (by  $O(\varepsilon)$ ). Rather than giving a rigorous statement and proof [3], we mention that two geometrical properties are responsible for this phenomenon: (1) the incompressibility of the Hamiltonian flow, and (2) the ergodicity of the motion of solutions of a frozen system along their energy curves. A beautiful and much earlier explanation of adiabatic invariance is due to Einstein, at least for the special case of a pendulum. This explanation is based on the observation that as we (say) shorten the length of the pendulum, we do not only raise the pendulum, but also work against the extra centrifugal tension of the string. Using the conservation of energy then leads to the conclusion of adiabatic invariance of the area, i.e. of the action. More details on this can be found in [43].

EXAMPLE 10 (Adiabatic Invariants for all Time). Another application of KAM theory is the following. Consider a slowly changing time-periodic Hamiltonian system (31), where  $H(x, y, \tau) = H(x, y, \tau + 1)$ . Assume, for simplicity, that  $H(0, 0, \tau) = 0$  for all  $\tau$  and that the level curves  $H(x, y, \tau) = c$  are closed for all  $\tau$  and for all  $c > 0$ . Assume also that the frozen system has twist, i.e. that  $\partial T(H, \tau)/\partial H \neq 0$  for all  $H > 0$  and all  $\tau$ , where  $T(H, \tau)$  is the period of the solution with energy  $H$ . Finally, assume that  $H$  is sufficiently smooth – say, real analytic. Under these assumptions Arnold's theorem states that, for  $\varepsilon$  sufficiently small, the system possesses an adiabatic invariant valid for all time (and not just for time of order  $1/\varepsilon$ ). For more details we refer to [1,5] and references therein.

## 5. Arnold diffusion

In conclusion, we mention briefly the phenomenon of 'Arnold diffusion': the 'drift' of action for near-integrable systems. This drift can happen only to the orbits that do not lie on KAM tori, and for small perturbations these drifting orbits form a small portion of the phase space. In fact, full proof of existence of such orbits in 'typical' systems is still not available, although great progress has been made and Mather announced a solution. There is by now a large literature on the subject; we mention [2,7–10,12,13,15,16,20,29,31,33–38,44,45,49,50].

## References

- [1] V. Arnold, *On the behavior of an adiabatic invariant under a slow periodic change of the Hamiltonian*, Soviet Math. Dokl. **3** (1962), 136–139.
- [2] V. Arnold, *Instabilities in dynamical systems with several degrees of freedom*, Sov. Math. Dokl. **5** (1964), 581–585.
- [3] V. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd edn, Springer-Verlag, New York, Berlin, Heidelberg (1988).
- [4] V. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd edn, Graduate Texts in Mathematics, Vol. 60 (1989).
- [5] V.I. Arnold, V.V. Kozlov and A.I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*, Dynamical Systems III, 3rd edn, Encyclopaedia of Mathematical Sciences, Vol. 3, Springer-Verlag, Berlin (2006); Translated from the Russian original by E. Khukhro.
- [6] S. Aubry and P.Y. LeDaeron, *The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground states*, Physica **8D** (1983), 381–422.
- [7] P. Bernard and G. Contreras, *A generic property of families of Lagrangian systems*, Ann. of Math. (2) **167** (3) (2008), 1099–1108.
- [8] M. Berti and Ph. Bolle, *A functional analysis approach to Arnold diffusion*, Ann. Inst. H. Poincaré **19** (4) (2002), 395–450.
- [9] U. Bessi, *An approach to Arnold's diffusion through the calculus of variations*, Nonlinear Anal. **26** (6) (1996), 1115–1135.
- [10] U. Bessi, L. Chierchia and E. Valdinoci, *Upper bounds on Arnold diffusion times via Mather theory*, J. Math. Pures Appl. **80** (1) (2001), 105–129.
- [11] L. Bin, *Boundedness for solutions of nonlinear Hill's equations with Periodic Forcing Terms via Moser's twist Theorem*, J. Differential Equations **79** (1989), 304–315.
- [12] J. Bourgain and V. Kaloshin, *On diffusion in high-dimensional Hamiltonian systems*, J. Funct. Anal. **229** (1) (2005), 1–61.
- [13] C.-Q. Cheng and J. Yan, *Existence of diffusion orbits in a priori unstable Hamiltonian systems*, J. Differential Geom. **67** (3) (2004), 457–517.

- [14] R. de la Llave, *A tutorial on KAM theory*, Smooth Ergodic Theory and its Applications (Seattle, WA, 1999), A. Katok, R. de la Llave, Ya. Pesin and H. Weiss, eds, Proc. Sympos. Pure Math., Vol. 69, Amer. Math. Soc., Providence, RI (2001), 175–292.
- [15] A. Delshams, R. de la Llave and T. Seara, *A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model*, Mem. Amer. Math. Soc. **179** (844) (2006), viii+141 pp.
- [16] R. Douady, *Stabilité ou instabilité des points fixes elliptiques*, Ann. Sci. de l'É.N.S. 4<sup>e</sup> série **21** (1) (1988), 1–46.
- [17] R. Douady, Ph. D. Thesis, Ecole Polytechnique (1988).
- [18] J. Féjoz, *Démonstration du théorème d'Arnol'd" sur la stabilité du système planétaire (d'après Michael Herman)*, Michael Herman Memorial Issue, Ergodic Theory Dynam. Systems **24** (5) (2004), 1521–1582. Updated version at <http://www.institut.math.jussieu.fr/~fejoz/>.
- [19] M.R. Herman, *Sur les courbes invariantes par les difféomorphismes de l'anneau  $\mathbb{I}$* , Asterisque **144** (1986).
- [20] V. Kaloshin and M. Levi, *An example of Arnold diffusion for near-integrable Hamiltonians*, Bull. Amer. Math. Soc. (N.S.) **45** (3) (2008), 409–427.
- [21] S. Laederich and M. Levi, *Invariant curves and time-periodic potentials*, Ergodic Theory Dynam. Systems **11** (1991), 365–378.
- [22] V.F. Lazutkin, *Existence of caustics for the billiard problem in a convex domain*, Math. USSR-Izvestiya **7** (1973), 185–214.
- [23] M. Levi, *Quasiperiodic motions in superquadratic time-periodic potentials*, Commun. Math. Phys. **143** (1991), 43–83.
- [24] M. Levi, *On Littlewood's counterexample of unbounded motions in superquadratic potentials*, Dynamics Reported, 1 (New Series) (1992), 113–124.
- [25] M. Levi and J. Moser, *A Lagrangian proof of the invariant curve theorem for twist mappings*, Proc. Symp. Pure Math. **69** (2001), 733–746.
- [26] M. Levi and E. Zehnder, *Boundedness of solutions for quasiperiodic potentials*, SIAM J. Math. Anal. **26** (5) (1995), 1233–1256.
- [27] J.E. Littlewood, *Unbounded solutions of an equation  $\ddot{y} + g(y) = p(t)$  with  $p(t)$  periodic and  $g(y)/y \rightarrow \infty$  as  $y \rightarrow \pm\infty$* , J. London Math. Soc. **41** (1966), 497–507.
- [28] P. Lochak, *Canonical perturbation theory: an approach based on joint approximations*, Uspekhi Mat. Nauk **47** (6(288)) (1992), 59–140; Translation in Russian Math. Surveys **47** (6) (1992), 57–133.
- [29] P. Lochak and J.-P. Marco, *Diffusion times and stability exponents for nearly integrable analytic systems*, Cent. Eur. J. Math. **3** (3) (2005), 342–397.
- [30] Y. Long, *An unbounded solution of a superlinear Duffing's equation*, Acta Math. Sinica (N.S.) **7** (4) (1991), 360–369.
- [31] J.-P. Marco and D. Sauzin, *Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems*, Publ. Math. Inst. Hautes Etudes Sci. (96) (2002), 199–275.
- [32] J. Mather, *Existence of quasi-periodic orbits for twist homeomorphisms of the annulus*, Topology **21** (1985), 457–476.
- [33] J. Mather, *Modulus of continuity of Peierls' barrier*, Periodic Solutions Hamiltonian Systems and Related Topics, P. Rabinowitz, ed., NATO ASI Series C, Vol. 209 (1987), 177–202.
- [34] J. Mather, *Action minimizing invariant measures for positive definite Lagrangians*, Math. Z. **207** (1991), 169–207.
- [35] J. Mather, *Variational construction of connecting orbits*, Ann. Inst. Fourier **43** (1993), 1349–1386.
- [36] J. Mather, *Arnold diffusion. I: Announcement of results*, J. Math. Sci. **124** (5) (2004), 5275–5289.
- [37] J. Mather, *Arnold diffusion*, II, preprint, 2006, 160 pp.
- [38] J. Mather, *Total disconnectedness of the quotient Aubry set in low dimensions*, Comm. Pure Appl. Math. **56** (8), 1178–1183.
- [39] G.R. Morris, *A case of boundedness in Littlewood's problem on oscillatory differential equations*, Bull. Austr. Math. Soc. **14** (1976), 71–93.
- [40] J. Moser, *On invariant curves of area-preserving mappings of an annulus*, Nachrichten der Akademie der Wissenschaften, Göttingen, Math.-Phys. Klasse IIa, 1962, pp. 1–20.
- [41] J. Moser, *A stability theorem for minimal foliations of the torus*, Ergodic Theory Dynam. Systems **8** (1988), 151–188.

- [42] J. Moser and E. Zehnder, *Notes on Dynamical Systems*, Courant Lecture Notes in Mathematics, Vol. 12 (2005).
- [43] K. Nakamura, *Quantum Chaos*, Cambridge University Press (1993).
- [44] N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Uspekhi Math. Nauk **32** (1) (1977), 5–66.
- [45] L. Niederman, *Prevalence of exponential stability among nearly integrable Hamiltonian systems*, Ergodic Theory Dynam. Systems **27** (2007), 905–928.
- [46] J. Pöschel, *Integrability of Hamiltonian Systems on Cantor Sets*, Comm. Pure Appl. Math. **35** (1982), 653–696.
- [47] L.D. Pustyl'nikov, *Stable and oscillating motions in non-autonomous dynamical systems*, Trans. Moscow Math. Soc. **14** (1978), 1–101.
- [48] D. Salamon and E. Zehnder, *KAM theory in configuration space*, Comm. Math. Helv. **64** (1989), 84–132.
- [49] D. Treschev, *Evolution of slow variables in a priori unstable Hamiltonian systems*, Nonlinearity **17** (5) (2004), 1803–1841.
- [50] J. Xia, *Arnold diffusion: a variational construction*, Proc of the ICM, Vol. II (Berlin, 1998) (1998), 867–877.
- [51] V. Zharnitsky, *Breakdown of stability of motion in superquadratic potentials*, Comm. Math. Phys. **189** (1) (1997), 165–204.
- [52] V. Zharnitsky, *Instability in Fermi-Ulam “ping-pong” problem*, Nonlinearity **11** (6) (1998), 1481–1487.
- [53] V. Zharnitsky, *Invariant tori in Hamiltonian systems with impacts*, Comm. Math. Phys. **211** (2) (2000), 289–302.
- [54] V. Zharnitsky, *Invariant curve theorem for quasiperiodic twist mappings and stability of motion in the Fermi-Ulam problem*, Nonlinearity **13** (4) (2000), 1123–1136.



## CHAPTER 6

# KAM Theory: Quasi-periodicity in Dynamical Systems

H.W. Broer

*Johann Bernoulli Institute for Mathematics and Computer Science,  
P.O. Box 407 9700 AK Groningen, The Netherlands  
Bernoulliborg, Building 5161, Nijenborgh 9, 9747 AG Groningen, The Netherlands  
E-mail: [h.w.broer@rug.nl](mailto:h.w.broer@rug.nl)*

Mikhail B. Sevryuk

*Institute of Energy Problems of Chemical Physics, The Russia Academy of Sciences, Leninskii prospect 38,  
Bldg. 2, Moscow 119334, Russia  
E-mail: [sevryuk@mccme.ru](mailto:sevryuk@mccme.ru)*

### Contents

|                                                                      |     |
|----------------------------------------------------------------------|-----|
| 1. Introduction . . . . .                                            | 251 |
| 1.1. The ‘classical’ KAM theorem . . . . .                           | 251 |
| 1.2. Related developments: outline . . . . .                         | 252 |
| 1.3. Discussion . . . . .                                            | 253 |
| 2. Complex linearization . . . . .                                   | 254 |
| 2.1. Formal solution . . . . .                                       | 255 |
| 2.2. Convergence and small divisors . . . . .                        | 255 |
| 2.3. Measure and category . . . . .                                  | 255 |
| 2.4. Discussion . . . . .                                            | 256 |
| 3. KAM Theory for circle and annulus maps . . . . .                  | 257 |
| 3.1. Circle maps . . . . .                                           | 257 |
| 3.2. Area preserving annulus maps . . . . .                          | 263 |
| 4. KAM Theory for flows . . . . .                                    | 266 |
| 4.1. Introduction . . . . .                                          | 266 |
| 4.2. Families of normally hyperbolic quasi-periodic tori . . . . .   | 270 |
| 4.3. KAM Theory for Lagrangean tori in Hamiltonian systems . . . . . | 272 |
| 4.4. Applications of the Lagrangean KAM Theorem 6 . . . . .          | 277 |
| 5. Further developments in KAM Theory . . . . .                      | 280 |
| 5.1. Background . . . . .                                            | 280 |
| 5.2. Parametrized KAM Theory . . . . .                               | 283 |

|                                                                  |     |
|------------------------------------------------------------------|-----|
| 6. Quasi-periodic bifurcations: dissipative setting . . . . .    | 288 |
| 6.1. Quasi-periodic Hopf bifurcation . . . . .                   | 290 |
| 6.2. Discussion . . . . .                                        | 294 |
| 7. Quasi-periodic bifurcation theory in other settings . . . . . | 296 |
| 7.1. Hamiltonian cases . . . . .                                 | 297 |
| 7.2. Discussion . . . . .                                        | 299 |
| 8. Further Hamiltonian KAM Theory . . . . .                      | 299 |
| 8.1. Exponential condensation . . . . .                          | 299 |
| 8.2. Destruction of resonant tori . . . . .                      | 301 |
| 8.3. Lower dimensional isotropic invariant tori . . . . .        | 303 |
| 8.4. Excitation of elliptic normal modes . . . . .               | 314 |
| 8.5. Higher dimensional coisotropic invariant tori . . . . .     | 315 |
| 8.6. Atropic invariant tori . . . . .                            | 317 |
| 9. Whitney smooth bundles of KAM tori . . . . .                  | 318 |
| 9.1. Motivation . . . . .                                        | 318 |
| 9.2. Formulation of the global KAM theorem . . . . .             | 319 |
| 9.3. Applications . . . . .                                      | 321 |
| 9.4. Discussion . . . . .                                        | 323 |
| 10. Conclusion . . . . .                                         | 324 |
| Acknowledgments . . . . .                                        | 325 |
| References . . . . .                                             | 325 |

## 1. Introduction

Kolmogorov–Arnold–Moser (or KAM) Theory was developed for conservative (Hamiltonian) dynamical systems that are nearly integrable. Integrable systems in their phase space contain lots of invariant tori and KAM Theory establishes persistence of such tori, which carry quasi-periodic motions. We present this theory which begins with Siegel’s and Kolmogorov’s pioneering work in the 1940s and 50s.

Since Moser’s results from the 1960s it is known that KAM Theory extends outside the world of Hamiltonian systems. Indeed, as will be explained below, families of quasi-periodic attractors can be dealt with in the same way as quasi-periodic Lagrangean invariant tori in Hamiltonian systems. In both cases a Kolmogorov-like nondegeneracy condition is needed on the way frequencies vary with the unperturbed tori. The background is Moser’s Lie algebra version of KAM Theory. There are other types of nondegeneracy conditions as well, for instance, the so-called Rüssmann condition. All our formulations include Whitney differentiable conjugations with collections of Diophantine quasi-periodic tori in integrable approximations. This part of the theory was initiated by Lazutkin and Pöschel in the 1970s and 80s. From this, for a large class of KAM tori, uniqueness follows.

A general type of nondegeneracy, involving unfolding parameters and transversality, was developed in the late 1980s by Broer, Huitema, and Takens. It can be shown to encompass (in a sense to be made precise) both the Kolmogorov and Rüssmann nondegeneracy. Also it is at the basis of the quasi-periodic bifurcation theory. It turns out that the standard (semi-algebraic) bifurcation diagrams, as known for equilibria and periodic solutions, in the quasi-periodic setting, occur in a ‘Cantorized’ and, sometimes, ‘frayed’ way. These developments took place during the 1990s and round the turn of the century.

Recently a global KAM Theory was constructed, which leads to Whitney smooth bundles of invariant tori.

### 1.1. The ‘classical’ KAM theorem

At the International Congress of Mathematicians in 1954, held in Amsterdam, A.N. Kolmogorov gave a closing lecture with the title ‘The general theory of dynamical systems and classical mechanics’ [250]. Among many other things he discussed his paper [251]. The event took place in the Amsterdam Concertgebouw and it has played a major role in the developments of the Dynamical System Theory and of Mathematical Physics, in particular of what is now called Kolmogorov–Arnold–Moser (or KAM) Theory. We like to note that the term ‘KAM Theory’ was first used in [239,469].

In this lecture Kolmogorov considered the occurrence of multi- or quasi-periodic motions, which in the phase space, are confined to invariant tori. He restricted himself to conservative, or Hamiltonian, dynamical systems, as these are generally used for modelling in classical mechanics. Invariant Lagrangean tori that carry quasi-periodic motions were well-known to occur in Liouville integrable systems, and Kolmogorov’s paper [251] and lecture [250] dealt with the persistence of these tori under small, non-integrable perturbations of the Hamiltonian. Due to so-called small divisors, the corresponding perturbation series diverge on a dense set.



In broad terms, KAM Theory states that generically, in small perturbations of integrable systems the union of quasi-periodic Lagrangean invariant tori has positive Liouville measure both in the phase space and in the energy hypersurfaces. Here two related nondegeneracy conditions come in play, dealing with the way frequencies or frequency ratios vary with the tori in the phase space.

‘This theorem is often said to be the first and perhaps foremost result of modern non-linear dynamics of conservative systems’ [361, p. 487] (compare with [201]).

## 1.2. *Related developments: outline*

Already in the 19th century (in fact, even earlier) the problem of small divisors was met, notably in perturbation series related to a three-body problem [9,25,28,155,195,323,374,414]. This fact initiated many developments, partly ending up in KAM Theory as reported on here. H. Poincaré played a central role in these early developments.

Poincaré was also one of the founders of the linearization program to which the example of the next section belongs. This concerns the linearization near a fixed point of a holomorphic diffeomorphism of the complex plane and leads to the first solution of a small divisor problem by C.L. Siegel [413], compare with [19]. Here we first meet the theme of ‘measure versus category’ [338] which is so central in KAM Theory.

Next we turn to the dynamics of circle maps, which goes back to Arnold [11,19], followed by a discussion on area preserving twist maps [218,316,322,373].

After this the flow case is considered. First we show that in general quasi-periodic invariant tori have a natural affine structure, which in the Liouville integrable Hamiltonian setting coincides with that given by the Liouville–Arnold Theorem [13,20,22,25,138,305]. However, it was already known to J.K. Moser [317–319,323] that KAM Theory admits a much greater generality than the world of Hamiltonian systems. This will be illustrated by treating families of quasi-periodic attractors [333,369], exactly like the ‘classical’ case of Lagrangean invariant tori in the Hamiltonian systems we next describe.

All KAM theorems below are given in the ‘structural stability’ form, where a conjugation is produced between the Diophantine quasi-periodic tori in the integrable and the nearly integrable cases. In the present setting we speak of ‘quasi-periodic stability’. In the spirit of Pöschel [354], these conjugations are Whitney differentiable; also compare with [122,267,268,423,470,471]. This accounts for the fact that it is a typical property to have a union of quasi-periodic invariant tori of positive Hausdorff measure of the appropriate dimension [76,77,132]. A property is called ‘typical’ if it occurs on an open set of the dynamical systems at hand. Regarding the topology on the ‘function space’, one may think of the (weak) Whitney topology for differentiable systems [225,325], or of the compact-open topology for holomorphic extensions of real analytic systems, compare with [93]. Moreover, Whitney differentiability enables us to show uniqueness for a large class of the perturbed KAM tori in several situations [91].

We next describe the Parametrized KAM Theory, inspired by Moser [318,319] (see also [30,387]), where a general nondegeneracy concept has been developed in the late 1980s by Broer, Huitema, and Takens. This BHT nondegeneracy involves a certain (uni)versality of parametrized systems [76–78,236]. By considering the geometry and the number theory of the (Diophantine) quasi-periodic frequency vectors, this notion can

be shown to encompass both the Kolmogorov and Rüssmann nondegeneracies [250,251,376,378,381] (in the latter case, the number of parameters can be drastically reduced). This theory is developed in a structure preserving way, using KAM Theory as formulated for certain Lie algebras of vector fields [71,78,227,236,319]. In many Hamiltonian and reversible settings, all parameters can be ‘compensated’ by phase space variables.

The BHT nondegeneracy also is at the basis of the quasi-periodic bifurcation theory [44,78]. It turns out that the standard (semi-algebraic) bifurcation diagrams in the product of the phase space and the parameter space, as known for equilibria and periodic solutions since Whitney, Thom, Mather, and Arnold [19,21,23,24,428], occur in the quasi-periodic setting in a ‘Cantorized’ way: near the dense set of resonances KAM Theory does not work. To show this we deal with the quasi-periodic Hopf bifurcation in some detail, also indicating certain Hamiltonian and reversible analogues [57,58,63,65,66,68,69,71,208–212,227,337]. Here the ‘conventional’ Hopf bifurcation [228] (or the Poincaré–Andronov phenomenon [19]) for equilibrium points plays a central role.

We next dwell upon other branches of KAM Theory, confining ourselves to the Hamiltonian case. The central theme here is quasi-periodic invariant tori whose dimension is not equal to the number  $n$  of degrees of freedom. Lower dimensional (of dimensions  $< n$ ) isotropic tori studied first by Melnikov [309,310] and Moser [318,319] have been explored in great detail by now, which enables us to formulate the corresponding KAM theorems on the torus persistence in the ‘structural stability’ form. Another type of theorem on lower dimensional tori concerns families of tori of dimensions  $\ell + 1, \ell + 2, \dots, n$  around  $\ell$ -tori with partially ‘elliptic’ normal behaviour, and we consider this topic (the so-called ‘excitation of elliptic normal modes’) as well. Higher dimensional (of dimensions  $> n$ ) coisotropic tori introduced into KAM Theory by Parasyuk [341,346–348] have been understood worse than lower dimensional tori, and we present here just a review without precise statements. In 2000 Huang, Cong, and Li [231,232] started examining quasi-periodic invariant tori which are ‘atropic’, i.e. neither isotropic nor coisotropic. The dimension of such atropic tori can be smaller than, equal to, or greater than the number of degrees of freedom. Apart from lower dimensional isotropic, higher dimensional coisotropic, and atropic tori, we also return to the ‘classical’ case of Lagrangean invariant KAM tori, and describe their ‘exponential condensation’ and ‘superexponential stickiness’ in the analytic category [314], as well as discussing the destruction of resonant unperturbed Lagrangean tori into finite collections of lower dimensional tori (the phenomenon first investigated in detail by Treshchëv [429]).

We end with the description of the global KAM Theory [60] in the ‘classical’ Hamiltonian setting, as this leads to Whitney smooth bundles of invariant tori that inherit the corresponding geometry of the integrable bundle, involving monodromy, Chern classes, etc. Here the conjugations of the Lagrangean KAM theorem are glued together with a Partition of Unity [225,325,418].

### 1.3. Discussion

The Kolmogorov and BHT nondegeneracies are formulated in local terms that, by the Inverse Function Theorem, give rise to open domains in the product of the phase space and parameter space on which certain frequency maps are submersions. Most of the results

have been formulated in terms of these open domains, which in applications may turn out to be quite large.

Below, KAM Theory is developed mostly for Hamiltonian systems (which is justified partly by historical reasons and partly by reasons of applications), whereas parallel results exist for other classes of dynamical systems as well. In particular, the reversible KAM Theory (starting with Moser's paper [317]) is to a great extent parallel to the Hamiltonian one, see e.g. [18,26,31,57,58,75–77,319,323,354,389,392,398,404,409–411] (in the case of reversible diffeomorphisms, however, some special effects are exhibited [359]), and the weakly reversible KAM Theory has been developed in [26,389].<sup>1</sup> The present set-up is even more general, also including the class of 'dissipative' systems, where no structure has to be preserved. As already mentioned, in these cases often parameters are needed for the persistence of quasi-periodic tori. For earlier, partial overviews of KAM Theory in the same general spirit, see [55,56,76–78,234,319,398]. We note that, although the present theory is mostly being developed for flows, a completely analogous approach exists for diffeomorphisms. Also we like to mention that our bibliography is partly complementary to those of [25,77,147,168,212,357,408].

For a first acquaintance with KAM Theory, the introductory texts [55,56,92,146,234,450], manuals [20,22,25,323,414,431], and reviews [40,51,398] also are useful. The detailed survey of the Hamiltonian KAM Theory presented in [25] is especially recommended.

Finally it should be mentioned that we will not touch the so-called converse KAM Theory in this survey. The converse theorems assert that, under appropriate hypotheses, dynamical systems admit no invariant quasi-periodic tori, or the measure of the union of those tori is small. The papers [213,248,296,298,299,306,402,424] exemplify this theory, see also [77,218,221,397].

## 2. Complex linearization

We deal with the linearization problem for a holomorphic map near a fixed point, for a description see V.I. Arnold's manual [19] or J.W. Milnor's monograph [311].

To be precise, consider a local holomorphic map (or a germ)  $F : (\mathbb{C}, 0) \rightarrow (\mathbb{C}, 0)$  of the form  $F(z) = \lambda z + f(z)$  with  $f(0) = f'(0) = 0$ . The question is whether or not there exists a local biholomorphic transformation  $\Phi : (\mathbb{C}, 0) \rightarrow (\mathbb{C}, 0)$  such that

$$\Phi \circ F = \lambda \cdot \Phi. \quad (1)$$

We say that  $\Phi$  linearizes  $F$  near its fixed point 0.

---

<sup>1</sup>A vector field  $X$  is said to be *reversible* with respect to a phase space diffeomorphism  $G$  if  $G$  conjugates  $X$  to  $-X$ , that is  $G_*X = -X$ . This means that  $G(a(-t))$  is a solution of the corresponding system  $\dot{a} = X(a)$  of ordinary differential equations whenever  $a(t)$  is. The classical example is the Newtonian equations of motion  $\ddot{u} = F(u)$ ,  $u \in \mathbb{R}^N$ , which can be written in the form  $\dot{u} = v$ ,  $\dot{v} = F(u)$ ; here  $G : (u, v) \mapsto (u, -v)$ . Similarly, a diffeomorphism  $A$  of a certain manifold is said to be *reversible* with respect to another diffeomorphism  $G$  of the same manifold if  $GAG^{-1} = A^{-1}$ . Following Arnold's note [18], one often speaks of *weakly reversible* systems in the case where  $G$  is *not* an involution (i.e.  $G^2$  is not the identity transformation). For general references on reversible dynamical systems, see [262,364].

### 2.1. Formal solution

First consider the problem at a formal level. Given a series expansion  $f(z) = \sum_{j \geq 2} f_j z^j$  we look for another series  $\Phi(z) = z + \sum_{j \geq 2} \phi_j z^j$ , such that the conjugation relation (1) holds formally. It turns out that a formal solution exists whenever  $\lambda \neq 0$  is not a root of 1 (clearly, it suffices to consider the case  $0 < |\lambda| \leq 1$ , otherwise one can examine  $F^{-1}$  in place of  $F$ ). Indeed, the coefficients  $\phi_j$  can be determined recursively by the following equations:

$$\begin{aligned} \lambda(1 - \lambda)\phi_2 &= f_2, \\ \lambda(1 - \lambda^2)\phi_3 &= f_3 + 2\lambda f_2\phi_2, \\ \lambda(1 - \lambda^{n-1})\phi_n &= f_n + \text{already known terms}, \quad n \geq 3. \end{aligned} \tag{2}$$

From this the claim directly follows.

### 2.2. Convergence and small divisors

In the hyperbolic case where  $0 < |\lambda| \neq 1$ , the series for  $\Phi$  has positive radius of convergence. This was proven by Poincaré, not by considering the series but by a direct iteration method, compare with [19]. So there remains the elliptic case with  $\lambda \in \mathbb{T}^1 \subset \mathbb{C}$ , the unit circle on the complex plane. What is important for an analysis of Equations (2) in this case is that, even if  $\lambda$  is not a root of unity, its powers do accumulate on 1. This would give *small divisors* in the formal series of  $\Phi$ , which casts doubt on its convergence.

This problem was successfully solved by C.L. Siegel [413] in 1942. To this purpose, writing  $\lambda = e^{2\pi i \beta}$ , the following *Diophantine conditions* were introduced: for some  $\tau > 1$  and  $\gamma > 0$  it is required that

$$\left| \beta - \frac{p}{q} \right| \geq \frac{\gamma}{q^{\tau+1}}, \tag{3}$$

for all rationals  $p/q$  (with  $q > 0$ ). It turns out that this is sufficient for convergence of the formal solution for  $\Phi$ .

For the moment, let it be enough to say that the set of all  $\lambda \in \mathbb{T}^1$ , such that the corresponding  $\beta$  are Diophantine for some  $\tau > 1$  and  $\gamma > 0$ , has full measure in  $\mathbb{T}^1$ . In the next section we shall give a more elaborate discussion on Diophantine sets.

### 2.3. Measure and category

From J.C. Oxtoby's manual [338] it is known that the real number line  $\mathbb{R}$  contains subsets that are large in measure and yet topologically small. The following example, due to H. Cremer [137] in 1927, illustrates this in the present situation. For a nice description of this example in a somewhat different context see [35].

**EXAMPLE 1** (Linearization of a Quadratic Map [137]). Consider the map  $F : \mathbb{C} \rightarrow \mathbb{C}$  given by

$$F(z) = \lambda z + z^2,$$

where  $\lambda \in \mathbb{T}^1$  is not a root of unity. We shall see that there is a topologically large subset of  $\lambda$  for which this map has periodic points in any neighbourhood of 0. For such  $\lambda$  it follows that the formal conjugation diverges. Indeed, since the linear map  $z \mapsto \lambda z$  is the rotation over an angle incommensurable with  $2\pi$ , and all the orbits (everywhere dense in  $\mathbb{T}^1$ ) of such a rotation are infinite for  $|z| > 0$ , the existence of periodic orbits in every neighbourhood of 0 implies that the formal linearization must have zero radius of convergence.

To examine periodic points of period  $q$ , we consider the equation

$$F^q(z) = z. \quad (4)$$

Using

$$F^q(z) = \lambda^q z + \dots + z^{2^q},$$

it follows that

$$F^q(z) - z = z \left( \lambda^q - 1 + \dots + z^{2^q-1} \right).$$

Abbreviating  $N(q) = 2^q - 1$ , let  $z_1, z_2, \dots, z_N$  be the non-zero solutions of (4). Their product satisfies the relation

$$z_1 z_2 \dots z_N = 1 - \lambda^q,$$

since  $N(q)$  is odd. From this we see that there exists at least one solution within radius

$$|\lambda^q - 1|^{1/N(q)}$$

of  $z = 0$ . Now consider the set of  $\lambda \in \mathbb{T}^1$  satisfying

$$\liminf_{q \rightarrow \infty} |\lambda^q - 1|^{1/N(q)} = 0.$$

This set turns out to be residual,<sup>2</sup> in fact, it resembles the set of Liouville numbers (we recall that a residual set contains a countable intersection of dense-open sets, which expresses the fact that the set is large in the sense of topology). Notice that this set is necessarily of measure zero, since it has all Diophantine numbers in its complement.

We conclude that for all  $\lambda$  contained in this residual set, periodic points of  $F$  occur in any neighbourhood of  $z = 0$ , which implies that for such  $\lambda$  the formal normal form transformation has zero radius of convergence.

## 2.4. Discussion

By the end of the 20th century J.-C. Yoccoz [465,466] completely solved the elliptic case, based on the so-called *Bruno condition* [97,98,372] on associated continued fractions. To be precise, let  $p_j/q_j \in \mathbb{Q}$ ,  $j \in \mathbb{N}$ , be the sequence of convergents<sup>3</sup> to  $\beta \notin \mathbb{Q}$ . The Bruno

<sup>2</sup>In other words: a dense  $G_\delta$  or a set of second Baire category [338].

<sup>3</sup>The *convergents* are the initial segments of a continued fraction. If  $\beta = [a_0; a_1, a_2, \dots]$  is the continued fraction representation of  $\beta$  (in our case  $a_0 = 0$  since  $0 < \beta < 1$ ) then the  $j$ th convergent to  $\beta$  is  $p_j/q_j = [a_0; a_1, a_2, \dots, a_j]$ . The theory of continued fractions is expounded in e.g. [247,366].

condition then requires

$$\sum_{j \geq 1} \frac{\log q_{j+1}}{q_j} < +\infty. \quad (5)$$

For further information see e.g. Milnor's monograph [311], also see Devaney's article [2].

The above discussion reveals that in the setting of holomorphic maps, there is nontrivial occurrence of dynamics that, up to a biholomorphic change of variables, is a rigid rotation over a Diophantine irrational angle<sup>4</sup> on an open region of the complex plane (a neighbourhood of 0). In the above case this occurs on so-called Siegel discs. Similar results hold for so-called Arnold–Herman rings [215], compare with [169,295]. For an overview of holomorphic dynamics, see [311]. For other matters on holomorphic dynamics, compare with [2].

### 3. KAM Theory for circle and annulus maps

In the previous section we witnessed a typical occurrence of rigid rotations on open regions of the complex plane. Slightly paraphrasing this, we might say that the unperturbed system is a rigid rotation over a Diophantine angle, which corresponds to a large measure set in the parameter space  $\mathbb{T}^1 = \{\lambda\}$ . Siegel's theorem [19,215,311,413] implies that this situation is persistent under all sufficiently small perturbations.

Keeping this in mind, we now turn to maps of the circle that are close to rigid rotations in a suitable topology on the corresponding function space. For simplicity all future results are formulated in the  $C^\infty$ -topology [225,325]. We mention that these topologies are compatible with the real analytic case under the compact-open topology on holomorphic extensions [93]. Moreover, the present theory admits generalizations in the  $C^k$ -topology for  $k \in \mathbb{N}$  sufficiently large. For details, see Section 4.2.1.

One important feature of the present problem is that parameters are needed for persistence of conjugations to rigid rotations. Therefore we shall speak of 1-parameter families of circle maps, that will be regarded as ‘vertical’ maps of the cylinder, i.e. as circle bundle maps.

#### 3.1. Circle maps

Consider a  $C^\infty$ -family of circle diffeomorphisms

$$P_{\beta,\varepsilon} : \mathbb{T}^1 \rightarrow \mathbb{T}^1; \quad x \mapsto x + 2\pi\beta + \varepsilon a(x, \beta, \varepsilon), \quad (6)$$

where  $\beta$  is a real parameter varying over an open finite interval  $\Gamma \subset \mathbb{R}$ . We regard  $\varepsilon$  as a perturbation parameter, on which (6) also depends smoothly. For  $\varepsilon = 0$ , we deal with the family  $P_{\beta,0}$  of rigid circle rotations.

---

<sup>4</sup>Strictly speaking, here and henceforth the words ‘rational’, ‘irrational’, ‘Diophantine’ refer not to the angle  $2\pi\beta$  itself but to the number  $\beta$ .

For convenience we rewrite the family (6) as a ‘vertical’ cylinder map

$$P_\varepsilon : \mathbb{T}^1 \times \Gamma \rightarrow \mathbb{T}^1 \times \Gamma; \quad (x, \beta) \mapsto (x + 2\pi\beta + \varepsilon a(x, \beta, \varepsilon), \beta).$$

Naively speaking, the problem is to find a diffeomorphism  $\Phi_\varepsilon$  that conjugates  $P_\varepsilon$  and  $P_0$ . To be precise, we require that the following diagram commutes:

$$\begin{array}{ccc} \mathbb{T}^1 \times \Gamma & \xrightarrow{P_\varepsilon} & \mathbb{T}^1 \times \Gamma \\ \uparrow \Phi_\varepsilon & & \uparrow \Phi_\varepsilon \\ \mathbb{T}^1 \times \Gamma & \xrightarrow{P_0} & \mathbb{T}^1 \times \Gamma, \end{array}$$

meaning that

$$P_\varepsilon \circ \Phi_\varepsilon = \Phi_\varepsilon \circ P_0, \quad (7)$$

compare with (1).

**3.1.1. Small divisors again.** We proceed by formally solving Equation (7). In order to respect the ‘verticality’ of the cylinder maps  $P_0$  and  $P_\varepsilon$ , we assume that  $\Phi_\varepsilon$  has the skew form

$$\Phi_\varepsilon(x, \beta) = (x + \varepsilon U(x, \beta, \varepsilon), \beta + \varepsilon \sigma(\beta, \varepsilon)), \quad (8)$$

i.e. that it preserves projections to the parameter space  $\Gamma$ . It follows that (7) can be rewritten as

$$\begin{aligned} U(x + 2\pi\beta, \beta, \varepsilon) - U(x, \beta, \varepsilon) \\ = 2\pi\sigma(\beta, \varepsilon) + a(x + \varepsilon U(x, \beta, \varepsilon), \beta + \varepsilon\sigma(\beta, \varepsilon), \varepsilon). \end{aligned}$$

Expanding in powers of  $\varepsilon$  and comparing lowest order coefficients, Equation (7) leads to the linear equation

$$U_0(x + 2\pi\beta, \beta) - U_0(x, \beta) = 2\pi\sigma_0(\beta) + a_0(x, \beta),$$

also called the *homological* equation. The latter equation can be directly solved in Fourier series. Indeed, writing

$$\begin{aligned} a_0(x, \beta) &= \sum_{k \in \mathbb{Z}} a_{0k}(\beta) e^{ikx} \quad \text{and} \\ U_0(x, \beta) &= \sum_{k \in \mathbb{Z}} U_{0k}(\beta) e^{ikx}, \end{aligned}$$

we find that  $\sigma_0 = -\frac{1}{2\pi} a_{00}$ , which yields a parameter shift, and that, for  $k \in \mathbb{Z} \setminus \{0\}$ ,

$$U_{0k}(\beta) = \frac{a_{0k}(\beta)}{e^{2\pi i k \beta} - 1},$$

while  $U_{00}(\beta)$  can be taken arbitrarily, which roughly corresponds to a circle translation.

As in Section 2.2, we observe that generally there exists a formal solution only for irrational  $\beta$ . As before, the powers of  $e^{2\pi i\beta}$  still accumulate on 1. This gives *small divisors* in the Fourier series, which threatens its convergence.

**3.1.2. A KAM theorem for circle maps.** To overcome the problem of small divisors, *Diophantine conditions* are introduced, as before, see (3), requiring that for given numbers  $\tau > 1$  and  $\gamma > 0$ , and for all rationals  $p/q$  with  $q > 0$  one has

$$\left| \beta - \frac{p}{q} \right| \geq \frac{\gamma}{q^{\tau+1}}. \quad (9)$$

Let us denote the set of all such  $\beta$  by  $\mathbb{R}_{\tau,\gamma} \subset \mathbb{R}$ , noting that  $\mathbb{R}_{\tau,\gamma}$  is closed. Recalling that  $\Gamma \subset \mathbb{R}$  is an open interval, we also consider the closed interval  $\Gamma_\gamma = \{\beta \in \Gamma \mid \text{dist}(\beta, \partial\Gamma) \geq \gamma\}$  and next  $\Gamma_{\tau,\gamma} = \Gamma_\gamma \cap \mathbb{R}_{\tau,\gamma}$ . It follows that  $\Gamma_{\tau,\gamma}$  is closed (and hence even compact). Thus, by the Cantor–Bendixson theorem [214] it follows that  $\Gamma_{\tau,\gamma}$  is the union of a perfect and a countable set. The perfect set is a Cantor set, since it is also compact and totally disconnected (or zero-dimensional). The latter holds since the dense set of rationals is in the complement of  $\Gamma_{\tau,\gamma}$ . We conclude that  $\Gamma_{\tau,\gamma}$  is nowhere dense and hence topologically ‘small’. However, the Lebesgue measure of  $\Gamma_{\tau,\gamma}$ , for  $\gamma \downarrow 0$ , is large. Indeed, a small computation shows that

$$\text{meas}(\Gamma \setminus \Gamma_{\tau,\gamma}) \leq \text{const } \gamma \sum_{q \geq 1} q^{-\tau} = O(\gamma), \quad (10)$$

as  $\gamma \downarrow 0$ , where we have used the fact that  $\tau > 1$ , compare with e.g. [19,76–78,236]. Note that the estimate (10) implies that the union

$$\bigcup_{\tau,\gamma} \Gamma_{\tau,\gamma}$$

is of full measure in  $\Gamma$ , compare with remarks in the previous section. Again we refer to [338] for a discussion of measure vs category. As a first KAM theorem we now formulate

**THEOREM 2 (KAM for Circle Maps).** *In the above circumstances assume that  $\tau > 1$  and that  $\gamma > 0$  is sufficiently small. Then, if the family  $P_\varepsilon$  is sufficiently close to  $P_0$  in the  $C^\infty$ -topology, there exists a  $C^\infty$ -diffeomorphism of the cylinder  $\Phi_\varepsilon : \mathbb{T}^1 \times \Gamma \rightarrow \mathbb{T}^1 \times \Gamma$ , of the skew form (8) with the following properties.*

1.  $\Phi_\varepsilon$  is a  $C^\infty$ -near the identity map and depends  $C^\infty$ -ly on  $\varepsilon$ .
2. The image of the  $P_0$ -invariant union of circles  $\mathbb{T}^1 \times \Gamma_{\tau,\gamma}$  under  $\Phi_\varepsilon$  is  $P_\varepsilon$ -invariant, and the restricted map  $\widehat{\Phi}_\varepsilon = \Phi_\varepsilon|_{\mathbb{T}^1 \times \Gamma_{\tau,\gamma}}$  conjugates  $P_0$  to  $P_\varepsilon$ , that is

$$P_\varepsilon \circ \widehat{\Phi}_\varepsilon = \widehat{\Phi}_\varepsilon \circ P_0.$$

That  $\Phi_\varepsilon$  is a  $C^\infty$ -near the identity map means that whenever  $P_\varepsilon \rightarrow P_0$  in the  $C^\infty$ -topology, also  $\Phi_\varepsilon \rightarrow \text{Id}$  in the  $C^\infty$ -topology (here and henceforth,  $\text{Id}$  denotes the identity mapping). Theorem 2 goes back to V.I. Arnold [11], also compare with [19]. The present



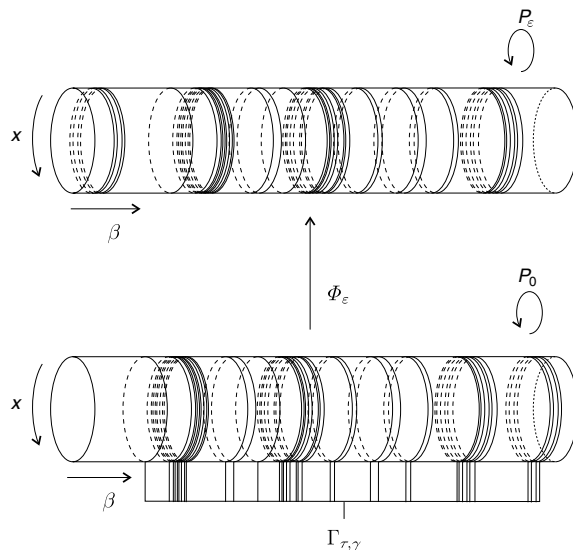


Fig. 1. Conjugation between the maps  $P_0$  and  $P_\varepsilon$  on  $\mathbb{T}^1 \times \Gamma_{\tau,\gamma}$ .

formulation closely fits with [76–78,236]. Concerning the smoothness of the map  $\Phi_\varepsilon$ , compare with J. Pöschel’s results [354] and with [452,470,471].

The proof of [Theorem 2](#) does not directly deal with the power series in  $\varepsilon$ . Instead it uses a Newton-like iterative method and an approximation property (by analytic maps) of Whitney smooth maps defined on closed sets. Finally, the diffeomorphism  $\Phi_\varepsilon$  is obtained by the Whitney Extension Theorem [452,453]. This approach does need a formulation of the present KAM theorem in terms of ‘vertical’ cylinder maps, compare with [Figure 1](#).

- REMARKS.
1. For sufficiently small  $\gamma > 0$ , the maximal size of the perturbation (the size of the difference between  $P_\varepsilon$  and  $P_0$ ) depends on  $\gamma$  in a linear way. From the measure-theoretical point of view, it is optimal to choose  $\gamma$  as small as the perturbation allows. In certain cases it is even possible to consider the limit as  $\gamma \downarrow 0$ , thereby creating so-called Lebesgue density points of quasi-periodicity [77,316,354].
  2. [Theorem 2](#), like many other KAM theorems, has a perturbative character, since it only applies to small perturbations of the rigid rotation family  $P_0$ . In contrast to this, M.R. Herman [215,220] and J.-C. Yoccoz [462,465] have proven a non-perturbative version of this theorem, in terms of the rotation number being Diophantine (also see [304]).
  3. No KAM theorem in this survey (except for [Theorem 19](#)) will be provided with a proof. The standard scheme of proving various KAM statements, as invented by Kolmogorov [251] and refined further by Arnold [14] and Moser [316] and subsequently by many other authors, is based on constructing an infinite sequence of coordinate transformations whose domains of definition shrink down to the invariant tori sought for. This cumbersome iterative procedure is similar to Newton’s method of tangents for solving algebraic equations. There are also approaches

using ‘hard’ Implicit Function Theorems in infinite-dimensional spaces (see e.g. [471]) or the Schauder fixed-point theory (see e.g. [218]). On the other hand, the authors of the papers [117,235,440] present detailed proofs of Kolmogorov’s original theorem [251] that follow rather closely the lines of [251]. Recently, the so-called ‘direct methods’ in proving the existence and persistence theorems for quasi-periodic motions were developed that deal directly with (Poincaré–Lindstedt) series in the perturbation parameter and exploit techniques usual in Quantum Field Theory, like the multiscale decomposition, tree expansions, and renormalization groups (see e.g. [39,46,119–121,166,174,181,182,185–187,189–191] and references therein). These techniques allow one to find explicitly delicate cancellations (‘compensations’) among large terms of the Lindstedt series (absolutely divergent due to the small divisors) and obtain estimates implying convergence.

**3.1.3. Discussion.** A circle diffeomorphism smoothly conjugated to a rigid rotation  $P_{\beta,0} : x \mapsto x + 2\pi\beta$  with  $\beta$  irrational is said to be *quasi-periodic*. It is well-known that each orbit of such a map fills the circle densely, see e.g. [19,20,154]. For  $\beta \in \Gamma_{\tau,\gamma}$ , the rigid rotation  $P_{\beta,0}$  certainly is quasi-periodic. A first consequence of Theorem 2 is that the circle maps  $P_{\beta,\varepsilon}$  that are conjugated to one of the Diophantine rigid rotations  $P_{\beta',0}$ , are still quasi-periodic. In fact, since  $\Phi_\varepsilon$  is near identity in the  $C^\infty$ -topology, it follows that, for  $\varepsilon \neq 0$  small, the measure of the union of circles with Diophantine rotation numbers is still large (for a definition of the rotation number see e.g. [19,154]). The conclusion is that quasi-periodicity typically occurs with positive measure in the parameter space. Moreover, the fact that a Cantor set is perfect, meaning that it contains no isolated points, implies that quasi-periodicity *almost never* occurs as an isolated phenomenon.

We formulated Theorem 2 in its (structural) stability form, which for this occasion is called *quasi-periodic stability* [76–78,236]. This term was chosen in reminiscence of the so-called  $\Omega$ -stability, which refers to structural stability when restricted to the non-wandering set [246].

**The Arnold family** An example is given by the Arnold family

$$P_{\beta,\varepsilon}(x) = x + 2\pi\beta + \varepsilon \sin x \quad (11)$$

of circle maps, where we consider the  $(\beta, \varepsilon)$ -plane of parameters. We restrict ourselves to  $|\varepsilon| < 1$ , to ensure that  $P_\varepsilon$  is a diffeomorphism. For  $|\varepsilon| \ll 1$  the family  $P_\varepsilon$  is close to  $P_0$  in the  $C^\infty$ -topology. This even holds in the compact-open topology on holomorphic extensions, as mentioned earlier. It is known [11,19,62,86,154] that from the points  $(\beta, \varepsilon) = (p/q, 0)$ , resonance tongues (otherwise called *Arnold tongues*) emanate into the two open half-planes  $\varepsilon \neq 0$ , in such a way that for small  $\varepsilon$  an open and dense subset is covered. In the tongue emanating from  $(\beta, \varepsilon) = (p/q, 0)$ , the dynamics is asymptotically periodic with rotation number  $p/q$ . Compare with Figure 2. Theorem 2 implies that in the complement of this union of tongues, there exists a union of smooth curves that fill out positive measure. For parameter values on these curves, the dynamics is Diophantine quasi-periodic.

**Applications** Theorem 2 has applications for systems of ordinary differential equations with a 2-torus attractor, of which  $P_\varepsilon$  is a Poincaré map close to a rigid rotation family.

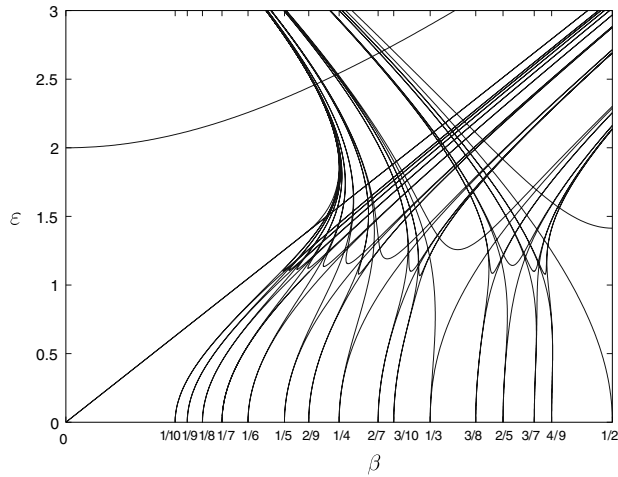


Fig. 2. Resonance tongues in the Arnold family [86].

In that case the Diophantine quasi-periodic subsystem of  $P_\varepsilon$  is referred to as a family of *quasi-periodic attractors* [369].

The basic ingredient of the examples to follow is a nonlinear oscillator with equation of motion

$$\ddot{u} + c\dot{u} + au + f(u, \dot{u}) = 0, \quad (12)$$

where  $u \in \mathbb{R}$  and  $\dot{u} = du/dt$ , which is assumed to have a hyperbolic periodic attractor, i.e. a periodic solution with a negative Floquet exponent. For the moment we consider coefficients like  $a$  and  $c$  as positive constants, but later on some of them occasionally will act as parameters. A classical example of such a system is the Van der Pol oscillator, where the nonlinearity is given by  $f(u, \dot{u}) = bu^2\dot{u}$ , with  $b$  a real constant.

As a first example with quasi-periodic attractors, consider the oscillator (12) subject to a weak time-periodic forcing:

$$\ddot{u} + c\dot{u} + au + f(u, \dot{u}) = \varepsilon g(u, \dot{u}, t), \quad (13)$$

$u, t \in \mathbb{R}$ , where  $g(u, \dot{u}, t + 2\pi/\Omega) \equiv g(u, \dot{u}, t)$ , and where  $\varepsilon$  is a small perturbation parameter. As usual we take the time  $t$  as an extra state variable, introducing the 3-dimensional (generalized) phase space  $\mathbb{R}^2 \times \mathbb{T}^1$  with coordinates  $(u, \dot{u}) \in \mathbb{R}^2$  and  $t \in \mathbb{T}^1 = \mathbb{R}/((2\pi/\Omega)\mathbb{Z})$ . Here the non-autonomous oscillator (13) defines the vector field  $X_\varepsilon$  given by

$$\begin{aligned} \dot{u} &= v \\ \dot{v} &= -au - cv - f(u, v) + \varepsilon g(u, v, t) \\ \dot{t} &= 1. \end{aligned} \quad (14)$$

In the unperturbed case,  $\varepsilon = 0$ , the oscillator is free and in (14) decouples from the third equation  $\dot{t} = 1$ . Combining the periodic attractor of the free oscillator with this third equation gives rise to an invariant 2-torus attractor to be denoted by  $T_0$ . By direct

techniques from ordinary differential equations [19,204] one can show that the Poincaré return map or stroboscopic map of the section  $t = 0 \bmod(2\pi/\Omega)\mathbb{Z}$  has the appropriate circle map format for  $C^k$ -versions of [Theorem 2](#) (for  $k \in \mathbb{N}$  large), for details compare with e.g. [61,77]. Here we use the fact that the torus attractor  $T_0$  is normally hyperbolic by hyperbolicity of the periodic orbit. Thus, according to the Centre Manifold Theorem, see [126,175,226,434],  $T_0$  persists as an invariant manifold. This means that, for  $|\varepsilon| \ll 1$ , the vector field  $X_\varepsilon$  has a smooth invariant 2-torus  $T_\varepsilon$  (close to  $T_0$ ), also depending smoothly on  $\varepsilon$ . Here ‘smooth’ means ‘ $C^k$ ’ for an appropriate  $k \in \mathbb{N}$ , which in this case tends to  $\infty$  as  $\varepsilon \rightarrow 0$ . Note that we need to regard a coefficient like  $a$  in (13) as a parameter, to obtain a family of quasi-periodic attractors in the Poincaré map. We summarize by saying that [Theorem 2](#) provides a family of quasi-periodic attractors in the system (14).

As a second example consider two nonlinear oscillators with a weak coupling

$$\begin{aligned}\ddot{u}_1 + c_1\dot{u}_1 + a_1u_1 + f_1(u_1, \dot{u}_1) &= \varepsilon g_1(u_1, u_2, \dot{u}_1, \dot{u}_2) \\ \ddot{u}_2 + c_2\dot{u}_2 + a_2u_2 + f_2(u_2, \dot{u}_2) &= \varepsilon g_2(u_1, u_2, \dot{u}_1, \dot{u}_2),\end{aligned}$$

$u_1, u_2 \in \mathbb{R}$ . This yields the following vector field  $X_\varepsilon$  on the 4-dimensional phase space  $\mathbb{R}^2 \times \mathbb{R}^2 = \{(u_1, v_1; u_2, v_2)\}$ :

$$\begin{aligned}\dot{u}_j &= v_j \\ \dot{v}_j &= -a_ju_j - c_jv_j - f_j(u_j, v_j) + \varepsilon g_j(u_1, u_2, v_1, v_2),\end{aligned}$$

$j = 1, 2$ . Note that for  $\varepsilon = 0$  the system decouples to a system of two independent oscillators, which has an attractor in the form of a 2-torus  $T_0$ . This torus arises as the product of two circles, along which each of the oscillators has its periodic solution. (The circles lie in the two-dimensional planes given by  $v_2 = u_2 = 0$  and  $v_1 = u_1 = 0$  respectively.) The persistence of  $T_0$  for  $|\varepsilon| \ll 1$  runs exactly as before and the Poincaré map is defined accordingly, again yielding a family of quasi-periodic attractors.

It may be clear that a similar coupling of  $n$  nonlinear oscillators gives rise to an attracting  $n$ -torus inside  $\mathbb{R}^{2n}$ . Below we shall obtain a formulation for this problem (see [Theorem 3](#)), which is more appropriate for generalization to higher dimensions (compare with [22]). There are higher dimensional analogues of the present situation, where next to periodicity and quasi-periodicity, also chaotic dynamics coexist [61,77,78]. This scenario and the transitions or bifurcations between the various kinds of dynamics have been associated with the onset of turbulence in fluid dynamics, see Ruelle, Takens, and Newhouse [333,370,371]. The quasi-periodic state then is seen as intermediate between very orderly and chaotic, also see [229,263,264]. In Section 4.4.2 we shall return to this subject.

### 3.2. Area preserving annulus maps

Close to the setting of circle maps (or of ‘vertical’ cylinder maps), is that of maps of the annulus that preserve area. This set-up relates to conservative dynamics by the Liouville Theorem [20,25]. First we develop Moser’s Twist Mapping Theorem [316,322] as an analogue of [Theorem 2](#).

**3.2.1. Moser's Twist Mapping Theorem.** Consider an annulus with ‘polar’ coordinates  $(x, y) \in \mathbb{T}^1 \times A$ , where  $A \subset \mathbb{R}$  is open. We endow  $\mathbb{T}^1 \times A$  with the area form  $\sigma = dx \wedge dy$ . Consider a  $C^\infty$ -map  $P_\varepsilon : \mathbb{T}^1 \times A \rightarrow \mathbb{T}^1 \times A$  of the form

$$P_\varepsilon(x, y) = (x + 2\pi\beta(y), y) + \varepsilon(f(x, y), g(x, y)), \quad (15)$$

that preserves the area  $\sigma$ , meaning that  $\det DP_\varepsilon \equiv 1$ . Note that for  $\varepsilon = 0$  this map leaves the family of circles  $y = \text{const}$  invariant, and again the problem is the persistence of this family when  $P_\varepsilon$  is  $C^\infty$ -near  $P_0$ , i.e. when  $|\varepsilon| \ll 1$ . We say that  $P_0$  is a (pure) twist map if the map  $y \mapsto \beta(y)$  is strictly monotonic<sup>5</sup> (hence a diffeomorphism) on  $A$ . We impose Diophantine conditions as before, see (3) and (9). To be more precise, for given constants  $\tau > 1$  and  $\gamma > 0$  we require again that

$$\left| \beta(y) - \frac{p}{q} \right| \geq \frac{\gamma}{q^{\tau+1}}, \quad (16)$$

for all rationals  $p/q$  with  $q > 0$ , or, in other words, that  $\beta(y) \in \mathbb{R}_{\tau, \gamma}$ . Define  $\Gamma = \beta(A)$  as well as subsets  $\Gamma_\gamma$  and  $\Gamma_{\tau, \gamma} \subset \Gamma$ , as in Section 3.1.2. We note that the map  $\beta$  pulls  $\Gamma_{\tau, \gamma}$  back to a subset  $A_{\tau, \gamma} \subset A$ , which, for  $\gamma > 0$  sufficiently small, ‘is’ a Cantor set of large measure.

**THEOREM 3 (Twist).** *In the above circumstances assume that  $\tau > 1$  and that  $\gamma > 0$  is sufficiently small. Then, if  $P_\varepsilon$  is sufficiently close to  $P_0$  in the  $C^\infty$ -topology, there exists a  $C^\infty$ -diffeomorphism of the annulus  $\Phi_\varepsilon : \mathbb{T}^1 \times A \rightarrow \mathbb{T}^1 \times A$  with the following properties.*

1.  $\Phi_\varepsilon$  is a  $C^\infty$ -near the identity map and depends  $C^\infty$ -ly on  $\varepsilon$ .
2. The image of the  $P_0$ -invariant union of circles  $\mathbb{T}^1 \times A_{\tau, \gamma}$  under  $\Phi_\varepsilon$  is  $P_\varepsilon$ -invariant, and the restricted map  $\widehat{\Phi}_\varepsilon = \Phi_\varepsilon|_{\mathbb{T}^1 \times A_{\tau, \gamma}}$  conjugates  $P_0$  to  $P_\varepsilon$ , that is

$$P_\varepsilon \circ \widehat{\Phi}_\varepsilon = \widehat{\Phi}_\varepsilon \circ P_0.$$

This theorem was first proven by J.K. Moser [316] for maps of the class  $C^{333}$ , for additional comments see [322]. Subsequently, H. Rüssmann reduced  $C^{333}$  to  $C^5$  [373], while F. Takens proved that  $C^1$  is not enough [424]. As was finally shown by M.R. Herman [218], Theorem 3 is carried over to  $C^3$ -mappings (with  $\tau = 1$ , however) but not to  $C^2$ -mappings whose second derivatives belong to the Hölder class  $C^{1-\delta}$ , however small  $\delta > 0$  is. The present formulation is close to Theorem 2 and, concerning the smoothness of  $\Phi_\varepsilon$ , the same comments apply again. For a nice exposition of the real analytic version of Theorem 3 (accompanied by a complete proof), see [414].

**REMARK.** The remark on the perturbation size depending on  $\gamma$  (see the first remark after the formulation of Theorem 2) also applies here. One example, where the limit as  $\gamma \downarrow 0$  can be taken, is near a generic elliptic fixed point, which thereby becomes a density point of quasi-periodicity. This situation often is referred to as a ‘small twist’ [316, 354].

<sup>5</sup>Or, equivalently, if  $d\beta/dy$  is nowhere zero.

**3.2.2. Discussion.** [Theorem 3](#) (on the annulus) is quite close to [Theorem 2](#) (on the cylinder), where the role of the parameter  $\beta$  has been taken by the action variable  $y$ . In the same spirit as in [Section 3.1.3](#), we conclude that for the area preserving case, typically quasi-periodicity occurs with positive measure in the phase space.

A difference in the settings of [Theorems 2](#) and [3](#) is that the former deals with ‘vertical’ maps while the latter does not. This means that generally in [Theorem 3](#), projections to the action space  $A = \{y\}$  are *not* preserved by the conjugation  $\Phi_\varepsilon$ . Moreover,  $\Phi_\varepsilon$  generally is not symplectic.

**Applications** We mimic the dissipative discussion presented in [Section 3.1.3](#). As a first example consider the mathematical (planar) pendulum with a (weak) periodic forcing. As possible equations of motion one may take

$$\ddot{u} + \omega^2 \sin u = \varepsilon \cos t \quad \text{or} \\ \ddot{u} + (\omega^2 + \varepsilon \cos t) \sin u = 0,$$

which give rise to volume preserving 3-dimensional vector fields. For simplicity we only write down the first example:

$$\begin{aligned} \dot{u} &= v \\ \dot{v} &= -\omega^2 \sin u + \varepsilon \cos t \\ \dot{t} &= 1. \end{aligned}$$

As is usual in mechanics, see e.g. [\[20,22,25\]](#) as well as [\[138,305\]](#), we introduce angle-action variables  $(x, y)$  for  $\varepsilon = 0$ , i.e. for the autonomous planar pendulum.<sup>6</sup> In fact, denoting the energy by  $H_0(u, v) = \frac{1}{2}v^2 - \omega^2 \cos u$ , we restrict ourselves to the oscillatory region where  $H_0(u, v) < \omega^2$ . Next consider any level set  $\{H_0(u, v) = h\}$ , with  $|h| < \omega^2$ . The action variable  $y$  then is defined by

$$y(h) = \frac{1}{2\pi} \oint_{H_0(u,v)=h} v \, du,$$

which is proportional to the area enclosed by the level set. The angle variable  $x$  is obtained by taking the time parametrization of the periodic motion inside this level set scaled to period  $2\pi$ . Thus one obtains the canonical equations

$$\dot{y} = 0, \quad \dot{x} = \beta(y)$$

for the oscillatory motions of the planar pendulum.

One easily sees that the Poincaré (or stroboscopic) map  $P_\varepsilon$  for a perturbed pendulum has the form [\(15\)](#) and is area preserving. A direct computation, involving an elliptic integral, shows that  $P_0$  is a pure twist map. According to [Theorem 3](#), the conclusion of quasi-periodicity occurring with positive measure in the phase space does apply here.

---

<sup>6</sup>In the literature, such angle-action variables often are denoted by  $(x, y) = (\varphi, I)$  or  $(\alpha, a)$ .

A related application deals with two coupled oscillators

$$\begin{aligned}\ddot{u}_1 &= -\omega_1^2 \sin u_1 - \varepsilon \frac{\partial U}{\partial u_1}(u_1, u_2) \\ \ddot{u}_2 &= -\omega_2^2 \sin u_2 - \varepsilon \frac{\partial U}{\partial u_2}(u_1, u_2),\end{aligned}$$

leading to a 4-dimensional Hamiltonian vector field

$$\begin{aligned}\dot{u}_j &= v_j \\ \dot{v}_j &= -\omega_j^2 \sin u_j - \varepsilon \frac{\partial U}{\partial u_j}(u_1, u_2),\end{aligned}$$

$j = 1, 2$ , with the Hamilton function

$$H_\varepsilon(u_1, u_2, v_1, v_2) = \frac{1}{2}v_1^2 + \frac{1}{2}v_2^2 - \omega_1^2 \cos u_1 - \omega_2^2 \cos u_2 + \varepsilon U(u_1, u_2).$$

In this case, it is the iso-energetic Poincaré maps that obtain the form  $P_\varepsilon$  of (15). Now Theorem 3 yields the conclusion of quasi-periodicity occurring with positive measure in the energy hypersurfaces of  $H_\varepsilon$ .

This discussion also leads to higher dimensional explorations in KAM Theory.

#### 4. KAM Theory for flows

We turn to the context of smooth vector fields on manifolds (locally corresponding to systems of ordinary differential equations). First we give a formal definition of quasi-periodicity, next considering a few conceptual aspects.

##### 4.1. Introduction

Let  $M$  be a  $C^\infty$ -manifold and  $X$  a  $C^\infty$ -vector field on  $M$ . Fix  $n \in \mathbb{N}$  with  $n \geq 2$ . Also consider the standard  $n$ -torus  $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$  endowed with coordinates  $x_1, x_2, \dots, x_n$  counted modulo  $2\pi\mathbb{Z}$ . For any vector  $\omega \in \mathbb{R}^n$ , on  $\mathbb{T}^n$  we consider the constant vector field

$$\mathbb{X}_\omega = \sum_{j=1}^n \omega_j \frac{\partial}{\partial x_j}, \quad (17)$$

in the system form given by

$$\dot{x}_j = \omega_j, \quad 1 \leq j \leq n.$$

Assume that  $T \subseteq M$  is an  $X$ -invariant  $n$ -torus. We say that the restriction  $X|_T$  is *parallel* or *conditionally periodic* whenever there exist a vector  $\omega \in \mathbb{R}^n$  and a  $C^\infty$ -diffeomorphism

$\Phi : T \rightarrow \mathbb{T}^n$  conjugating  $X|_T$  with a constant vector field  $\mathbb{X}_\omega$ , which means that  $\Phi_*(X|_T) = \mathbb{X}_\omega$ . We refer to the  $\omega_j$  as the *frequencies* of  $X|_T$  and to  $\omega = (\omega_1, \omega_2, \dots, \omega_n)$  as the *frequency vector* of  $X|_T$ .

We first note that  $\omega$  is not uniquely determined by  $T$ , but depends also on  $\Phi$ . If we compose  $\Phi$  by a translation on  $\mathbb{T}^n$ ,  $\omega$  does not change. However, let us consider an invertible linear map  $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , then  $S$  projects to a torus diffeomorphism if and only if  $S \in GL(n, \mathbb{Z})$ .<sup>7</sup> Now if  $\Phi$  is composed by  $S$ , the frequency vector changes to  $S\omega$ , since  $S_*\mathbb{X}_\omega = \mathbb{X}_{S\omega}$ .

The dynamics on  $X|_T$  is said to be *quasi-periodic* (or *non-resonant*) whenever the frequencies  $\omega_1, \omega_2, \dots, \omega_n$  are independent over the rationals  $\mathbb{Q}$ , compare with Section 3.1.3. We also call  $T$  a quasi-periodic invariant  $n$ -torus of  $X$ . Observe that such a quasi-periodic  $n$ -torus  $T$  is densely filled by each of its  $X$ -trajectories. On the other hand, if the frequencies  $\omega_1, \omega_2, \dots, \omega_n$  satisfy  $l$  independent resonance relations ( $1 \leq l \leq n$ ), then the torus  $T$  is foliated into invariant  $(n - l)$ -tori of  $X$  (which are quasi-periodic for  $l \leq n - 2$ ), and each  $X$ -trajectory on  $T$  densely fills one of these tori.

REMARKS. 1. Note that the definitions of parallel and quasi-periodic dynamics on  $T$  do not depend on the choice of the conjugation  $\Phi$ . Moreover, the so-called *frequency module* (or *lattice of frequencies*)

$$\mathcal{L}(\omega) = \{\omega_1 k_1 + \omega_2 k_2 + \dots + \omega_n k_n \mid k \in \mathbb{Z}^n\}$$

of an invariant  $n$ -torus with parallel dynamics and frequencies  $\omega_1, \omega_2, \dots, \omega_n$  is determined uniquely [245, 318, 319]. To be more precise, it is not hard to verify that  $\mathcal{L}(\omega) = \mathcal{L}(\omega')$ , if and only if there exists an operator  $S \in GL(n, \mathbb{Z})$  such that  $\omega' = S\omega$ .

2. A similar formal definition of quasi-periodicity can be given for diffeomorphisms of a torus. In the previous sections we implicitly used such a definition in the case of circle diffeomorphisms and of holomorphic maps, compare with Section 3.1.3.
3. The structure of orbits of  $GL(n, \mathbb{Z})$  or of  $SL(n, \mathbb{Z}) = \{S \in GL(n, \mathbb{Z}) \mid \det S = 1\}$  in  $\mathbb{R}^n$  is known: J.S. Dani's theorem [143] states that if  $\omega \in \mathbb{R}^n$  is not proportional to an integer vector (i.e. if the trajectories of  $\mathbb{X}_\omega$  are not closed), then the orbit  $SL(n, \mathbb{Z})\omega$  is *dense* in  $\mathbb{R}^n$ , and vice versa. There are generalizations of this result to the action of the group  $SL(n, \mathbb{Z})$  on the space of  $d$ -frames in  $\mathbb{R}^n$ ,  $2 \leq d \leq n - 1$  [144, 199]. So, the set of all the frequency vectors that can be assigned to a given conditionally periodic motion on an  $n$ -torus  $T$  is in fact *dense* in  $\mathbb{R}^n$ , provided that this motion is not 'maximally' resonant (see also a discussion in [409]).

**4.1.1. Affine structure.** We shall argue that on any quasi-periodic  $X$ -invariant  $n$ -torus  $T$ , a natural affine structure is defined. To this end first observe that the self-conjugations of the constant vector field  $\mathbb{X}_\omega$ , with rationally independent frequencies  $\omega_1, \omega_2, \dots, \omega_n$ , are exactly the translations of the standard torus  $\mathbb{T}^n$ . This directly follows from the fact that

<sup>7</sup>  $GL(n, \mathbb{Z})$  is the group of linear operators in  $\mathbb{R}^n$  represented by matrices with integer entries and determinant  $\pm 1$ .



each trajectory of  $\mathbb{X}_\omega$  is dense. Note that these translations determine the affine structure on  $\mathbb{T}^n$ .

**LEMMA 4** (Affine Structure [60,91]). *In the above setting with a  $C^\infty$ -diffeomorphism  $\Phi : T \rightarrow \mathbb{T}^n$  such that  $\Phi_*(X|_T) = \mathbb{X}_\omega$ , the self-conjugations of  $X|_T$  determine a natural affine structure on  $T$ . The translations of  $T$  are self-conjugations and the conjugation  $\Phi$  is unique modulo torus translations on  $T$  and  $\mathbb{T}^n$ .*

**REMARK.** Note that the structure on  $T$  is a bit stronger than affine, since the transition maps are not general elements of  $GL(n, \mathbb{R})$ , but are restricted to  $GL(n, \mathbb{Z})$ .

**4.1.2. The perturbation problem.** We start preparing the perturbation problem. As before, for simplicity we formulate all the results in the  $C^\infty$ -topology [225,325].

We consider an unperturbed vector field as *integrable*, when it is invariant under a suitable action of the  $n$ -torus  $\mathbb{T}^n$  ( $n \geq 2$ ), whereas the invariant tori in question are orbits of this action, compare with [76–78,236]. In this case we can restrict our attention to a subset of the phase space  $M$  diffeomorphic to  $\mathbb{T}^n \times \mathbb{R}^m$  ( $m \geq 0$ ). We note that any (isolated) parallel  $n$ -torus  $T$ , by definition, is integrable.

Now for any  $A \subset \mathbb{R}^m$  open (and bounded), we may consider the  $C^k$ -norm  $\|\cdot\|_{k,\bar{A}}$  for  $C^\infty$ -functions on the closure  $\mathbb{T}^n \times \bar{A}$ . The  $C^\infty$ -topology on  $C^\infty(\mathbb{T}^n \times \mathbb{R}^m)$  then is generated by all such norms. This induces similar topologies on all spaces of smooth dynamical systems [225,325].

As in the previous section (see Sections 3.1.2 and 3.2.1), the KAM theorems will be formulated in terms of conjugations between subsystems of an integrable system and its perturbation, which are nearby in the  $C^\infty$ -topology. This is the structural stability formulation, for this occasion called *quasi-periodic stability*, as already mentioned in Section 3.1.3. The subsystems will be (unions of) quasi-periodic invariant  $n$ -tori, so with non-resonant frequencies  $\omega_1, \omega_2, \dots, \omega_n$ . In fact and as before, see inequalities (3), (9), (16), we shall need stronger nonresonance conditions to be introduced now.

**REMARK.** In many applications, the open set  $A \subset \mathbb{R}^m$  is given by a local nondegeneracy condition via the Inverse Function Theorem [225,325,418].

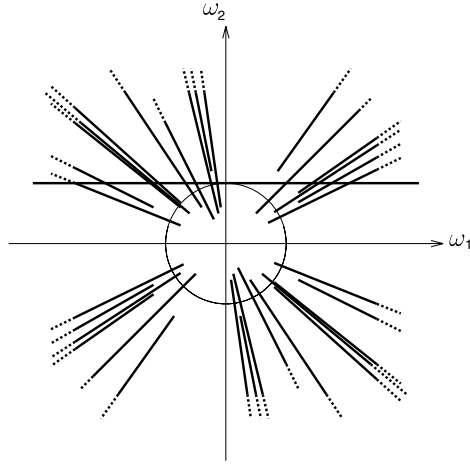
**Diophantine conditions** The strong nonresonance conditions on the frequencies, as mentioned above, are Diophantine in the following sense. Let  $\tau > n - 1$  and  $\gamma > 0$  be constants. Set

$$D_{\tau,\gamma}(\mathbb{R}^n) = \left\{ \omega \in \mathbb{R}^n \mid |\langle \omega, k \rangle| \geq \gamma |k|^{-\tau}, \text{ for all } k \in \mathbb{Z}^n \setminus \{0\} \right\}. \quad (18)$$

Here and henceforth,

$$\langle \omega, k \rangle = \sum_{j=1}^n \omega_j k_j \quad \text{and} \quad |k| = \sum_{j=1}^n |k_j|.$$

Elements of  $D_{\tau,\gamma}(\mathbb{R}^n)$  are called  $(\tau, \gamma)$ -*Diophantine* frequency vectors. One easily sees that  $D_{\tau,\gamma}(\mathbb{R}^n)$  is a closed subset with the following closed half-line property: whenever

Fig. 3. Sketch of the set  $D_{\tau, \gamma}(\mathbb{R}^2)$ .

$\omega \in D_{\tau, \gamma}(\mathbb{R}^n)$  and  $s \geq 1$ , then also the product  $s\omega \in D_{\tau, \gamma}(\mathbb{R}^n)$ , compare with Figure 3. The intersection  $\mathbb{S}_{\tau, \gamma} = D_{\tau, \gamma}(\mathbb{R}^n) \cap \mathbb{S}^{n-1}$  of  $D_{\tau, \gamma}(\mathbb{R}^n)$  with the unit sphere, again, is a closed (even a compact) set. An application of the Cantor–Bendixson theorem [214] yields that  $\mathbb{S}_{\tau, \gamma}$  is the union of a perfect and a countable set. Since the resonant hyperplanes (with equations  $\langle \omega, k \rangle = 0$ ,  $k \in \mathbb{Z}^n \setminus \{0\}$ ) give a dense web in the complement  $\mathbb{S}^{n-1} \setminus \mathbb{S}_{\tau, \gamma}$ , it follows that this perfect set is totally disconnected. Summing up, we conclude that the perfect subset of  $\mathbb{S}_{\tau, \gamma}$  is a Cantor set. This implies that  $D_{\tau, \gamma}(\mathbb{R}^n)$  is nowhere dense. Moreover, the measure of  $\mathbb{S}^{n-1} \setminus \mathbb{S}_{\tau, \gamma}$  is of the order of  $\gamma$  as  $\gamma \downarrow 0$ , compare with the discussion on the Diophantine condition (9) in Section 3.1.2.

In Figure 3 we roughly sketch the set  $D_{\tau, \gamma}(\mathbb{R}^2)$ . Considering the intersection with the line  $\omega_2 = 1$  directly gives the connection of (18) with (3), (9) and (16), compare with [132].

In the sequel, we shall regard  $\tau$  as a *fixed* number greater than  $n - 1$ .

**Mild Diophantine conditions** Diophantine conditions like (3), (9), (16) and (18), as well as future versions (30) (in Section 5.2.1), (33) (in Section 6.1.1), (36) (in Section 6.1.3), and (52) (in Section 8.3.1), serve to overcome the small divisor problems that are inherent to KAM Theory as described in Sections 2.2 and 3.1.1. For a more detailed treatment of small divisors we refer ahead to Section 5.1.2.

In the literature (see e.g. [189, 355, 376–380]), also ‘milder’ Diophantine conditions occur, like

$$|\langle \omega, k \rangle| \geq \frac{\gamma}{\Delta(|k|)} \quad \text{for all } k \in \mathbb{Z}^n \setminus \{0\}.$$

Here  $\Delta$  is an *approximation function*, i.e. an arbitrary monotonically increasing (or just nondecreasing) continuous function  $\Delta : [1, \infty) \rightarrow \mathbb{R}_+ = [0, \infty)$  such that  $\Delta(1) > 0$  and

$$\int_1^\infty \frac{\log \Delta(u) \, du}{u^2} < +\infty.$$

Note that the usual Diophantine condition (18) corresponds to the case  $\Delta(u) = u^\tau$ .

Yet another mild version of the Diophantine condition is given by [189]

$$\sum_{j \geq 0} 2^{-j} \log \frac{1}{\min_{0 < |k| \leq 2^j} |\langle \omega, k \rangle|} < +\infty,$$

which, in reminiscence of (5), also is called the Bruno condition, compare with e.g. [97, 98, 355, 372, 375–380, 465, 466].

For a similar discussion also see [77, §1.5.2].

**Shrinking the frequency domain** Let  $\Gamma \subset \mathbb{R}^n$  be an open subset of frequency vectors. We also consider the shrunken version of  $\Gamma$  defined by

$$\Gamma_\gamma = \{\omega \in \Gamma \mid \text{dist}(\omega, \partial\Gamma) \geq \gamma\}.$$

Let  $D_{\tau,\gamma}(\Gamma_\gamma) = \Gamma_\gamma \cap D_{\tau,\gamma}(\mathbb{R}^n)$ . In the sequel we always take  $\gamma$  sufficiently small to ensure that the nowhere dense set  $D_{\tau,\gamma}(\Gamma_\gamma)$  has positive measure. Recall that the measure of  $\Gamma \setminus D_{\tau,\gamma}(\Gamma_\gamma)$  tends to zero in  $\Gamma$  as  $\gamma \downarrow 0$ , compare with [14, 15, 19, 77, 78, 236, 354].

## 4.2. Families of normally hyperbolic quasi-periodic tori

Now we are in a position to deal with the KAM theorem for normally hyperbolic invariant tori. By the Centre Manifold Theory [126, 175, 226, 434] such tori, *as invariant manifolds*, persist under small perturbations. We shall restrict to such ‘centre manifold’ tori, setting up the above perturbation program for this case. We recall that parallelity of such a torus implies integrability.

So, we assume parallelity of certain unperturbed ‘centre manifold’ tori, which implies  $\infty$ -normal hyperbolicity. The perturbed tori then are diffeomorphic to the unperturbed ones. The diffeomorphism has a finite degree, say  $k$ , of smoothness [437]. Moreover, the perturbed tori are unique, i.e. independent of  $k$ , and we can take  $k \rightarrow \infty$  as the perturbation becomes small.

Summarizing, we assume that the phase space of both the unperturbed and the perturbed systems consists of the  $n$ -torus  $\mathbb{T}^n$ , where, as usual, we take  $n \geq 2$ . The interest again is the persistence of parallelity. Referring to similar discussions in Section 3.1.3, we can see that we have to be restricted to Diophantine quasi-periodic invariant tori. As said before at the beginning of Section 3, the KAM Theory in question has  $C^k$ -versions for sufficiently large  $k \in \mathbb{N}$ , but for simplicity we formulate all KAM theorems in the sequel in the  $C^\infty$ -setting, in a similar way to Theorems 2 and 3 above. As in the case of circle maps, see Section 3.1, we need parameters for persistence of quasi-periodic  $n$ -tori.

**4.2.1. Formulation of the normally hyperbolic KAM theorem.** Let  $P \subseteq \mathbb{R}^s$  be an open set of parameters and consider families of vector fields  $X = X_\mu(x)$ , with  $x \in \mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$  and  $\mu \in P$ . We shall treat such a family as a ‘vertical’ vector field on the product  $\mathbb{T}^n \times P$ . Throughout we assume a  $C^\infty$ -dependence of all the vector fields on both

$x$  and  $\mu$ . As before we often use the vector field notation, writing

$$f(x, \mu) \frac{\partial}{\partial x} \quad \text{instead of} \quad \dot{x} = f(x, \mu),$$

compare with (17). The starting point is an integrable family

$$X_\mu(x) = \omega(\mu) \frac{\partial}{\partial x}, \quad (19)$$

$x \in \mathbb{T}^n$ ,  $\mu \in P$ , where integrability amounts to  $x$ -independence, which expresses invariance under the natural  $\mathbb{T}^n$ -action. Our interest is the family of  $X$ -invariant  $n$ -tori  $\mathbb{T}^n \times \{\mu\}$ , where  $\mu \in P$ . The smooth map  $\omega : P \rightarrow \mathbb{R}^n$  is called the *frequency map*.

The family  $X$  is said to be *nondegenerate* at  $\mu_0 \in P$  if the derivative  $D_{\mu_0}\omega$  is surjective (in particular, this implies  $s \geq n$ ). As in Section 3, we are especially interested in the fate of the  $X$ -invariant tori  $\mathbb{T}^n \times \{\mu\}$ ,  $\mu \in P$ , under smooth perturbations

$$\tilde{X}_\mu = X_\mu + \tilde{f}(x, \mu) \frac{\partial}{\partial x}, \quad (20)$$

when  $\mu$  is near  $\mu_0$ , and where the size of  $\tilde{X} - X$  (i.e. the size of  $\tilde{f}$ ) is small in the  $C^\infty$ -topology. Here we confine our attention to the tori that are Diophantine in the sense of (18).

By the Inverse Function Theorem [225,325,418], the nondegeneracy condition implies that the point  $\mu_0 \in P$  has an open bounded neighbourhood  $A \subseteq P$ , restricted to which the map  $\omega$  is a submersion, i.e. is conjugated to the projection on a lower dimensional subspace. We then say that  $X$  is *nondegenerate* on  $\mathbb{T}^n \times A$ . Defining  $\Gamma = \omega(A)$ , we consider shrunken versions  $\Gamma_\gamma$  and  $A_\gamma = \omega^{-1}(\Gamma_\gamma)$  of the domains  $\Gamma$  and  $A$ , respectively, see Section 4.1.2 above. Accordingly we consider the set of Diophantine frequency vectors  $D_{\tau,\gamma}(\Gamma_\gamma)$  and its pull-back  $D_{\tau,\gamma}(A_\gamma)$  along the frequency map. Note that the measure of  $A \setminus D_{\tau,\gamma}(A_\gamma)$  tends to 0 as  $\gamma \downarrow 0$ .

**THEOREM 5** (Normally Hyperbolic KAM [76–78,132,236]). *Let  $n \geq 2$ . Let the integrable  $C^\infty$ -family  $X = X_\mu(x)$  of vector fields (19) be nondegenerate on  $\mathbb{T}^n \times A$ , with  $A \subseteq P$  open. Then, for  $\gamma > 0$  sufficiently small, there exists a neighbourhood  $\mathcal{O}$  of  $X$  in the  $C^\infty$ -topology, such that for any perturbed family  $\tilde{X} \in \mathcal{O}$  as in (20), there exists a mapping  $\Phi : \mathbb{T}^n \times A \rightarrow \mathbb{T}^n \times A$  with the following properties.*

1.  $\Phi$  is a  $C^\infty$ -diffeomorphism onto its image and a  $C^\infty$ -near the identity map. Also,  $\Phi$  preserves projections to  $P$ .
2. The image of the  $X$ -invariant torus union  $\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma)$  under  $\Phi$  is  $\tilde{X}$ -invariant, and the restricted map  $\hat{\Phi} = \Phi|_{\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma)}$  conjugates  $X$  to  $\tilde{X}$ , that is

$$\hat{\Phi}_* X = \tilde{X}.$$

**REMARKS.** 1. As in the case of Theorems 2 and 3, the fact that  $\Phi$  is a  $C^\infty$ -near the identity map means that in the  $C^\infty$ -topology, whenever  $\tilde{X} \rightarrow X$  also  $\Phi \rightarrow \text{Id}$ .

2. The map  $\Phi$  takes  $\mathbb{T}^n \times A_\gamma$  into  $\mathbb{T}^n \times A$ . [Theorem 5](#) asserts that the integrable system (19) is (locally) quasi-periodically stable [76–78,236], compare with the discussion in Section 3.1.3 following [Theorem 2](#).
3. As in Section 3.1, the diffeomorphism  $\Phi$  is taken of a skew form, which means that the projection  $\mathbb{T}^n \times A \rightarrow A$  is preserved, or, in other words, that the transformation in the  $\mu$ -direction is independent of the angles  $x \in \mathbb{T}^n$ . This transformation in the  $\mu$ -direction gives a diffeomorphic image of the union of closed half-lines in  $D_{\tau,\gamma}(A_\gamma)$ .
4. There are direct generalizations of [Theorem 5](#) to the world of  $C^k$ -systems endowed with the  $C^k$ -topology [225,325] for  $k \in \mathbb{N}$  sufficiently large. For  $C^k$ -versions of the present KAM Theory, see [10,118,124,198,218,267,268,316,354,373,382,383,423,471]; for discussions on this subject also see [15,77,78,236,322,406]. To give an idea, for  $k > 4\tau + 2$ , the conjugation is at least of class  $C^{k-2\tau}$ , also compare with Section 5.1.2 below. Therefore in the  $C^\infty$ -case, no losses of differentiability occur and the conjugations also are of class  $C^\infty$ . In the real analytic case, the conjugations are even Gevrey smooth [350,351,408,444,460,473]. More generally, the conjugations also are Gevrey smooth as soon as the original system is [352,353,474,475].

**4.2.2. Discussion.** In [Theorem 5](#), the restriction of  $\Phi$  to the union of the Diophantine quasi-periodic tori  $\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma)$  preserves the natural affine structure of the quasi-periodic tori, see Section 4.1.1. In the complement of this nowhere dense set the diffeomorphism  $\Phi$  has *no* dynamical meaning.

The set-up (and proof) of [Theorem 5](#) is very close to the ‘classical’ KAM theorem for Lagrangean invariant tori in nearly integrable Hamiltonian systems (see Section 4.3) in the formulation of J. Pöschel [354], also compare with [76–78,132,236]. In fact [Theorem 5](#) relates to this ‘classical’ KAM theorem as [Theorem 2](#) does to [Theorem 3](#).

The normally hyperbolic KAM [Theorem 5](#), particularly, is relevant in the case where the tori are attractors, in which case we are dealing with *families of quasi-periodic attractors* [369]. In the discussion in Section 3.1.3 following [Theorem 2](#), we considered the case  $n = 2$ , where generically on the tori in between the quasi-periodic ones, only regular dynamics could occur with a Poincaré map that is Kupka–Smale [339]. In particular this refers to the oscillator models in Section 3.1.3. For  $n \geq 3$  the situation ‘in between’ the quasi-periodic tori can be a little different, since invariant 3-tori can contain strange attractors [333].

Below, in Section 6, we shall discuss several scenarios where families of quasi-periodic attractors undergo bifurcations. One of these is the quasi-periodic Hopf bifurcation, where the torus attractors lose stability and a higher dimensional torus family branches off.

Computational and numerical aspects of normally hyperbolic quasi-periodic tori have been dealt with in [86,101–103,388], also see [416].

### 4.3. KAM Theory for Lagrangean tori in Hamiltonian systems

We switch to the world of Hamiltonian systems of class  $C^\infty$ . This means, first of all, that the phase space  $M$  is now a symplectic manifold, say of dimension  $2n$ . First consider

a Liouville integrable Hamiltonian system with  $n$  degrees of freedom, recalling that this means that there exist  $n$  independent integrals in involution, including the energy  $H$ . If the energy level locally is compact, the Liouville–Arnold Integrability Theorem [13,20,22,25,138,305] provides us with angle-action variables  $(x, y) \in \mathbb{T}^n \times \mathbb{R}^n$ . This implies that the corresponding Hamiltonian vector field  $X = X_H$  possesses  $n$ -parameter families of Lagrangean invariant  $n$ -tori  $T_y$ , parametrized by  $y$ . The word ‘Lagrangean’ means that the dimension of the tori is equal to the number of degrees of freedom and that the restriction of the symplectic form to these tori vanishes.

This situation fits well in the general approach of this section, since the Liouville–Arnold Integrability Theorem and its proof provide a smooth conjugation of  $X_H|_{T_y}$  and a constant vector field on the standard torus  $\mathbb{T}^n$ , as well as an affine structure that fits with Lemma 4. Note that in this Hamiltonian setting, both the conjugation and the affine structure extend to all parallel (or conditionally periodic) invariant  $n$ -tori.

**4.3.1. Formulation of the Lagrangean KAM theorem.** For  $n \geq 2$  consider  $P \subseteq \mathbb{R}^n$  as an open domain. As before, let  $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$  denote the standard  $n$ -torus. The product  $M = \mathbb{T}^n \times P$  is endowed with coordinates  $(x, y) = (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n)$ , where the  $x_j$  are counted modulo  $2\pi\mathbb{Z}$ , and with the symplectic form  $\sigma = dx \wedge dy = \sum_{j=1}^n dx_j \wedge dy_j$ .

Assume that the Hamiltonian  $H : M \rightarrow \mathbb{R}$  does not depend on the angle variable  $x$ . The corresponding Hamiltonian vector field  $X_H$ , defined by the relation  $\iota_{X_H}\sigma = dH$ ,<sup>8</sup> then takes the form

$$X_H(x, y) = \omega(y) \frac{\partial}{\partial x} = \sum_{j=1}^n \omega_j(y) \frac{\partial}{\partial x_j}, \quad (21)$$

where  $\omega(y) = \partial H(y)/\partial y$  is the frequency vector. Formula (21) expresses that  $(x, y)$  is a collection of angle-action variables for  $X_H$  [1,9,20,22,25,138]. As in Section 4.2.1, we call  $\omega : P \rightarrow \mathbb{R}^n$  the *frequency map*, saying that  $H$  is *Kolmogorov nondegenerate* at  $y_0 \in P$  if the derivative  $D_{y_0}\omega$  is invertible.

As in Sections 3.1, 3.2 and 4.2, we are particularly interested in the fate of the  $X$ -invariant Lagrangean tori  $\mathbb{T}^n \times \{y\}$ ,  $y \in P$ , under smooth Hamiltonian perturbations  $\tilde{X}$  of  $X = X_H$ , when  $y$  is near  $y_0$  and where the size of  $\tilde{X} - X$  is small in the  $C^\infty$ -topology. Again, as before, we confine our attention to the tori that are Diophantine in the sense of (18).

By the Implicit Function Theorem [225,325,418], the nondegeneracy condition implies that  $y_0 \in P$  has an open bounded neighbourhood  $A \subseteq P$ , restricted to which the map  $\omega$  is a diffeomorphism. We then say that  $X = X_H$  is *Kolmogorov nondegenerate* on  $\mathbb{T}^n \times A$ . Defining  $\Gamma = \omega(A)$ , we consider shrunken versions  $\Gamma_\gamma$  and  $A_\gamma = \omega^{-1}(\Gamma_\gamma)$  of the domains  $\Gamma$  and  $A$ , respectively, see Section 4.1.2 above. Accordingly we consider the set of Diophantine frequency vectors  $D_{\tau,\gamma}(\Gamma_\gamma)$  and its pull-back  $D_{\tau,\gamma}(A_\gamma)$  along the

<sup>8</sup>Recall that  $(\iota_X\sigma)(Y) = \sigma(X, Y)$  for any vector field  $Y$  [1,9,20,25,77,360]. In the literature, Hamiltonian vector fields  $X_H$  are often defined by  $\iota_{X_H}\sigma = -dH$ .

frequency map. Note that the frequency map pulls back the union of closed half-lines in  $D_{\tau,\gamma}(\Gamma_\gamma)$  in a diffeomorphic way. Also note that the measure of  $A \setminus D_{\tau,\gamma}(A_\gamma)$  tends to 0 as  $\gamma \downarrow 0$ .

Staying within the class of Hamiltonian systems, we perturb the Hamiltonian vector field  $X = X_H$  in the  $C^\infty$ -topology as described in Section 4.1.2.

**THEOREM 6** (Lagrangian KAM [76–78,236,354]). *Suppose that the integrable Hamiltonian  $C^\infty$ -system  $X = X_H$  is Kolmogorov nondegenerate on  $\mathbb{T}^n \times A$ , for  $A \subseteq P$  open. Then, for sufficiently small  $\gamma > 0$ , there exists a neighbourhood  $\mathcal{O}$  of  $X$  in the  $C^\infty$ -topology such that for each Hamiltonian vector field  $\tilde{X} \in \mathcal{O}$ , there exists a mapping  $\Phi : \mathbb{T}^n \times A \rightarrow \mathbb{T}^n \times A$  with the following properties.*

1.  $\Phi$  is a  $C^\infty$ -diffeomorphism onto its image and a  $C^\infty$ -near the identity map.
2. The image of the  $X$ -invariant torus union  $\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma)$  under  $\Phi$  is  $\tilde{X}$ -invariant, and the restricted map  $\hat{\Phi} = \Phi|_{\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma)}$  conjugates  $X$  to  $\tilde{X}$ , that is

$$\hat{\Phi}_* X = \tilde{X}.$$

**REMARKS.** 1. The map  $\Phi$ , that takes  $\mathbb{T}^n \times A_\gamma$  into  $\mathbb{T}^n \times A$ , generally is not symplectic. [Theorem 6](#) asserts that the integrable system [\(21\)](#) is (locally) quasi-periodically stable [76–78,236], compare with the discussion in Section 3.1.3 following [Theorem 2](#) and with Remark 2 in Section 4.2.1 following [Theorem 5](#).

2. Unlike in the case of [Theorem 5](#), here the projection  $\mathbb{T}^n \times A \rightarrow A$  is *not* preserved by the map  $\Phi$ . Nevertheless, the perturbed union of invariant tori  $\Phi(\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma))$ , up to a diffeomorphism, is organized in terms of closed half-lines as described in Remark 3 following [Theorem 5](#).

3. The discussion on  $C^k$ -generalizations of [Theorem 6](#), for  $k \in \mathbb{N}$  sufficiently large, runs as in Remark 4 following [Theorem 5](#).

**4.3.2. Discussion.** We now discuss several variations on the ‘classical’ Lagrangian KAM [Theorem 6](#). Note that each  $\tilde{X}$ -invariant  $n$ -torus given by [Theorem 6](#) is usually called a KAM torus.

**A typicality conclusion** [Theorem 6](#) is the ‘classical’ KAM theorem in its stability formulation, for earlier formulations see e.g. [14,251] (this theorem also admits formulations quite different from the one we presented, see e.g. [148,382]). It implies that in Hamiltonian systems with  $n$  degrees of freedom, typically quasi-periodic Lagrangian  $n$ -tori occur with positive Liouville measure in the phase space. As said before in Section 1.2, ‘typically’ here means that classes of examples exist that are  $C^\infty$ -ly open. These examples are close to certain Liouville integrable systems [20,22,25,138,305] or are locally so. For instance, this description applies to any Hamiltonian system with two or more degrees of freedom near a so-called Birkhoff nondegenerate elliptic equilibrium point. Here strong resonances are forbidden, which by Normal Form Theory [19,20,25,34,47,88,98,99,127,129,203,320,323,379,414] implies local near-integrability. Moreover, an appropriate local Kolmogorov nondegeneracy condition has to be satisfied, amounting to the nonvanishing of a specific normal form coefficient. As a consequence, in a neighbourhood of such an

equilibrium, there are many KAM tori. Their union is of positive Liouville measure, where the equilibrium point even is a Lebesgue density point of quasi-periodicity [77,354].

**Iso-energetic nondegeneracy** It should be mentioned here that the above conclusions generically also apply when restricted to the energy levels. This is a consequence of the so-called *iso-energetic* KAM theorem, which is a slight variation of [Theorem 6](#). The difference is mainly due to the nondegeneracy conditions imposed on the integrable approximation.

Indeed, for the ‘standard’ KAM [Theorem 6](#), the Kolmogorov nondegeneracy condition [14,251] requires that the derivative of the frequency map  $\omega : A \rightarrow \mathbb{R}^n$  should have maximal rank  $n$ ; this implies that the frequency map  $y \mapsto \omega(y)$  locally is a diffeomorphism.

The Arnold condition for iso-energetic nondegeneracy [14,15] similarly requires that  $\omega$  should nowhere vanish in  $A$  and the derivative of the corresponding frequency-ratio map

$$A \ni y \mapsto [\omega(y)] = [\omega_1(y) : \omega_2(y) : \dots : \omega_n(y)] \in \mathbb{P}^{n-1}(\mathbb{R})$$

should have maximal rank  $n - 1$  on each energy hypersurface  $H^{-1}(c)$ , where  $\mathbb{P}^{n-1}(\mathbb{R})$  is the  $(n - 1)$ -dimensional real projective space, also see [20,25,74,77,151,236,408,409,414,431]. In coordinates, this condition means that the so-called Arnold determinant

$$\det \begin{pmatrix} \partial\omega/\partial y & \omega \\ \omega & 0 \end{pmatrix}$$

of order  $n+1$  vanishes nowhere in  $A$ . Several equivalent reformulations of the iso-energetic nondegeneracy condition are compiled in [151,409].

**Rüssmann nondegeneracy** Another companion KAM theorem was proven by H. Rüssmann in the mid 1980s. In fact, Rüssmann announced his result in [376] and presented the proof in a number of talks; a detailed written account of the proof, however, appeared only in a 1998 preprint which was published in 2001 [378]—three years later still (see also [381]). Meanwhile, different proofs were published in the mid 1990s by other authors [77,114,397,400]. Again, let  $\mathbb{T}^n \times A = \{(x, y)\}$ , for  $A \subseteq \mathbb{R}^n$  open, be the domain of definition of a Hamilton function  $H$ , where integrability amounts to  $x$ -independence of  $H$ . As before, let  $\omega(y) = \partial H(y)/\partial y$  be the corresponding frequency vector. We now say that the integrable Hamiltonian system  $X = X_H$  (as well as its frequency map  $\omega$ ) is *Rüssmann nondegenerate* on  $A$ , if there exists a positive integer  $Q$  such that for each  $y \in A$ , the collection of all the  $\binom{n+Q}{n}$  partial derivatives

$$D^q \omega(y) = \frac{\partial^{q_1+\dots+q_n}}{\partial y_1^{q_1} \dots \partial y_n^{q_n}} \omega(y)$$

of the frequency map  $\omega : A \rightarrow \mathbb{R}^n$  at  $y$  of all the orders from 0 to  $Q$  spans  $\mathbb{R}^n$  (i.e. the linear hull of these derivatives is  $\mathbb{R}^n$ ). Roughly speaking, the manifold  $\omega(A) \subset \mathbb{R}^n$  winds and curves enough, to have a measure-theoretically large intersection with the Diophantine set  $D_{\tau,\gamma}(\mathbb{R}^n)$ . Such varieties are studied in the theory of Diophantine approximations



on submanifolds of Euclidean spaces, also known as Diophantine approximations of ‘dependent quantities’. It directly follows that Kolmogorov nondegenerate or iso-energetically nondegenerate systems are Rüssmann nondegenerate with  $Q = 1$ .

If  $A$  is connected and  $H$  is real analytic then Rüssmann nondegeneracy of  $X$  on  $A$  is ‘almost’ equivalent to saying that the set  $\omega(A)$  does not lie in any linear hyperplane in  $\mathbb{R}^n$  passing through the origin. To be more precise, the latter property of the frequency map  $\omega$  is equivalent to saying that  $X$  is Rüssmann nondegenerate on any open and bounded subset  $A' \subset A$  whose closure is contained in  $A$  (with the number  $Q$  possibly depending on  $A'$ ) [76,77,376,378,461].

Rüssmann’s theorem states that Rüssmann nondegeneracy of  $X$  implies the presence of many perturbed Diophantine quasi-periodic Lagrangean  $n$ -tori in any Hamiltonian system  $\tilde{X}$  sufficiently close to  $X$  (so that the union of these tori fills up positive measure). However, in the Rüssmann case, there is, in general, *no* connection between the unperturbed and the perturbed frequencies (see a detailed discussion in [25,77,397,408,409]), so we cannot speak of the persistence of the unperturbed tori  $\mathbb{T}^n \times \{y\}$ . Actually, Rüssmann nondegeneracy is much weaker than the other two nondegeneracy conditions we have considered. For instance, the image  $\omega(A)$  of the frequency map for a Rüssmann nondegenerate integrable Hamiltonian system can be a submanifold of  $\mathbb{R}^n$  of *any* positive dimension  $d \leq n$ , see [397,400,408] and [77, Example 4.7]. Finally, it is worthwhile to note that if  $\omega(A)$  is contained in some linear hyperplane in  $\mathbb{R}^n$  passing through the origin then *all* the invariant tori  $\mathbb{T}^n \times \{y\}$  can be destroyed by an arbitrarily small perturbation of  $X$  [77,397,408].

That Diophantine approximations on submanifolds of Euclidean spaces are required in many problems in mathematics and in mathematical physics was first pointed out by V.I. Arnold in 1968 in his lecture ‘Problems of Diophantine approximations in analysis’ at a symposium in the Russian city of Vladimir (see also a discussion in [19]). To the best of the authors’ knowledge, the first application of Diophantine approximations of dependent quantities in KAM Theory was due to I.O. Parasyuk [340] in 1982.

REMARKS. 1. The conclusions drawn in the example of two coupled oscillators of Section 3.2.2 also directly follow as an application of the iso-energetic KAM theorem.

2. In the case of Kolmogorov nondegeneracy, the Diophantine set is pulled back along the frequency map in a locally diffeomorphic way. Arnold’s iso-energetic nondegeneracy means that the energy hypersurfaces  $H^{-1}(c)$  are transversal to the Diophantine half-line ‘bundle’ of the  $n$ -dimensional analogue of Figure 3. With help of these ideas, a straightforward proof of the iso-energetic KAM theorem is possible from Theorem 6 [74,236]. The two nondegeneracy conditions—Kolmogorov and iso-energetic—are independent, as simple examples show, see e.g. [409] and [77, §4.2.3] (examples for  $n = 2$  are also given in [151,414]). In typical cases, however, *both* nondegeneracy conditions are satisfied, implying that the union of quasi-periodic Lagrangean invariant tori has positive measure in the phase space, in such a way that the conditional measure within the energy hypersurfaces is also positive.
3. It is also possible to derive Rüssmann’s KAM theorem from Theorem 6 (to be more precise, from its analogue admitting external parameters), considering the

geometry and number theory of the Diophantine set  $D_{\tau,\gamma}(\mathbb{R}^n)$  in more detail [397, 400], compare with Section 8.3.2 below.

4. All the three nondegeneracy conditions above are open in the  $C^k$ -topology for  $k$  sufficiently large, including  $k = \infty$ , compare with Section 4.1.2. The same holds for the compact-open topology in the real analytic case. We like to add here that, by the Analytic Unicity Theorem, in real analytic systems certain nondegeneracies are easily seen to be satisfied almost everywhere.

**Normal ‘triviality’** In Sections 4.2.1 and 4.3.1, we dealt with integrable systems, where the invariant  $n$ -tori are normally ‘trivial’ (or even normally ‘irrelevant’), namely within centre manifolds [77,78,236], and with the ‘classical’ case of Lagrangean tori in Hamiltonian systems [10,14,29,40,118,119,196,251,330,354,357,377,383,440]. A similar situation occurs in certain reversible systems [18,26,31,75,236,317,319,323,389] or in the volume preserving case for codimension 1 tori [43,47,48,76–78,236,318].

#### 4.4. Applications of the Lagrangean KAM Theorem 6

Although pure Liouville integrability is quite degenerate, integrable systems occur a lot as approximations. We have already met the example of a Birkhoff normal form truncation near a nondegenerate elliptic equilibrium (see the beginning of Section 4.3.2). In fact, we have already mentioned several times that invariant Lagrangean tori with Diophantine quasi-periodic dynamics occur in Hamiltonian systems in a typical way. There are quite a few examples of classical, nearly integrable systems that have received a lot of attention in the literature [12,15,20,22,25,28,155,206,207,320,323,374,414], therefore we will be restrictive here.

**4.4.1. Applications in Classical, Quantum, and Statistical Mechanics.** We shall briefly deal with the stability problem of the Solar System, with the Anderson localization, and with the Ergodic Hypothesis.

**The Solar System** As a historical application of KAM Theory in Classical Mechanics, consider the Solar System, seen as a perturbation of the integrable system obtained when neglecting the interaction between the planets. If Theorem 6 applies, it would follow that the actual evolution has positive probability to be quasi-periodic, when assuming the initial conditions to have been chosen at random. In that case the Solar System would be called ‘stable’. Much has been said about this example [12,15,25,104–106,155,173,266,283,320,323,328,365,374], and here we just give a few remarks (see also Section 1.2).

Firstly the Solar System contains quite strong resonances [8,15,25], which necessitate a more suitable integrable approximation than the one described here. Secondly the interaction between the planets probably is far too strong for an actual application of Theorem 6 as a perturbation result. The third remark refers to recent numerical work by J. Laskar [265], which seems to show that the Solar System is entirely chaotic, mankind just does not exist long enough to have noticed. . .

**Quantum Mechanics** Other applications of the KAM techniques occur in Quantum Mechanics, in particular in the study of the so-called electron (Anderson) localization.

If one considers the Schrödinger equations with spatially ergodic potentials, a localized (non-conducting) state is an eigenfunction of the Schrödinger operator for some energy eigenvalue. Such a function will decay exponentially in space.

Since in the localized regime the Schrödinger operator typically has a dense point spectrum, one may develop a perturbation theory of the corresponding resolvent operator, which diverges at this dense set of eigenvalues. These problems are very similar to those in the KAM set-up, where the dense set of resonances gives rise to divergent perturbation expansions.

Various KAM-inspired proofs and analyses of localization have been proposed for typical realizations of random potentials in arbitrary dimensions for large interaction strengths or high energies, see e.g. [179], as well as for quasi-periodic potentials in one dimension, describing electrons in quasicrystals, see e.g. [81,156,240,268,324]. For further developments on spectral properties of the Schrödinger operators with periodic and quasi-periodic potentials also see [79,81,84,85,167,170,358,375]. It turns out that KAM Theory can be developed to show the existence of a Cantor spectrum. Moreover, applications of Singularity Theory give indications for a generic theory of gap-closing.

Yet another field of physics, which is notorious for divergent perturbation theory problems and where KAM-like ideas are starting to play a significant role, is Quantum Field Theory [46,166,174,187].

### Ergodic Hypothesis

Statistical Mechanics deals with particle systems that are large, often infinitely large. The taking of limits as the number of particles tends to infinity is a notoriously difficult subject. Here we discuss a few direct consequences of [Theorem 6](#) for many degrees of freedom. This discussion starts with Kolmogorov's papers [250,251], which we now present in a slightly rephrased form. First, we recall that for Hamiltonian systems (say, with  $n$  degrees of freedom), typically the union of Diophantine quasi-periodic Lagrangean invariant  $n$ -tori fills up positive measure in the phase space and also in the energy hypersurfaces. Second, such a collection of KAM tori immediately gives rise to non-ergodicity, since it clearly implies the existence of distinct invariant sets of positive measure. For background on Ergodic Theory, see e.g. [22,33,34,183,184,301]. Apparently the KAM tori form an 'obstruction' to ergodicity, and a question is how bad is this obstruction as  $n \rightarrow \infty$ ? To fix thoughts we give an example.

**EXAMPLE 7 (A Lattice System).** Consider the 1-dimensional lattice  $\mathbb{Z} \subset \mathbb{R}$ , at the vertices of which identical nonlinear oscillators are situated. For simplicity, think of the lattice being situated on a horizontal line, where at all the vertices identical pendula are suspended, subject to constant vertical gravity. Also we connect nearest neighbour oscillators by weak springs, the spring constants can either be the same for all the oscillator pairs or decay at infinity.

Let  $\Lambda_N \subset \mathbb{Z}$  be the box with vertices in the interval  $[-N, N]$ . Then, for  $M \leq N$  consider two of these boxes  $\Lambda_M \subseteq \Lambda_N$ . We 'truncate' the infinite system by ignoring all the pendula outside the larger box  $\Lambda_N$ .

First consider the integrable system associated to  $\Lambda_N$ , where all interactions are neglected. Suppose that the oscillators situated at the vertices in  $\Lambda_M$  are in motion, while

the others are at rest. In the phase space this corresponds to an invariant  $(2M + 1)$ -torus, which is normally elliptic (see Sections 8.2 and 8.3 below for a rigorous definition of normally elliptic invariant tori in Hamiltonian systems). Moreover, the normal frequencies of this torus are in  $1:1:\dots:1$  resonance.

Then we ‘turn on’ the activity of the interaction springs. A suitable adaptation [145] of Theorem 6 and of the results of [355] for this case yields the persistence of these elliptic tori for small values of the spring constants. The corresponding kind of motion is a ‘quasi-periodic breather’; for a similar type of motion in the periodic case see [297]. The union of elliptic tori has positive  $2(2M + 1)$ -dimensional Hausdorff measure in the phase space.

The question now is what are the asymptotics of the density of this measure as  $N$  (and  $M$ )  $\rightarrow \infty$ . A partial answer to this question [145] says that this density decays at least exponentially fast in  $N$ , while there is a further at least polynomial decay in  $M$ .

What conclusions can be drawn from this example? Although KAM Theory gives typical counterexamples to the Ergodic Hypothesis, it seems that the corresponding ‘obstruction’ to ergodicity is not too bad as the size of the system tends to infinity. This is in the same spirit as an earlier result by Arnold [16].

As we have already pointed out, taking the limit as  $N \rightarrow \infty$  is an extremely difficult problem. Another question is what happens if the limit  $N \rightarrow \infty$  is really attained? The KAM Theory for infinite systems is fully in development (see [6,41,245,256–258,321,355] and references therein), but infinite lattice systems still have many secrets, compare with e.g. [362].

**4.4.2. Discussion.** We conclude this section with a few remarks on the general dynamics in a neighbourhood of Hamiltonian KAM tori. In particular this concerns so-called ‘superexponential stickiness’ of the KAM tori and adiabatic stability of the action variables, involving the so-called Nekhoroshev estimate.

To begin with, emphasize the following difference between the cases  $n = 2$  and  $n \geq 3$  in Theorem 6. For  $n = 2$  the level surfaces of the Hamiltonian are three-dimensional, while the Lagrangean tori have dimension two and hence codimension one in the energy hypersurfaces. This means that for open sets of initial conditions, the evolution curves are forever trapped in between KAM tori, as these tori foliate over nowhere dense sets. This implies perpetual adiabatic stability of the action variables. In contrast, for  $n \geq 3$  the Lagrangean tori have codimension  $n - 1 > 1$  in the energy hypersurfaces, and evolution curves may escape.

**REMARK.** This actually occurs in the case of so-called *Arnold diffusion*. The literature on Arnold diffusion is immense, and we here just quote [16,22,116,123,125,150,193,283,285,286,302,303,307,328,386,432] for results, details, and references.

Next we consider the motion in a neighbourhood of the KAM tori, in the case where the systems are real analytic or at least Gevrey smooth.

First we mention that, measured in terms of the distance to the KAM torus, nearby evolution curves generically stay nearby over a *superexponentially* long time [314]. This property often is referred to as ‘superexponential stickiness’ of the KAM tori, see Section 8.1 below for more details.

Second, nearly integrable Hamiltonian systems, in terms of the perturbation size, generically exhibit *exponentially* long adiabatic stability of the action variables, see e.g. [19,77,151,194,197,282–284,286–288,302,313,315,326,328,329,336,356,386]. This property is referred to as the *Nekhoroshev estimate* or the *Nekhoroshev theorem*. For related work on perturbations of so-called superintegrable systems, see [171]. The exponential stability of elliptic equilibria has been studied in e.g. [172,195,334].

## 5. Further developments in KAM Theory

In this section we deal with Parametrized KAM Theory as this was initiated by J.K. Moser in the 1960s. This theory was set up in a unifying Lie algebra format, thereby covering many classes of dynamical systems characterized by preservation of a certain structure, like the general ‘dissipative’ case, Hamiltonian and volume preserving cases, etc.

### 5.1. Background

We here discuss uniqueness of KAM tori, heavily using Whitney differentiability. Also we present a few elements of the Paley–Wiener theory on Fourier series, which are fundamental for the background mathematics.

**5.1.1. Unicity of KAM tori.** The ‘classical’ KAM Theorem 6 establishes persistence of invariant Lagrangean tori in nearly integrable Hamiltonian systems. These tori are quasi-periodic with Diophantine frequency vectors and their union is a nowhere dense set of positive measure in the phase space. It is a long standing question of how far the perturbed tori are unique. Using the fact that at the level of the tori, there exists a Whitney smooth conjugation between the integrable approximation and its perturbation, this unicity follows on a closed subset of the Diophantine torus union of full measure [91] (see also [383]). We first introduce this subset  $D_{\tau,\gamma}^d(\mathbb{R}^n) \subseteq D_{\tau,\gamma}(\mathbb{R}^n)$ .

To explain this, in general let  $K \subseteq \mathbb{R}^n$  be a closed set. We say that  $a \in K$  is a *density point* precisely if any  $C^\infty$ -function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ , such that  $F|_K \equiv 0$ , has an infinite-jet  $j^\infty(F)(a) = 0$ . The set of all density points of  $K$  is denoted by  $K^d$ . Moreover, in general we say that the closed set  $K \subseteq \mathbb{R}^n$  possesses the closed half-line property if the following holds: whenever  $p \in K$  and  $s \geq 1$ , then also  $sp \in K$ .

LEMMA 8 (Properties of  $K^d$  [91]). *Let  $K \subseteq \mathbb{R}^n$  be a closed set. Then*

1.  $K^d \subseteq K$  is a closed set;
2.  $K \setminus K^d$  has Lebesgue measure zero;
3. If  $K$  possesses the closed half-line property, then so does  $K^d$ .

The proof is rather direct, item 2 using the Fubini theorem. Applying this construction to  $K = D_{\tau,\gamma}(\mathbb{R}^n)$  gives the subset  $D_{\tau,\gamma}^d(\mathbb{R}^n) \subseteq D_{\tau,\gamma}(\mathbb{R}^n)$ . Recall from Section 4.1.2 that  $D_{\tau,\gamma}(\mathbb{R}^n)$  and hence  $D_{\tau,\gamma}^d(\mathbb{R}^n)$  possess the closed half-line property.

**THEOREM 9** (Unicity of KAM Tori [91]). *Assuming the set-up of the ‘classical’ KAM Theorem 6, there exist  $C^\infty$ -neighbourhoods  $\mathcal{U}_2$  of  $X_H$  and  $\mathcal{V}_2$  of the identity map  $\text{Id}_M$ , such that the following holds. If  $\Phi \in \mathcal{V}_2$ , restricted to  $\mathbb{T}^n \times D_{\tau,\gamma}(A_\gamma)$ , is a conjugation between the vector fields  $X_H$  and  $X_{H+\tilde{F}} \in \mathcal{U}_2$ , then the further restriction of  $\Phi$ , to  $\mathbb{T}^n \times D_{\tau,\gamma}^d(A_\gamma)$ , is unique up to a torus translation.*

**REMARKS.** 1. We note that the above definition of a density point does not coincide with that of a Lebesgue density point. A general problem is to characterize  $D_{\tau,\gamma}^d(\mathbb{R}^n) \subseteq D_{\tau,\gamma}(\mathbb{R}^n)$ , compare with Section 4.2.1.

2. Again the somewhat smaller closed half-line structure of  $X_H$  is inherited by the perturbation  $X_{H+\tilde{F}}$ , up to a diffeomorphism.

3. It is conjectured [91] that Theorem 9 generalizes to certain other KAM theorems as well, e.g. to Theorem 3.

**5.1.2. Paley–Wiener estimates and Diophantine frequencies.** As before, let  $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$  be the standard  $n$ -torus with coordinates  $x_1, \dots, x_n$  counted modulo  $2\pi\mathbb{Z}$ . We consider functions  $h : \mathbb{T}^n \rightarrow \mathbb{R}$  of class  $C^r$ , for  $r \geq 0$ . One of the main tools in the proofs of the KAM theorems is the solution of linear (or homological) partial differential equations of the form

$$\sum_{j=1}^n \omega_j \frac{\partial h(x)}{\partial x_j} = H(x), \quad (22)$$

where  $H : \mathbb{T}^n \rightarrow \mathbb{R}$  is given with  $\mathbb{T}^n$ -average 0. These equations have to be solved for  $h$ , also compare with Section 3.1.1. In the proofs of all the KAM theorems for flows, as discussed here, this linear problem is central in a Newtonian iteration process that solves a nonlinear conjugation equation under Diophantine conditions. We now discuss this problem in terms of Fourier series, for details refer to e.g. [77,91,132,354].

For  $k \in \mathbb{Z}^n$  the  $k$ th Fourier coefficient of  $h$  is given by

$$h_k = \frac{1}{(2\pi)^n} \oint_{\mathbb{T}^n} e^{-i\langle k, x \rangle} h(x) \, dx, \quad (23)$$

and we consider the formal Fourier series

$$h(x) \doteq \sum_{k \in \mathbb{Z}^n} h_k e^{i\langle k, x \rangle}.$$

We have the following familiar norms in terms of the Fourier coefficients:

$$\begin{aligned} \|h\|_\infty &= \max_k |h_k|; \\ \|h\|_2 &= \left( \sum_k |h_k|^2 \right)^{1/2}; \end{aligned}$$

$$\|h\|_1 = \sum_k |h_k|.$$

For any (continuous) function  $h : \mathbb{T}^n \rightarrow \mathbb{R}$  we recall that

$$\|h\|_\infty \leq \|h\|_2 \leq \|h\|_0 \leq \|h\|_1. \quad (24)$$

Since we did not require the Fourier series to converge, some of these norms may be infinite. For any continuous function  $h : \mathbb{T}^n \rightarrow \mathbb{R}$  we decompose

$$h = h_0 + \tilde{h},$$

with  $h_0$  as in (23), i.e. with  $h_0$  equal to the  $\mathbb{T}^n$ -average of  $h$ . We call  $\tilde{h}$  the *variable part* of  $h$ .

LEMMA 10 (Paley–Wiener Estimates). *Let  $h : \mathbb{T}^n \rightarrow \mathbb{R}$  be of class  $C^r$ , with variable part  $\tilde{h}$ .*

1. *Then there exists a positive constant  $C_{r,n}^{(1)}$  such that for all  $k \in \mathbb{Z}^n \setminus \{0\}$*

$$|k|^r |h_k| \leq C_{r,n}^{(1)} \|\tilde{h}\|_r.$$

2. *Moreover, for  $r \geq n + 1$  there exists a positive constant  $C_{r,n}^{(2)}$  such that*

$$\|\tilde{h}\|_0 \leq C_{r,n}^{(2)} \max_{k \neq 0} |k|^r |h_k|.$$

REMARKS. 1. We mention that the first item of Lemma 10 is the familiar Paley–Wiener decay for the Fourier coefficients of a  $C^r$ -function, which directly follows by partial integration. For the second item of the lemma, it suffices to take  $C_{r,n}^{(2)} = \sum_{k \neq 0} 1/|k|^r$ .

2. We conclude that if  $h$  is of class  $C^\infty$ , its Fourier coefficients decay faster than any positive power of  $|k|^{-1}$ . Similarly, when  $h$  is real analytic this decay is exponentially fast.

LEMMA 11 (Small Divisors). *Assume that  $h : \mathbb{T}^n \rightarrow \mathbb{R}$  satisfies the differential equation (22) for a given  $H : \mathbb{T}^n \rightarrow \mathbb{R}$  of class  $C^r$  with  $\mathbb{T}^n$ -average 0, where  $\omega \in D_{\tau,\gamma}(\mathbb{R}^n)$ , see (18). Then:*

1. *For all  $k \in \mathbb{Z}^n \setminus \{0\}$  one has that*

$$|h_k| \leq \frac{|k|^\tau}{\gamma} |H_k|.$$

2. *Moreover, for  $r \geq n + p + \tau$  with  $p \in \mathbb{N}$  one has that  $h \in C^{p-1}(\mathbb{T}^n, \mathbb{R})$  with*

$$\|\tilde{h}\|_{p-1} \leq \frac{C_{r,n,p,\tau}^{(3)}}{\gamma} \|H\|_r,$$

*for a positive constant  $C_{r,n,p,\tau}^{(3)}$ .*

REMARKS. 1. The first estimate follows by comparing terms in the Fourier series, using the Diophantine condition (18). To illustrate the next estimate we consider the case  $p = 1$ . It follows from the last inequality in (24) and Lemma 10, that for the  $C^0$ -norm of  $h$

$$\begin{aligned} \|\tilde{h}\|_0 &\leq \|\tilde{h}\|_1 \leq \sum_{k \neq 0} \frac{|k|^\tau}{\gamma} |H_k| \\ &\leq \sum_{k \neq 0} \frac{|k|^\tau}{\gamma} |k|^{-r} C_{r,n}^{(1)} \|H\|_r \leq \frac{C}{\gamma} \sum_{\ell \geq 1} \ell^{\tau-r} \|H\|_r \ell^{n-1} \\ &= C \frac{\|H\|_r}{\gamma} \sum_{\ell \geq 1} \ell^{n+\tau-r-1} = C' \frac{\|H\|_r}{\gamma} \end{aligned}$$

( $\ell = |k|$ ), where we have used the fact that  $n + \tau - r - 1 \leq -2$ , which gives convergence of the last sum.

2. We conclude that if  $H$  is of class  $C^\infty$ , then so is  $h$ . Similarly whenever  $H$  is real analytic, then so is  $h$ .

## 5.2. Parametrized KAM Theory

Already in Sections 3.1 and 4.2 we met parameters in KAM Theory, which were needed to get persistence of quasi-periodic invariant tori of the integrable approximation. Below, we shall generalize this approach by developing Parametrized KAM Theory in a more systematic way and taking the normal linear dynamics into account.

In [318,319], J.K. Moser presented two directions in which one may generalize the set-up of Section 4, that are logically connected. Firstly, it turns out that KAM Theorem 6 can be directly carried over to the reversible setting, to the general dissipative setting, etc., compare with Theorem 5 and see [78,236,323,389]. These cases are characterized by the fact that they are normally ‘trivial’ or ‘irrelevant’, see the end of Section 4.3.2. Secondly, the possibly nontrivial normal behaviour of the invariant tori can be taken into account within a ‘modifying term’ formalism. This means that, in order to obtain a conjugation between the unperturbed and the perturbed tori, the system has to be changed ‘at lowest order’, a suitable nondegeneracy condition having to be fulfilled. This result gives e.g. an alternative approach to KAM Theory of lower dimensional isotropic invariant tori in Hamiltonian systems (see Section 8.3). Here, apart from the internal frequencies of the quasi-periodic tori, also the normal frequencies play a role and enter the Diophantine conditions. The latter procedure turns out to allow also for a further generalization of the non-Hamiltonian settings mentioned above. In fact, there exists a quite general formulation of this theory in terms of Lie algebras of vector fields, encompassing the Hamiltonian, volume preserving, and several equivariant cases, etc., as well as the reversible set-up. For an axiomatic approach to these ‘admissible’ Lie algebras, see [71,78,227,236,319], the reversible counterpart is treated in [58,75,130].

REMARK. Compare this with the Formal Normal Form Theory at equilibria, periodic solutions, and quasi-periodic tori as developed in a structure preserving Lie algebra formalism [47,50], that similarly also covers the dissipative normal forms [19,31,98–100,



127,129,202,260,454], the Hamiltonian (Birkhoff) normal forms [20,25,34,98,99,127,129,203,320,379,380,414], etc.

This line of thought was picked up in [76–78,236] as follows. Instead of considering a given system that has to be modified *later*, the theory *starts off* with families of systems, i.e. with the set-up where the systems depend on (external) parameters. This determines a fixed universe of parametrized systems that can be either integrable or not. Generally integrability is defined as invariance under the natural action of a torus group, where the invariant tori in question are orbits of the action, compare with Section 4.2.1. It turns out that Moser’s nondegeneracy condition now translates to (trans)versality of the integrable family, to be called the BHT (Broer–Huitema–Takens) nondegeneracy. Moreover, perturbations of this family contain the perturbed tori, for which the connecting, parameter dependent conjugation, preserves not only the internal frequencies of the tori, but their entire normal linear part as well. Furthermore, the (trans)versality of the integrable family also affects this normal linear part, which now takes, to a large extent, the role of the ‘lowest order’ terms mentioned above.

**5.2.1. The parametrized dissipative KAM theorem.** For simplicity we consider in this section the following setting. As the phase space we take  $M = \mathbb{T}^n \times \mathbb{R}^m = \{(x, y)\}$ , and as the parameter space  $P \subseteq \mathbb{R}^s = \{\mu\}$ , an open subset. The starting point is an integrable  $C^\infty$ -family  $X = X_\mu(x, y)$  of vector fields, given by

$$\begin{aligned}\dot{x} &= \omega(\mu) + O(|y|) \\ \dot{y} &= \Omega(\mu)y + O(|y|^2),\end{aligned}\tag{25}$$

where the  $O$ -estimates are (locally) uniform in  $\mu$ . Here  $\omega(\mu) \in \mathbb{R}^n$  and  $\Omega(\mu) \in gl(m, \mathbb{R})$ , for each value of  $\mu$ . Integrability of  $X$  again amounts to  $x$ -independence, compare with Section 4.2.1. The interest is with persistence properties of the family  $T_\mu = \mathbb{T}^n \times \{0\}$  of invariant  $n$ -tori with parallel dynamics. The local nondegeneracy condition roughly means that the pair of maps consisting of the internal frequency vector  $\omega(\mu)$  and the ‘normal’ matrix  $\Omega(\mu)$ , near  $\mu = \mu_0$  has to vary sufficiently with  $\mu \in P$ . In particular, the map  $\mu \mapsto \omega(\mu)$  is a submersion as in Section 4.2.1, while at the same time the map  $\mu \mapsto \Omega(\mu)$  is a *versal unfolding* of  $\Omega(\mu_0)$ , so it is transversal to the orbit of  $\Omega(\mu_0)$  under the adjoint action of  $GL(m, \mathbb{R})$ . Note that versal unfoldings with a minimal number of parameters are said to be *miniversal*, compare with [17,19], also see e.g. [192]. In coordinates the normal linear part of (25) is given by

$$\omega(\mu) \frac{\partial}{\partial x} + \Omega(\mu)y \frac{\partial}{\partial y}.\tag{26}$$

For a coordinate-free definition of the normal linear part of integrable systems at invariant tori, using the normal bundle of the tori, see [78, I Section 2] and [236, Section 2].

**BHT nondegeneracy** We shall confine ourselves to the case where  $\Omega$  has only simple eigenvalues. Suppose that these eigenvalues are given by

$$(\delta_1, \dots, \delta_{N_1}, \alpha_1 \pm i\beta_1, \dots, \alpha_{N_2} \pm i\beta_{N_2})\tag{27}$$

with  $\beta_j > 0$  for  $1 \leq j \leq N_2$ . Note that  $N_1 + 2N_2 = m$ . We call  $\beta = (\beta_1, \beta_2, \dots, \beta_{N_2})$  the *normal frequencies* of the invariant torus. Next consider the map

$$\text{spec} : gl(m, \mathbb{R}) \rightarrow \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \times \mathbb{R}^{N_2}; \quad \Omega \mapsto (\delta, \alpha, \beta), \quad (28)$$

which, by the simpleness of the eigenvalues, parametrizes the  $GL(m, \mathbb{R})$  orbit space near  $\Omega(\mu_0)$ . The BHT nondegeneracy condition in this case boils down to saying that at  $\mu = \mu_0$  the map

$$P \ni \mu \mapsto (\omega \times (\text{spec} \circ \Omega))(\mu) \in \mathbb{R}^n \times \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \times \mathbb{R}^{N_2} \quad (29)$$

is a submersion [76–78,236]. As in Section 4.2.1, we can use the Inverse Function Theorem [225,325,418] to re-parametrize

$$\mu \leftrightarrow (\omega, \delta, \alpha, \beta)$$

on an open subset  $A \subseteq P$ , suppressing, for simplicity, extra parameters that may possibly occur. We then say that the family  $X_\mu$  is nondegenerate on the torus union  $\mathbb{T}^n \times \{0\} \times A = \bigcup_{\mu \in A} T_\mu$ , recall that here  $0 \in \mathbb{R}^m$ .

**Diophantine conditions** Now we need Diophantine conditions on the internal and normal frequencies. For  $\tau > n - 1$  and  $\gamma > 0$  define the set of  $(\tau, \gamma)$ -Diophantine normal-internal frequency vectors by

$$\begin{aligned} D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2}) = \left\{ (\omega, \beta) \in \mathbb{R}^n \times \mathbb{R}^{N_2} \mid |\langle \omega, k \rangle + \langle \beta, \ell \rangle| \geq \gamma |k|^{-\tau}, \right. \\ \left. \text{for all } k \in \mathbb{Z}^n \setminus \{0\} \text{ and for all } \ell \in \mathbb{Z}^{N_2} \text{ with } |\ell| \leq 2 \right\}, \quad (30) \end{aligned}$$

compare with (18). This set is again a nowhere dense set of positive measure (for  $\gamma$  sufficiently small) with the closed half-line property, see Sections 4.1.2 and 5.1.1; compare with [76–78,236,318,319]. Defining  $\Gamma = (\omega \times (\text{spec} \circ \Omega))(A)$ , without any restriction we may assume that  $\Gamma$  has the ‘product form’  $\Gamma = \Gamma_\omega \times \Gamma_\delta \times \Gamma_\alpha \times \Gamma_\beta$ . Furthermore we define the shrunken version

$$\Gamma_\gamma = \{(\omega, \delta, \alpha, \beta) \in \Gamma \mid \text{dist}((\omega, \delta, \alpha, \beta), \partial\Gamma) \geq \gamma\}$$

of  $\Gamma$  as well as

$$D_{\tau, \gamma}(\Gamma_\gamma) = \Gamma_\gamma \cap \left( D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2}) \times \Gamma_\delta \times \Gamma_\alpha \right)$$

and  $D_{\tau, \gamma}(A_\gamma) \subset A$ , compare with Sections 4.2.1 and 4.3.1. Note that the closed half-lines of  $D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2})$  now turn into closed linear half-spaces of dimension  $1 + N_1 + N_2$ . Again these geometrical structures, up to a diffeomorphism, are inherited by the perturbations. This is a consequence of the following theorem.

**THEOREM 12** (Parametrized KAM – Dissipative Case [71,78,132,227,236]). *Let  $n \geq 2$ . Let the integrable  $C^\infty$ -family  $X = X_\mu(x, y)$  of vector fields (25) be BHT nondegenerate on*

$\mathbb{T}^n \times \{0\} \times A$ , with  $A \subseteq P$  open. Also assume that for  $\mu \in A$  we have  $\det \Omega(\mu) \neq 0$ . Then, for  $\gamma > 0$  sufficiently small, there exists a neighbourhood  $\mathcal{O}$  of  $X$  in the  $C^\infty$ -topology, such that for any perturbed (not necessarily integrable) family  $\tilde{X} \in \mathcal{O}$  there exists a  $C^\infty$ -mapping  $\Phi : \mathbb{T}^n \times \mathbb{R}^m \times A \rightarrow \mathbb{T}^n \times \mathbb{R}^m \times A$ , defined near  $\mathbb{T}^n \times \{0\} \times A$ , with the following properties:

1.  $\Phi$  is a  $C^\infty$ -near the identity map preserving projections to the parameter space  $P$ .
2. The image of the  $X$ -invariant torus union  $\mathbb{T}^n \times \{0\} \times D_{\tau,\gamma}(A_\gamma)$  under  $\Phi$  is  $\tilde{X}$ -invariant, and the restricted map  $\hat{\Phi} = \Phi|_{\mathbb{T}^n \times \{0\} \times D_{\tau,\gamma}(A_\gamma)}$  conjugates  $X$  to  $\tilde{X}$ , that is,

$$\hat{\Phi}_* X = \tilde{X}.$$

3.  $\Phi$  preserves the normal linear behaviour of the tori  $T_\mu$  for  $\mu \in D_{\tau,\gamma}(A_\gamma)$  with respect to  $X$ . This means the following. Let

$$\Phi(x, y, \mu) = (\Psi[\mu](x, y), \Upsilon[\mu]),$$

where  $\Psi[\mu](x, y) \in \mathbb{T}^n \times \mathbb{R}^m$  and  $\Upsilon[\mu] \in P$  (the fact that the  $\mu$ -component  $\Upsilon[\mu]$  of  $\Phi$  does not depend on the phase space variables  $x$  and  $y$  just expresses the preservation of projections to the  $\mu$ -space). Then for each  $\mu \in D_{\tau,\gamma}(A_\gamma)$  the vector field

$$(\Psi[\mu])_*^{-1} \tilde{X}_{\Upsilon[\mu]}$$

is given by (25), where now the  $O$ -terms are, generally speaking,  $x$ -dependent.

REMARKS. 1. The conclusion of Theorem 12 first of all expresses that the family  $X$  is quasi-periodically stable on the union

$$\bigcup_{\mu \in D_{\tau,\gamma}(A_\gamma)} T_\mu = \mathbb{T}^n \times \{0\} \times D_{\tau,\gamma}(A_\gamma)$$

of Diophantine quasi-periodic invariant  $n$ -tori [76–78, 236], compare with Sections 3.1.3, 4.2.1 and 4.3.1 above. Including item 3 of Theorem 12, we also speak of *normal linear stability* of this torus union.

2. Theorem 12 was also stated and proven in the more general setting with preservation of certain structures. This includes, next to the general dissipative case of Theorem 12 itself, the Hamiltonian and the volume preserving cases, as well as certain equivariant and reversible cases, all of these with external parameters if necessary. In fact, the parametrized KAM theorem was stated and proven for certain ‘admissible’ Lie algebras of vector fields and for reversible analogues of such algebras, as we already mentioned at the beginning of Section 5.2; for an axiomatic approach see [58, 71, 75, 78, 130, 227, 236, 319]. In many set-ups, in the definition of the map spec we have to refrain from counting double. This refers, for instance, to the Hamiltonian and reversible settings, where complex eigenvalues can show up in quadruplets while non-zero real eigenvalues come in pairs (see Section 8.3.1). For an elaborate discussion see [78, I Section 2], [236, Section 2], and [75]. The

Hamiltonian counterpart of [Theorem 12](#) will be presented in [Section 8.3.1](#) below, see [Theorem 17](#).

3. The assumption that  $\Omega(\mu)$  has only simple eigenvalues for each  $\mu$  can be dropped, even in the full generality of the structure preserving settings mentioned in the previous item. Here we use the fact that Arnold's theory of matrices depending on parameters [[17,19](#)] can be carried over to a large class of 'admissible' Lie subalgebras of  $gl(m, \mathbb{R})$  under the adjoint action of the corresponding subgroups of  $GL(m, \mathbb{R})$ , and to reversible analogues of such actions. In this general case versal unfoldings can be normalized to the so-called *linear centralizer unfolding*; for definitions and other mathematical details see e.g. [[17,19,127,180,192,249,349](#)]. Note that this definition encompasses the above case of simple eigenvalues. Also compare with [[1,339](#)] concerning genericity in terms of transversality. The corresponding Parametrized KAM Theory has been worked out in detail in [[57,58,63,71,130,131,227](#)] for Hamiltonian and reversible variants of the normal 1:–1 resonance.
4. In the above set-up the condition that  $\det \Omega(\mu) \neq 0$  cannot be omitted, although it was not needed in [[318,319](#)]. This problem has been overcome by Wagener [[445,446](#)].

**5.2.2. Direct consequences of the parametrized approach.** Partly recalling [[76–78,236](#)], we briefly discuss here a number of straightforward consequences of the parametrized approach. First of all, the normally 'trivial' cases (see the end of [Section 4.3.2](#)) without external parameters, like the case of Lagrangean tori in the Hamiltonian setting ([Theorem 6](#)), the case of codimension 1 tori in the volume preserving setting, as well as the 'standard' reversible set-up [[18,26,31,236,319,323,389](#)], directly follow using so-called *localization*, see [[78, I Section 5](#)] and [[236, Section 5](#)] (compare also with Remark 2 after [Theorem 17](#) below). This means that by introducing an extra local multi-parameter  $\mu$  we end up examining a family  $T_\mu$  of invariant tori as before, where we have to consider *only one torus* for each value of  $\mu$ . Moreover, since  $\Omega(\mu) \equiv 0$ , we only need the constant unfolding in this direction. After application of the corresponding version of Parametrized KAM Theory, we can project back to obtain a persistence result without parameters. In particular this holds for [Theorem 6](#), which was actually also proven in this way [[319,357](#)]. Also the centre manifold situation of [Theorem 5](#) can be seen as a particular case of parametrized KAM [Theorem 12](#); for a more elaborate discussion see [[78, I Section 7](#)] and [[236, Section 7](#)].

- REMARKS.
1. As explained in [[78, I Section 7](#)] and [[236, Section 7](#)], all cases with varying frequency ratios  $[\omega_1:\omega_2:\dots:\omega_n] \in \mathbb{P}^{n-1}(\mathbb{R})$  fall under this approach, leading to *weak quasi-periodic stability*. Indeed, a scaling of the time gives an extra parameter, after which Parametrized KAM Theory applies. After projecting to the original setting, the conjugation  $\Phi$  turns into an equivalence, for definitions see e.g. [[339](#)].
  2. A similar discussion applies to the Hamiltonian iso-energetic setting [[20,25,74,151,236,408,409,431](#)], for a geometric discussion see [Section 4.3.2](#) above.
  3. The set-up of Parametrized KAM Theory is particularly suitable for studying quasi-periodic bifurcations. In the next two sections we shall come back to this subject.

**Rüssmann nondegeneracy revisited** Preservation of the normal linear behaviour surely may require a lot of external parameters, and a clear aim always is to get rid of as many parameters as possible, compare with [78, I Section 7c] and [236, Section 7c], using the geometry of the Diophantine sets (see the previous discussion as an example). In fact, [Theorem 12](#) and all its structure preserving analogues possess ‘miniparameter’ versions [76,77] with Rüssmann-like nondegeneracy conditions on the unperturbed frequency maps. These conditions are slight generalizations of the Rüssmann nondegeneracy condition of [Section 4.3.2](#) for the case where the frequency map depends on external parameters. The ‘miniparameter’ KAM theorems can be most easily obtained using the so-called Herman method [76,77,397,399,400,405,409–411]. In [Section 8.3.2](#), we shall illustrate Herman’s approach for the Hamiltonian counterpart of [Theorem 12](#).

**5.2.3. Reducibility issues.** The present setting assumes that the integrable family  $X$  has the form (25) on the phase space  $M = \mathbb{T}^n \times \mathbb{R}^m$ , while its normal linear part has the form (26). This means that the system is in Floquet form (the coefficients of the variational equation along each invariant torus do not depend on the point on this torus). Of course, the results also hold for all the cases reducible to this form.

First of all we observe that it is an immediate consequence of parametrized KAM [Theorem 12](#), that reducibility is a persistent property on Diophantine sets of parameters, compare with [78, I Section 7] and [236, Section 7].

However, it is known that this reducibility is not always possible, see e.g. [164,167,217,219,252,253]. In [89,90,96,427,442], also compare with [132], the skew Hopf bifurcation was treated as a first example of non-reducible KAM Theory. For more details see [Section 6.2.3](#) below. Concerning non-reducibility in the Hamiltonian lower dimensional context, see [Section 8.3.3](#) below.

## 6. Quasi-periodic bifurcations: dissipative setting

The bifurcation theory of equilibrium points is widely developed in the general dissipative setting, see e.g. [19,21,127,202,260,332,454]. As generic codimension 1 bifurcations, we mention the saddle-node (or fold) bifurcation and the Hopf bifurcation [228] (also called the Poincaré–Andronov phenomenon [19]). This theory has a direct extension to bifurcations of periodic solutions. The Hopf bifurcation then translates to the Hopf–Neĭmark–Sacker bifurcation which, by the Poincaré map, relates to the Hopf bifurcation for fixed points of diffeomorphisms [425], for a discussion also see [132]. Here normal-internal resonances already play a role in the interaction between the (internal) frequency of the periodic solution and the normal frequency, which gives rise to a pattern of Arnold resonance tongues comparable to [Figure 2](#). In order to see this pattern, one needs to keep track of the normal frequency, which requires more parameters. It turns out that the Arnold family of circle maps (11), mentioned in [Section 3.1.3](#), to some extent is a good model for this Hopf–Neĭmark–Sacker scenario. For periodic solutions also the period doubling bifurcation occurs as a generic codimension 1 bifurcation. We note that saddle-node, period doubling, and Hopf bifurcations all are related to loss of (normal) hyperbolicity. Presently the interest is with the analogue of this theory in the case of quasi-periodic tori; compare with [44,49,52,77,78,132].

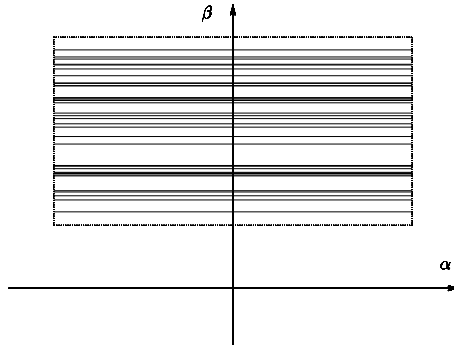


Fig. 4. Parametrized section of the Diophantine set  $D_{\tau, \gamma}(\Gamma_\gamma)$ , where  $\omega \in \mathbb{R}^n$  has a fixed Diophantine value. The Hopf line  $\alpha = 0$  is Cantorized by normal-internal resonances.

To study this we return to the phase space  $M = \mathbb{T}^n \times \mathbb{R}^m = \{(x \pmod{2\pi}, y)\}$ , where we consider  $C^\infty$ -families  $X$  of vector fields of the general form

$$X_\mu(x, y) = F_\mu(x, y) \frac{\partial}{\partial x} + G_\mu(x, y) \frac{\partial}{\partial y}$$

( $\mu$  being a suitable multi-parameter). As before, integrability translates to a  $\mathbb{T}^n$ -symmetry, which simply means that the coefficients  $F_\mu$  and  $G_\mu$  are  $x$ -independent.

If  $X$  is such an integrable family, then, by dividing out the  $\mathbb{T}^n$ -symmetry, we can reduce it to a  $C^\infty$ -family  $X_{\text{red}, \mu}$  of vector fields

$$X_{\text{red}, \mu}(y) = G_\mu(y) \frac{\partial}{\partial y}$$

on  $\mathbb{R}^m$ . Notice that equilibria of  $X_{\text{red}}$  correspond to invariant  $n$ -tori of the integrable family  $X$ ; this is why these equilibria are called *relative*. Similar remarks go for (relative) periodic solutions of  $X_{\text{red}}$ . In the reduced family  $X_{\text{red}}$  we can meet the bifurcations described above. For the integrable family  $X$  this gives a direct translation in terms of torus bifurcations.

The problem addressed in this section is what happens to the integrable bifurcations when perturbing to nearly integrable families  $\tilde{X} = \tilde{X}_\mu(x, y)$ ? It turns out [78] that all the three cases (saddle-node, period doubling, and Hopf) lead to typical (i.e.  $C^\infty$ -open) *quasi-periodic bifurcation* scenarios. This results from a combination of KAM Theory and Bifurcation Theory (and Singularity Theory) as this goes back to Whitney, Thom, Mather, and Arnold, see e.g. [19, 21, 23, 24, 428]. Generally speaking, it turns out that the (real) semi-algebraic stratifications that occur as bifurcation sets in the product of the phase space and the parameter space, are ‘Cantorized’ in a systematic way, compare with Figure 4.

The KAM Theory of quasi-periodic versions of period doubling and Hopf bifurcations is a direct application of Theorem 12, while the quasi-periodic saddle-node bifurcation is more involved, see [78, II Section 5] and [446]. An extension regarding the so-called quasi-periodic  $d$ -fold degenerate bifurcation, based on the so-called ‘translated torus’ theorem [216, 218, 373, 463], was carried out in [445].

We shall illustrate the general approach on the quasi-periodic Hopf bifurcation, compare with [44,51,54,77,78,132].

### 6.1. Quasi-periodic Hopf bifurcation

The unperturbed, integrable  $C^\infty$ -family  $X = X_\mu(x, y)$  on  $\mathbb{T}^n \times \mathbb{R}^2$  has the form (25)

$$X_\mu(x, y) = [\omega(\mu) + f(y, \mu)] \frac{\partial}{\partial x} + [\Omega(\mu)y + g(y, \mu)] \frac{\partial}{\partial y}, \quad (31)$$

with  $f = O(|y|)$  and  $g = O(|y|^2)$ . Moreover,  $\mu \in P$  is a multi-parameter with  $P \subseteq \mathbb{R}^s$  open, while  $\omega : P \rightarrow \mathbb{R}^n$  and  $\Omega : P \rightarrow gl(2, \mathbb{R})$  are smooth maps. Without loss of generality we may assume that  $\Omega$  has the form

$$\Omega(\mu) = \begin{pmatrix} \alpha(\mu) & -\beta(\mu) \\ \beta(\mu) & \alpha(\mu) \end{pmatrix}.$$

REMARK. Note that for  $\alpha \neq 0$ , we return to the normally hyperbolic situation of Theorem 5 of Section 4.2.1, but here we are particularly interested in the transition occurring when  $\alpha$  passes through 0. In addition, the dissipative parametrized KAM Theorem 12 of Section 5.2.1 is applicable to small perturbations of (31) for any  $\alpha$  (with  $m = 2$ ,  $N_1 = 0$ ,  $N_2 = 1$ ).

The assumption of BHT nondegeneracy leads to the existence of an open subset  $A \subseteq P$  on which the map

$$P \ni \mu \mapsto (\omega \times (\text{spec} \circ \Omega))(\mu) = (\omega(\mu), \alpha(\mu), \beta(\mu)) \in \mathbb{R}^n \times \mathbb{R}^2$$

is a submersion [78]; compare with (29). As in Section 5.2.1, we re-parametrize

$$A \ni \mu \leftrightarrow (\omega, \alpha, \beta) \in \Gamma \subset (\mathbb{R}^n \times \mathbb{R}^2),$$

suppressing, for simplicity, the possible occurrence of other parameters. Here we denote by  $\Gamma$  the open  $(\omega, \alpha, \beta)$ -parameter domain, without loss of generality, taking it of the product form  $\Gamma = \Gamma_\omega \times \Gamma_\alpha \times \Gamma_\beta$ , as in Section 5.2.1, and assuring that  $0 \in \Gamma_\alpha$ .

We furthermore assume that the reduced system

$$X_{\text{red},(\omega,\alpha,\beta)}(y) = [\Omega(\alpha, \beta)y + g(y, \omega, \alpha, \beta)] \frac{\partial}{\partial y} \quad (32)$$

exhibits a standard supercritical Hopf bifurcation [19,21,127,202,260,454] at  $\alpha = 0$ . This means that the (relative) equilibrium  $y = 0$  is attracting for  $\alpha < 0$  and repelling for  $\alpha > 0$ , while for  $\alpha > 0$  a (relative) periodic solution that is attracting branches off, also compare with [1]. These statements hold for ‘a half’ of all nonlinearities  $g$ : depending on the sign of the coefficient of a certain third order term in the normal form for  $g$ , the system (32) admits either an attracting periodic solution for  $\alpha > 0$  (a supercritical Hopf bifurcation), or a repelling periodic solution for  $\alpha < 0$  (a subcritical Hopf bifurcation).

For the integrable family  $X = X_{\omega,\alpha,\beta}(x, y)$  this scenario directly translates to  $n$ - and  $(n + 1)$ -tori with parallel dynamics, and the question is: what is their fate when in the  $C^\infty$ -topology we perturb  $X$  to a nearly integrable family  $\tilde{X} = \tilde{X}_{\omega,\alpha,\beta}(x, y)$ ?

**6.1.1. Persistent quasi-periodic  $n$ -tori.** We start answering the question of persistent invariant  $n$ -tori by applying [Theorem 12](#) in the present setting. Therefore, for  $\tau > n - 1$  and  $\gamma > 0$ , as a special case of [\(30\)](#), we consider the set

$$\begin{aligned} D_{\tau,\gamma}(\mathbb{R}^n; \mathbb{R}) = \{(\omega, \beta) \in \mathbb{R}^n \times \mathbb{R} \mid |(\omega, k) + \beta\ell| \geq \gamma|k|^{-\tau}, \\ \text{for all } k \in \mathbb{Z}^n \setminus \{0\} \text{ and for all } \ell \in \mathbb{Z} \text{ with } |\ell| \leq 2\}. \end{aligned} \quad (33)$$

The sets  $\Gamma_\gamma$  and  $D_{\tau,\gamma}(\Gamma_\gamma)$  are now defined as in [Section 5.2.1](#). For a sketch of a section of  $D_{\tau,\gamma}(\Gamma_\gamma)$  for fixed Diophantine  $\omega$ , see [Figure 4](#); compare with [\[44,421\]](#). We also introduce the full measure subset of density points  $D_{\tau,\gamma}^d(\mathbb{R}^n; \mathbb{R}) \subseteq D_{\tau,\gamma}(\mathbb{R}^n; \mathbb{R})$ , see [Section 5.1.1](#). We take  $\gamma > 0$  sufficiently small, so that the projection  $\Gamma_{\alpha\gamma}$  of  $\Gamma_\gamma$  on  $\Gamma_\alpha$  has  $\alpha = 0$  as an interior point. Also we take  $\gamma > 0$  sufficiently small for the nowhere dense set  $D_{\tau,\gamma}(\Gamma_\gamma)$  to have positive measure. In addition, we consider  $D_{\tau,\gamma}^d(\Gamma_\gamma)$ , obtained by taking the product with the interval  $\Gamma_{\alpha\gamma}$ . As a consequence of [Theorem 12](#), for any family  $\tilde{X}$  on  $\mathbb{T}^n \times \mathbb{R}^2 \times P$ , sufficiently close to  $X$  in the  $C^\infty$ -topology, a near-identity  $C^\infty$ -diffeomorphism  $\Phi : \mathbb{T}^n \times \mathbb{R}^2 \times \Gamma \rightarrow \mathbb{T}^n \times \mathbb{R}^2 \times \Gamma$  exists, defined near  $\mathbb{T}^n \times \{0\} \times \Gamma$ , that conjugates  $X$  to  $\tilde{X}$  when further restricted to  $\mathbb{T}^n \times \{0\} \times D_{\tau,\gamma}(\Gamma_\gamma)$ .

Next consider the perturbed family  $\tilde{X}$  in the coordinates provided by the inverse  $\Phi^{-1}$ . In other words, we study the pull-back vector field  $\Phi^*\tilde{X} = (\Phi^{-1})_*\tilde{X}$ , that, on the nowhere dense set  $\mathbb{T}^n \times \{0\} \times D_{\tau,\gamma}(\Gamma_\gamma)$ , coincides with the integrable family  $X$ . We directly conclude that  $\Phi^*\tilde{X}$  for parameter values in  $D_{\tau,\gamma}(\Gamma_\gamma)$  has  $\mathbb{T}^n \times \{0\}$  as a quasi-periodic invariant  $n$ -torus; this torus is attracting for  $\alpha < 0$  and repelling for  $\alpha > 0$ . As in [Section 5.1.1](#), we further restrict to the set of density points  $D_{\tau,\gamma}^d(\Gamma_\gamma) \subset D_{\tau,\gamma}(\Gamma_\gamma)$  as a full measure subset, which leads to the normal form decomposition

$$\left(\Phi^*\tilde{X} - X\right)_{\omega,\alpha,\beta}(x, y) = O(|y|)\frac{\partial}{\partial x} + O(|y|^2)\frac{\partial}{\partial y} + Q_{\omega,\alpha,\beta}(x, y), \quad (34)$$

as  $y \rightarrow 0$ . For bounded  $\Gamma$  the  $O$ -estimates are uniform in  $x$  and  $\omega, \alpha, \beta$ . The  $C^\infty$ -family of vector fields  $Q$  is uniformly flat on  $\mathbb{T}^n \times \Delta \times D_{\tau,\gamma}^d(\Gamma_\gamma) \subset \mathbb{T}^n \times \mathbb{R}^2 \times \Gamma$ , where  $\Delta$  is a small neighbourhood of  $0 \in \mathbb{R}^2$ . This means that the Taylor series of  $Q$  completely vanishes. Indeed, for small  $\Delta$  we can arrange that  $Q$  vanishes on the  $\mathbb{T}^n \times \Delta \times D_{\tau,\gamma}^d(\Gamma_\gamma)$ , which by the definition of a density point implies that on the open set all the derivatives of  $Q$  vanish. This is what the flatness expresses.

**6.1.2. Fattening the parameter domain of invariant  $n$ -tori.** We keep studying the perturbed system  $\tilde{X}$  in its pull-back form, so we are still considering  $\Phi^*\tilde{X}$ . For  $\alpha \neq 0$ , the invariant  $n$ -tori in question are normally hyperbolic. By the Centre Manifold Theorem [\[126,175,226,434\]](#) we conclude that the parameter domain inside  $\Gamma$ , where invariant  $n$ -tori exist, contains a neighbourhood of the parameter values corresponding to



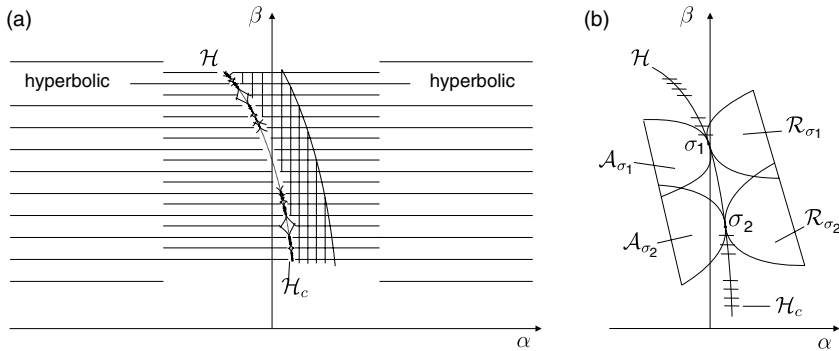


Fig. 5. Sketch of a ‘frayed’ Hopf boundary, where  $\mathcal{H}_c \subset \mathcal{H}$  is nowhere dense of positive measure, while the boundary  $\mathcal{H}$  is a smooth curve. (a). Global impression of the domains with attracting or repelling  $n$ - and  $(n+1)$ -tori (doubly shaded), compare with [44]. (b). Formation of one resonance ‘bubble’ in between discs attached to  $\sigma_1, \sigma_2 \in \mathcal{H}_c$ .

the Diophantine quasi-periodic tori. In other words, the nowhere dense parameter domain  $D_{\tau, \gamma}^d(\Gamma_\gamma)$  of  $\Phi^* \tilde{X}$ -invariant  $n$ -tori, for  $\alpha \neq 0$ , can be *fattened* to an *open* subset of  $\Gamma$ . We note that outside  $D_{\tau, \gamma}^d(\Gamma_\gamma)$ , the dynamics on invariant  $n$ -tori does *not* have to be conditionally periodic.

The fattening by hyperbolicity can be carried out using a more or less well-known contraction principle, see e.g. [126]; for a detailed construction using a variation of constants operator see [44]. Here we are restricted to describing the result of the fattening operation. To this purpose we proceed as follows.

1. As before assume that  $\Gamma = \Gamma_\omega \times \Gamma_\alpha \times \Gamma_\beta$ , i.e. that  $\Gamma$  is of the product form; compare with Figure 4.
2. In the frequency space  $\mathbb{R}^n \setminus \{0\} = \{\omega\}$ , define  $\bar{\omega} = \omega/|\omega| \in \mathbb{S}^{n-1} \subset \mathbb{R}^n$ . Also, let  $\varrho : \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \rightarrow \mathbb{R}_+$  be the metric  $\mathbb{S}^{n-1}$  inherits from  $\mathbb{R}^n$ .
3. Finally consider any monotonically increasing  $C^\infty$ -function  $p : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  that is (infinitely) flat at 0.

For any fixed  $\omega_0 = |\omega_0|\bar{\omega}_0 \in \Gamma_\omega$  and  $\beta_0 \in \Gamma_\beta$ , such that  $(\omega_0, \alpha, \beta_0) \in D_{\tau, \gamma}(\Gamma_\gamma)$  for all  $\alpha \in \Gamma_{\alpha\gamma}$ , consider sets of the form

$$\left\{ (\omega, \alpha, \beta) \in \Gamma \mid 0 < |\alpha| < C \text{ and } p(\varrho(\bar{\omega}, \bar{\omega}_0) + |\beta - \beta_0|) < D|\alpha|^K \right\}, \quad (35)$$

where  $C$ ,  $D$ , and  $K$  are positive constants. Notice (compare with Figure 5) that this is the union of two open discs  $\mathcal{A}_{\omega_0, \beta_0}$  (occurring for  $\alpha < 0$ ) and  $\mathcal{R}_{\omega_0, \beta_0}$  (occurring for  $\alpha > 0$ ), each with a piecewise smooth boundary, that at  $\beta = \beta_0$  have an infinite order of contact with the bifurcation hyperplane  $\alpha = 0$ .<sup>9</sup>

<sup>9</sup>In [44, 78], for historical reasons, instead of ‘disc’ the term ‘(blunt or flat conic) cusp’ was used.

**PROPOSITION 13** (Fattened Domain of  $n$ -tori [44,78]). *In the above situation, given  $r \in \mathbb{N}$ , there exist positive constants  $C$  and  $D$  with the following property. For each  $(\omega_0, \beta_0)$  such that  $(\omega_0, \alpha, \beta_0) \in D_{\tau, \gamma}^d(\Gamma_\gamma)$  for all  $\alpha \in \Gamma_{\alpha\gamma}$ , the corresponding discs  $\mathcal{A}_{\omega_0, \beta_0}$  and  $\mathcal{R}_{\omega_0, \beta_0}$  (35) with  $K = 3$  are contained in the parameter domain with normally hyperbolic  $\Phi^* \tilde{X}$ -invariant  $n$ -tori of class  $C^r$ . These tori are attracting in  $\mathcal{A}_{\omega_0, \beta_0}$  and repelling in  $\mathcal{R}_{\omega_0, \beta_0}$ .*

**REMARKS.** 1. The discs  $\mathcal{A}_{\omega_0, \beta_0}$  and  $\mathcal{R}_{\omega_0, \beta_0}$  become larger as the degree of differentiability  $r$  decreases.

2. The domain of invariant tori is an uncountable union of discs, leaving open a countable number of ‘bubbles’ centred around the pure resonances  $(\omega, 0, \beta) \in \Gamma$  with  $\langle \omega, k \rangle + \beta \ell = 0$  for some  $k \in \mathbb{Z}^n \setminus \{0\}$  and  $\ell = -2, -1, 0, 1, 2$ .

3. The resonant dynamics inside certain bubbles of the quasi-periodic Hopf bifurcation also has been widely studied, see e.g. [27,80,188,261,384,443]. For similar studies related to the quasi-periodic saddle-node bifurcation, see [107–109].

**6.1.3. The parameter domain of invariant  $(n+1)$ -tori.** In order to find invariant  $(n+1)$ -tori, we first develop a new pull-back of the perturbed system  $\tilde{X}$  that has a  $\mathbb{T}^{n+1}$ -symmetric normal form truncation and which is related to both the planar supercritical Hopf family  $X_{\text{red}}$  and to the quasi-periodic normal form (34). To this purpose, given  $N \in \mathbb{N}$ , consider the subset

$$E_{\tau, \gamma; N}(\mathbb{R}^n; \mathbb{R}) = \left\{ (\omega, \beta) \in \mathbb{R}^n \times \mathbb{R} \mid |\langle \omega, k \rangle + \beta \ell| \geq \gamma |k|^{-\tau}, \right. \\ \left. \text{for all } k \in \mathbb{Z}^n \setminus \{0\} \text{ and for all } \ell \in \mathbb{Z} \text{ with } |\ell| \leq N \right\}, \quad (36)$$

compare with (33), which again is a nowhere dense set of positive measure (for  $\gamma$  sufficiently small), with the closed half-line property. Note that  $E_{\tau, \gamma; 2}(\mathbb{R}^n; \mathbb{R}) = D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R})$ . Accordingly, one may define  $E_{\tau, \gamma; N}(\Gamma_\gamma)$  and  $E_{\tau, \gamma; N}^d(\Gamma_\gamma)$ . In these circumstances we can roughly paraphrase Theorem 12 as follows. For  $\alpha$  sufficiently small, there exists a near-identity  $C^\infty$ -diffeomorphism  $\Phi$  defined near  $\mathbb{T}^n \times \{0\} \times \Gamma \subset \mathbb{T}^n \times \mathbb{R}^2 \times \Gamma$ , such that the following normal form decomposition holds:

$$\begin{aligned} \left( \Phi^* \tilde{X} \right)_{\omega, \alpha, \beta}(x, y) &= \left[ \omega + |y|^2 f(|y|^2, \omega, \alpha, \beta) + O(|y|^N) \right] \frac{\partial}{\partial x} \\ &+ \left[ \beta + |y|^2 g(|y|^2, \omega, \alpha, \beta) + O(|y|^{N+1}) \right] \left[ -y_2 \frac{\partial}{\partial y_1} + y_1 \frac{\partial}{\partial y_2} \right] \\ &+ \left[ \alpha + |y|^2 h(|y|^2, \omega, \alpha, \beta) + O(|y|^{N+1}) \right] \left[ y_1 \frac{\partial}{\partial y_1} + y_2 \frac{\partial}{\partial y_2} \right] \\ &+ Q(x, y, \omega, \alpha, \beta), \end{aligned} \quad (37)$$

where the family  $Q$  of vector fields is uniformly flat on  $\mathbb{T}^n \times \{0\} \times E_{\tau, \gamma; N}(\Gamma_\gamma)$ .

Indeed, decomposition (37) for  $N \geq 2$  is obtained by initially applying Theorem 12, followed by a standard formal normal form procedure as developed in [44,47,78,127,426]. For  $N = 2$  we recover (34). Thus, the  $\mathbb{T}^{n+1}$ -symmetry of the normal linear part for  $\alpha = 0$  is pushed over the formal series in  $y$ .

In our application we take  $N = 7$ . Note that then the  $(\partial/\partial y)$ -component of (37) is close to the standard planar Hopf normal form [47,260,426]. The invariant  $(n + 1)$ -tori now can be found in a straightforward manner.

**PROPOSITION 14** (Fattened Domain of  $(n + 1)$ -tori [78]). *In the above situation, given  $r \in \mathbb{N}$ , there exist positive constants  $C$  and  $D$  with the following property. For each  $(\omega_0, \beta_0)$  such that  $(\omega_0, \alpha, \beta_0) \in E_{\tau, \gamma; 7}^d(\Gamma_\gamma)$  for all  $\alpha \in \Gamma_{\alpha\gamma}$ , the corresponding disc  $\mathcal{R}_{\omega_0, \beta_0}$ , i.e. the set (35) for  $\alpha > 0$ , with  $K = 7/2$ , is contained in the parameter domain with normally hyperbolic  $\Phi^* \tilde{X}$ -invariant  $(n + 1)$ -tori of class  $C^r$ . These tori are attracting.*

Mutatis mutandis, the same remarks apply as those following Proposition 13.

## 6.2. Discussion

For an overview of the quasi-periodic Hopf bifurcation, see [44,51,54,77,78,132]. The quasi-periodic saddle-node and period doubling bifurcations have a similar structure [78], although the saddle-node case is more involved. For a Hamiltonian version of the latter case see [208]. For an early treatment of such torus bifurcations with only one parameter, see [110,111].

**6.2.1. Fraying.** We summarize the above results as follows. The quasi-periodic bifurcations to some extent are similar to their periodic analogues. However, as already seen in the Hopf–Neĭmark–Sacker bifurcation, the phenomenon of normal-internal resonances (which in this case leads to Arnold resonance tongues) is only visible when an extra parameter is taken into account. For quasi-periodic bifurcations even more parameters may be needed.

In fact, the main difference from the periodic theory is that the presence of resonances leads to Cantorization; compare with Figure 4. To be more precise, in the periodic theory, the subsets of the parameter space corresponding to non-hyperbolicity are piecewise smooth manifolds. The same holds true also in the torus case when we would consider only integrable systems. In the nearly integrable case, however, the dense set of resonances really interrupts these bifurcation boundaries.

In the present dissipative setting, the domains of hyperbolicity can be fattened to open subsets of the parameter space, that near the normal-internal resonances leave over strands of bubbles. The total effect of this is called *fraying* of the bifurcation boundaries; compare with Figure 5.

**6.2.2. Non-parallel dynamics.** The parameter domains with quasi-periodic tori are nowhere dense and of positive measure. In the open domains with normally hyperbolic tori, several types of dynamics can occur: already in 3-tori, next to quasi-periodicity one meets periodicity (phase lock) and chaos [333].

Inside a bubble we are closer to a resonance of the form  $\langle \omega, k \rangle + \beta \ell = 0$ . For  $\ell = 0$  this is an internal resonance, while for  $\ell \neq 0$  the resonance is normal-internal. Compare with [27,80,108,109,188,238,270–272,289,384,415,427,443] for research in this direction; some of these works are case studies, while others have a more generic point of view. For a Hamiltonian analogue see [65,66]. One important aspect is the repetition of the whole scenario within bubbles. The near-resonance dynamics is quite rich, and may involve cantori and chaos.

**6.2.3. Final remarks.** The quasi-periodic bifurcation theory sketched above has been extended and applied in various directions. A direct generalization of the quasi-periodic saddle-node bifurcation to the case of cusps and higher order degenerate bifurcations is given in [445]. One can also develop a bifurcation theory for invariant tori in dissipative systems without reference to quasi-periodicity (see e.g. [32]).

**The quasi-periodic response problem** A widely used context for applications of KAM Theory is that of response solutions in quasi-periodically forced oscillators. Again, here we consider only the dissipative case, which goes back to J.J. Stoker [421]. To be definite, we identify the leading example, where a free Duffing–Van der Pol oscillator is forced as follows:

$$\ddot{u} + (a + cu^2)\dot{u} + bu + du^3 = \varepsilon f(\omega_1 t, \dots, \omega_m t, u, \dot{u}, a, b, \varepsilon),$$

also compare with Section 3.1.3. Here the perturbation  $f$  is assumed to be  $2\pi$ -periodic in each of its first  $m$  arguments. This non-autonomous second order differential equation can be written as a vector field (in the system form)

$$\begin{aligned} \dot{x}_j &= \omega_j, & j &= 1, \dots, m, \\ \dot{u} &= v, \\ \dot{v} &= -(a + cu^2)v - bu - du^3 + \varepsilon f(x_1, \dots, x_m, u, v, a, b, \varepsilon), \end{aligned} \tag{38}$$

defined on the phase space  $\mathbb{T}^m \times \mathbb{R}^2 = \{(x_1, \dots, x_m; u, v)\}$ . We consider  $(a, b) \in \mathbb{R}^2$  (varying over an open domain) as a multi-parameter. In this setting the frequency vector  $\omega \in \mathbb{R}^m$  is fixed, with rationally independent components, which is why the forcing is said to be quasi-periodic. The function  $f$  is assumed to be of class  $C^\infty$ . Finally  $\varepsilon \in \mathbb{R}$ , with  $|\varepsilon| \ll 1$ , is as usual a perturbation parameter.

The response problem asks for the existence of quasi-periodic solutions with the fixed frequency vector  $\omega$ . This problem reduces to that of an invariant  $m$ -torus of (38) in the phase space  $\mathbb{T}^m \times \mathbb{R}^2$ , which is a graph  $y = y(x)$  over  $\mathbb{T}^m \times \{0\}$ ; each such  $m$ -torus thus corresponds to an  $m$ -parameter family of response solutions  $y = y(x(t))$ . Here we denote  $y = (u, v)$ . In the case of strong damping  $|a| \gg 1$  (with  $a$  either positive or negative) the problem is solved in [421]: in a more contemporary language, it reduces to the persistence of normally hyperbolic invariant  $m$ -tori close to  $\mathbb{T}^m \times \{0\} \subset \mathbb{T}^m \times \mathbb{R}^2$ ; compare with Section 4.2. In the case of small damping  $|a|$  a quasi-periodic Hopf bifurcation occurs, for details see [44] and for a discussion [51,77,78,132]. The dynamics associated with normal-internal resonance bubbles is treated in [80,188,384,443].

REMARK. The response problem has many analogues in the Hamiltonian and reversible contexts; compare with e.g. [73,94,233,242,317,318,410,436,439].

**The skew Hopf bifurcation** An extension of the above theory is formed by the *skew* Hopf bifurcation, in its simplest form taking place as a diffeomorphism of the solid torus  $\mathbb{T}^1 \times \mathbb{C}$ :

$$P_{\beta,c} : \mathbb{T}^1 \times \mathbb{C} \rightarrow \mathbb{T}^1 \times \mathbb{C};$$

$$(x, z) \mapsto (x + 2\pi\beta, cze^{ikx}) + \text{higher order terms}, \quad (39)$$

where  $\beta$  and  $c$  are real parameters and we take  $c > 0$ . Moreover,  $k \in \mathbb{Z}$  is a discrete ‘parameter’. The system (39) turns out to be non-reducible whenever  $k \neq 0$ .

Integrability in this case amounts to rotational symmetry in the  $z$ -direction. In the integrable case we see that the circle  $T_{\beta,c} = \mathbb{T}^1 \times \{0\}$  is invariant, being a hyperbolic attractor for  $0 < c < 1$  and a hyperbolic repeller for  $c > 1$ . Moreover, it turns out that for  $c > 1$ , a 2-torus attractor or repeller  $T'$  is born, where near  $c = 1$  resonance problems occur, leading to both Cantorization and fraying in the  $(\beta, c)$ -plane. The corresponding perturbation problem was studied both in the integrable [89,90] and in the nearly integrable [90,96,427,442] cases. As said before in Section 5.2.3, this was an early, successful attempt to develop KAM Theory for non-reducible systems.

REMARK. In the integrable case, system (39) turns out to have a mixed power spectrum, which may have some interest for certain experiments with rotational symmetry, where a mixed spectrum occurs [89].

**Onset of turbulence** The quasi-periodic Hopf bifurcation has an interest for the onset of turbulence as described by the theories of Landau–Hopf–Lifshitz and of Ruelle–Takens [229,263,264,369–371]. The idea is to view this aspect of fluid dynamics in finitely many dimensions, the number of which is increasing when the turbulence gets more developed. For instance, a stationary fluid flow corresponds to equilibrium dynamics. In the initial Landau–Hopf–Lifshitz theory [229,263,264], there are repeated Hopf bifurcations, where the dynamics stays quasi-periodic, but picks up more and more frequencies, thereby complicating the phase portrait. Later the Ruelle–Takens theory [369–371] modified this picture by pointing out that already in 3-tori there can be strange attractors with chaotic dynamics.

We observe that both scenarios have been unified in the present generic theory. For another example of how to incorporate quasi-periodicity with  $n$  frequencies (for  $n \in \mathbb{N}$  arbitrary) in an infinite dimensional dynamical system, see [54,176]. For further discussion also see [77,78,132].

## 7. Quasi-periodic bifurcation theory in other settings

The marriage of KAM Theory with Bifurcation Theory extends far beyond the dissipative setting, but largely follows the same approach.

### 7.1. Hamiltonian cases

As an example consider the Hamiltonian centre-saddle bifurcation, a Singularity Theory model of which is given by

$$H(p, q) = \frac{1}{2}p^2 + V_\mu(q), \quad \text{where } V_\mu(q) = \frac{1}{3}q^3 - \mu q,$$

with the phase space  $\mathbb{R}^2 = \{(p, q)\}$  and with one real parameter  $\mu$ ; compare with [21, 212, 428]. The corresponding system reads

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\frac{dV_\mu}{dq}(q). \end{aligned} \tag{40}$$

The equilibria are given by the equations

$$p = 0, \quad q^2 - \mu = 0, \tag{41}$$

which determine a curve in the product  $\{(p, q, \mu)\}$  of the phase space and the parameter space. This curve is smoothly parametrized by  $q$ . The equilibria are elliptic for  $q > 0$  and hyperbolic for  $q < 0$ . For  $\mu = 0$ , a parabolic singularity (fold) occurs at  $q = 0$ . The problem is: what happens to this scenario when (40) is changed to the integrable Hamiltonian system

$$\begin{aligned} \dot{x} &= \omega(y, p, q) \\ \dot{y} &= 0 \\ \dot{q} &= p \\ \dot{p} &= -\frac{dV_\mu}{dq}(q), \end{aligned} \tag{42}$$

with  $x \in \mathbb{T}^n$  and  $y \in \mathbb{R}^n$ , and even more, what happens when nearly integrable Hamiltonian perturbations of (42) are considered? As in the previous section, one directly sees that in the integrable case the (relative) equilibria of (40) correspond to invariant  $n$ -tori of (42), which can be either elliptic or hyperbolic. Incorporating nearly integrable perturbations again requires KAM Theory [208].

It turns out that the bifurcation set (41) becomes Cantorized by the resonances and the corresponding Diophantine conditions

$$\begin{aligned} |\langle \omega(y, p, q), k \rangle| &\geq \gamma |k|^{-\tau} \quad \text{for } q < 0 \quad \text{and} \\ |\langle \omega(y, p, q), k \rangle + \beta \ell| &\geq \gamma |k|^{-\tau} \quad \text{for } q > 0, \end{aligned}$$

as usual, for all  $k \in \mathbb{Z}^n \setminus \{0\}$  and for all  $\ell \in \mathbb{Z}$  with  $|\ell| \leq 2$ , which apparently differ for the hyperbolic case ( $q < 0$ ) and the elliptic case ( $q > 0$ ). Here  $\beta = \sqrt{2\sqrt{\mu}} = \sqrt{2q}$  is the normal frequency in the elliptic case, as before being the positive imaginary part of the purely imaginary eigenvalue  $i\beta$  associated with the normal linear part. For a sketch of

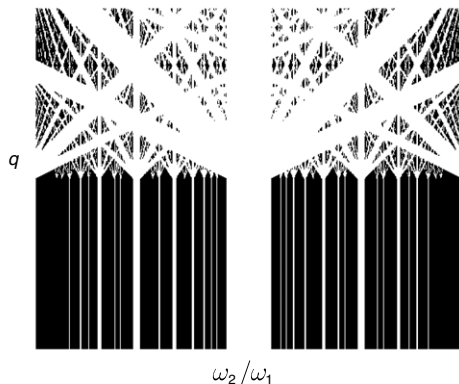


Fig. 6. Sketch of the Cantorized bifurcation set of the quasi-periodic centre-saddle bifurcation for  $n = 2$  [208, 209, 212], where the horizontal axis indicates the frequency ratio  $\omega_2 : \omega_1$ . The lower part of the figure corresponds to hyperbolic tori and the upper part to elliptic tori.

the bifurcation set, see Figure 6. In fact, as before, there is a Whitney smooth conjugation between the integrable and nearly integrable tori, slightly deforming the nowhere dense domains of Diophantine quasi-periodic tori. We observe that, in the Hamiltonian setting, even in the hyperbolic case, *no fattening occurs*. This is due to the equation  $\dot{y} = 0$  in (42): the ‘hyperbolic’ invariant tori in Hamiltonian systems are *not* ‘normally hyperbolic’ in the sense of the Centre Manifold Theory [126, 175, 226, 434] (see Sections 8.2 and 8.3.1 below, as well as a detailed discussion in [36]).

The phase space variables  $(y, q)$  can ‘absorb’ the external parameters  $(\omega, \mu)$ , under suitable nondegeneracy conditions of Kolmogorov (including the iso-energetic counterpart) or Rüssmann types, as described in Sections 4.3.2 and 5.2.2.

The study of quasi-periodic bifurcations in the Hamiltonian setting has been developed further for normally parabolic and normally umbilic torus bifurcations related to simple singularities and to singularities involving model parameters [68, 69, 208–212]. We emphasize that these results do not include a strict application of the structure preserving Parametrized KAM Theory as mentioned above in Section 5.2.1 and described in detail below (for the Hamiltonian case) in Section 8.3.1, since the analogue of the condition  $\det \Omega(\mu_0) \neq 0$  is violated. The bifurcation sets of the integrable approximations exhibit the familiar hierarchy in the stratifications of Catastrophe Theory [19, 21, 23, 24, 72, 95, 192, 294, 428], and the nearly integrable perturbations yield a Cantorized version of these bifurcation sets.

Another branch of this theory studies the quasi-periodic Hamiltonian Hopf bifurcation [63, 64, 227, 337], built on the normal  $1:-1$  resonance and illustrated by the Lagrange top [142] subject to periodic and quasi-periodic forcing. Here Parametrized KAM Theory can be directly used. Within this bifurcation, higher dimensional tori branch off, for an overview see [212]. For a case study of the Hamiltonian normal  $1:1:\dots:1$  resonance, see [145]. For a case study of dynamical effects of Hamiltonian normal-internal resonances, see [65, 66], which is a conservative counterpart of [80, 384]. The standard reference on the ‘conventional’ (non-quasi-periodic) Hamiltonian Hopf bifurcation is the book [435].

## 7.2. Discussion

We already mentioned reversible KAM Theory several times in Sections 5.2.1 and 5.2.2, so now we just refer to [57,58,130,131] for a study in the normal 1:1 resonance, that generically involves the quasi-periodic reversible Hopf bifurcation.<sup>10</sup>

We summarize that normal-internal resonances give rise to Cantorization of the corresponding ‘classical’ bifurcation diagrams in the case of (relative) equilibria or periodic solutions. These ‘classical’ diagrams, up to diffeomorphisms, are real semi-algebraic sets obtained from Singularity Theory. The Cantorized objects sometimes have been called *Cantor stratifications*. Moreover, in certain hyperbolic cases, the nowhere dense domain of Diophantine quasi-periodic tori, as obtained so far, can be fattened. We however emphasize that this yields invariant tori that do *not* have to be quasi-periodic.

Quasi-periodic bifurcation theory becomes important in higher dimensional modelling when dense sets of resonances can lead to Cantorization of bifurcation sets. For examples see [54,87,107–109,238,289,438].

## 8. Further Hamiltonian KAM Theory

In this section we discuss KAM Theory for invariant tori in Hamiltonian systems, other than the Lagrangean case, which is familiar from Section 4.3. This concerns lower dimensional isotropic tori, higher dimensional coisotropic tori, and ‘atropic’ tori (which are neither isotropic nor coisotropic). We treat the case of lower dimensional isotropic tori in great detail and consider also the possible excitation of elliptic normal modes of these tori. However, we start with certain properties of the Lagrangean tori and describe their ‘exponential condensation’ and ‘superexponential stickiness’ as well as the mechanisms of the destruction of resonant tori.

### 8.1. Exponential condensation

In Sections 8.1 and 8.2, our concern is some addenda to the Lagrangean KAM Theorem 6, and we retain all the notations and concepts of Sections 4.3 and 4.4 (see especially Section 4.3.1). Consider again a Hamiltonian vector field  $\tilde{X}$  (with  $n \geq 2$  degrees of freedom) close to an integrable Hamiltonian field  $X = X_H$  defined on  $\mathbb{T}^n \times P$ ,  $P \subseteq \mathbb{R}^n$  being open and connected. Let us assume  $X$  and  $\tilde{X}$  to satisfy the following.

1. Both the  $X$  and  $\tilde{X}$  are real analytic.
2.  $X$  is Kolmogorov nondegenerate on  $\mathbb{T}^n \times A$  for an open subset  $A \subseteq P$ .
3. The Hamiltonian  $H$  is *quasi-convex* on  $A$ . This means that  $\langle (D_y \omega) \eta, \eta \rangle \neq 0$  whenever  $y \in A$ ,  $\eta \in \mathbb{R}^n \setminus \{0\}$ , and  $\langle \eta, \omega(y) \rangle = 0$ ; recall that  $D_y \omega$  is the derivative of the frequency map  $\omega : P \rightarrow \mathbb{R}^n$ ,  $\omega = \partial H / \partial y$ , and  $(x \in \mathbb{T}^n, y \in P)$  are the angle-action coordinates for  $X$ .

<sup>10</sup>In the reversible setting, there is no difference between the 1:1 and the 1:–1 resonance.



The concept of quasi-convex integrable Hamiltonians was introduced by N.N. Nekhoroshev [328]. Quasi-convexity of  $H$  means strict convexity of the corresponding unperturbed energy hypersurfaces  $H^{-1}(c)$  regarded as surfaces in  $A$  (rather than in  $\mathbb{T}^n \times A$ ). The properties of Kolmogorov nondegeneracy and quasi-convexity are independent (as simple examples show, see [77, §4.2.3]) but quasi-convexity implies iso-energetic nondegeneracy [283] (see Section 4.3.2) and is equivalent to iso-energetic nondegeneracy for  $n = 2$  [283, 328].

The system  $X$  is Kolmogorov nondegenerate. Consequently, according to Theorem 6, if  $\tilde{X}$  lies in a sufficiently small neighbourhood  $\mathcal{O}$  of  $X$  (say, in the  $C^\infty$ -topology), then  $\tilde{X}$  possesses a family of Diophantine quasi-periodic Lagrangean invariant  $n$ -tori (KAM tori) in  $\mathbb{T}^n \times A$ . These tori are close to the unperturbed tori  $\mathbb{T}^n \times \{y\}$ , and their union  $\mathcal{W}$  fills up positive measure. Actually, since  $X$  is also iso-energetically nondegenerate, the union of  $\tilde{X}$ -invariant tori fills up positive measure even on each perturbed energy hypersurface in the phase space.

**THEOREM 15** (Exponential Condensation and Superexponential Stickiness of KAM Tori [314]). *Under the three conditions above, let  $\tilde{X}$  be fixed and sufficiently close to  $X$  and let  $\mathcal{T}$  be an arbitrary fixed KAM torus of  $\tilde{X}$ . Denote by  $\mathcal{U}_\rho(\mathcal{T})$  the  $\rho$ -neighbourhood of  $\mathcal{T}$  in the phase space. Then the measure of  $\mathcal{U}_\rho(\mathcal{T}) \setminus \mathcal{W}$  is at most of the order of  $\exp(-c^*/\rho)$  as  $\rho \downarrow 0$ , where  $c^* > 0$  is a certain constant. Moreover, if the frequency vector of torus  $\mathcal{T}$  is  $(\tau, \gamma)$ -Diophantine, then all the  $\tilde{X}$ -evolution curves starting at a distance  $\rho < \rho^*$  from  $\mathcal{T}$  stay near  $\mathcal{T}$  over an exceedingly long time  $T_{\text{hold}}$  of the order of*

$$\exp \left\{ \exp \left[ \left( \frac{\rho^*}{\rho} \right)^{1/(\tau+1)} \right] \right\},$$

where  $\rho^* > 0$  is again a certain constant.

One says that the KAM tori ‘exponentially condense’ to each of them [403] and that each torus is ‘superexponentially sticky’. The latter property was already mentioned at the end of Section 4.4.2.

- REMARKS.**
1. We like to note that the phenomena described here are related to the following. We consider the completely general set-up of Theorems 5 and 12 and the ensuing remarks. In all these cases KAM tori are density points of quasi-periodicity [77, 78]. If the ambient setting is of class  $C^k$ , for  $k \in \mathbb{N} \cup \{\infty\}$  large, the corresponding estimates will be only polynomial. In the real analytic category, these estimates can be directly sharpened to exponential or superexponential [82, 83, 331, 417].
  2. Any Cantor set  $\mathcal{C}$  is perfect, which means that each of its points  $a$  is an accumulation point of the complement  $\mathcal{C} \setminus \{a\}$ . Every torus among the KAM tori that ‘condense’ to  $\mathcal{T}$  in Theorem 15 is, in turn, a ‘condensation point’ of other KAM tori, and so on. This hierarchy (which seems to be not completely understood yet) is described and discussed in [197, 313, 315].<sup>11</sup>

<sup>11</sup>Similar open questions exist in the characterization of the Diophantine sets  $D_{\tau, \gamma}^d(\mathbb{R}^n) \subseteq D_{\tau, \gamma}(\mathbb{R}^n)$  as mentioned in Section 5.1.1.

3. The ‘exponential condensation’ of KAM tori proven in [314] was confirmed numerically (for a model problem of area preserving maps of the plane, i.e. in the context of Moser’s Twist Mapping [Theorem 3](#)) in [178,269], see also a discussion in [177].
4. Most probably, the quasi-convexity condition imposed on the unperturbed Hamiltonian  $H$  in [Theorem 15](#) is a purely technical one and can be omitted. There is also no doubt that [Theorem 15](#) can be carried over to Gevrey smooth systems (perhaps with a somewhat worse estimate of the time  $T_{\text{hold}}$ ). To the best of the authors’ knowledge, these improvements have not been proven yet. Note, however, that the closely related question of the validity of Nekhoroshev estimates (see the end of [Section 4.4.2](#)) for Gevrey smooth systems has been solved affirmatively in [302, 386].

## 8.2. Destruction of resonant tori

Here we again consider a small Hamiltonian perturbation  $\tilde{X}$  of an integrable Hamiltonian system  $X = X_H$  (with  $n \geq 2$  degrees of freedom) assumed to be Kolmogorov nondegenerate on  $\mathbb{T}^n \times A$ ,  $A \subseteq \mathbb{R}^n$ . According to [Theorem 6](#), Diophantine quasi-periodic invariant  $n$ -tori  $\mathbb{T}^n \times \{y^*\}$  of  $X$  ( $y^* \in A$ ) give rise to Lagrangean invariant  $n$ -tori of  $\tilde{X}$  (KAM tori) with the same frequency vectors  $\omega(y^*)$ , where  $\omega = \partial H / \partial y$ . In between the KAM tori, there lie the so-called *resonant zones* of  $\tilde{X}$ . What is the fate of an unperturbed torus  $\mathbb{T}^n \times \{y^*\}$  in the opposite case, where the frequencies  $\omega_1(y^*), \omega_2(y^*), \dots, \omega_n(y^*)$  are rationally dependent (so that the torus in question is resonant and is foliated into invariant tori of a smaller dimension)? To fix our thoughts, suppose that among the  $n$  components of the vector  $\omega(y^*) = \omega^*$ , there are  $n-l$  ( $1 \leq l \leq n-1$ ) strongly incommensurable numbers

$$\omega_{i_1}^*, \omega_{i_2}^*, \dots, \omega_{i_{n-l}}^*, \quad (43)$$

whereas the remaining  $l$  components  $\omega_{j_1}^*, \omega_{j_2}^*, \dots, \omega_{j_l}^*$  are rational combinations of the numbers (43) — so that the frequencies of the torus  $\mathbb{T}^n \times \{y^*\}$  satisfy  $l$  independent resonance relations. Then the following statement holds.

**THEOREM 16 (Destruction of Resonant Unperturbed Tori).** *Let  $X$  be Kolmogorov nondegenerate. Then for generic sufficiently small perturbations  $\tilde{X}$  of  $X$ , the invariant  $n$ -torus  $\mathbb{T}^n \times \{y^*\}$  of  $X$ , with the properties indicated above, breaks up into a finite collection of Diophantine quasi-periodic invariant  $(n-l)$ -tori of  $\tilde{X}$  which lie in a resonant zone.*

We will not try to attach any exact meaning to the word ‘generic’ here (various precise versions of [Theorem 16](#) can be found in [115,133,149,278,279,429,448]). In fact, in the resonant zones, as a rule, all the complexity of Singularity Theory occurs. This leads to a Cantor stratification in the neighbourhood of any single resonance; compare with [70]. Let us note the following, regarding the codimension 0 strata where the tori are elliptic or hyperbolic or of mixed type, compare with a discussion in [77,408]. Under some mild conditions, via a certain rather standard averaging and truncation procedure described in

detail in e.g. [25], one can reduce the perturbed Hamiltonian of the system  $\tilde{X}$  near the torus  $\mathbb{T}^n \times \{y^*\}$  to the form

$$\widehat{\mathcal{H}}(\chi, J, K) = \langle \widehat{\omega}^*, K \rangle + \mathcal{H}(\chi, J), \quad \mathcal{H}(\chi, J) = \frac{1}{2} \langle BJ, J \rangle + V(\chi),$$

where  $\widehat{\omega}^* \in \mathbb{R}^{n-l}$  is the vector with components (43),  $\chi \in \mathbb{T}^l$ ,  $J \in \mathbb{R}^l$ ,  $K \in \mathbb{R}^{n-l}$ ,  $B$  is a real symmetric  $l \times l$  matrix ( $\det B \neq 0$ ), and  $|V| \ll 1$ . It turns out that generically to each nondegenerate critical point  $\chi^*$  of the function  $V$ , there ‘corresponds’ (in the sense to be made precise) a Diophantine quasi-periodic invariant  $(n-l)$ -torus of the system  $\tilde{X}$ . This torus lies in a small vicinity of the original  $n$ -torus  $\mathbb{T}^n \times \{y^*\}$ . The word ‘nondegenerate’ here means that

$$\det \frac{\partial^2 V(\chi^*)}{\partial \chi^2} \neq 0.$$

On the other hand,  $\chi^* \in \mathbb{T}^l$  is a nondegenerate critical point of the potential  $V$  if and only if the point  $(\chi^*, 0)$  is a nondegenerate equilibrium of the Hamiltonian system with  $l$  degrees of freedom and the Hamiltonian  $\mathcal{H}$  (the symplectic form being  $d\chi \wedge dJ$ ). In the latter case, the word ‘nondegenerate’ means that all the eigenvalues of this equilibrium, i.e. the eigenvalues of the  $2l \times 2l$  matrix

$$\begin{pmatrix} 0 & B \\ -\partial^2 V(\chi^*)/\partial \chi^2 & 0 \end{pmatrix}, \quad (44)$$

are other than zero.

A nondegenerate equilibrium of a Hamiltonian system can be *elliptic* (all the eigenvalues are purely imaginary), *hyperbolic* (all the eigenvalues lie outside the imaginary axis), and *of mixed type*. According to the type of equilibria  $(\chi^*, 0)$  of the system with Hamiltonian  $\mathcal{H}$ , the corresponding invariant  $(n-l)$ -tori of the system  $\tilde{X}$  are also said to be *elliptic*, *hyperbolic*, and *of mixed type* (see [36] for a detailed discussion on the relations between this definition of hyperbolicity of invariant tori in Hamiltonian flows and the general concept of hyperbolic invariant manifolds in dynamical systems as presented in e.g. [126,175,226,434]). The case of hyperbolic  $(n-l)$ -tori in Theorem 16 turns out to be much easier than the case of non-hyperbolic tori (i.e. elliptic tori and tori of mixed type).

For instance, suppose that the Hamiltonian of  $\tilde{X}$  has the form  $H(y) + \varepsilon h(x, y, \varepsilon)$ . Then the potential  $V$  has the form  $V(\chi) = \varepsilon v(\chi, \varepsilon)$ , and for every sufficiently small  $\varepsilon > 0$ , the system  $\tilde{X}$  possesses a hyperbolic invariant  $(n-l)$ -torus ‘emerging’ from a given hyperbolic equilibrium  $(\chi^*, 0)$  of the system with  $l$  degrees of freedom and the Hamiltonian

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \mathcal{H}(\chi, \varepsilon^{1/2} J) = \frac{1}{2} \langle BJ, J \rangle + v(\chi, 0) \quad (45)$$

(provided that the function  $h(\cdot, \cdot, 0)$  is generic). In the analytic category this torus depends on  $\varepsilon$  analytically [429]. A non-hyperbolic equilibrium  $(\chi^*, 0)$  of the system with the

Hamiltonian (45) gives rise to a non-hyperbolic invariant  $(n-l)$ -torus of the system  $\tilde{X}$  only for the most values (in the sense of Lebesgue measure) of the perturbation parameter  $\varepsilon$ .

**Historical remarks** The case  $l = n - 1$  of ‘maximal’ resonance (where the tori in question are in fact circles) in [Theorem 16](#) was considered already by Poincaré (for a modern presentation see e.g. [77]). This classical result lies outside KAM Theory because it does not involve small divisors. The case of arbitrary  $l$  was examined only in 1989 (that is 35 years after KAM Theory was founded in Kolmogorov’s paper [251]!) by D.V. Treshchëv [429]. But Treshchëv treated only hyperbolic invariant  $(n-l)$ -tori. The hyperbolic  $l = 1$  case was also explored independently in subsequent papers [112,165,335,367,433,447]. Non-hyperbolic invariant tori in [Theorem 16](#) were first constructed in [113] for the case  $l = 1$  (note that if  $l = 1$  then the tori are either hyperbolic or elliptic). The general case of [Theorem 16](#)—an arbitrary  $l$  and arbitrary type of the invariant  $(n-l)$ -tori—was announced by Ch.-Q. Cheng and Sh. Wang [115,448]. In fact, Cheng and Wang [115,448] considered only the case where the eigenvalues of the matrix (44) are either real or purely imaginary (quadruplets  $\pm a \pm ib$  of complex eigenvalues were excluded). Finally, F. Cong, T. Küpper, Y. Li, and J. You [133], and somewhat later on Y. Li and Y. Yi [278,279], proved [Theorem 16](#) for arbitrary  $l$  and arbitrary type of the invariant  $(n-l)$ -tori (and arbitrary collections of eigenvalues). Some degenerate cases were examined in [205]. Another approach to constructing invariant  $(n-l)$ -tori with arbitrary  $l$  and arbitrary normal behaviour in the context of [Theorem 16](#) was proposed in [149]. In [185,186,189,190], elliptic, hyperbolic, and mixed type invariant tori in [Theorem 16](#) were obtained by a new method (the paper [186] is devoted to the case  $l = 1$  in the presence of some degeneracies). Thus, now we possess a complete picture of the destruction of resonant tori of integrable Hamiltonian systems under small perturbations.

The papers [112,113,133,165,185,186,189,190,205,278,279,335,367,429,433,447], cited above, studied the analytic situation, whereas the articles [115,149,448] dealt with finitely smooth systems. In the works [149,165,335,367,429,433], special attention was paid to the  $n$ -dimensional separatrix stable and unstable manifolds (‘whiskers’) of the hyperbolic invariant  $(n-l)$ -tori one looks for. Such ‘whiskers’ are of great importance in the Arnold diffusion mechanism (see Section 4.4.2).

[Theorem 16](#) admits reversible analogues [273,451]. The papers [273,451] consider an arbitrary number of resonance relations and arbitrary type of the tori. Some analogues of [Theorem 16](#) for symplectic diffeomorphisms are presented in [241].

### 8.3. Lower dimensional isotropic invariant tori

Our discussion of KAM Theory for lower dimensional isotropic invariant tori in Hamiltonian systems is based on the parametrized approach as expounded in detail for the dissipative case in Section 5.2. Here the words ‘lower dimensional’ mean that the dimensions of the tori are smaller than the number of degrees of freedom and the word ‘isotropic’ means that the restriction of the symplectic form to the tori in question vanishes.

We start with the Hamiltonian analogue of [Theorem 12](#). In Remark 2 after the formulation of this theorem in Section 5.2.1 we already pointed out that [Theorem 12](#) has

analogues for various structure preserving settings. Next we explain how one can get rid of external parameters and obtain a KAM theorem for lower dimensional invariant tori in individual Hamiltonian systems satisfying Rüssmann-like nondegeneracy conditions via the so-called Herman method already mentioned at the end of Section 5.2.2. Historical remarks will conclude the topic.

**8.3.1. The parametrized Hamiltonian KAM theorem.** Take  $M = \mathbb{T}^n \times P_{\text{act}} \times \mathbb{R}^{2m} = \{(x, y, z)\}$  as the phase space endowed by the symplectic form

$$\sigma = \sum_{j=1}^n dx_j \wedge dy_j + \sum_{j=1}^m dz_{j+m} \wedge dz_j,$$

where  $P_{\text{act}} \subseteq \mathbb{R}^n = \{y\}$  is an open and connected subset. Let also  $P_{\text{par}} \subseteq \mathbb{R}^s = \{\mu\}$ , an open subset, be the parameter space. Here the subscripts ‘act’ and ‘par’ are for ‘action’ and ‘parameter’, respectively. On  $M$ , consider a  $C^\infty$ -family of Hamiltonians

$$H = H_\mu(x, y, z) = E(y, \mu) + \frac{1}{2} \langle B(y, \mu)z, z \rangle + h(x, y, z, \mu), \quad (46)$$

where  $E : P_{\text{act}} \times P_{\text{par}} \rightarrow \mathbb{R}$  and  $B : P_{\text{act}} \times P_{\text{par}} \rightarrow gl(2m, \mathbb{R})$  are certain  $C^\infty$ -mappings, while  $h = O(|z|^3)$  with the O-estimate uniform in  $y$  and  $\mu$  (the uniformity in  $x$  is automatic due to the compactness of  $\mathbb{T}^n$ ). The  $2m \times 2m$  matrix  $B(y, \mu)$  is supposed to be symmetric for all the values of  $y$  and  $\mu$ . The corresponding family of vector fields  $X = X_{H_\mu}(x, y, z)$  is given by

$$\begin{aligned} \dot{x} &= \partial H_\mu / \partial y = \omega(y, \mu) + O(|z|^2) \\ \dot{y} &= -\partial H_\mu / \partial x = O(|z|^3) \\ \dot{z} &= \mathcal{J}(\partial H_\mu / \partial z) = \Omega(y, \mu)z + O(|z|^2), \end{aligned} \quad (47)$$

compare with (25), where

$$\omega(y, \mu) = \partial E(y, \mu) / \partial y, \quad \Omega(y, \mu) = \mathcal{J}B(y, \mu), \quad (48)$$

and  $\mathcal{J} \in GL(2m, \mathbb{R})$  denotes the ‘symplectic unit’:

$$\begin{pmatrix} m \times m \text{ zero matrix} & -(m \times m \text{ identity matrix}) \\ m \times m \text{ identity matrix} & m \times m \text{ zero matrix} \end{pmatrix}.$$

Note that  $\mathcal{J}^T = \mathcal{J}^{-1} = -\mathcal{J}$  where the superscript T means transpose.

The vector fields (47) are not integrable in the sense of Sections 4.2.1, 4.3.1 and 5.2.1 because we allow the remainder  $h$  to be  $x$ -dependent, but this is irrelevant for our purposes since the normal linear parts of (47) given by

$$\omega(y, \mu) \frac{\partial}{\partial x} + \Omega(y, \mu)z \frac{\partial}{\partial z}, \quad (49)$$

compare with (26), are integrable. In particular, on the invariant  $2n$ -dimensional surface  $\Pi = \{z = 0\} = \mathbb{T}^n \times P_{\text{act}} \times \{0\}$  we obtain Liouville integrable Hamiltonian dynamics (see Section 4.3.1). Indeed, the  $n$ -tori  $T_{y\mu} = \mathbb{T}^n \times \{y\} \times \{0\}$  are  $X_{H_\mu}$ -invariant for all  $\mu$  and the restriction of  $X_{H_\mu}$  to these tori is conditionally periodic (see Section 4.1), the corresponding frequency map  $\omega : P_{\text{act}} \times P_{\text{par}} \rightarrow \mathbb{R}^n$  being defined by (48). Note that the tori  $T_{y\mu}$  are Lagrangean within  $\Pi$  and are isotropic when considered as submanifolds of the whole phase space  $M$ .

As usual, we are interested in the fate of the tori  $\mathbb{T}^n \times \{y\} \times \{0\}$ ,  $y \in P_{\text{act}}$ , under small Hamiltonian  $C^\infty$ -perturbations  $\tilde{X}$  of  $X$ , for the ‘action’ variable  $y$  ranging near a certain point  $y_0 \in P_{\text{act}}$  and for the parameter values  $\mu$  near a certain point  $\mu_0 \in P_{\text{par}}$ . In the sequel, we mimic the exposition of Section 5.2.1.

**BHT nondegeneracy** As in the dissipative case treated in Section 5.2.1, the suitable BHT nondegeneracy condition for the unperturbed vector fields (47) in the present Hamiltonian set-up involves the frequencies  $\omega_1(y, \mu), \dots, \omega_n(y, \mu)$  of the tori  $T_{y\mu}$  (the internal frequencies) as well as the eigenvalues of the corresponding ‘normal’ matrices  $\Omega(y, \mu)$ . The latter matrices are Hamiltonian (= infinitesimally symplectic) in the sense that  $\Omega\mathcal{J} + \mathcal{J}\Omega^T = 0$  for each  $y$  and  $\mu$ :

$$\Omega\mathcal{J} + \mathcal{J}\Omega^T = \mathcal{J}B\mathcal{J} + \mathcal{J}B^T\mathcal{J}^T = \mathcal{J}B\mathcal{J} - \mathcal{J}B\mathcal{J} = 0.$$

Thus,  $\Omega^T = \mathcal{J}\Omega\mathcal{J} = -\mathcal{J}\Omega\mathcal{J}^{-1}$ , whence  $-\Omega$  has the same spectrum and the same Jordan structure as  $\Omega^T$  and, consequently, as  $\Omega$ . In particular, the eigenvalues of  $\Omega$  come in pairs  $\pm\lambda$ .

As in Section 5.2.1, we will confine ourselves to the case where all these eigenvalues are simple for  $(y, \mu)$  near  $(y_0, \mu_0)$ , then automatically  $\det \Omega(y, \mu) \neq 0$  for such  $(y, \mu)$ . Suppose that the eigenvalues of  $\Omega$  are given by

$$(\pm\delta_1, \dots, \pm\delta_{N_1}, \pm i\zeta_1, \dots, \pm i\zeta_{N_2}, \pm\alpha_1 \pm i\beta_1, \dots, \pm\alpha_{N_3} \pm i\beta_{N_3}) \quad (50)$$

with positive  $\delta_j, \zeta_j, \alpha_j, \beta_j$ , compare with (27). Note that  $N_1 + N_2 + 2N_3 = m$ . We call

$$(\zeta, \beta) = (\zeta_1, \dots, \zeta_{N_2}, \beta_1, \dots, \beta_{N_3})$$

the *normal frequencies* of the invariant torus. The map

$$\text{spec} : \Omega \mapsto (\delta, \zeta, \alpha, \beta) \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \times \mathbb{R}^{N_3} \times \mathbb{R}^{N_3},$$

compare with (28), parametrizes near  $\Omega(y_0, \mu_0)$  the orbit space of the group of  $2m \times 2m$  symplectic matrices (= the group of  $2m \times 2m$  matrices  $S$  subject to the equality  $S\mathcal{J}S^T = \mathcal{J}$ ) acting on the set of  $2m \times 2m$  Hamiltonian matrices. The BHT nondegeneracy condition, in this case, consists in the fact that the map

$$P_{\text{act}} \times P_{\text{par}} \ni (y, \mu) \mapsto (\omega \times (\text{spec} \circ \Omega))(y, \mu) \in \mathbb{R}^n \times \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \times \mathbb{R}^{N_3} \times \mathbb{R}^{N_3}, \quad (51)$$

compare with (29), is a submersion in an open neighbourhood  $A \subseteq (P_{\text{act}} \times P_{\text{par}})$  of  $(y_0, \mu_0)$  [76–78,236]. According to the Inverse Function Theorem [225,325,418], this condition is equivalent to saying that the derivative of the map (51) is surjective at  $y = y_0$ ,  $\mu = \mu_0$ . We then say that the family  $X = X_{H_\mu}$  is nondegenerate on the torus union  $\mathbb{T}^n \times \{0\} \times A = \bigcup_{(y,\mu) \in A} T_{y\mu}$ , recall that here  $0 \in \mathbb{R}^{2m}$ . Contrary to Section 5.2.1, we will not try to suppress extra parameters that may possibly occur because such a suppression could a priori affect not only  $\mu$  but the ‘action’ variable  $y$  as well. Without loss of generality the domain  $A$  can be assumed to have the form  $A = A_{\text{act}} \times A_{\text{par}}$  with  $A_{\text{act}} \subseteq P_{\text{act}}$  and  $A_{\text{par}} \subseteq P_{\text{par}}$ .

**Diophantine conditions** To formulate appropriate Diophantine conditions on the internal and normal frequencies, for  $\tau > n - 1$  and  $\gamma > 0$  we define the set of  $(\tau, \gamma)$ -Diophantine normal-internal frequency vectors by

$$\begin{aligned} D_{\tau,\gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3}) \\ = \left\{ (\omega, \zeta, \beta) \in \mathbb{R}^n \times \mathbb{R}^{N_2} \times \mathbb{R}^{N_3} \mid |\langle \omega, k \rangle + \langle \zeta, t \rangle + \langle \beta, \ell \rangle| \geq \gamma |k|^{-\tau}, \right. \\ \left. \text{for all } k \in \mathbb{Z}^n \setminus \{0\} \text{ and for all } t \in \mathbb{Z}^{N_2}, \ell \in \mathbb{Z}^{N_3} \text{ with } |t| + |\ell| \leq 2 \right\}, \end{aligned} \quad (52)$$

compare with (30). This is again a nowhere dense set of positive measure (for  $\gamma$  sufficiently small) that possesses the closed half-line property, see Sections 4.1.2 and 5.1.1; compare with [76–78,236,318,319]. Let  $\Gamma = (\omega \times (\text{spec} \circ \Omega))(A)$ , as in Section 5.2.1,

$$\Gamma_\gamma = \{(\omega, \delta, \zeta, \alpha, \beta) \in \Gamma \mid \text{dist}((\omega, \delta, \zeta, \alpha, \beta), \partial\Gamma) \geq \gamma\},$$

and

$$D_{\tau,\gamma}(\Gamma_\gamma) = \Gamma_\gamma \cap \left( D_{\tau,\gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3}) \times \mathbb{R}^{N_1} \{\delta\} \times \mathbb{R}^{N_3} \{\alpha\} \right).$$

The further definition of  $D_{\tau,\gamma}(A_\gamma) \subset A$  is now obvious, compare with Sections 4.2.1 and 4.3.1. Again, note that the closed half-lines of  $D_{\tau,\gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3})$  now turn into closed linear half-spaces of dimension  $1 + N_1 + N_3$  and that these geometrical structures, up to a diffeomorphism, are inherited by the perturbations.

**THEOREM 17** (Parametrized KAM – Hamiltonian Isotropic Case [71,78,227,236]). *Let  $n \geq 2$ . Let the  $C^\infty$ -family  $X = X_{H_\mu}(x, y, z)$  of vector fields (47) be BHT nondegenerate on  $\mathbb{T}^n \times \{0\} \times A$ , with  $A = A_{\text{act}} \times A_{\text{par}} \subseteq (P_{\text{act}} \times P_{\text{par}})$  open. Then, for  $\gamma > 0$  sufficiently small, there exists a neighbourhood  $\mathcal{O}$  of  $X$  in the  $C^\infty$ -topology, such that for any perturbed family of Hamiltonian fields  $\tilde{X} \in \mathcal{O}$  there exists a  $C^\infty$ -mapping*

$$\Phi : \mathbb{T}^n \times A_{\text{act}} \times \mathbb{R}^{2m} \times A_{\text{par}} \times A_{\text{act}} \rightarrow \mathbb{T}^n \times A_{\text{act}} \times \mathbb{R}^{2m} \times A_{\text{par}},$$

defined near  $\mathbb{T}^n \times A_{\text{act}} \times \{0\} \times A_{\text{par}} \times A_{\text{act}}$ , with the following properties:

1. For each  $y' \in A_{\text{act}}$ , the mapping

$$\Phi[y'] : \mathbb{T}^n \times A_{\text{act}} \times \mathbb{R}^{2m} \times A_{\text{par}} \rightarrow \mathbb{T}^n \times A_{\text{act}} \times \mathbb{R}^{2m} \times A_{\text{par}}$$

defined as

$$\Phi[y'](x, y, z, \mu) = \Phi(x, y, z, \mu, y')$$

is a  $C^\infty$ -near the identity map preserving projections to the parameter space  $P_{\text{par}}$ .

2. For each  $y' \in A_{\text{act}}$ , the image of the  $X$ -invariant torus union

$$\mathbb{T}^n \times \{y'\} \times \{0\} \times \{\mu \in A_{\text{par}} \mid (y', \mu) \in D_{\tau, \gamma}(A_\gamma)\}$$

under  $\Phi[y']$  is  $\tilde{X}$ -invariant, and the restricted map

$$\widehat{\Phi}[y'] = \Phi[y']|_{\mathbb{T}^n \times \{y'\} \times \{0\} \times \{\mu \in A_{\text{par}} \mid (y', \mu) \in D_{\tau, \gamma}(A_\gamma)\}}$$

conjugates  $X$  to  $\tilde{X}$ , that is

$$\widehat{\Phi}[y']_* X = \tilde{X}.$$

3. Moreover,  $\Phi[y']$  preserves the normal linear behaviour of the tori  $T_{y'\mu}$  for  $(y', \mu) \in D_{\tau, \gamma}(A_\gamma)$  with respect to  $X$  (see item 3 of [Theorem 12](#) for an explanation).

REMARKS. 1. Roughly speaking, [Theorem 17](#) relates to [Theorem 12](#) of Section 5.2.1 as the Lagrangean KAM [Theorem 6](#) of Section 4.3.1 does to the normally hyperbolic KAM [Theorem 5](#) of Section 4.2.1 (and as the twist [Theorem 3](#) of Section 3.2.1 does to the circle-map [Theorem 2](#) of Section 3.1.2). The conclusion of [Theorem 17](#) first of all expresses that the family  $X$  (47) is quasi-periodically stable on the union

$$\bigcup_{(y, \mu) \in D_{\tau, \gamma}(A_\gamma)} T_{y\mu} = \mathbb{T}^n \times \{0\} \times D_{\tau, \gamma}(A_\gamma)$$

of Diophantine quasi-periodic invariant  $n$ -tori [76–78, 236]; compare with Sections 3.1.3, 4.2.1, 4.3.1 and 5.2.1 above. Taking item 3 of [Theorem 17](#) into account, we also speak of *normal linear stability* of this torus union; compare with Remark 1 after [Theorem 12](#).

2. Introducing the additional parameter  $y' \in A_{\text{act}}$  is called a *localization procedure*. This trick is explained in [71, 78, 227, 236] in detail, see also the beginning of Section 5.2.2. Note that for some particular values of  $y'$ , the Lebesgue measure  $\text{meas}[y']$  of the set  $\{\mu \in A_{\text{par}} \mid (y', \mu) \in D_{\tau, \gamma}(A_\gamma)\}$  can be zero and this set can even be empty, however small  $\gamma$  is. Nevertheless,

$$\int_{A_{\text{act}}} \text{meas}[y'] dy'$$

tends to the measure of  $A$  as  $\gamma \downarrow 0$  according to the Fubini theorem.

3. The maps  $\Phi[y']$  (to be more precise, their phase space components) are, in general, not symplectic (compare with [Theorem 6](#)).
4. As we already indicated in Remark 3 after [Theorem 12](#), the assumption of a simple spectrum of  $\Omega(y_0, \mu_0)$  can be dropped, the key references for this in the present Hamiltonian setting being [71, 227].



The unperturbed  $X$ -invariant tori  $T_{y\mu}$ ,  $(y, \mu) \in A$ , as well as the perturbed  $\tilde{X}$ -invariant tori  $\Phi[y](T_{y\mu})$ ,  $(y, \mu) \in D_{\tau, \gamma}(A_\gamma)$ , are said to be

1. *elliptic*, if all the eigenvalues of  $\Omega(y, \mu)$  are purely imaginary (i.e.  $N_1 = N_3 = 0$ ,  $N_2 = m$ );
2. *hyperbolic*, if all the eigenvalues of  $\Omega(y, \mu)$  lie outside the imaginary axis (i.e.  $N_2 = 0$ ,  $N_1 + 2N_3 = m$ );
3. *of mixed type*, if some eigenvalues of  $\Omega(y, \mu)$  are purely imaginary and some are not (i.e.  $N_2 > 0$  and  $N_1 + N_3 > 0$ ),

compare with Section 8.2. The relation between this definition of hyperbolicity of lower dimensional invariant tori in Hamiltonian flows and the general concept of hyperbolic invariant manifolds in dynamical systems (as presented in e.g. [126,175,226,434]) is discussed in detail in [36]. It is important to emphasize that each of these three types of normal behaviour of invariant tori is an *open* property: if a certain torus is elliptic (hyperbolic, of mixed type), so are all the nearby tori of all the nearby systems.

**8.3.2. Lower dimensional tori in individual Hamiltonian systems.** Quasi-periodic and normal linear stability of invariant  $n$ -tori  $T_{y\mu}$  in Theorem 17, for  $(y, \mu) \in D_{\tau, \gamma}(A_\gamma)$ , requires a lot of external parameters: the parameter space  $P_{\text{par}}$  should be of dimension  $s \geq m$  to make it possible for the map (51) to be a submersion. Now we will show how one can get rid of *all* these parameters and deduce a KAM statement about lower dimensional invariant tori in *individual* Hamiltonian systems.

We consider the same symplectic phase space  $M = \{(x, y, z)\}$  as in Section 8.3.1 (but with the range domain for  $y$  denoted by  $P$  rather than by  $P_{\text{act}}$ ) and a  $C^\infty$ -Hamiltonian  $H$  on  $M$  of the same form as (46), but without the parameter  $\mu$ :

$$H = H(x, y, z) = E(y) + \frac{1}{2} \langle B(y)z, z \rangle + h(x, y, z), \quad (53)$$

where  $E : P \rightarrow \mathbb{R}$  and  $B : P \rightarrow gl(2m, \mathbb{R})$  are certain  $C^\infty$ -mappings, while  $h = O(|z|^3)$  with a  $y$ -uniform  $O$ -estimate. The  $2m \times 2m$  matrix  $B(y)$  is supposed to be symmetric for all the values of  $y$ . The corresponding vector field  $X = X_H(x, y, z)$  is given by

$$\begin{aligned} \dot{x} &= \partial H / \partial y = \omega(y) + O(|z|^2) \\ \dot{y} &= -\partial H / \partial x = O(|z|^3) \\ \dot{z} &= \mathcal{J}(\partial H / \partial z) = \Omega(y)z + O(|z|^2), \end{aligned} \quad (54)$$

compare with (47), where

$$\omega(y) = \partial E(y) / \partial y, \quad \Omega(y) = \mathcal{J}B(y),$$

compare with (48). The normal linear part of (54) is

$$\omega(y) \frac{\partial}{\partial x} + \Omega(y)z \frac{\partial}{\partial z}, \quad (55)$$

compare with (49). Our interest is Diophantine quasi-periodic invariant  $n$ -tori of small Hamiltonian  $C^\infty$ -perturbations  $\tilde{X}$  of  $X$  near the surface  $\Pi = \{z = 0\}$ , to be more precise, near the invariant tori  $\mathbb{T}^n \times \{y\} \times \{0\}$  of  $X$ ,  $y \in P$ .

**Diophantine stability** To this end, assume again that for  $y$  ranging in an open (and bounded) subset  $A \subseteq P$ , the spectrum of the Hamiltonian matrix  $\Omega$  is simple and given by (50) with positive  $\delta_j, \zeta_j, \alpha_j, \beta_j$ . Suppose also that the map

$$\omega \times \zeta \times \beta : P \rightarrow \mathbb{R}^n \times \mathbb{R}^{N_2} \times \mathbb{R}^{N_3}; \quad y \mapsto (\omega(y), \zeta(y), \beta(y)) \quad (56)$$

possesses the following property of *Diophantine stability*, related to the theory of Diophantine approximations on submanifolds of Euclidean spaces, see Section 4.3.2.

**DEFINITION 18** (Diophantine Stability). The map (56) is said to be *Diophantine stable* if there exists a neighbourhood  $\mathcal{Q}$  of this map in the  $C^\infty$ -topology such that for any sufficiently large positive  $\tau$  and for any perturbed map

$$\tilde{\omega} \times \tilde{\zeta} \times \tilde{\beta} \in \mathcal{Q}; \quad y \mapsto (\tilde{\omega}(y), \tilde{\zeta}(y), \tilde{\beta}(y)),$$

the Lebesgue measure of the set

$$\left\{ y \in A \mid (\tilde{\omega}(y), \tilde{\zeta}(y), \tilde{\beta}(y)) \notin D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3}) \right\}$$

tends to 0 as  $\gamma \downarrow 0$  uniformly in  $\tilde{\omega} \times \tilde{\zeta} \times \tilde{\beta} \in \mathcal{Q}$ . Recall that the set  $D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3})$  of  $(\tau, \gamma)$ -Diophantine normal-internal frequency vectors is defined in (30) and (52).

The explicit conditions on (56) sufficient for Diophantine stability are known [76,77,405] and can be formulated in terms of the partial derivatives of  $\omega, \zeta$ , and  $\beta$  of all the orders from 0 to some positive integer  $Q$ . In particular, these conditions include Rüssmann nondegeneracy of  $\omega$  discussed in Section 4.3.2. We will not reproduce these conditions here because they are somewhat cumbersome and not illuminative, but we would like to note the following technical error persistent in many formulas in the works [76,77,405] as well as in the papers [399,401,403,404,407] and in the Russian 2002 edition of the book [25] (due to an aberration in some calculations of the second author of the present survey): *all* the expressions of the form

$$\sum_{|q|=J} D^q f(a) b^q \quad \text{or} \quad \sum_{|q|=J} \langle D^q f(a), g \rangle b^q$$

in *all* these works should be replaced by

$$J! \sum_{|q|=J} D^q f(a) \frac{b^q}{q!} \quad \text{and} \quad J! \sum_{|q|=J} \langle D^q f(a), g \rangle \frac{b^q}{q!},$$

respectively. Here  $J$  is a non-negative integer,  $f : \mathbb{R}^\nu \rightarrow \mathbb{R}^l$  is a  $C^J$ -mapping,  $a \in \mathbb{R}^\nu$ ,  $b \in \mathbb{R}^\nu$ ,  $g \in \mathbb{R}^l$ , all the components of the vector  $q \in \mathbb{Z}^\nu$  are non-negative, and the standard

multi-index notation is used:

$$D^q f(a) = \frac{\partial^{q_1+\dots+q_v}}{\partial a_1^{q_1} \dots \partial a_v^{q_v}} f(a), \quad b^q = b_1^{q_1} \dots b_v^{q_v}, \quad q! = q_1! \dots q_v!.$$

**THEOREM 19** (Hamiltonian Lower Dimensional KAM without Parameters). *Let  $n \geq 2$ . Let the map (56) corresponding to the  $C^\infty$ -vector field  $X = X_H(x, y, z)$  of the form (54) be Diophantine stable on an open and bounded subset  $A \subseteq P$ . Then, for any  $\epsilon > 0$ , there exists a neighbourhood  $\mathcal{O}$  of  $X$  in the  $C^\infty$ -topology, such that for any perturbed Hamiltonian field  $\tilde{X} \in \mathcal{O}$  there exist*

- a subset  $A^\sharp \subset A$ ,
- a vector field

$$X_{\text{lin}}^\sharp(x, y, z) = \omega^\sharp(y) \frac{\partial}{\partial x} + \Omega^\sharp(y) z \frac{\partial}{\partial z}, \quad (57)$$

- compare with (55), with  $C^\infty$ -maps  $\omega^\sharp : A \rightarrow \mathbb{R}^n$  and  $\Omega^\sharp : A \rightarrow gl(2m, \mathbb{R})$ ,
- and a  $C^\infty$ -mapping  $\Phi : \mathbb{T}^n \times A \times \mathbb{R}^{2m} \times A \rightarrow \mathbb{T}^n \times A \times \mathbb{R}^{2m}$ , defined near  $\mathbb{T}^n \times A \times \{0\} \times A$ ,

that possess the following properties.

1. The Lebesgue measure of  $A \setminus A^\sharp$  is less than  $\epsilon$ .
2. The maps  $\omega^\sharp$  and  $\Omega^\sharp$  are  $C^\infty$ -close to the maps  $\omega$  and  $\Omega$ , respectively. Moreover, the  $2m \times 2m$  matrix  $\Omega^\sharp(y)$  is Hamiltonian for each  $y \in A$  ( $\Omega^\sharp \mathcal{J} + \mathcal{J} \Omega^{\sharp T} \equiv 0$ ).
3. For each  $y' \in A$ , the mapping  $\Phi[y'] : \mathbb{T}^n \times A \times \mathbb{R}^{2m} \rightarrow \mathbb{T}^n \times A \times \mathbb{R}^{2m}$  defined as  $\Phi[y'](x, y, z) = \Phi(x, y, z, y')$  is a  $C^\infty$ -near the identity map.
4. For each  $y' \in A^\sharp$ , the image of the  $X$ - and  $X_{\text{lin}}^\sharp$ -invariant torus  $\mathbb{T}^n \times \{y'\} \times \{0\}$  under  $\Phi[y']$  is  $\tilde{X}$ -invariant, and the restricted map  $\widehat{\Phi}[y'] = \Phi[y']|_{\mathbb{T}^n \times \{y'\} \times \{0\}}$  conjugates  $X_{\text{lin}}^\sharp$  to  $\tilde{X}$ , that is

$$\widehat{\Phi}[y']_* X_{\text{lin}}^\sharp = \tilde{X}.$$

5. Finally, for each  $y' \in A^\sharp$ ,  $\Phi[y']$  preserves the normal linear behaviour of the torus  $\mathbb{T}^n \times \{y'\} \times \{0\}$  with respect to  $X_{\text{lin}}^\sharp$ .

This theorem is an almost immediate corollary of [Theorem 17](#).

**SKETCH OF THE PROOF OF THEOREM 19.** For  $y \in A$ , one can easily construct  $C^\infty$ -mappings  $E^\dagger = E^\dagger(y, \mu)$  and  $B^\dagger = B^\dagger(y, \mu)$ , with  $\mu$  ranging near the origin 0 of  $\mathbb{R}^s$  ( $s$  being sufficiently large), such that

- $\sim E^\dagger(y, 0) \equiv E(y)$  and  $B^\dagger(y, 0) \equiv B(y)$ ;
- $\sim B^\dagger(y, \mu)$  is symmetric for all the values of  $y$  and  $\mu$ ;
- $\sim$  the  $C^\infty$ -family of Hamiltonian vector fields  $X_{H_\mu^\dagger}^\dagger$  on  $\mathbb{T}^n \times A \times \mathbb{R}^{2m}$  afforded by the Hamiltonians

$$H^\dagger = H_\mu^\dagger(x, y, z) = E^\dagger(y, \mu) + \frac{1}{2} \langle B^\dagger(y, \mu) z, z \rangle + h(x, y, z),$$

compare with (46) and (53), is BHT nondegenerate on the torus union  $\bigcup_{\mu \text{ near } 0 \in \mathbb{R}^s} \mathbb{T}^n \times A \times \{0\}$ .

That the spectrum of  $\Omega(y)$  is simple for  $y \in A$  is sufficient for this, independently of whether the map (56) is Diophantine stable or not. Note that  $H_0^\dagger \equiv H$ ,  $X_{H_0^\dagger}^\dagger = X_H$ , and for the maps

$$\omega^\dagger(y, \mu) = \partial E^\dagger(y, \mu) / \partial y, \quad \Omega^\dagger(y, \mu) = \mathcal{J} B^\dagger(y, \mu),$$

compare with (48), one has  $\omega^\dagger(y, 0) \equiv \omega(y)$ ,  $\Omega^\dagger(y, 0) \equiv \Omega(y)$ . The eigenvalues of  $\Omega^\dagger$  are given by

$$\left( \pm \delta_1^\dagger, \dots, \pm \delta_{N_1}^\dagger, \pm i \zeta_1^\dagger, \dots, \pm i \zeta_{N_2}^\dagger, \pm \alpha_1^\dagger \pm i \beta_1^\dagger, \dots, \pm \alpha_{N_3}^\dagger \pm i \beta_{N_3}^\dagger \right)$$

with positive  $\delta_j^\dagger, \zeta_j^\dagger, \alpha_j^\dagger, \beta_j^\dagger$ , compare with (50), and  $\delta_j^\dagger(y, 0) \equiv \delta_j(y)$ ,  $\zeta_j^\dagger(y, 0) \equiv \zeta_j(y)$ ,  $\alpha_j^\dagger(y, 0) \equiv \alpha_j(y)$ ,  $\beta_j^\dagger(y, 0) \equiv \beta_j(y)$ .

Fix positive  $\tau$  sufficiently large and apply Theorem 17 to the families of Hamiltonian vector fields  $X^\dagger = X_{H_\mu^\dagger}^\dagger$  and  $X_{H_\mu^\dagger}^\dagger + \tilde{X} - X_H$ . For  $\gamma > 0$  sufficiently small, and for  $\tilde{X}$  sufficiently  $C^\infty$ -close to  $X = X_H$ , we obtain a  $C^\infty$ -mapping

$$\Phi^\dagger = \Phi^\dagger(x, y, z, \mu, y') = (\Psi[\mu, y'](x, y, z), \Upsilon[\mu, y']), \quad (58)$$

where  $x \in \mathbb{T}^n$ ,  $y \in A$ ,  $y' \in A$ , the variable  $z$  ranges near  $0 \in \mathbb{R}^{2m}$ , the parameter  $\mu$  ranges near  $0 \in \mathbb{R}^s$ , and  $\Psi[\mu, y'](x, y, z) \in \mathbb{T}^n \times A \times \mathbb{R}^{2m}$ ,  $\Upsilon[\mu, y'] \in \mathbb{R}^s$ . The mapping (58) possesses the following properties:

1. For each  $y' \in A$ , the mapping  $(x, y, z, \mu) \mapsto \Phi^\dagger(x, y, z, \mu, y')$  is  $C^\infty$ -near the identity (the fact that the  $\mu$ -component  $\Upsilon[\mu, y']$  of this mapping does not depend on the phase space variables  $x, y, z$  just expresses the preservation of projections to the  $\mu$ -space).
2. For each  $y'_0 \in A$  and  $\mu_0$  near  $0 \in \mathbb{R}^s$  such that  $y'_0$  is not too close to the boundary  $\partial A$  and

$$\left( \omega^\dagger(y'_0, \mu_0), \zeta^\dagger(y'_0, \mu_0), \beta^\dagger(y'_0, \mu_0) \right) \in D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3}),$$

the  $n$ -torus

$$\Psi[\mu_0, y'_0](\mathbb{T}^n \times \{y'_0\} \times \{0\}) \subset (\mathbb{T}^n \times A \times \mathbb{R}^{2m}) \quad (59)$$

is invariant under the vector field  $X_{H_{\Upsilon[\mu_0, y'_0]}^\dagger}^\dagger + \tilde{X} - X_H$ .

3. Moreover, the restriction of  $\Psi[\mu_0, y'_0]$  to the torus  $\mathbb{T}^n \times \{y'_0\} \times \{0\}$  conjugates  $X_{H_{\mu_0}^\dagger}^\dagger$

to  $X_{H_{\Upsilon[\mu_0, y'_0]}^\dagger}^\dagger + \tilde{X} - X_H$ .

4. Moreover,  $\Psi[\mu_0, y'_0]$  preserves the normal linear behaviour of the torus  $\mathbb{T}^n \times \{y'_0\} \times \{0\}$  with respect to  $X_{H_{\mu_0}}^\dagger$ .

One can solve the equation  $\Upsilon[\mu, y'] = 0$  with respect to  $\mu$  and obtain  $\mu = \xi(y')$  where the function  $\xi : A \rightarrow \mathbb{R}^s$  is  $C^\infty$ -small. Set  $\mu_0 = \xi(y'_0)$  in the above list of the properties of mapping (58). For  $\mu_0 = \xi(y'_0)$ , the torus (59) is  $\tilde{X}$ -invariant. Since the map (56) is Diophantine stable, the measure of the set  $A^\parallel$  of points  $y' \in A$  such that

$$\left( \omega^\dagger(y', \xi(y')), \zeta^\dagger(y', \xi(y')), \beta^\dagger(y', \xi(y')) \right) \in D_{\tau, \gamma}(\mathbb{R}^n; \mathbb{R}^{N_2+N_3})$$

tends to the measure of  $A$  as  $\gamma \downarrow 0$ . Now it suffices to set

$$\begin{aligned} A^\sharp &= A^\parallel \setminus (\text{a narrow neighbourhood of } \partial A), \\ \omega^\sharp(y) &= \omega^\dagger(y, \xi(y)), \quad \Omega^\sharp(y) = \Omega^\dagger(y, \xi(y)), \\ \Phi(x, y, z, y') &= \Psi[\xi(y'), y'](x, y, z). \end{aligned}$$

The proof is completed.  $\square$

The central idea of this proof is artificially introducing parameters into the system (or adding extra parameters) to achieve BHT nondegeneracy and subsequently eliminating these parameters. This idea was proposed by M.R. Herman in his talk at an international conference on dynamical systems in Lyons in 1990, where Herman gave an exceedingly simple proof of the fact that Rüssmann nondegeneracy is sufficient for the presence of many perturbed Lagrangean tori. Herman did not publish his proof, although his ‘parameter reduction’ method was used in [463] in a problem outside Hamiltonian mechanics (that problem concerned the so-called ‘vertically translated’  $n$ -tori in  $\mathbb{T}^n \times \mathbb{R}$ ). In the mid 1990s, the authors of the present survey together with G.B. Huitema applied systematically Herman’s method (to be treated as a special tool within Parametrized KAM Theory) to various KAM contexts, including Lagrangean [76,77,397,400] and lower dimensional [76,77,405] invariant tori in Hamiltonian systems, as well as invariant tori in reversible [76,77,399], volume preserving, and dissipative [76,77] systems. In [410], Herman’s idea is used to construct invariant tori in quasi-periodic non-autonomous perturbations of dynamical systems preserving various structures. In [409,411], the same method is exploited to study the so-called partial preservation of frequencies (as well as of eigenvalues of the ‘normal’ matrices in [411]) in KAM Theory explored previously in [128,274,275,280]. Some similar approaches were used by A.A. Kubichka, Yu.V. Loveĭkin and I.O. Parasyuk [255,290,291,293] (see also their works [254,292,347,348] and J. Féjoz’ paper [173]).

REMARKS. 1. The conclusion of Theorem 19 expresses the fact that any Hamiltonian system  $\tilde{X}$  sufficiently close to  $X$  admits a Whitney differentiable family of invariant  $n$ -tori close to the unperturbed tori  $\mathbb{T}^n \times \{y\} \times \{0\}$  and carrying quasi-periodic motions. However, there is, in general, *no* connection between the frequencies and the normal behaviour of the perturbed  $\tilde{X}$ -invariant tori, on the one hand, and the frequencies and the normal behaviour of the unperturbed  $X$ -invariant tori, on the other hand (this is typical for all the KAM theorems with Rüssmann-like nondegeneracy conditions, compare with Section 4.3.2). The frequency vector and

‘normal’ matrix of every perturbed torus are equal, respectively, to  $\omega^\sharp(y)$  and  $\Omega^\sharp(y)$  for some  $y \in A^\sharp \subset A$  rather than to  $\omega(y)$  and  $\Omega(y)$ . Therefore, in the framework of Theorem 19, we can speak neither of quasi-periodic stability of the unperturbed tori nor of normal linear stability. The unperturbed tori in Theorem 19 do not persist under small Hamiltonian perturbations, one may just assert the existence of many invariant tori in a perturbed system.

2. Of course, the concept of Diophantine stability and all the setting of Theorem 19 can be carried over to the case where the unperturbed system already depends on some external parameters [76,77].
3. Generically, the mapping  $y \mapsto \omega^\sharp(y) = \omega^\sharp(y, \xi(y))$  in Theorem 19 cannot be represented, even locally, as the frequency map of any integrable Hamiltonian system with  $n$  degrees of freedom: the 1-form  $\omega^\sharp(y) dy = \sum_{j=1}^n \omega_j^\sharp(y) dy_j$  is not closed. Consequently, the vector field (57) is, in general, not Hamiltonian.
4. As in Theorem 17, the maps  $\Phi[y']$  in Theorem 19 are generically not symplectic.
5. The assumption of a simple spectrum of  $\Omega(y)$  in Theorem 19 can also be omitted, see [281,455,468]. There are versions of this theorem where  $\det \Omega(y)$  is allowed to vanish [281,467].

**8.3.3. Historical remarks.** Studies of lower dimensional isotropic invariant  $n$ -tori in Hamiltonian systems were started by V.K. Melnikov [309,310] and J.K. Moser [318,319] in the mid 1960s. Since then, this topic (in the context where the unperturbed system possesses, for each value of the external parameter if the latter is present, a  $2n$ -dimensional surface foliated into unperturbed invariant  $n$ -tori) has attracted a wide attention, the most important references probably being [30, 42,71,77,78,99,124,134,136,163,200,205,227,236,242,243,274,280,281,355,378,439,449, 455,456,458,467,468,471,473,474], see also [25,36,63,76,120,121,123,173,212,233,244, 337,368,403,405,409–411,430,459,470]. These works assume quite different nondegeneracy and nonresonance conditions imposed on the mappings  $\omega$  and  $\Omega$ . For instance, Rüssmann-like nondegeneracy conditions for  $\omega$  are used in [25,76,77,134,136,274,280, 378,403,405,409–411,449,473,474]. Resonances between internal and normal frequencies of elliptic tori are considered in the papers [42,449,456,458,459] (recall that other dynamical settings where normal-internal resonances come into play are examined in e.g. [65,66, 80,188,384,443], see Sections 6 and 7).

As a rule, non-hyperbolic lower dimensional invariant tori in KAM Theory turn out to be much harder to construct than the hyperbolic ones, compare with Section 8.2. For example, first general preservation theorems for hyperbolic lower dimensional tori were obtained in 1973 and 1974 [30,200], and those for elliptic lower dimensional tori only 15 years later [163,355]. Theorem 17 treats all the nondegenerate types of the normal behaviour of invariant tori in a unified way, and Herman’s technique exploited in the proof of Theorem 19 enables one to explore non-hyperbolic lower dimensional tori in individual systems as easily as hyperbolic ones.

By the way, in the hyperbolic case, the Centre Manifold reduction allows one to deduce Theorem 19 from the Lagrangean KAM theorem (with Rüssmann nondegeneracy conditions) in the finitely smooth setting, compare with Section 4.2. For Kolmogorov nondegeneracy, this approach was mentioned in [200] and discussed in detail in [233].

Finitely differentiable Hamiltonian KAM Theory for elliptic lower dimensional tori has been developed in [124].

All the unperturbed as well as perturbed invariant tori in Theorems 17 and 19 are *reducible*, i.e. the variational equation along each of these tori can be reduced to a form with constant coefficients, see Section 5.2.3. The reducibility of the unperturbed tori is expressed by the fact that the matrices  $B$  in (46) and (53) and, consequently, the matrices  $\Omega = \mathcal{J}B$  in (47) and (54), are  $x$ -independent. However, the Hamiltonian KAM Theory for hyperbolic lower dimensional invariant tori can be generalized to non-reducible perturbed [200,233] and even unperturbed [134,136,280,471] tori. In certain cases such a generalization can also be obtained for elliptic lower dimensional perturbed invariant tori [42]. The separatrix stable and unstable manifolds ('whiskers') of the non-reducible perturbed hyperbolic tori are constructed in [200,471].

'Exponential condensation', as described in Section 8.1, also takes place at least in some lower dimensional settings in the analytic category [242,439].

For the reversible counterpart to the Hamiltonian lower dimensional KAM Theory, see [58,75,77,273,359,387,391,393,396,399,404,457] as well as [57,71,76,390,392,409–411]. These works also use quite different nondegeneracy and nonresonance conditions on the unperturbed systems. For instance, Rüssmann-like nondegeneracy conditions for the unperturbed frequency map are assumed in [76,77,399,404,409–411].

#### 8.4. Excitation of elliptic normal modes

As was already pointed out, the hyperbolic case of the lower dimensional Hamiltonian KAM settings treated in Section 8.3, where the  $2m \times 2m$  matrices  $\Omega$  in (47) and (54) have no purely imaginary eigenvalues ( $N_2 = 0$ ), is generally believed to be much simpler than the non-hyperbolic case. However, consider more closely the latter case, where the purely imaginary eigenvalues  $\pm i\zeta_1, \dots, \pm i\zeta_{N_2}$  of the matrices  $\Omega$  do exist ( $N_2 > 0$ ). These eigenvalues are sometimes called the *elliptic normal modes* of the unperturbed invariant  $n$ -tori  $\mathbb{T}^n \times \{y\} \times \{0\}$  and yield the problem of the occurrence of Diophantine quasi-periodic invariant tori of dimensions from  $n + 1$  to  $n + N_2$  near the surface  $\Pi = \{z = 0\}$  in the unperturbed system  $X = X_H$  (or in the unperturbed family of systems  $X = X_{H_\mu}$ ) as well as in perturbed systems  $\tilde{X}$ . If such tori of dimensions  $n + 1, \dots, n + N_2$  exist, one sometimes says that the elliptic normal modes of the unperturbed  $n$ -tori  $\mathbb{T}^n \times \{y\} \times \{0\}$  *excite*.

REMARKS. 1. Here we also suppose that  $n \geq 2$ . The presence of quasi-periodic invariant tori of dimensions  $\kappa$  from 2 to  $n + m$  (the number of degrees of freedom) near non-hyperbolic equilibria ( $n = 0$ ) or periodic trajectories ( $n = 1$ ) also belongs to the realm of KAM Theory, of course, but this question is much easier and more 'conventional', and one usually does not speak of excitation of elliptic normal modes in this set-up. Instead, this topic is referred to as the 'local' Lagrangean KAM Theory (for  $\kappa = n + m$ ) or the 'local' lower dimensional KAM Theory (for  $\kappa < n + m$ ). By the way, 'exponential condensation' of invariant  $m$ -tori near elliptic equilibria of analytic Hamiltonian systems with  $m$  degrees of freedom is established in [152,153].

2. In the Hamiltonian and reversible settings, the occurrence of purely imaginary normal eigenvalues is open (see the end of Section 8.3.1). In the general dissipative set-up, partial normal ellipticity is a codimension 1 phenomenon, meaning that it has a persistent occurrence in generic 1-parameter families of systems, where it then takes place for isolated values of the parameter. The dissipative analogue of the present Hamiltonian discussion for non-hyperbolic invariant  $n$ -tori therefore is the branching off of  $(n + 1)$ -tori in a quasi-periodic Hopf bifurcation, as described in Section 6. Here we also keep track of the internal and normal frequencies of the  $n$ -torus, which presupposes the presence of sufficiently many parameters. In contrast, when examining excitation of elliptic normal modes in the Hamiltonian setting, we do not study any metamorphoses in the phase portrait of the system as the external parameter varies (even if the latter is present). In fact, this excitation can well occur in an individual system in the absence of any parameters.

It turns out that under certain nondegeneracy and nonresonance conditions, invariant  $\kappa$ -tori of all the dimensions in the range  $n + 1 \leq \kappa \leq n + N_2$  do exist in the unperturbed system(s) as well as in perturbed systems within the framework of the lower dimensional Hamiltonian KAM Theory. We shall not formulate the corresponding theorems even vaguely and will confine ourselves to the relevant references. Up to now, the phenomenon of excitation of elliptic normal modes has been explored for analytic Hamiltonian systems only (although it undoubtedly takes place for  $C^\infty$ - and finitely smooth systems as well). The first excitation results were obtained in 1962 and 1963 by V.I. Arnold [12,15] who considered the particular case  $\kappa - n = N_2 = m$ . This case was recently revisited by M.R. Herman, see [105,173]. A.D. Bruno [99] examined the general case of arbitrary  $N_2 \leq m$  and  $\kappa \leq n + N_2$  and constructed analytic families of invariant  $\kappa$ -tori. General theorems describing Whitney differentiable families of invariant  $\kappa$ -tori (for arbitrary  $N_2$  and  $\kappa$ ) were proven in [77,401] and independently in [243,439] (see also [25,244]). Analytic families of tori found by Bruno are subfamilies of these Whitney smooth families. Å. Jorba and J. Villanueva [243,244,439] also established the ‘exponential condensation’ of invariant  $\kappa$ -tori (which, of course, is impossible in the finitely smooth and  $C^\infty$ -categories). Various versions of the excitation theorem have been surveyed in detail in the review [403].

The phenomenon of the excitation of elliptic normal modes is also known in reversible [77,359,394,395,399,404] and volume preserving [407] set-ups. The works [77, 394,399,404] consider the excitation of elliptic normal modes in reversible flows and the papers [359,395,399], in reversible diffeomorphisms.

### 8.5. Higher dimensional coisotropic invariant tori

Apart from the KAM Theory for lower dimensional isotropic invariant tori in Hamiltonian systems we discussed in Section 8.3, there exists the somewhat ‘dual’ theory for higher dimensional coisotropic invariant tori. The words ‘higher dimensional’ mean that the dimensions of the tori are larger than the number of degrees of freedom and the word ‘coisotropic’ means that the tangent space  $T_p\mathcal{T}$  to a torus  $\mathcal{T}$  at any point  $p \in \mathcal{T}$  contains the skew-orthogonal complement of  $T_p\mathcal{T}$  with respect to the symplectic form (in the isotropic case, the space  $T_p\mathcal{T}$  is contained in its skew-orthogonal complement).



The observation crucial for the higher dimensional Hamiltonian KAM Theory was made by M.R. Herman in 1988.

LEMMA 20 (Automatic Isotropy [221,222]). *Any quasi-periodic invariant torus of a Hamiltonian system is isotropic provided that the corresponding symplectic form is exact.*

In fact, Herman [221,222] proved this lemma for a particular case of invariant  $n$ -tori of symplectic diffeomorphisms of  $2n$ -dimensional symplectic manifolds, but the general case (verified in [77,173,408]) of invariant tori of arbitrary dimensions is not harder at all.

By the way, Lemma 20 implies that all the perturbed invariant tori in the ‘conventional’ KAM Theory are *automatically* isotropic (in particular, Lagrangean, if their dimension is equal to the number of degrees of freedom). This refers e.g. to Theorems 6, 16, 17, 19 and 21 (in the case of Theorem 21 below, the symplectic form is exact in a neighbourhood of each torus).

Lemma 20 shows that the symplectic form in the higher dimensional KAM Theory should be non-exact (compare also with Theorem 5 in [318]). Moreover, it turns out that the periods of the symplectic form (its integrals over the two-dimensional cycles within the tori in question) should satisfy certain Diophantine-like conditions: all the theorems on coisotropic tori proven by now, include such Diophantine hypotheses. The coisotropic Hamiltonian KAM Theory was founded by I.O. Parasyuk [341] in 1984, see also subsequent papers [254,255,346–348] by Parasyuk and Kubichka (as well as the note [290] by Parasyuk and Loveĭkin). Coisotropic invariant  $n$ -tori of Hamiltonian systems with  $N < n$  degrees of freedom were also studied by Herman [223,224] (see also [321, 463,464,472]) and by F. Cong and Y. Li [135].

In Parasyuk’s theory, one starts with an unperturbed Hamiltonian system with  $N \geq 2$  degrees of freedom, whose phase space is smoothly foliated into coisotropic invariant  $n$ -tori carrying conditionally periodic dynamics ( $N + 1 \leq n \leq 2N - 1$ ). Then, as in the Lagrangean case  $n = N$  we considered in Section 4.3, one can prove that, under certain conditions on the symplectic form and the unperturbed Hamilton function, perturbed systems still admit many Diophantine quasi-periodic coisotropic invariant  $n$ -tori. The proofs unavoidably involve Diophantine approximations of dependent quantities, see Section 4.3.2. The measure of the complement to the union of the perturbed tori, of course, vanishes as the perturbation size tends to zero. The symplectic form here is usually supposed to be fixed, as in the ‘conventional’ Lagrangean or lower dimensional isotropic Hamiltonian KAM Theory. The theory of angle-action coordinates in the coisotropic framework is developed by Parasyuk in [345].

Herman’s papers [223,224] (compare with [472]) are devoted to the particular case  $n = 2N - 1$ . In this case, each unperturbed as well as perturbed energy hypersurface (to be more precise, each connected component of each energy hypersurface) is obviously an  $n$ -dimensional torus. It turns out that the frequency vectors of all the unperturbed tori in the set-up of [223,224] are proportional to the same Diophantine vector  $\omega^0 \in \mathbb{R}^n$  and, moreover, *all* the perturbed energy hypersurfaces still carry Diophantine quasi-periodic dynamics with the frequency vectors *also* proportional to the *same* vector  $\omega^0$ . This amazing picture (discussed also in [77, §1.4.2]) is a direct consequence of the fact that the Hamiltonian nature of a vector field imposes very severe restrictions on the motion on invariant tori of small codimensions in the phase space.

The most important application of the coisotropic Hamiltonian KAM Theory is counterexamples to the so-called Quasi-Ergodic Hypothesis [25,77,321,463,464]: ‘on (almost) every compact and connected energy hypersurface of a generic Hamiltonian system, there is an everywhere dense evolution curve’. The fallacy of this conjecture for the case of two degrees of freedom follows from the Lagrangean KAM Theory (see Section 4.4.2): Lagrangean KAM 2-tori divide three-dimensional energy hypersurfaces and exclude everywhere dense evolution curves. For the case of  $N \geq 3$  degrees of freedom, exactly the same obstruction to quasi-ergodicity is provided by the coisotropic KAM Theory: coisotropic invariant  $(2N - 2)$ -tori divide  $(2N - 1)$ -dimensional energy hypersurfaces and again exclude an everywhere dense evolution curve: all the evolution curves are forever trapped in between those tori. This reason is also due to Herman (see [321,463,464]). In addition, one obtains perpetual adiabatic stability of the two-dimensional ‘action’ variable. However, whether the Quasi-Ergodic Hypothesis is valid for  $N \geq 3$  degrees of freedom and exact symplectic forms is still an open question.

### 8.6. Atropic invariant tori

Now suppose that a Hamiltonian system  $X$  with  $N + m$  degrees of freedom ( $N \geq 2, m \geq 1$ ) possesses a  $2N$ -dimensional normally hyperbolic invariant surface  $\Pi$  with the following properties:

1. the restriction of the symplectic form to  $\Pi$  is a symplectic form on  $\Pi$ ;
2. the induced Hamiltonian dynamics on  $\Pi$  exemplifies the unperturbed system for some coisotropic KAM theorem (in the finitely smooth set-up).

In particular, the surface  $\Pi$  is foliated into  $X$ -invariant  $n$ -tori for some  $n$  in the range  $N + 1 \leq n \leq 2N - 1$  which are coisotropic *within*  $\Pi$ . In the ambient phase space, these tori are neither coisotropic nor isotropic: they are, as one says, *atropic*. Each Hamiltonian system  $\tilde{X}$  sufficiently close to  $X$  will have a normally hyperbolic invariant surface  $\tilde{\Pi}$  that is close to  $\Pi$  [126,175,226,434] and contains many  $\tilde{X}$ -invariant Diophantine quasi-periodic  $n$ -tori coisotropic within  $\tilde{\Pi}$  and atropic in the ambient phase space. The dimension of these tori can be

1. smaller than the number of degrees of freedom (if  $m \geq 2$  and  $N + 1 \leq n < N + \min\{m, N\}$ );
2. equal to the number of degrees of freedom (if  $1 \leq m \leq N - 1$  and  $n = N + m$ );
3. greater than the number of degrees of freedom (if  $N \geq 3, 1 \leq m \leq N - 2$ , and  $N + m < n \leq 2N - 1$ ).

One concludes that the Hamiltonian KAM Theory can be developed for atropic invariant tori. This was first observed by Q. Huang, F. Cong, and Y. Li [231,232] who in fact considered analytic systems and proved the existence of perturbed tori (with hyperbolic [231] as well as elliptic [232] normal behaviour) by entirely different methods. The review [408] presents a detailed discussion of this new and promising branch of KAM Theory. In a sense, the atropic context is to the coisotropic one as the isotropic context is to the Lagrangean one [412].

Moreover, it turns out that the coisotropic and atropic Hamiltonian KAM Theory (but not the Hamiltonian KAM Theory for Lagrangean or lower dimensional isotropic invariant tori!) can be carried over to the so-called locally Hamiltonian systems. Recall that a vector field  $X$  on a symplectic manifold is said to be *locally Hamiltonian*[20] if the 1-form  $\iota_X\sigma$  is *closed* (but not necessarily exact) where  $\sigma$  is the symplectic form (cf. Section 4.3.1). If the form  $\iota_X\sigma$  is exact and equal to  $dH$  then the vector field  $X = X_H$  is Hamiltonian. The locally Hamiltonian KAM Theory was constructed by I.O. Parasyuk and Yu.V. Loveikin, see their papers concerning coisotropic [291,292,342–344] and atropic [293] tori in locally Hamiltonian systems. This theory has been briefly reviewed in [412]. By the way, Lemma 20 is valid for invariant tori of locally Hamiltonian systems as well.

REMARK. Bifurcational aspects of the theory of coisotropic invariant tori (under small perturbations of both the vector field and the symplectic form) are treated in [255,290,347] for the case of Hamiltonian systems and in [291,292] for the case of locally Hamiltonian ones. Generalizations to atropic invariant tori in locally Hamiltonian systems are obtained in [293].

Note finally that the Hamiltonian KAM Theory can be generalized to the case where the phase space is a Poisson manifold rather than a symplectic one [275–277,282]. Poisson manifolds are equipped with a Poisson bracket (of functions) which is allowed to be degenerate.

## 9. Whitney smooth bundles of KAM tori

In classical mechanics many torus bundles are known to be nontrivial while the ‘standard’ KAM Theorem 6 of Section 4.3.1 only applies to trivial torus bundles of the form  $\mathbb{T}^n \times A$ . In this section we develop a global KAM Theory, that can be applied to nontrivial torus bundles, thereby obtaining a Cantorized bundle of which we can still speak of nontriviality. Note that the problem of constructing such a theory was mentioned by V.I. Arnold in a talk at the beginning of the 1990s, who referred to the paper [157].

### 9.1. Motivation

A motivation for globalizing the Lagrangean KAM theorem is the nontriviality of certain torus fibrations in Liouville integrable systems, for example in the spherical pendulum. Here an obstruction to the triviality of the fibration by Liouville tori is given by *monodromy*, see [138,157,158]. For a geometrical discussion of all obstructions for a toric fibration of an integrable Hamiltonian system to be trivial, see e.g. [37,38,157,327,476,477]. A natural question is whether (nontrivial) monodromy also can be defined for non-integrable perturbations of e.g. the spherical pendulum. Answering this question is of interest in the study of semi-classical versions of such classical systems, see [140,141,312]. The results discussed in this section imply that for an open set of Liouville integrable Hamiltonian systems, under a sufficiently small perturbation, the geometry of the fibration is largely preserved by a Whitney smooth diffeomorphism. Consequently, monodromy can be

defined in the nearly integrable case as well [53,59,60]. In particular, our approach applies to the spherical pendulum. For a similar result in the case of two degrees of freedom near a focus-focus singularity (or complex saddle point), see [363]. We expect that a suitable reformulation of our results will be valid in the general Lie algebra setting of [76–78,236,319].

Our goal is to establish a global quasi-periodic stability result for fibrations or bundles of Lagrangean tori, by gluing together local conjugations obtained from the classical ‘local’ KAM Theorem 6. This gluing uses a Partition of Unity [225,325,418] and the fact that invariant tori of the unperturbed integrable system have a natural affine structure [20,60,91,138,171], see Section 4.1.1. The global conjugation is obtained as an appropriate convex linear combination of the local conjugations. This construction is reminiscent of the one used to build connections or Riemannian metrics in differential geometry.

REMARK. In [60] it is also shown that the Whitney Extension Theorem [300,325,418,453] can be globalized to manifolds.

## 9.2. Formulation of the global KAM theorem

We now give a precise formulation of the global KAM theorem [60] in the world of  $C^\infty$ -systems. Consider a  $2n$ -dimensional, connected, smooth symplectic manifold  $(M, \sigma)$  with a surjective smooth map  $\pi : M \rightarrow B$ , where  $B$  is an  $n$ -dimensional smooth manifold. We assume that the map  $\pi$  defines a smooth locally trivial fibre bundle, the fibres of which are Lagrangean  $n$ -tori. As before,  $\mathbb{T}^n = \mathbb{R}^n / (2\pi\mathbb{Z})^n$  is the standard  $n$ -torus.

By the Liouville–Arnold Integrability Theorem [13,20,22,25,138,305] it follows that for every  $b \in B$  there is a neighbourhood  $U^b \subseteq B$  and a symplectic diffeomorphism

$$\varphi^b : V^b = \pi^{-1}(U^b) \rightarrow \mathbb{T}^n \times A^b; \quad m \mapsto (x^b(m), y^b(m)),$$

with  $A^b \subseteq \mathbb{R}^n$  an open set and with the symplectic form  $\sum_{j=1}^n dx_j^b \wedge dy_j^b$  on  $\mathbb{T}^n \times A^b$ , such that  $y^b = (y_1^b, y_2^b, \dots, y_n^b)$  is constant on the fibres of  $\pi$ . We call  $(x^b, y^b)$  angle-action variables and  $(V^b, \varphi^b)$  an angle-action chart.

Now consider a smooth Hamilton function  $H : M \rightarrow \mathbb{R}$ , which is constant on the fibres of  $\pi$ , that is,  $H$  is an integral of  $\pi$ . Then the corresponding Hamiltonian vector field  $X_H$ , defined by  $\iota_{X_H}\sigma = dH$ , is tangent to these fibres, for this notation see [1,9,20,25,77,360] and Section 4.3.1 above. This leads to a vector field

$$(\varphi_*^b X_H)(x^b, y^b) = \sum_{j=1}^n \omega_j^b(y^b) \frac{\partial}{\partial x_j^b} = \sum_{j=1}^n \frac{\partial(\varphi_*^b H)}{\partial y_j^b} \frac{\partial}{\partial x_j^b}$$

on  $\mathbb{T}^n \times A^b$  with the frequency vector  $\omega^b(y^b) = (\omega_1^b(y^b), \dots, \omega_n^b(y^b))$ . We call  $\omega^b : A^b \rightarrow \mathbb{R}^n$  the local frequency map. We say that  $H$  is a *globally nondegenerate* integral of  $\pi$  (in the sense of Kolmogorov), if for a collection  $\{(V^b, \varphi^b)\}_{b \in B}$  of angle-action charts whose domains  $V^b$  cover  $M$ , each local frequency map  $\omega^b : A^b \rightarrow \mathbb{R}^n$  is a diffeomorphism onto its image.

REMARKS. 1. According to a remark of J.J. Duistermaat [157], there is a natural affine structure on the space of actions  $B$ . We note that, as in Section 4.1.1, the affine structure is restricted somewhat further by the transition maps. Using this affine structure on  $B$ , for each  $b \in B$  the second derivative of  $h^b = (H|_{V^b}) \circ (\varphi^b)^{-1}$  on  $\mathbb{T}^n \times A^b$  is well-defined. By global nondegeneracy we mean that the second derivative  $D^2h^b$  has maximal rank  $n$  everywhere on  $V^b$  for each  $b \in B$ . Because of the affine structure on  $B$ , it follows that on overlapping angle-action charts the rank is independent of the choice of chart.

2. Often the regular  $n$ -torus bundle is a part of a larger structure containing singularities, see Section 3.2.2 for an example (the planar pendulum) and compare with the Stefan–Sussmann theory [419,422], also see [37,38].

Suppose that  $H$  is a globally nondegenerate integral on  $B$ . If  $B' \subseteq B$  is a relatively compact subset of  $B$ , that is, the closure of  $B'$  is compact, then there is a finite subcover  $\{U^b\}_{b \in \mathcal{F}}$  of  $\{U^b\}_{b \in B'}$  such that for every  $b \in \mathcal{F}$  the local frequency map  $\omega^b$  is a diffeomorphism onto its image. Accordingly, let  $M' = \pi^{-1}(B')$  and consider the corresponding bundle  $\pi' : M' \rightarrow B'$ .

We shall take perturbations of  $H$  in the  $C^\infty$ -topology on  $M$  [225,325], for a description see Section 4.1.2. We now formulate the global counterpart of Theorem 6.

**THEOREM 21** (Lagrangian KAM for Bundles [60]). *Let  $(M, \sigma)$  be a smooth  $2n$ -dimensional symplectic manifold with  $\pi : M \rightarrow B$  a smooth locally trivial Lagrangian  $n$ -torus bundle. Let  $B' \subseteq B$  be an open and relatively compact subset and let  $M' = \pi^{-1}(B')$ . Suppose that  $H : M \rightarrow \mathbb{R}$  is a smooth integral of  $\pi$ , which is globally nondegenerate. Finally let  $\tilde{F} : M \rightarrow \mathbb{R}$  be a smooth function. If  $\tilde{F}|_{\overline{M'}}$  is sufficiently small in the  $C^\infty$ -topology, then there is a subset  $C \subset B'$  and a map  $\Phi : M' \rightarrow M'$  with the following properties.*

1. *The subset  $C \subset B'$  is nowhere dense, and the measure of  $B' \setminus C$  tends to 0 as the size of the perturbation  $\tilde{F}$  tends to zero.*
2. *The subset  $\pi^{-1}(C) \subset M'$  is a union of Diophantine  $X_H$ -invariant Lagrangian  $n$ -tori.*
3. *The map  $\Phi$  is a  $C^\infty$ -diffeomorphism onto its image and is  $C^\infty$ -near the identity.*
4. *The restriction  $\hat{\Phi} = \Phi|_{\pi^{-1}(C)}$  conjugates  $X_H$  to  $X_{H+\tilde{F}}$ , that is,*

$$\hat{\Phi}_* X_H = X_{H+\tilde{F}}.$$

*Note that the Hamiltonian  $H + \tilde{F}$  need not be an integral of  $\pi$ .*

REMARKS. 1. The map  $\Phi$  generally is not symplectic, compare with the local situation of Section 4.3.1.

2. Item 4 of Theorem 21 can be also expressed by saying that  $X_H$  is (globally) quasi-periodically stable on  $M'$ . Note that, by the smoothness of  $\Phi$ , the measure of the nowhere dense set  $\Phi(\pi^{-1}(C))$ , which is the union of the perturbed  $n$ -tori, is large.
3. The restriction of  $\Phi$  to  $\pi^{-1}(C) \subset M'$  preserves the affine structure of the quasi-periodic tori, see Sections 4.1–4.3. In the complement  $M' \setminus \pi^{-1}(C)$  the

diffeomorphism  $\Phi$  has *no* dynamical meaning. Still, the push forward of the integrable bundle  $\pi' : M' \rightarrow B'$  by  $\Phi$  is a smooth  $n$ -torus bundle which interpolates the tori in  $\Phi(\pi^{-1}(C))$ .

4. Regarding the closed half-line structure of the subbundle  $\pi' : \pi^{-1}(C) \rightarrow C$ , we can largely repeat the remarks made after [Theorems 5 and 6](#). We like to add that by the gluing of various angle-action charts, the half-line structure is largely preserved [60]; one could speak here of a *Cantor manifold*.

### 9.3. Applications

As said earlier in [Section 9.1](#), many torus bundles occurring in Liouville integrable Hamiltonian systems are nontrivial. A notorious example is the spherical pendulum [138, 157], for more details see [Section 9.3.1](#) below. Other examples are met in the Hamiltonian Hopf bifurcation, the champagne bottle, etc.; for an overview see [160].

We first explain the global KAM [Theorem 21](#) on the example of the spherical pendulum, later on showing how monodromy can also be defined for Cantorized torus bundles as these occur in nearly integrable Hamiltonian systems. Also the connection with quantum monodromy is briefly discussed.

**9.3.1. Example: the spherical pendulum.** Here we consider the spherical pendulum [20, 138, 157, 363]. Dynamically, the spherical pendulum is the motion of a unit mass particle restricted to the unit sphere in  $\mathbb{R}^3$  in a constant vertically downward gravitational field. The configuration space of the spherical pendulum is the 2-sphere  $\mathbb{S}^2 = \{q \in \mathbb{R}^3 \mid |q| = 1\}$  and the phase space, the cotangent bundle

$$\tilde{M} = T^*\mathbb{S}^2 = \left\{ (q, p) \in \mathbb{R}^6 \mid |q| = 1 \text{ and } \langle q, p \rangle = 0 \right\}$$

of  $\mathbb{S}^2$ . Here  $q = (q_1, q_2, q_3)$  and  $p = (p_1, p_2, p_3)$ , while  $\langle \cdot, \cdot \rangle$ , as usual, denotes the standard inner product in  $\mathbb{R}^3$ .

The spherical pendulum is a Liouville integrable system. By Noether's Theorem [20, 25, 138] the rotational symmetry about the vertical axis gives rise to the angular momentum  $I(q, p) = q_1 p_2 - q_2 p_1$ , which is a second integral of motion, in addition to the energy  $E = H(q, p) = \frac{1}{2}|p|^2 + q_3$ . The energy-momentum map of the spherical pendulum is

$$\mathcal{EM} : T^*\mathbb{S}^2 \rightarrow \mathbb{R}^2; \quad (q, p) \mapsto (I, E) = \left( q_1 p_2 - q_2 p_1, \frac{1}{2}|p|^2 + q_3 \right).$$

Its fibres corresponding to regular values give rise to a fibration of the phase space by Lagrangean 2-tori. The image  $\tilde{B}$  of  $\mathcal{EM}$  is the closed part of the plane lying in between the two curves meeting at a corner, see [Figure 7](#). The set of singular values of  $\mathcal{EM}$  consists of the two boundary curves and the points  $(I, E) = (0, \pm 1)$ . These points correspond to the equilibria  $(q, p) = (0, 0, \pm 1, 0, 0, 0)$ , whereas the boundary curves correspond to the horizontal periodic motions of the pendulum discovered by Huygens [237]. Therefore the set  $B$  of regular  $\mathcal{EM}$ -values consists of the interior of  $\tilde{B}$  minus the point  $(I, E) = (0, 1)$ ,

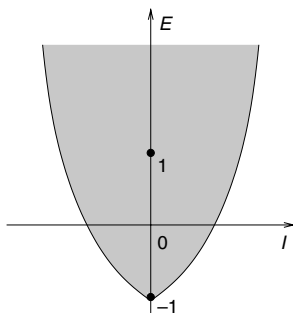


Fig. 7. Range of the energy-momentum map of the spherical pendulum.

corresponding to the unstable equilibrium point  $(0, 0, 1, 0, 0, 0)$ . This point is the centre of the nontrivial monodromy. The corresponding fibre  $\mathcal{EM}^{-1}(0, 1)$  is a once pinched 2-torus. Note that  $\mathcal{EM} : \tilde{M} \rightarrow \tilde{B}$  is a singular foliation in the sense of Stefan–Sussmann [419, 422].

On  $B$ , one of the two components of the frequency map is single valued while the other is multi-valued [138, 157]. E. Horozov [230] established global nondegeneracy of  $H$  on  $B$ . Thus the global KAM Theorem 21 can be applied to any relatively compact open subset  $B' \subseteq B$ . Consequently, the integrable dynamics on the 2-torus bundle  $\mathcal{EM}' : M' \rightarrow B'$  of the spherical pendulum is quasi-periodically stable.

This means that any sufficiently small Hamiltonian perturbation of the spherical pendulum has a nowhere dense union of Diophantine invariant tori of large measure, which admits a smooth interpolation by a push forward of the integrable bundle  $\mathcal{EM}' : M' \rightarrow B'$ . In Section 9.3.2 we shall argue that this allows for a definition of (nontrivial) monodromy for the perturbed torus bundle.

What precisely happens when the spherical pendulum system is slightly perturbed within the world of Hamiltonian systems? First, when the perturbation preserves the axial symmetry, the perturbed system remains Liouville integrable by Noether's Theorem [20, 25, 138] and hence the monodromy is preserved [308, 478].

The question now is what happens to the monodromy when by the perturbation the axial symmetry is broken, thereby yielding a nearly integrable system?

**9.3.2. Monodromy in the nearly integrable case.** Here we indicate how to define the concept of monodromy for a nearly integrable nowhere dense torus bundle [60]. Our construction, however, is independent of any integrable approximation.

**A regular union of tori** Let  $M$  be a manifold endowed with a (smooth) metric  $\varrho$ . Then for any two subsets  $A, B \subseteq M$  we define  $\varrho(A, B) = \inf_{x \in A, y \in B} \varrho(x, y)$ . Note that in general this does not define a metric on the set of all subsets:  $\varrho(A, B) = 0$  does not imply  $A = B$ , and  $\varrho(A, C)$  can well exceed  $\varrho(A, B) + \varrho(B, C)$ . Let  $M' \subseteq M$  be compact and  $\{T_\lambda\}_{\lambda \in \Lambda}$  a collection of pairwise disjoint  $n$ -tori in  $M$ . We require the following *regularity* properties. There exist positive constants  $\varepsilon$  and  $\delta$  such that

1. For all  $\lambda \in \Lambda$ , each continuous map  $h_\lambda : T_\lambda \rightarrow T_\lambda$  with  $\varrho(x, h(x)) < 2\varepsilon$  for all  $x \in T_\lambda$ , is homotopic with the identity map  $\text{Id}_{T_\lambda}$ ;



2. For each  $\lambda, \lambda' \in \Lambda$ , such that  $\varrho(T_\lambda, T_{\lambda'}) < \delta$ , there exists a homeomorphism  $h_{\lambda', \lambda} : T_\lambda \rightarrow T_{\lambda'}$ , such that  $\varrho(x, h_{\lambda', \lambda}(x)) < \frac{1}{2}\varepsilon$  for all  $x \in T_\lambda$ ;
3. For each  $x \in M'$  there exists  $\lambda \in \Lambda$  such that  $\varrho(\{x\}, T_\lambda) < \frac{1}{2}\delta$ .

Note that the homeomorphism  $h_{\lambda', \lambda}$ , as required to exist by item 2, by item 1 is unique modulo homotopy. We also observe the following. In the present situation, we started with a globally nondegenerate Liouville integrable Hamiltonian system, such that an open and dense part of its phase space  $M$  is foliated by invariant Lagrangean  $n$ -tori. Let  $M'$  be a compact union of such tori, then, under sufficiently small, non-integrable perturbations, the remaining Lagrangean KAM tori in  $M'$ , as discussed in this section, form a regular collection in the above sense.

**Construction of a  $\mathbb{Z}^n$ -bundle** We now construct a  $\mathbb{Z}^n$ -bundle, first only over the union  $M'' = \bigcup_\lambda T_\lambda$ , which is assumed to be regular in the above sense. For each point  $x \in T_\lambda$  the fibre is defined by  $F_x = H_1(T_\lambda, \mathbb{Z}) \approx \mathbb{Z}^n$ , the first homology group of  $T_\lambda \approx \mathbb{T}^n$  over  $\mathbb{Z}$ . Using the regularity property 2 and the fact that  $h_{\lambda', \lambda}$  is unique modulo homotopy, it follows that this bundle is locally trivial. Let  $E''$  denote the total space of this bundle and  $\pi'' : E'' \rightarrow M''$  the bundle projection.

Now this bundle is extended over  $M' \cup M''$  as follows. For each  $x \in M'$  we define

$$\Lambda(x) = \left\{ \lambda \in \Lambda \mid \varrho(\{x\}, T_\lambda) < \frac{1}{2}\delta \right\}.$$

Then we consider the set of pairs  $(\lambda, \alpha)$ , where  $\lambda \in \Lambda(x)$  and  $\alpha \in H_1(T_\lambda, \mathbb{Z})$ . On this set we have the following equivalence relation:

$$(\lambda, \alpha) \sim (\lambda', \alpha') \iff (h_{\lambda', \lambda})_* \alpha = \alpha',$$

where  $h_*$  denotes the action of  $h$  on the homology. The set of equivalence classes is defined as the fibre  $F_x$  at  $x$ . The fibre  $F_x$  is isomorphic to  $F_{x'}$  in a natural (and unique) way for any  $x' \in T_\lambda$  with  $\lambda \in \Lambda(x)$ . This extended bundle again is locally trivial.

We conclude by observing that the monodromy of the initial torus bundle is exactly the obstruction to global triviality of the  $\mathbb{Z}^n$ -bundle just constructed. Moreover, by the global KAM Theorem 21, in the integrable case one obtains the same  $\mathbb{Z}^n$ -bundle as in the nearly integrable case.

#### 9.4. Discussion

We described global quasi-periodic stability for bundles of Lagrangean invariant tori, under the assumption of global Kolmogorov nondegeneracy on the integrable approximation. We emphasize that our approach works for arbitrarily many degrees of freedom and that it is independent of the integrable geometry one starts with. The global Whitney smooth conjugation  $\widehat{\Phi}$  between the Diophantine tori of the integrable and the nearly integrable systems can be suitably extended to a smooth map  $\Phi$ , which serves to smoothly interpolate the nearly integrable, only Whitney smooth Diophantine torus bundle  $\Phi(\pi^{-1}(C))$ , defined over the nowhere dense set  $C$ . We observe that if the extended diffeomorphisms are



sufficiently close to the identity, they are also isotopic to the identity, see [60]. Therefore the interpolation becomes bundle isomorphic with its integrable counterpart  $\pi' : M' \rightarrow B'$ . This means that the global geometry is as in the integrable case.

In this sense, the lack of unicity of Whitney extensions plays no role and we can generalize concepts like monodromy directly to the nearly integrable case. For a topological discussion of the corresponding  $n$ -torus bundles, see [157,327,476–478].

The present global quasi-periodic stability result directly carries over to the general setting of [76–78,236,319]. Within the world of Hamiltonian systems, this leads to applications at the level of lower dimensional isotropic tori, see Section 8.3. However, the approach also applies to dissipative, volume preserving, or reversible systems, compare with [57,58,75].

REMARK. The asymptotic considerations of B.W. Rink [363] near focus-focus singularities (i.e. complex saddles) in Liouville integrable Hamiltonian systems with two degrees of freedom, allow for a similar application to the spherical pendulum where  $B'$  is a small annular region around the point  $(I, E) = (0, 1)$ . The conclusion again is that the nearly integrable systems have nontrivial monodromy.

It is tempting to combine the global KAM Theory with that of quasi-periodic bifurcations [57,63,337]. In the ensuing Cantorization, apart from closed half-lines, also higher dimensional closed half-spaces do occur. It is to be noted that in the quasi-periodic reversible Hopf bifurcation, also nontrivial monodromy occurs [57].

One of the main motivations for the interest in nontriviality of symplectic torus bundles in Hamiltonian systems, is the connection with certain spectral properties in related semi-classical systems. In particular this refers to the so-called spectral defect. This connection is proven by Vũ Ngọc San [385]. Many applications have been reported in [139,140,159–162,441], for overviews see [53,140,420].

By Theorem 21 and its consequences, classical monodromy also exists in the nearly integrable case. As far as we know, quantum monodromy is only well-defined for integrable systems, often obtained by just truncating a nearly integrable system. Similarly in a number of the Quantum Theory applications, the classical limit is not integrable, but only nearly so. It is an open question as to whether quantum monodromy can also be defined in nearly integrable cases, but we expect that the present approach may be useful in this.

## 10. Conclusion

The lasting influence of Kolmogorov, Arnold, and Moser on the present state of the art in mathematics, physics and other sciences is enormous and this review only sketches part of this legacy. Nevertheless we believe that it is an important part, which is still fully in development. The role of the Ergodic Hypothesis in Statistical Mechanics has turned out to be much more subtle than was expected, see e.g. [45,183,184]. Regarding KAM Theory, for further reading we mention the introductory texts [118,132,147,304,357]. Also reading of [7,25,77,168,259,268,408] is recommended. The discussion around quantum monodromy and its relationship to KAM Theory, in particular to the Cantorization of Whitney smooth bundles, has not yet come to its conclusion [60,140,160,420]. The

role of quasi-periodic bifurcation theory [67,77,78,212,445,446] and the corresponding Cantorization and fraying undoubtedly will become increasingly more relevant, especially since now higher dimensional modelling is practiced more often.

For general reference we also refer to the companion Handbook volumes [3–5].

## Acknowledgments

The authors thank Richard Cushman, Konstantinos Efsthathiou, Aernout van Enter, Francesco Fassò, Heinz Hanßmann, Jun Hoo, George Huitema, Sergei Kuksin, Olga Lukina, Jan van Maanen, Anatoliĭ Neĭshtadt, Hinke Osinga, Joaquim Puig, Bob Rink, Khairul Saleh, Carles Simó, Floris Takens, Dmitriĭ Treshchĕv, Renato Vitolo, and Florian Wagener for helpful discussions. The first author is grateful to Jürgen Moser for his guidance in KAM Theory. The second author is very much obliged to Vladimir Igorevich Arnold who taught him KAM Theory. Moreover, the first author acknowledges hospitality of the Universitat de Barcelona and the Université de Bourgogne and partial support of the Dutch FOM program Mathematical Physics (MF-G-b) and of European Community funding for the Research and Training Network MASIE (HPRN-CT-2000-00113). The second author is grateful to the University of Maryland for its hospitality and acknowledges partial support of the Council for Grants of President of the Russia Federation, Grants No. NSh-1972.2003.1, NSh-4719.2006.1, NSh-709.2008.1 and NSh-8462.2010.1.

## References

### *Surveys in volumes 1–3*

- [1] H.W. Broer and F. Takens, *Preliminaries of dynamical systems theory*, Handbook of Dynamical Systems, Vol. 3, H.W. Broer, B. Hasselblatt and F. Takens, eds, Elsevier, Amsterdam (2010).
- [2] R.L. Devaney, *Complex 1-dimensional mappings and their bifurcations*, Handbook of Dynamical Systems, Vol. 3, H.W. Broer, B. Hasselblatt and F. Takens, eds, Elsevier, Amsterdam (2010).
- [3] B. Fiedler, ed., Handbook of Dynamical Systems, Vol. 2, Elsevier, Amsterdam (2002).
- [4] B. Hasselblatt and A. Katok, eds, Handbook of Dynamical Systems, Vol. 1A, Elsevier, Amsterdam (2002).
- [5] B. Hasselblatt and A. Katok, eds, Handbook of Dynamical Systems, Vol. 1B, Elsevier, Amsterdam (2006).
- [6] S.B. Kuksin, *Hamiltonian PDEs*, with an appendix by D. Bambusi, Handbook of Dynamical Systems, Vol. 1B, B. Hasselblatt and A. Katok, eds, Elsevier, Amsterdam (2006), 1087–1133.
- [7] M. Levi, *Some applications of Moser's twist theorem*, Handbook of Dynamical Systems, Vol. 3, H.W. Broer, B. Hasselblatt and F. Takens, eds, Elsevier, Amsterdam (2010).

### *Other sources*

- [8] K. Abdullah and A. Albouy, *On a strange resonance noticed by M. Herman*, Regul. Chaotic Dyn. **6** (2001), 421–432.
- [9] R.H. Abraham and J.E. Marsden, *Foundations of Mechanics*, with an appendix by A.N. Kolmogorov, 2nd edn, Benjamin/Cummings, Reading, MA (1978).
- [10] J. Albrecht, *On the existence of invariant tori in nearly-integrable Hamiltonian systems with finitely differentiable perturbations*, Regul. Chaotic Dyn. **12** (2007), 281–320.
- [11] V.I. Arnold, *Small divisors I: On mappings of the circle onto itself*, Izvestiya Akad. Nauk SSSR, Ser. Mat. **25** (1961), 21–86 (in Russian); English translation: Amer. Math. Soc. Transl., Ser. 2 **46** (1965), 213–284; Erratum: Izvestiya Akad. Nauk SSSR, Ser. Mat. **28** (1964), 479–480 (in Russian).

- [12] V.I. Arnold, *On the classical perturbation theory and the problem of stability of planetary systems*, Soviet Math. Dokl. **3** (1962), 1008–1012 (English; Russian original).
- [13] V.I. Arnold, *On Liouville's theorem concerning integrable problems of dynamics*, Sibirsk. Mat. Zh. **4** (1963), 471–474 (in Russian); English translation: Amer. Math. Soc. Transl., Ser. 2 **61** (1967), 292–296.
- [14] V.I. Arnold, *Proof of a theorem by A.N. Kolmogorov on the persistence of conditionally periodic motions under a small change of the Hamilton function*, Russian Math. Surveys **18** (5) (1963), 9–36 (English; Russian original).
- [15] V.I. Arnold, *Small denominators and problems of stability of motion in classical and celestial mechanics*, Russian Math. Surveys **18** (6) (1963), 85–191 (English; Russian original); Erratum: Uspekhi Mat. Nauk **23** (6) (1968), 216 (in Russian).
- [16] V.I. Arnold, *On the instability of dynamical systems with many degrees of freedom*, Soviet Math. Dokl. **5** (1964), 581–585 (English; Russian original).
- [17] V.I. Arnold, *On matrices depending on parameters*, Russian Math. Surveys **26** (2) (1971), 29–43 (English; Russian original).
- [18] V.I. Arnold, ed., *Reversible systems*, Nonlinear and Turbulent Processes in Physics, Vol. 3 (Kiev, 1983), R.Z. Sagdeev, ed., Harwood Academic, Chur, New York (1984), 1161–1174.
- [19] V.I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, 2nd edn, Springer, New York (1988) (English; Russian original).
- [20] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd edn, Springer, New York (1989) (English; Russian original).
- [21] V.I. Arnold, ed., *Dynamical Systems. V. Bifurcation Theory and Catastrophe Theory*, Encyclopædia of Mathematical Sciences, Vol. 5, Springer, Berlin (1994) (English; Russian original).
- [22] V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, 2nd edn, Addison-Wesley, Redwood City, CA (1989) (English; French original).
- [23] V.I. Arnold, V.V. Goryunov, O.V. Lyashko and V.A. Vasil'ev, *Singularity Theory. I. Local and Global Theory*, Dynamical Systems. VI, Encyclopædia of Mathematical Sciences, Vol. 6, Springer, Berlin (1993) (English; Russian original).
- [24] V.I. Arnold, V.V. Goryunov, O.V. Lyashko and V.A. Vasil'ev, *Singularity Theory. II. Classification and Applications*, Dynamical Systems. VIII, Encyclopædia of Mathematical Sciences, Vol. 39, Springer, Berlin (1993) (English; Russian original).
- [25] V.I. Arnold, V.V. Kozlov and A.I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, 3rd edn, Dynamical Systems. III, Encyclopædia of Mathematical Sciences, Vol. 3, Springer, Berlin (2006) (English; the Russian original of 2002 is briefer).
- [26] V.I. Arnold and M.B. Sevryuk, *Oscillations and bifurcations in reversible systems*, Nonlinear Phenomena in Plasma Physics and Hydrodynamics, R.Z. Sagdeev, ed., Mir, Moscow (1986), 31–64.
- [27] C. Baesens, J. Guckenheimer, S. Kim and R.S. MacKay, *Three coupled oscillators: mode-locking, global bifurcations and toroidal chaos*, Physica D **49** (1991), 387–475.
- [28] J. Barrow-Green, *Poincaré and the Three Body Problem*, History of Mathematics, Vol. 11, Amer. Math. Soc., Providence, RI (1997); London Math. Soc., London.
- [29] G. Benettin, L. Galgani, A. Giorgilli and J.-M. Strelcyn, *A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method*, Nuovo Cimento B **79** (2) (1984), 201–223.
- [30] Yu.N. Bibikov, *A sharpening of a theorem of Moser*, Soviet Math. Dokl. **14** (1973), 1769–1773 (English; Russian original).
- [31] Yu.N. Bibikov, *Local Theory of Nonlinear Analytic Ordinary Differential Equations*, Lecture Notes in Math., Vol. 702, Springer, Berlin (1979).
- [32] Yu.N. Bibikov, *Construction of invariant tori of systems of differential equations with a small parameter*, Trudy Leningrad. Mat. Obshch. **1** (1990), 26–53 (in Russian); English translation: Amer. Math. Soc. Transl., Ser. 2 **155** (1993), 19–46.
- [33] G.D. Birkhoff, *What is the ergodic theorem?* Amer. Math. Monthly **49** (1942), 222–226.
- [34] G.D. Birkhoff, *Dynamical Systems* (Second Edition with an addendum by J.K. Moser), Amer. Math. Soc. Colloquium Publications, Vol. 9, Amer. Math. Soc., Providence, RI (1966) (First Edition: 1927).
- [35] P. Blanchard, *Complex analytic dynamics on the Riemann sphere*, Bull. Amer. Math. Soc. (N.S.) **11** (1984), 85–141.
- [36] S.V. Bolotin and D.V. Treshchëv, *Remarks on the definition of hyperbolic tori of Hamiltonian systems*, Regul. Chaotic Dyn. **5** (2000), 401–412.

- [37] A.V. Bolsinov and A.T. Fomenko, *Integrable Geodesic Flows on Two-Dimensional Surfaces*, Consultants Bureau, New York (2000).
- [38] A.V. Bolsinov and A.T. Fomenko, *Integrable Hamiltonian Systems. Geometry, Topology, Classification*, with appendices by V.V. Kalashnikov, Chapman & Hall/CRC, Boca Raton, FL (2004) (English; Russian original).
- [39] F. Bonetto, G. Gallavotti, G. Gentile and V. Mastropietro, *Lindstedt series, ultraviolet divergences and Moser's theorem*, Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., IV Ser. **26** (1998), 545–593.
- [40] J.-B. Bost, *Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnold, Moser, Russmann, Zehnder, Herman, Pöschel, ...)*, Séminaire Bourbaki, Vol. 1984/85, Exp. 639, Astérisque **133–134** (1986), 113–157.
- [41] J. Bourgain, *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Int. Math. Res. Not. (11) (1994), 475–497.
- [42] J. Bourgain, *On Melnikov's persistency problem*, Math. Res. Lett. **4** (1997), 445–458.
- [43] B.L.J. Braaksma and H.W. Broer, *Quasi-periodic flow near a codimension one singularity of a divergence free vector field in dimension four*, Bifurcation, Théorie Ergodique et Applications (Dijon, 1981), Astérisque **98–99** (1982), 74–142.
- [44] B.L.J. Braaksma and H.W. Broer, *On a quasi-periodic Hopf bifurcation*, Ann. Inst. H. Poincaré Anal. Non Linéaire **4** (2) (1987), 115–168.
- [45] J. Bricmont, *Science of chaos or chaos in science? The Flight from Science and Reason* (New York, 1995), P.R. Gross, N. Levitt and M.W. Lewis, eds, Ann. New York Academy of Sciences, Vol. 775, New York Academy of Sciences, New York (1996), 131–175; Also appeared in: Physica Magazine **17** (1995), 159–208.
- [46] J. Bricmont, K. Gawędzki and A. Kupiainen, *KAM theorem and Quantum Field Theory*, Comm. Math. Phys. **201** (1999), 699–727.
- [47] H.W. Broer, *Formal normal form theorems for vector fields and some consequences for bifurcations in the volume preserving case*, Dynamical Systems and Turbulence (Warwick, 1980), D.A. Rand and L.S. Young, eds, Lecture Notes in Math., Vol. 898, Springer, Berlin (1981), 54–74.
- [48] H.W. Broer, *Quasi-periodic flow near a codimension one singularity of a divergence free vector field in dimension three*, Dynamical Systems and Turbulence (Warwick, 1980), D.A. Rand and L.S. Young, eds, Lecture Notes in Math., Vol. 898, Springer, Berlin (1981), 75–89.
- [49] H.W. Broer, *Quasi-periodic bifurcations, applications*, Proceedings of the XI CEDYA (Congreso de Ecuaciones Diferenciales y Aplicaciones) (1989), Univ. Málaga, Málaga (1990), 3–21.
- [50] H.W. Broer, *Notes on perturbation theory* (1991), Erasmus ICP Mathematics and Fundamental Applications, Aristotle Univ., Thessaloniki (1993).
- [51] H.W. Broer, *KAM-Theory: Thesi-periodicity in conservative and dissipative systems*, Nieuw Arch. Wisk., Ser. 4 **14** (1) (1996), 65–79.
- [52] H.W. Broer, *Quasi-periodicity in dissipative systems*, MIHMI (J. Indonesian Math. Soc.) **7** (3) (2001), 7–33.
- [53] H.W. Broer, *A global KAM-theorem: Monodromy in near-integrable perturbations of the spherical pendulum*, Proc. ITB **34** (2–3) (2003), 309–324.
- [54] H.W. Broer, *Coupled Hopf-bifurcations: Persistent examples of  $n$ -quasiperiodicity determined by families of 3-jets*, Astérisque **286** (2003), 223–229.
- [55] H.W. Broer, *Kolmogorov: la 'κ' de KAM*, Butl. Soc. Catalana Mat. **18** (2) (2003), 39–57 (Catalan; English original).
- [56] H.W. Broer, *KAM Theory: the legacy of Kolmogorov's 1954 paper*, Bull. Amer. Math. Soc. (N.S.) **41** (2004), 507–521.
- [57] H.W. Broer, M.-C. Ciocci and H. Hanßmann, *The quasi-periodic reversible Hopf bifurcation*, Internat. J. Bifur. Chaos **17** (2007), 2605–2623.
- [58] H.W. Broer, M.-C. Ciocci, H. Hanßmann and A. Vanderbauwhede, *Quasi-periodic stability of normally resonant tori*, Physica D **238** (2009), 309–318.
- [59] H.W. Broer, R.H. Cushman and F. Fassò, *A Hamiltonian KAM theorem for bundles of Lagrangean tori*, EQUADIFF 2003 (Proc. Intern. Conf. Differential Equations) (Hasselt/Diepenbeek, 2003), F. Dumortier, H.W. Broer, J. Mawhin, A. Vanderbauwhede and S.M. Verduyn Lunel, eds, World Scientific, Hackensack, NJ (2005), 696–701.

- [60] H.W. Broer, R.H. Cushman, F. Fassò and F. Takens, *Geometry of KAM tori for nearly integrable Hamiltonian systems*, Ergodic Theory Dynam. Systems **27** (2007), 725–741.
- [61] H.W. Broer, F. Dumortier, S.J. van Strien and F. Takens, *Structures in Dynamics: Finite Dimensional Deterministic Studies*, E.W.C. van Groesen and E.M. de Jager, eds, Studies in Mathematical Physics, Vol. 2, North-Holland, Amsterdam (1991).
- [62] H.W. Broer, M. Golubitsky and G. Vegter, *The geometry of resonance tongues: a singularity theory approach*, Nonlinearity **16** (2003), 1511–1538.
- [63] H.W. Broer, H. Hanßmann and J. Hoo, *The quasi-periodic Hamiltonian Hopf bifurcation*, Nonlinearity **20** (2007), 417–460.
- [64] H.W. Broer, H. Hanßmann, J. Hoo and V. Naudot, *Nearly-integrable perturbations of the Lagrange top: applications of KAM-theory*, Dynamics and Stochastics: Festschrift in Honor of M.S. Keane on the Occasion of his 65th Birthday, D. Denteneer, F. den Hollander and E. Verbitskiy, eds, IMS Lecture Notes–Monograph Series, Vol. 48, Institute of Mathematical Statistics, Beachwood, OH (2006), 286–303.
- [65] H.W. Broer, H. Hanßmann, À. Jorba, J. Villanueva and F.O.O. Wagener, *Normal-internal resonances in quasi-periodically forced oscillators: a conservative approach*, Nonlinearity **16** (2003), 1751–1791.
- [66] H.W. Broer, H. Hanßmann, À. Jorba, J. Villanueva and F.O.O. Wagener, *Quasi-periodic response solutions at normal-internal resonances*, EQUADIFF 2003 (Proc. Intern. Conf. Differential Equations) (Hasselt/Diepenbeek, 2003), F. Dumortier, H.W. Broer, J. Mawhin, A. Vanderbauwhede and S.M. Verduyn Lunel, eds, World Scientific, Hackensack, NJ (2005), 702–707.
- [67] H.W. Broer, H. Hanßmann and F.O.O. Wagener, *Quasi-Periodic Bifurcation Theory* (in preparation).
- [68] H.W. Broer, H. Hanßmann and J. You, *Bifurcations of normally parabolic tori in Hamiltonian systems*, Nonlinearity **18** (2005), 1735–1769.
- [69] H.W. Broer, H. Hanßmann and J. You, *Umbilical torus bifurcations in Hamiltonian systems*, J. Differential Equations **222** (2006), 233–262.
- [70] H.W. Broer, H. Hanßmann and J. You, *On the destruction of resonant Lagrangean tori in Hamiltonian systems* (submitted).
- [71] H.W. Broer, J. Hoo and V. Naudot, *Normal linear stability of quasi-periodic tori*, J. Differential Equations **232** (2007), 355–418.
- [72] H.W. Broer, I. Hoveijn, G.A. Lunter and G. Vegter, *Bifurcations in Hamiltonian Systems: Computing Singularities by Gröbner Bases*, Lecture Notes in Math., Vol. 1806, Springer, Berlin (2003).
- [73] H.W. Broer, I. Hoveijn, M. van Noort, C. Simó and G. Vegter, *The parametrically forced pendulum: a case study in  $1\frac{1}{2}$  degree of freedom*, J. Dynam. Differential Equations **16** (2004), 897–947.
- [74] H.W. Broer and G.B. Huitema, *A proof of the isoenergetic KAM-theorem from the “ordinary” one*, J. Differential Equations **90** (1991), 52–60.
- [75] H.W. Broer and G.B. Huitema, *Unfoldings of quasi-periodic tori in reversible systems*, J. Dynam. Differential Equations **7** (1995), 191–212.
- [76] H.W. Broer, G.B. Huitema and M.B. Sevryuk, *Families of quasi-periodic motions in dynamical systems depending on parameters*, Nonlinear Dynamical Systems and Chaos (Groningen, 1995), H.W. Broer, S.A. van Gils, I. Hoveijn and F. Takens, eds, Progress in Nonlinear Differential Equations and their Applications, Vol. 19, Birkhäuser, Basel (1996), 171–211.
- [77] H.W. Broer, G.B. Huitema and M.B. Sevryuk, *Quasi-Periodic Motions in Families of Dynamical Systems: Order amidst Chaos*, Lecture Notes in Math., Vol. 1645, Springer, Berlin (1996).
- [78] H.W. Broer, G.B. Huitema, F. Takens and B.L.J. Braaksma, *Unfoldings and bifurcations of quasi-periodic tori*, Mem. Amer. Math. Soc. **83** (421) (1990), i–vii and 1–175.
- [79] H.W. Broer and M. Levi, *Geometrical aspects of stability theory for Hill’s equations*, Arch. Ration. Mech. Anal. **131** (1995), 225–240.
- [80] H.W. Broer, V. Naudot, R. Roussarie, K. Saleh and F.O.O. Wagener, *Organising centres in the semi-global analysis of dynamical systems*, Internat. J. Appl. Math. Stat. (IJAMAS) **12** (D07) (2007), 7–36.
- [81] H.W. Broer, J. Puig and C. Simó, *Resonance tongues and instability pockets in the quasi-periodic Hill–Schrödinger equation*, Comm. Math. Phys. **241** (2003), 467–503.
- [82] H.W. Broer and R. Roussarie, *Exponential confinement of chaos in the bifurcation sets of real analytic diffeomorphisms*, Global Analysis of Dynamical Systems. Festschrift Dedicated to Floris Takens for his 60th Birthday (Leiden, 2001), H.W. Broer, B. Krauskopf and G. Vegter, eds, Inst. Physics, Bristol (2001), 167–210.

- [83] H.W. Broer, R. Roussarie and C. Simó, *Invariant circles in the Bogdanov–Takens bifurcation for diffeomorphisms*, Ergodic Theory Dynam. Systems **16** (1996), 1147–1172.
- [84] H.W. Broer and C. Simó, *Hill's equation with quasi-periodic forcing: resonance tongues, instability pockets and global phenomena*, Bol. Soc. Brasil. Mat. (N.S.) **29** (1998), 253–293.
- [85] H.W. Broer and C. Simó, *Resonance tongues in Hill's equations: a geometric approach*, J. Differential Equations **166** (2000), 290–327.
- [86] H.W. Broer, C. Simó and J.C. Tatjer, *Towards global models near homoclinic tangencies of dissipative diffeomorphisms*, Nonlinearity **11** (1998), 667–770.
- [87] H.W. Broer, C. Simó and R. Vitolo, *Bifurcations and strange attractors in the Lorenz-84 climate model with seasonal forcing*, Nonlinearity **15** (2002), 1205–1267.
- [88] H.W. Broer and F. Takens, *Formally symmetric normal forms and genericity*, Dynamics Reported **2** (1989), 39–59.
- [89] H.W. Broer and F. Takens, *Mixed spectra and rotational symmetry*, Arch. Ration. Mech. Anal. **124** (1993), 13–42.
- [90] H.W. Broer, F. Takens and F.O.O. Wagener, *Integrable and non-integrable deformations of the skew Hopf bifurcation*, Regul. Chaotic Dyn. **4** (2) (1999), 16–43.
- [91] H.W. Broer and F. Takens, *Unicity of KAM tori*, Ergodic Theory Dynam. Systems **27** (2007), 713–724.
- [92] H.W. Broer and F. Takens, *Dynamical Systems and Chaos*, Epsilon Uitgaven, Vol. 64 (2009); Springer Appl. Math. Sci., Vol. 172 (2010).
- [93] H.W. Broer and F.M. Tangerman, *From a differentiable to a real analytic perturbation theory, applications to the Kupka Smale theorems*, Ergodic Theory Dynam. Systems **6** (1986), 345–362.
- [94] H.W. Broer, M. van Noort and C. Simó, *Existence and measure of 2-quasiperiodicity in Hamiltonian one-and-a-half degree of freedom systems*, EQUADIFF 2003 (Proc. Intern. Conf. Differential Equations) (Hasselt/Diepenbeek, 2003), F. Dumortier, H.W. Broer, J. Mawhin, A. Vanderbauwhede and S.M. Verduyn Lunel, eds, World Scientific, Hackensack, NJ (2005), 595–600.
- [95] H.W. Broer and G. Vegter, *Bifurcational aspects of parametric resonance*, Dynamics Reported (N.S.) **1** (1992), 1–53.
- [96] H.W. Broer and F.O.O. Wagener, *Quasi-periodic stability of subfamilies of an unfolded skew Hopf bifurcation*, Arch. Ration. Mech. Anal. **152** (2000), 283–326.
- [97] A.D. Bruno, *On the convergence of transformations of differential equations to the normal form*, Soviet Math. Dokl. **6** (1965), 1536–1538 (English; Russian original).
- [98] A.D. Bruno, *Analytic form of differential equations, I and II*, Trans. Moscow Math. Soc. **25** (1971), 131–288; **26** (1972), 199–239 (English; Russian original).
- [99] A.D. Bruno, *Local Methods in Nonlinear Differential Equations*, Springer, Berlin (1989) (English; Russian original).
- [100] A.D. Bruno, *Power Geometry in Algebraic and Differential Equations*, North-Holland Mathematical Library, Vol. 57, North-Holland, Amsterdam (2000) (English; Russian original).
- [101] X. Cabré, E. Fontich and R. de la Llave, *The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces*, Indiana Univ. Math. J. **52** (2003), 283–328.
- [102] X. Cabré, E. Fontich and R. de la Llave, *The parameterization method for invariant manifolds. II. Regularity with respect to parameters*, Indiana Univ. Math. J. **52** (2003), 329–360.
- [103] X. Cabré, E. Fontich and R. de la Llave, *The parameterization method for invariant manifolds. III. Overview and applications*, J. Differential Equations **218** (2005), 444–515.
- [104] A. Celletti and L. Chierchia, *KAM stability for a three-body problem of the Solar System*, Z. Angew. Math. Phys. **57** (2006), 33–41.
- [105] A. Celletti and L. Chierchia, *KAM tori for N-body problems: a brief history*, Celest. Mech. Dynam. Astronomy **95** (2006), 117–139.
- [106] A. Celletti and L. Chierchia, *KAM stability and Celestial Mechanics*, Mem. Amer. Math. Soc. **187** (878) (2007), i–vii and 1–134.
- [107] A. Chenciner, *Bifurcations de points fixes elliptiques. I. Courbes invariantes*, Publ. Math. IHÉS **61** (1985), 67–127.
- [108] A. Chenciner, *Bifurcations de points fixes elliptiques. II. Orbites périodiques et ensembles de Cantor invariants*, Invent. Math. **80** (1985), 81–106.
- [109] A. Chenciner, *Bifurcations de points fixes elliptiques. III. Orbites périodiques de “petites” périodes et élimination résonnante des couples de courbes invariantes*, Publ. Math. IHÉS **66** (1988), 5–91.

- [110] A. Chenciner and G. Iooss, *Bifurcations de tores invariants*, Arch. Ration. Mech. Anal. **69** (1979), 109–198.
- [111] A. Chenciner and G. Iooss, *Persistence et bifurcation de tores invariants*, Arch. Ration. Mech. Anal. **71** (1979), 301–306.
- [112] Ch.-Q. Cheng, *Birkhoff–Kolmogorov–Arnold–Moser tori in convex Hamiltonian systems*, Comm. Math. Phys. **177** (1996), 529–559.
- [113] Ch.-Q. Cheng, *Lower dimensional invariant tori in the regions of instability for nearly integrable Hamiltonian systems*, Comm. Math. Phys. **203** (1999), 385–419.
- [114] Ch.-Q. Cheng and Y.-S. Sun, *Existence of KAM tori in degenerate Hamiltonian systems*, J. Differential Equations **114** (1994), 288–335.
- [115] Ch.-Q. Cheng and Sh. Wang, *The surviving of lower dimensional tori from a resonant torus of Hamiltonian systems*, J. Differential Equations **155** (1999), 311–326.
- [116] Ch.-Q. Cheng and J. Yan, *Existence of diffusion orbits in a priori unstable Hamiltonian systems*, J. Differential Geom. **67** (2004), 457–517.
- [117] L. Chierchia, *Kolmogorov’s 1954 paper on nearly-integrable Hamiltonian systems*, Regul. Chaotic Dyn. **13** (2008), 130–139.
- [118] L. Chierchia, *KAM lectures*, Dynamical Systems, Part I. Hamiltonian Systems and Celestial Mechanics (Pisa, 2002), Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa (2003), 1–55.
- [119] L. Chierchia and C. Falcolini, *A direct proof of a theorem by Kolmogorov in Hamiltonian systems*, Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., IV Ser. **21** (1994), 541–593.
- [120] L. Chierchia and C. Falcolini, *Compensations in small divisor problems*, Comm. Math. Phys. **175** (1996), 135–160.
- [121] L. Chierchia and C. Falcolini, *On the convergence of formal series containing small divisors*, Hamiltonian Systems with Three or More Degrees of Freedom (S’Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 345–349.
- [122] L. Chierchia and G. Gallavotti, *Smooth prime integrals for quasi-integrable Hamiltonian systems*, Nuovo Cimento B **67** (2) (1982), 277–295.
- [123] L. Chierchia and G. Gallavotti, *Drift and diffusion in phase space*, Ann. Institut Henri Poincaré, Physique Théorique **60** (1) (1994), 1–144; Erratum: **68** (1) (1998), 135.
- [124] L. Chierchia and D. Qian, *Moser’s theorem for lower dimensional tori*, J. Differential Equations **206** (2004), 55–93.
- [125] B.V. Chirikov, *A universal instability of many-dimensional oscillator systems*, Phys. Rep. **52** (5) (1979), 263–379.
- [126] S.-N. Chow and J.K. Hale, *Methods of Bifurcation Theory*, Springer, New York (1982).
- [127] S.-N. Chow, C.Z. Li and D. Wang, *Normal Forms and Bifurcation of Planar Vector Fields*, Cambridge Univ. Press, Cambridge (1994).
- [128] S.-N. Chow, Y. Li and Y. Yi, *Persistence of invariant tori on submanifolds in Hamiltonian systems*, J. Nonlinear Sci. **12** (2002), 585–617.
- [129] G. Cicogna and G. Gaeta, *Symmetry and Perturbation Theory in Nonlinear Dynamics*, Lecture Notes in Phys. (Monographs Series), Vol. 57, Springer, Berlin (1999).
- [130] M.-C. Ciocci, *Bifurcation of periodic orbits and persistence of quasi-periodic orbits in families of reversible systems*, Ph.D. Thesis, Univ. Gent (2003).
- [131] M.-C. Ciocci, *Persistence of quasi-periodic orbits in families of reversible systems with a 1 : 1 resonance*, EQUADIFF 2003 (Proc. Intern. Conf. Differential Equations) (Hasselt/Diepenbeek, 2003), F. Dumortier, H.W. Broer, J. Mawhin, A. Vanderbauwhede and S.M. Verduyn Lunel, eds, World Scientific, Hackensack, NJ (2005), 720–725.
- [132] M.-C. Ciocci, A. Litvak-Hinzenon and H.W. Broer, *Survey on dissipative KAM theory including quasi-periodic bifurcation theory, based on lectures by Henk Broer*, Geometric Mechanics and Symmetry. The Peyresq Lectures (Peyresq, 2000–01), J. Montaldi and T.S. Ratiu, eds, London Math. Soc. Lecture Note Series, Vol. 306, Cambridge Univ. Press, Cambridge (2005), 303–355.
- [133] F. Cong, T. Küpper, Y. Li and J. You, *KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems*, J. Nonlinear Sci. **10** (2000), 49–68.
- [134] F. Cong and Y. Li, *Invariant hyperbolic tori for Hamiltonian systems with degeneracy*, Discrete Contin. Dynam. Systems **3** (1997), 371–382.

- [135] F. Cong and Y. Li, *Existence of higher dimensional invariant tori for Hamiltonian systems*, J. Math. Anal. Appl. **222** (1998), 255–267.
- [136] F. Cong, Y. Li and D. Jin, *Invariant hyperbolic tori for Hamiltonian systems with Rüssmann nondegeneracy conditions*, Rocky Mountain J. Math. **29** (1999), 831–851.
- [137] H. Cremer, *Zum Zentrumproblem*, Math. Ann. **98** (1927), 151–163.
- [138] R.H. Cushman and L.M. Bates, *Global Aspects of Classical Integrable Systems*, Birkhäuser, Basel (1997).
- [139] R.H. Cushman and J.J. Duistermaat, *The quantum mechanical spherical pendulum*, Bull. Amer. Math. Soc. (N.S.) **19** (1988), 475–479.
- [140] R.H. Cushman, H.R. Dullin, A. Giacobbe, D.D. Holm, M. Joyeux, P. Lynch, D.A. Sadovskii and B.I. Zhilinskii, *CO<sub>2</sub> molecule as a quantum realization of the 1 : 1 : 2 resonant swing-spring with monodromy*, Phys. Rev. Lett. **93** (2004), 024302.
- [141] R.H. Cushman and D.A. Sadovskii, *Monodromy in the hydrogen atom in crossed fields*, Physica D **142** (2000), 166–196.
- [142] R.H. Cushman and J.-C. van der Meer, *The Hamiltonian Hopf bifurcation in the Lagrange top*, Géométrie Symplectique et Mécanique (La Grande Motte, 1988), C. Albert, ed., Lecture Notes in Math., Vol. 1416, Springer, Berlin (1990), 26–38.
- [143] J.S. Dani, *Density properties of orbits under discrete groups*, J. Indian Math. Soc. (N.S.) **39** (1975), 189–217.
- [144] S.G. Dani and S. Raghavan, *Orbits of Euclidean frames under discrete linear groups*, Israel J. Math. **36** (1980), 300–320.
- [145] H.H. de Jong, *Quasiperiodic breathers in systems of weakly coupled pendulums: Applications of KAM theory to classical and statistical mechanics*, Ph.D. Thesis, Univ. Groningen (1999).
- [146] R. de la Llave, *Introduction to KAM Theory*, Computational Physics (Almuñécar, 1992), P.L. Garrido and J. Marro, eds, World Scientific, River Edge, NJ (1993), 73–105.
- [147] R. de la Llave, *A tutorial on KAM Theory*, Smooth Ergodic Theory and its Applications (Seattle, 1999), A. Katok, R. de la Llave, Ya. Pesin and H. Weiss, eds, Proc. Symposia Pure Mathematics, Vol. 69, Amer. Math. Soc., Providence, RI (2001), 175–292.
- [148] R. de la Llave, A. González, Æ. Jorba and J. Villanueva, *KAM theory without action-angle variables*, Nonlinearity **18** (2005), 855–895.
- [149] R. de la Llave and C.E. Wayne, *Whiskered and low dimensional tori in nearly integrable Hamiltonian systems*, Math. Phys. Electronic J. **10** (2004), paper 5.
- [150] A. Delshams, R. de la Llave and T.M. Seara, *A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model*, Mem. Amer. Math. Soc. **179** (844) (2006), i–vii and 1–141.
- [151] A. Delshams and P. Gutiérrez, *Effective stability and KAM theory*, J. Differential Equations **128** (1996), 415–490.
- [152] A. Delshams and P. Gutiérrez, *Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system*, J. Differential Equations **131** (1996), 277–303.
- [153] A. Delshams and P. Gutiérrez, *Exponentially small estimates for KAM theorem near an elliptic equilibrium point*, Hamiltonian Systems with Three or More Degrees of Freedom (S’Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 386–390.
- [154] R.L. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd edn, Addison-Wesley, Redwood City, CA (1989).
- [155] F. Diacu and P. Holmes, *Celestial Encounters. The Origins of Chaos and Stability*, Princeton Univ. Press, Princeton, NJ (1996).
- [156] E.I. Dinaburg and Ya.G. Sinai, *The one-dimensional Schrödinger equation with a quasiperiodic potential*, Funct. Anal. Appl. **9** (1975), 279–289 (English; Russian original).
- [157] J.J. Duistermaat, *On global action-angle coordinates*, Comm. Pure Appl. Math. **33** (1980), 687–706.
- [158] J.J. Duistermaat, *The monodromy in the Hamiltonian Hopf bifurcation*, Z. Angew. Math. Phys. **49** (1998), 156–161.
- [159] H.R. Dullin, A. Giacobbe and R.H. Cushman, *Monodromy in the resonant swing spring*, Physica D **190** (2004), 15–37.
- [160] K. Efstathiou, *Metamorphoses of Hamiltonian Systems with Symmetries*, Lecture Notes in Math., Vol. 1864, Springer, Berlin (2005).



- [161] K. Efsthathiou, R.H. Cushman and D.A. Sadovskii, *Hamiltonian Hopf bifurcation of the hydrogen atom in crossed fields*, *Physica D* **194** (2004), 250–274.
- [162] K. Efsthathiou and D.A. Sadovskii, *Perturbations of the 1 : 1 : 1 resonance with tetrahedral symmetry: a three degree of freedom analogue of the two degree of freedom Hénon–Heiles Hamiltonian*, *Nonlinearity* **17** (2004), 415–446.
- [163] L.H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, *Ann. Sc. Norm. Super. Pisa, Sci. Fis. Mat., IV Ser.* **15** (1988), 115–147.
- [164] L.H. Eliasson, *Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation*, *Comm. Math. Phys.* **146** (1992), 447–482.
- [165] L.H. Eliasson, *Biasymptotic solutions of perturbed integrable Hamiltonian systems*, *Bol. Soc. Brasil. Mat. (N.S.)* **25** (1994), 57–76.
- [166] L.H. Eliasson, *Absolutely convergent series expansions for quasi periodic motions*, *Math. Phys. Electronic J.* **2** (1996), paper 4.
- [167] L.H. Eliasson, *Ergodic skew-systems on  $\mathbb{T}^d \times SO(3, \mathbb{R})$* , *Ergodic Theory Dynam. Systems* **22** (2002), 1429–1449.
- [168] L.H. Eliasson, S.B. Kuksin, S. Marmi and J.-C. Yoccoz, *Dynamical Systems and Small Divisors (Cetraro, 1998)*, S. Marmi and J.-C. Yoccoz, eds, *Lecture Notes in Math.*, Vol. 1784, Springer, Berlin; Centro Internazionale Matematico Estivo, Florence (2002).
- [169] A.Ë. Erëmenko and M.Yu. Lyubich, *The dynamics of analytic transformations*, *Leningrad Math. J.* **1** (1990), 563–634 (English; Russian original).
- [170] R. Fabbri, R. Johnson and R. Pavani, *On the nature of the spectrum of the quasi-periodic Schrödinger operator*, *Nonlinear Anal. Real World Appl.* **3** (1) (2002), 37–59.
- [171] F. Fassò, *Hamiltonian perturbation theory on a manifold*, *Celest. Mech. Dynam. Astronomy* **62** (1995), 43–69.
- [172] F. Fassò, M. Guzzo and G. Benettin, *Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems*, *Comm. Math. Phys.* **197** (1998), 347–360.
- [173] J. Féjoz, *Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après Herman)*, *Ergodic Theory Dynam. Systems* **24** (2004), 1521–1582; A revised version is available at <http://www.math.jussieu.fr/~fejoz>.
- [174] J. Feldman and E. Trubowitz, *Renormalization in classical mechanics and many-body quantum field theory*, *J. Anal. Math.* **58** (1992), 213–247.
- [175] N. Fenichel, *Persistence and smoothness of invariant manifolds for flows*, *Indiana Univ. Math. J.* **21** (1971), 193–226.
- [176] B. Fiedler and P. Poláčik, *Complicated dynamics of scalar reaction diffusion equations with a nonlocal term*, *Proc. Roy. Soc. Edinburgh, Sect. A* **115** (1990), 167–192.
- [177] C. Froeschlé, A. Giorgilli, E. Lega and A. Morbidelli, *On the measure of the structure around an invariant KAM torus – Analytical and numerical investigation*, *Dynamics, Ephemerides and Astrometry of the Solar System – IAU Symposia* **172** (1996), 293–298.
- [178] C. Froeschlé and E. Lega, *On the measure of the structure around the last KAM torus before and after its break-up*, *Celest. Mech. Dynam. Astronomy* **64** (1996), 21–31.
- [179] J. Fröhlich and T. Spencer, *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, *Comm. Math. Phys.* **88** (1983), 151–184.
- [180] D.M. Galin, *Versal deformations of linear Hamiltonian systems*, *Trudy Sem. Imeni I.G. Petrovskogo* **1** (1975), 63–74 (in Russian); English translation: *Amer. Math. Soc. Transl., Ser. 2* **118** (1982), 1–12.
- [181] G. Gallavotti, *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review*, *Rev. Math. Phys.* **6** (1994), 343–411.
- [182] G. Gallavotti, *Lindstedt series and Kolmogorov theorem*, *Hamiltonian Systems with Three or More Degrees of Freedom (S’Agaró, 1995)*, C. Simó, ed., *NATO ASI Series C: Math. Phys. Sci.*, Vol. 533, Kluwer Academic, Dordrecht (1999), 62–71.
- [183] G. Gallavotti, *Statistical Mechanics. A Short Treatise*, Springer, Berlin (1999).
- [184] G. Gallavotti, F. Bonetto and G. Gentile, *Aspects of Ergodic. Qualitative and Statistical Theory of Motion*, Springer, Berlin (2004).
- [185] G. Gallavotti and G. Gentile, *Hyperbolic low-dimensional invariant tori and summations of divergent series*, *Comm. Math. Phys.* **227** (2002), 421–460.

- [186] G. Gallavotti, G. Gentile and A. Giuliani, *Fractional Lindstedt series*, J. Math. Phys. **47** (2006), 012702.
- [187] G. Gallavotti, G. Gentile and V. Mastropietro, *Field theory and KAM tori*, Math. Phys. Electronic J. **1** (1995), paper 5.
- [188] J.-M. Gambaudo, *Perturbation of a Hopf bifurcation by an external time-periodic forcing*, J. Differential Equations **57** (1985), 172–199.
- [189] G. Gentile, *Degenerate lower-dimensional tori under the Bryuno condition*, Ergodic Theory Dynam. Systems **27** (2007), 427–457.
- [190] G. Gentile and G. Gallavotti, *Degenerate elliptic resonances*, Comm. Math. Phys. **257** (2005), 319–362.
- [191] G. Gentile and V. Mastropietro, *Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications*, Rev. Math. Phys. **8** (1996), 393–444.
- [192] C.G. Gibson, *Singular Points of Smooth Mappings*, Research Notes in Math., Vol. 25, Pitman, Boston, MA (1979).
- [193] M. Gidea and R. de la Llave, *Topological methods in the instability problem of Hamiltonian systems*, Discrete Contin. Dynam. Systems, Ser. A **14** (2006), 295–328.
- [194] A. Giorgilli, *Exponential stability of Hamiltonian systems*, Dynamical Systems, Part I. Hamiltonian Systems and Celestial Mechanics (Pisa, 2002), Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa (2003), 87–198.
- [195] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani and C. Simó, *Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem*, J. Differential Equations **77** (1989), 167–198.
- [196] A. Giorgilli and U. Locatelli, *A classical self-contained proof of Kolmogorov's theorem on invariant tori*, Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 72–89.
- [197] A. Giorgilli and A. Morbidelli, *Invariant KAM tori and global stability for Hamiltonian systems*, Z. Angew. Math. Phys. **48** (1997), 102–134.
- [198] A. González-Enríquez and R. de la Llave, *Analytic smoothing of geometric maps with applications to KAM theory*, J. Differential Equations **245** (2008), 1243–1298.
- [199] A. Gorodnik, *Uniform distribution of orbits of lattices on spaces of frames*, Duke Math. J. **122** (2004), 549–589.
- [200] S.M. Graff, *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations **15** (1974), 1–69.
- [201] J.M. Greene, *The status of KAM Theory from a physicist's point of view*, Chaos in Australia (Sydney, 1990), G. Brown and A. Opie, eds, World Scientific, River Edge, NJ (1993), 8–23.
- [202] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, 5th edn, Appl. Math. Sciences, Vol. 42, Springer, New York (1997).
- [203] F. Gustavson, *On constructing formal integrals of a Hamiltonian system near an equilibrium point*, Astrophys. J. **71** (1966), 670–686.
- [204] J.K. Hale, *Ordinary Differential Equations*, 2nd edn, Robert E. Krieger, Huntington, NY (1980).
- [205] Y. Han, Y. Li and Y. Yi, *Degenerate lower-dimensional tori in Hamiltonian systems*, J. Differential Equations **227** (2006), 670–691.
- [206] H. Hanßmann, *Quasi-periodic motions of a rigid body: A case study on perturbations of superintegrable systems*, Ph.D. Thesis, Univ. Groningen (1995).
- [207] H. Hanßmann, *Quasi-periodic motions of a rigid body I. Quadratic Hamiltonians on the sphere with a distinguished parameter*, Regul. Chaotic Dyn. **2** (2) (1997), 41–57; Erratum: **4** (4) (1999), 100.
- [208] H. Hanßmann, *The quasi-periodic centre-saddle bifurcation*, J. Differential Equations **142** (1998), 305–370.
- [209] H. Hanßmann, *A survey on bifurcations of invariant tori*, New Advances in Celestial Mechanics and Hamiltonian Systems (Guanajuato, 2001), J. Delgado, E.A. Lacomba, J. Llibre and E. Pérez-Chavela, eds, Kluwer Academic/Plenum Press, New York (2004), 109–121.
- [210] H. Hanßmann, *Hamiltonian torus bifurcations related to simple singularities*, Dynamic Systems and Applications, Vol. 4 (Atlanta, 2003), G.S. Ladde, N.G. Medhin and M. Sambandham, eds, Dynamic Publishers, Atlanta, GA (2004), 679–685.
- [211] H. Hanßmann, *On Hamiltonian bifurcations of invariant tori with a Floquet multiplier  $-1$* , Dyn. Syst. **21** (2) (2006), 115–145.

- [212] H. Hanßmann, *Local and Semi-Local Bifurcations in Hamiltonian Dynamical Systems. Results and Examples*, Lecture Notes in Math., Vol. 1893, Springer, Berlin (2007).
- [213] À. Haro, *Converse KAM theory for monotone positive symplectomorphisms*, *Nonlinearity* **12** (1999), 1299–1322.
- [214] F. Hausdorff, *Set Theory*, 2nd edn, Chelsea, New York (1962) (English; German original).
- [215] M.R. Herman, *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations*, *Publ. Math. IHÉS* **49** (1979), 5–233.
- [216] M.R. Herman, *Démonstration de théorème des courbes translatées par les difféomorphismes de l’anneau; démonstration du théorème des tores invariants*, Manuscript (1980).
- [217] M.R. Herman, *Construction d’un difféomorphisme minimal d’entropie topologique non nulle*, *Ergodic Theory Dynam. Systems* **1** (1981), 65–76.
- [218] M.R. Herman, *Sur les courbes invariantes par les difféomorphismes de l’anneau*, Vol. 1 and 2, *Astérisque* **103–104** (1983), i and 1–221; **144** (1986), 1–248.
- [219] M.R. Herman, *Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d’un théorème d’Arnold et de Moser sur le tore de dimension 2*, *Comment. Math. Helvetici* **58** (1983), 453–502.
- [220] M.R. Herman, *Simple proofs of local conjugacy theorems for diffeomorphisms of the circle with almost every rotation number*, *Bol. Soc. Brasil. Mat.* **16** (1985), 45–83.
- [221] M.R. Herman, *Existence et non existence de tores invariants par des difféomorphismes symplectiques*, *Séminaire sur les Équations aux Dérivées Partielles*, Vol. 1987/88, Exp. 14, École Polytechnique, Centre de Mathématiques, Palaiseau (1988).
- [222] M.R. Herman, *Inégalités “a priori” pour des tores lagrangiens invariants par des difféomorphismes symplectiques*, *Publ. Math. IHÉS* **70** (1989), 47–101.
- [223] M.R. Herman, *Différentiabilité optimale et contre-exemples à la fermeture en topologie  $C^\infty$  des orbites récurrentes de flots hamiltoniens*, *C. R. Acad. Sci. Paris, Sér. I Math.* **313** (1991), 49–51.
- [224] M.R. Herman, *Exemples de flots hamiltoniens dont aucune perturbation en topologie  $C^\infty$  n’a d’orbites périodiques sur un ouvert de surfaces d’énergies*, *C. R. Acad. Sci. Paris, Sér. I Math.* **312** (1991), 989–994.
- [225] M.W. Hirsch, *Differential Topology*, 2nd edn, Springer, New York (1994).
- [226] M.W. Hirsch, C.C. Pugh and M. Shub, *Invariant Manifolds*, Lecture Notes in Math., Vol. 583, Springer, Berlin (1977).
- [227] J. Hoo, *Quasi-periodic bifurcations in a strong resonance: Combination tones in gyroscopic stabilisation*, Ph.D. Thesis, Univ. Groningen (2005).
- [228] E. Hopf, *Repeated branching through loss of stability, an example*, *Proceedings of the Conference on Differential Equations* (Dedicated to A. Weinstein), J.B. Diaz, ed., Univ. Maryland Book Store, College Park, MD (1956), 49–56.
- [229] E. Hopf, *A mathematical example displaying features of turbulence*, *Comm. Appl. Math.* **1** (1948), 303–322.
- [230] E. Horozov, *Perturbations of the spherical pendulum and abelian integrals*, *J. Reine Angew. Math.* **408** (1990), 114–135.
- [231] Q. Huang, F. Cong and Y. Li, *Persistence of hyperbolic invariant tori for Hamiltonian systems*, *J. Differential Equations* **164** (2000), 355–379.
- [232] Q. Huang, F. Cong and Y. Li, *Persistence of elliptic invariant tori for Hamiltonian systems*, *Nonlinear Anal.* **45** (2001), 241–260.
- [233] D. Huang and Z. Liu, *On the persistence of lower-dimensional invariant hyperbolic tori for smooth Hamiltonian systems*, *Nonlinearity* **13** (2000), 189–202.
- [234] J.H. Hubbard, *The KAM theorem*, *Kolmogorov’s Heritage in Mathematics*, É. Charpentier, A. Lesne and N.K. Nikolski, eds, Springer, Berlin (2007), 215–238.
- [235] J.H. Hubbard and Yu.S. Il’yashenko, *A proof of Kolmogorov’s theorem*, *Discrete Contin. Dynam. Systems, Ser. A* **10** (2004), 367–385.
- [236] G.B. Huitema, *Unfoldings of quasi-periodic tori*, Ph.D. Thesis, Univ. Groningen (1988).
- [237] Chr. Huygens, *Appendix I to “De vi centrifuga”* (1659), 302–311; *Œuvres Complètes de Christiaan Huygens*, published by the Société Hollandaise des Sciences, Vol. 16, Martinus Nijhoff, The Hague (1929), 306–309.<sup>12</sup>

<sup>12</sup>The text “De vi centrifuga” was published posthumously in 1704, but the appendix remained unpublished until the edition of the *Œuvres*. So, during Huygens’ lifetime the *Horologium Oscillatorium* of 1673 was the publication which came nearest to his knowledge of circular motion and centrifugal force.

- [238] G. Iooss and J.E. Los, *Quasi-genericity of bifurcations to high dimensional invariant tori for maps*, Comm. Math. Phys. **119** (1988), 453–500.
- [239] F.M. Izrailev and B.V. Chirikov, *Stochasticity of the simplest dynamical model with divided phase space*, The USSR Academy of Sciences, Novosibirsk Institute for Nuclear Physics, Preprint 191 (1968) (in Russian).
- [240] R. Johnson and J.K. Moser, *The rotation number for almost periodic potentials*, Comm. Math. Phys. **84** (1982), 403–438; Erratum: **90** (1983), 317–318.
- [241] À. Jorba, R. de la Llave and M. Zou, *Lindstedt series for lower dimensional tori*, Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 151–167.
- [242] À. Jorba and J. Villanueva, *On the persistence of lower dimensional invariant tori under quasi-periodic perturbations*, J. Nonlinear Sci. **7** (1997), 427–473.
- [243] À. Jorba and J. Villanueva, *On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems*, Nonlinearity **10** (1997), 783–822.
- [244] À. Jorba and J. Villanueva, *The fine geometry of the Cantor families of invariant tori in Hamiltonian systems*, European Congress of Mathematics, Vol. 2, (Barcelona, 2000), C. Casacuberta, R.M. Miró-Roig, J. Verdera and S. Xambó-Descamps, eds, Progr. Math., Vol. 202, Birkhäuser, Basel (2001), 557–564.
- [245] T. Kappeler and J. Pöschel, *KdV & KAM*, Springer, Berlin (2003).
- [246] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, with a supplementary chapter by A. Katok and L. Mendoza, Encyclopedia of Mathematics and its Applications, Vol. 54, Cambridge Univ. Press, Cambridge (1995).
- [247] A.Ya. Khinchin, *Continued Fractions*, 2nd edn, Dover Publications, Mineola, NY (1997) (English; Russian original).
- [248] A. Knauf, *Closed orbits and converse KAM theory*, Nonlinearity **3** (1990), 961–973.
- [249] H. Koçak, *Normal forms and versal deformations of linear Hamiltonian systems*, J. Differential Equations **51** (1984), 359–407.
- [250] A.N. Kolmogorov, *The general theory of dynamical systems and classical mechanics*, Proceedings of the International Congress of Mathematicians, Vol. 1 (Amsterdam, 1954), J.C.H. Gerretsen and J. de Groot, eds, North-Holland, Amsterdam (1957), 315–333 (in Russian); Reprinted in: International Mathematical Congress in Amsterdam, 1954 (Plenary Lectures), Fizmatgiz, Moscow (1961), 187–208; English translation as Appendix D in: R.H. Abraham, *Foundations of Mechanics*, Benjamin, New York (1967), 263–279; Reprinted as Appendix in: R.H. Abraham and J.E. Marsden, *Foundations of Mechanics*, 2nd edn, Benjamin/Cummings, Reading, MA (1978), 741–757.
- [251] A.N. Kolmogorov, *On the persistence of conditionally periodic motions under a small change of the Hamilton function*, Dokl. Akad. Nauk SSSR **98** (1954), 527–530 (in Russian); English translation in: Stochastic Behavior in Classical and Quantum Hamiltonian Systems, Volta Memorial Conference (Como, 1977), G. Casati and J. Ford, eds, Lecture Notes in Phys., Vol. 93, Springer, Berlin (1979), 51–56; Reprinted in: Bai Lin Hao, ed, *Chaos*, World Scientific, Singapore (1984), 81–86.
- [252] R. Krikorian,  *$C^0$ -densité globale des systèmes produits-croisés sur le cercle réductibles*, Ergodic Theory Dynam. Systems **19** (1999), 61–100.
- [253] R. Krikorian, *Réductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts*, Ann. Sci. École Norm. Super. (4) **32** (2) (1999), 187–240.
- [254] A.A. Kubichka and I.O. Parasyuk, *A Whitney differentiable family of coisotropic invariant tori for a nearly degenerate Hamiltonian system*, Visn. Mat. Mekh., Kyiv. Univ. Imeni Tarasa Shevchenka (4) (2000); 20–29 (in Ukrainian).
- [255] A.A. Kubichka and I.O. Parasyuk, *Bifurcation of a Whitney smooth family of coisotropic invariant tori of a Hamiltonian system under a small deformation of the symplectic structure*, Ukrain. Math. J. **53** (2001), 701–718 (English; Ukrainian original).
- [256] S.B. Kuksin, *Nearly Integrable Infinite-Dimensional Hamiltonian Systems*, Lecture Notes in Math., Vol. 1556, Springer, Berlin (1993).
- [257] S.B. Kuksin, *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its Applications, Vol. 19, Oxford Univ. Press, Oxford (2000).
- [258] S.B. Kuksin, *Fifteen years of KAM for PDE*, Geometry, Topology, and Mathematical Physics. S.P. Novikov's Seminar: 2002–2003, V.M. Buchstaber and I.M. Krichever, eds, Amer. Math. Soc. Transl., Ser. 2 **212**; Adv. Math. Sci. **55** (2004), 237–258.

- [259] S.B. Kuksin, V.F. Lazutkin and J. Pöschel, eds, *Seminar on Dynamical Systems (St. Petersburg, 1991)*, Progress in Nonlinear Differential Equations and their Applications, Vol. 12, Birkhäuser, Basel (1994).
- [260] Yu.A. Kuznetsov, *Elements of Applied Bifurcation Theory*, 3rd edn, Appl. Math. Sciences, Vol. 112, Springer, New York (2004).
- [261] Yu.A. Kuznetsov and H.G.E. Meijer, *Numerical normal forms for codim 2 bifurcations of fixed points with at most two critical eigenvalues*, SIAM J. Sci. Comput. **26** (2005), 1932–1954.
- [262] J.S.W. Lamb and J.A.G. Roberts, *Time-reversal symmetry in dynamical systems: A survey*, Physica D **112** (1998), 1–39.
- [263] L.D. Landau, *On the problem of turbulence*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **44** (1944), 311–314 (English; Russian original).
- [264] L.D. Landau and E.M. Lifshitz, *Fluid Mechanics*, 2nd edn, Pergamon, Oxford (1987) (English; Russian original).
- [265] J. Laskar, *Large scale chaos and marginal stability in the Solar System*, XIth International Congress of Mathematical Physics (Paris, 1994), D. Iagolnitzer, ed., Internat. Press, Cambridge, MA (1995), 75–120.
- [266] J. Laskar and P. Robutel, *Stability of the planetary three-body problem. I. Expansion of the planetary Hamiltonian*, Celest. Mech. Dynam. Astronomy **62** (1995), 193–217.
- [267] V.F. Lazutkin, *Existence of a continuum of closed invariant curves for a convex billiard*, Uspekhi Mat. Nauk **27** (3) (1972), 201–202 (in Russian).
- [268] V.F. Lazutkin, *KAM Theory and Semiclassical Approximations to Eigenfunctions*, with an addendum by A.I. Shnirelman, Springer, Berlin (1993).
- [269] E. Lega and C. Froeschlé, *Numerical investigations of the structure around an invariant KAM torus using the frequency map analysis*, Physica D **95** (1996), 97–106.
- [270] A. Litvak-Hinenzon and V. Rom-Kedar, *Parabolic resonances in 3 degree of freedom near-integrable Hamiltonian systems*, Physica D **164** (2002), 213–250; Erratum: **172** (2002), 217.
- [271] A. Litvak-Hinenzon and V. Rom-Kedar, *Resonant tori and instabilities in Hamiltonian systems*, Nonlinearity **15** (2002), 1149–1177.
- [272] A. Litvak-Hinenzon and V. Rom-Kedar, *On energy surfaces and the resonance web*, SIAM J. Appl. Dynam. Syst. **3** (2004), 525–573.
- [273] B. Liu, *On lower dimensional invariant tori in reversible systems*, J. Differential Equations **176** (2001), 158–194.
- [274] Zh. Liu, *Persistence of lower dimensional invariant tori on sub-manifolds in Hamiltonian systems*, Nonlinear Anal. **61** (2005), 1319–1342.
- [275] Zh. Liu, D. Yihe and Q. Huang, *Persistence of hyperbolic tori in generalized Hamiltonian systems*, Northeast. Math. J. **21** (2005), 447–464.
- [276] B.-f. Liu, W.-zh. Zhu and Y. Han, *Persistence of lower-dimensional hyperbolic invariant tori for generalized Hamiltonian systems*, J. Math. Anal. Appl. **322** (2006), 251–275.
- [277] Y. Li and Y. Yi, *Persistence of invariant tori in generalized Hamiltonian systems*, Ergodic Theory Dynam. Systems **22** (2002), 1233–1261.
- [278] Y. Li and Y. Yi, *A quasi-periodic Poincaré’s theorem*, Math. Ann. **326** (2003), 649–690.
- [279] Y. Li and Y. Yi, *On Poincaré-Treshchëv tori in Hamiltonian systems*, EQUADIFF 2003 (Proc. Intern. Conf. Differential Equations) (Hasselt/Diepenbeek, 2003), F. Dumortier, H.W. Broer, J. Mawhin, A. Vanderbauwhede and S.M. Verduyn Lunel, eds, World Scientific, Hackensack, NJ (2005), 136–151.
- [280] Y. Li and Y. Yi, *Persistence of hyperbolic tori in Hamiltonian systems*, J. Differential Equations **208** (2005), 344–387.
- [281] Y. Li and Y. Yi, *Persistence of lower dimensional tori of general types in Hamiltonian systems*, Trans. Amer. Math. Soc. **357** (2005), 1565–1600.
- [282] Y. Li and Y. Yi, *Nekhoroshev and KAM stabilities in generalized Hamiltonian systems*, J. Dynam. Differential Equations **18** (2006), 577–614.
- [283] P. Lochak, *Canonical perturbation theory via simultaneous approximation*, Russian Math. Surveys **47** (6) (1992), 57–133.
- [284] P. Lochak, *Hamiltonian perturbation theory: periodic orbits, resonances and intermittency*, Nonlinearity **6** (1993), 885–904.
- [285] P. Lochak, *Arnold diffusion; a compendium of remarks and questions*, Hamiltonian Systems with Three or More Degrees of Freedom (S’Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 168–183.

- [286] P. Lochak and J.-P. Marco, *Diffusion times and stability exponents for nearly integrable analytic systems*, Central European J. Math. **3** (2005), 342–397.
- [287] P. Lochak and A.I. Neishtadt, *Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian*, Chaos **2** (1992), 495–499.
- [288] P. Lochak, A.I. Neishtadt and L. Niederman, *Stability of nearly integrable convex Hamiltonian systems over exponentially long times*, Seminar on Dynamical Systems (St. Petersburg, 1991), S.B. Kuksin, V.F. Lazutkin and J. Pöschel, eds, Progress in Nonlinear Differential Equations and their Applications, Vol. 12, Birkhäuser, Basel (1994), 15–34.
- [289] J.E. Los, *Dédoublement de courbes invariantes sur le cylindre: petits diviseurs*, Ann. Inst. H. Poincaré, Anal. Non Linéaire **5** (1) (1988), 37–95.
- [290] Yu.V. Loveĭkin and I.O. Parasyuk, *General KAM theorem for coisotropic invariant tori of Hamiltonian systems*, Visn. Mat. Mekh., Kyiv. Univ. Imeni Tarasa Shevchenka (11) (2004), 53–58 (in Ukrainian).
- [291] Yu.V. Loveĭkin and I.O. Parasyuk, *Theorem on a perturbation of coisotropic invariant tori of locally Hamiltonian systems and its applications*, Nonlinear Oscillations **8** (2005), 487–512 (English; Ukrainian original).
- [292] Yu.V. Loveĭkin and I.O. Parasyuk, *Bifurcation of coisotropic invariant tori under locally Hamiltonian perturbations of integrable systems and a nondegenerate deformation of the symplectic structure*, Nonlinear Oscillations **9** (2006), 215–225 (English; Ukrainian original).
- [293] Yu.V. Loveĭkin and I.O. Parasyuk, *Invariant tori of locally Hamiltonian systems close to conditionally integrable ones*, Ukrain. Math. J. **59** (2007), 70–99 (English; Ukrainian original).
- [294] G.A. Lunter, *Bifurcations in Hamiltonian systems: Computing singularities by Gröbner bases*, Ph.D. Thesis, Univ. Groningen (1999).
- [295] M.Yu. Lyubich, *The dynamics of rational transformations: a topological picture*, Russian Math. Surveys **41** (4) (1986), 43–117 (English; Russian original).
- [296] R.S. MacKay, *A criterion for nonexistence of invariant tori for Hamiltonian systems*, Physica D **36** (1989), 64–82.
- [297] R.S. MacKay and S. Aubry, *Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators*, Nonlinearity **7** (1994), 1623–1643.
- [298] R.S. MacKay, J.D. Meiss and J. Stark, *Converse KAM theory for symplectic twist maps*, Nonlinearity **2** (1989), 555–570.
- [299] R.S. MacKay and I.C. Percival, *Converse KAM: theory and practice*, Comm. Math. Phys. **98** (1985), 469–512.
- [300] B. Malgrange, *Ideals of Differentiable Functions*, Oxford Univ. Press, London (1967).
- [301] R. Mañé, *Ergodic Theory and Differentiable Dynamics*, Springer, Berlin (1987) (English; Portuguese original).
- [302] J.-P. Marco and D. Sauzin, *Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems*, Publ. Math. IHÉS **96** (2003), 199–275.
- [303] J.-P. Marco and D. Sauzin, *Wandering domains and random walks in Gevrey near-integrable systems*, Ergodic Theory Dynam. Systems **24** (2004), 1619–1666.
- [304] S. Marmi, *An Introduction to Small Divisors Problems*, Dipartimento di Matematica dell'Univ. Pisa, Pisa (2000).
- [305] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems*, 2nd edn, Springer, New York (1999).
- [306] J.N. Mather, *A criterion for the nonexistence of invariant circles*, Publ. Math. IHÉS **63** (1986), 153–204.
- [307] J.N. Mather, *Arnold diffusion. I: Announcement of results*, J. Math. Sci. **124** (2004), 5275–5289.
- [308] V.S. Matveev, *Integrable Hamiltonian systems with two degrees of freedom. Topological structure of saturated neighbourhoods of points of focus-focus and saddle-saddle types*, Sb. Math. **187** (1996), 495–524 (English; Russian original).
- [309] V.K. Melnikov, *On some cases of the persistence of conditionally periodic motions under a small change of the Hamilton function*, Soviet Math. Dokl. **6** (1965), 1592–1596 (English; Russian original).
- [310] V.K. Melnikov, *On a certain family of conditionally periodic solutions of a Hamiltonian system*, Soviet Math. Dokl. **9** (1968), 882–886 (English; Russian original).
- [311] J.W. Milnor, *Dynamics in One Complex Variable*, 3rd edn, Ann. Math. Studies, Vol. 160, Princeton Univ. Press, Princeton, NJ (2006).

- [312] R. Montgomery, *The connection whose holonomy is the classical adiabatic angles of Hannay and Berry and its generalization to the non-integrable case*, Comm. Math. Phys. **120** (1988), 269–294.
- [313] A. Morbidelli, *Bounds on diffusion in phase space: Connection between Nekhoroshev and KAM theorems and superexponential stability of invariant tori*, Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 514–517.
- [314] A. Morbidelli and A. Giorgilli, *Superexponential stability of KAM tori*, J. Stat. Phys. **78** (1995), 1607–1617.
- [315] A. Morbidelli and A. Giorgilli, *On a connection between KAM and Nekhoroshev's theorems*, Physica D **86** (1995), 514–516.
- [316] J.K. Moser, *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen, Math.-Phys. Kl. II **1** (1962), 1–20.
- [317] J.K. Moser, *Combination tones for Duffing's equation*, Comm. Pure Appl. Math. **18** (1965), 167–181.
- [318] J.K. Moser, *On the theory of quasiperiodic motions*, SIAM Rev. **8** (2) (1966), 145–172.
- [319] J.K. Moser, *Convergent series expansions for quasi-periodic motions*, Math. Ann. **169** (1967), 136–176.
- [320] J.K. Moser, *Lectures on Hamiltonian systems*, Mem. Amer. Math. Soc. **81** (1968), 1–60.
- [321] J.K. Moser, *Old and new applications of KAM Theory*, Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 184–192.
- [322] J.K. Moser, *Remark on the paper "On invariant curves of area-preserving mappings of an annulus"*, Regul. Chaotic Dyn. **6** (2001), 337–338.
- [323] J.K. Moser, *Stable and Random Motions in Dynamical Systems, with Special Emphasis on Celestial Mechanics*, 2nd edn, Ann. Math. Studies, Vol. 77, Princeton Univ. Press, Princeton, NJ (2001).
- [324] J.K. Moser and J. Pöschel, *An extension of a result by Dinaburg and Sinaĭ on quasi-periodic potentials*, Comment. Math. Helvetici **59** (1984), 39–85.
- [325] J.R. Munkres, *Elementary Differential Topology*, 2nd edn, Ann. Math. Studies, Vol. 54, Princeton Univ. Press, Princeton, NJ (1966).
- [326] N.N. Nekhoroshev, *On the behaviour of Hamiltonian systems close to integrable ones*, Funct. Anal. Appl. **5** (1971), 338–339 (English; Russian original).
- [327] N.N. Nekhoroshev, *Action-angle variables and their generalizations*, Trans. Moscow Math. Soc. **26** (1972), 180–198 (English; Russian original).
- [328] N.N. Nekhoroshev, *An exponential estimate of the stability time of nearly integrable Hamiltonian systems. I*, Russian Math. Surveys **32** (6) (1977), 1–65 (English; Russian original).
- [329] N.N. Nekhoroshev, *An exponential estimate of the stability time of nearly integrable Hamiltonian systems. II*, Trudy Sem. Imeni I.G. Petrovskogo **5** (1979), 5–50 (in Russian); English translation: Topics in Modern Mathematics, Petrovskii Seminar No. 5, O.A. Oleĭnik, ed, Consultants Bureau, New York (1985), 1–58.
- [330] A.I. Neĭshadt, *Estimates in the Kolmogorov theorem on the persistence of conditionally periodic motions*, J. Appl. Math. Mech. **45** (1981), 766–772 (English; Russian original).
- [331] A.I. Neĭshadt, *On the separation of motions in systems with a rapidly rotating phase*, J. Appl. Math. Mech. **48** (1984), 133–139 (English; Russian original).
- [332] S.E. Newhouse, J. Palis and F. Takens, *Bifurcations and stability of families of diffeomorphisms*, Publ. Math. IHÉS **57** (1983), 5–71.
- [333] S.E. Newhouse, D. Ruelle and F. Takens, *Occurrence of strange Axiom A attractors near quasi-periodic flows on  $\mathbb{T}^m$ ,  $m \geq 3$* , Comm. Math. Phys. **64** (1978), 35–40.
- [334] L. Niederman, *Nonlinear stability around an elliptic equilibrium point in a Hamiltonian system*, Nonlinearity **11** (1998), 1465–1479.
- [335] L. Niederman, *Dynamics around simple resonant tori in nearly integrable Hamiltonian systems*, J. Differential Equations **161** (2000), 1–41.
- [336] L. Niederman, *Exponential stability for small perturbations of steep integrable Hamiltonian systems*, Ergodic Theory Dynam. Systems **24** (2004), 593–608.
- [337] M. Ollé, J.R. Pacha and J. Villanueva, *Kolmogorov–Arnold–Moser aspects of the periodic Hamiltonian Hopf bifurcation*, Nonlinearity **21** (2008), 1759–1811.
- [338] J.C. Oxtoby, *Measure and Category. A Survey of the Analogies between Topological and Measure Spaces*, 2nd edn, Springer, New York (1980).

- [339] J. Palis and W.C. de Melo, *Geometric Theory of Dynamical Systems. An Introduction*, Springer, New York (1982) (English; Portuguese original).
- [340] I.O. Parasyuk, *Persistence of quasi-periodic motions in reversible multifrequency systems*, Dokl. Akad. Nauk Ukrain. SSR, Ser. A (9) (1982), 19–22 (in Russian).
- [341] I.O. Parasyuk, *On the persistence of multidimensional invariant tori of Hamiltonian systems*, Ukrain. Math. J. **36** (1984), 380–385 (English; Russian original).
- [342] I.O. Parasyuk, *Coisotropic invariant tori of locally Hamiltonian systems*, Asymptotic Integration of Differential Equations, Yu.A. Mitropolskiĭ, ed., Institut Matematiki Akad. Nauk USSR, Kiev (1985), 129–133 (in Russian).
- [343] I.O. Parasyuk, *On the preservation of coisotropic invariant tori of locally Hamiltonian systems*, Some Problems in the Theory of Asymptotic Methods in Nonlinear Mechanics, Yu.A. Mitropolskiĭ, A.A. Berezovskii, V.G. Kolomiets, A.K. Lopatin and O.B. Lykova, eds, Institut Matematiki Akad. Nauk USSR, Kiev (1986), 150–154 (in Russian).
- [344] I.O. Parasyuk, *Coisotropic invariant tori of Hamiltonian systems in the quasiclassical theory of the motion of a conduction electron*, Ukrain. Math. J. **42** (1990), 308–312 (English; Russian original).
- [345] I.O. Parasyuk, *Variables of the action-angle type on symplectic manifolds foliated into coisotropic tori*, Ukrain. Math. J. **45** (1993), 85–93 (English; Russian original).
- [346] I.O. Parasyuk, *Coisotropic quasi-periodic motions near the relative equilibrium of a Hamiltonian system*, Nonlinear Math. Phys. **1** (1994), 340–357.
- [347] I.O. Parasyuk, *Bifurcation of a Cantor set of coisotropic invariant tori of a Hamiltonian system under a perturbation of the symplectic structure*, Nelin. Kolyvannya **1** (2) (1998), 81–89 (in Ukrainian).
- [348] I.O. Parasyuk, *Perturbations of degenerate coisotropic invariant tori of Hamiltonian systems*, Ukrain. Math. J. **50** (1998), 83–99 (English; Ukrainian original).
- [349] J. Patera, C. Rousseau and D. Schlomiuk, *Versal deformations of elements of real classical Lie algebras*, J. Phys. A: Math. Gen. **15** (1982), 1063–1086.
- [350] G. Popov, *KAM tori and quantum Birkhoff normal forms*, Séminaire sur les Équations aux Dérivées Partielles, Vol. 1999/2000, Exp. 20, École Polytechnique, Centre de Mathématiques, Palaiseau (2000).
- [351] G. Popov, *Invariant tori, effective stability, and quasimodes with exponentially small error terms, I. — Birkhoff normal forms*, Ann. Henri Poincaré **1** (2000), 223–248.
- [352] G. Popov, *KAM theorem for Gevrey Hamiltonians*, Ergodic Theory Dynam. Systems **24** (2004), 1753–1786.
- [353] G. Popov, *KAM theorem and quasimodes for Gevrey Hamiltonians*, Mat. Contemp. **26** (2004), 87–107.
- [354] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*, Comm. Pure Appl. Math. **35** (1982), 653–696.
- [355] J. Pöschel, *On elliptic lower dimensional tori in Hamiltonian systems*, Math. Z. **202** (1989), 559–608.
- [356] J. Pöschel, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, Math. Z. **213** (1993), 187–216.
- [357] J. Pöschel, *A lecture on the classical KAM theorem*, Smooth Ergodic Theory and its Applications (Seattle, 1999), A. Katok, R. de la Llave, Ya. Pesin and H. Weiss, eds, Proc. Symposia Pure Mathematics, Vol. 69, Amer. Math. Soc., Providence, RI (2001), 707–732.
- [358] J. Puig, *Cantor spectrum for the almost Mathieu operator*, Comm. Math. Phys. **244** (2004), 297–309.
- [359] G.R.W. Quispel and M.B. Sevryuk, *KAM theorems for the product of two involutions of different types*, Chaos **3** (1993), 757–769.
- [360] T.S. Ratiu, R. Tudoran, L. Sbano, E. Sousa Dias and G. Terra, *A crash course in geometric mechanics. Notes of the courses given by Tudor Ratiu*, Geometric Mechanics and Symmetry. The Peyresq Lectures (Peyresq, 2000–01), J. Montaldi and T.S. Ratiu, eds, London Math. Soc. Lecture Note Series, Vol. 306, Cambridge Univ. Press, Cambridge (2005), 23–156.
- [361] W.P. Reinhardt, *Chaos and collisions: introductory concepts*, Collision Theory for Atoms and Molecules (Cortona, 1987), F.A. Gianturco, ed., NATO ASI Series B: Physics, Vol. 196, Plenum Press, New York (1989), 465–518.
- [362] B.W. Rink, *Geometry and dynamics in Hamiltonian lattices*, Ph.D. Thesis, Univ. Utrecht (2003).
- [363] B.W. Rink, *A Cantor set of tori with monodromy near a focus-focus singularity*, Nonlinearity **17** (2004), 347–356.
- [364] J.A.G. Roberts and G.R.W. Quispel, *Chaos and time-reversal symmetry. Order and chaos in reversible dynamical systems*, Phys. Rep. **216** (2–3) (1992), 63–177.



- [365] P. Robutel, *Stability of the planetary three-body problem. II. KAM theory and existence of quasiperiodic motions*, Celest. Mech. Dynam. Astronomy **62** (1995), 219–261; Erratum: **84** (2002), 317.
- [366] A.M. Rockett and P. Szűsz, *Continued Fractions*, World Scientific, River Edge, NJ (1992).
- [367] M. Rudnev and S. Wiggins, *KAM theory near multiplicity one resonant surfaces in perturbations of a-priori stable Hamiltonian systems*, J. Nonlinear Sci. **7** (1997), 177–209.
- [368] M. Rudnev and S. Wiggins, *On a partially hyperbolic KAM theorem*, Regul. Chaotic Dyn. **4** (4) (1999), 39–58.
- [369] D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory*, Academic Press, Boston, MA (1989).
- [370] D. Ruelle and F. Takens, *On the nature of turbulence*, Comm. Math. Phys. **20** (1971), 167–192.
- [371] D. Ruelle and F. Takens, *Note concerning our paper “On the nature of turbulence”*, Comm. Math. Phys. **23** (1971), 343–344.
- [372] H. Rüssmann, *Über die Iteration analytischer Funktionen*, J. Math. Mech. **17** (1967), 523–532.
- [373] H. Rüssmann, *Kleine Nenner I: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes*, Nachr. Akad. Wiss. Göttingen, Math.-Phys. Kl. II **5** (1970), 67–105.
- [374] H. Rüssmann, *Konvergente Reihenentwicklungen in der Störungstheorie der Himmelsmechanik*, Selecta Mathematica. Vth volume, K. Jacobs, ed., Heidelberger Taschenbücher, Vol. 201, Springer, Berlin (1979), 93–260.
- [375] H. Rüssmann, *On the one-dimensional Schrödinger equation with a quasiperiodic potential*, Nonlinear Dynamics (New York, 1979), R.H.G. Helleman, ed., Ann. New York Academy of Sciences, Vol. 357, New York Academy of Sciences, New York (1980), 90–107.
- [376] H. Rüssmann, *Non-degeneracy in the perturbation theory of integrable dynamical systems*, Number Theory and Dynamical Systems (York, 1987), M.M. Dodson and J.A.G. Vickers, eds, London Math. Soc. Lecture Note Series, Vol. 134, Cambridge Univ. Press, Cambridge (1989), 5–18; Reprinted in: Stochastics, Algebra and Analysis in Classical and Quantum Dynamics (Marseille, 1988), S. Albeverio, P. Blanchard and D. Testard, eds, Math. and its Appl., Vol. 59, Kluwer Academic, Dordrecht (1990), 211–223.
- [377] H. Rüssmann, *On the frequencies of quasi periodic solutions of analytic nearly integrable Hamiltonian systems*, Seminar on Dynamical Systems (St. Petersburg, 1991), S.B. Kuksin, V.F. Lazutkin and J. Pöschel, eds, Progress in Nonlinear Differential Equations and their Applications, Vol. 12, Birkhäuser, Basel (1994), 160–183.
- [378] H. Rüssmann, *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regul. Chaotic Dyn. **6** (2001), 119–204.
- [379] H. Rüssmann, *Stability of elliptic fixed points of analytic area-preserving mappings under the Bruno condition*, Ergodic Theory Dynam. Systems **22** (2002), 1551–1573.
- [380] H. Rüssmann, *Convergent transformations into a normal form in analytic Hamiltonian systems with two degrees of freedom on the zero energy surface near degenerate elliptic singularities*, Ergodic Theory Dynam. Systems **24** (2004), 1787–1832.
- [381] H. Rüssmann, *Addendum to “Invariant tori in non-degenerate nearly integrable Hamiltonian systems”*, Regul. Chaotic Dyn. **10** (2005), 21–31.
- [382] D.A. Salamon and E. Zehnder, *KAM theory in configuration space*, Comment. Math. Helvetici **64** (1989), 84–132.
- [383] D.A. Salamon, *The Kolmogorov–Arnold–Moser theorem*, Math. Phys. Electronic J. **10** (2004), paper 3.
- [384] K. Saleh and F.O.O. Wagener, *Semi-global analysis of a strong normal-internal resonance*, Univ. Groningen, preprint (2005).
- [385] Vũ Ngọc San, *Quantum monodromy in integrable systems*, Comm. Math. Phys. **203** (1999), 465–479.
- [386] D. Sauzin, *Nekhoroshev estimates and instability for Gevrey class Hamiltonians*, Dynamical Systems, Part I. Hamiltonian Systems and Celestial Mechanics (Pisa, 2002), Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa (2003), 199–217.
- [387] J. Scheurle, *Bifurcation of quasi-periodic solutions from equilibrium points of reversible dynamical systems*, Arch. Ration. Mech. Anal. **97** (1987), 103–139.
- [388] F. Schilder, H.M. Osinga and W. Vogt, *Continuation of quasi-periodic invariant tori*, SIAM J. Appl. Dyn. Syst. **4** (2005), 459–488.
- [389] M.B. Sevryuk, *Reversible Systems*, Lecture Notes in Math., Vol. 1211, Springer, Berlin (1986).
- [390] M.B. Sevryuk, *On invariant tori of reversible systems in a neighbourhood of an equilibrium point*, Russian Math. Surveys **42** (4) (1987), 147–148 (English; Russian original).

- [391] M.B. Sevryuk, *Invariant  $m$ -tori of reversible systems whose phase space is of dimension greater than  $2m$* , Trudy Sem. Imeni I.G. Petrovskogo **14** (1989), 109–124 (in Russian); English translation: J. Soviet Math. **51** (1990), 2374–2386.
- [392] M.B. Sevryuk, *Lower-dimensional tori in reversible systems*, Chaos **1** (1991), 160–167.
- [393] M.B. Sevryuk, *Invariant tori of reversible systems in the presence of additional even coordinates*, Russian Acad. Sci. Dokl. Math. **46** (1993), 286–289 (English; Russian original).
- [394] M.B. Sevryuk, *Invariant tori of reversible systems of intermediate dimensions*, Russian Acad. Sci. Dokl. Math. **47** (1993), 129–133 (English; Russian original); Erratum: Dokl. Akad. Nauk **346** (1996), 576 (in Russian).
- [395] M.B. Sevryuk, *New cases of quasiperiodic motions in reversible systems*, Chaos **3** (1993), 211–214.
- [396] M.B. Sevryuk, *New results in the reversible KAM theory*, Seminar on Dynamical Systems (St. Petersburg, 1991), S.B. Kuksin, V.F. Lazutkin and J. Pöschel, eds, Progress in Nonlinear Differential Equations and their Applications, Vol. 12, Birkhäuser, Basel (1994), 184–199.
- [397] M.B. Sevryuk, *KAM-stable Hamiltonians*, J. Dyn. Control Syst. **1** (1995), 351–366.
- [398] M.B. Sevryuk, *Some problems of KAM theory: conditionally periodic motions in typical systems*, Russian Math. Surveys **50** (2) (1995), 341–353 (English; Russian original).
- [399] M.B. Sevryuk, *The iteration-approximation decoupling in the reversible KAM theory*, Chaos **5** (1995), 552–565.
- [400] M.B. Sevryuk, *Invariant tori of Hamiltonian systems nondegenerate in the sense of Rüssmann*, Dokl. Math. **53** (1996), 69–72 (English; Russian original).
- [401] M.B. Sevryuk, *Excitation of elliptic normal modes of invariant tori in Hamiltonian systems*, Topics in Singularity Theory. V.I. Arnold's 60th Anniversary Collection, A.G. Khovanskiĭ, A.N. Varchenko and V.A. Vasil'ev, eds, Amer. Math. Soc. Transl., Ser. 2 **180**, Adv. Math. Sci. **34** (1997), 209–218.
- [402] M.B. Sevryuk, *Invariant sets of degenerate Hamiltonian systems near equilibria*, Regul. Chaotic Dyn. **3** (3) (1998), 82–92.
- [403] M.B. Sevryuk, *Invariant tori of intermediate dimensions in Hamiltonian systems*, Regul. Chaotic Dyn. **3** (1) (1998), 39–48 (English; Russian original).
- [404] M.B. Sevryuk, *The finite-dimensional reversible KAM theory*, Physica D **112** (1998), 132–147.
- [405] M.B. Sevryuk, *The lack-of-parameters problem in the KAM theory revisited*, Hamiltonian Systems with Three or More Degrees of Freedom (S' Agaró, 1995), C. Simó, ed., NATO ASI Series C: Math. Phys. Sci., Vol. 533, Kluwer Academic, Dordrecht (1999), 568–572.
- [406] M.B. Sevryuk, *On the convergence of coordinate transformations in the KAM procedure*, Regul. Chaotic Dyn. **5** (2000), 181–188.
- [407] M.B. Sevryuk, *Excitation of elliptic normal modes of invariant tori in volume preserving flows*, Global Analysis of Dynamical Systems. Festschrift Dedicated to Floris Takens for his 60th Birthday (Leiden, 2001), H.W. Broer, B. Krauskopf and G. Vegter, eds, Inst. Physics, Bristol (2001), 339–352.
- [408] M.B. Sevryuk, *The classical KAM theory at the dawn of the twenty-first century*, Moscow Math. J. **3** (2003), 1113–1144.
- [409] M.B. Sevryuk, *Partial preservation of frequencies in KAM theory*, Nonlinearity **19** (2006), 1099–1140.
- [410] M.B. Sevryuk, *Invariant tori in quasi-periodic non-autonomous dynamical systems via Herman's method*, Discrete Contin. Dynam. Systems, Ser. A **18** (2007), 569–595.
- [411] M.B. Sevryuk, *Partial preservation of frequencies and Floquet exponents in KAM theory*, Proc. Steklov Inst. Math. **259** (2007), 167–195 (English; Russian original).
- [412] M.B. Sevryuk, *KAM tori: persistence and smoothness*, Nonlinearity **21** (2008), T177–T185.
- [413] C.L. Siegel, *Iteration of analytic functions*, Ann. Math. (2) **43** (1942), 607–612.
- [414] C.L. Siegel and J.K. Moser, *Lectures on Celestial Mechanics*, 3rd edn, Springer, Berlin (1995) (English; partially German original).
- [415] C. Simó, *Perturbations of translations in the two-dimensional torus: the case near resonance*, Proceedings of the VI CEDYA (Congreso de Ecuaciones Diferenciales y Aplicaciones), Univ. Zaragoza, Zaragoza (1984).
- [416] C. Simó, *On the analytical and numerical computation of invariant manifolds*, Modern Methods in Celestial Mechanics (Goutelas, 1989), 2nd edn, D. Benest and C. Froeschlé, eds, Frontières, Gif-sur-Yvette (1992), 285–330.
- [417] C. Simó, *Averaging under fast quasiperiodic forcing*, Hamiltonian Mechanics: Integrability and Chaotic Behavior (Toruń, 1993), J. Seimenis, ed., NATO ASI Series B: Physics, Vol. 331, Plenum Press, New York (1994), 13–34.

- [418] M. Spivak, *A Comprehensive Introduction to Differential Geometry, Vol. I*, 2nd edn, Publish or Perish, Wilmington, DE (1979).
- [419] P. Stefan, *Accessible sets, orbits, and foliations with singularities*, Proc. London Math. Soc., Ser. 3 **29** (1974), 699–713.
- [420] I. Stewart, *Quantizing the classical cat*, Nature **430** (2004), 731–732.
- [421] J.J. Stoker, *Nonlinear Vibrations in Mechanical and Electrical Systems*, 2nd edn, John Wiley & Sons, New York (1992) (First Edition: 1950).
- [422] H.J. Sussmann, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. Soc. **180** (1973), 171–188.
- [423] N.V. Svanidze, *Small perturbations of an integrable dynamical system with an integral invariant*, Proc. Steklov Inst. Math. (2) (1981), 127–151 (English; Russian original).
- [424] F. Takens, *A  $C^1$  counterexample to Moser's twist theorem*, Indag. Math. **33** (1971), 378–386.
- [425] F. Takens, *Forced oscillations and bifurcations*, Applications of Global Analysis, I (Utrecht, 1973), Comm. Math. Inst. Univ. Utrecht **3** (1974), 1–59; Reprinted in: Global Analysis of Dynamical Systems. Festschrift Dedicated to Floris Takens for his 60th Birthday (Leiden, 2001), H.W. Broer, B. Krauskopf and G. Vegter, eds, Inst. Physics, Bristol (2001), 1–61.
- [426] F. Takens, *Singularities of vector fields*, Publ. Math. IHÉS **43** (1974), 47–100.
- [427] F. Takens and F.O.O. Wagener, *Resonances in skew and reducible quasi-periodic Hopf bifurcations*, Nonlinearity **13** (2000), 377–396.
- [428] R. Thom, *Structural Stability and Morphogenesis. An Outline of a General Theory of Models*, 2nd edn, Addison-Wesley, Redwood City, CA (1989) (English; French original).
- [429] D.V. Treshchëv, *The mechanism of destruction of resonant tori in Hamiltonian systems*, Math. USSR–Sb. **68** (1991), 181–203 (English; Russian original).
- [430] D.V. Treshchëv, *Hyperbolic tori and asymptotic surfaces in Hamiltonian systems*, Russian J. Math. Phys. **2** (1994), 93–110.
- [431] D.V. Treshchëv, *Introduction to the Theory of Perturbations of Hamiltonian Systems*, PHASIS, Moscow (1998) (in Russian).
- [432] D.V. Treshchëv, *Evolution of slow variables in a priori unstable Hamiltonian systems*, Nonlinearity **17** (2004), 1803–1841.
- [433] E. Valdinoci, *Families of whiskered tori for a-priori stable/unstable Hamiltonian systems and construction of unstable orbits*, Math. Phys. Electronic J. **6** (2000), paper 2.
- [434] A. Vanderbauwhede, *Centre manifolds, normal forms and elementary bifurcations*, Dynamics Reported **2** (1989), 89–169.
- [435] J.-C. van der Meer, *The Hamiltonian Hopf Bifurcation*, Lecture Notes in Math., Vol. 1160, Springer, Berlin (1985).
- [436] M. van Noort, *Global coherent dynamics of the parametrically forced pendulum: a case study in one-and-a-half degrees of freedom*, Ph.D. Thesis, Univ. Groningen (2001).
- [437] S.J. van Strien, *Centre manifolds are not  $C^\infty$* , Math. Z. **166** (1979), 143–145.
- [438] L. van Veen, *The quasi-periodic doubling cascade in the transition to weak turbulence*, Physica D **210** (2005), 249–261.
- [439] J. Villanueva, *Normal forms around lower dimensional tori of Hamiltonian systems*, Ph.D. Thesis, Univ. Politècnica de Catalunya (1997).
- [440] J. Villanueva, *Kolmogorov Theorem revisited*, J. Differential Equations **244** (2008), 2251–2276.
- [441] H. Waalkens, A. Junge and H.R. Dullin, *Quantum monodromy in the two-centre problem*, J. Phys. A: Math. Gen. **36** (2003), L307–L314.
- [442] F.O.O. Wagener, *On the skew Hopf bifurcation*, Ph.D. Thesis, Univ. Groningen (1998).
- [443] F.O.O. Wagener, *Semi-local analysis of the  $k : 1$  and  $k : 2$  resonances in quasi-periodically forced systems*, Global Analysis of Dynamical Systems. Festschrift Dedicated to Floris Takens for his 60th Birthday (Leiden, 2001), H.W. Broer, B. Krauskopf and G. Vegter, eds, Inst. Physics, Bristol (2001), 113–129.
- [444] F.O.O. Wagener, *A note on Gevrey regular KAM theory and the inverse approximation lemma*, Dyn. Syst. **18** (2) (2003), 159–163.
- [445] F.O.O. Wagener, *On the quasi-periodic  $d$ -fold degenerate bifurcation*, J. Differential Equations **216** (2005), 261–281.
- [446] F.O.O. Wagener, *A parametrized version of Moser's modifying terms theorem*, Univ. Amsterdam, preprint (2005). To appear in Discrete Contin. Dynam. Systems, Ser. S (2011).

- [447] Sh. Wang and Ch.-Q. Cheng, *Birkhoff lower-dimensional tori in Hamiltonian systems*, Chinese Sci. Bull. **42** (1997), 1866–1870.
- [448] Sh. Wang and Ch.-Q. Cheng, *Lower-dimensional tori for generic Hamiltonian systems*, Chinese Sci. Bull. **44** (1999), 1187–1191.
- [449] Sh. Wang, D. Zhang and J. Xu, *On the persistence of elliptic lower-dimensional tori in Hamiltonian systems under the first Melnikov condition and Rüssmann's non-degeneracy condition*, Nonlinear Anal. **66** (2007), 1675–1685.
- [450] C.E. Wayne, *An introduction to KAM Theory*, Dynamical Systems and Probabilistic Methods in Partial Differential Equations (Berkeley, 1994), P. Deift, C.D. Levermore and C.E. Wayne, eds, Lectures in Appl. Math., Vol. 31, Amer. Math. Soc., Providence, RI (1996), 3–29.
- [451] B. Wei, *Perturbations of lower dimensional tori in the resonant zone for reversible systems*, J. Math. Anal. Appl. **253** (2001), 558–577.
- [452] H. Whitney, *Differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934), 369–387.
- [453] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934), 63–89.
- [454] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 2nd edn, Springer, New York (2003).
- [455] J. Xu, *Persistence of elliptic lower dimensional invariant tori for small perturbation of degenerate integrable Hamiltonian systems*, J. Math. Anal. Appl. **208** (1997), 372–387.
- [456] J. Xu, *Persistence of lower dimensional invariant tori for nearly integrable Hamiltonian systems*, Nonlinear Anal. **52** (2003), 383–403.
- [457] J. Xu, *Normal form of reversible systems and persistence of lower dimensional tori under weaker nonresonance conditions*, SIAM J. Math. Anal. **36** (2004), 233–255.
- [458] J. Xu and J. You, *Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition*, J. Math. Pures Appl., Sér. 9 **80** (2001), 1045–1067.
- [459] J. Xu and J. You, *A symplectic map and its application to the persistence of lower dimensional invariant tori*, Sci. China A **45** (2002), 598–603.
- [460] J. Xu and J. You, *Gevrey-smoothness of invariant tori for analytic nearly integrable Hamiltonian systems under Rüssmann's non-degeneracy condition*, J. Differential Equations **235** (2007), 609–622.
- [461] J. Xu, J. You and Q. Qiu, *Invariant tori for nearly integrable Hamiltonian systems with degeneracy*, Math. Z. **226** (1997), 375–387.
- [462] J.-C. Yoccoz,  *$C^1$ -conjugaison des difféomorphismes du cercle*, Geometric Dynamics (Rio de Janeiro, 1981), J. Palis, ed., Lecture Notes in Math., Vol. 1007, Springer, Berlin (1983), 814–827.
- [463] J.-C. Yoccoz, *Travaux de Herman sur les tores invariants*, Séminaire Bourbaki, Vol. 1991/92, Exp. 754, Astérisque **206** (1992), 311–344.
- [464] J.-C. Yoccoz, *Recent developments in dynamics*, Proceedings of the International Congress of Mathematicians, Vol. 1, (Zürich, 1994), S.D. Chatterji, ed., Birkhäuser, Basel (1995), 246–265; Reprinted in: Fields Medallists' Lectures, M.F. Atiyah and D. Iagolnitzer, eds, World Scientific, River Edge, NJ; Singapore Univ. Press, Singapore (1997), 588–609.
- [465] J.-C. Yoccoz, *Théorème de Siegel, nombres de Bruno et polynômes quadratiques*, Astérisque **231** (1995), 3–88.
- [466] J.-C. Yoccoz, *Analytic linearization of circle diffeomorphisms*, Dynamical Systems and Small Divisors (Cetraro, 1998), S. Marmi and J.-C. Yoccoz, eds, Lecture Notes in Math., Vol. 1784, Springer, Berlin; Centro Internazionale Matematico Estivo, Florence (2002), 125–173.
- [467] J. You, *A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems*, Comm. Math. Phys. **192** (1998), 145–168.
- [468] J. You, *Perturbations of lower dimensional tori for Hamiltonian systems*, J. Differential Equations **152** (1999), 1–29.
- [469] G.M. Zaslavskii and B.V. Chirikov, *Stochastic instability of nonlinear oscillations*, Soviet Phys. Uspekhi **14** (1972), 549–568 (English; Russian original).
- [470] E. Zehnder, *An implicit function theorem for small divisor problems*, Bull. Amer. Math. Soc. **80** (1974), 174–179.
- [471] E. Zehnder, *Generalized implicit function theorems with applications to some small divisor problems, I and II*, Comm. Pure Appl. Math. **28** (1975), 91–140; **29** (1976), 49–111.

- [472] E. Zehnder, *Remarks on periodic solutions on hypersurfaces*, Periodic Solutions of Hamiltonian Systems and Related Topics (Il Ciocco, 1986), P.H. Rabinowitz, A. Ambrosetti, I. Ekeland and E.J. Zehnder, eds, NATO ASI Series C: Math. Phys. Sci., Vol. 209, Reidel, Dordrecht (1987), 267–279.
- [473] D. Zhang and J. Xu, *Gevrey-smoothness of elliptic lower-dimensional invariant tori in Hamiltonian systems under Rüssmann's non-degeneracy condition*, J. Math. Anal. Appl. **323** (2006), 293–312.
- [474] D. Zhang and J. Xu, *On elliptic lower dimensional tori for Gevrey-smooth Hamiltonian systems under Russmann's non-degeneracy condition*, Discrete Contin. Dynam. Systems, Ser. A **16** (2006), 635–655.
- [475] D. Zhang and J. Xu, *Invariant tori for Gevrey-smooth Hamiltonian systems under Rüssmann's non-degeneracy condition*, Nonlinear Anal. **67** (2007), 2240–2257.
- [476] Nguyen Tien Zung, *Symplectic topology of integrable Hamiltonian systems. II. Topological classification*, Compositio Math. **138** (2003), 125–156.
- [477] Nguyen Tien Zung, *Symplectic topology of integrable Hamiltonian systems. I. Arnold–Liouville with singularities*, Compositio Math. **101** (1996), 179–215.
- [478] Nguyen Tien Zung, *A note on focus-focus singularities*, Diff. Geom. Appl. **7** (1997), 123–130.

## CHAPTER 7

# Reconstruction Theory and Nonlinear Time Series Analysis

Floris Takens

*Johann Bernoulli Institute for Mathematics and Computer Science,  
P.O. Box 407 9700 AK Groningen, The Netherlands  
Bernoulliborg, Building 5161, Nijenborgh 9, 9747 AG Groningen, The Netherlands*

### Contents

|                                                                                                                                         |     |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1. Introduction . . . . .                                                                                                               | 347 |
| 2. An experimental example: the dripping tap . . . . .                                                                                  | 347 |
| 3. The reconstruction theorem . . . . .                                                                                                 | 348 |
| 4. The reconstruction theorem and nonlinear time series analysis: discrimination between deterministic and random time series . . . . . | 352 |
| 4.1. Box counting dimension and its numerical estimation . . . . .                                                                      | 354 |
| 5. Stationarity and reconstruction measures . . . . .                                                                                   | 357 |
| 5.1. Measures defined by relative frequencies . . . . .                                                                                 | 358 |
| 5.2. Definition of stationarity and reconstruction measures . . . . .                                                                   | 358 |
| 5.3. Examples . . . . .                                                                                                                 | 359 |
| 6. Correlation dimensions and entropies . . . . .                                                                                       | 360 |
| 7. Numerical estimation of correlation integrals and the corresponding dimensions and entropies . . . . .                               | 363 |
| 8. Classical time series analysis, the analysis in terms of correlation integrals, and predictability . . . . .                         | 366 |
| 8.1. Classical time series analysis . . . . .                                                                                           | 366 |
| 8.2. Determinism and auto covariances . . . . .                                                                                         | 368 |
| 8.3. Predictability and correlation integrals . . . . .                                                                                 | 370 |
| 9. Miscellaneous subjects . . . . .                                                                                                     | 372 |
| 9.1. Liapunov exponents . . . . .                                                                                                       | 373 |
| 9.2. The Kantz–Diks test – discriminating between time series and testing for reversibility . . . . .                                   | 375 |
| References . . . . .                                                                                                                    | 376 |



## 1. Introduction

Since it was realized that even simple and deterministic dynamical systems can produce trajectories which look like the result of a random process, there were some obvious questions: How can one distinguish ‘deterministic chaos’ from ‘randomness’ and, in the case of determinism, how can we extract from a time series relevant information about the dynamical system generating it? It turned out that these questions could not be answered in terms of the conventional linear time series analysis, i.e. in terms of auto covariances and power spectra. In this chapter we review the new methods of time series analysis which were introduced in order to answer these questions. We shall restrict ourselves to the methods which are based on embedding theory and correlation integrals: these methods, and the relations of these methods with the linear methods, are now well understood and supported by mathematical theory. Though they are applicable to a large class of time series, it is often necessary to supplement them for special applications by ad hoc refinements. We shall not deal in this chapter with such ad hoc refinements, but concentrate on the main concepts. In this way we hope to make the general theory as clear as possible. For the special methods which are required for the various applications, we refer to the monographs [27,14,20] as well as to the collections of research papers [53,31]. This chapter is an expanded version of Chapter 6 in [7].

We start in the next section with an experimental example, a dripping tap, taken from [10] which gives a good heuristic idea of the method of reconstruction. Then, in Section 3, we describe the reconstruction theorem which gives, amongst others, a justification for this analysis of the dripping tap. It also motivates a (still rather primitive) method for deriving numerical invariants which can be used to characterize the dynamics of the system generating a time series; this is discussed in Section 4. After that, we introduce, in Section 5, some notions related to stationarity, in particular the so-called reconstruction measures, which describe the statistics of the behaviour of a time series as far as subsegments (of small lengths) are concerned. After that, we discuss the correlation integrals and the dimensions and entropies which are based on them in Section 6 and their estimation in Section 7. In Section 8 we relate the methods in terms of correlation integrals with linear methods. Finally, in Section 9, we give a short discussion of related topics which, however, do not strictly fall within the framework of the analysis of time series in terms of correlation integrals.

## 2. An experimental example: the dripping tap

Here we give an account of a heuristic way in which the idea of reconstruction was applied to a simple physical system which was first published in Scientific American [10] and which was based on earlier work in [32]. We also use this example to introduce the notion of a (deterministic) dynamical system as a generator of a time series.

We consider a sequence of successive measurements from a simple physical experiment: One measures the lengths of the time intervals between successive drops falling from a dripping tap. They form a finite, but long, sequence of real numbers, a *time series*  $\{y_0, y_1, \dots\}$ . The question is whether this sequence of real numbers contains an indication about whether it is generated by a deterministic system or not (the tap was set in such a



way that the dripping was not ‘regular’, i.e. not periodic; for many taps such a setting is hard to obtain). In order to answer this question, the authors of [10] considered the set of 3-vectors  $\{(y_{i-1}, y_i, y_{i+1})\}_i$ , each obtained by taking three successive values of the time series. These vectors form a ‘cloud of points’ in  $\mathbb{R}^3$ . It turned out that, in the experiment reported in [10], this cloud of points formed a (slightly thickened) curve  $C$  in 3-space. This gives an indication that the process is deterministic (with small noise corresponding to the thickening of the curve). The argument is based on the fact that the values  $y_{n-1}$  and  $y_n$  of the time series, together with the (thickened) curve  $C$ , enable, in general, a good prediction for the next value  $y_{n+1}$  of the time series. This prediction of  $y_{n+1}$  is obtained in the following way: we consider the line  $l_n$  in 3-space which is in the direction of the third coordinate axis and whose first two coordinates are  $y_{n-1}$  and  $y_n$ . This line  $l_n$  has to intersect the cloud of 3-vectors, since  $(y_{n-1}, y_n, y_{n+1})$  belongs to this cloud (even though the cloud is composed of vectors which were collected in the past, we assume the past to be so extensive that new vectors do not substantially *extend* the cloud). If we ‘idealize’ this cloud to a curve  $C$ , then *generically*  $l_n$  and  $C$  should only have this one point in common. We shall make all this later mathematically rigorous, but the idea is that two (smooth) curves in 3-space have *in general*, or *generically*, no intersection (in dimension 3 there is enough space for two curves to become disjoint by a small perturbation); by construction,  $C$  and  $l_n$  must have one point in common, namely  $(y_{n-1}, y_n, y_{n+1})$ , but there is no reason for other intersections. This means that we expect that we can predict the value  $y_{n+1}$  from the previous values  $y_{n-1}$  and  $y_n$  as the third coordinate of the intersection  $C \cap l_n$ . Even if we do not idealize the cloud of points in  $\mathbb{R}^3$  to the curve, we conclude that the next value  $y_{n+1}$  should be in the range given by the intersection of this cloud with  $l_n$ ; usually this range is small and hence the expected error of the prediction will be small. The fact that such predictions are possible means that we deal with a deterministic system or an *almost* deterministic system if we take into account the fact that the cloud of vectors is really a *thickened* curve.

So, in this example, we can consider the cloud of vectors in 3-space, or its idealization to the curve  $C$ , as the *statistical information* concerning the time series which is extracted from a *long* sequence of observations. It describes the dynamical law generating the time series in the sense that it enables us to predict, from a *short* sequence of recent observations, the next value.

A mathematical justification of the above method to detect a deterministic structure in an apparently random time series, is given by the reconstruction theorem; see Section 3. This was the main starting point of what is now called nonlinear time series analysis.

The fact that many systems are not strictly deterministic but can be interpreted as deterministic with a small contamination of noise gives extra complications to which we will return later.

### 3. The reconstruction theorem

In this section we formulate the reconstruction theorem and discuss some of its generalizations. In the next section we give a first example of how it is relevant for the questions raised in the introduction.

We consider a dynamical system, given by a diffeomorphism  $\varphi : M \rightarrow M$  on a compact manifold  $M$ , together with a smooth ‘read out’ function  $f : M \rightarrow \mathbb{R}$ . This set up is proposed as the general form of mathematical models for deterministic dynamical systems generating time series in the following sense. Possible *evolutions*, or *orbits*, of the dynamical system are sequences of the form  $\{x_n = \varphi^n(x_0)\}$ , where usually the index  $n$  runs through the non-negative integers. We assume that what is observed, or measured, at each time  $n$  is not the whole state  $x_n$  but just one real number  $y_n = f(x_n)$ , where  $f$  is the read out function which assigns to each state the value which is measured when the system is in that state. So corresponding to each evolution  $\{x_n = \varphi^n(x_0)\}$  there is a time series  $Y$  of successive measurements  $\{y_n = f(\varphi^n(x_0)) = f(x_n)\}$ .

In the above setting of a deterministic system with measurements, we made some restrictions which are not completely natural: the time could have been continuous, i.e. the dynamics could have been given by a differential equation (or vector field) instead of a diffeomorphism (in other words, the time set could have been the reals  $\mathbb{R}$  instead of the natural numbers  $\mathbb{N}$ ); instead of measuring only one value, one could measure several values (or a finite dimensional vector); also the dynamics could have been given by an endomorphism instead of a diffeomorphism; finally we could allow  $M$  to be non-compact. We shall return to such generalizations after treating the situation as proposed, namely with the dynamics given by a diffeomorphism on a compact manifold and a 1-dimensional read out function. In this situation we have:

**THEOREM.** *For a compact  $m$ -dimensional manifold  $M$ ,  $\ell \geq 1$ , and  $k > 2m$  there is an open and dense subset  $U$  in  $\text{Diff}^\ell(M) \times C^\ell(M)$ , the product of the space of  $C^\ell$ -diffeomorphisms on  $M$  and the space of  $C^\ell$ -functions on  $M$ , such that for  $(\varphi, f) \in U$  the following map is an embedding of  $M$  into  $\mathbb{R}^k$ :*

$$M \ni x \mapsto (f(x), f(\varphi(x)), \dots, f(\varphi^{k-1}(x))) \in \mathbf{R}^k.$$

*So the conclusion of this theorem holds for generic pairs  $(\varphi, f)$ .* □

The proof of this theorem first appeared in [41], see also [1], and is an adaptation of Whitney’s embedding theorem [54]. For a later and more didactical version, with some extensions, see [40]. For the proof we refer to these references.

We introduce the following notation: the map from  $M$  to  $\mathbb{R}^k$ , given by  $x \mapsto (f(x), f(\varphi(x)), \dots, f(\varphi^{k-1}(x)))$ , is denoted by  $\text{Rec}_k$ ; vectors of the form  $(f(x), f(\varphi(x)), \dots, f(\varphi^{k-1}(x)))$  are called  $k$ -dimensional *reconstruction vectors* of the system defined by  $(\varphi, f)$ . The image of  $M$  under  $\text{Rec}_k$  is denoted by  $\mathcal{X}_k$ .

The meaning of the reconstruction theorem is the following. For a time series  $Y = \{y_n\}$  we consider the sequence of its  $k$ -dimensional *reconstruction vectors*  $\{(y_n, y_{n+1}, \dots, y_{n+k-1}) \in \mathbf{R}^k\}_n$ ; this sequence is ‘diffeomorphic’ to the evolution of the deterministic system producing the time series  $Y$  provided that the following conditions are satisfied:

- $k$  is sufficiently large, e.g. larger than twice the dimension of the state space  $M$ ;
- the pair  $(\varphi, f)$ , consisting of the deterministic system and the read out function, is generic in the sense of the above theorem.

The sense in which the sequence of reconstruction vectors of the time series is diffeomorphic to the underlying evolution is as follows. The relation between the time series  $Y = \{y_n\}$  and the underlying evolution  $\{x_n = \varphi^n(x_0)\}$  is that  $y_n = f(x_n)$ . This means that the reconstruction vectors of the time series  $Y$  are just the images under the map  $Rec_k$  of the successive points of the evolution  $\{x_n\}$ . So  $Rec_k$ , which is, by the theorem, an embedding of  $M$  onto its image  $\mathcal{X}_k \subset \mathbb{R}^k$ , sends the orbit of  $\{x_n\}$  to the sequence of reconstruction vectors  $\{(y_n, y_{n+1}, \dots, y_{n+k-1})\}$  of the time series  $Y$ . We note that a consequence of this fact, and the compactness of  $M$ , is that for any metric  $d$  on  $M$  (derived from some Riemannian metric), the evolution  $\{x_n\}$  and the sequence of reconstruction vectors of  $Y$  are metrically equal up to a *bounded distortion*. This means that the quotients

$$\frac{d(x_n, x_m)}{\|(y_n, \dots, y_{n+k-1}) - (y_m, \dots, y_{m+k-1})\|}$$

are uniformly bounded and bounded away from zero. So, all the recurrence properties of the evolution  $\{x_n\}$  are still present in the corresponding sequence of reconstruction vectors of  $Y$ .

## Generalizations

*Continuous time.* Suppose now that we have a dynamical system with continuous time, i.e. a dynamical system given by a vector field (or a differential equation)  $X$  on a compact manifold  $M$ . For each  $x \in M$  we denote the corresponding orbit by  $t \mapsto \varphi^t(x)$  for  $t \in \mathbb{R}$ . In that case there are two alternatives for the definition of the reconstruction map  $Rec_k$ . One is based on a discretization: we take a (usually small) time interval  $h > 0$  and define the reconstruction map  $Rec_k$  in terms of the diffeomorphism  $\varphi^h$ , the time  $h$  map of the vector field  $X$ . Also then one can show that for  $k > 2m$  and generic triples  $(X, f, h)$ , the reconstruction map  $Rec_k$  is an embedding. Another possibility is to define the reconstruction map in terms of the derivatives of the function  $t \mapsto f(\varphi^t(x))$  at  $t = 0$ , i.e. by taking

$$Rec_k(x) = \left( f(\varphi^t(x)), \frac{\partial}{\partial t} f(\varphi^t(x)), \dots, \frac{\partial^{k-1}}{\partial t^{k-1}} f(\varphi^t(x)) \right) \Big|_{t=0}.$$

Also for this definition of the reconstruction map the conclusion is the same: for  $k > 2m$  and generic pairs  $(X, f)$ , the reconstruction map is an embedding of  $M$  into  $\mathbb{R}^k$ . The proof for both these versions is in [41]; in fact in [1] it was just the case of continuous time that was treated through the discretization mentioned above.

We note that the second form of reconstruction is not very useful for experimental time series since the evaluation of the derivatives introduces, in general, a considerable amount of noise.

*Multidimensional measurements.* In the case that at each time one measures a multidimensional vector, so that we have a read out function with values in  $\mathbb{R}^s$ , the reconstruction map  $Rec_k$  has values in  $\mathbb{R}^{sk}$ . In this case, the conclusion of the

reconstruction theorem still holds for generic pairs  $(\varphi, f)$  whenever  $ks > 2m$ . As far as I know there is no reference for this result, but the proof is the same as in the case with 1-dimensional measurements.

*Endomorphisms.* If we allow the dynamics to be given by an endomorphism instead of a diffeomorphism, then the obvious generalization of the reconstruction theorem is *not* correct. A proper generalization in this case, which can serve as a justification for the type of analysis which we described in Section 2, and also of the numerical estimation of dimensions and entropies in Section 7, was given in [44]. For  $k > 2m$  one proves that there is, under the usual generic assumptions, a map  $\pi_k : \mathcal{X}_k \rightarrow M$  such that  $\pi_k \circ \text{Rec}_k = \varphi^{k-1}$ . This means that a sequence of  $k$  successive measurements determines the state of the system at the *end* of the sequence of measurements. However, there may be different reconstruction vectors determining the same final state; in this case  $\mathcal{X}_k \subset \mathbb{R}^k$  is, in general, not a sub-manifold, but still the map  $\pi_k$  is differentiable, in the sense that it can be extended to a differentiable map from a neighbourhood of  $\mathcal{X}_k$  in  $\mathbb{R}^k$  to  $M$ .

*Compactness.* If  $M$  is not compact the reconstruction theorem remains true, but not the remark about ‘bounded distortion’. The reason for this restriction was however the following. When talking about predictability, as in Section 2, based on the knowledge of a (long) segment of a time series, an implicit assumption is always: what will happen next has already happened (exactly or approximately) in the past. So this idea of predictability is only applicable to evolutions  $\{x_n\}_{n \geq 0}$  which have the property that for each  $\varepsilon > 0$  there is an  $N(\varepsilon)$  such that for each  $n > N(\varepsilon)$ , there is some  $0 < n'(n) < N(\varepsilon)$  with  $d(x_n, x_{n'(n)}) < \varepsilon$ . This is a mathematical formulation of the condition that after a sufficiently long segment of the evolution (here length  $N(\varepsilon)$ ) every new state, like  $x_n$ , is approximately equal (here equal up to a distance  $\varepsilon$ ) to one of the past states, here  $x_{n'(n)}$ . It is not hard to see that this condition on  $\{x_n\}$ , assuming that the state space is a complete metric space, is equivalent to the condition that the positive evolution  $\{x_n\}_{n \geq 0}$  has a compact closure. Since such an assumption is basic for the main applications of the reconstruction theorem, it is no great restriction of the generality to assume, as we did, that the state space  $M$  itself is a compact manifold, since we only want to deal with a compact part of the state space anyway. In this way we avoid also the complications of defining the topology on spaces of functions (and diffeomorphisms) on non-compact manifolds.

There is another generalization of the reconstruction theorem which is related to the above remark. For any evolution  $\{x_n\}_{n \geq 0}$  of a (differentiable) dynamical system with compact state space, one defines its  $\omega$ -limit  $\omega(x_0)$  as:

$$\omega(x_0) = \{x | \exists n_i \rightarrow \infty, \text{ such that } x_{n_i} \rightarrow x\}.$$

(Often such an  $\omega$ -limit is an attractor, but that is not of immediate concern to us here.) Using the compactness of  $M$ , or of the closure of the evolution  $\{x_n\}_{n \geq 0}$ , one can prove that for each  $\varepsilon > 0$  there is some  $N'(\varepsilon)$  such that for each  $n > N'(\varepsilon)$ ,  $d(x_n, \omega(x_0)) < \varepsilon$ . So the  $\omega$ -limit describes the dynamics of the evolution starting in  $x_0$  without the peculiarities (transients) which are only due to the specific choice of the initial state. For a number of applications, one does not need the reconstruction map  $\text{Rec}_k$  to be an embedding of

the whole of  $M$ , but only of the  $\omega$ -limit of the evolution under consideration. Often the dimension of such an  $\omega$ -limit is much lower than the dimension of the state space. For this reason it was important that the reconstruction theorem was extended in [40] to this case, where the condition on  $k$  could be weakened:  $k$  only has to be bigger than twice the dimension of the  $\omega$ -limit. It should however be observed that these  $\omega$ -limit sets are in general not very nice spaces (like manifolds) and that for that reason the notion of dimension is not so obvious (one has to take, in this case, the box counting dimension, see Section 4.1); in the conclusion of the theorem, one gets an injective map of the  $\omega$ -limit into  $\mathbb{R}^k$ , but the property of bounded distortion has not been proven for this case (and is probably false).

*Historical note.* This reconstruction theorem was obtained independently by Aeyels and myself. In fact it was D.L. Elliot who pointed out to me, at the Warwick conference in 1980 where I presented the reconstruction theorem, that Aeyels had obtained the same type of results in his thesis (Washington University, 1978) which was published as [1]. His motivation came from systems theory, and in particular from the observability problem for generic nonlinear systems.

#### 4. The reconstruction theorem and nonlinear time series analysis: discrimination between deterministic and random time series

In this section we show how one can obtain, from the distribution of reconstruction vectors of a given time series, an indication as to whether the time series was generated by a deterministic or a random process. In some cases this is very simple, see Section 2. As an example we consider two time series: one is obtained from the Hénon system and the second is a randomized version of the first one. We recall that the Hénon system has a 2-dimensional state space  $\mathbb{R}^2$  and is given by the map  $(x_1, x_2) \mapsto (1 - ax_1^2 + bx_2, x_1)$ ; we take for  $a$  and  $b$  the usual values 1.4 and 0.3. As a read out function we take  $f(x_1, x_2) = x_1$ . For an evolution  $(x_1(n), x_2(n))$ , with  $(x_1(0), x_2(0))$  close to  $(0, 0)$ , we consider the time series  $y_n = f(x_1(n), x_2(n)) = x_1(n)$  with  $0 \leq n \leq N$  for some large  $N$ . In order to get a time series of an evolution *inside* the Hénon attractor (within the limits of the resolution of the graphical representation) we omitted the first 100 values. This is our first time series. Our second time series is obtained by applying a random permutation to the first time series. So the second time series is essentially a random iid (identically and independently distributed) time series with the same histogram as the first time series. From the time series itself it is not obvious which of the two is deterministic, see Figure 1. However, if we plot the ‘cloud of 2-dimensional reconstruction vectors’, the situation is completely different, see Figure 2. In the case of the reconstruction vectors from the Hénon system we clearly see the well known picture of the Hénon attractor. In the case of the reconstruction vectors of the randomized time series we just see a structureless cloud.

Of course the above example is rather special, but, in terms of the reconstruction theorem, it is clear that there is a more general method behind this observation. In the case where we use a time series that is generated by a deterministic process, the reconstruction

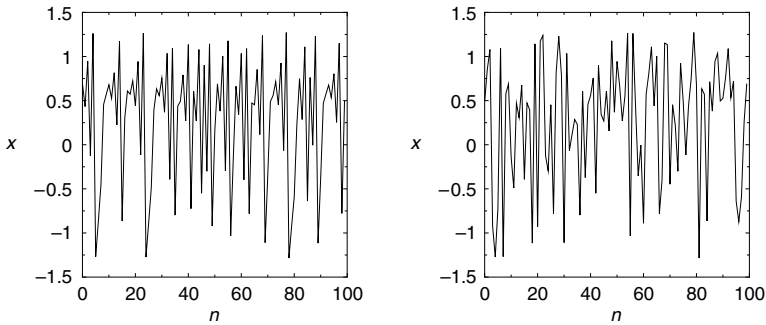


Fig. 1. Time series: (left) from the Hénon system with  $n$  running from 100 to 200 and (right) a randomized version of this time series.

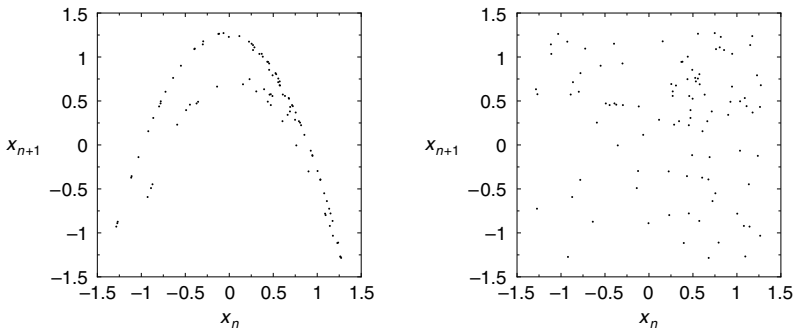


Fig. 2. The clouds of 2-dimensional reconstruction vectors for the two time series in Figure 1.

vectors fill a ‘dense’<sup>1</sup> subset of the differentiable image (under the reconstruction map) of the closure of the corresponding evolution of the deterministic system (or of its  $\omega$ -limit if we omit a sufficiently long segment from the beginning of the time series). If the reconstruction vectors are of sufficiently high dimension, one generically gets even a diffeomorphic image. In the case of the Hénon system the reconstruction vectors form a diffeomorphic image of the Hénon attractor which has, according to numerical evidence, a dimension strictly between 1 and 2 (for the notion of dimension see below). This also means that the reconstruction vectors densely fill a subset of dimension strictly smaller than 2; this explains that the 2-dimensional reconstruction vectors are concentrated on a ‘thin’ subset of the plane (this will also be explained below). Such a concentration on a ‘thin’ subset is clearly different from the diffuse distribution of reconstruction vectors which one sees in the case of the randomized time series. A diffuse distribution, filling out some open region, is typical for reconstruction vectors generated by random processes,

<sup>1</sup>It would only be really dense if we would consider an infinite set of reconstruction vectors.

e.g. this is what one gets for time series generated by any (non-deterministic) Gaussian process.

In the above discussion the, still not defined, notion of dimension played a key role. The notion of dimension, which we have to use here, is not the usual one from topology or linear algebra which can only have integer values: we will use here so called fractal dimensions. In this section we will limit ourselves to the box counting dimension. We first discuss its definition, or rather a number of equivalent definitions, and some of its basic properties. It then turns out that the numerical estimation of these dimensions provides a general method for discrimination between time series as discussed here, namely those generated by a low dimensional attractor of a deterministic system, and those generated by a random process.

#### 4.1. Box counting dimension and its numerical estimation

For introductory information on the various dimensions which are used in dynamical systems we refer the reader to [33], and to the references given there. For a more advanced treatment see [18,19,34].

The box counting dimension, which we discuss here, is certainly not the most convenient for numerical estimation, but it is conceptually the simpler one. The computationally more relevant dimensions are discussed in Section 6.

Let  $K$  be some compact metric space. We say that a subset  $A \subset K$  is  $\varepsilon$ -spanning if each point in  $K$  is contained in the  $\varepsilon$ -neighbourhood of one of the points of  $A$ . The smallest cardinality of an  $\varepsilon$ -spanning subset of  $K$  is denoted by  $a_\varepsilon(K)$ ; note that this number is finite due to the compactness of  $K$ .

For the same compact metric space a subset  $B \subset K$  is called  $\varepsilon$ -separated if, for any two points  $b_1 \neq b_2 \in B$ , the distance between  $b_1$  and  $b_2$  is at least  $\varepsilon$ . The greatest cardinality of an  $\varepsilon$ -separated subset of  $K$  is denoted by  $b_\varepsilon(K)$ .

It is not hard to prove that:

$$a_{\frac{\varepsilon}{2}}(K) \geq b_\varepsilon(K) \geq a_\varepsilon(K).$$

Also it is clear that, as  $\varepsilon$  tends to 0,  $a_\varepsilon$  and  $b_\varepsilon$  tend to infinity, except if  $K$  consists of a finite number of points only. The speed of this growth turns out to provide a definition of dimension.

DEFINITION. The *box counting dimension* of a compact metric space  $K$  is defined as

$$d(K) = \limsup_{\varepsilon \rightarrow 0} -\frac{\ln(a_\varepsilon(K))}{\ln(\varepsilon)} = \limsup_{\varepsilon \rightarrow 0} -\frac{\ln(b_\varepsilon(K))}{\ln(\varepsilon)}.$$

(The equality of the two expressions follows from the above inequalities.) □

It is not hard to show that if  $K$  is the unit cube, or any compact subset with non-empty interior, in the Euclidean  $n$ -dimensional space, then its box counting dimension is  $d(K) = n$ ; also, for any compact subset  $K \subset \mathbf{R}^n$ , we have  $d(K) \leq n$ . This is what we should expect from any quantity which is some sort of a dimension. There are however

also examples where this dimension is not an integer. For example the mid-third Cantor set has box counting dimension  $\ln(2)/\ln(3)$ .

Next we consider a compact subset  $K \subset \mathbb{R}^k$ . With the Euclidean distance function this is also a metric space. For each  $\varepsilon > 0$  we consider the partition of  $\mathbb{R}^k$  in  $\varepsilon$ -boxes of the form

$$\{n_1\varepsilon \leq x_1 < (n_1 + 1)\varepsilon, \dots, n_k\varepsilon \leq x_k < (n_k + 1)\varepsilon\},$$

with  $(n_1, \dots, n_k) \in \mathbb{Z}^k$ . The number of boxes of this partition, which contain at least one point of  $K$ , is denoted by  $c_\varepsilon(K)$ . We then have the following inequalities:

$$a_{\varepsilon\sqrt{k}}(K) \leq c_\varepsilon(K) \leq 3^k a_\varepsilon(K).$$

This means that we may replace in the definition of box counting dimension the quantities  $a_n$  or  $b_n$  by  $c_n$  whenever  $K$  is a subset of a Euclidean space. This explains the name ‘box counting dimension’; also other names are used for this dimension; for a discussion, with some history, see [19].

It follows easily from the definitions that if  $K'$  is the image of  $K$  under a map  $f$  with bounded expansion, or a Lipschitz map, i.e. a map such that for some constant  $C$  we have for all pairs  $k_1, k_2 \in K$  that

$$\rho'(f(k_1), f(k_2)) \leq C\rho(k_1, k_2),$$

where  $\rho, \rho'$  denote the metrics in  $K$  and  $K'$  respectively, then  $d(K') \leq d(K)$ . A differentiable map, restricted to a compact set, has bounded expansion. So:

**LEMMA.** *Under a differentiable map the box counting dimension of a compact set does not increase, and under a diffeomorphism it remains the same.*  $\square$

## Remarks

*Differentiability and the Peano curve.* One should not expect that this lemma remains true without the assumption of differentiability. This can be seen from the well known Peano curve which is a continuous map from the (1-dimensional) unit interval onto the (2-dimensional) square.

*Numerical estimation of the box counting dimension.* Suppose  $K \subset \mathbb{R}^k$  is a compact subset and  $k_1, k_2, \dots \in K$  is a dense sequence in  $K$ . If such a dense sequence is numerically given, in the sense that we have (good approximations of)  $k_1, \dots, k_N$  for some large value  $N$ , this can be used to estimate the box counting dimension of  $K$  in the following way: first we estimate the quantities  $c_\varepsilon(K)$  for various values of  $\varepsilon$  by counting the number of  $\varepsilon$ -boxes containing at least one of the points  $k_1, \dots, k_N$ . Then we use these quantities to estimate the limit (or limsup) of  $-\ln(c_\varepsilon(K))/\ln(\varepsilon)$  (which is the box counting dimension).

We have to use the notion of ‘estimating’ here in a loose sense: there is certainly no error bound. We can however say somewhat more: first, we should not try to make  $\varepsilon$  smaller than



the precision with which we know the successive points  $k_i$ ; then there is another lower bound for  $\varepsilon$  which is much harder to assess, and which depends on  $N$ : if we take  $\varepsilon$  very small, then  $c_\varepsilon(K)$  will usually become bigger than  $N$  which implies that we will get too low an estimate for  $c_\varepsilon$  (since our estimate will at most be  $N$ ). These limitations on  $\varepsilon$  imply that the estimation of the limit (or limsup) of  $-\ln(c_\varepsilon(K))/\ln(\varepsilon)$  cannot be more than rather ‘speculative’. The ‘speculation’ which one usually assumes is that  $\ln(c_\varepsilon(K))$ , as a function of  $-\ln(\varepsilon)$  approaches a straight line for decreasing  $\varepsilon$ , i.e. for increasing  $-\ln(\varepsilon)$ . Whenever this speculation is supported by numerical evidence, one takes the slope of the estimated straight line as the estimate for the box counting dimension. It turns out that this procedure often leads to reproducible results in situations which one encounters in the analysis of time series of (chaotic) dynamical systems (at least if the dimension of the state space is sufficiently low).

Later, we shall see other definitions of dimension which lead to estimation procedures which are better than the present estimate of the box counting dimension, in the sense that we have a somewhat better idea of the variance of the estimates. Still, dimension estimations have been carried out along the line as described above, see [23]. In that paper, dimension estimates were obtained for a 1-parameter family of dimensions, the so called  $D_q$  dimensions introduced by Renyi, see [36]. The box counting dimension as defined here coincides with  $D_0$ . The estimated value of this dimension for the Hénon attractor is  $D_0 = 1.272 \pm .006$ ; the error bound is of course not rigorous: there is even no proof of the fact that  $D_0$  is greater than 0 in this case!

*Box counting dimension as indication for ‘thin’ subsets.* As we remarked above, a compact subset of the plane which has a box counting dimension smaller than 2 cannot have interior points. So the ‘thin appearance’ of the reconstructed Hénon attractor is in agreement with the fact that the box counting dimension is estimated to be smaller than 2.

From the above arguments it is clear that, if we apply the dimension estimation to a sequence of reconstruction vectors from a time series which is generated by a deterministic system with a low dimensional attractor, then we should find a dimension which does not exceed the dimension of the attractor. So an estimated dimension, based on a sequence of reconstruction vectors  $\{(y_i, \dots, y_{i+k-1}) \in \mathbb{R}^k\}$ , which is significantly lower than the dimension  $k$  in which the reconstruction takes place, is an indication in favour of a deterministic system generating the time series.

Here we have to come back to the remark made at the end of Section 2, also in relation to the above remark warning against too small values of  $\varepsilon$ . Suppose we consider a sequence of reconstruction vectors generated by a system which we expect to be essentially deterministic, but contaminated with some small noise (as in the case of the dripping tap in Section 2). If we consider in these cases values of  $\varepsilon$  which are smaller than the amplitude of the noise, we should find that the ‘cloud’ of  $k$ -dimensional reconstruction vectors has dimension  $k$ . So such values of  $\varepsilon$  should be avoided.

*Estimation of topological entropy.* The topological entropy of a dynamical system is a measure of the sensitiveness of evolutions on the initial states, or of its unpredictability. This quantity can be estimated from a time series by a procedure which is very much related to the above procedure for estimating the box counting dimension. It can also be

used as an indication as to whether we deal with a time series generated by a deterministic or by a random system, see [41]. We shall, however, not discuss this matter here. The reason is that, as we observed before, the actual estimation of these quantities is difficult and often unreliable. We shall introduce in Section 6 quantities which are easier to estimate and carry the same type of information. In the context of these quantities we shall also discuss a form of entropy.

*Distinguishing between deterministic and random time series.* So far, we have given a partial answer to how this question can be attacked. In Section 8, and in particular the Sections 8.2 and 8.3, we will be able to discuss this question in greater depth.

## 5. Stationarity and reconstruction measures

In this section we discuss some general notions concerning time series. These time series  $\{y_i\}_{i \geq 0}$  are supposed to be of infinite length; the elements  $y_i$  can be in  $\mathbb{R}$ , in  $\mathbb{R}^k$ , or in some manifold. This also means that a (positive) evolution of a dynamical system can be considered as a time series from this point of view. On the other hand we do not assume that our time series are generated by deterministic processes. Also, we do not consider our time series as realizations of some stochastic process, as is commonly done in texts in probability theory, i.e. we do not consider  $y_i$  as a function of  $\omega \in \Omega$ , where  $\Omega$  is some probability space; in Section 5.2 we shall, however, indicate how our approach can be interpreted in terms of stochastic processes.

The main notion which we discuss here is the notion of *stationarity*. This notion is related to the notion of predictability as we discussed it before: it means that the ‘statistical behaviour’ of the time series stays constant in time, so that knowledge of the past can be used to predict the future, or to conclude that the future cannot be (accurately) predicted. Equivalently, stationarity means that all kinds of averages, quantifying the ‘statistical behaviour’ of the time series, are well defined.

Before giving a formal definition, we give some (counter) examples. Standard examples of non-stationary time series are economical time series like prices as a function of time (say measured in years): due to inflation such a time series will, in average, increase, often at an exponential rate. This means that e.g. the average of such a time series (average over the infinite future) is not defined, at least not as a finite number. Quotients of prices in successive years have a much better chance of forming a stationary time series. These examples have to be taken with a grain of salt: such time series are not (yet) known for the whole infinite future. More sophisticated mathematical examples can be given. For example we consider the time series  $\{y_i\}$  with:

- $y_i = 0$  if the integer part of  $\ln(i)$  is even;
- $y_i = 1$  if the integer part of  $\ln(i)$  is odd.

It is not hard to verify that  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^n y_i$  does not exist. This non-existence of the average for the infinite future means that this time series is not stationary. Though this example may look pathological, we shall see that examples of this type occur as evolutions of dynamical systems.

### 5.1. Measures defined by relative frequencies

We consider a topological space  $M$  with a countable sequence of points  $\{p_i \in M\}_{i \in \mathbb{Z}_+}$ . We want to use this sequence to define a Borel probability measure  $\mu$  on  $M$ , which assigns to each (reasonable) subset  $A \subset M$  the average fraction of points of the sequence  $\{p_i\}$  which are in  $A$ , so

$$\mu(A) = \lim_{n \rightarrow \infty} \frac{\text{number of indices } 0 \leq i < n \text{ such that } p_i \in A}{n}.$$

If this measure is well defined we call it the *measure of relative frequencies* of the sequence  $\{p_i\}$ . One thing is that these limits may not be defined; another thing is that the specification of the ‘reasonable’ sets is not so obvious. This is the reason that, from a mathematical point of view, another definition of this measure of relative frequencies is preferred. For this definition we need our space  $M$  to be metrizable and to have a countable basis of open sets (these properties hold for all the spaces one encounters as state spaces of dynamical systems). Next we assume that for each continuous function  $\psi : M \rightarrow \mathbb{R}$  with compact support the limit

$$\hat{\mu}(\psi) = \lim_{n \rightarrow \infty} \frac{\sum_{i=0}^{n-1} \psi(p_i)}{n}$$

exists; if this limit does not exist for some continuous function with compact support, then the measure of relative frequencies is not defined. If all these limits exist, then  $\hat{\mu}$  is a continuous positive linear functional on the space of continuous functions with compact support. According to the Riesz representation theorem, e.g. see [37], there is then a unique (positive) Borel measure  $\mu$  on  $M$  such that for each continuous function with compact support  $\psi$  we have

$$\hat{\mu}(\psi) = \int \psi d\mu.$$

It is not hard to see that for this measure we have  $\mu(M) \leq 1$ . If  $\mu(M) = 1$ , we call  $\mu$  the measure of relative frequencies defined by  $\{p_i\}$ . Note that it may happen that  $\mu(M)$  is strictly smaller than 1, e.g. if we have  $M = \mathbb{R}$  and  $p_i \rightarrow \infty$ , for  $i \rightarrow \infty$  (like prices as a function of time in the above example): then  $\hat{\mu}(\psi) = 0$  for each function  $\psi$  with compact support, so that  $\mu(M) = 0$ . Also, in these cases where  $\mu(M) < 1$ , we say that the measure of relative frequencies is not defined.

### 5.2. Definition of stationarity and reconstruction measures

We first consider a time series  $\{y_i\}$  with values in  $\mathbb{R}$ . For each  $k$  we have a corresponding sequence of  $k$ -dimensional reconstruction vectors  $\{Y_i^k = (y_i, \dots, y_{i+k-1}) \in \mathbb{R}^k\}$ . We say that the time series  $\{y_i\}$  is stationary if for each  $k$  the measure of relative frequencies of  $\{Y_i^k\}$

is well defined. So here we need the time series to be of infinite length. In the case where the time series has values in a manifold  $M$  (or in a vector space  $E$ ), the definition is completely analogous. The only adaptation one has to make is that the elements  $Y_i^k = (y_i, \dots, y_{i+k-1})$  are not  $k$ -vectors, but elements of  $M^k$  (or of  $E^k$ ).

Assuming that our time series is stationary, the measures of relative frequencies of  $k$ -dimensional reconstruction vectors are called *k-dimensional reconstruction measures* and are denoted by  $\mu^k$ . These measures satisfy some obvious compatibility relations. For time series with values in  $\mathbb{R}$  these are: for any  $A \subset \mathbb{R}^k$ :

$$\mu^{k+1}(\mathbb{R} \times A) = \mu^k(A) = \mu^{k+1}(A \times \mathbb{R});$$

For any sequence of measures  $\mu^k$  on  $\mathbb{R}^k$ , satisfying these compatibility relations (and an ergodicity property, which is usually satisfied for reconstruction measures), there is a well defined time series in the sense of stochastic processes. This is explained in [5, Chapter 2.11].

### 5.3. Examples

We have seen that the existence of probability measures of relative frequencies, and hence the validity of stationarity, can be violated in two ways: one is that the elements of the sequence, or of the time series, under consideration, move off to infinity; the other is that limits of the form

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=0}^{n-1} \psi(p_i)}{n}$$

do not exist. If we restrict ourselves to time series generated by a dynamical system with compact state space, then nothing can move off to infinity, so then there is only the problem of existence of the limits of these averages. There is a general belief that for evolutions of generic dynamical systems with generic initial state, the probability measures of relative frequency are well defined. A mathematical justification of this belief has not yet been given. In my opinion this is one of the challenging problems in the ergodic theory of dynamical systems; see also [39,45].

There are however non-generic counterexamples. One is the map  $\varphi(x) = 3x$  modulo 1, defined on the circle  $\mathbb{R}$  modulo 1. If we take an appropriate initial point, then we get an evolution which is much like the 0-, 1-sequence which we gave as an example of a non-stationary time series. The construction is the following. As the initial state we take the point  $x_0$  which has, in the ternary system, the form  $0.\alpha_1\alpha_2\dots$  with

- $\alpha_i = 0$  if the integer part of  $\ln(i)$  is even;
- $\alpha_i = 1$  if the integer part of  $\ln(i)$  is odd.

If we consider the partition of  $\mathbb{R}$  modulo  $\mathbb{Z}$ , given by  $I_i = [\frac{i}{3}, \frac{i+1}{3})$ ,  $i = 0, 1, 2$ , then  $\varphi^i(x_0) \in I_{\alpha_i}$ ; both  $I_0$  and  $I_1$  have a unique expanding fixed point of  $\varphi$ : 0 and 0.5 respectively; long sequences of 0's or 1's correspond to long sojourns very near 0

respectively 0.5. This implies that for any continuous function  $\psi$  on the circle, which has different values in the two fixed points 0 and 0.5, the partial averages

$$\frac{\sum_{i=0}^{n-1} \psi(\varphi^i(x_0))}{n}$$

do not converge for  $n \rightarrow \infty$ .

For other examples see [43] and the references given in that paper.

## 6. Correlation dimensions and entropies

As we observed before, the box counting dimension of an attractor (or  $\omega$ -limit) can be estimated from a time series, but the estimation procedure is not very reliable. In this section we shall describe the correlation dimension and entropy, which admits better estimation procedures. Correlation dimension and entropy are special members of 1-parameter families of dimensions and entropies, which we will define at the end of this section. The present notion of dimension (and also of entropy) is only defined for a metric space *with a Borel probability measure*. Roughly speaking, the definition is based on the probabilities for two randomly and independently chosen points, for a given probability distribution, to be at a distance of, at most  $\varepsilon$ , for  $\varepsilon \rightarrow 0$ . In an  $n$ -dimensional space with probability measure having a positive and continuous density, one expects this probability to be of the order  $\varepsilon^n$ : once the first of the two points is chosen, the probability that the second point is within distance  $\varepsilon$  is equal to the measure of the  $\varepsilon$ -neighbourhood of the first point, and this is of the order  $\varepsilon^n$ .

This is formalized in the following definitions, in which  $K$  is a (compact) metric space, with metric  $\rho$  and Borel probability measure  $\mu$ .

DEFINITION. The  $\varepsilon$ -correlation integral  $C(\varepsilon)$  of  $(K, \rho, \mu)$  is the  $\mu \times \mu$  measure of

$$\Delta_\varepsilon = \{(p, q) \in K \times K \mid \rho(p, q) < \varepsilon\} \subset K \times K.$$

We note that the above definition is equivalent to

$$C(\varepsilon) = \int \mu(D(x, \varepsilon)) d\mu(x),$$

where  $D(x, \varepsilon)$  is the  $\varepsilon$ -neighbourhood of  $x$ , i.e. it is the  $\mu$ -average of the  $\mu$ -measure of  $\varepsilon$ -neighbourhoods. □

DEFINITION. The *correlation dimension*, or  $D_2$ -dimension, of  $(K, \rho, \mu)$  is

$$D_2(K) = \limsup_{\varepsilon \rightarrow 0} \frac{\ln(C_\varepsilon)}{\ln \varepsilon}. \quad \square$$

As in the case of the box counting dimension, one can also prove that here the unit cube in  $\mathbb{R}^n$ , with the Lebesgue measure, has  $D_2$ -dimension  $n$ ; any probability measure on  $\mathbb{R}^n$  has

its  $D_2$ -dimension at most equal to  $n$ . For more information on this and other dimensions, in particular in relation with dynamical systems, we refer to [34]. This  $D_2$ -dimension is a member of a 1-parameter family of dimensions defined by Renyi, and later generalized in several directions, which we will introduce at the end of this section in terms of the order  $q$  correlation integrals.

Next we come to the notion of entropy. This notion quantifies the sensitive dependence of evolutions on initial states. For the definition we need a compact metric space  $(K, \rho)$  with Borel probability measure  $\mu$  and a (continuous) map  $\varphi : K \rightarrow K$  defining the dynamics. Though not necessary for the definitions which we consider here, we note that usually one assumes that the probability measure  $\mu$  is invariant under the map  $\varphi$ , i.e. one assumes that for each open set  $U \subset K$  we have  $\mu(U) = \mu(\varphi^{-1}(U))$ . We use the map  $\varphi$  to define a family of metrics  $\rho_n$  on  $K$  in the following way:

$$\rho_n(p, q) = \max_{i=0, \dots, n-1} \rho(\varphi^i(p), \varphi^i(q)).$$

So  $\rho_1 = \rho$  and, if  $\varphi$  is continuous, then each one of the metrics  $\rho_n$  defines the same topology on  $K$ . If, moreover,  $\varphi$  is a map with bounded expansion, in the sense that there is a constant  $C > 1$  such that  $\rho(\varphi(p), \varphi(q)) \leq C\rho(p, q)$  for all  $p, q \in K$ , then  $\rho(p, q) \leq \rho_n(p, q) \leq C^{n-1}\rho(p, q)$ . Note that any  $C^1$ -map, defined on a compact set, has bounded expansion.

In this situation we consider the  $\varepsilon, n$ -correlation integral  $C^{(n)}(\varepsilon)$ , which is just the ordinary correlation integral, with the metric  $\rho$  replaced by  $\rho_n$ . Then we define the  $H_2$ -entropy as

$$H_2(K, \varphi) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{-\ln(C^{(n)}(\varepsilon))}{n}.$$

Recall that  $C^{(n)}(\varepsilon)$  is the  $\mu$ -average of  $\mu(D^{(n)}(x, \varepsilon))$  and that  $D^{(n)}(x, \varepsilon)$  is the  $\varepsilon$ -neighbourhood of  $x$  with respect to the metric  $\rho_n$ .

As we shall explain, this entropy is also a member of a 1-parameter family of entropies. For this and related definitions, see [46]. In order to see the relation between the entropy as defined here and the entropy as originally defined by Kolmogorov, we recall the Brin–Katok theorem.

**THEOREM (Brin–Katok, [6]).** *In the above situation with  $\mu$  a non-atomic Borel probability measure which is invariant and ergodic with respect to  $\varphi$ , for  $\mu$  almost every point  $x$  the following limit exists and equals the Kolmogorov entropy:*

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{-\ln(\mu(D^{(n)}(x, \varepsilon)))}{n}. \quad \square$$

This means that the Kolmogorov entropy is roughly equal to the exponential rate at which the measure of an  $\varepsilon$ -neighbourhood with respect to  $\rho_n$  decreases as a function of  $n$ . This is indeed a measure for the sensitive dependence of evolutions on initial states: it measures the decay, with increasing  $n$ , of the fraction of the initial states whose evolutions

stay within distance  $\varepsilon$  after  $n$  iterations. In our definition it is however the exponential decay of the *average* measure of  $\varepsilon$ -neighbourhoods with respect to the  $\rho_n$  metric which is taken as entropy. In general the Kolmogorov entropy is greater than or equal to the  $H_2$  entropy as we defined it.

*Compatibility of the definitions of dimension and entropy with reconstruction.* In the following section we consider the problem of estimating dimension and entropy, as defined here, based on the information of one time series. One then deals not with the original state space, but with the set of reconstruction vectors. One has to assume then that the (generic) assumptions in the reconstruction theorem are satisfied so that the reconstruction map  $Rec_k$ , for  $k$  sufficiently big, is an embedding, and hence a map of bounded distortion. Then the dimension and entropy of the  $\omega$ -limit of an orbit (with probability measure defined by relative frequencies) are equal to the dimension and entropy of that same limit set *after reconstruction*. This is based on the fact, which is easy to verify, that if  $h$  is a homeomorphism  $h : K \rightarrow K'$  between metric spaces  $(K, \rho)$  and  $(K', \rho')$ , such that the quotients

$$\frac{\rho(p, q)}{\rho'(h(p), h(q))},$$

with  $p \neq q$ , are bounded and bounded away from zero, and if  $\mu, \mu'$  are Borel probability measures on  $K, K'$ , respectively, such that  $\mu(U) = \mu'(h(U))$  for each open  $U \subset K$ , then the  $D_2$ -dimensions of  $K$  and  $K'$  are equal. If, moreover,  $\varphi : K \rightarrow K$  defines a dynamical system on  $K$ , and if  $\varphi' = h\varphi h^{-1} : K' \rightarrow K'$ , then also the  $H_2$ -entropies of  $\varphi$  and  $\varphi'$  are the same.

In the case where the dynamics is given by an endomorphism instead of a diffeomorphism, similar results hold, see [44]. If the assumptions in the reconstruction theorem are not satisfied, this may lead to underestimation of both dimension and entropy.

*Generalized correlation integrals, dimensions and entropies.* As we mentioned before, both the dimension and the entropy, as we defined it, are members of a 1-parameter family of dimensions, respectively, entropies. These notions are defined in a way which is very similar to the definition of the Rényi information of order  $q$ ; hence they are referred to as Rényi dimensions and entropies. These 1-parameter families can be defined in terms of generalized correlation integrals. The (generalized) correlation integral of order  $q$  of  $(K, \rho, \mu)$ , for  $q > 1$ , are defined as:

$$C_q(\varepsilon) = \sqrt[q-1]{\int (\mu(D(x, \varepsilon)))^{q-1} d\mu(x)},$$

where  $D(x, \varepsilon)$  is the  $\varepsilon$ -neighbourhood of  $x$ . We see that  $C_2(\varepsilon) = C(\varepsilon)$ . This definition is also used for  $q < 1$  but that is not very important for our purposes. For  $q = 1$  a different definition is needed, which can be justified by a continuity argument:

$$C_1(\varepsilon) = \exp \left( \int \ln(\mu(D(x, \varepsilon))) d\mu(x) \right).$$

The definitions of the generalized dimensions and entropies of order  $q$ , denoted by  $D_q$  and  $H_q$ , are obtained by replacing in the definitions of dimension and entropy, the ordinary correlation integrals by the order  $q$  correlation integrals. For a discussion of the literature on these generalized quantities and their estimation, also for  $q < 1$ , see at the end of Section 7.

## 7. Numerical estimation of correlation integrals and the corresponding dimensions and entropies

We are now in the position to describe how the above theory can be applied to analyze a time series which is generated by a deterministic system. As described before, in Section 3, the general form of a model generating such time series is a dynamical system, given by a compact manifold  $M$  with a diffeomorphism (or endomorphism)  $\varphi : M \rightarrow M$ , and a read out function  $f : M \rightarrow \mathbb{R}$ . We assume the generic conditions in the reconstruction theorem to be satisfied. A time series generated by such a system is then determined by the initial state  $x_0 \in M$ . The corresponding evolution is  $\{x_n = \varphi^n(x_0)\}$  with time series  $\{y_n = f(x_n)\}$ . It is clear that from such a time series we can only get information about this particular orbit, and, in particular, about its  $\omega$ -limit set  $\omega(x_0)$  and the dynamics on that limit set given by  $\varphi|_{\omega(x_0)}$ . As discussed in the previous section, for the definition of the  $D_2$ -dimension and the  $H_2$ -entropy we need some ( $\varphi$ -invariant) measure concentrated on the set to be investigated. Preferably this measure should be related to the dynamics. The only natural choice seems to be the measure of relative frequencies, as discussed in Section 5.2, of the evolution  $\{x_n\}$ , *if that measure exists*. We now assume that this measure of relative frequencies does exist for the evolution under consideration. We note that this measure of relative frequencies is  $\varphi$ -invariant and that it is concentrated on the  $\omega$ -limit set  $\omega(x_0)$ .

Given the time series  $\{y_n\}$ , we consider the corresponding time series of reconstruction vectors  $\{Y_n^{k_0} = (y_n, \dots, y_{n+k_0-1})\}$  for  $k_0$  sufficiently big, i.e. so that the sequence of reconstruction vectors and the evolution  $\{x_n\}$  are metrically equal up to bounded distortion, see Section 3. Then we also have the measure of relative frequencies for this sequence of reconstruction vectors. This measure is denoted by  $\mu$ ; its support is the limit set

$$\Omega = \left\{ Y \mid \text{for some } n_i \rightarrow \infty, \lim_{i \rightarrow \infty} Y_{n_i}^{k_0} = Y \right\}.$$

Clearly,  $\Omega = \text{Rec}_{k_0}(\omega(x_0))$  and the map  $\Phi$  on  $\Omega$ , defined by  $\text{Rec}_{k_0} \circ \varphi \circ (\text{Rec}_{k_0}|_{\omega(x_0)})^{-1}$  is the unique continuous map which sends each  $Y_n^{k_0}$  to  $Y_{n+1}^{k_0}$ .

Both the dimension and the entropy have been defined in terms of correlation integrals. As observed before, we may substitute the correlation integrals of  $\omega(x_0)$  by those of  $\Omega$ , as long as the transformation  $\text{Rec}_{k_0}$ , restricted to  $\omega(x_0)$ , has bounded distortion. An estimate of a correlation integral of  $\Omega$ , based on the first  $N$  reconstruction vectors, is given by:

$$\hat{C}(\varepsilon, N) = \frac{(\text{number of pairs } (i, j) \text{ such that } i < j \text{ and } \|Y_i^{k_0} - Y_j^{k_0}\| < \varepsilon)}{(N(N-1)/2)}.$$



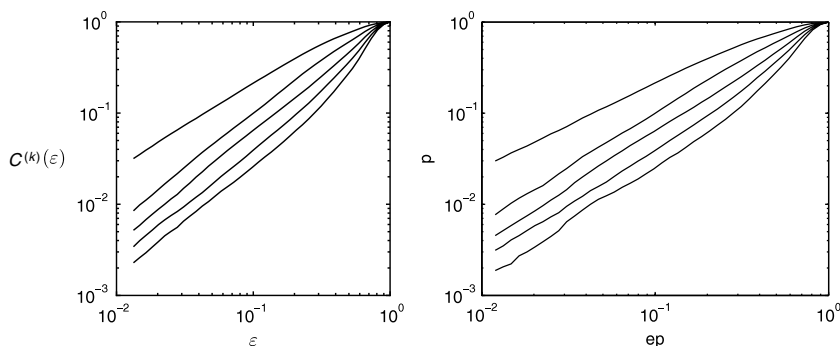


Fig. 3. Estimates for the correlation integrals of the Hénon attractor. For  $k = 1, \dots, 5$  the estimates of  $C^{(k)}(\varepsilon)$  are given as function of  $\varepsilon$ , with logarithmic scale along both axis. (This time series was normalized so that the difference between the maximal and the minimal values is 1.) Since it follows from the definitions that  $C^{(k)}(\varepsilon)$  decreases for increasing values of  $k$  the highest curve corresponds to  $k = 1$  and the lower curves to  $k = 2, \dots, 5$ , respectively.

It was proved in [13] that this estimate converges for  $N \rightarrow \infty$ , in the statistical sense, to the true value. In that paper information about the variance of this estimator was also obtained. In the definition of this estimate, one may take for  $\|\cdot\|$  the Euclidean norm, but in view of the usage of these estimates for the estimation of the entropy, it is better to take the *maximum norm*:  $\|(z_1, \dots, z_{k_0})\|_{\max} = \max_{i=1 \dots k_0} |z_i|$ . The reason is the following. If the distance function on  $\Omega$  is denoted by  $d$ , then, as in the previous section, when discussing the definition of entropy, we use the metrics

$$d_n(Y, Y') = \max_{i=0, \dots, n-1} d(\Phi^i(Y), \Phi^i(Y')).$$

If we now take for the distance function  $d$  on  $\Omega$  the maximum norm (for  $k_0$ -dimensional reconstruction vectors), then  $d_n$  is the corresponding maximum norm for  $(k_0 + n - 1)$ -dimensional reconstruction vectors. Note also that the transition from the Euclidean norm to the maximum norm is a ‘transformation with bounded distortion’, so that it should not influence the dimension estimates. When necessary, we express the dimension of the reconstruction vectors used in the estimation  $\hat{C}(\varepsilon, N)$  by writing  $\hat{C}^{(k)}(\varepsilon, N)$ ; as in Section 6, the correlation integrals based on  $k$ -dimensional reconstruction vectors are denoted by  $C^{(k)}(\varepsilon)$ .

In order to obtain estimates for the  $D_2$ -dimension and the  $H_2$ -entropy of the dynamics on the  $\omega$ -limit set, one needs to have estimates of  $C^{(k)}(\varepsilon)$  for many values of both  $k$  and  $\varepsilon$ . A common form to display such information graphically is a diagram showing the estimates of  $\ln(C^{(k)}(\varepsilon))$ , or rather the logarithms of the estimates of  $C^{(k)}(\varepsilon)$ , for the various values of  $k$ , as a function of  $\ln(\varepsilon)$ . In Figure 3 we show such estimates based on a time series of length 4000 which was generated with the Hénon system (with the usual parameter values  $a = 1.4$  and  $b = 0.3$ ). We discuss various aspects of this figure and then give references to the literature for the algorithmic aspects for extracting numerical information from the estimates as displayed.

First we observe that we normalized the time series in such a way that the difference between the maximal and minimal value equals 1. For  $\varepsilon$  we took the values  $1, .9, .9^2, \dots, .9^{42} \sim .011$ . It is clear that for  $\varepsilon = 1$  the correlation integrals have to be 1, independently of the reconstruction dimension  $k$ . Since  $C^{(k)}(\varepsilon)$  is decreasing as a function of  $k$ , the various graphs of (the estimated values of)  $\ln(C^{(k)}(\varepsilon))$  as a function of  $\ln(\varepsilon)$ , are ordered monotonically in  $k$ : the higher graphs belong to the lower values of  $k$ . For  $k = 1$  we see that  $\ln(C^{(1)}(\varepsilon))$ , as a function of  $\ln(\varepsilon)$ , is very close to a straight line, at least for values of  $\varepsilon$  smaller than say 0.2; the slope of this straight line is approximately equal to 1 (N.B.: to see this, one has to re-scale the figure in such a way that a factor of 10 along the horizontal and vertical axis correspond to the same length!). This is an indication that the set of 1-dimensional reconstruction vectors has a  $D_2$ -dimension equal to 1. This is to be expected: since this set is contained in a 1-dimensional space, the dimension is at most 1, and the dimension of the Hénon attractor is, according to numerical evidence, somewhat bigger than 1. Now we consider the curves for  $k = 2, \dots, 5$ . Also these curves approach straight lines for small values of  $\varepsilon$  (with fluctuations due, at least in part, to estimation errors). The slope of these lines is somewhat bigger than in the case of  $k = 1$  indicating that the reconstruction vectors, in the dimensions  $k > 1$ , form a set with  $D_2$ -dimension also somewhat bigger than 1. The slope indicates a dimension around 1.2. Also we see that the slopes of these curves do not seem to depend significantly on  $k$  for  $k \geq 2$ ; this is an indication that this slope corresponds to the true dimension of the Hénon attractor itself. This agrees with the fact that the reconstruction maps  $rec_k$ ,  $k > 1$ , for the Hénon system are embeddings (if we use one of the coordinates as a read out function). We point out again, that the estimation of dimensions, as carried out on the basis of estimated correlation integrals, can never be rigorously justified: The definition of the dimension is in terms of the behaviour of  $\ln(C^{(k)}(\varepsilon))$  for  $\varepsilon \rightarrow 0$  and if we have a time series of a given (finite) length, these estimates become more and more unreliable as  $\varepsilon$  tends to zero because the number of pairs of reconstruction vectors, which are very close, will become extremely small (except in the trivial case where the time series is periodic). Due to these limitations, one tries to find an interval, a so-called *scaling interval*, of  $\varepsilon$  values in which the estimated values of  $\ln(C^{(k)}(\varepsilon))$  are linear, up to small fluctuations, in  $\ln(\varepsilon)$ . Then the dimension estimate is based on the slope of the estimated  $\ln(C^{(k)}(\varepsilon))$  vs  $\ln(\varepsilon)$  curves in that interval. The justification for such a restriction is that for large values of  $\varepsilon$  no linear dependence is to be expected because correlation integrals are, by definition, at most equal to 1, that for small values of  $\varepsilon$ , the estimation errors become too big, and that for many systems we know that there is, in good approximation, such a linear dependence of  $\ln(C^{(k)}(\varepsilon))$  on  $\ln(\varepsilon)$  for intermediate values of  $\varepsilon$ . So the procedure is just based on an extrapolation to smaller values of  $\varepsilon$ . In the case of a time series generated by some (physical) experiment one also has to take into account that small fluctuations are always present due to thermal fluctuations or other noise; this provides another argument for disregarding values of  $\varepsilon$  which are too small.

Also an estimate for the  $H_2$ -entropy can be obtained from the estimated correlation integrals. The information in Figure 3 suggests here how to proceed. We observe that the parallel lines formed by the estimated  $\ln(C^{(k)}(\varepsilon))$  vs  $\ln(\varepsilon)$  curves, for  $k \geq 2$ , are not only parallel, but also at approximately equal distances. This is an indication

that the differences

$$\ln(C^{(k)}(\varepsilon)) - \ln(C^{(k+1)}(\varepsilon))$$

are approximately independent of  $\varepsilon$  and  $k$ , at least for small values of  $\varepsilon$ , here smaller than say 0.2, and big values of  $k$ , here bigger than 1. Assuming that this constancy holds in good approximation, this difference should be equal to the  $H_2$ -entropy as defined above.

We note that this behaviour of the estimated  $\ln(C^{(k)}(\varepsilon))$  vs  $\ln(\varepsilon)$  curves, namely that they approach equidistant parallel lines for small values of  $\varepsilon$  and big values of  $k$ , turns out to occur quite generally for time series generated by deterministic systems. It may however happen that, in order to reach the ‘good’  $\varepsilon$  and  $k$  values with reliable estimates of the correlation integrals, one has to take extremely long time series (with extremely small noise). For a quantification of these limitations, see [38].

This way of estimating the dimension and entropy was proposed in [42] and applied in [21,22]. There is an extensive literature on the question of how to improve these estimations. We refer to the (survey) publications and proceedings [17,48,11,50,12], and the textbook [27]. As we mentioned above, the  $D_2$ -dimension and  $H_2$ -entropy are both members of families of dimensions and entropies respectively. The theory of these generalized dimensions and entropies, also for  $q < 1$ ), and their estimation was considered in a number of publications; see [25,3,4,24,48,47,52] and the references in these works.

## 8. Classical time series analysis, the analysis in terms of correlation integrals, and predictability

What we here consider as classical time series analysis is the analysis of time series in terms of auto covariances and power spectra. This is a rather restricted view, but still it makes sense to compare that type of analysis with the analysis in terms of correlation integrals which we discussed so far. In the first section we give a résumé of this classical time series analysis in so far as we will need it; for more information see e.g. [35]. Then we show that the correlation integrals provide information which cannot be extracted from the classical analysis, in particular, the auto covariances cannot be used to distinguish whether a time series is generated by a deterministic system or a stochastic system. Finally we discuss predictability in terms of correlation integrals.

### 8.1. Classical time series analysis

We consider a stationary time series  $\{y_i\}$ . Without loss of generality we may assume that the average of this time series is 0. The  $k^{th}$  auto covariance is defined as the average  $R_k = \text{average}_i(y_i \cdot y_{i+k})$ . Note that  $R_0$  is called the variance of the time series. From the definition it follows that for each  $k$  we have  $R_k = R_{-k}$ .

We shall not make use of the power spectrum but, since it is used very much as a means to give graphical information about a signal (or time series) we just give a brief description of its meaning and its relation to the auto covariances. The power spectrum

gives information about how strong each (angular) frequency  $\omega$  is represented in the signal:  $P(\omega)$  is the energy, or squared amplitude, of the (angular) frequency  $\omega$ . Note that for a discrete time series, which we are discussing here, these frequencies only have values in  $[-\pi, \pi]$  and that  $P(\omega) = P(-\omega)$ . The formal definition is somewhat complicated due to possible difficulties with convergence: in general the power spectrum only exists as a generalized function, in fact a measure. It can be shown that the auto covariances are the Fourier transform of the power spectrum:

$$R_k = \int_{-\pi}^{\pi} e^{i\omega k} P(\omega) d\omega.$$

This means that not only are the auto covariances determined by the power spectrum, but that also, by the inverse Fourier transform, the power spectrum is determined by the auto covariances. So the information, which this classical theory can provide, is exactly the information which is contained in the auto covariances.

We note that under very general assumptions the auto covariances  $R_k$  tend to zero as  $k$  tends to infinity. (The most important exceptions are (quasi)-periodic time series.) This is what we assume from now on. If the auto covariances converge sufficiently fast to zero, then the power spectrum will be continuous, smooth or even analytic. For time series, generated by a deterministic system, the auto covariances converge to zero if the system is mixing, see [2].

*Optimal linear predictors.* For a time series as discussed above, one can construct the optimal linear predictor of order  $k$ . It is given by coefficients  $\alpha_1^{(k)}, \dots, \alpha_k^{(k)}$ ; it makes a prediction of the  $n^{\text{th}}$  element as a *linear* combination of the  $k$  preceding elements:

$$\hat{y}_n^{(k)} = \alpha_1^{(k)} y_{n-1} + \dots + \alpha_k^{(k)} y_{n-k}.$$

The values  $\alpha_1^{(k)}, \dots, \alpha_k^{(k)}$  are chosen in such a way that the average value  $\sigma_k^2$  of the squares of the prediction errors  $(\hat{y}_n^{(k)} - y_n)^2$  is minimal. This last condition explains why we call this an *optimal* linear predictor. It can be shown that the auto covariances  $R_0, \dots, R_k$  determine, and are determined by,  $\alpha_1^{(k)}, \dots, \alpha_k^{(k)}$  and  $\sigma_k^2$ . Also if we generate a time series  $\{z_i\}$  by using the autoregressive model (or system):

$$z_n = \alpha_1^{(k)} z_{n-1} + \dots + \alpha_k^{(k)} z_{n-k} + \varepsilon_n,$$

where the values of  $\varepsilon_n$  are independently chosen from a probability distribution, usually a normal distribution, with zero mean and variance  $\sigma_k^2$ , then the auto covariances  $\tilde{R}_i$  of this new time series satisfy, with probability 1,  $\tilde{R}_i = R_i$  for  $i = 0, \dots, k$ . This new time series, with  $\varepsilon_n$  taken from a normal distribution, can be considered as the ‘most unpredictable’ time series with the given first  $k+1$  auto covariances. By performing the above construction with increasing values of  $k$  one obtains time series which are, as far as power spectrum or auto covariances are concerned, converging to the original time series  $\{y_i\}$ .

We note that the analysis in terms of power spectra and auto covariances is, in particular, useful for *Gaussian time series*. These time series are characterized by the fact that their

reconstruction measures are normal; one can show then that:

$$\mu_k = |B_k|^{-1/2} (2\pi)^{-k/2} e^{-\langle x, B_k^{-1} x \rangle / 2} dx,$$

where  $B_k$  is a symmetric and invertible matrix which can be expressed in terms of  $R_0, \dots, R_{k-1}$ :

$$B_k = \begin{pmatrix} R_0 & R_1 & \cdot & \cdot & R_{k-1} \\ R_1 & R_0 & \cdot & \cdot & R_{k-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ R_{k-1} & R_{k-2} & \cdot & \cdot & R_0 \end{pmatrix}.$$

(For an exceptional situation where these formulae are not valid see the footnote in Section 8.2.) So these time series are, as far as reconstruction measures are concerned, completely determined by their auto covariances. Such time series are generated e.g. by the above type of autoregressive systems whenever the  $\varepsilon_n$ 's are normally distributed.

## 8.2. Determinism and auto covariances

We have discussed the analysis of time series in terms of correlation integrals in the context of time series which are generated by deterministic systems, i.e. systems with the property that the state at a given time determines all future states and for which the space of all possible states is compact and has a low, or at least finite, dimension. We call such time series 'deterministic'. The Gaussian time series belong to a completely different class. These time series are prototypes of what we call 'random' time series. One can use the analysis in terms of correlation integrals to distinguish between deterministic and Gaussian time series in the following way.

If one tries to estimate numerically the dimension of an underlying attractor (or  $\omega$ -limit) as in Section 7, one makes estimates of the dimension of the set of  $k$ -dimensional reconstruction vectors for several values of  $k$ . If one applies this algorithm to a Gaussian time series, then one should find that the set of  $k$ -dimensional reconstruction vectors has dimension  $k$  for all  $k$ . This contrasts with the case of deterministic time series where these dimensions are bounded by a finite number, namely the dimension of the relevant attractor (or  $\omega$ -limit). Also if one tries to estimate the entropy of a Gaussian time series by estimating the limits

$$H_2(\varepsilon) = \lim_{k \rightarrow \infty} (\ln(C^{(k)}(\varepsilon)) - \ln(C^{(k+1)}(\varepsilon))),$$

one finds that this quantity tends to infinity as  $\varepsilon$  tends to zero, even for a fixed value of  $k$ . This contrasts with the deterministic case, where this quantity has a finite limit, at least if the map defining the time evolution has bounded expansion.

It should be clear that the above arguments are valid for a much larger class of non-deterministic time series than just the Gaussian one.

Next we want to argue that the auto covariances cannot be used to make the distinction between random and deterministic systems. Suppose that we are given a time series (of which we still assume that the average is 0) with auto covariances  $R_0, \dots, R_k$ . Then such a time series, i.e. a time series with these same first  $k + 1$  auto covariances, can be generated by an autoregressive model, which is of course random.<sup>2</sup> On the other hand, one can also generate a time series with the same auto covariances by a deterministic system. This can be done by taking in the autoregressive model for the values  $\varepsilon_n$ , not random (and independent) values from some probability distribution, but values generated by a suitable deterministic system. Though deterministic,  $\{\varepsilon_n\}$  should still have (1) average equal to zero, (2) variance equal to the variance of the prediction errors of the optimal linear predictor of order  $k$ , and (3) all the other auto covariances equal to zero, i.e.  $R_i = 0$  for  $i > 0$ . These conditions, namely, imply that the linear theory cannot ‘see’ the difference between such a deterministic time series and a truly random time series of independent and identically distributed noise. The first two of the above conditions are easy to satisfy: one just has to apply an appropriate affine transformation to all elements of the time series.

In order to make an explicit example where all these conditions are satisfied, we start with the *tent map*. This is a dynamical system given by the map  $\varphi$  on  $[-1/2, 1/2]$ :

$$\begin{aligned}\varphi(x) &= +2x + 1/2 & \text{for } x \in [-1/2, 0) \\ \varphi(x) &= -2x + 1/2 & \text{for } x \in [0, 1/2).\end{aligned}$$

It is well known that for almost any initial point  $x \in [-1/2, 1/2]$ , in the sense of Lebesgue, the measure defined by the relative frequencies of  $x_n = \varphi^n(x)$ , is just the Lebesgue measure on  $[-1/2, +1/2]$ . We show that such a time series satisfies (up to scalar multiplication) the above conditions. For this we have to evaluate

$$R_i = \int_{-1/2}^{+1/2} x \varphi^i(x) dx.$$

Since for  $i > 0$  we have that  $\varphi^i(x)$  is an even function of  $x$ , this integral is 0. The variance of this time series is  $1/12$ . So in order to make the variance equal to  $\sigma^2$  one has to transform the state space, i.e. the domain of  $\varphi$ , to  $[-\sqrt{12}\sigma^2, +\sqrt{12}\sigma^2]$ . We denote this re-scaled evolution map by  $\varphi_{\sigma^2}$ . Now we can write down explicitly the (deterministic) system which generates a time series with the prescribed auto covariances as the time series generated by the autoregressive model:

$$y_n = \alpha_1^{(k)} y_{n-1} + \dots + \alpha_k^{(k)} y_{n-k} + \varepsilon_n,$$

---

<sup>2</sup>There is a degenerate exception where the variance of the optimal linear predictor of order  $k$  is zero. In that case the time series is multi-periodic and has the form

$$y_n = \sum_{i \leq s} (a_i \sin(\omega_i n) + b_i \cos(\omega_i n)),$$

with  $2s \leq k + 1$ . Then also the matrices  $B_k$ , see Section 8.1, are not invertible.

where the variance of  $\varepsilon_n$  is  $\sigma^2$ . The dynamics is given by:

$$\begin{aligned} u_1 &\mapsto \sum_{i=1}^k \alpha_i^{(k)} u_i + x \\ u_2 &\mapsto u_1 \\ u_3 &\mapsto u_2 \\ &\dots \\ u_k &\mapsto u_{k-1} \\ x &\mapsto \varphi_{\sigma^2}(x). \end{aligned}$$

The time series consists of the successive values of  $u_1$ , i.e. the read out function is given by  $f(u_1, \dots, u_k, x) = u_1$ .

We note that, instead of the tent map, we could have used the Logistic map  $x \mapsto 1 - 2x^2$  for  $x \in [-1, 1]$ .

We produced this argument in detail because it proves that the correlation integrals, which can be used to detect the difference between time series generated by deterministic systems and Gaussian time series, contain information which is not contained in the auto covariances, and hence also not in the power spectrum. We should expect that deterministic time series are much more predictable than random (or Gaussian) time series with the same auto covariances. In the next subsection we give an indication as to how the correlation integrals give an indication about the predictability of a time series.

### 8.3. Predictability and correlation integrals

The prediction problem for chaotic time series has attracted a lot of attention. The following proceedings volumes were devoted, at least in part, to it: [50,9,53,49].

The prediction methods proposed in these papers all strongly use the stationarity of the time series: the generating model is supposed to be unknown, so the only way to make predictions about the future is to use the behaviour in the past, assuming (by stationarity) that the future behaviour will be like the behaviour in the past. Roughly speaking there are two ways to proceed.

First, one can try to estimate a (nonlinear) model for the system generating the time series and then use this model to generate predictions. The disadvantage of this method is that estimating such a model is problematic. This is, amongst other things, due to the fact that the orbit, corresponding to the time series which we know, explores only a (small) part of the state space of the system so that only a (small) part of the evolution equation can be known.

Second, one can compare, for each prediction, the most recent values of the time series with values in the past: if one finds in the past a segment which is very close to the segment of the most recent values, one predicts that the continuation in the future will be like the continuation in the past (we give a more detailed discussion below). The disadvantage is that for *each* prediction one has to inspect the *whole* past, which can be very time consuming.

Though the second method also has its disadvantages it is less ad hoc than the first one and hence more suitable for a theoretical investigation. So we shall analyze the second method and show how the correlation integrals, and in particular the estimates of dimension and entropy, give indications about the predictability.

Note that in this latter procedure there are two ways in which we use the past: we have the ‘complete known past’ to obtain information about the statistical properties, i.e. the reconstruction measures, of the time series; using these statistical properties, we base our prediction of the next value on a short sequence of the most recent values. Recall also our discussion of the dripping tap in Section 2.

In order to discuss the predictability in terms of this second procedure in more detail, we assume that we know a finite, but long, segment  $\{y_0, \dots, y_N\}$  of a stationary time series. The question is how to predict the next value  $y_{N+1}$  of the time series. For this procedure, in its simplest form, we have to choose an integer  $k$  and a positive real number  $\varepsilon$ . We come back to the question of how to choose these values. (Roughly speaking, the  $k$  most recent values  $y_{N-k+1}, \dots, y_N$  should contain the relevant information as far as predicting the future is concerned, given the statistical information about the time series which can be extracted from the whole past;  $\varepsilon$  should be of the order of the prediction error which we are willing to accept.) We now inspect the past, searching for a segment  $y_m, \dots, y_{m+k-1}$  which is  $\varepsilon$ -close to the last  $k$  values, in the sense that  $|y_{m+i} - y_{N-k+1+i}| < \varepsilon$  for  $i = 0, \dots, k-1$ . If there is no such segment, we should try again with a smaller value of  $k$  or a bigger value of  $\varepsilon$ . If there is such a segment, then we predict the next value  $y_{N+1}$  to be  $\hat{y}_{N+1} = y_{m+k}$ .

This is about the most primitive way to use the past to predict the future and later we will discuss an improvement; see the discussion below of local linear predictors. For the moment we concentrate on this prediction procedure and try to find out how to assess the reliability of the prediction. We shall assume, for simplicity, that a prediction which is correct within an error of at most  $\varepsilon$  is considered to be correct; in the case of a bigger error, the prediction is considered to be incorrect. Now the question is: what is the probability of the prediction to be correct? This probability can be given in terms of correlation integrals. Namely the probability that two segments of length  $k$ , respectively,  $k+1$ , are equal up to an error of at most  $\varepsilon$  is  $C^{(k)}(\varepsilon)$ , respectively,  $C^{(k+1)}(\varepsilon)$ . So the probability that two segments of length  $k+1$ , given that the first  $k$  elements are equal up to an error of at most  $\varepsilon$ , also have their last elements equal up to an error of at most  $\varepsilon$ , is  $C^{(k+1)}(\varepsilon)/C^{(k)}(\varepsilon)$ . Note that this quotient decreases in general with increasing values of  $k$ . So in order to get good predictions we should take  $k$  so big that this quotient has essentially its limit value. This is the main restriction on how to take  $k$ , which we want otherwise to be as small as possible.

We have seen this quotient, or rather its logarithm, in the context of the estimation of entropy. We define:

$$H_2(\varepsilon) = \lim_{k \rightarrow \infty} (\ln(C^{(k)}(\varepsilon)) - \ln(C^{(k+1)}(\varepsilon))) = \lim_{k \rightarrow \infty} -\ln \left( \frac{C^{(k+1)}(\varepsilon)}{C^{(k)}(\varepsilon)} \right).$$

So assuming that  $k$  is chosen as above, the probability of giving a correct prediction is  $e^{-H_2(\varepsilon)}$ . This means that  $H_2(\varepsilon)$  and hence also  $H_2 = \lim_{\varepsilon \rightarrow 0} H_2(\varepsilon)$  is a measure of the *unpredictability* of the time series.



Also the estimate for the dimension gives an indication about the (un)predictability in the following way. Suppose we have fixed the value of  $k$  in the prediction procedure so that  $\frac{C^{(k+1)}(\varepsilon)}{C^{(k)}(\varepsilon)}$  is essentially at its limiting value (assuming for simplicity that this value of  $k$  is independent of  $\varepsilon$ ). Then, in order to have a reasonable probability of finding in the past a sequence of length  $k$  which agrees sufficiently with the sequence of the last  $k$  values, we should have that  $1/N$ , the inverse length of the time series, is small compared with  $C^{(k)}(\varepsilon)$ . This is the main limitation when choosing  $\varepsilon$ , which we otherwise want to be as small as possible. The dependence of  $C^{(k)}(\varepsilon)$  on  $\varepsilon$  is, in the limiting situation, i.e. for small  $\varepsilon$  and sufficiently big  $k$ , proportional to  $\varepsilon^{D_2}$ , so in the case of a high dimension, we cannot take  $\varepsilon$  very small and hence we cannot expect predictions to be very accurate. To be more precise, if we want to improve the accuracy of our predictions by a factor 2, then we expect to need a segment of past values which is  $2^{D_2}$  times longer.

*Local linear predictors.* As we mentioned before, the above procedure of using the past to predict the future is rather primitive. We discuss here briefly a refinement which is based on a combination of optimal linear prediction and the above procedure. This type of prediction was the subject of [8] which appeared in [50]. In this case we start in the same way as above, but now we collect *all* the segments from the past which are  $\varepsilon$ -close to the last segment of  $k$  elements. Let  $m_1, \dots, m_s$  be the first indices of these segments. We have then  $s$  different values, namely  $y_{m_1+k}, \dots, y_{m_s+k}$ , which we can use as predictions for  $y_{N+1}$ . This collection of possible predictions already gives a better idea of the variance of the possible prediction errors.

We can however go further. Assuming that there should be some functional dependence  $y_n = F(y_{n-1}, \dots, y_{n-k})$  with differentiable  $F$  (and such an assumption is justified by the reconstruction theorem if we have a time series which is generated by a smooth and deterministic dynamical system and if  $k$  is sufficiently big), then  $F$  admits locally a good linear, or rather affine, approximation (given by its derivative). In the case that  $s$ , the number of nearby segments of length  $k$ , is sufficiently large, one can estimate such a linear approximation. This means estimating the constants  $\alpha_0, \dots, \alpha_k$  such that the variance of  $\{\hat{y}_{m_i+k} - y_{m_i+k}\}$  is minimal, where  $\hat{y}_{m_i+k} = \alpha_0 + \alpha_1 y_{m_i+k-1} + \dots + \alpha_k y_{m_i}$ . The determination of the constants  $\alpha_0, \dots, \alpha_k$  is done by linear regression. The estimation of the next value is then  $\hat{y}_{N+1} = \alpha_0 + \alpha_1 y_N + \dots + \alpha_k y_{N-k+1}$ . This means that we use essentially a linear predictor, which is however only based on ‘nearby segments’ (the fact that we have here a term  $\alpha_0$ , which was absent in the discussion of optimal linear predictors, comes from the fact that here there is no analogue of the assumption that the average is zero).

Finally we note that if nothing is known about how a (stationary) time series is generated, it is not a priori clear that this method of local linear estimation will give better results. Also, a proper choice of  $k$  and  $\varepsilon$  is less obvious in that case; the [8].

## 9. Miscellaneous subjects

In this concluding section we consider two additional methods for the analysis of time series which are of a somewhat different nature. Both methods can be motivated by the

fact that the methods which we discussed up to now cannot be used to detect the difference between a time series and its time reversal. The *time reversal* of a time series  $\{y_i\}_{i \in \mathbb{Z}}$  is the time series  $\{\tilde{y}_i = y_{-i}\}_{i \in \mathbb{Z}}$  (for finite segments of a time series the definition is similar). It is indeed obvious from the definitions that autocorrelations and correlation integrals do not change if we reverse time. Still we can show with a simple example that a time series may become much more unpredictable by reversing the time — this also indicates that the predictability, as measured by the entropy, is a somewhat defective notion.

Our example is a time series generated by the tent map (using a generic initial point); for the definition see Section 8.2. If we know the initial state of this system, we can predict the future, but not the past. This is a first indication that time reversal has drastic consequences in this case. We make this more explicit. If we know the initial state of the tent map up to an error  $\varepsilon$ , then we know the next state up to an error  $2\varepsilon$ , so that a prediction of the next state has a probability of 50% to be correct if we allow for an error of at most  $\varepsilon$ . (This holds only for small  $\varepsilon$  and even then to first approximation due to the effect of boundary points and the ‘turning point’ at  $x = 0$ .) For predicting the previous state, which corresponds to ordinary prediction of the time series with the time reversed, we also expect a probability of 50% of a correct prediction, i.e. with an error not more than  $\varepsilon$ , but in this latter case the prediction error is, in general, much bigger: the inverse image of an  $\varepsilon$ -interval consists, in general, of two intervals of length  $\varepsilon/2$  at a distance which is in average approximately  $1/2$ !

In the next subsection we discuss the *Liapunov exponents* and their estimation from a time series. They give another indication for the (un)predictability of a time series, this time not invariant under time reversal.

In the second subsection we consider the (*smoothed*) *mixed correlation integrals* which were originally introduced in order to detect the possible differences between a time series and its time reversal. It can however be used much more generally as a means to test whether two time series have the same statistical behaviour or not.

### 9.1. Liapunov exponents

A good survey on this subject is the third chapter of [17] to which we refer for the details. We give here only a heuristic definition of these exponents and a sketch their numerical estimation.

We consider a dynamical system given by a differentiable map  $f : M \rightarrow M$ , where  $M$  is a compact  $m$ -dimensional manifold. For an initial state  $x \in M$  the Liapunov exponents describe how evolutions, starting infinitesimally close to  $x$ , diverge from the evolution of  $x$ . For these Liapunov exponents to be defined, we have to assume that there are subspaces of the tangent space at  $x$ :  $T_x = E_1(x) \supset E_2(x) \supset \cdots \supset E_s(x)$  and real numbers  $\lambda_1(x) > \lambda_2(x) > \cdots > \lambda_s(x)$  such that for a vector  $v \in T_x$  the following are equivalent:

$$\lim_{n \rightarrow \infty} \frac{\ln(\|df^n(v)\|)}{n} = \lambda_i;$$

$$v \in (E_i(x) - E_{i+1}(x)).$$

The norm  $\|\cdot\|$  is with respect to some Riemannian metric on  $M$ . The existence of these subspaces (and hence the convergence of these limits) for generic initial states is the subject

of Oseledets' multiplicative ergodic theorem, see [30,17]. Here generic does *not* mean belonging to an open and dense subset, but belonging to a subset which has full measure, with respect to any  $f$ -invariant measure. The numbers  $\lambda_1, \dots, \lambda_s$  are called the Liapunov exponents at  $x$ . They are, to a large extent, independent of  $x$ : in general they only depend on the attractor to which the orbit through  $x$  converges. In the notation we shall drop the dependence of the Liapunov exponents on the point  $x$ . We say that the Liapunov exponent  $\lambda_i$  has multiplicity  $k$  if  $\dim(E_i(x)) - \dim(E_{i+1}(x))$  is equal to  $k$ . Often each Liapunov exponent with multiplicity  $k > 1$  is repeated  $k$  times, so that we have exactly  $m$  Liapunov exponents with  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ .

When we know the map  $f$  and its derivative explicitly, we can estimate the Liapunov exponents using the fact that, for a generic basis  $v_1, \dots, v_m$  of  $T_x$ , and for each  $s = 1, \dots, m$

$$\sum_{i=1}^s \lambda_i = \lim_{n \rightarrow \infty} \frac{\ln(\|df^n(v_1, \dots, v_s)\|_s)}{n},$$

where  $\|df^n(v_1, \dots, v_s)\|_s$  denotes the  $s$ -dimensional volume of the parallelepiped spanned by the  $s$  vectors  $df^n(v_1), \dots, df^n(v_s)$ .

It should be clear that in the case of sensitive dependence on initial states one expects the first (and biggest) Liapunov exponent  $\lambda_1$  to be positive. There is an interesting conjecture relating Liapunov exponents to the dimension of an attractor, the Kaplan–Yorke conjecture [28], which can be formulated as follows. Consider the function  $L$ , defined on the interval  $[0, m]$  such that  $L(s) = \sum_{i=0}^s \lambda_i$  for integers  $s$  (and  $L(0) = 0$ ) and which interpolates linearly on the intervals in between. Since Liapunov exponents are by convention non-increasing as functions of their index, this function is concave. We define the Kaplan–Yorke dimension  $D_{KY}$  as the largest value of  $t \in [0, m]$  for which  $L(t) \geq 0$ . The conjecture claims that  $D_{KY} = D_1$ , where  $D_1$  refers to the  $D_1$ -dimension, as defined in Section 6, of the attractor (or  $\omega$ -limit) to which the evolution in question converges. For some 2-dimensional systems this conjecture was proven, see [29]; in higher dimensions there are counter examples, but so far no *persistent* counter examples have been found, so that it might be generically true. Another important relation in terms of Liapunov exponents is that, under ‘weak’ hypotheses (for details see [17] or [29]), the Kolmogorov (or  $H_1$ ) entropy of an *attractor* is the sum of its positive Liapunov exponents.

*Estimation of Liapunov exponents from a time series.* The main method of estimating Liapunov exponents, from the time series of a deterministic system, is due to Eckmann et al. [16]; for references to later improvements, see [31].

This method makes use of the reconstructed orbit. The main problem here is to obtain estimates of the derivatives of the mapping (in the space of reconstruction vectors). This derivative is estimated in a way which is very similar to the construction of the local linear predictors, see Section 8.3. Working with  $k$ -dimensional reconstruction, and assuming that  $Rec_k$  is an embedding from the state space  $M$  into  $\mathbb{R}^k$ , we try to estimate the derivative at each reconstruction vector  $Y_i^k$  by collecting a sufficiently large set of reconstruction vectors  $\{Y_{j(i,s)}^k\}_{1 \leq s \leq s(i)}$  which are close to  $Y_i^k$ . Then we determine, using a least squares

method, a linear map  $A_i$  which sends each  $Y_{j(i,s)}^k - Y_i^k$  or  $Y_{j(i,s)+1}^k - Y_{i+1}^k$  plus a small error  $\varepsilon_{i,s}$  (of course so that the sum the squares of these  $\varepsilon_{i,s}$  is minimized). These maps  $A_i$  are used as estimates of the derivative of the dynamics at  $Y_i^k$ . The Liapunov exponents are then estimated, based on an analysis of the growth of the various vectors under the maps  $A_N A_{N-1} \dots A_1$ . A main problem with this method, and also with its later improvements, is that the linear maps which we are estimating are too high dimensional: we should use the derivative of the map on  $\mathcal{X}_k$  defined by  $F = \text{Rec}_k f \text{Rec}_k^{-1}$ , with  $\mathcal{X}_k = \text{Rec}_k(M)$ , while we are estimating the derivative of a (non defined) map on the higher dimensional  $\mathbb{R}^k$ . In fact the situation is even worse: the reconstruction vectors only give information on the map  $F$  restricted to the reconstruction of one orbit (and its limit set) which usually is of even lower dimension. Still, the method often gives reproducible results, at least as far as the biggest Liapunov exponent(s) is (are) concerned.

## 9.2. The Kantz–Diks test – discriminating between time series and testing for reversibility

The main idea of the Kantz–Diks test can be described in a context which is more general than that of time series. We assume we have two samples of vectors in  $\mathbb{R}^k$ , namely  $\{X_1, \dots, X_s\}$  and  $\{Y_1, \dots, Y_t\}$ . We assume that the  $X$ -vectors are chosen independently from a distribution  $P_X$  and the  $Y$ -vectors from a distribution  $P_Y$ . The question is whether these samples indicate, in a statistically significant way, that the distributions  $P_X$  and  $P_Y$  are different. The original idea of Kantz was to use the correlation integrals in combination with a ‘mixed correlation integral’ [26]. We denote the correlation integrals of the distributions  $P_X$  and  $P_Y$ , respectively, by  $C_X(\varepsilon)$  and  $C_Y(\varepsilon)$  — they are as defined in Section 6. The mixed correlation integral  $C_{XY}(\varepsilon)$  is defined as the probability that the distance between a  $P_X$ -random vector and a  $P_Y$ -random vector is component wise smaller than  $\varepsilon$ . If the distributions  $P_X$  and  $P_Y$  are the same, then these three correlation integrals are equal. If not, one expects the mixed correlation integral to be smaller (that need not be the case, but this problem is removed by using the smoothed correlation integrals which we define below). So if an estimate of  $C_X(\varepsilon) + C_Y(\varepsilon) - 2C_{XY}(\varepsilon)$  differs in a significant way from zero, this is an indication that  $P_X$  and  $P_Y$  are different.

A refinement of this method was given in [15]. It is based on the notion of *smoothed correlation integrals*. In order to explain this notion, we first observe that the correlation integral  $C_X(\varepsilon)$  can be defined as the expectation value, with respect to the measure  $P_X \times P_X$ , of the function  $H_\varepsilon(v, w)$  on  $\mathbb{R}^k \times \mathbb{R}^k$  which is 1 if the vectors  $v$  and  $w$  differ, component wise, less than  $\varepsilon$  and 0 otherwise. In the definition of the smoothed integral  $S_\varepsilon(X)$ , the only difference is that the function  $H_\varepsilon$  is replaced by

$$G_\varepsilon(v, w) = e^{-\frac{\|v-w\|}{\varepsilon^2}}$$

where  $\|\cdot\|$  denotes the Euclidean distance. The smoothed correlation integrals  $S_Y(\varepsilon)$  and  $S_{XY}(\varepsilon)$  are similarly defined. This modification has the following consequences:

- the quantity  $S_X(\varepsilon) + S_Y(\varepsilon) - 2S_{XY}(\varepsilon)$  is always non negative;
- $S_X(\varepsilon) + S_Y(\varepsilon) - 2S_{XY}(\varepsilon) = 0$  if and only if the distributions  $P_X$  and  $P_Y$  are equal.

The estimation of  $S_X(\varepsilon) + S_Y(\varepsilon) - 2S_{XY}(\varepsilon)$ , as a test for the distribution  $P_X$  and  $P_Y$  to be equal or not, can be used for various purposes. In the present context it is used to test whether two (finite segments of) time series show a significantly different behaviour, by applying it to the distributions of their reconstruction vectors. One has to be aware however of the fact that these vectors cannot be considered as statistically independent; this is discussed in [14].

Originally this method was proposed as a test for ‘reversibility’, i.e. as a test of whether or not a time series and its time reversal are significantly different. Another application is a test for stationarity. This is used when monitoring a process in order to obtain a warning whenever the dynamics changes. Here one uses a segment during which the dynamics is as desired. Then one collects, say every 10 minutes, a segment and applies the test to see whether there is a significant difference between the last segment and the standard segment. See [51] for an application of this method in chemical engineering.

Finally we note that for Gaussian time series the  $k$ -dimensional reconstruction measure, and hence the correlation integrals up to order  $k$ , are completely determined by the auto covariances  $\rho_0, \dots, \rho_{k-1}$ . There is however no explicit formula for correlation integrals in terms of these auto covariances. For the smoothed correlation integrals such formulas are given in [T,2003] where also the dimensions and entropies in terms of smoothed correlation integrals are discussed.

## References

- [1] D. Aeyels, *Generic observability of differentiable systems*, SIAM J. Control Optim. **19** (1981), 595–603.
- [2] V.I. Arnold and A. Avez, *Problèmes ergodiques de la mécanique classique*, Gautier-Villars (1967); English translation: Addison-Wesley, 1988.
- [3] R. Badii and A. Politi, *Hausdorff dimension and uniformity factor of strange attractors*, Phys. Rev. Lett. **52** (1984), 1661–1664.
- [4] R. Badii and A. Politi, *Statistical description of chaotic attractors: the dimension function*, J. Stat. Phys. **40** (1985), 725–750.
- [5] D.R. Brillinger, *Time Series*, McGraw-Hill (1981).
- [6] M. Brin and A. Katok, *On local entropy*, Geometric Dynamics, J. Palis, ed., LNM, Vol. 1007 (1983).
- [7] H.W. Broer and F. Takens, *Dynamical Systems and Chaos*, Epsilon Uitgaven, Vol. 64 (2009); Springer Appl. Math. Sci., Vol. 172 (2010).
- [8] M. Casdagli, *Chaos and deterministic versus stochastic non-linear modelling*, J. R. Statist. Soc. B **54** (1992), 303–328.
- [9] M. Casdagli and S. Eubank, eds, *Nonlinear Modelling and Forecasting*, Addison-Wesley (1992).
- [10] J.P. Crutchfield, J.D. Farmer, N.H. Packard and R.S. Shaw, *Chaos, Scientific American (December)*, (1986), pp. 38–49.
- [11] C.D. Cutler, *Some results on the behaviour and estimation of the fractal dimension of distributions on attractors*, J. Stat. Phys. **62** (1991), 651–708.
- [12] C.D. Cutler and D.T. Kaplan, eds, *Nonlinear dynamics and time series*, Fields Inst. Communications, Vol. 11, AMS (1997).
- [13] M. Denker and G. Keller, *Rigorous statistical procedures for data from dynamical systems*, J. Stat. Phys. **44** (1986), 67–93.
- [14] C. Diks, *Nonlinear Time Series Analysis*, World Scientific (1999).
- [15] C. Diks, W.R. van Zwet, F. Takens and J. DeGoede, *Detecting differences between delay vector distributions*, Phys. Rev. E **53** (1996), 2169–2176.
- [16] J.-P. Eckmann, S. Oliffson Kamphorst, D. Ruelle and S. Ciliberto, *Liapunov exponents from time series*, Phys. Rev. A **34** (1986), 4971–4979.

- [17] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Rev. Mod. Phys. **57** (1985), 617–656.
- [18] K.J. Falconer, *The Geometry of Fractal Sets*, Cam. Univ. Press (1985).
- [19] K. Falconer, *Fractal Geometry*, Wiley (1990).
- [20] A. Galka, *Topics in Nonlinear Time Series Analysis, with Implications for EEG Analysis*, World Scientific (2000).
- [21] P. Grassberger and I. Procaccia, *Characterisation of strange attractors*, Phys. Rev. Lett. **50** (1983), 346–349.
- [22] P. Grassberger and I. Procaccia, *Estimation of Kolmogorov entropy from a chaotic signal*, Phys. Rev. A **28** (1983), 2591–2593.
- [23] P. Grassberger and I. Procaccia, *Dimensions and entropies of strange attractors from a fluctuating dynamics approach*, Physica D **13** (1984), 34–54.
- [24] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia and B.I. Shraiman, *Fractal measures and their singularities: the characterisation of strange sets*, Phys. Rev. A **33** (1986), 1141–1151.
- [25] H.G.E. Hentschel and I. Procaccia, *The infinite number of generalised dimensions of fractals and strange attractors*, Physica D **8** (1983), 435–444.
- [26] H. Kantz, *Quantifying the closeness of fractal measures*, Phys. Rev. E **49** (1994), 5091–5097.
- [27] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Camb. Univ. Press (1997).
- [28] J.L. Kaplan and J.A. Yorke, *Chaotic behaviour of multidimensional difference equations*, Functional Differential Equations and Approximations of Fixed Points, H.-O. Peitgen and H.O. Walter, eds, LNM, Vol. 730, Springer (1979).
- [29] F. Ledrappier and L.S. Young, *The metric entropy of diffeomorphisms, Part I and II*, Ann. of Math. **122** (1985), 509–539 and 540–574.
- [30] V.I. Oseledets, *A multiplicative ergodic theorem*, Trans. Moscow Math. Soc. **19** (1968), 197–221.
- [31] E. Ott, T. Sauer and J.A. Yorke, eds, *Coping with Chaos*, Wiley (1994).
- [32] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, *Geometry from time series*, Phys. Rev. Lett. **45** (1980), 712–716.
- [33] H.-O. Peitgen, H. Jürgens and D. Saupe, *Chaos and Fractals*, Springer (1992).
- [34] Y.B. Pesin, *Dimension Theory in Dynamical Systems*, University of Chicago Press (1997).
- [35] M.B. Priesley, *Spectral Analysis and Time Series*, Academic Press (1981).
- [36] A. Renyi, *Probability Theory*, North-Holland (1970).
- [37] W. Rudin, *Real and Complex Analysis*, McGraw Hill (1970).
- [38] D. Ruelle, *Deterministic chaos: the science and the fiction*, Proc. R. Soc. London A **427** (1990), 241–248.
- [39] D. Ruelle, *Historical behaviour in smooth dynamical systems*, Global Analysis of Dynamical Systems, H.W. Broer, B. Krauskopf and G. Vegter, eds, IoP Publishing (2001).
- [40] T. Sauer, J.A. Yorke and M. Casdagli, *Embedology*, J. Stat. Phys. **65** (1991), 579–616.
- [41] F. Takens, *Detecting strange attractors in turbulence*, Dynamical Systems and Turbulence, Warwick 1980, D. Rand and L.-S. Young, eds, LNM, Vol. 898, Springer (1981).
- [42] F. Takens, *Invariants related to dimension and entropy*, Atas do 13<sup>o</sup> colóquio bras. de mat. 1981, IMPA, Rio de Janeiro (1983).
- [43] F. Takens, *Heteroclinic attractors: time averages and moduli of topological conjugacy*, Bol. Soc. Bras. Mat. **25** (1994), 107–120.
- [44] F. Takens, *The reconstruction theorem for endomorphisms*, Bol. Bras. Math. Soc. **33** (2002), 231–262.
- [45] F. Takens, *Orbits with historic behaviour, or non-existence of averages*, Nonlinearity **21** (2008), T33–T36.
- [46] F. Takens and E. Verbitski, *Generalised entropies: Rényi and correlation integral approach*, Nonlinearity **11** (1998), 771–782.
- [47] F. Takens and E. Verbitski, *General multi-fractal analysis of local entropies*, Fund. Math. **165** (2000), 203–237.
- [48] J. Theiler, *Estimating fractal dimension*, J. Opt. Soc. Am. **7** (1990), 1055–1073.
- [49] H. Tong, ed., *Chaos and Forecasting*, World Scientific (1995).
- [50] H. Tong and R.L. Smith, eds, Royal statistical society meeting on chaos, J. R. Stat. Soc. B **54** (1992), 301–474.
- [51] J.R. van Ommen, M.-O. Coppens, J.C. Schouten and C.M. van den Bleek, *Early warnings of agglomeration in fluidised beds by attractor comparison*, A.I.Ch.E. Journal **46** (2000), 2183–2197.
- [52] E.A. Verbitski, *Generalised entropies and dynamical systems*, Thesis, Rijksuniversiteit Groningen (2000).
- [53] A.S. Weigend and N.A. Gershenfeld, eds, *Time Series Prediction*, Addison-Wesley (1994).
- [54] H. Whitney, *Differentiable manifolds*, Ann. of Math. **37** (1936), 645–680.



## CHAPTER 8

# Homoclinic and Heteroclinic Bifurcations in Vector Fields

Ale Jan Homburg

*Korteweg-de Vries Institute for Mathematics, University of Amsterdam, 1018 TV Amsterdam, The Netherlands*

Björn Sandstede

*Division of Applied Mathematics, Brown University, Providence, RI 02906, USA*

### Contents

|                                                                                    |     |
|------------------------------------------------------------------------------------|-----|
| 1. Introduction . . . . .                                                          | 381 |
| 2. Homoclinic and heteroclinic orbits, and their geometry . . . . .                | 383 |
| 2.1. Homoclinic orbits to hyperbolic equilibria . . . . .                          | 383 |
| 2.2. Homoclinic orbits to nonhyperbolic equilibria . . . . .                       | 389 |
| 2.3. Heteroclinic cycles with hyperbolic equilibria . . . . .                      | 391 |
| 3. Analytical and geometric approaches . . . . .                                   | 393 |
| 3.1. Normal forms and linearizability . . . . .                                    | 394 |
| 3.2. Shil'nikov variables . . . . .                                                | 396 |
| 3.3. Lin's method . . . . .                                                        | 399 |
| 3.4. Homoclinic centre manifolds . . . . .                                         | 400 |
| 3.5. Stable foliations . . . . .                                                   | 401 |
| 3.6. Case study: The creation of periodic orbits from a homoclinic orbit . . . . . | 403 |
| 4. Phenomena . . . . .                                                             | 406 |
| 4.1. N-pulses and N-periodic orbits . . . . .                                      | 406 |
| 4.2. Robust singular dynamics . . . . .                                            | 407 |
| 4.3. Singular horseshoes . . . . .                                                 | 410 |
| 4.4. The boundary of Morse–Smale flows . . . . .                                   | 414 |
| 4.5. Homoclinic-doubling cascades . . . . .                                        | 416 |
| 4.6. Intermittency . . . . .                                                       | 418 |
| 5. Catalogue of homoclinic and heteroclinic bifurcations . . . . .                 | 420 |
| 5.1. Homoclinic orbits in generic systems . . . . .                                | 420 |
| 5.2. Heteroclinic cycles in generic systems . . . . .                              | 440 |
| 5.3. Conservative and reversible systems . . . . .                                 | 451 |
| 5.4. Homoclinic orbits arising through local bifurcations . . . . .                | 467 |
| 5.5. Equivariant systems . . . . .                                                 | 476 |



- 6. Related topics . . . . . 497
  - 6.1. Topological indices . . . . . 498
  - 6.2. Moduli . . . . . 500
  - 6.3. Existence results . . . . . 504
  - 6.4. Numerical techniques . . . . . 505
  - 6.5. Variational methods . . . . . 506
  - 6.6. Singularly perturbed systems . . . . . 506
  - 6.7. Infinite-dimensional systems . . . . . 507
- References . . . . . 509

## 1. Introduction

Our goal in this paper is to review the existing literature on homoclinic and heteroclinic bifurcation theory for flows. More specifically, we shall focus on bifurcations from homoclinic and heteroclinic orbits between equilibria in autonomous ordinary differential equations (ODEs)

$$\frac{du}{dt} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^n \times \mathbb{R}^m, \quad t \in \mathbb{R}. \quad (1.1)$$

Throughout the entire survey, we shall assume that the nonlinearity  $f$  is sufficiently smooth for the results to hold (the precise requirements can be found in the cited references). We write  $\mathcal{X}(\mathbb{R}^n)$  for the space of ODEs on  $\mathbb{R}^n$  endowed with the  $C^\infty$  Whitney topology.

Equilibria  $p$  of (1.1) are time-independent solutions that therefore satisfy  $f(p, \mu) = 0$ . We say that a solution  $h(t)$  of (1.1) is a heteroclinic orbit if  $h(t) \rightarrow p_\pm$  as  $t \rightarrow \pm\infty$  for equilibria  $p_\pm \in \mathbb{R}^n$ . If  $p_- = p_+$ , we say that  $h(t)$  is a homoclinic orbit (assuming tacitly that  $h(t)$  is not the equilibrium solution itself). We will also consider heteroclinic cycles which, by definition, consist of several heteroclinic orbits  $h_j(t)$  labelled by the index  $j = 1, \dots, \ell$  so that

$$\lim_{t \rightarrow \infty} h_j(t) = p_{j+1} = \lim_{t \rightarrow -\infty} h_{j+1}(t), \quad j = 1, \dots, \ell$$

with the understanding that  $h_{\ell+1} := h_1$  and  $p_{\ell+1} = p_1$ . Illustrations of homoclinic and heteroclinic orbits can be found in Figure 1.1.

Homoclinic and heteroclinic orbits play an important role in applications. For instance, we may be interested in modelling action potentials in nerve axons by an ODE of the form (1.1): in this case, we can think of  $u$  as representing the electric potential and certain ion concentrations in the nerve axon, with  $t$  being physical time. The rest state of the axon determines a natural equilibrium of (1.1), and action potentials in the axon correspond then to homoclinic orbits to this equilibrium. Heteroclinic orbits typically arise when a system can cycle through several different states. For instance, in certain Rayleigh–Bénard convection experiments, roll patterns may arise that can orient themselves at angles of 0, 120 or 240 degrees: as time progresses, the roll pattern cycles through this set of angles, but stays at each angle for a long time, followed by a fast transition to the next angle. Thus, in an appropriate ODE model with three equilibria corresponding to the rolls oriented at angles of 0, 120 and 240 degrees, heteroclinic orbits correspond to transitions from one equilibrium to another.

In the above examples, the independent variable  $t$  corresponds to physical time, whence we refer to such applications as describing *temporal dynamics*. Another class of applications, referred to as *spatial dynamics*, are problems where  $t$  represents a spatial direction. Important examples in this respect are the travelling-wave solutions of partial differential equations (PDEs) on unbounded domains. Consider, for instance, the reaction-diffusion system

$$U_t = DU_{xx} + F(u), \quad U \in \mathbb{R}^N, \quad x \in \mathbb{R}, \quad (1.2)$$

where  $t$  and  $x$  represent physical time and space, respectively. A travelling wave of (1.2) is

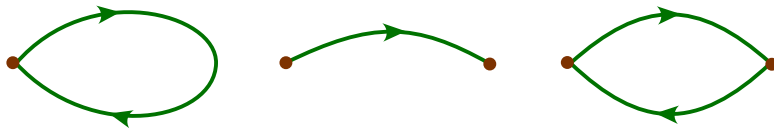


Fig. 1.1. The panels contain a homoclinic orbit, which connects an equilibrium to itself [left], a heteroclinic orbit that connects two different equilibria [centre], and a heteroclinic cycle with two connecting orbits [right].

a solution of the form  $U(x, t) = U_*(x - ct)$  corresponding to a fixed profile  $U_*$  that travels to the right (for  $c > 0$ ) or the left (for  $c < 0$ ) as a function of time  $t$ . Upon substituting the ansatz  $U(x, t) = U_*(x - ct)$  into (1.1) and using  $\xi = x - ct$  as the new independent variable, we find that  $(U, V) = (U_*, \partial_\xi U_*)$  must satisfy the ODE

$$\frac{d}{d\xi} \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} V \\ -D^{-1}(cV + F(U)) \end{pmatrix} \quad (1.3)$$

which is of the form (1.1) with  $\mu$  representing the wave speed  $c$ . Any solution of (1.3) gives a travelling wave with speed  $c$  of the PDE (1.2). In particular, homoclinic orbits of (1.3) correspond to pulses, i.e. to travelling waves that are localized in space, while heteroclinic orbits correspond to fronts which are waves that become constant as  $x \rightarrow \pm\infty$ . Pulses and fronts are of particular importance in applications. We remark that, once their existence is established, it is of interest to determine whether or not these structures are stable with respect to the PDE dynamics associated with (1.2): we refer the reader to [6] for a survey of this topic.

With these motivating examples in mind, we now turn to a discussion of the relevant issues and questions that surround homoclinic and heteroclinic orbits. One such issue is, of course, to establish the existence of these orbits in the first place, for instance by analytical or numerical techniques. The topic we shall focus on in this survey paper, however, is the dynamics of (1.1) near given homoclinic or heteroclinic orbits, particularly under changes of a systems parameter  $\mu$ . This includes the persistence of homoclinic and heteroclinic orbits under parameter variations but also, more importantly, the characterization of all recurrent orbits, that is, of all solutions that stay in a fixed tubular neighbourhood of a given homoclinic orbit or heteroclinic cycle for all times. Particularly interesting recurrent orbits are  $N$ -homoclinic and  $N$ -periodic orbits: these are solutions that follow the original homoclinic orbit or heteroclinic cycle  $N$ -times before closing up. In other words, these solutions have winding number  $N$  when considered as loops inside a tubular cylindrical neighbourhood of a homoclinic orbit or a heteroclinic cycle.

In the spirit of local bifurcation theory, we are then interested in identifying bifurcation scenarios at which the recurrent dynamics near a set of connecting orbits changes qualitatively. At such bifurcation points,  $N$ -homoclinic orbits may spin off or complicated dynamics may set in. An important characteristic feature of bifurcation points is their codimension, which determines how many parameters we need to adjust before being able to observe a given bifurcation scenario. The codimension typically depends strongly on whether the underlying ODE has any additional structure such as respecting a group of symmetries or being time-reversible.

These are roughly the questions and topics that we wish to review and discuss in this survey paper. In Section 2, we shall review geometric properties of homoclinic orbits and introduce various hypotheses that will be used throughout the remainder of the paper. Section 3 contains a discussion of the analytical and geometric techniques that have been developed to investigate connecting orbits. Several common phenomena that arise in many different bifurcation scenarios are summarized in Section 4. The core of this paper is Section 5 where we give a catalogue of homoclinic and heteroclinic bifurcations for vector fields. We conclude with a brief discussion of related topics in Section 6.

References to original publications as well as books and review papers relevant to the topics of this article will be included in the sections below. Here, among books devoted to global bifurcation theory, we single out [375,376] by L.P. Shil'nikov, A.L. Shil'nikov, Turaev and Chua and [207] by Il'yashenko and Li. Further useful books containing sections on global bifurcation theory include [12,247] and, for equivariant systems, [93].

## 2. Homoclinic and heteroclinic orbits, and their geometry

This section serves as an introduction to homoclinic and heteroclinic orbits and to set up many of the hypotheses and assumptions that we shall refer to later when discussing homoclinic and heteroclinic bifurcations. Specifically, we will use this section to illustrate various geometric notions and how they are encoded and reflected analytically. To set the scene, consider again the ODE

$$\dot{u} = f(u, \mu), \quad (2.1)$$

where  $u \in \mathbb{R}^n$  and  $\mu = (\mu_1, \dots, \mu_d) \in \mathbb{R}^d$ .

### 2.1. Homoclinic orbits to hyperbolic equilibria

Assume that  $h(t)$  is a given homoclinic orbit of (2.1), say for  $\mu = 0$ , which converges to the equilibrium  $p$  as  $t \rightarrow \pm\infty$ . One key assumption that we will often impose is that the equilibrium  $p$  itself does not undergo a local bifurcation at  $\mu = 0$ . A sufficient condition is hyperbolicity of  $p$  which, by definition, means that the matrix  $f_u(p, 0)$ , which is obtained by linearizing the right-hand side of (2.1) with respect to  $u$  at  $u = p$  for  $\mu = 0$ , has no eigenvalues on the imaginary axis. We summarize this assumption as follows:

**HYPOTHESIS 2.1 (Hyperbolicity).** *The equilibrium  $p$  is hyperbolic at  $\mu = 0$ , that is, the linearization  $f_u(p, 0)$  has no eigenvalues on the imaginary axis.*

For the remainder of this section, we assume that  $p$  is a hyperbolic equilibrium of (2.1) and refer the reader to Section 2.2 for material on the geometry of the flow near homoclinic orbits to nonhyperbolic equilibria. The stable and unstable manifolds of a hyperbolic equilibrium  $p$  are defined by

$$\begin{aligned} W^s(p, 0) &= \{u(0); u(t) \text{ satisfies (2.1) and } u(t) \rightarrow p \text{ as } t \rightarrow \infty\} \\ W^u(p, 0) &= \{u(0); u(t) \text{ satisfies (2.1) and } u(t) \rightarrow p \text{ as } t \rightarrow -\infty\}; \end{aligned}$$

see [7]. These sets turn out to be smooth immersed manifolds that are invariant under the flow, and the convergence towards  $p$  is, in fact, exponential in  $t$ .

If  $h(t)$  is a homoclinic orbit to  $p$  for  $\mu = 0$ , then its entire orbit must lie in the intersection of stable and unstable manifolds of  $p$  so that  $h(t) \in W^s(p, 0) \cap W^u(p, 0)$  for all  $t$ . In particular, the tangent spaces of stable and unstable manifolds evaluated at  $u = h(t)$  intersect in an at least one-dimensional subspace that contains  $\dot{h}(t)$ . The analytical interpretation of the tangent spaces of stable and unstable manifolds is as follows. Consider the variational equation

$$\dot{v} = f_u(h(t), 0)v \quad (2.2)$$

obtained by linearizing (2.1) about  $h(t)$ . We then have

$$\begin{aligned} T_{h(t)}W^s(p, 0) &= \{v(t); v(\cdot) \text{ satisfies (2.2) and } v(s) \rightarrow 0 \text{ as } s \rightarrow \infty\} \\ T_{h(t)}W^u(p, 0) &= \{v(t); v(\cdot) \text{ satisfies (2.2) and } v(s) \rightarrow 0 \text{ as } s \rightarrow -\infty\}. \end{aligned}$$

In addition, we may consider the adjoint variational equation defined by

$$\dot{w} = -f_u(h(t), 0)^*w \quad (2.3)$$

where  $A^*$  denotes the transpose of a matrix  $A$ . If  $\Phi(t, s)$  denotes the evolution of (2.2), then  $\Phi(s, t)^*$  is the solution operator of (2.3) (just differentiate the equation  $\Phi(t, s)\Phi(s, t) = \text{id}$ ). Using the identity

$$\frac{d}{dt} \langle v(t), w(t) \rangle = 0$$

which holds for any solutions  $v(t)$  of (2.2) and  $w(t)$  of (2.3), we conclude that

$$\begin{aligned} [T_{h(t)}W^s(p, 0)]^\perp &= \{w(t); w(\cdot) \text{ satisfies (2.3) and } w(s) \rightarrow 0 \text{ as } s \rightarrow \infty\} \\ [T_{h(t)}W^u(p, 0)]^\perp &= \{w(t); w(\cdot) \text{ satisfies (2.3) and } w(s) \rightarrow 0 \text{ as } s \rightarrow -\infty\}. \end{aligned}$$

In particular, we obtain

$$\begin{aligned} T_{h(t)}W^s(p, 0) \cap T_{h(t)}W^u(p, 0) \\ = \{v(t); v(\cdot) \text{ satisfies (2.2) and } v(s) \rightarrow 0 \text{ as } |s| \rightarrow \infty\} \end{aligned} \quad (2.4)$$

$$\begin{aligned} [T_{h(t)}W^s(p, 0) + T_{h(t)}W^u(p, 0)]^\perp \\ = \{w(t); w(\cdot) \text{ satisfies (2.3) and } w(s) \rightarrow 0 \text{ as } |s| \rightarrow \infty\}, \end{aligned} \quad (2.5)$$

and both spaces have the same dimension. We shall assume that these spaces are one-dimensional. In particular,  $\dot{h}(t)$  will span the intersection (2.4) of the stable and unstable tangent spaces, and we choose a solution  $\psi(t)$  of (2.3) that spans the complement (2.5) of their sum.

Next, we discuss the persistence of the homoclinic orbit  $h(t)$  if we change the parameter  $\mu$  near  $\mu = 0$ . Firstly, since we assumed that  $p$  is hyperbolic, there will be a unique

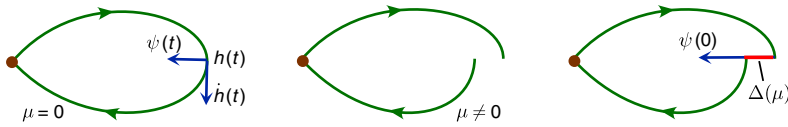


Fig. 2.1. Varying a parameter  $\mu$  may result in nonintersecting stable and unstable manifolds, so that the homoclinic orbit ceases to exist: the distance between the nonintersecting manifolds is measured by the quantity  $\Delta(\mu)$  defined in (2.6).

equilibrium  $p(\mu)$  near  $p$  for all  $\mu$  near zero, and  $p(\mu)$  and its stable and unstable manifolds will depend smoothly on  $\mu$ . However,  $W^s(p(\mu), \mu)$  and  $W^u(p(\mu), \mu)$  may no longer intersect; see Figure 2.1. To measure the distance between stable and unstable manifolds, we seek solutions near the homoclinic orbit in these manifolds that are closest to each other:

LEMMA 2.1. Assume that Hypothesis 2.1 is met and that  $T_{h(t)}W^s(p, 0) \cap T_{h(t)}W^u(p, 0) = \mathbb{R}\dot{h}(t)$ . For each  $\mu$  close to zero, there are unique orbits  $h^s(\cdot; \mu) \in W^s(p(\mu), \mu)$  and  $h^u(\cdot; \mu) \in W^u(p(\mu), \mu)$  of (2.1) with  $h^s(0; 0) = h^u(0; 0) = h(0)$  so that

$$h^u(0; \mu) - h^s(0; \mu) \in \mathbb{R}\psi(0) \quad \forall \mu.$$

The functions  $h^s(\cdot; \mu)$  and  $h^u(\cdot; \mu)$ , considered with values in  $C^0(\mathbb{R}^+, \mathbb{R}^n)$  and  $C^0(\mathbb{R}^-, \mathbb{R}^n)$ , respectively, are smooth in  $\mu$ .

PROOF. Let  $\Sigma$  be a small open ball centred at  $h(0)$  in the hyperplane  $h(0) + \dot{h}(0)^\perp$ . The sets  $\Sigma \cap [W^s(p(\mu), \mu) \oplus \mathbb{R}\psi(0)]$  and  $\Sigma \cap [W^u(p(\mu), \mu) \oplus \mathbb{R}\psi(0)]$  are manifolds that intersect transversely along a line for all sufficiently small  $\mu$ . This line intersects stable and unstable manifolds in unique points which are the initial conditions for the desired orbits.  $\square$

In particular, the stable and unstable manifolds of  $p(\mu)$  intersect near  $h(0)$  for  $\mu \approx 0$  if, and only if,  $\langle \psi(0), h^u(0; \mu) - h^s(0; \mu) \rangle = 0$ . It therefore makes sense to define the distance between stable and unstable manifolds near  $h(0)$  as

$$\Delta(\mu) := \langle \psi(0), h^u(0; \mu) - h^s(0; \mu) \rangle.$$

Using the variation-of-constant formula, it can be shown that the distance function  $\Delta(\mu)$  is given by

$$\begin{aligned} \Delta(\mu) &= \langle \psi(0), h^u(0; \mu) - h^s(0; \mu) \rangle \\ &= \left[ \int_{\mathbb{R}} \langle \psi(t), f_\mu(h(t), 0) \rangle dt \right] \mu + O(|\mu|^2) =: M\mu + O(|\mu|^2). \end{aligned} \quad (2.6)$$

The quantity  $M$  is commonly referred to as the Melnikov integral: the homoclinic orbit will not persist if  $M \neq 0$ . In the following hypothesis, we summarize the assumption made in Lemma 2.1 together with the condition that  $M \neq 0$ .

**HYPOTHESIS 2.2** (Nondegeneracy). *Consider the following nondegeneracy conditions:*

- (1) *Stable and unstable manifolds intersect as transversely as possible:*

$$T_{h(0)}W^s(p, 0) \cap T_{h(0)}W^u(p, 0) = \mathbb{R}\dot{h}(0).$$

- (2) *Stable and unstable manifolds unfold generically with respect to the parameter  $\mu_1$ :*

$$M := \int_{\mathbb{R}} \langle \psi(t), f_{\mu_1}(h(t), 0) \rangle dt \neq 0.$$

It turns out that, for most of the bifurcation scenarios that will be discussed below, more detailed information about the asymptotic behaviour of the homoclinic orbit  $h(t)$  than mere convergence will be needed. The asymptotics of  $h(t)$  for  $|t| \gg 1$  are, to a large extent, determined by the linearization  $\dot{v} = f_u(p, 0)v$  of (2.1) about the hyperbolic equilibrium  $p$ : in particular, the exponential decay rate of  $\|h(t) - p\|$  will be given by the real part of one of the eigenvalues of the Jacobian  $f_u(p, 0)$ . The eigenvalues closest to the imaginary axis will typically dominate the asymptotics as they give the slowest possible exponential rates: we therefore call them the leading stable and unstable eigenvalues of  $f_u(p, 0)$ , denoted by  $v^s$  and  $v^u$ , respectively. More precisely, denote the eigenvalues of  $f_u(p, 0)$  by  $v_j$  with  $j = 1, \dots, n$ , repeated with multiplicity and ordered by increasing real part so that

$$\operatorname{Re} v_1 \leq \operatorname{Re} v_2 \leq \dots \leq \operatorname{Re} v_k < 0 < \operatorname{Re} v_{k+1} \leq \dots \leq \operatorname{Re} v_{n-1} \leq \operatorname{Re} v_n. \quad (2.7)$$

The eigenvalues  $v_j$  with  $\operatorname{Re} v_j = \operatorname{Re} v_k$  are the leading stable eigenvalues, while the leading unstable eigenvalues  $v_j$  are those that satisfy  $\operatorname{Re} v_j = \operatorname{Re} v_{k+1}$ . We expect that the leading eigenvalues are simple and unique (up to complex conjugation), which leads to the following assumptions on the leading eigenvalues that are often imposed:

**HYPOTHESIS 2.3** (Leading Eigenvalues). *Consider the following eigenvalue conditions:*

- (1) *The unique leading unstable eigenvalue  $v^u$  is real and simple, and we have  $|\operatorname{Re} v^s| > v^u$ .*
- (2) *The leading stable and unstable eigenvalues are unique, real and simple.*
- (3) *The leading unstable eigenvalues is unique, real and simple. There are precisely two leading stable eigenvalues  $v^s$  and  $\overline{v^s}$ , and these are complex<sup>1</sup> and simple.*
- (4) *The leading stable and unstable eigenvalues are unique (up to complex conjugation) and simple.*

The quotient  $-v^s/v^u$  is often referred to as the saddle quantity. For future use, we shall also define real numbers  $\lambda^{ss}$  and  $\lambda^{uu}$  that separate the real parts of leading eigenvalues from those of the remaining strong eigenvalues so that

$$\operatorname{Re} v_1 \leq \dots \leq \operatorname{Re} v_{k-n_{ls}} < \lambda^{ss} < \operatorname{Re} v_{k-n_{ls}+1} = \dots = \operatorname{Re} v_k < 0 \quad (2.8)$$

and analogously for the unstable eigenvalues.

---

<sup>1</sup>Throughout the entire paper, we say that eigenvalues are complex if they are not real.

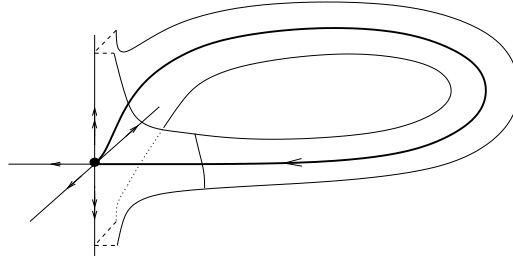


Fig. 2.2. Following the unstable manifold backwards along the homoclinic orbit, we expect that this manifold tends towards the strong unstable directions. This property is violated when the unstable manifold is in an inclination-flip configuration.

Having defined the leading stable and unstable eigenvalues, we return to the asymptotic behaviour of the homoclinic orbit and assume from now on that [Hypothesis 2.3\(4\)](#) is met. Using the solutions  $h^s(t; \mu)$  and  $h^u(t; \mu)$  from [Lemma 2.1](#), we define<sup>2</sup>

$$\begin{aligned} v^s(\mu) &:= \lim_{t \rightarrow \infty} e^{-v^s(\mu)t} [h^s(t; \mu) - p(\mu)] \\ v^u(\mu) &:= \lim_{t \rightarrow -\infty} e^{-v^u(\mu)t} [h^u(t; \mu) - p(\mu)], \end{aligned} \quad (2.9)$$

where  $v^s(\mu)$  and  $v^u(\mu)$  denote the leading eigenvalues of  $f_u(p(\mu), \mu)$ . It can be shown that these limits exist and are smooth in  $\mu$ . Furthermore,  $v^j(\mu)$  is a multiple of the eigenvector of  $f_u(p(\mu), \mu)$  associated with  $v^j(\mu)$  for  $j = s, u$ . In addition, there is a constant  $\epsilon > 0$  so that

$$h^s(t; \mu) = p(\mu) + e^{v^s(\mu)t} v^s(\mu) + O\left(e^{(\operatorname{Re} v^s(\mu) - \epsilon)t}\right), \quad t \rightarrow \infty$$

and analogously for  $h^u(t; \mu)$ . Thus, we can think of  $v^s(0)$  and  $v^u(0)$  as determining the effective dynamical components of  $h(t)$  in the leading eigendirections. In particular, we see that  $v^s(0) = 0$  if, and only if,  $h(t)$  lies in the strong stable manifold  $W^{ss}(p; 0)$  of the equilibrium  $p$ , which consists, by definition, of all solutions  $u(t)$  to [\(2.1\)](#) that satisfy  $\|u(t) - p\| = O(e^{\lambda^{ss}t})$ ; see [\[7\]](#).

Other important geometric properties associated with a homoclinic orbit  $h(t)$  are the inclination properties of the stable and unstable manifolds when they are transported along the homoclinic orbit. As illustrated in [Figure 2.2](#), we expect that the tangent space  $T_{h(t)} W^u(p, 0)$  converges as  $t \rightarrow \infty$  to the sum of  $T_p W^{uu}(p, 0)$  and the eigendirection associated with the leading stable eigenvalue<sup>3</sup>: in this case, we say that the unstable manifold is not in an inclination-flip configuration. The asymptotic behaviour as  $t \rightarrow \infty$  of the unstable manifold along the homoclinic orbit can be encoded similarly to the way we encoded in  $v^u(0)$  the property that  $h(t)$  should not lie in the strong unstable manifold

<sup>2</sup>If the eigenvalues are not real, then consider the limits in [\(2.9\)](#) using complex coordinates.

<sup>3</sup>We assume here that [Hypothesis 2.3\(3\)](#) is met so that the leading eigenvalues are real and simple.



of  $p$ : indeed, define

$$v_*^s := \lim_{t \rightarrow -\infty} e^{\nu^s t} \psi(t), \quad v_*^u := \lim_{t \rightarrow \infty} e^{\nu^u t} \psi(t),$$

then  $v_*^s$  and  $v_*^u$  are multiples of the leading stable and unstable eigenvectors of the adjoint Jacobian  $f_u(p, 0)^*$ . Furthermore, as desired,  $v_*^u \neq 0$  if, and only if, the unstable manifold is not in an inclination-flip configuration.

Generic homoclinic orbits do not lie in the strong stable or unstable manifolds, and the stable and unstable manifolds are not in inclination-flip configurations:

**HYPOTHESIS 2.4** (Inclination and Orbit Properties). *The following conditions exclude inclination-flip and orbit-flip configurations:*

- (1) *The stable manifold along the homoclinic orbit is not in an inclination-flip configuration, that is,  $v_*^s \neq 0$ .*
- (2) *The homoclinic orbit is not in an orbit-flip configuration within the stable manifold, that is, it does not lie in the strong-stable manifold  $W^{ss}(p, 0)$ :  $v^s \neq 0$ .*
- (3) *The unstable manifold along the homoclinic orbit is not in an inclination-flip configuration, that is,  $v_*^u \neq 0$ .*
- (4) *The homoclinic orbit is not in an orbit-flip configuration within the unstable manifold, that is,  $v^u \neq 0$ .*

Inclination and orbit flips have the following geometric interpretation. Assume again that the leading eigenvalues are real and simple. If **Hypothesis 2.4** is met, then

$$\mathcal{O} := \text{sign} [\langle v_*^s, v^s(0) \rangle \langle v_*^u, v^u(0) \rangle] \quad (2.10)$$

satisfies  $\mathcal{O} = \pm 1$ ; in other words, **Hypothesis 2.4** fails precisely when  $\mathcal{O} = 0$ . Geometrically,  $\mathcal{O}$  is an orientation index that encodes the orientability of the two-dimensional homoclinic centre manifold  $W_{\text{hom}}^c(\mu)$  that we shall discuss in Section 3.4; see **Figure 3.3** for an illustration. In particular, generically, the two-dimensional homoclinic centre manifold changes at an inclination or orbit flip from orientable to non-orientable, or vice versa. We say that a homoclinic orbit is orientable if  $\mathcal{O} = 1$  and call it non-orientable when  $\mathcal{O} = -1$ .

Alternatively, inclination-flip conditions, **Hypotheses 2.4**(1) and **2.4**(3), can also be stated in the following slightly different geometric terms. First, invariant manifold theory provides invariant manifolds  $W^{\text{ls},u}(p)$  near  $p$  whose tangent space at  $p$  consists of the generalized unstable eigenspace of  $f_u(p, 0)$  plus the leading stable eigendirections; similarly, there is an invariant manifold  $W^{\text{s},\text{lu}}(p)$  whose tangent space consists of the generalized stable eigenspace plus the leading stable eigendirections. The smoothness of these manifolds depends on spectral gap conditions: in general, they are only continuously differentiable. Furthermore, even though their tangent space at  $p$  is unique, the manifolds themselves are not uniquely defined. The tangent bundle of  $W^{\text{s},\text{lu}}(p)$  along the stable manifold is a smooth and uniquely defined vector bundle. Similarly, the tangent bundle of  $W^{\text{ls},u}(p)$  along the unstable manifold is a smooth and uniquely defined vector bundle. We now have the following characterization:

LEMMA 2.2. *The stable manifold along the homoclinic orbit is not in an inclination-flip configuration precisely if  $W^{\text{ls,u}}(p) \cap W^s(p)$  along the homoclinic orbit. The unstable manifold along the homoclinic orbit is not in an inclination-flip configuration precisely, if  $W^{\text{s,lu}}(p) \cap W^u(p)$  along the homoclinic orbit.*

If  $v_*^s$  or  $v_*^u$  vanish, then we would like to characterize how the degenerate inclination of stable or unstable manifolds is unfolded as  $\mu$  varies near zero. To this end, we need the following lemma whose proof is similar to that of Lemma 2.1.

LEMMA 2.3. *Assume that Hypotheses 2.1 and 2.2(1) are met. For  $j = s, u$ , there are unique solutions  $\psi^j(\cdot; \mu)$  of*

$$\dot{w} = -f_u(h^j(t; \mu), \mu)^* w, \quad j = s, u$$

with  $\|\psi^j(0; \mu)\| = 1$  and  $\psi^j(t; \mu) \perp T_{h^j(t; \mu)} W^j(p(\mu), \mu)$  so that  $\psi^u(0; \mu) - \psi^s(0; \mu) \in \mathbb{R}\dot{h}(0)$ . The functions  $\psi^s(\cdot; \mu)$  and  $\psi^u(0; \mu)$ , considered with values in  $C^0(\mathbb{R}^+, \mathbb{R}^n)$  and  $C^0(\mathbb{R}^-, \mathbb{R}^n)$ , respectively, are smooth in  $\mu$ .

We then define

$$v_*^s(\mu) := \lim_{t \rightarrow -\infty} e^{v^s(\mu)t} \psi^s(t; \mu), \quad v_*^u(\mu) := \lim_{t \rightarrow \infty} e^{v^u(\mu)t} \psi^u(t; \mu), \quad (2.11)$$

and it can again be shown that these limits exist and are smooth in  $\mu$ . We may now impose that derivatives of  $v_*^j(\mu)$  are nonzero at  $\mu = 0$  whenever  $v_*^j(0) = 0$  for  $j = s$  or  $j = u$ .

## 2.2. Homoclinic orbits to nonhyperbolic equilibria

Like hyperbolic equilibria, nonhyperbolic equilibria may admit homoclinic solutions. In systems with additional structure, such as reversibility or a Hamiltonian structure, this may be typical or at least of low codimension. We now discuss various geometric notions that we shall use later when we review bifurcations of homoclinic orbits that converge to nonhyperbolic equilibria.

From now on, let  $p$  be a nonhyperbolic equilibrium of  $\dot{u} = f(u, 0)$ , so that  $f_u(p, 0)$  has at least one eigenvalue on the imaginary axis. Centre manifolds and normal forms can then be used to study the local bifurcations near the equilibrium  $p$ , and we refer the reader to [7,410] for their properties and various examples.

First, consider the case where  $p$  is a saddle-node equilibrium: its linearization  $f_u(p, 0)$  has a simple real eigenvalue  $\nu = 0$  and no further eigenvalues on the imaginary axis. In generic systems, homoclinic orbits to a saddle-node equilibrium occur as a codimension-one phenomenon. Define  $v_c$  and  $w_c$  to be the right and left eigenvectors of the eigenvalue 0 of  $f_u(p, 0)$ .

HYPOTHESIS 2.5 (Codimension-one Saddle-node Bifurcation). *The following conditions define a generic saddle-node bifurcation:*

- (1) *The saddle-node equilibrium is not degenerate:  $\langle w_c, f_{uu}(p, 0)[v_c, v_c] \rangle \neq 0$ .*
- (2) *The unfolding is generic:  $\langle w_c, f_\mu(p, 0) \rangle \neq 0$ .*

If this hypothesis is met, then the vector field on the one-dimensional centre manifold can be brought into normal form  $\dot{x}^c = b(\mu) + a(\mu)(x^c)^2 + O(|x^c|^3)$ , where  $a(0) \neq 0$  and  $b_\mu(0) \neq 0$ . Associated with the saddle-node equilibrium  $p$  at  $\mu = 0$  is its stable manifold  $W^s(p)$ , which consists of orbits that converge towards  $p$  exponentially in  $t$  as  $t \rightarrow \infty$ , and its unstable manifold  $W^u(p)$ , which consists of orbits that converge towards  $p$  exponentially in  $t$  as  $t \rightarrow -\infty$ . Invariant-manifold theory [7] also gives the existence of a centre-stable manifold  $W^{cs}(p)$  and a centre-unstable manifold  $W^{cu}(p)$  so that a centre manifold is obtained as the transverse intersection  $W^c(p) = W^{cs}(p) \pitchfork W^{cu}(p)$ . The stable set  $\mathcal{M}^s(p)$  that consists of orbits which converge (not necessarily exponentially) to  $p$  as  $t \rightarrow \infty$  is a submanifold of  $W^{cs}(p)$  with boundary  $W^s(p)$ . Similarly, the unstable set  $\mathcal{M}^u(p)$  that consists of orbits converging to  $p$  for  $t \rightarrow -\infty$  is a manifold with boundary  $W^u(p)$  inside  $W^{cu}(p)$ .

As for hyperbolic equilibria, we can classify the hyperbolic part of the spectrum of  $f_u(p, 0)$  into different categories.

**HYPOTHESIS 2.6 (Leading Hyperbolic Eigenvalues).** *Consider the following conditions on the eigenvalues of  $f_u(p, 0)$  with nonzero real part:*

- (1) *The leading stable and unstable eigenvalues are unique, real and simple.*
- (2) *There are precisely two leading stable eigenvalues  $v^s$  and  $\bar{v}^s$ , and these are complex and simple.*
- (3) *The leading stable and unstable eigenvalues are unique (up to complex conjugation) and simple.*

A homoclinic orbit to  $p$  lies in the intersection  $\mathcal{M}^s(p) \cap \mathcal{M}^u(p)$ , and the following hypothesis excludes orbit-flip configurations.

**HYPOTHESIS 2.7 (Orbit Properties).** *The following conditions exclude orbit-flip configurations:*

- (1) *The homoclinic orbit is not in an orbit-flip configuration within the centre-stable manifold, that is, it does not lie in the stable manifold.*
- (2) *The homoclinic orbit is not in an orbit-flip configuration within the centre-unstable manifold, that is, it does not lie in the unstable manifold.*

Next, we discuss Hopf bifurcations of  $p$ , where the linearized vector field about the nonhyperbolic equilibrium  $p$  has complex conjugate eigenvalues on the imaginary axis when  $\mu = 0$ . Homoclinic orbits to a Hopf equilibrium occur, in generic systems, as a codimension-two phenomenon. This gives rise to the Shil'nikov–Hopf bifurcation, which we discuss in Section 5.1.10. We suppose that all eigenvalues of  $f_u(p, 0)$  are away from the imaginary axis except for two simple eigenvalues  $v^c, \bar{v}^c$  on the imaginary axis. Using the complex coordinate  $z$  on the two-dimensional local centre manifold, we can write the vector field on the centre manifold as  $\dot{z} = v^c(\mu)z + g(z, \bar{z}, \mu)$ . Recall that a smooth coordinate change transforms this equation into the normal form  $\dot{w} = v^c(\mu)w + c_1(\mu)|w|^2w + O(|w|^5)$ ; see [7,410].

**HYPOTHESIS 2.8 (Codimension-one Hopf Bifurcation).** *The following conditions define a generically unfolded supercritical Hopf bifurcation:*

- (1) *The Hopf equilibrium is not degenerate:  $c_1(0) \neq 0$ .*
- (2) *The unfolding is generic:  $v_\mu^c(0) \neq 0$ .*
- (3) *The Hopf bifurcation is supercritical:  $c_1(0) < 0$  and  $v_\mu^c(0) > 0$ .*

### 2.3. Heteroclinic cycles with hyperbolic equilibria

A heteroclinic cycle is an invariant set for (2.1) that consists of disjoint equilibria  $p_1, \dots, p_\ell$  and heteroclinic orbits  $h_1(t), \dots, h_\ell(t)$  that connect  $p_i$  to  $p_{i+1}$  so that

$$\lim_{t \rightarrow -\infty} h_i(t) = p_i, \quad \lim_{t \rightarrow \infty} h_i(t) = p_{i+1}$$

for  $i = 1, \dots, \ell$ , where  $p_{\ell+1} = p_1$ . A connected invariant set that can be written as a finite union of heteroclinic cycles (possibly including homoclinic loops) is called a polycycle.

As with homoclinic orbits we start with the key assumption that the equilibria do not undergo bifurcations, which is again guaranteed by a hyperbolicity condition.

**HYPOTHESIS 2.9 (Hyperbolicity).** *The equilibria  $p_i$  are hyperbolic at  $\mu = 0$  for  $1 \leq i \leq \ell$ .*

If the preceding assumption is met, then  $h_i(t) \in W^u(p_i) \cap W^s(p_{i+1})$  for each  $i$ . We now explore different possible configurations that depend on Morse indices and leading eigenvalues at the equilibria. As the Morse index  $\text{ind}(p_i) := \dim W^u(p_i)$  of  $p_i$  may differ from the index  $\text{ind}(p_{i+1})$  of  $p_{i+1}$ , heteroclinic orbits can occur robustly (or even in families) or may occur only if a sufficient number of parameters is varied. If  $\text{ind}(p_i) > \text{ind}(p_{i+1})$ , then transverse intersections of  $W^u(p_i)$  and  $W^s(p_{i+1})$  form a manifold of dimension  $\text{ind}(p_i) - \text{ind}(p_{i+1})$ . If  $\text{ind}(p_i) \leq \text{ind}(p_{i+1})$ , the codimension of  $h_i$  is typically equal to  $d = \text{ind}(p_{i+1}) - \text{ind}(p_i) + 1$ : for a family  $\dot{u} = f(u, \mu)$  of vector fields with  $\mu \in \mathbb{R}^d$ , transverse intersections of  $\bigcup_\mu (W^u(p_i(\mu), \mu), \mu)$  and  $\bigcup_\mu (W^s(p_{i+1}(\mu), \mu), \mu)$  in  $\mathbb{R}^n \times \mathbb{R}^d$  yield isolated heteroclinic orbits at isolated parameter values. Here,  $p_i(\mu)$  is the continuation of  $p_i$ , and  $W^u(p_i(\mu), \mu)$  is the unstable manifold of  $p_i(\mu)$  for  $\dot{u} = f(u, \mu)$ .

Following Section 2.1, consider the adjoint variational equation

$$\dot{w} = -f_u(h_i(t), 0)^* w. \quad (2.12)$$

Suppose  $T_{h_i(0)} W^u(p_i, 0) + T_{h_i(0)} W^s(p_{i+1}, 0)$  has codimension  $d$ , then there are  $d$  linearly independent bounded solutions  $\psi_i^1(t), \dots, \psi_i^d(t)$  to (2.12). We can now formulate nondegeneracy conditions akin to Hypothesis 2.2.

**HYPOTHESIS 2.10 (Nondegeneracy).** *Consider the following nondegeneracy conditions on  $h_i(t)$ :*

- (1) *The heteroclinic orbit  $h_i$  is of codimension  $d_i \geq 0$ :  $\text{ind}(p_{i+1}) - \text{ind}(p_i) + 1 = d_i$ , and the codimension of  $T_{h_i(0)} W^u(p_i, 0) + T_{h_i(0)} W^s(p_{i+1}, 0)$  is  $d_i$ .*
- (2) *The heteroclinic orbit  $h_i$  is of codimension  $d_i$ , and stable and unstable manifolds unfold generically along  $h_i$  with respect to the parameter  $\mu = (\mu_1, \dots, \mu_{d_i}) \in \mathbb{R}^{d_i}$*

so that the matrix  $M \in \mathbb{R}^{d_i \times d_i}$  with entries

$$M_{kl} = \int_{\mathbb{R}} \langle \psi_i^k(t), f_{\mu_i}(h_i(t), 0) \rangle dt$$

has full rank.

For  $i = 1, \dots, \ell$ , we define the cross sections  $\Sigma_i$  via

$$\Sigma_i = h_i(0) + Y_i, \quad Y_i = f(h_i(0), 0)^\perp. \quad (2.13)$$

The cross section  $\Sigma_i$  is a hyperplane that intersects the orbit  $h_i(t)$  transversally at  $h_i(0)$ . Let

$$Z_i = (T_{h_i(0)} W^u(p_i, 0) + T_{h_i(0)} W^s(p_{i+1}, 0))^\perp \quad (2.14)$$

be the subspace of  $Y_i$  spanned by  $\psi_i^1(0), \dots, \psi_i^{d_i}(0)$  and note that  $h_i(0) + Z_i \subset \Sigma_i$ . The argument used to prove [Lemma 2.1](#) also gives the following corresponding lemma for heteroclinic orbits.

**LEMMA 2.4.** *Suppose  $h_i$  is a heteroclinic orbit of codimension one (in particular,  $\dim Z_i = 1$ ). For each  $\mu$  close to 0, there are unique orbits  $h_i^s(\cdot; \mu) \in W^s(p_{i+1}(\mu), \mu)$  and  $h_i^u(\cdot; \mu) \in W^u(p_i(\mu), \mu)$  such that*

$$h_i^u(0; \mu) - h_i^s(0; \mu) \in Z_i, \quad \forall \mu.$$

Next, we state conditions on the leading eigenvalues.

**HYPOTHESIS 2.11 (Leading Eigenvalues).** *Consider the following eigenvalue conditions:*

- (1) *The unique leading unstable eigenvalue  $v_i^u$  at  $p_i$  is real and simple, and we have  $|\operatorname{Re} v_i^s| > v_i^u$ .*
- (2) *The leading stable and unstable eigenvalues at  $p_i$  are unique, real and simple.*
- (3) *The leading stable eigenvalue of  $f_u(p_{i+1}, 0)$  and unstable eigenvalues of  $f_u(p_i, 0)$  are unique (up to complex conjugation) and simple.*

Assume that [Hypotheses 2.11\(3\)](#) and [2.10\(1\)](#) are met with  $d_i = 1$ . As in [Section 2.1](#), it can be shown that

$$v_i^s(\mu) := \lim_{t \rightarrow \infty} e^{-v_{i+1}^s(\mu)t} [h_i^s(t; \mu) - p_{i+1}(\mu)] \quad (2.15)$$

$$v_i^u(\mu) := \lim_{t \rightarrow -\infty} e^{-v_i^u(\mu)t} [h_i^u(t; \mu) - p_i(\mu)]$$

exist and are smooth in  $\mu$ , where  $v_j^s(\mu)$  and  $v_j^u(\mu)$  denote the leading eigenvalues of  $f_u(p_j(\mu), \mu)$ . Likewise, we define

$$v_{i,*}^s := \lim_{t \rightarrow -\infty} e^{v_i^s(0)t} \psi_i(t), \quad v_{i,*}^u := \lim_{t \rightarrow \infty} e^{v_{i+1}^u(0)t} \psi_i(t), \quad (2.16)$$

then  $v_{i,*}^s$  and  $v_{i,*}^u$  are multiples of the leading stable and unstable eigenvectors of the adjoint linearizations  $f_u(p_i, 0)^*$  and  $f_u(p_{i+1}, 0)^*$ , respectively.

**HYPOTHESIS 2.12** (Inclination and Orbit Properties). *The following conditions exclude inclination-flip and orbit-flip configurations along a codimension-one heteroclinic orbit  $h_i$ :*

- (1) *The stable manifold  $W^s(p_{i+1}, 0)$  along  $h_i$  is not in an inclination-flip configuration, that is,  $v_{i,*}^s \neq 0$ .*
- (2)  *$h_i$  is not in an orbit-flip configuration within the stable manifold so that it does not lie in the strong-stable manifold  $W^{ss}(p_{i+1}, 0)$ :  $v_i^s(0) \neq 0$ .*
- (3) *The unstable manifold  $W^u(p_i, 0)$  along  $h_i$  is not in an inclination-flip configuration, that is,  $v_{i,*}^u \neq 0$ .*
- (4)  *$h_i$  is not in an orbit-flip configuration within the unstable manifold, that is,  $v_i^u(0) \neq 0$ .*

There is a geometric description of the inclination-flip property akin to [Lemma 2.2](#). Recall the definition of the invariant manifolds  $W^{ls,u}(p_i)$  and  $W^{s,lu}(p_{i+1})$  from [Section 2.1](#).

**LEMMA 2.5.** *The stable manifold  $W^s(p_{i+1})$  along  $h_i$  is not in an inclination-flip configuration precisely if  $W^{ls,u}(p_i) \pitchfork W^s(p_{i+1})$ . The unstable manifold  $W^u(p_i)$  along  $h_i$  is not in an inclination-flip configuration precisely if  $W^{s,lu}(p_{i+1}) \pitchfork W^u(p_i)$ .*

### 3. Analytical and geometric approaches

The goal of homoclinic bifurcation theory is to investigate the recurrent dynamics near a homoclinic orbit  $h(t)$ . In other words, we are interested in finding all orbits that stay in a fixed tubular neighbourhood of a given homoclinic orbit or heteroclinic cycle for all times.

A natural way for approaching this problem is to use Poincaré or first-return maps. Denote by  $\Sigma$  a cross section placed at  $h(0)$  which, by definition, means that  $\Sigma$  is the ball of radius  $\epsilon > 0$  centred at  $h(0)$  in the hyperplane  $h(0) + \dot{h}(0)^\perp$ :

$$\Sigma = B_\epsilon(h(0)) \subset h(0) + \dot{h}(0)^\perp, \quad (3.1)$$

where  $\epsilon > 0$  is chosen to be so small that the flow is transverse to  $\Sigma$ . Starting with an initial condition  $u_0$  in  $\Sigma$ , we then follow  $u(t)$  until it hits  $\Sigma$  again, say at time  $t = T$ , and define the first-return map  $\Pi$  via  $\Pi(u_0) := u(T) \in \Sigma$ . The main issue is that a solution that starts near  $h(0)$  may not return to a neighbourhood of  $h(0)$ ; see [Figure 3.1](#). Thus, in general, the domain of the Poincaré map  $\Pi$  does not contain an open neighbourhood of  $h(0)$  in  $\Sigma$ . Furthermore, any solution that does return will spend a very long time near the equilibrium, thus spoiling most finite-time error estimates for  $\Pi$  from standard variation-of-constant formulae.

Several techniques have been developed to overcome these difficulties. We distinguish, to some extent artificially, two different approaches that treat homoclinic bifurcations from

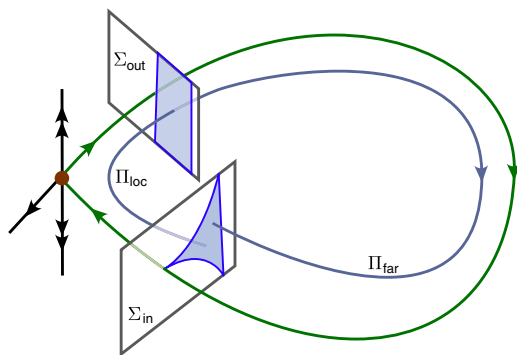


Fig. 3.1. The domain of the first-return map on a cross section does not contain an open neighbourhood of  $h(0)$  (one often has wedge shaped domains as shown here). For computational purposes, it is useful to consider two cross sections  $\Sigma_{in}$  and  $\Sigma_{out}$  and write the first-return map  $\Pi$  as a composition of transition maps between the cross sections.

somewhat different viewpoints: The first approach computes the Poincaré map by writing it as a composition of a local transition map near the equilibrium with a global transition map; see Figure 3.1. The main difficulty here is to get expansions of the local transition map. The second approach is to seek orbits that stay near the homoclinic orbit as a solution to some abstract functional-analytic system using Lyapunov–Schmidt reduction. Both of these methods can be used in conjunction with a geometric reduction to a low-dimensional invariant homoclinic centre manifold that contains all recurrent dynamics. Each of these approaches has its own advantages and disadvantages that we shall comment on in the following when we discuss them in more detail.

### 3.1. Normal forms and linearizability

Given two cross sections  $\Sigma_{in}$  and  $\Sigma_{out}$  that are transverse to, respectively, the local stable and local unstable manifolds of  $p$ , the local transition map  $\Pi_{loc}$  is the first-return map

$$\Pi_{loc} : \Sigma_{in} \longrightarrow \Sigma_{out}. \quad (3.2)$$

Understanding  $\Pi_{loc}$  requires solving

$$\dot{u} = f(u, \mu) = f_u(p, \mu)(u - p) + O(\|u - p\|^2) \quad (3.3)$$

for  $u$  close to  $p$ . To calculate  $\Pi_{loc}$ , it is often advantageous to simplify the vector field on the right-hand side of (3.3). Of course, the best possible outcome is that the nonlinear terms can simply be transformed away which renders the vector field linear.

The Hartman–Grobman theorem states that there is a coordinate change, which is continuous in  $(u, \mu)$ , near each hyperbolic equilibrium  $p$  that transforms (3.3) into the

linear system

$$\dot{v} = f_u(p, \mu)v. \quad (3.4)$$

Using such coordinates, the return map on a cross section is a homeomorphism but it is not clear how the expansions in  $u$  can be derived that are needed for a bifurcation study. Belitskii [37] derived eigenvalue conditions which ensure the existence of a continuously differentiable coordinate change that linearizes a fixed flow. This linearization theorem can be applied to obtain statements on the stability or the existence of hyperbolic sets near homoclinic solutions.

**THEOREM 3.1 ([37]).** *Consider the equation  $\dot{u} = f(u)$  on  $\mathbb{R}^n$  near a hyperbolic equilibrium  $p$  and assume that  $f$  is of class  $C^2$ . With the eigenvalue ordering given in (2.7), if*

$$\operatorname{Re} v_i \neq \operatorname{Re} v_{j_1} + \operatorname{Re} v_{j_2} \quad (3.5)$$

*for each  $i$  and all  $1 \leq j_1 \leq \dim W^s(p) < j_2 \leq n$ , then there is a local coordinate transformation of class  $C^1$  that transforms the ODE into its linearization  $\dot{v} = f_u(p)v$  near  $u = p$ .*

For the study of bifurcations that involve periodic orbits, parameter-dependent versions and higher degrees of differentiability are needed. The next theorem, a parameter-dependent version of Sternberg's linearization result, gives conditions under which (3.3) can be transformed into (3.4) by an appropriate smooth coordinate transformation (see also Section 3.6.3 for a complementary linearization result). More generally, Takens [388] constructs partial linearizations near equilibria with centre directions.

**THEOREM 3.2 ([353,385,388]).** *Assume that  $f(u, \mu)$  is  $C^\infty$  in  $(u, \mu)$ . Fix any  $\ell \geq 1$ , then there exist numbers  $N = N(\ell, f_u(0, 0)) \geq 1$  and  $\epsilon > 0$  with the following property: if*

$$v_i \neq \sum_{j=1}^n N_j v_j \quad (3.6)$$

*for each  $i$  and all natural numbers  $N_j$  with  $2 \leq \sum_{j=1}^n N_j \leq N$ , then we can  $C^\ell$ -linearize (3.3) in  $B_\epsilon(p)$  for  $|\mu| < \epsilon$ , i.e. there is a coordinate transformation of class  $C^\ell$  in  $u \in B_\epsilon(p)$  and  $|\mu| < \epsilon$  that transforms (3.3) into (3.4).*

If the system can be linearized, we can compute  $\Pi_{\text{loc}}$  explicitly once the transition time from  $\Sigma_{\text{in}}$  to  $\Sigma_{\text{out}}$  has been computed.

If the non-resonance condition (3.6) is not met, the vector field may still be transformed into normal form [86], and we refer the reader to [208] for an exposition of results on finitely smooth normal forms for families of vector fields and to [54] for analytic local normal forms for families of ODEs. Structure preserving normal forms (e.g. in the class of equivariant, conservative or reversible ODEs) are another important topic: we will not



discuss this here but refer to [38] for Sternberg's theorem in an equivariant setting and to [53] for a general approach.

Useful expansions for  $\Pi_{\text{loc}}$  may be hard to obtain if the normal form is not linear though problems involving one-dimensional unstable separatrices are often tractable. Another approach in this situation is to use Shil'nikov variables which we shall discuss next.

### 3.2. Shil'nikov variables

Shil'nikov variables were introduced by Shil'nikov in 1968 to compute the local transition map near equilibria to leading order. Instead of solving an initial-value problem, solutions near the equilibrium are found using an appropriate boundary-value problem. The analysis of the resulting integral formulae leads to asymptotic expansions for the solutions. We concentrate on hyperbolic equilibria and refer the reader to [106] for results in the non-hyperbolic case. We remark that, in the case of a hyperbolic equilibrium with one-dimensional unstable directions, the approach leads to asymptotic expansions for solutions of an initial-value problem. Further information on Shil'nikov variables can be found, for instance, in the books [375,376] where considerable space is devoted to their analysis and use.

Assume that  $p = 0$  is a hyperbolic equilibrium for all  $\mu$ . We may also assume that the stable and unstable eigenspaces of  $f_u(0, \mu)$  do not depend on  $\mu$ . We denote these spaces by  $E_0^s$  and  $E_0^u$ , respectively, and choose local coordinates  $u = (x, y) \in E_0^s \oplus E_0^u$ . In these coordinates, (3.3) becomes

$$\begin{aligned}\dot{x} &= A^s(\mu)x + g^s(x, y; \mu), \quad \text{spec}(A^s(\mu)) = \text{spec}(f_u(0, \mu)) \cap \{\text{Re } \nu < 0\}, \\ \dot{y} &= A^u(\mu)y + g^u(x, y; \mu), \quad \text{spec}(A^u(\mu)) = \text{spec}(f_u(0, \mu)) \cap \{\text{Re } \nu > 0\}.\end{aligned}\quad (3.7)$$

For each fixed  $(x_0, y_1)$  close to zero and each  $\tau \gg 1$ , we seek solutions  $(x, y)(t)$  of (3.7) that satisfy

$$x(0) = x_0, \quad y(\tau) = y_1. \quad (3.8)$$

We proceed, as follows, to construct these solutions. First, we separate the eigenvalues of  $f_u(0, 0)$  into leading and strong directions, and choose numbers  $\lambda^{\text{ss}}$  and  $\lambda^{\text{uu}}$  for  $\mu = 0$  as in (2.8). Next, we write

$$x = (x^{\text{ls}}, x^{\text{ss}}) \in E_0^s, \quad y = (y^{\text{lu}}, y^{\text{uu}}) \in E_0^u,$$

where  $x^{\text{ls}}$  and  $x^{\text{ss}}$  lie in the generalized eigenspaces of  $f_u(0, \mu)$  that belong to the stable eigenvalues of  $f_u(0, \mu)$  whose real parts are, respectively, larger and smaller than  $\lambda^{\text{ss}}$ . Similarly,  $y^{\text{lu}}$  and  $y^{\text{uu}}$  lie in the generalized eigenspaces of  $f_u(0, \mu)$  that belong to the unstable eigenvalues of  $f_u(0, \mu)$  whose real parts are, respectively, smaller and larger than  $\lambda^{\text{uu}}$ . Note that the leading stable and unstable eigenvalues may acquire slightly different real parts upon changing  $\mu$ . An appropriate coordinate transformation [110,375] brings (3.7) into the following normal form. More detailed normal forms may be obtained by similar methods if appropriate spectral conditions are met.

PROPOSITION 3.1. *A smooth coordinate change brings the ODE near  $p$  into the form*

$$\begin{pmatrix} \dot{x}^{\text{ls}} \\ \dot{x}^{\text{ss}} \\ \dot{y}^{\text{lu}} \\ \dot{y}^{\text{uu}} \end{pmatrix} = \begin{pmatrix} A^{\text{ls}}(\mu) & 0 & 0 & 0 \\ 0 & A^{\text{ss}}(\mu) & 0 & 0 \\ 0 & 0 & A^{\text{lu}}(\mu) & 0 \\ 0 & 0 & 0 & A^{\text{uu}}(\mu) \end{pmatrix} \begin{pmatrix} x^{\text{ls}} \\ x^{\text{ss}} \\ y^{\text{lu}} \\ y^{\text{uu}} \end{pmatrix} + \begin{pmatrix} O(|x^{\text{ls}}|^2 + |x^{\text{ss}}||y|) \\ O(|x^{\text{ls}}|^2 + |x^{\text{ss}}|(|x| + |y|)) \\ O(|x^{\text{lu}}|^2 + |x^{\text{uu}}||x|) \\ O(|y^{\text{lu}}|^2 + |y^{\text{uu}}|(|x| + |y|)) \end{pmatrix}.$$

PROOF. We may choose smooth coordinates  $x = (x^{\text{ss}}, x^{\text{ls}}, y^{\text{lu}}, y^{\text{uu}})$  near  $p$  so that

$$\begin{aligned} W^s(p) &= \{y = 0\}, & W^u(p) &= \{x = 0\}, \\ W^{s,\text{lu}}(p) \cap W^s(p) &= \{y^{\text{uu}} = 0\}, & W^{\text{ls},u}(p) \cap W^u(p) &= \{x^{\text{ss}} = 0\}, \end{aligned}$$

where the notation  $W \bar{\cap}_q V$  means that  $W$  is tangent to  $V$  at  $q$ . This brings (3.7) into the form

$$\begin{pmatrix} \dot{x}^{\text{ls}} \\ \dot{x}^{\text{ss}} \\ \dot{y}^{\text{lu}} \\ \dot{y}^{\text{uu}} \end{pmatrix} = \begin{pmatrix} A^{\text{ls}}(\mu) & 0 & 0 & 0 \\ 0 & A^{\text{ss}}(\mu) & 0 & 0 \\ 0 & 0 & A^{\text{lu}}(\mu) & 0 \\ 0 & 0 & 0 & A^{\text{uu}}(\mu) \end{pmatrix} \begin{pmatrix} x^{\text{ls}} \\ x^{\text{ss}} \\ y^{\text{lu}} \\ y^{\text{uu}} \end{pmatrix} + \begin{pmatrix} O(|x^s|(|x| + |y|)) \\ O(|x^{\text{ls}}|^2 + |x^{\text{ss}}|(|x| + |y|)) \\ O(|y^u|(|x| + |y|)) \\ O(|y^{\text{lu}}|^2 + |y^{\text{uu}}|(|x| + |y|)) \end{pmatrix}.$$

The remaining coordinate changes are described in detail in [303], and we give here only a brief overview. A polynomial coordinate change removes quadratic terms  $x^{\text{ls}}x^{\text{lu}}$  from the differential equations for  $x^{\text{ls}}$ . Consider next a change of coordinates of the form

$$\tilde{x}^{\text{ls}} = x^{\text{ls}} + p^{\text{ls}}(y)x^{\text{ls}}, \quad \tilde{x}^{\text{ss}} = x^{\text{ss}}, \quad \tilde{y}^{\text{lu}} = y^{\text{lu}}, \quad \tilde{y}^{\text{uu}} = y^{\text{uu}} \quad (3.9)$$

for a function  $p^{\text{ls}}$  that vanishes along  $y = 0$ . Write the differential equation for  $x^{\text{ls}}$  in the new coordinates (skipping the tildes) as

$$\dot{x}^{\text{ls}} = A^{\text{ls}}(\mu)x^{\text{ls}} + P^{\text{ls}}(x, y)x^{\text{ls}} + g^{\text{ls}}(x, y)x^{\text{ss}}.$$

Along the unstable manifold  $x = 0$  we find

$$P^{\text{ls}}(0, y) = \dot{p}^{\text{ls}} + p^{\text{ls}}A^{\text{ls}}(\mu) - A^{\text{ls}}(\mu)p^{\text{ls}},$$

where the higher-order terms are of at least quadratic order. Consider  $\dot{p}^{\text{ls}}$  as a variable and

construct the local unstable manifold tangent to  $\{p^{\text{ls}} = 0\}$  at the origin for the resulting differential equations

$$\begin{aligned}\dot{p}^{\text{ls}} &= A^{\text{ls}}(\mu)p^{\text{ls}} - p^{\text{ls}}A^{\text{ls}}(\mu) + \text{h.o.t.}, \\ \dot{y}^{\text{lu}} &= A^{\text{lu}}(\mu)y^{\text{lu}} + \text{h.o.t.}, \\ \dot{y}^{\text{uu}} &= A^{\text{uu}}(\mu)y^{\text{uu}} + \text{h.o.t.}\end{aligned}$$

along the unstable manifold  $\{x = 0\}$ . The resulting coordinate change will transform  $f^{\text{ls}}$  to a map with expansion

$$f^{\text{ls}}(x) = O(|x^{\text{ls}}|^2 + |x^{\text{ss}}|(|x| + |y|)).$$

As the coordinate change leaves  $y^{\text{lu}}$  unaltered, a similar coordinate change for  $f^{\text{lu}}$  can be performed. We have now achieved the following form:

$$\begin{pmatrix} \dot{x}^{\text{ls}} \\ \dot{x}^{\text{ss}} \\ \dot{y}^{\text{lu}} \\ \dot{y}^{\text{uu}} \end{pmatrix} = \begin{pmatrix} A^{\text{ls}}(\mu) & 0 & 0 & 0 \\ 0 & A^{\text{ss}}(\mu) & 0 & 0 \\ 0 & 0 & A^{\text{lu}}(\mu) & 0 \\ 0 & 0 & 0 & A^{\text{uu}}(\mu) \end{pmatrix} \begin{pmatrix} x^{\text{ls}} \\ x^{\text{ss}} \\ y^{\text{lu}} \\ y^{\text{uu}} \end{pmatrix} + \begin{pmatrix} O(|x^{\text{ls}}|^2 + |x^{\text{ss}}|(|x| + |y|)) \\ O(|x^{\text{ls}}|^2 + |x^{\text{ss}}|(|x| + |y|)) \\ O(|y^{\text{lu}}|^2 + |y^{\text{uu}}|(|x| + |y|)) \\ O(|y^{\text{lu}}|^2 + |y^{\text{uu}}|(|x| + |y|)) \end{pmatrix}.$$

Within the stable manifold, there is a smooth strong stable foliation, which can be transformed into an affine foliation with leaves  $\{x^{\text{ss}} = \text{constant}\}$ . The differential equation for  $x^{\text{ls}}$  in the set  $\{y = 0\}$  depends only on  $x^{\text{ls}}$  and no longer on  $x^{\text{ss}}$ . Since eigenvalues of  $A^{\text{ls}}(\mu)$  have the same real part, one can smoothly linearize this differential equation.  $\square$

In the coordinates of the preceding proposition, there exists an  $\epsilon > 0$  such that the boundary-value problem (3.8) has a unique solution  $(x, y)(t)$  for each  $(x_0, y_1)$ ,  $\mu$  and  $\tau$  that satisfy  $|x_0| + |y_0| < \epsilon$ ,  $|\mu| < \epsilon$  and  $\tau > 1/\epsilon$ . This solution depends smoothly on the data  $(x_0, y_1, \mu, \tau)$ .

**PROPOSITION 3.2.** *Assuming the normal form from Proposition 3.1, there is an  $\eta > 0$  so that the solution  $(x, y)(t)$  admits the expansion*

$$\begin{aligned}x^{\text{ls}}(t) &= e^{A^{\text{ls}}(\mu)t} [x_0^{\text{ls}} + O(e^{-\eta t})], & x^{\text{ss}}(t) &= O(e^{(\text{Re } v^{\text{s}} - \eta)t}), \\ y^{\text{lu}}(t) &= e^{A^{\text{lu}}(\mu)(t-\tau)} [y_0^{\text{lu}} + O(e^{\eta(t-\tau)})], & y^{\text{uu}}(t) &= O(e^{(\text{Re } v^{\text{u}} + \eta)(t-\tau)}).\end{aligned}\tag{3.10}$$

These asymptotic expansions are derived by considering integral formulae for solutions obtained using variation of constants. Similar estimates also hold for the derivatives of the solution with respect to the data. Note the similarity with the expansions for

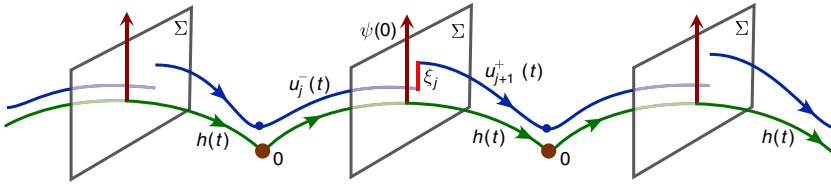


Fig. 3.2. Orbit pieces in Lin's method with jumps in a specified direction in the cross section.

linear differential equations. We refer the reader to [105,110,366,375] for proofs and generalizations of this result.

### 3.3. Lin's method

In this section, we discuss a functional-analytic approach due to Lin [261]. For clarity of exposition, we assume that the equilibrium  $p$  of

$$\dot{u} = f(u, \mu) \quad (3.11)$$

is hyperbolic and that the homoclinic orbit  $h(t)$  satisfies [Hypothesis 2.1\(1\)](#). We pick a nontrivial bounded solution  $\psi(t)$  of the adjoint variational equation (2.3) about  $h(t)$  and the cross section  $\Sigma$  from (3.1). The idea of Lin's method is to construct a sequence  $u_j$  of solutions to (3.11) that begin and end in  $\Sigma$  after spending a given number of time units near the homoclinic orbit. The key feature of these solutions is that the difference between the end point of the  $j$ th solution and the initial condition for the  $(j + 1)$ th solution lies in  $\mathbb{R}\psi(0)$ : In particular, the individual solutions can be spliced together to form a solution of (3.11) if, and only if, the jumps in  $\mathbb{R}\psi(0)$  vanish for all  $j$ . This is further illustrated in [Figure 3.2](#).

In detail, it has been shown in [261,338] that there are constants  $0 < \epsilon \ll 1$  and  $T_* \gg 1$  such that the boundary-value problem

$$\begin{aligned} \dot{u}_j^- &= f(u_j^-, \mu) & t \in (-T_j, 0) \\ \dot{u}_j^+ &= f(u_j^+, \mu) & t \in (0, T_j) \\ u_j^\pm(0) &\in \Sigma \\ u_j^-(-T_j) &= u_j^+(T_j) \\ \|u_j^\pm(t) - h(t)\| &< \epsilon & t \in [-T_j, 0] \text{ and } [0, T_j], \text{ respectively} \\ u_j^-(0) - u_{j+1}^+(0) &\in \mathbb{R}\psi(0) \end{aligned} \quad (3.12)$$

with  $j \in \mathbb{Z}$  has a unique solution  $\{u_j\}_{j \in \mathbb{Z}}$  for given data  $\{T_j\}_{j \in \mathbb{Z}}$  and  $\mu$  with  $T_j > T_*$  and  $|\mu| < \epsilon$ , and the solution is smooth in those data. In particular, if the bifurcation functions

$$\xi_j := \langle \psi(0), u_j^-(0) - u_{j+1}^+(0) \rangle$$

all vanish, then the concatenation of the solutions  $u_j$  gives a solution  $u$  of (3.11) that follows the homoclinic orbit  $h(t)$  for all times. Using the notation introduced in Section 2, the bifurcation functions  $\xi_j$  admit the expansion

$$\begin{aligned} \xi_j = & \langle \psi^u(-T_j, \mu), h^s(T_j, 0) - p(\mu) \rangle - \langle \psi^s(T_{j+1}, \mu), h^u(-T_{j+1}, 0) - p(\mu) \rangle \\ & + \Delta(\mu) + O\left(e^{-(2\min(|v^s|, v^u) + \eta) \min_{k \in \mathbb{Z}} T_k}\right) \end{aligned} \quad (3.13)$$

for some  $\eta > 0$ . We refer the reader to [6, Section 5.1.1] for a geometric explanation of the quantities that appear in the above expression. The error estimate in (3.13) can be greatly improved, and we refer to [261, 338] for details.

EXAMPLE. If Hypothesis 2.3(1) is met, then we can use (2.6), (2.9) and (2.11) to simplify the expression (3.13) to get the bifurcation equations

$$\begin{aligned} \langle v_*^s(\mu), v^s(\mu) \rangle e^{2v^s T_j} - \langle v_*^u(\mu), v^u(\mu) \rangle e^{-2v^u T_{j+1}} + M\mu \\ + O\left(e^{-(2\min(|v^s|, v^u) + \eta) \min_{k \in \mathbb{Z}} T_k} + |\mu|^2\right) = 0. \end{aligned} \quad (3.14)$$

### 3.4. Homoclinic centre manifolds

Instead of directly investigating the full  $n$ -dimensional ODE

$$\dot{u} = f(u, \mu) \quad (3.15)$$

near a given homoclinic orbit  $h(t)$ , it may be desirable to first reduce the dimensionality of the system by constructing a locally invariant, normally hyperbolic, manifold that contains the homoclinic orbit and all solutions staying close to it for all times. We refer to such a manifold as a *homoclinic centre manifold*: normal hyperbolicity implies robustness under parameter perturbations, while the property that it contains all recurrent dynamics shows that it plays indeed a role similar to that of centre manifolds in local bifurcation theory. The theorem stated below asserts that linear normal hyperbolicity along the homoclinic orbit implies the existence of a homoclinic centre manifold.

**HYPOTHESIS 3.1 (Linear Normal Hyperbolicity).** *Assume that  $h(t)$  is a homoclinic orbit of (3.15) for  $\mu = 0$  that converges to the hyperbolic equilibrium  $p$ . Suppose further that*

$$\begin{aligned} \text{spec}(f_u(p, 0)) = \sigma^s \cup \sigma^c \cup \sigma^u, \quad \max \text{Re } \sigma^s < \min \text{Re } \sigma^c, \\ \max \text{Re } \sigma^c < \min \text{Re } \sigma^u, \end{aligned}$$

and denote by  $E_p^s \oplus E_p^c \oplus E_p^u$  the associated decomposition of  $\mathbb{R}^n$  into generalized spectral eigenspaces of  $f_u(p, 0)$ . In this setting, we assume that there are subspaces  $E^j(t)$  of  $\mathbb{R}^n$  for  $j = s, c, u$ , defined and continuous in  $t \in \mathbb{R}$ , so that  $E^s(t) \oplus E^c(t) \oplus E^u(t) = \mathbb{R}^n$  for

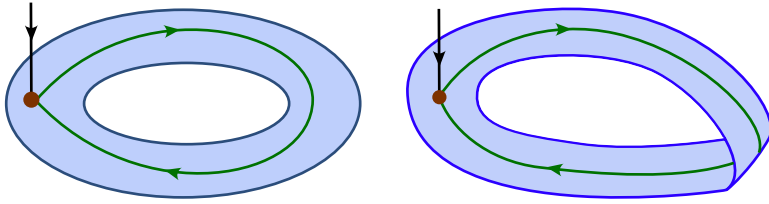


Fig. 3.3. If the bundle  $E^c$  is two-dimensional, a homoclinic centre manifold is a two-dimensional surface that is diffeomorphic to either an annulus [left] or a Möbius band [right]. The associated orientation index  $\mathcal{O}$  defined in (2.10) is  $\mathcal{O} = 1$  [left] or  $\mathcal{O} = -1$  [right].

all  $t \in \mathbb{R}$  and  $E^j(t) \rightarrow E_p^j$  as  $|t| \rightarrow \infty$  for each  $j$ , such that the evolution  $\Phi(t, s)$  of

$$\dot{v} = f_u(h(t), 0)v$$

maps  $E^j(s)$  into  $E^j(t)$  for all  $t, s \in \mathbb{R}$  and each  $j$ . Lastly, we assume that  $\dot{h}(t) \in E^c(t)$ .

The following result has been proved in [187,338,342], see also [193,333,360,400].

**THEOREM 3.3.** Assume that *Hypothesis 3.1* is met. Pick any integer  $l \geq 1$  and a number  $\alpha \in (0, 1)$  so that

$$l + \alpha < \min \left\{ \frac{\min \operatorname{Re} \sigma^c}{\max \operatorname{Re} \sigma^s}, \frac{\min \operatorname{Re} \sigma^u}{\max \operatorname{Re} \sigma^c} \right\},$$

then there are a constant  $\epsilon > 0$  and a locally invariant, normally hyperbolic, homoclinic centre manifold  $W_{\text{hom}}^c(\mu)$  associated with  $h(t)$  and defined for  $|\mu| < \epsilon$  with the following properties:  $W_{\text{hom}}^c(\mu)$  is of class  $C^{l,\alpha}$  jointly in  $(u, \mu)$  and has dimension equal to  $\dim E_p^c$ .

The tangent bundle  $E^c(t) = T_{h(t)} W_{\text{hom}}^c(0)$  along the homoclinic orbit is uniquely defined and, in fact, given by the intersection of  $T_{h(t)} W^{s,\text{lu}}(p, 0)$  and  $T_{h(t)} W^{ls,u}(p, 0)$ .

If *Hypotheses 2.2(1)*, *2.3(2)*, and *2.4* are met, then  $E^c(t)$  is a continuous bundle of planes that limit on the eigenspace of  $f_u(p, 0)$  associated with the leading eigenvalues, and a two-dimensional homoclinic centre manifold therefore exists in this situation; see [Figure 3.3](#). In this case, the orientability of  $W_{\text{hom}}^c(\mu)$  is determined by the index  $\mathcal{O}$  defined in (2.10).

### 3.5. Stable foliations

For some global bifurcations, reductions to homoclinic centre manifolds or other global centre manifolds are not possible. In such situations, stable foliations may still provide reductions to semiflows on branched manifolds. This applies, in particular, to flows that contain Lorenz-like attractors, but also to studies of annihilation processes of suspended horseshoes through homoclinic bifurcations. Whether such reductions are helpful depends on the smoothness of the foliation (i.e. the smoothness of the holonomy map along the leaves of the foliation).

Geometric models for Lorenz-like attractors depend on the existence of a stable foliation  $\mathcal{F}^{\text{ss}}$  with one-dimensional leaves for the flow. Stable foliations can be constructed by graph transform techniques. Their smoothness, however, depends on spectral gap conditions at the equilibrium: if  $\lambda^{\text{ss}} < \lambda^{\text{ls}}$  are the stable eigenvalues, and  $\lambda^{\text{lu}}$  is the unstable eigenvalue, then the stable foliation is  $C^l$  for  $l < (\lambda^{\text{ls}} - \lambda^{\text{ss}})/\lambda^{\text{lu}}$ ; see [326]. In particular, for open sets of eigenvalues, the stable foliation is merely continuous. A stable foliation  $\mathcal{G}^{\text{ss}}$  for a return map on a cross section is obtained by projecting  $\mathcal{F}^{\text{ss}}$  along flow lines into the section. This projection can be expected to increase smoothness and map  $\mathcal{F}^{\text{ss}}$  to a continuously differentiable foliation. A general result along these lines has been obtained by Homburg [187] in the context of bifurcations of singular horseshoes: we shall present it later in Proposition 4.3. Similar statements apply, for instance, to Lorenz-like attractors that occur in the unfolding of two homoclinic loops to an equilibrium with resonant eigenvalues; a continuous stable foliation for the flow projects to a continuously differentiable stable foliation for the return map on a cross section.

Alternatively, one could construct foliations directly for the return map defined on some cross section, rather than constructing them for flows and then projecting along flow lines. We shall now present a theorem that gives continuously differentiable stable foliations for return maps with prescribed asymptotic expansions. Consider a map  $\Pi = (f, g)$  from  $D = ([-1, 1] \setminus \{0\}) \times [-1, 1]^n \subset \mathbb{R} \times \mathbb{R}^n$  to itself of the form

$$\begin{aligned} f(x, y) &= \begin{cases} x_*^- + |x|^\alpha (A^- + \phi^-(x, y)), & x < 0, \\ x_*^+ + |x|^\alpha (A^+ + \phi^+(x, y)), & x > 0, \end{cases} \\ g(x, y) &= \begin{cases} y_*^- + |x|^\alpha \psi^-(x, y), & x < 0, \\ y_*^+ + |x|^\alpha \psi^+(x, y), & x > 0. \end{cases} \end{aligned}$$

We make the following assumption.

**HYPOTHESIS 3.2** (Asymptotic Expansions). *Assume that the following is true for some  $\eta > 0$ :*

- (1)  $A^-, A^+ \neq 0$ .
- (2)  $\phi^\pm, \psi^\pm$  have continuous derivatives up to order two in  $D$ .
- (3) For some  $\varepsilon > 0$ ,  $\|\frac{\partial^{k+l}}{\partial x^k \partial y^l} \phi^\pm\| \leq \varepsilon |x|^{\eta-k}$  and  $\|\frac{\partial^{k+l}}{\partial x^k \partial y^l} \psi^\pm\| \leq \varepsilon |x|^{\eta-k}$ .

The following result due to Shashkov and Shil'nikov is stated for maps with sufficiently small higher-order terms; sharper results can be found in [358].

**THEOREM 3.4** ([358]). *There exists an  $\varepsilon_0 > 0$  so that, if Hypothesis 3.2 is met for  $\varepsilon < \varepsilon_0$ , then  $\Pi$  admits a stable  $C^1$ -smooth foliation on  $D$  with  $C^2$  leaves, that contains  $\{x = 0\}$ .*

The preceding result is similar to a result by Rychlik [336], who considered maps on  $D$  that are close to  $(x, y) \mapsto (1 - c \operatorname{sign}(x)|x|^\alpha, 0)$  with  $c \in (1, 2)$  and  $c\alpha > 2$ ; closeness is expressed by estimates as in Hypothesis 3.2(3) with  $\alpha + \eta > 1$ .

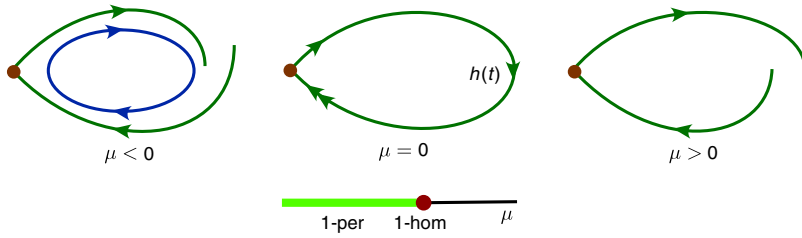


Fig. 3.4. The phase [top row] and bifurcation [bottom panel] diagrams of a nondegenerate homoclinic orbit  $h(t)$  to an equilibrium with real leading eigenvalues  $|\nu^s| > \nu^u$  is shown. A unique attracting single-round periodic orbit bifurcates to one side of  $\mu = 0$  (for  $\mu < 0$ , say), and its period goes to infinity as  $\mu$  tends to zero.

### 3.6. Case study: The creation of periodic orbits from a homoclinic orbit

We illustrate the different approaches to homoclinic bifurcation theory by applying them to a homoclinic bifurcation of codimension one with real leading eigenvalues: specifically, we assume that [Hypotheses 2.1, 2.2, 2.3\(4\) and 2.4](#) are all satisfied. First, we derive bifurcation equations that capture periodic orbits using, separately, Shil'nikov variables and Lin's method. Afterwards, we discuss how homoclinic centre manifolds can be used to derive similar results. The resulting phase and bifurcation diagrams are summarized in [Figure 3.4](#). We remark that this bifurcation is an example of a blue sky catastrophe where a periodic orbit disappears in a bifurcation at which its period goes to infinity.

**3.6.1. Shil'nikov variables.** For Shil'nikov variables, it is convenient to introduce a single cross section  $\Sigma$  that is transverse to the homoclinic solution  $h$  at  $\mu = 0$ . Proceeding in this way will also illuminate the differences and similarities between the approaches via Shil'nikov variables and Lin's method.

Let  $\Pi(\cdot, \mu) : \Sigma \rightarrow \Sigma$  be the first-return map given by the flow of  $\dot{u} = f(u, \mu)$  and consider the dynamical system inclination-flip configuration, then a parameter-dependent coordinate  $x_{j+1} = \Pi(x_j, \mu)$  for  $x_j \in \Sigma$ . If both the stable and unstable manifolds of  $p$  along the homoclinic orbit are not in an system  $x = (x^{ss}, x^{lu}, x^{uu})$  on  $\Sigma$  can be chosen so that

$$\begin{aligned} W^s(p) \cap \Sigma &= \{x^{lu}, x^{uu} = 0\}, \\ W^u(p) \cap \Sigma &= \{x^{ss}, x^{lu} = 0\}, \\ W^{s,lu}(p) \cap \Sigma &\cap \overline{W^s(p) \cap \Sigma} \quad \{x^{uu} = 0\}, \\ W^{ls,u}(p) \cap \Sigma &\cap \overline{W^u(p) \cap \Sigma} \quad \{x^{ss} = 0\}, \end{aligned}$$

where the notation  $W \cap_q V$  means that  $W$  is tangent to  $V$  at  $q$ . The following result provides a normal form for the return map on  $\Sigma$ . Its proof combines [Proposition 3.2](#) for solutions between the local cross sections  $\Sigma_{\text{in}}$  and  $\Sigma_{\text{out}}$  with expansions for the transition maps between  $\Sigma$  and  $\Sigma_{\text{in}}$  and between  $\Sigma_{\text{out}}$  and  $\Sigma$ ; the transition time from  $\Sigma_{\text{in}}$  to  $\Sigma_{\text{out}}$  is solved for as a function of the initial data on these cross sections.



PROPOSITION 3.3 ([193]). Assume that *Hypothesis 2.4* is met. In the coordinates constructed above,  $x_{j+1} = \Pi(x_j, \mu)$  and  $x_j$  are related by

$$(x_{j+1}^{ss}, x_{j+1}^{lu}, x_j^{uu}) = \tilde{\Pi}(x_j^{ss}, x_j^{lu}, x_{j+1}^{uu}, \mu)$$

for some map  $\tilde{\Pi}$  with the asymptotics

$$\begin{aligned} x_{j+1}^{ss} &= O(|x_j^{lu}|^{-v^s/v^u+\eta}), \\ x_{j+1}^{lu} &= a(\mu) + \varphi(\mu)[x_j^{lu}]^{-v^s/v^u} + O(|x_j^{lu}|^{-v^s/v^u+\eta}), \\ x_j^{uu} &= O(|x_{j+1}^{lu}|^{1+\eta}), \end{aligned} \quad (3.16)$$

where  $a$  and  $\varphi$  are smooth functions of  $\mu$ .

Consider a sequence  $\{x_j\}_{j \in \mathbb{Z}}$  in  $\Sigma$ . These points lie on the same orbit of  $\Pi(\cdot, \mu)$  if, and only if,

$$(x_{j+1}^{ss}, x_{j+1}^{lu}, x_j^{uu}) - \tilde{\Pi}(x_j^{ss}, x_j^{lu}, x_{j+1}^{uu}, \mu) = 0, \quad j \in \mathbb{Z} \quad (3.17)$$

or, more detailed,

$$\begin{aligned} x_{j+1}^{ss} - O(|x_j^{lu}|^{-v^s/v^u+\eta}) &= 0, \\ x_{j+1}^{lu} - a(\mu) - \varphi(\mu)[x_j^{lu}]^{-v^s/v^u} + O(|x_j^{lu}|^{-v^s/v^u+\eta}) &= 0, \\ x_j^{uu} - O(|x_{j+1}^{lu}|^{1+\eta}) &= 0. \end{aligned} \quad (3.18)$$

If we denote by  $l_{\mathbb{R}^N}^\infty$  be the space of bi-infinite sequences with entries in  $\mathbb{R}^N$  equipped with the supremum norm, then Equation (3.17) can be considered as an equation in  $l_{\mathbb{R}^{n_{ss}}}^\infty \times l_{\mathbb{R}}^\infty \times l_{\mathbb{R}^{n_{uu}}}^\infty$ . Since the first and third equations in (3.18) depend smoothly on  $(x^{ss}, x^{uu})$ , we can use an implicit function theorem in  $l^\infty$ , see [44] or [95], to solve the first and third equations in (3.18) for  $(x^{ss}, x^{uu})$ . Substitution into the second equation of (3.18) gives the reduced bifurcation equations

$$x_{j+1}^{lu} - a(\mu) - \varphi(\mu)[x_j^{lu}]^{-v^s/v^u} + O(|x_j^{lu}|^{-v^s/v^u+\eta}) = 0. \quad (3.19)$$

Note that the higher-order terms depend on the entire sequence  $x^{lu} = \{x_j^{lu}\}_{j \in \mathbb{Z}}$ . Imposing periodicity on the entire sequence, we obtain  $N$  reduced bifurcation equations if we wish to investigate  $N$ -periodic orbits of  $\Pi(\cdot, \mu)$  that have period  $N$ . In particular, the reduced bifurcation equation for a single-round periodic orbit takes the form

$$x^{lu} - a(\mu) - \varphi(\mu)[x^{lu}]^{-v^s/v^u} - O(|x^{lu}|^{-v^s/v^u+\eta}) = 0,$$

where  $x^{lu} \in \mathbb{R}$  with  $|x^{lu}| \ll 1$ . Solving is straightforward: if  $-v^s > v^u$ , then the solution is  $x^{lu} = a'(0)\mu + o(\mu)$  where  $\mu$  is such that  $a'(0)\mu > 0$ ; similarly, if  $-v^s < v^u$ , then the solution is  $[x^{lu}]^{-v^s/v^u} = -a'(0)\mu/\varphi(0) + o(\mu)$  where  $\mu$  is such that  $a'(0)\mu/\varphi(0) < 0$ .

**3.6.2. Lin's method.** To apply the technique outlined in Section 3.3 to the bifurcation of single-round periodic orbits from homoclinic orbits, we assume  $|v^s| > v^u$  and consider the bifurcation equations (3.14)

$$-\langle v_*^u(\mu), v^u(\mu) \rangle e^{-2v^u T_{j+1}} + M\mu + O\left(e^{-(2v^u+\eta)\min_{k \in \mathbb{Z}} T_k} + |\mu|^2\right) = 0 \quad (3.20)$$

for  $j \in \mathbb{Z}$  that describe the existence of solutions that follow the homoclinic orbit with prescribed return times  $T_j > T_*$ , where the Melnikov integral  $M$  has been defined in Hypothesis 2.2. To find single-round periodic orbits, we set  $T_j = T$  for all  $j$ : With this choice, (3.20) is the same equation for all  $j$  due to the uniqueness of solutions to (3.12), and (3.20) therefore reduces to

$$-\langle v_*^u(\mu), v^u(\mu) \rangle e^{-2v^u T} + M\mu + O\left(e^{-(2v^s+\eta)T} + |\mu|^2\right) = 0. \quad (3.21)$$

For  $M \neq 0$ , this equation can be solved uniquely for  $\mu$  as a function of  $T$  to get

$$\mu = \mu_*(T) = \frac{1}{M} \langle v_*^u(0), v^u(0) \rangle e^{-2v^u T} + O\left(e^{-(2v^s+\eta)T}\right), \quad (3.22)$$

which proves that periodic orbits bifurcate from the homoclinic orbit and shows how period and system parameter are related for large periods.

**3.6.3. Homoclinic centre manifolds.** As in local bifurcation theory, where Lyapunov–Schmidt reductions and centre manifold reductions each have their own advantages and disadvantages, analytic and geometric approaches to global bifurcation theory can complement each other: some aspects and questions are easier to investigate from a geometric viewpoint while, in other situations, it might be advantageous to use analytical techniques.

The assumptions we made at the beginning of Section 3.6 guarantee that the homoclinic centre manifold given in Theorem 3.3 is two-dimensional and that it is of class  $C^{1+\alpha}$  for some  $\alpha > 0$  and depends in a  $C^{1+\alpha}$  fashion on the parameter  $\mu$ . The lack of smoothness of homoclinic centre manifolds is often an obstacle for deriving and solving bifurcation equations near homoclinic orbits. On the other hand, the dimension reduction gives much insight into the geometry of the flow, which is helpful when studying stability and hyperbolicity. In our case, the existence of a two-dimensional homoclinic centre manifold  $W_{\text{hom}}^c$  immediately excludes the existence of  $N$ -periodic solutions for  $N > 2$  if  $W_{\text{hom}}^c$  is non-orientable, and for  $N > 1$  for orientable  $W_{\text{hom}}^c$ .

Even though the homoclinic centre manifold is only  $C^{1+\alpha}$ , it is possible to find continuously differentiable linearizing coordinates of the flow on  $W_{\text{hom}}^c(\mu)$  [187]. The argument to prove this statement utilizes the smoothness of the flow in  $\mathbb{R}^n$  and differs as such from a parameter-dependent version of linearization results by Belitskii [37].

**PROPOSITION 3.4.** *There exist constants  $\alpha > 0$  and  $\varepsilon > 0$  so that, for  $|\mu| < \varepsilon$ , there are local  $C^{1+\alpha}$  coordinates on  $W_{\text{hom}}^c(\mu)$  near the equilibrium  $p(\mu)$ , which are also  $C^{1+\alpha}$  in*

$\mu$ , for which the local transition map  $\Pi_{\text{loc}} : W_{\text{hom}}^c(\mu) \cap \Sigma_{\text{in}} \rightarrow W_{\text{hom}}^c(\mu) \cap \Sigma_{\text{out}}$  has the expression  $\Pi_{\text{loc}}(x, \mu) = x^{-v^s(\mu)/v^u(\mu)}$ .

Up to a re-parameterization, the first-return map  $\Pi$  on  $W_{\text{hom}}^c(\mu) \cap \Sigma_{\text{in}}$  is therefore given by

$$\Pi(x, \mu) = \mu + x^{-v^s/v^u}(a(\mu) + O(x^\eta))$$

for some  $\eta > 0$ , and the bifurcation of single-round periodic orbits follows trivially from this expression. We mention again that the geometric approach via homoclinic centre manifolds immediately excludes the existence of multi-round homoclinic and periodic solutions.

## 4. Phenomena

Although there are a great many types of homoclinic and heteroclinic bifurcations, as evidenced by the considerable list in Section 5, a number of features are common to several of them. In this section, we discuss such common features. We also survey a number of transitions through global bifurcations from Morse–Smale flows with finitely many critical elements (equilibria and periodic orbits) to complicated dynamics involving suspended transitive hyperbolic sets, singular hyperbolic attractors, and suspended Hénon-like attractors. Homoclinic-doubling cascades, the analogue of period-doubling cascades for homoclinic orbits, provide another mechanism for the transition from Morse–Smale to non–Morse–Smale flows. Finally, we review the creation of intermittent time series through homoclinic bifurcations.

### 4.1. $N$ -pulses and $N$ -periodic orbits

The occurrence of multi-round homoclinic orbits is of central importance in applications to spatial dynamics, where they correspond to travelling or standing multi-pulses. A large collection of homoclinic bifurcation results have been derived with these applications in mind.

Multi-round homoclinic orbits may be created in bifurcations from a homoclinic orbit, which we refer to as the primary or single-round homoclinic orbit. The definition of a multi-round homoclinic orbit is given with reference to a tubular neighbourhood of the primary homoclinic orbit. Let  $\dot{u} = f(u, \mu)$  be a family of ODEs with a homoclinic orbit  $h = \{h(t)\}_{t \in \mathbb{R}}$  at  $\mu = 0$  and denote by  $\mathcal{U}$  a small tubular neighbourhood of the closure of the homoclinic orbit  $h$ . If  $\Sigma$  is a cross section placed at  $h(0)$  that is transverse to the orbit  $h$  when  $\mu = 0$ , then a homoclinic loop contained in  $\mathcal{U}$  is called an  $N$ -homoclinic orbit or  $N$ -round homoclinic loop if it intersects  $\Sigma$  precisely  $N$  times. Multi-round periodic orbits are defined analogously by counting the number of intersections with  $\Sigma$ . Figure 4.1 illustrates a 2-round homoclinic orbit and a 2-round periodic orbit.

Near Shil'nikov saddle-focus homoclinic orbits,  $N$ -homoclinic orbits for all  $N$  appear in the unfolding, provided the leading eigenvalues satisfy a certain condition (in this case, the homoclinic orbit is called a wild saddle-focus homoclinic orbit). Although perturbations

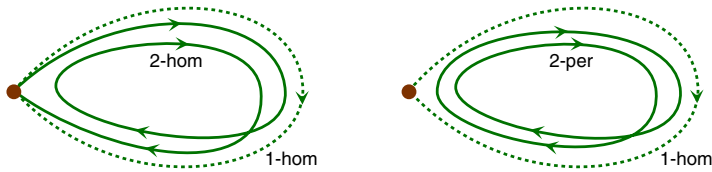


Fig. 4.1. Shown are a 2-round homoclinic orbit [left] and a 2-round periodic orbit [right], relative to a given primary homoclinic orbit.  $N$ -homoclinic and  $N$ -periodic orbits for larger  $N$  are defined analogously: these solutions make  $N$  rounds near the primary, or 1-homoclinic, orbit.

from a codimension-one homoclinic bifurcation with real leading eigenvalues do not give rise to multi-round homoclinic orbits, various codimension-two homoclinic bifurcations with real leading eigenvalues do. Consider a two-parameter family  $\dot{u} = f(u, \mu)$  of ODEs that have a homoclinic solution  $h(t)$  to a hyperbolic equilibrium with real leading eigenvalues when  $\mu = 0$ . As long as [Hypotheses 2.1](#) and [2.2](#) hold, the homoclinic orbit persists along a curve in parameter space (see [Section 6.1](#) for a topological continuation theory for homoclinic orbits). At isolated parameter points along such a curve, and assuming that the leading eigenvalues stay real, one of the conditions in [Hypotheses 2.3](#) or [2.4](#) may be violated. At these points, multi-round homoclinic solutions may bifurcate from the curve of primary homoclinic solutions, and we refer to the bifurcation theorems in [Sections 5.1.5–5.1.7](#) for the results. Subject to further conditions on the spectrum and geometry of the flow, a double-round homoclinic orbit will be created in all these cases.

If the dynamics contains  $N$ -periodic orbits for all  $N$ , then the overall dynamics is often organized around suspended horseshoes: we say that the system has a suspended horseshoe if the first-return map  $\Pi$  to a cross section  $\Sigma$  of the homoclinic orbit admits a Smale's horseshoe, that is, a compact hyperbolic invariant set on which the dynamics coming from  $\Pi$  is conjugated to a shift on two symbols; see, for instance, [\[313\]](#) for further details. Suspended horseshoes are found in the unfoldings of various homoclinic bifurcations. Wild saddle-focus homoclinic orbits and bi-focus homoclinic orbits have suspended horseshoes in each neighbourhood of the homoclinic orbit. Homoclinic orbits to equilibria with real leading eigenvalues can give rise to suspended horseshoes in cases of higher codimension. In particular, suspended horseshoes can be found in the unfolding of codimension-two inclination-flip and orbit-flip bifurcations, provided certain eigenvalue conditions are met. In all these scenarios, suspended horseshoes are created through homoclinic tangencies (akin to the Hénon family), which implies that it is impossible to give a complete description of the bifurcations in finitely many parameters: we refer to [\[50, 313\]](#) for an overview of the underlying theory.

#### 4.2. Robust singular dynamics

Within the framework of perturbations in the  $C^1$  topology, a comprehensive theory of generic dynamical properties of flows exists, at least for three-dimensional flows, which is still under active development. This program is in the spirit of Palis's papers [\[309,310\]](#) in which conjectures for the dynamics of typical systems are sketched out. We will only

skim through the relevant results, instead referring to the books [20,50] and the review article [4] for further discussion.

The central technique is the  $C^1$ -connection lemma by Hayashi [179,180]. For critical elements<sup>4</sup>  $\sigma$ , the connection lemma entails the following. Let  $\sigma$  be a hyperbolic critical element of a differential equation  $\dot{u} = f(u)$  and suppose there exists a point  $q \notin \sigma$  in  $(\overline{W^s(\sigma)} \cap W^u(\sigma)) \cup (\overline{W^u(\sigma)} \cap W^s(\sigma))$ ; such a point is called an almost homoclinic point. There is then a differential equation  $C^1$  close to  $f$  that coincides with  $f$  in a neighbourhood of  $\sigma$  and has a homoclinic solution to  $\sigma$ .

Making use of his connection lemma, Hayashi finished in [179,180] the proof of the  $C^1$  stability conjecture for flows, which states that a differential equation on a three-dimensional compact manifold is  $C^1$  structurally stable if, and only if, it is uniformly hyperbolic and all stable and unstable manifolds are transverse.

There are, however, open sets of flows that are not  $C^1$  structurally stable, so that the search for typical dynamical properties remains. Arroyo and Rodriguez-Hertz [31] proved that a differential equation on a three-dimensional compact manifold can be  $C^1$  approximated either by a system that is uniformly hyperbolic or that has a homoclinic tangency between stable and unstable manifolds of periodic orbits or a singular cycle; a *singular cycle* is a heteroclinic cycle between critical elements among which is at least one equilibrium.

In this context, transitive but non-hyperbolic sets can exist. The primary example is the Lorenz attractor, which is the ‘butterfly’ attractor in the Lorenz equations [265,384] given by

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= \rho x - y - xz, \\ \dot{z} &= -\beta z + xy.\end{aligned}$$

To understand the geometry of the Lorenz attractor, geometric models have been developed [10,167,423]. The robust strange attractors in these models are called *geometric Lorenz attractors* or *Lorenz-like attractors*. Tucker provided a computer-assisted proof for the existence of a robust strange attractor in the Lorenz equations whose geometry is that of the robust strange attractors in geometric Lorenz models.

**THEOREM 4.1** ([397,398]). *The Lorenz equations support a robust strange attractor for the classical parameter values  $\sigma = 10$ ,  $\rho = 28$ , and  $\beta = 8/3$ .*

It turns out that the relevant notion for a general theory of robust transitive sets is that of dominated splitting. A compact invariant set  $\Lambda$  of  $\dot{u} = f(u, \mu)$  is a *partially hyperbolic set* if, up to time reversal, there is an invariant dominated splitting  $T\Lambda = E^s \oplus E^c$ : this means that there are positive constants  $K, \lambda$  such that

- (1)  $E^s$  is contracting:  $\|\partial_x \varphi_t|_{E_x^s}\| \leq K e^{-\lambda t}$ , for all  $x \in \Lambda$  and  $t > 0$ ;
- (2)  $E^s$  dominates  $E^c$ :  $\|\partial_x \varphi_t|_{E_x^s}\| \|\partial_x \varphi_t|_{E_{\varphi_t(x)}^c}\| \leq K e^{\lambda t}$ , for all  $x \in \Lambda$  and  $t > 0$ .

---

<sup>4</sup>A critical element is a periodic orbit or an equilibrium.

The central direction  $E^c$  of  $\Lambda$  is said to be volume expanding if the additional condition

$$|\det(\partial_x \varphi_t|_{E_x^c})| \geq K e^{\lambda t}$$

holds for all  $x \in \Lambda$  and  $t > 0$ . Let  $\Lambda$  be a compact invariant set of  $\dot{u} = f(u)$  that contains at least one equilibrium, then  $\Lambda$  is called a *singular hyperbolic set* if it is partially hyperbolic with volume expanding central directions, and  $\Lambda$  is called a *singular hyperbolic attractor* if, in addition, it attracts all points in some open neighbourhood.

A compact invariant set  $\Lambda$  for a flow  $\varphi_t$  of  $\dot{u} = f(u)$  is robust transitive if it is the maximal invariant set  $\bigcap_{t \in \mathbb{R}} \varphi_t(U)$  inside an open neighbourhood  $U$  of  $\Lambda$  and, for the flow  $\psi_t$  of any differential equation  $\dot{u} = g(u)$  with  $gC^1$ -close to  $f$ ,  $\bigcap_{t \in \mathbb{R}} \psi_t(U)$  is a nontrivial transitive set (nontrivial means that it does not consist of a critical element). Morales, Pacifico and Pujals demonstrated for three-dimensional flows that  $C^1$  robust strange attractors which contain equilibria are singular hyperbolic sets. An invariant set is proper if it is not the whole manifold.

**THEOREM 4.2 ([286]).** *A robust transitive set for a three-dimensional flow that contains an equilibrium is a proper singular hyperbolic attractor or repeller.*

There are ODEs in  $\mathbb{R}^3$  that possess singular hyperbolic attractors with any number of equilibria, and Section 5.5.5 provides theorems that can be used to construct such examples; see also Figure 5.28. An example of a singular hyperbolic attractor from a fluid convection model that contains two equilibria can be found in [297]. We refer to [13, 19, 22, 267] for further results, addressing primarily ergodic properties, Lorenz-like and other singular hyperbolic attractors.

A different class of strange attractors that contain an equilibrium is found in contracting Lorenz models. Like the strange attractors encountered in the Hénon family, these strange attractors persist only in a measure theoretic sense. Contracting Lorenz models are geometric Lorenz models but with contracting instead of expanding central directions (the saddle quantity, i.e. the quotient  $-\nu^s/\nu^u$  of the leading stable and unstable eigenvalues at the equilibrium, is larger than 1). Arneodo, Coullet and Tresser [26] noted that contracting Lorenz models contain dynamics that can be described by interval maps (in fact unimodal maps if one restricts to  $\mathbb{Z}_2$ -symmetric flows). Indeed, in geometric (contracting) Lorenz models, an invariant strong stable foliation enables a reduction to a semiflow on a branched manifold and to interval maps through a first-return map. We recall a result by Rovella on the dynamics of contracting Lorenz models, focusing on strange attractors that contain an equilibrium.

**THEOREM 4.3 ([335]).** *There exists a contracting Lorenz model  $\dot{u} = f(u)$  in  $\mathbb{R}^3$  with an attractor  $\Lambda$  containing a hyperbolic equilibrium so that the following properties hold:*

- (1) *There exists a local basin of attraction  $\mathcal{B}$  for  $\Lambda$ , a neighbourhood  $\mathcal{V}$  of  $f$  in the  $C^3$  topology, and an open and dense subset  $\mathcal{V}_0 \subset \mathcal{V}$  such that, for an ODE  $\dot{u} = g(u)$  with  $g \in \mathcal{V}_0$ , the maximal invariant set in  $\mathcal{B}$  consists of the equilibrium, one or two attracting periodic orbits, and a hyperbolic basic set.*

- (2) *In generic two-parameter families  $\dot{u} = f(u, \mu)$  with  $f(\cdot, 0) = f$ , there is a set of positive measure containing  $\mu = 0$  as a density point for which an attractor in  $\mathcal{B}$  containing the equilibrium exists.*

The proof of the above result uses further eigenvalue conditions at the equilibrium to ensure  $C^3$  strong stable foliations. Such strange attractors can occur in unfoldings of (symmetric) resonant homoclinic loops [288,329]; see also Section 5.5.5.

Another type of strange attractor in three-dimensional flows that contain an equilibrium would be formed by the strange attractors for which there is no strong stable foliation near the attractor and hence no reduction to a branched manifold. Apart from spiral attractors (see Section 5.1.2), these can be found in hooked Lorenz models [50]. Such attractors may also occur in unfoldings from certain homoclinic and heteroclinic bifurcations; see [233, 298]. For an analysis of the dynamics of interval maps with both singularities and critical points that yield approximate one-dimensional models for these attractors, see [115,268, 269].

Examples of higher-dimensional robust strange attractors that contain equilibria are described in [52,406].

### 4.3. Singular horseshoes

In this section we review the creation, or disappearance, of a suspended Smale horseshoe through sequences of homoclinic bifurcations; each periodic orbit in the horseshoe will be created in a homoclinic bifurcation. We will point out relations with bifurcation theory of singular cycles between an equilibrium and a periodic orbit. The disappearance of suspended horseshoes through homoclinic bifurcations has been shown to occur near the codimension-two bifurcations of homoclinic orbits in inclination-flip or orbit-flip configurations, which will be discussed in Section 5. The appearance in these bifurcations makes this phenomenon a common feature in the bifurcation diagrams of models of ODEs with two or more parameters.

The starting point is the disappearance through a homoclinic bifurcation of a periodic orbit that is part of a larger transitive invariant set. This can be expected to trigger additional bifurcations of orbits within the invariant set. The basic case illustrating this phenomenon is where the periodic orbit lies in the suspension of a hyperbolic horseshoe, and we review this scenario in detail below, following [187]. It is instructive to compare the bifurcation scenario with other scenarios in which horseshoes break up, in particular the involved scenarios triggered by homoclinic tangencies [313] or saddle-node bifurcations of periodic orbits [100,102,116,431].

To fix ideas, consider a differential equation that admits an invariant set which is a suspended Smale horseshoe  $\Lambda$ . Orbits in  $\Lambda$  are coded by doubly infinite sequences  $\{0, 1\}^{\mathbb{Z}}$ . Write  $q$  for the periodic orbit that corresponds to the coding  $0^\infty$ . The stable manifold  $W^s(q)$  is a one-sided boundary leaf of the lamination  $W^s(\Lambda)$ ; near a point  $x \in W^s(q)$ , other leaves in  $W^s(\Lambda)$  can be found only on one side of  $W^s(q)$ . The unstable manifold  $W^u(q)$  is likewise a one-sided boundary leaf of the lamination  $W^u(\Lambda)$ . Note that both  $W^s(q)$  and  $W^u(q)$  are orientable surfaces.

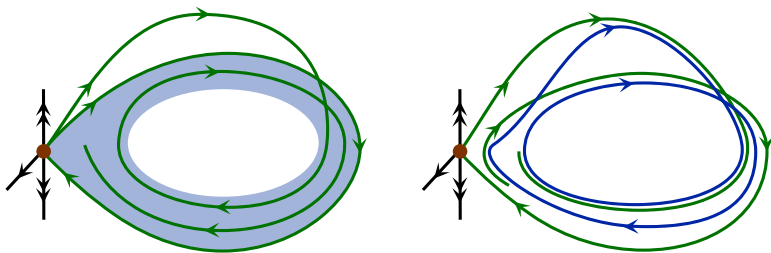


Fig. 4.2. A homoclinic orbit that coexists with a superhomoclinic orbit, which connects the equilibrium to the homoclinic orbit, implies the existence of a singular horseshoe [left]. Singular cycles between the equilibrium and a periodic orbit appear in the unfolding of singular horseshoes [right].

Bifurcations that possess a suspended horseshoe for  $\mu < 0$  and undergo a homoclinic bifurcation of the periodic orbit  $q(t)$  at  $\mu = 0$  can be identified in the following set-up. Assume  $\dot{u} = f(u, \mu)$  is a one-parameter family of ODEs on  $\mathbb{R}^n$  that has a homoclinic solution  $h(t)$  at  $\mu = 0$  which satisfies [Hypotheses 2.1, 2.2, 2.3\(2\)](#), and [2.4](#). In particular, the linearization  $f_u(p, 0)$  has unique real leading eigenvalues and is not in an inclination-flip or orbit-flip configuration.

**HYPOTHESIS 4.1.** *Consider the following spectral and geometric conditions:*

- (1)  $-\nu^s/\nu^u > 1$ .
- (2) *The two-dimensional homoclinic centre manifold  $W_{\text{hom}}^c$  near  $h$  is an annulus.*

Under these hypotheses one can identify a nonempty stable set of the homoclinic orbit.

**LEMMA 4.1.** *If  $\dot{u} = f(u, 0)$  has a homoclinic orbit  $h$  to an equilibrium  $p$  so that [Hypotheses 2.1, 2.2, 2.3\(2\)](#), and [2.4](#) are met and, in addition, [Hypothesis 4.1](#) is satisfied, then the stable set*

$$\mathcal{M}^s(\bar{h}) = \{x \in \mathbb{R}^n; \omega(x) = \bar{h}\},$$

where  $\bar{h} = h \cup \{p\}$  is the closure of  $h$ , is a manifold with boundary  $W^s(p)$ , and  $\dim \mathcal{M}^s(\bar{h}) = 1 + \dim W^s(p)$ .

Under the above assumptions, so-called superhomoclinic orbits may exist which are transverse intersections of  $\mathcal{M}^s(\bar{h})$  and  $W^u(p)$  [[187,377,401](#)]. The following proposition connects the resulting bifurcations to bifurcations of singular cycles between an equilibrium and a periodic orbit as further illustrated in [Figure 4.2](#).

**PROPOSITION 4.1.** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs on  $\mathbb{R}^n$  with a homoclinic orbit  $h$  to an equilibrium  $p$  so that [Hypotheses 2.1, 2.2, 2.3\(2\)](#), and [2.4](#) are met. If [Hypothesis 4.1](#) is met and  $\mathcal{M}^s(\bar{h}) \pitchfork W^u(p)$ , then there are bifurcation values that accumulate onto  $\mu = 0$  at which  $\dot{u} = f(u, \mu)$  possesses a singular cycle between  $p$  and a saddle periodic orbit  $q(t)$  such that  $\dim W^u(q) = \dim W^u(p)$  and  $W^s(q) \pitchfork W^u(p)$ .*

In the reverse direction, if a singular cycle as stated exists, the  $\lambda$ -lemma implies that  $W^u(p)$  accumulates onto  $W^u(q)$ , and small perturbations of the differential equation



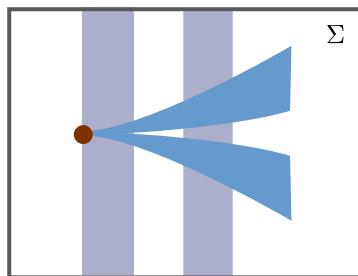


Fig. 4.3. The return map for a singular horseshoe (pictured here in a two-dimensional section  $\Sigma$ ) maps two wedge-shaped regions to vertical strips, roughly contracting horizontal directions and expanding vertical directions.

will create homoclinic orbits to  $p$ . Bifurcations from singular cycles form an interesting problem in their own right, and we review relevant material on this subject, which was initiated in [35], in Section 5.2.4. We note that the singular cycles in Proposition 4.1 are called expanding (and remark that the definitions in [35] consider the time reversed flow).

Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs on  $\mathbb{R}^n$  with a homoclinic orbit  $h$  to an equilibrium  $p$  as in the above proposition, so that  $\mathcal{M}^s(\bar{h}) \cap W^u(p)$  along a solution  $h_1(t)$ . Identify a small neighbourhood  $\mathcal{U}$  of the closed set  $\overline{h \cup h_1}$ . We will formulate a bifurcation theorem for orbits in  $\mathcal{U}$ . Along the unstable separatrices of  $p$  one can identify a bundle of centre directions  $E^c$  that are invariant under the variational equation and converge to the sum of the leading stable and leading unstable directions over  $p$  as the base point approaches  $p$  (see Section 3.4). This bundle yields a continuous plane bundle over  $\overline{h \cup h_1}$ .

**HYPOTHESIS 4.2.** *Different orientations of the bundle of centre directions  $E^c$  occur:*

- (1) *The centre bundle  $E^c$  along  $\overline{h \cup h_1}$  is orientable.*
- (2) *The centre bundle  $E^c$  along  $\overline{h \cup h_1}$  is non-orientable.*

**THEOREM 4.4** ([187]). *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs on  $\mathbb{R}^n$  with a homoclinic orbit  $h$  to the equilibrium  $p$  and a generalized homoclinic orbit  $h_1$  in  $\mathcal{M}^s(\bar{h}) \cap W^u(p)$  so that Hypotheses 2.1, 2.2, 2.3(2), and 2.4 are met. On one side of  $\mu = 0$ , say for  $\mu < 0$ , the invariant set in  $\mathcal{U}$  is then a suspended horseshoe. For  $\mu \geq 0$ , the bifurcation set is a set of Lebesgue measure zero with the following properties:*

- (1) *If Hypothesis 4.2(1) is met, then the bifurcation set is a Cantor set in which homoclinic bifurcation values lie dense.*
- (2) *If Hypothesis 4.2(2) is met, then the bifurcation set is the union of a Cantor set in which homoclinic bifurcation values lie dense, and infinitely many sequences of homoclinic bifurcation values that converge to points in the Cantor set.*

First, we describe the dynamics for  $\mu < 0$ . Let  $\Sigma$  be a small cross section transverse to the homoclinic orbit  $h$ . The return map on  $\Sigma$  maps two horizontal wedge-like strips that emerge from the same point to two vertical strips; see Figure 4.3. The invariant set of

$\dot{u} = f(u, 0)$  in  $\mathcal{U}$  is a singular hyperbolic set (as introduced in Section 4.2 for flows in  $\mathbb{R}^3$  after reversing the direction of time) which we call a *singular horseshoe* (we note that our usage of this term differs from [251]).

We write  $\Pi^{-1}$  for the continuous extension of the inverse of the first-return map on  $\Sigma$  and remark that  $\Pi^{-1}$  has a fixed point at  $W^s(p) \cap \Sigma$ ; note that  $\Pi^{-1}$  is not invertible. Let  $\Lambda$  be the maximal invariant set of  $\Pi^{-1}$ , then we claim that iterates of  $\Pi^{-1}$  restricted to  $\Lambda$  are conjugate to a factor of a full shift on two symbols. To make this precise, let  $\Omega = \{0, 1\}^{\mathbb{Z}}$  equipped, as usual, with the product topology. Let  $\tau$  be the right shift  $\tau s(i) = s(i-1)$  on  $\Omega$ . Define an equivalence relation on  $\Omega$  by  $s \sim s'$  if  $s(i) = s'(i) = 0$  for all  $i \leq 0$ . Let  $\Omega_* = \Omega / \sim$ , and let  $\tau_*$  be the map on  $\Omega_*$  induced by  $\tau$ ; observe that  $\tau_*$  is not invertible. Denote the two vertical strips in  $\Pi(\Sigma)$  by  $V_0, V_1$  with  $W^s(p) \cap \Sigma$  lying in the closure of  $V_0$ . For  $x \in \Lambda$ , let  $h(x) \in \Omega_*$  be the itinerary of  $x$  under  $\Pi^{-1}$ :  $h(x)(i) = j$  if  $[\Pi^{-1}]^i(x) \in \overline{V_j}$ , then  $h$  gives a conjugacy of  $\Pi^{-1}$  on  $\Lambda$  with  $\tau_*$  on  $\Omega_*$ , see [108].

PROPOSITION 4.2.  $h \circ \Pi^{-1} = \tau_* \circ h$ .

The orbits in the suspended horseshoe, which exists for  $\mu < 0$ , disappear in bifurcations for  $\mu > 0$  as stated in Theorem 4.4. The homoclinic bifurcations that occur for  $\mu > 0$  come in two types, depending on the orientability of the homoclinic centre manifolds. The homoclinic bifurcations of non-orientable homoclinic orbits are the locally isolated bifurcation values that appear in one of the two cases.

The analysis leading to Theorem 4.4 proceeds through dimension reductions using invariant manifolds and foliations. This reduction leads to interval maps, whose study then gives the theorem. We will continue with some details of the constructions. Choose a coordinate chart  $(x, y)$  on  $\Sigma$  so that  $W^s(p)$  intersects  $\Sigma$  in  $(0, 0)$  and  $W^u(p) \cap \Sigma = \{x = 0\}$ . The first-return map  $\Pi$  on  $\Sigma$  admits a continuously differentiable unstable foliation that contains  $\{x = 0\}$  as a leaf.

PROPOSITION 4.3. *Under the conditions of Theorem 4.4, there exists a normally hyperbolic centre-unstable manifold  $W^{cu}(\overline{h \cup h_1})$  in  $\mathcal{U}$  that is  $C^{1+\alpha}$  for some  $\alpha > 0$  jointly in  $(u, \mu)$ . Moreover, there are  $C^{1+\alpha}$  coordinates  $(x, y)$  on  $W^{cu}(\overline{h \cup h_1}) \cap \Sigma$  so that  $\{x = \text{constant}\}$  defines an unstable foliation  $\mathcal{F}^u$  for the first-return map  $\Pi$  on  $\Sigma$ .*

The invariant manifolds and foliations are constructed by the usual graph transform techniques. We note that the centre unstable manifold admits an invariant unstable foliation for the flow which is, in general, only continuous. Its projection along flow lines to the cross section  $\Sigma$  defines an invariant foliation for the first-return map that is continuously differentiable. The return map  $\Pi$  acts on the space of leaves of  $\mathcal{F}^u$  as a multi-valued interval map, while its inverse acts as a piecewise expanding interval map  $\pi$ . The domain of  $\pi$  is the union of two intervals  $I_1, I_2$  with the left boundary  $I_1$  equal to 0. In fact, one finds the following asymptotic expansions for  $\pi$ :

$$\pi(x) = \begin{cases} a(\mu) + b(\mu)x^{-v^u/v^s}(1 + O(x^\eta)), & x \in I_1, \\ d(\mu)x^{-v^u/v^s}(1 + O(x^\eta)), & x \in I_2, \end{cases}$$

for some  $\eta > 0$ ; see Figure 4.4.

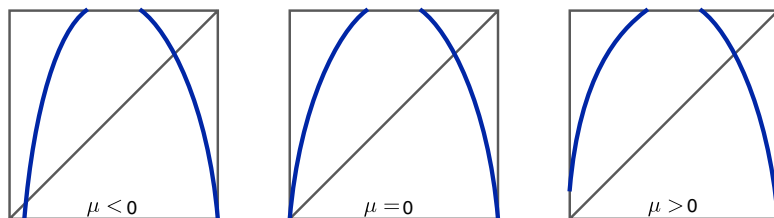


Fig. 4.4. The map  $\pi$  on the leaves of the unstable foliation is an expanding interval map for the case described in Hypothesis 4.2(2).

#### 4.4. The boundary of Morse–Smale flows

Morse–Smale flows are flows with finitely many equilibria and periodic orbits that are all hyperbolic and whose stable and unstable manifolds intersect transversally. Morse–Smale flows are robust, and their dynamics can therefore only change when a family crosses the boundary of the set of all Morse–Smale flows.

Write  $\mathcal{X}(M)$  for the set of smooth vector fields on a compact manifold  $M$  endowed with the Whitney topology. A bifurcation on the boundary of the set of Morse–Smale flows in  $\mathcal{X}(M)$  is accessible if there exists a path  $\chi : [0, 1] \mapsto \mathcal{X}(M)$  of Morse–Smale flows except for the bifurcation at the endpoint  $\chi(i)$ . We are interested in accessible bifurcation points in whose vicinity the Morse–Smale flows are not everywhere dense. In particular, we review bifurcations directly from Morse–Smale flows to flows with suspended horseshoes or strange attractors. Local bifurcations with such transitions are considered in Section 5.4.

Consider a one-parameter family of ODEs  $\dot{u} = f(u, \mu)$  that unfold a saddle-node bifurcation of an equilibrium  $p$  occurring at  $\mu = 0$ . Shil’nikov established that the unfolding of a saddle-node bifurcation can create suspended horseshoes if there are multiple homoclinic solutions to the saddle-node equilibrium. The bifurcation is on the boundary of Morse–Smale flows and is accessible from the set of Morse–Smale flows.

**THEOREM 4.5** ([367]). *Consider a one-parameter family of ODEs that has a saddle-node equilibrium, which satisfies Hypothesis 2.5, so that  $\mathcal{M}^s(p)$  has transverse intersections with  $\mathcal{M}^u(p)$  along more than one orbit. On one side of the parameter  $\mu = 0$ , the family then has a hyperbolic transitive set with infinitely many periodic orbits.*

There exist families as in the above theorem for which the periodic orbits in the hyperbolic transitive set span every possible knot and link type [153].

Later papers have produced several constructions of codimension-one bifurcations leading directly from a Morse–Smale flow to strange attractors of different types. Afraïmovich, Chow and Liu [11] describe a codimension-one bifurcation from a Morse–Smale flow to Lorenz-like attractors, that involve singular cycles between a hyperbolic equilibrium  $p$  and a periodic solution  $q$  of saddle-node type: the equilibrium  $p$  is assumed to have two-dimensional stable and one-dimensional unstable directions and leading eigenvalues that satisfy  $-\nu^s/\nu^u < 1$ , while the periodic orbit  $q$  has an attracting normally hyperbolic direction. Its stable set  $\mathcal{M}^s(q)$  is thus an open region of  $\mathbb{R}^3$  bounded by the stable manifold  $W^s(q)$  of  $q$ , and the unstable set  $\mathcal{M}^u(q)$  is a surface bounded by  $q$ .

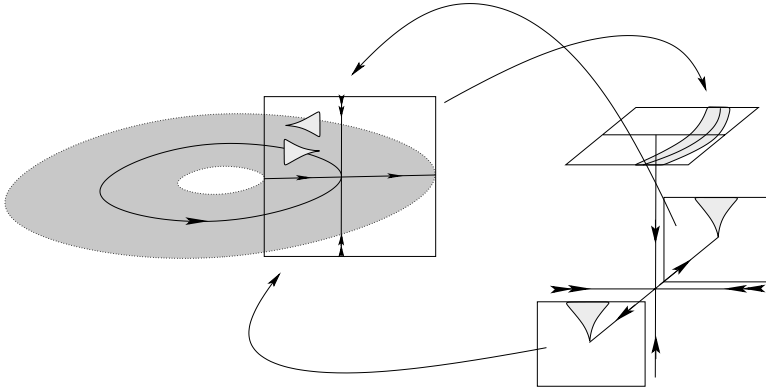


Fig. 4.5. Lorenz-like attractors are created directly from a Morse–Smale flow at the unfolding of certain singular cycles between a hyperbolic equilibrium and a saddle-node periodic orbit.

The differential equation at the moment of bifurcation satisfies  $W^u(p) \subset \mathcal{M}^s(q)$  and admits a transverse intersection of  $\mathcal{M}^u(q)$  with  $W^s(p)$ . Observe that both separatrices that form  $W^u(p)$  tend to the periodic orbit  $q$ . One-parameter unfoldings from an open set of vector fields with the given properties are shown to possess Lorenz-like attractors on one side of the bifurcation value  $\mu = 0$ ; see Figure 4.5 for an illustration. The creation of Lorenz-like attractors from codimension-two homoclinic bifurcations in  $\mathbb{Z}_2$ -equivariant flows is considered in Section 5.5.5.

Additional analysis by Morales, Pacífico and Pujals [285] produced further examples of transitions from Morse–Smale flows, through bifurcations of singular cycles between a hyperbolic equilibrium and a saddle-node periodic orbit, to robust singular attractors. Their work contains examples where these last attractors are not Lorenz-like. In the example of Afraïmovich, Chow and Liu,  $\mathcal{M}^u(q)$  intersects the strong stable foliation  $\mathcal{F}^{ss}$  of  $\mathcal{M}^s(q)$  transversally. In critical cycles, when a tangency of  $\mathcal{M}^u(q)$  with some leaf of  $\mathcal{F}^{ss}$  exists, transitions from Morse–Smale flows to flows with suspended Hénon-like attractors are possible [289]. We mention that Morales [281] has examples of direct transitions in one parameter families of ODEs from Lorenz-like attractors to suspended Plykin attractors, i.e. from a singular hyperbolic attractor to a uniformly hyperbolic attractor.

In two-parameter families of ODEs, one can study the creation of multiple homoclinic solutions to a saddle-node equilibrium through a quadratic tangency of  $\mathcal{M}^s(p)$  and  $\mathcal{M}^u(p)$ . Let  $\psi(t)$  be the unique bounded nonzero solution to the adjoint variational equation

$$\dot{w} = -f_\mu(h(t), 0)^* w$$

so that  $\psi(t) \in (T_{h(t)} W^{cs}(p) + T_{h(t)} W^{cu})^\perp$  exists and decays to zero exponentially. Similar to Hypothesis 2.5, we define the two vectors

$$M = \int_{-\infty}^{\infty} \langle \psi(t), f_\mu(h(t), 0) \rangle dt, \quad N = \langle w_c, f_\mu(p, 0) \rangle$$

in  $\mathbb{R}^2$ .

**HYPOTHESIS 4.3.** *Consider the following nondegeneracy conditions:*

- (1)  $\langle w_c, f_{uu}(p, 0)[v_c, v_c] \rangle \neq 0$ ,
- (2)  $M$  and  $N$  are linearly independent.

The following result from [79,282] establishes the existence of suspended Hénon-like strange attractors, arbitrarily close to the boundary of Morse–Smale systems in codimension-two unfoldings involving a saddle-node bifurcation of an equilibrium.

**THEOREM 4.6** ([79,282]). *Assume [Hypothesis 2.5](#) and  $-v^s/v^u > 1$ . A generic two-parameter family  $\dot{u} = f(u, \mu)$  in  $\mathbb{R}^3$  for which [Hypothesis 4.3](#) is met then admits parameter values with suspended Hénon-like strange attractors.*

#### 4.5. Homoclinic-doubling cascades

Cascades of period-doubling bifurcations in one-parameter families of ODEs have attracted much interest as it is one of the routes to onset of chaos. Scaling properties of the period-doubling bifurcations, universal in the sense that they do not depend on details of the family, are the most noticeable feature. Recall that the universal scalings are explained by renormalization theory and, in particular, by the existence of a fixed point for the renormalization operator with a single unstable eigenvalue. One can ask the general question: how this scenario can change if a second parameter is varied? One way involves the disappearance of the periodic orbits through homoclinic bifurcations and gives rise to cascades of homoclinic-doubling bifurcations. In this section, we review the relevant literature and discuss approximations by interval maps: universal scalings in the bifurcation diagram for the interval maps turn out to be related to the appearance of a fixed point of a renormalization operator with two unstable eigenvalues. An extensive numerical investigation of homoclinic-doubling cascades can be found in [302], and further numerical evidence for the existence of homoclinic-doubling cascades has been provided in the Shimizu–Morioka model

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= x - ay - xz, \\ \dot{z} &= -bz + x^2;\end{aligned}$$

see [376].

Under certain conditions, an orbit homoclinic to a hyperbolic equilibrium can undergo a homoclinic-doubling bifurcation that creates a double-round homoclinic orbit, see Sections 5.1.5–5.1.7. In a two-parameter family of ODEs, one can continue homoclinic solutions along curves in the parameter plane. A two-parameter family of ODEs is said to possess a cascade of homoclinic bifurcations if there exists a connected set of parameter values in the parameter plane, corresponding to homoclinic orbits, that contains a cascade  $\{\mu_n\}_{n \in \mathbb{N}}$  of homoclinic-doubling bifurcations in which a  $2^n$ -homoclinic orbit is created. The actual existence of this phenomenon is confirmed in [195], following earlier work in [231,232].

**THEOREM 4.7 ([195]).** *In the space of two-parameter families of smooth vector fields on  $\mathbb{R}^3$ , there is an open set consisting of families that possess a cascade of homoclinic-doubling bifurcations.*

Next, we outline the bifurcation theory of homoclinic-doubling cascades that occur in unfoldings of codimension-three homoclinic bifurcations, which lead to strongly dissipative return maps that are small perturbations of interval maps.

Consider a three-parameter family  $\dot{u} = f(u, \mu)$  of ODEs with  $\mu = (\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3$  that have a hyperbolic equilibrium at  $p$ . Write  $v^{\text{ss}} < v^s$  for the two real negative eigenvalues at  $p$ , and  $v^u$  for the single real positive eigenvalue, and define

$$\alpha = -v^{\text{ss}}/v^u, \quad \beta = -v^s/v^u.$$

Suppose the differential equation has a homoclinic orbit to  $p$  when  $\mu_2 = 0$ , which is an inclination-flip homoclinic orbit when  $\mu_1 = \mu_2 = 0$ . We let  $\mu_3 = 2 - \frac{1}{\beta}$ , so that  $\beta = \frac{1}{2}$  when  $\mu_3 = 0$ . We assume that the eigenvalue index  $\alpha$  will satisfy the open condition  $\alpha > 1$ . Furthermore, assume that the two-dimensional invariant manifold  $W^{\text{ls},u}(p)$  of  $p$  has a quadratic tangency with the stable manifold  $W^s(p)$  along the homoclinic orbit when  $\mu = 0$ . Take a cross section  $\Sigma$  transverse to the homoclinic orbit.

Following [195, 196], there are coordinates  $(x^{\text{ss}}, x^u)$  on  $\Sigma$  so that the first-return map  $\Pi(\cdot, \mu) : \Sigma \rightarrow \Sigma$  has the asymptotics

$$\Pi(x^{\text{ss}}, x^u, \mu) = \begin{pmatrix} a_0(\mu) + a_1(\mu)(x^u)^\beta + a_2(\mu)(x^u)^{2\beta} + O((x^u)^{2\beta+\eta}) \\ \mu_2 + \mu_1(x^u)^\beta + a_3(\mu)(x^u)^{2\beta} + O(|\mu_1|(x^u)^{\beta+\eta} + (x^u)^{2\beta+\eta}) \end{pmatrix} \quad (4.1)$$

for some  $\eta > 0$  when  $\mu$  varies near zero. The coefficients  $a_j(\mu)$  depend smoothly on  $\mu$ , and the higher-order terms can be differentiated for  $x^u > 0$ .

**THEOREM 4.8.** *Let  $\dot{u} = f(u, \mu)$  be as above and assume that  $a_3(0) > 1$ . For each small fixed  $\mu_3 > 0$ , the resulting two-parameter family  $\dot{u} = f(u, \mu)$  possesses a cascade of homoclinic-doubling bifurcations.*

Related scenarios near other codimension-three bifurcation points are treated in [196].

The proof of the preceding theorem utilizes return maps. First, a rescaling transforms the first-return map to a map that is a small perturbation of an interval map. Let  $\Pi$  be as in (4.1) and define rescaled coordinates  $(\hat{x}^{\text{ss}}, \hat{x}^u)$  by

$$\begin{aligned} x^{\text{ss}} - a_0(\mu) &= |\mu_1| \hat{x}^{\text{ss}}, \\ x^u &= \left| \frac{\mu_1}{2a_3(\mu)} \right|^{1/\beta} \hat{x}^u. \end{aligned}$$

The following proposition, which gives expansions for the first-return map in rescaled coordinates, is proved by a direct computation.

**PROPOSITION 4.4.** *Let  $\hat{\Pi}$  be the first-return map in the rescaled coordinates  $(\hat{x}^{\text{ss}}, \hat{x}^u)$  and write  $b_2(\mu) = |\mu_1|^{\mu_3}/4a_3(\mu)$  and  $b_1(\mu) = b_2(\mu)(4a_3(\mu)\mu_2|\mu_1|^{-2} - 1)$ . For some  $\eta > 0$ ,*

we then have

$$\hat{\Pi}(\hat{x}^{\text{ss}}, \hat{x}^{\text{u}}) = \begin{pmatrix} \frac{a_1(\mu)}{2a_3(\mu)}(\hat{x}^{\text{u}})^{\beta} + \mathcal{O}(|\mu_1|^{\eta}(\hat{x}^{\text{u}})^{\beta}) \\ b_1(\mu) + b_2(\mu)(1 - (\hat{x}^{\text{u}})^{\beta})^2 + \mathcal{O}(|\mu_1|^{\eta}(\hat{x}^{\text{u}})^{\beta}) \end{pmatrix}.$$

As  $\mu_1 \rightarrow 0$ , restricting  $\hat{x}^{\text{u}}$  to a compact interval and parameters  $(\mu_1, \mu_2, \mu_3)$  to a chart near the origin on which  $b = (b_1, b_2)$  is bounded,  $\hat{\Pi}$  converges to a map that depends only on  $\hat{x}^{\text{u}}$  and is given by

$$\hat{x}^{\text{u}} \mapsto f(\hat{x}^{\text{u}}) = b_1 + b_2(1 - (\hat{x}^{\text{u}})^{\beta})^2. \quad (4.2)$$

Scaling properties that exist in the bifurcation diagram of  $f$  in the  $(b_1, b_2)$  parameter plane can be investigated with renormalization theory [199]. The renormalization scheme involves a renormalization operator whose fixed point has two unstable directions, in contrast to the single unstable direction one finds in the theory of period-doubling cascades. A renormalization theory for the actual differential equations is not available.

What follows is a concise version of the results, and we refer to [199] for precise and complete statements and a discussion of the implications for scaling properties of the bifurcation diagram. Consider the class of unimodal functions on a fixed interval  $[0, R]$  for some  $R > 1$  of the form  $x \mapsto g(x^{\beta})$  with  $g$  smooth and a minimal value  $g(i) = 0$ . The renormalization operator  $\mathcal{R}$ , defined on a subset of these functions, maps  $g$  to a rescaled version of the second iterate  $g^2$ .

**THEOREM 4.9 ([199]).** *For  $\beta > \frac{1}{2}$  with  $|\beta - \frac{1}{2}| \ll 1$ , the renormalization operator  $\mathcal{R}$  possesses an isolated fixed point  $\phi$ . The function  $\phi$  depends continuously on  $\beta$  and converges to  $x \mapsto (1 - \sqrt{x})^2$  as  $\beta \rightarrow \frac{1}{2}$ . The linearization  $D\mathcal{R}$  at  $\phi$  has two unstable eigenvalues  $v_1, v_2$  which depend continuously on  $\beta$  and satisfy  $v_1 \rightarrow 2$ ,  $v_2 \rightarrow \infty$  as  $\beta \rightarrow \frac{1}{2}$ . The remainder of the spectrum of  $D\mathcal{R}(\phi)$  is strictly inside the unit disc.*

#### 4.6. Intermittency

Intermittency has been identified as one of the principal routes of the transition from a periodic state to chaos [45]. In this context, a time series is said to be intermittent if it is almost periodic apart from infrequent variations. Thus, intermittent time series consists of an almost periodic laminar phase and a chaotic burst or relaminarization phase. Saddle-node, Hopf or period-doubling bifurcations of periodic orbits can give rise to intermittency, which are labelled intermittency of type I, II, or III, respectively. Homoclinic bifurcations of homoclinic solutions to equilibria with real leading eigenvalues can also cause intermittent time series, and we review their characteristics in this section.

In equivariant systems, heteroclinic cycles can be robust under equivariant perturbations and attract orbits from an open neighborhood; see Section 5.5.1. Such stable heteroclinic cycles provide a mechanism for intermittency: a solution approaching the cycle spends long periods near equilibria and makes fast transitions from one equilibrium to the next. In a perfectly symmetric system, the return times increase monotonically and approach infinity,

thus making the intermittent behaviour uninteresting, though we remark that more complex switching phenomena can occur in more complicated heteroclinic networks. However, under small, symmetry-breaking perturbations, the cycling behaviour persists even though there may no longer be a cycle, and the transition times no longer converge to infinity. Alternatively, stochastic perturbations lead to boundedness of transition times [24,386]. Relevant references for these dynamical features include [90,92,186,277].

Suppose  $f(u, \mu)$  is a one-parameter family of vector fields that unfolds a homoclinic bifurcation with real leading eigenvalues at  $\mu = 0$ . We will consider codimension-one phenomena where a continuous bundle of centre directions along the homoclinic orbit exists (see also Section 4.3).

We assume the unstable manifold of the equilibrium  $p$  is one-dimensional. It is therefore the union of two separatrices and the equilibrium  $p$ : one separatrix in  $W^u(p)$  forms the homoclinic solution  $h$ , while we assume that the other separatrix converges to the homoclinic solution. This can only happen if  $-v^s/v^u > 1$  at  $\mu = 0$ . Such flows appear in unfoldings of gluing bifurcations where two homoclinic orbits coexist; see Section 5.1.8. More precisely, we shall assume the following.

**HYPOTHESIS 4.4.** *The unstable manifold  $W^u(p)$  of the hyperbolic equilibrium  $p$  is one-dimensional and therefore equal to the union of  $p$  and two separatrices  $W^\pm(p)$ . The leading eigenvalues  $v^s, v^u$  are unique and real and satisfy  $-v^s/v^u > 1$ . We assume that there is a homoclinic orbit  $h$  to  $p$  that lies in  $W^u_+(p)$ . The bundle  $E^c$  of centre directions is a continuous orientable plane bundle over the closure  $\bar{h} = h \cup \{p\}$  of  $h$ .*

The above hypothesis implies that the homoclinic centre manifold is an orientable annulus near the homoclinic orbit  $h$ . Due to the orientability of the plane bundle  $E^c$  and the condition on the saddle quantity  $-v^s/v^u$ , the stable set  $\mathcal{M}^s(h)$  of initial data whose  $\omega$ -limit set is  $\bar{h}$  is an open set bounded by  $W^s(p)$ .

**HYPOTHESIS 4.5** (Existence of a Superhomoclinic Orbit). *The  $\omega$ -limit set of points on the separatrix  $h_1(t) \in W^u_-(p)$  equals  $\bar{h}$  so that  $W^u_-(p) \subset \mathcal{M}^s(\bar{h})$ .*

We need additional information about the behaviour around the superhomoclinic orbit  $h_1(t)$ . The plane bundle  $E^c(h_1(t))$  along the superhomoclinic orbit  $h_1(t)$  extends to a continuous plane bundle  $E^c$  over  $\bar{h} \cup h_1$ .

**HYPOTHESIS 4.6.** *Consider the following properties for the plane bundle  $E^c$  over  $\bar{h} \cup h_1$ :*

- (1) *The bundle  $E^c$  over  $\bar{h} \cup h_1$  is an orientable plane bundle.*
- (2) *The bundle  $E^c$  over  $\bar{h} \cup h_1$  is a non-orientable plane bundle.*

**THEOREM 4.10** ([187]). *Consider a one-parameter family of ODEs as above that satisfies Hypotheses 4.4 and 4.5.*

- (1) *If Hypothesis 4.6(1) is met, then the bifurcation set in  $\mu \geq 0$  is a Cantor set of zero Lebesgue measure: the one-sided boundary points of the Cantor set are homoclinic bifurcations, while there is a unique periodic attractor for parameters outside the bifurcation set.*



- (2) If [Hypothesis 4.6\(2\)](#) is met, then the bifurcation set is a cascade  $\mu_n \downarrow 0$  of homoclinic bifurcations, while there are one or two periodic attractors for parameters outside the bifurcation set.

The geometry of the flow near the separatrices of  $p$  becomes clear by the following centre manifold theorem, which is similar to the homoclinic centre manifold theorem [187, 400].

**THEOREM 4.11.** *Assume that [Hypotheses 4.4](#) and [4.5](#) are met, then there is a two-dimensional locally invariant, normally hyperbolic centre manifold  $W^c(W^u(p) \cup \{p\})$ . The manifold is of class  $C^{l,\alpha}$  jointly in  $(u, \mu)$  for each  $l \geq 1$  and  $\alpha \in (0, 1)$  with*

$$l + \alpha < \min \left\{ \frac{\lambda^{ss}}{\nu^s}, \frac{\lambda^{uu}}{\nu^u} \right\}.$$

The manifold  $W^c(W^u(p) \cup \{p\})$  is homeomorphic to either a torus with a hole when [Hypothesis 4.6\(1\)](#) is met or a Klein bottle with a hole when [Hypothesis 4.6\(2\)](#) is met. We refer to [187] for a study of scaling properties of the bifurcation set; see also [270] for related results. The reduction result expressed by [Theorem 4.11](#) entails that the dynamics is described by interval maps that occur as first-return maps on a cross section inside the centre manifold. In the orientable case, this leads to the interval maps studied by Keener [217].

## 5. Catalogue of homoclinic and heteroclinic bifurcations

This section forms the core of this survey paper: it contains a catalogue of bifurcation results for homoclinic and heteroclinic orbits. We have subdivided the list into sections treating generic systems, conservative and reversible systems, and equivariant systems. Various homoclinic bifurcations, bifurcations from heteroclinic cycles, and the occurrence of homoclinic solutions from local bifurcations are reviewed.

We use the notation and the hypotheses that we introduced in [Section 2](#). We consider differential equations of the form  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^n$  with parameters  $\mu \in \mathbb{R}^d$  for some  $d \geq 1$  that admit a homoclinic orbit  $h$  to an equilibrium  $p$  when  $\mu = 0$ ; we restrict ourselves to three- or four-dimensional systems whenever results have only been proved for such lower-dimensional systems. A typical bifurcation result requires, apart from the defining conditions, a number of nondegeneracy and unfolding conditions that encapsulate a generic dependence on parameters.

### 5.1. Homoclinic orbits in generic systems

**5.1.1. Creation of 1-periodic orbits.** The problem of the birth of a limit cycle<sup>5</sup> from a homoclinic orbit to a hyperbolic equilibrium was solved for differential equations in

---

<sup>5</sup>A limit cycle is a periodic orbit that is the  $\omega$ - or  $\alpha$ -limit set of an orbit other than itself.

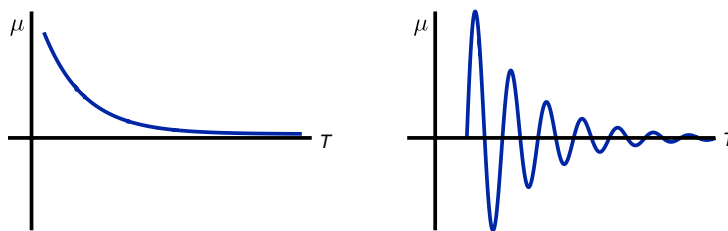


Fig. 5.1. Plotted are the period  $T$  versus the parameter  $\mu$  for the 1-periodic orbits that bifurcate from a homoclinic orbit with real leading eigenvalue (Theorem 5.1) [left] and complex leading eigenvalue (Theorem 5.4) [right]. See (3.22) for the expression of period versus parameter in the first case; a similar formula, now with  $v^u$  complex, holds in the latter case.

the plane by Andronov and Leontovich [18]. Shil'nikov extended this work to differential equations on  $\mathbb{R}^n$  [363,366]. We state a general result on the creation of a limit cycle in homoclinic bifurcations with saddle quantity  $-\text{Re } v^s/v^u$  larger than one [376]. We also refer to Section 3.6 and Figure 3.4 for the case of real leading eigenvalues.

**THEOREM 5.1.** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs with a homoclinic solution  $h(t)$  to a hyperbolic equilibrium  $p$  at  $\mu = 0$ . Assume that Hypotheses 2.2 and 2.3(1) are met, then a unique periodic solution  $q(t, \mu)$  bifurcates from the homoclinic orbit; this periodic solution is hyperbolic, it bifurcates either for  $\mu > 0$  or for  $\mu < 0$ , and  $q(t, \mu)$  converges to  $h(t)$  as  $\mu \rightarrow 0$  for each fixed  $t$ . Furthermore,  $W^s(q) = \dim W^s(p) + 1$ .*

Note that the periodic orbit  $q(t, \mu)$  is stable if  $p$  has one-dimensional unstable manifold and that the period of  $q(t, \mu)$  goes to infinity as  $\mu \rightarrow 0$ ; see Figure 5.1, and also Section 3.6 and Figure 3.4. If the saddle quantity is less than one, we can simply apply the preceding theorem to the time-reversed vector field (noting that the conclusion about the dimension of the stable manifold is then for the unstable manifold in the original time variable).

In [311] the question was posed as to whether other bifurcations were possible in which a periodic orbit develops infinite period and disappears in a so-called *blue sky catastrophe*. This question was answered affirmatively by Medvedev [275,276]; see also [55,56,207].

**THEOREM 5.2** ([275]). *There is a one-parameter family  $\dot{u} = f(u, \mu)$  of vector fields on the Klein bottle that has a saddle-node bifurcation of a periodic orbit at  $\mu = 0$  and an attracting periodic orbit for  $\mu > 0$  with unbounded arclength as  $\mu \downarrow 0$ .*

The example given in [275] can be embedded in a family of ODEs on  $\mathbb{R}^4$  that have a Klein bottle as the normally attracting invariant manifold. For  $\mu = 0$ , the Klein bottle coincides with the set of homoclinic solutions of the saddle-node periodic orbit. A different example has been constructed by Shil'nikov and Turaev [378,405]: again, a periodic attractor with unbounded period appears from a saddle-node bifurcation of a periodic orbit, but the unstable manifold of the saddle-node periodic orbit now forms a tube that spirals back towards the saddle-node periodic orbit.

**THEOREM 5.3** ([378,405]). *There is a one-parameter family  $\dot{u} = f(u, \mu)$  of vector fields on  $\mathbb{R}^3$  that has a saddle-node bifurcation of a periodic orbit at  $\mu = 0$  and an attracting periodic orbit for each  $\mu > 0$  with unbounded arclength as  $\mu \downarrow 0$ .*

### A numerical analysis of the explicit family of ODEs

$$\begin{aligned}\dot{x} &= x[2 + \mu - b(x^2 + y^2)] + x^2 + y^2 + 2y, \\ \dot{y} &= -z^3 - (y + 1)(z^2 + y^2 + 2y) - 4x + \mu y, \\ \dot{z} &= z^2(y + 1) + x^2 - \varepsilon\end{aligned}$$

reveals the existence of the latter type of blue sky catastrophe in low order ODEs [149]. We also refer to [372,374] for blue sky catastrophes in a singularly perturbed context and their role in explaining bursting phenomena in neuron models. Finally, we mention that the term blue sky catastrophe has also been used in reversible or conservative ODEs when sheets of periodic orbits are bounded by a homoclinic orbit; see [Theorem 5.38](#) below.

**5.1.2. Shil'nikov's saddle-focus homoclinic orbits.** A systematic study of the dynamics near saddle-focus homoclinic orbits was pioneered by Shil'nikov since the mid 1960s. Under an eigenvalue condition that states that the real leading eigenvalue dominates the complex conjugate leading eigenvalues, infinitely many periodic orbits of saddle type were shown to occur in each neighbourhood of the homoclinic orbit [364]. These periodic orbits are contained in suspended horseshoes that accumulate onto the homoclinic orbit [368]. In fact, the periodic orbits near two coexisting saddle-focus homoclinic orbits under the same eigenvalue condition are known to span every possible knot and link type [152]. Dynamical features beyond hyperbolic suspended horseshoes, including the existence of periodic and strange attractors accumulating onto the homoclinic orbit, were described in later papers [160,190,303].

Attractors with a spiral structure were envisaged to occur for perturbations of ODEs with two coexisting saddle-focus homoclinic orbits under suitable eigenvalue conditions [27]. Namely, a dissipative differential equation  $\dot{u} = f(u)$  with two saddle-focus homoclinic orbits to the same equilibrium will admit an invariant tubular neighbourhood  $\mathcal{U}$  that contains both homoclinic orbits. Any sufficiently small perturbation of  $f$  will then have an attractor inside  $\mathcal{U}$ . There is some evidence, from the study of models of interval maps, that spiral attractors exist and are robust in the sense of measure [21,306,307].

In this section, we discuss mostly three-dimensional ODEs, as much of the existing literature treats three-dimensional ODEs and the most complete picture arises in three dimensions. Several of these results extend readily to higher-dimensional systems. We start with some general results in  $\mathbb{R}^n$ . Consider the differential equation  $\dot{u} = f(u)$  with a hyperbolic equilibrium  $p$  that has a unique real leading unstable eigenvalue  $v^u$  and unique complex conjugate leading stable eigenvalues  $v^s, \bar{v}^s$ . Assume the differential equation has a homoclinic solution  $h(t)$  to  $p$ ; see [Figure 5.2](#) for an illustration. We will consider different conditions on the saddle quantity  $-\operatorname{Re} v^s/v^u$  as outlined in the following hypothesis.

**HYPOTHESIS 5.1 (Eigenvalue Conditions).** *Consider the following eigenvalue conditions:*

- (1) *The saddle-focus homoclinic orbit is tame :  $-\operatorname{Re} v^s/v^u > 1$ .*
- (2) *The saddle-focus homoclinic orbit is wild :  $-\operatorname{Re} v^s/v^u < 1$ .*
- (3) *We have  $-2\operatorname{Re} v^s/v^u > 1$  (which, in  $\mathbb{R}^3$ , means that the differential equation is dissipative near  $p$ ).*

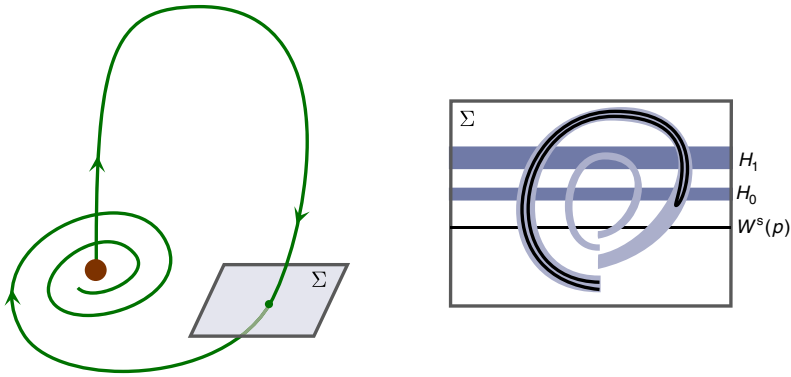


Fig. 5.2. A Shil'nikov saddle-focus homoclinic orbit is shown in the left panel. If the complex eigenvalues are closest to the imaginary axis, then the return map  $\Pi : \Sigma \rightarrow \Sigma$  contains infinitely many horseshoe maps obtained by restriction to strips  $H_i$ . The second iterate  $\Pi^2$  restricted to a union  $H_i \cup H_j$  may contain nonhyperbolic dynamics: this is illustrated in the right panel where the horseshoe-shaped images of two strips under  $\Pi$  and, with a darker shade, part of the image under  $\Pi^2$  of the union of the strips are shown.

**Hypothesis 5.1(1)** implies that an attracting periodic solution approaches the homoclinic solution  $h$  as  $\mu$  goes to zero from one side and disappears there, a bifurcation that we already discussed in Section 5.1.1. Shil'nikov discovered that, when **Hypothesis 5.1(2)** is met, the dynamics near the homoclinic solution involves infinitely many periodic orbits arbitrarily close to the homoclinic solution [364,368]. A geometric explanation of the organization of these periodic orbits into infinitely many horseshoes has been given in [396]. We note that, for three-dimensional flows, Belitskii's linearization theorem [37], which we stated in Section 3.1, allows a  $C^1$  linearization of the flow near the equilibrium, thus facilitating asymptotic expressions for the first-return map on a cross section. For higher-dimensional flows, the homoclinic centre manifold theorem [342] gives a three-dimensional homoclinic centre manifold, provided the homoclinic orbit is not in a flip configuration. This allows a geometric reduction to the three-dimensional case; we note that, although the homoclinic centre manifold is, in general, only continuously differentiable with Hölder continuous derivatives, we can still  $C^1$  linearize. Shashkov and Turaev [361] treat the dynamics near the saddle-focus homoclinic orbit for  $C^1$  vector fields.

Multi-round homoclinic orbits occur when unfolding the homoclinic orbit [132–134, 144,145].

**THEOREM 5.4.** *Assume that the system  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^n$  has a Shil'nikov saddle-focus homoclinic orbit for  $\mu = 0$  which satisfies **Hypotheses 2.2, 2.4** and **5.1(2)**. At  $\mu = 0$ , there are infinitely many suspended Smale horseshoes in each neighbourhood of the homoclinic solution. Furthermore, for each  $N > 0$ ,  $N$ -homoclinic orbits exist for infinitely many parameter values which accumulate onto  $\mu = 0$  from one side, say for  $\mu > 0$ . In fact, for  $\rho > -\text{Re } v^s / v^u$ , in any interval  $(0, \mu_+)$ , there is a set of parameter values corresponding to  $N$ -homoclinic loops that are indexed by  $\Gamma_N(\rho) = \{(k_1, \dots, k_N) \in \mathbb{N}^N; \rho k_{i-1} < k_i < k_1 / \rho\}$ . The parameter values corresponding to double-round homoclinic orbits form two*

sequences  $\mu_n^\alpha \downarrow 0$ ,  $\alpha = 0, 1$  with

$$\lim_{n \rightarrow \infty} \left| \ln \frac{\mu_{n+1}^\alpha}{\mu_n^\alpha} \right| = 2\pi \left| \frac{v^u}{\operatorname{Im} v^s} \right|.$$

For multi-round homoclinic orbits, Feroe [133] has derived estimates for the time differences between successive rounds near the principal homoclinic orbit which, equivalently, gives the spacing between pulses in the time profiles of the homoclinic orbits.

The dynamics near saddle-focus homoclinic solutions is considerably more complex than hinted at by the existence of infinitely many horseshoes. To outline the complexity, we restrict ourselves to three-dimensional ODEs, though several results have straightforward generalizations to more dimensions (see e.g. [109]). Dynamics and bifurcations are studied through the first-return map  $\Pi$  on a cross section  $\Sigma$  for which expressions can be derived by transforming the differential equation near the equilibrium into normal form and deriving expansions for solutions near the equilibrium by estimating integral formulae; see Section 3.2 on Shil'nikov variables. This procedure gives:

**PROPOSITION 5.1.** *Assume Hypothesis 5.1(2) is met, then there is a cross section  $\Sigma$  and smooth coordinates  $(\theta, z)$  on  $\Sigma$  so that  $\Pi(\cdot, \mu)$  has the following asymptotic expansion:*

$$\begin{aligned} \Pi(\theta, z, \mu) &= \begin{pmatrix} a + \phi_1(\theta) z^{\frac{-\operatorname{Re} v^s}{v^u}} \sin \left( -\frac{\operatorname{Im} v^s}{v^u} \ln z \right) + \phi_2(\theta) z^{\frac{-\operatorname{Re} v^s}{v^u}} \cos \left( -\frac{\operatorname{Im} v^s}{v^u} \ln z \right) + R_1(\theta, z, \mu) \\ b + \phi_3(\theta) z^{\frac{-\operatorname{Re} v^s}{v^u}} \sin \left( -\frac{\operatorname{Im} v^s}{v^u} \ln z \right) + \phi_4(\theta) z^{\frac{-\operatorname{Re} v^s}{v^u}} \cos \left( -\frac{\operatorname{Im} v^s}{v^u} \ln z \right) + R_2(\theta, z, \mu) \end{pmatrix}. \end{aligned}$$

The functions  $a, b, \phi_i$ ,  $i = 1, 2, 3, 4$ , are smooth functions of  $\theta$  and  $\mu$  (the dependence on  $\mu$  is suppressed in the notation) and satisfy  $\det \begin{pmatrix} \phi_1 & \phi_2 \\ \phi_3 & \phi_4 \end{pmatrix} \neq 0$ . Furthermore, there exist  $\eta > 0$  and positive constants  $C_i$  so that

$$\left| \frac{\partial^{k+l+m}}{\partial \theta^k \partial z^l \partial \mu^m} R_i(\theta, z, \mu) \right| \leq C_{k+l+m} z^{\frac{-\operatorname{Re} v^s}{v^u} + \eta}.$$

A topological invariant or a modulus is a function of the vector field that is invariant under topological equivalence; see also Section 6.2. The saddle quantity, whether larger than one or not, is always a topological invariant of saddle-focus homoclinic orbits [28, 73, 122, 395]. The proof of Togawa [395] uses link types of period orbits for saddle quantities smaller than one. Dufraine [122] proved that the absolute value of  $\operatorname{Im} v^s$  is also a modulus.

**THEOREM 5.5.** *Suppose  $f \in \mathcal{X}(\mathbb{R}^3)$  has a saddle-focus homoclinic orbit as above, then  $-\operatorname{Re} v^s / v^u$  and  $|\operatorname{Im} v^s|$  are topological invariants.*

The topological invariance of the saddle quantity reflects the dependence of the global configuration of the stable manifold of  $p$  on the saddle quantity. This fact has further implications for the bifurcation diagrams of bifurcations for two coexisting saddle-focus homoclinic orbits solutions [357, 359, 376]. It also suggests that it is natural to consider two-parameter families  $\dot{u} = f(u, \mu)$  with  $\mu \in \mathbb{R}^2$  to unfold the homoclinic bifurcation.

**HYPOTHESIS 5.2** (Generic Unfolding of Saddle Quantities). *The saddle quantity unfolds generically with respect to the parameter  $\mu_2$ :*

$$\frac{\partial}{\partial \mu_2} \left( \frac{-\operatorname{Re} v^s}{v^u} \right) \neq 0.$$

If **Hypothesis 2.2(2)** is met, we may assume that  $f(u, \mu)$  has homoclinic orbits when  $\mu_1 = 0$ . Setting  $\mu_1 = 0$  and varying  $\mu_2$  allows us to study variations in the dynamics when keeping the homoclinic orbit but changing eigenvalues. In [147,155,160], generic two-parameter families of differential equations near a Shil'nikov saddle-focus are studied, in particular with regard to curves of multi-round homoclinic orbits: they show that, in between the curves corresponding to double-round homoclinic orbits which, as stated in **Theorem 5.4**, occur along curves that accumulate onto  $\mu_1 = 0$ , there are curves of  $N$ -homoclinic orbits that fold back in  $\mu_2$ .

**THEOREM 5.6** ([160]). *Assume that  $\dot{u} = f(u, \mu)$  with  $(u, \mu) \in \mathbb{R}^3 \times \mathbb{R}^2$  has a Shil'nikov saddle-focus homoclinic orbit for  $\mu_1 = 0$  that satisfies **Hypotheses 2.2, 2.4, 5.1(2), and 5.2**. There is then a dense set of  $\mu_2$  values near  $\mu_2 = 0$  for which a curve of homoclinic orbits of  $N$ -homoclinic orbits with  $N \geq 3$  is tangent to the line  $\mu_2 = \text{constant}$  in the parameter plane.*

The horseshoes from **Theorem 5.4** are not isolated in the recurrence set. The existence of infinitely many suspended horseshoes and the nonhyperbolic dynamics around it, which we shall present below, is reminiscent of the unfolding of homoclinic tangencies. Ovsyannikov and Shil'nikov [303] established that the set of equations with nonhyperbolic dynamics is dense in the space of ODEs with Shil'nikov homoclinic orbits. More precisely, write  $\mathcal{X}_H(\mathbb{R}^3)$  for the space of ODEs in  $\mathcal{X}(\mathbb{R}^3)$  with a Shil'nikov homoclinic solution, then, for each  $\varepsilon > 0$ , the set of ODEs in  $\mathcal{X}_H(\mathbb{R}^3)$  that admit

- (1) a saddle-node bifurcation of a periodic orbit, and
- (2) a period-doubling bifurcation of a periodic orbit, and
- (3) a homoclinic tangency to a hyperbolic periodic orbit

inside an  $\varepsilon$ -tubular neighbourhood of the homoclinic solution is dense in  $\mathcal{X}_H(\mathbb{R}^3)$ .

For dissipative vector fields, the homoclinic tangencies give rise to suspended Hénon-like strange attractors. In fact, combining [303] and [99] (see also [321,322]) yields the following result.

**THEOREM 5.7.** *For each  $\epsilon > 0$ , there is a dense subset  $\mathcal{D}$  in the set of ODEs in  $\mathcal{X}_H(\mathbb{R}^3)$  for which **Hypothesis 5.1(2)–(3)** holds with the following property: for all  $f \in \mathcal{D}$ ,  $\dot{u} = f(u)$  has infinitely many coexisting strange attractors in an  $\epsilon$ -neighbourhood of  $\bar{h}$ .*

For ODEs with wild saddle-focus homoclinic orbits and  $-2\operatorname{Re} v^s/v^u = 1$  (on the boundary of the class of ODEs where **Hypothesis 5.1(3)** is met), renormalizations of a return map yield near area-preserving maps [47].

Next, we formulate a bifurcation result that makes the existence of attractors in parameterized families more precise [190]. To prove this result, one analyzes the first-return map  $\Pi$ ; the additional eigenvalue condition  $-2\operatorname{Re} v^s/v^u > 1$  assumed below allows for an improved asymptotic expansion.

**THEOREM 5.8 ([190]).** Consider a one-parameter family  $\dot{u} = f(u, 0, \mu_2)$  of ODEs in  $\mathbb{R}^3$  such that each differential equation has a saddle-focus homoclinic orbit. Suppose *Hypothesis 5.2(2)* is met and that *Hypothesis 5.1(2)–(3)* holds for each  $\mu_2$ . Let  $\Sigma$  be a cross section transverse to  $h$  and denote by  $\mathcal{U}_n$  a decreasing sequence of tubular neighbourhoods of  $\bar{h}$ .

Let  $P_{2,n}$  be the set of parameter values  $\mu_2$  for which  $f$  has an attracting 2-periodic orbit in  $\mathcal{U}_n$ , then, for  $n$  large enough,  $P_{2,n}$  is open and dense in  $I$  and  $\lim_{n \rightarrow \infty} |P_{2,n}| = 0$ , where  $|\cdot|$  denotes Lebesgue measure. Furthermore, the set of parameter values  $\gamma \in I$  for which  $X_\gamma$  has infinitely many 2-periodic attractors is dense in  $I$ , but has zero measure.

Let  $A_{2,n}$  be the set of parameter values in  $\gamma \in I$  for which  $X_\gamma$  has a Hénon-like strange attractor in  $\mathcal{U}_n$  that intersects  $\Sigma$  in 2 connected components. For  $n$  large enough,  $A_{2,n}$  has positive measure and is dense in  $I$ , and  $\lim_{n \rightarrow \infty} |A_{2,n}| = 0$ .

The complicated dynamics and the large number of attractors do not trap most orbits in a neighbourhood of the homoclinic orbit:

**THEOREM 5.9 ([190]).** Let  $\dot{u} = f(u)$  be a differential equation in  $\mathbb{R}^3$  with a saddle-focus homoclinic orbit and assume that *Hypothesis 5.1(2)–(3)* is met. Let  $\Sigma$  be a cross section transverse to the homoclinic orbit  $h$  and  $\mathcal{U}_n$  be a decreasing sequence of tubular neighbourhoods of the homoclinic orbit. If  $D_n$  is the set of points  $x$  in  $\Sigma \cap \mathcal{U}_n$  whose forward orbit leaves  $\mathcal{U}_n$ , then  $|D_n|/|\Sigma \cap \mathcal{U}_n| \rightarrow 1$  as  $n \rightarrow \infty$ .

**5.1.3. Bi-focus homoclinic orbits.** Homoclinic orbits to an equilibrium at which both the leading stable and the leading unstable eigenvalues are complex conjugate give dynamics similar to that near the wild saddle-focus homoclinic orbits that we discussed in the previous section. However, apart from a nonresonance condition, no eigenvalue conditions are needed. Also, in an unfolding, subsidiary homoclinic orbits typically accumulate on the primary bifurcation value from both sides.

The starting point is a one-parameter family  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^n$  with a homoclinic orbit to a hyperbolic equilibrium  $p$  for  $\mu = 0$ .

**HYPOTHESIS 5.3 (Leading Eigenvalues).** The leading stable eigenvalues of  $f_u(p, 0)$  are two simple complex eigenvalues  $v^s, \overline{v^s}$ ; similarly, the leading unstable eigenvalues are two simple complex eigenvalues  $v^u, \overline{v^u}$ .

A homoclinic orbit to an equilibrium that satisfies the preceding eigenvalue conditions is called a bi-focus homoclinic orbit. The existence of infinitely many suspended horseshoes accumulating onto a bi-focus homoclinic orbit has been proved by Shil'nikov [368], while bifurcations to subsidiary homoclinic orbits were studied in [154]. A discussion of the geometry of first-return maps for four-dimensional flows can be found in [141].

**THEOREM 5.10.** Assume that  $\dot{u} = f(u, \mu)$  with  $(u, \mu) \in \mathbb{R}^n \times \mathbb{R}$  has a Shil'nikov bi-focus homoclinic orbit for  $\mu = 0$  that satisfies *Hypotheses 2.2, 5.3 and 2.4*. For  $\mu = 0$ , there are infinitely many suspended Smale horseshoes in each neighbourhood of the homoclinic solution. If  $\text{Im } v^s / \text{Im } v^u \notin \mathbb{Q}$ , then, for any  $N > 0$ , there is an infinite number of parameter values near  $\mu = 0$  at which an  $N$ -homoclinic orbit exists and these parameter values accumulate from both sides onto  $\mu = 0$ .



As for the Shil'nikov saddle-focus homoclinic orbits discussed in Section 5.1.2, the horseshoes from Theorem 5.10 are not isolated in the recurrence set. Ovsyannikov and Shil'nikov [304] established that, in the space of ODEs with a Shil'nikov bi-focus homoclinic orbit in  $\mathbb{R}^4$ , one finds again dense subsets with nonhyperbolic dynamics. Write  $\mathcal{X}_H(\mathbb{R}^4)$  for the space of ODEs in  $\mathcal{X}(\mathbb{R}^4)$  with a nondegenerate Shil'nikov bi-focus homoclinic solution, then, for each  $\varepsilon > 0$ , the set of ODEs in  $\mathcal{X}_H(\mathbb{R}^4)$  for which

- (1) a saddle-node bifurcation of a periodic orbit, and
- (2) a period-doubling bifurcation of a periodic orbit, and
- (3) a homoclinic tangency to a hyperbolic periodic orbit

occur within an  $\varepsilon$ -tubular neighbourhood of the homoclinic solution, is dense in  $\mathcal{X}_H(\mathbb{R}^4)$ .

**5.1.4. Belyakov transitions.** Two different codimension-two homoclinic bifurcations are commonly referred to as Belyakov transitions [41]: the first involves an equilibrium with two real eigenvalues that collide and become complex [40], while the second [41] refers to the transition from tame to wild saddle-focus homoclinic orbits (see Section 5.1.2). Although we formulate conditions for two-parameter flows  $\dot{u} = f(u, \mu)$  with  $(u, \mu) \in \mathbb{R}^n \times \mathbb{R}^2$ , we state the results for three-dimensional flows as in Belyakov's papers.

**HYPOTHESIS 5.4** (Transition from Tame to Wild Homoclinic Loops). *At  $\mu = 0$ , we have  $\operatorname{Re} v^s + v^u = 0$  and  $[\operatorname{Re} v^s + v^u]_{\mu_2} \neq 0$ .*

**THEOREM 5.11** ([41]). *Let  $\dot{u} = f(u, \mu)$  with  $(u, \mu) \in \mathbb{R}^3 \times \mathbb{R}^2$  be a two-parameter family of ODEs on  $\mathbb{R}^3$  with a homoclinic solution  $h(t)$  to a hyperbolic equilibrium  $p$  at  $\mu = 0$ . Suppose that Hypotheses 2.2, 2.4 and 5.4 are met. Upon changing parameters, we may also assume that the primary homoclinic orbit exists for  $\mu_1 = 0$  and that  $[\operatorname{Re} v^s + v^u]_{\mu_2} > 0$  so that the primary homoclinic orbit is a wild saddle-focus homoclinic orbit for  $\{\mu_1 = 0, \mu_2 > 0\}$ . There is then a countable set of bifurcation curves for double-round homoclinic orbits that accumulate onto  $\{\mu_1 = 0, \mu_2 > 0\}$ .*

Belyakov [41] has further statements on curves of 3-round homoclinic orbits and curves of saddle-node bifurcations of periodic orbits, and it can also be shown, for instance using Lin's method, that  $N$ -homoclinic orbits bifurcate for each  $N \geq 2$ .

In the second Belyakov transition, two eigenvalues of the linearized vector field at the equilibrium collide on the real axis and become complex. Recall that the imaginary parts of the leading stable eigenvalues are nonzero if the discriminant  $\Delta(\mu)$  of  $f_u(p, \mu)$  restricted to the leading stable directions is negative.

**HYPOTHESIS 5.5** (Non-semisimple Leading Eigenvalues). *At  $\mu = 0$ , the real leading stable eigenvalue  $v^s$  has geometric multiplicity one and algebraic multiplicity two with  $\partial_{\mu_2} \Delta(0) \neq 0$  and  $-v^s/v^u < 1$ . Furthermore,*

$$\lim_{t \rightarrow \infty} \frac{1}{|t|} e^{|v^s t|} \|h(t)\| \neq 0, \quad \lim_{t \rightarrow -\infty} \frac{1}{|t|} e^{|v^s t|} \|\psi(t)\| \neq 0$$

at  $\mu = 0$ .



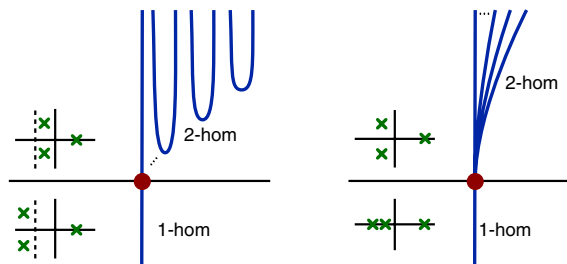


Fig. 5.3. Shown are the bifurcation diagrams of 2-homoclinic orbits for the Belyakov transitions described in [Theorem 5.11](#) [left] and [Theorem 5.12](#) [right]; the eigenvalues at the equilibrium  $p$  are shown as insets. The bifurcation curves for  $N$ -homoclinic orbits for any  $N \geq 2$  look similar.

**THEOREM 5.12** ([40,248]). *Let  $\dot{u} = f(u, \mu)$  with  $(u, \mu) \in \mathbb{R}^3 \times \mathbb{R}^2$  be a two-parameter family of ODEs on  $\mathbb{R}^3$  with a homoclinic solution  $h(t)$  to the hyperbolic equilibrium  $p$  at  $\mu = 0$ . Suppose [Hypotheses 2.2](#), [2.4\(3\)–\(4\)](#), and [5.5](#) are met. We may assume that the primary homoclinic orbit exists for  $\mu_1 = 0$  and that  $\partial_{\mu_2} \Delta(0) < 0$  so that complex conjugate leading stable eigenvalues occur for  $\mu_2 > 0$ . There are then infinitely many one-sided curves of 2-homoclinic orbits in the half plane  $\mu_1 > 0$  that emerge from  $\mu = 0$ , are tangent to  $\{\mu_1 = 0\}$  at  $\mu = (0, 0)$  and accumulate onto  $\mu_1 = 0$  from one side. Furthermore, there are infinitely many one-sided curves of saddle-node bifurcations and of period-doubling bifurcations of periodic orbits in  $\mu_1 > 0$  that emerge from  $\mu = 0$ , are tangent to  $\{\mu_1 = 0\}$  at  $\mu = (0, 0)$ , and accumulate onto  $\mu_1 = 0$  from both sides.*

In this situation, it can again be shown that  $N$ -homoclinic orbits bifurcate for each  $N \geq 2$ . See [Figure 5.3](#) for sketches of the bifurcation diagrams for the Belyakov transitions.

**5.1.5. Resonant homoclinic orbits.** The prototype homoclinic bifurcation theorem for homoclinic orbits with real leading eigenvalues, [Theorem 5.1](#), requires that the saddle quantity is not equal to one. A homoclinic orbit to an equilibrium with real leading eigenvalues for which the saddle quantity is equal to one is called a homoclinic orbit at resonance. Bifurcations from homoclinic orbits at resonance for planar vector fields have been investigated by Leontovich [259] and Nozdracheva [300]. Chow, Deng and Fiedler [95] have treated the general bifurcation problem in  $\mathbb{R}^n$ . We review these results here and remark that the case of complex conjugate eigenvalues that are at resonance is discussed in the previous section. We will consider two-parameter families  $\dot{u} = f(u, \mu)$  with  $\mu = (\mu_1, \mu_2)$  to unfold a homoclinic orbit at resonance.

**HYPOTHESIS 5.6 (Resonance Condition).** *We assume that the leading eigenvalues satisfy  $v^s(0) = v^u(0)$  and  $\partial_{\mu_2} v^s(0) \neq \partial_{\mu_2} v^u(0)$ .*

We shall also assume that the homoclinic orbit is not in an inclination-flip and/or an orbit-flip configuration: this implies the existence of a two-dimensional homoclinic centre manifold, as in [Section 3.4](#). It turns out that there are two cases, with different bifurcation diagrams, that depend on the orientability of the homoclinic centre manifold. Under [Hypothesis 2.4](#) there is a continuous bundle of planes  $E^c$  along the homoclinic

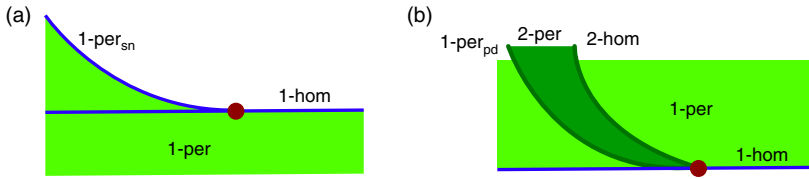


Fig. 5.4. Shown are the bifurcation diagrams at a resonant bifurcation described in [Theorem 5.13](#) for orientable (a) and non-orientable (b) homoclinic orbits.

orbit, invariant under the variational equation, so that  $\lim_{t \rightarrow \pm\infty} E^c(h(t))$  is the sum of the eigenspaces associated with the leading eigenvalues.

**HYPOTHESIS 5.7** (Generic Separatrix Value). *For  $\mu = 0$ , the separatrix quantity*

$$e^{\int_{-\infty}^{\infty} \text{div}_2(h(t)) dt} \neq 1,$$

*is not equal to one, where  $\text{div}_2$  denotes the rate of change of area within the plane field  $E^c(h(t))$ .*

**THEOREM 5.13** ([95]). *Let  $\dot{u} = f(u, \mu)$  with  $\mu = (\mu_1, \mu_2)$  be a two-parameter family of ODEs on  $\mathbb{R}^n$  with a homoclinic solution  $h(t)$  to the hyperbolic equilibrium  $p$  at  $\mu = 0$ . Suppose [Hypotheses 2.2](#), [2.3\(2\)](#), and [2.4](#) are met. Furthermore, assume that  $-v^s/v^u = 1$  at  $\mu = 0$  and that [Hypotheses 5.6](#) and [5.8](#) hold. If the homoclinic centre manifold  $W_{\text{hom}}^c$  is orientable, then the bifurcation diagram is as shown in [Figure 5.4\(a\)](#): a one-sided curve of saddle-node bifurcations of periodic orbits emerges from the curve of homoclinic orbits at  $\mu = 0$  in the parameter plane. If the homoclinic centre manifold  $W_{\text{hom}}^c$  is non-orientable, then the bifurcation diagram is as shown in [Figure 5.4\(b\)](#): a one-sided curve of period-doubling bifurcations of periodic orbits and a one-sided curve of 2-homoclinic orbits emerge from the curve of primary homoclinic orbits at  $\mu = 0$  in the parameter plane. All bifurcation curves are exponentially flat to the curve of primary homoclinic orbits at  $\mu = 0$ .*

The bifurcation equations for recurrent orbits take the form

$$x_{i+1} = \mu_1 + ax_i^{1+\mu_2} + R(\{x_i\}_{i \in \mathbb{Z}}, \mu)$$

with  $R(\{x_i\}_{i \in \mathbb{Z}}, \mu) = O(\|\{x_i\}_{i \in \mathbb{Z}}\|^{1+\eta})$  for some  $\eta > 0$ . In fact,  $|a|$  equals the separatrix quantity defined in [Hypothesis 5.7](#), and the sign of  $a$  reflects the orientation of the homoclinic centre manifold [327,328]. The bifurcation equations can be readily solved for  $N$ -homoclinic orbits and  $N$ -periodic orbits with  $N = 1, 2$ ; see also [338]. Note that the occurrence of  $N$ -homoclinic orbits and  $N$ -periodic orbits with  $N > 2$  is excluded by the existence of a two-dimensional homoclinic centre manifold.

Lorenz-like attractors can be found near differential equations with two homoclinic orbits to the same equilibrium at resonance. This has been researched by Robinson [327–329] whose results we review in [Theorem 5.80](#) in [Section 5.5.5](#). Other bifurcation phenomena that occur in resonant bifurcations from two homoclinic orbits are described in [151,191].

Higher codimension bifurcations of ODEs in the plane have been considered by Leontovich [259] and Roussarie [331]. In these studies, bounds are derived on the number of limit cycles that may appear in unfoldings. The following account follows [331]. Let  $\dot{u} = f(u, \mu)$  be a smooth family of ODEs on  $\mathbb{R}^2$  with a saddle type equilibrium for  $\mu = 0$  whose saddle quantity  $-v^s/v^u$  is equal to one. For  $\mu$  small and fixed  $\ell \in \mathbb{N}$ , the family near the equilibrium is then  $C^\ell$ -equivalent to

$$\begin{aligned}\dot{x}_s &= -x_s + \sum_{i=0}^{N(\ell)} \alpha_i(\mu) (x_s x_u)^i x_u, \\ \dot{x}_u &= x_u,\end{aligned}$$

where  $\alpha_i(\mu)$  are smooth functions of  $\mu$ . The  $C^\ell$ -equivalence is by rescaling time in an  $x$ -dependent  $C^\ell$ -smooth fashion and conjugating with diffeomorphisms of class  $C^\ell$  in  $(x_s, x_u, \mu)$ . Introduce the function

$$\omega(x, \varepsilon) = \frac{x^{-\varepsilon} - 1}{\varepsilon}. \quad (5.1)$$

Note that, for each  $k > 0$ ,  $x^k \omega$  tends to  $-x^k \ln(x)$  as  $\varepsilon \rightarrow 0$ , uniformly for  $x \in [0, X]$  for any  $X > 0$ . A Dulac map is a local transition map from a cross section transverse to the local stable manifold, to a cross section transverse to the local unstable manifold. We may assume that these cross sections are given by  $\{x_s = 1\}$  and  $\{x_u = 1\}$ . The Dulac map  $D(x, \mu)$  now has the following expansion:

$$\begin{aligned}D(x, \mu) &= x + \alpha_1(x\omega + \cdots) + \alpha_2(x^2\omega + \cdots) + \cdots \\ &\quad + \alpha_{N+1}(x^{N+1}\omega + \cdots) + \psi_\ell(x, \mu),\end{aligned} \quad (5.2)$$

for a  $C^\ell$  function  $\psi_\ell$  of  $(x, \mu)$  which is  $\ell$ -flat for  $x = 0$ . Each function between brackets is a finite combination of terms  $x^i \omega^j$  with  $0 \leq j \leq i$  in an increasing order ( $x^i \omega^j < x^{i'} \omega^{j'}$  precisely if  $i < i'$  or  $i = i'$  and  $j < j'$ ). For  $\mu = 0$ , the Dulac map is equivalent to either  $x \mapsto \beta_k x^k$  or  $x \mapsto \alpha_{k+1} x^{k+1} \ln(x)$ . Using this expansion, the following estimate on the number of limit cycles that appear in an unfolding from the homoclinic loop has been proved.

**THEOREM 5.14 ([331]).** *Let  $\dot{u} = f(u, \mu)$  be a family of ODEs on the plane with a homoclinic loop to the hyperbolic equilibrium for  $\mu = 0$ . Suppose that the saddle quantity  $v^s/v^u$  is equal to one at  $\mu = 0$  and that the Dulac map for  $\mu = 0$  is not flat. If the Dulac map for  $\mu = 0$  is equivalent to  $x \mapsto \beta_k x^k$ , then  $\dot{u} = f(u, \mu)$  has at most  $2k$  limit cycles for small  $\mu$ . If the Dulac map for  $\mu = 0$  is equivalent to  $x \mapsto \alpha_{k+1} x^{k+1} \ln(x)$ , then  $\dot{u} = f(u, \mu)$  has at most  $2k + 1$  limit cycles for small  $\mu$ .*

We remark that, for analytic ODEs on the plane, the Dulac map is known to be nonflat [205], so that a uniformly bounded number of limit cycles appears in the unfolding of a resonant homoclinic loop.

A homoclinic centre manifold reduces the general case of ODEs in  $\mathbb{R}^n$  to a two-dimensional flow. As the degree of differentiability of the homoclinic centre manifold is limited by gap conditions on the spectrum at the equilibrium, this reduction is of limited use for the study of bifurcations of periodic orbits. Bifurcations of higher codimension for higher-dimensional flows are considered in [169,170,333], and we remark that more complicated dynamics may occur: as stated in the next theorem, Turaev found an example of codimension three in which infinitely many limit cycles appear under perturbations in the non-planar case.

**THEOREM 5.15 ([402]).** *Let  $\dot{u} = f(u)$  be an ODE on  $\mathbb{R}^n$  for  $n \geq 4$  with a homoclinic solution  $h(t)$  to the hyperbolic equilibrium  $p$ . Suppose that [Hypotheses 2.2\(1\)](#), [2.3\(2\)](#), and [2.4](#) are met and that  $-v^s/v^u = 1$ . Assume, furthermore, that the separatrix quantity is equal to one and that the stable eigenvalue closest to  $v^s$  is complex. Under generic conditions, the homoclinic orbit is then the limit of infinitely many isolated periodic orbits.*

**5.1.6. Inclination-flips.** Yanagida [428] realized that 2-homoclinic solutions might appear in the unfolding of codimension-two bifurcations of homoclinic orbits to equilibria with real leading eigenvalues. He considered three different scenarios that are related to the three nondegeneracy conditions that we introduced in [Hypotheses 2.3](#) and [2.4](#). The first of these scenarios, homoclinic orbits at resonance, has been discussed in [Section 5.1.5](#). The remaining two bifurcations, which are concerned with homoclinic orbits that are in inclination-flip or orbit-flip configurations, are discussed in this and the following section. Depending on eigenvalue conditions, 2-homoclinic orbits may appear in the unfolding, or complicated dynamics may set in, involving  $N$ -homoclinic orbits for all  $N$  and suspended Hénon-like attractors.

We consider two-parameter families  $\dot{u} = f(u, \mu)$  of three-dimensional vector fields with  $(u, \mu) \in \mathbb{R}^3 \times \mathbb{R}^2$ . Throughout, we assume that the eigenvalues of  $f_u(p, 0)$  at the equilibrium  $p$  satisfy  $v^{ss} < v^s < 0 < v^u$  and that  $h(t)$  is an orbit homoclinic to  $p$  for  $\mu = 0$ .

**HYPOTHESIS 5.8 (Inclination Flip).** *Using the notation from [Hypothesis 2.4](#), we assume that  $v^s, v^u, v_*^u \neq 0$  at  $\mu = 0$ . Furthermore, we assume that  $v_*^s = 0$  at  $\mu = 0$  with  $\partial_{\mu_2} v_*^s(\mu)|_{\mu=0} \neq 0$ .*

As outlined in [Section 2.1](#), the orientation of the two-dimensional homoclinic centre manifold changes at an inclination flip. Define

$$\alpha = -v^{ss}/v^u, \quad \beta = -v^s/v^u$$

and observe  $\alpha > \beta > 0$ . We distinguish the following three cases:

Type A:  $\beta > 1$ ;

Type B:  $\alpha > 1$  and  $\frac{1}{2} < \beta < 1$ ;

Type C:  $\alpha < 1$  or  $\beta < \frac{1}{2}$ ,

and impose the following additional nondegeneracy condition for type C bifurcations:

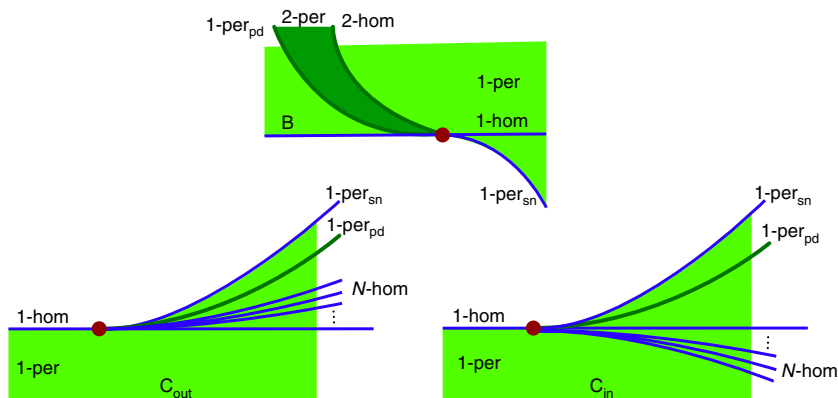


Fig. 5.5. The bifurcation diagrams at inclination- and orbit-flip bifurcations are shown for homoclinic-doubling bifurcations (type B) and the more involved type C bifurcations. For type C bifurcations, there are infinitely many branches of  $N$ -homoclinic orbits for each fixed  $N \geq 2$ .

**HYPOTHESIS 5.9** (Nondegeneracy Conditions for Type C). *For type C, we assume that*

- (1)  $\beta \neq \frac{1}{2}\alpha$ .
- (2) If  $\beta < \frac{1}{2}\alpha$ , the homoclinic orbit does not lie in the unique smooth leading stable manifold  $W^{ls}(p)$ .
- (3) If  $\beta > \frac{1}{2}\alpha$ ,  $W^{ls,u}(p)$  has a nondegenerate quadratic tangency with  $W^s(p)$  along the homoclinic orbit.

To expand on the above hypothesis: If  $\beta < \frac{1}{2}\alpha$ , a typical one-dimensional leading stable manifold is  $C^1$  but not  $C^2$ , but there exists a unique leading stable manifold that is smooth. If  $\beta > \frac{1}{2}\alpha$ , then  $W^{ls,u}(p)$  is a  $C^2$  manifold, and Hypothesis 5.9(3) is well defined.

For inclination-flips of type A, we refer to Theorem 5.1 in Section 5.1.1: a single periodic solution is created when crossing the curve of homoclinic orbits in the parameter plane; this result is also true in  $\mathbb{R}^n$  and does not require that  $\nu^s$  is real. The bifurcation diagrams in the remaining cases are shown in Figure 5.5. The unfolding for case B was treated in [222], see also [195,223,370,373,376], for vector fields in  $\mathbb{R}^n$ : it leads to homoclinic doubling, and the result holds also when  $\nu^{ss}$  is not simple or complex (in this case,  $\alpha$  is defined using  $\text{Re } \nu^{ss}$  in place of  $\nu^{ss}$ ).

**THEOREM 5.16.** *Assume that Hypotheses 2.2, 2.3(2) and 5.8 are met. Suppose further, that the inclination-flip is of type B, then the bifurcation diagram is as shown in Figure 5.5 with one-sided curves of saddle-node and period-doubling bifurcations of periodic orbits and a one-sided curve of 2-homoclinic orbits that emerge from the inclination-flip point at  $\mu = 0$  on the branch of primary homoclinic orbits.*

Finally, the unfolding for case C gives rise to  $N$ -homoclinic orbits for all  $N$  that are created through the unfolding of a singular horseshoe [194] (see also Section 4.3): There are two cases that differ in the global geometry of the stable and unstable manifolds of the equilibrium, and the existing proofs are limited to three-dimensional vector fields.

**THEOREM 5.17.** *Assume that [Hypotheses 2.2](#), [2.3\(2\)](#) and [5.8](#) are met. Suppose further that the inclination-flip is of type C and that [Hypothesis 5.9](#) is met. Depending on a global condition on the stable and unstable manifolds, the bifurcation diagram is then given by one of the two cases shown in [Figure 5.5](#). In either case, infinitely many one-sided curves of  $N$ -homoclinic orbits emerge for each  $N \geq 2$  from the inclination-flip point at  $\mu = 0$  on the branch of primary homoclinic orbits.*

Naudot [294] proved the existence of suspended Hénon-like attractors in the unfolding of type-C inclination-flip homoclinic orbits in  $\mathbb{R}^3$ . Results on inclination-flips in  $\mathbb{Z}_2$ -equivariant ODEs, in which Lorenz-like strange attractors appear, can be found in [Theorems 5.61](#) and [5.81](#) below. Applications in which inclination-flips appear include travelling waves in the FitzHugh–Nagumo equation [245], 1 : 2 spatial resonances in systems with broken  $O(2)$  symmetry [317], and models for instabilities in thermal convection [298].

**5.1.7. Orbit-flips.** At an orbit-flip bifurcation, the homoclinic orbit approaches the equilibrium along a strong stable or strong unstable direction. The most comprehensive study of the homoclinic orbit-flip bifurcation can be found in [338], and the bifurcation diagrams closely resemble those at inclination-flips. Geometrically, the orientation of the two-dimensional homoclinic centre manifold changes at an orbit-flip; see [Section 2.1](#).

We consider two-parameter families  $\dot{u} = f(u, \mu)$  of vector fields with  $(u, \mu) \in \mathbb{R}^n \times \mathbb{R}^2$ . Throughout, we assume that the leading eigenvalues of  $f_u(p, 0)$  at the hyperbolic equilibrium  $p$  are real and simple, and that  $h(t)$  is a homoclinic orbit to  $p$  for  $\mu = 0$ . We assume an orbit-flip configuration within the stable manifold  $W^s(p)$ .

**HYPOTHESIS 5.10 (Orbit-Flip).** *We assume that  $v_*^s, v_*^u, v_*^u \neq 0$  at  $\mu = 0$ , while  $v^s(0) = 0$  and  $\partial_{\mu_2} v^s(0) \neq 0$ .*

We write  $\lambda^{ss}$  for the largest real part of the stable eigenvalues that lie strictly to the left of the leading eigenvalue  $v^s$ , and define

$$\alpha = -\lambda^{ss}/v^u, \quad \beta = -v^s/v^u$$

so that  $\alpha > \beta > 0$ . We distinguish the cases

- Type A:  $\beta > 1$ ;
- Type B:  $\beta < 1$  and  $\alpha > 1$ ;
- Type C:  $\alpha < 1$ .

If the eigenvalues are of type C, we need the following additional genericity assumption, which is similar to [Hypothesis 5.9\(2\)](#):

**HYPOTHESIS 5.11 (Nondegeneracy Condition for Type C).** *There exists a unique eigenvalue  $v^{ss}$  of  $f_u(p, 0)$  with  $\operatorname{Re} v^{ss} = \lambda^{ss}$  (hence,  $v^{ss} = \lambda^{ss}$  is real and simple), and the homoclinic orbit  $h(t)$  satisfies  $\lim_{t \rightarrow \infty} e^{-v^{ss}t} h(t) \neq 0$ .*

For orbit-flips of type A, we refer to [Theorem 5.1](#) in [Section 5.1.1](#): a single periodic solution is created when crossing the curve of homoclinic orbits in the parameter plane.

Results for orbit-flips of type B and C are stated in the following two theorems and summarized in Figure 5.5.

**THEOREM 5.18** ([338]). *Assume that Hypotheses 2.2, 2.3(2) and 5.10 are met. If the orbit-flip is of type B, then the bifurcation diagram is as shown in Figure 5.5 with one-sided curves of saddle-node and period-doubling bifurcations of periodic orbits and a one-sided curve of 2-homoclinic orbits that emerge from the inclination-flip point at  $\mu = 0$  on the branch of primary homoclinic orbits.*

**THEOREM 5.19** ([338]). *Assume that Hypotheses 2.2, 2.3(2) and 5.10 are met. If the orbit-flip is of type C and that Hypothesis 5.11 is met. Depending on a global condition on the stable and unstable manifolds, the bifurcation diagram is then given by one of the two cases shown in Figure 5.5. In particular, infinitely many one-sided curves of  $N$ -homoclinic orbits emerge for each  $N \geq 2$  from the inclination-flip point at  $\mu = 0$  on the branch of primary homoclinic orbits.*

In [284], the occurrence of inclination-flips in perturbations from orbit-flips is discussed. We refer to Section 5.3.6 for orbit-flip bifurcations in conservative or reversible ODEs.

**5.1.8. Coexisting homoclinic orbits.** In this section, we review the unfolding of two-parameter families  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^n$  with a hyperbolic equilibrium at  $p$ . We assume that there are two homoclinic orbits  $h_0$  and  $h_1$  to  $p$  for  $\mu = 0$ .

Periodic orbits near the two coexisting homoclinic orbits can often be described completely by symbolic codes. Let  $q(t)$  be a periodic orbit that bifurcates in a gluing bifurcation, then we define its itinerary  $\chi(q)$  by listing the index  $i$  of the unstable separatrix  $h_i$  it follows in consecutive loops. More precisely, pick two cross sections  $\Sigma_0, \Sigma_1$  transverse to  $h_0$  and  $h_1$ , respectively, and list the sequence of indices 0, 1 that corresponds to which section the periodic orbit intersects at each return: this sequence is denoted by  $\chi(q)$ .

If the eigenvalue condition  $-\operatorname{Re} v^s/v^u > 1$  is met, the bifurcation involving two homoclinic orbits is often referred to as a gluing bifurcation. It turns out that the symbolic codes of periodic orbits that can be created in gluing bifurcations and the codes of periodic points for rigid rotations on the circle are closely related. Consider a circle rotation  $R_\alpha(x) = x + \alpha \bmod 1$  on  $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ . We divide the circle in two intervals  $I_0 = [0, \alpha)$  and  $I_1 = [\alpha, 1)$  and introduce a symbolic itinerary  $\chi_\alpha \in \{0, 1\}^{\mathbb{Z}}$  via  $\chi_\alpha(i) = j$  if  $R_\alpha^i \in I_j$ . Any sequence that occurs as an itinerary for some  $\alpha$  is called a rotation compatible sequence. A given rotation compatible sequence defines  $\alpha$  uniquely as the frequency with which the symbol zero occurs in it: in other words,  $\alpha$  is the *rotation number* of the itinerary.

**THEOREM 5.20** ([142, 143]). *We assume that the unstable manifold of  $p$  is one-dimensional and that there are two homoclinic orbits  $h_0$  and  $h_1$  to  $p$  for  $\mu = 0$ . Furthermore, suppose that  $-\operatorname{Re} v^s/v^u > 1$ . The itinerary of any periodic orbit created for nearby parameter values is then necessarily rotation compatible. Furthermore, each bifurcating periodic orbit is asymptotically stable, and there are at most two periodic orbits*

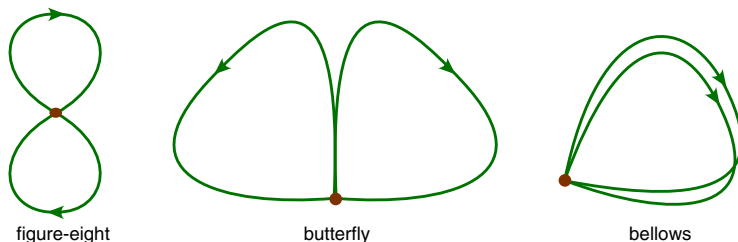


Fig. 5.6. Figure-eight, butterfly, and bellows configurations of two coexisting homoclinic orbits are illustrated (note that bellows can only exist in dimension four or greater).

for each given parameter value. If two periodic orbits exist for the same parameter value, then the rotation numbers of their itineraries are Farey neighbours.<sup>6</sup>

In the following, we assume that [Hypothesis 2.3\(2\)](#) holds so that the leading eigenvalues  $v^s, v^u$  at  $p$  are unique and real. Recall the definition of the vectors  $v_i^s(\mu)$  and  $v_i^u(\mu)$  with  $i = 0, 1$  from (2.9), where the subscript  $i$  indicates that these vectors are computed for the homoclinic orbit  $h_i$ . Different geometric configurations can now be distinguished as follows; see also [Figure 5.6](#).

**HYPOTHESIS 5.12 (Geometric Configurations).** Assume that  $p$  has unique real leading eigenvalues  $v^s, v^u$ .

- (1) *Figure eight*:  $v_0^s(0) = -v_1^s(0)$ ,  $v_0^u(0) = -v_1^u(0)$ .
- (2) *Expanding butterfly*:  $-v^s/v^u < 1$  and  $v_0^s(0) = v_1^s(0)$ ,  $v_0^u(0) = -v_1^u(0)$ .
- (3) *Contracting butterfly*:  $-v^s/v^u > 1$  and  $v_0^s(0) = v_1^s(0)$ ,  $v_0^u(0) = -v_1^u(0)$ .
- (4) *Bellows*:  $v_0^s(0) = v_1^s(0)$ ,  $v_0^u(0) = v_1^u(0)$ .

Bifurcations from two coexisting homoclinic solutions have been investigated by Turaev [399,403]. Near differential equations with two homoclinic orbits in the expanding butterfly configuration or the bellows configuration, one finds differential equations with suspended horseshoes (see [Theorem 5.79](#) in [Section 5.5.5](#)).

We now present the bifurcation diagrams of the remaining cases, where the flow remains Morse–Smale outside bifurcation curves, and refer to [376] for a more detailed description, including symbolic codings of orbits. We shall refer to a geometric configuration of two homoclinic orbits as orientable, semi-orientable, or non-orientable if, respectively, both, one, or none of the homoclinic centre manifolds  $W^c(h_i)$  with  $i = 1, 2$  are annuli. If a primary homoclinic orbit admits an orientable homoclinic centre manifold (and thus has a nonempty stable set), then one encounters the intermittency phenomenon surveyed in [Section 4.6](#) along any curve in a parameter plane that crosses the branch of primary homoclinic orbits transversally.

**THEOREM 5.21 ([399,403]).** Assume that the equilibrium  $p$  is hyperbolic and that [Hypothesis 2.3\(2\)](#) is met with  $-v^s/v^u > 1$ . Suppose furthermore that there are

<sup>6</sup>Recall that two rational numbers  $p/q$  and  $r/s$  are Farey neighbours if  $|ps - qr| = 1$ .



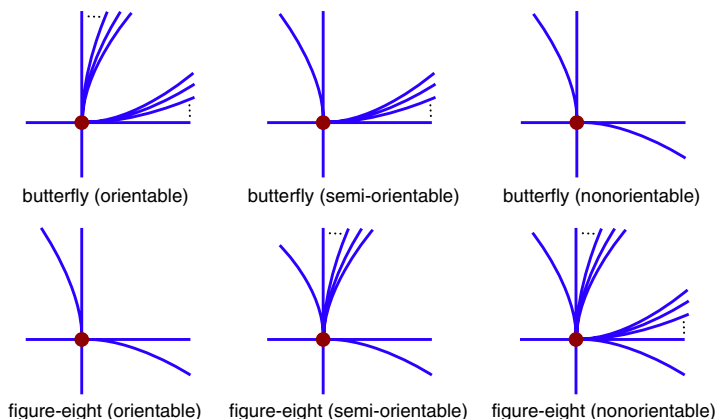


Fig. 5.7. Curves of homoclinic bifurcations in unfoldings of flows with two coexisting homoclinic solutions to an equilibrium with real leading eigenvalues: single one-sided curves correspond to 2-round homoclinic orbits, while sequences of one-sided curves correspond to multi-round homoclinic orbits.

distinct homoclinic orbits  $h_i(t)$  to  $p$  for  $i = 1, 2$  which exist when  $\mu_i = 0$ , satisfy [Hypotheses 2.2\(1\) and 2.4](#), and are unfolded generically with respect to the parameter  $\mu_i$ . The bifurcation diagrams for the figure-eight configuration ([Hypothesis 5.12\(1\)](#)) and for the contracting butterfly ([Hypothesis 5.12\(3\)](#)) are then as shown in [Figure 5.7](#). Homoclinic orbits with arbitrarily large arclength are found in the orientable and semi-orientable butterfly and the non-orientable figure-eight. For each parameter value, there exist at most two periodic orbits.

The next theorem treats the case when the leading stable eigenvalues are complex and the homoclinic orbits are tame. See also [\[185\]](#) for a description of dynamics near two tame saddle-focus homoclinic orbits.

**THEOREM 5.22** ([\[357,359\]](#)). Assume that the equilibrium  $p$  is hyperbolic and that [Hypothesis 2.3\(3\)](#) is met with  $-\operatorname{Re} v^s/v^u > 1$ . Suppose furthermore that there are distinct homoclinic orbits  $h_i(t)$  to  $p$  for  $i = 1, 2$  which exist when  $\mu_i = 0$ , satisfy [Hypotheses 2.2\(1\) and 2.4](#), and are unfolded generically with respect to the parameter  $\mu_i$ . The bifurcation diagram is then as shown in [Figure 5.8](#) and, for each parameter value, there are at most two periodic orbits.

Due to the moduli for topological equivalence (see [Theorem 5.5](#)), the bifurcation diagrams for different values of the saddle quantity  $-\operatorname{Re} v^s/v^u$  differ from each other.

**5.1.9. Degenerate homoclinic orbits.** A homoclinic orbit to a hyperbolic equilibrium  $p$  is called degenerate if the tangent spaces of the stable and unstable manifolds  $W^s(p)$  and  $W^u(p)$  along the homoclinic orbit  $h(t)$  have more than the vector field direction in common. Typically, this defines a homoclinic bifurcation of codimension three at which  $T_{h(0)}W^s(p) \cap T_{h(0)}W^u(p)$  is two-dimensional. We consider an equation  $\dot{u} = f(u, \mu)$  in  $\mathbb{R}^n$  with  $\mu \in \mathbb{R}^3$  for which a homoclinic orbit  $h(t)$  to a hyperbolic equilibrium  $p$  exists at  $\mu = 0$ .

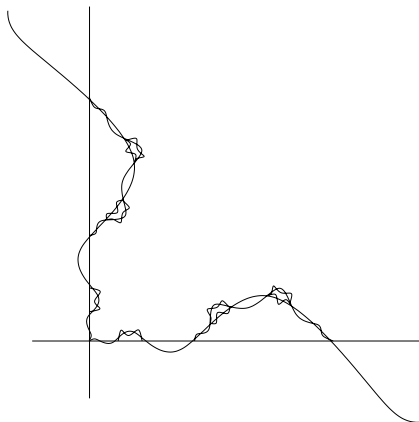


Fig. 5.8. Curves of homoclinic bifurcations in unfoldings of flows with two coexisting tame saddle-focus homoclinic solutions.

**HYPOTHESIS 5.13.** Assume that  $T_{h(0)}W^s(p, 0) \cap T_{h(0)}W^u(p, 0)$  is two-dimensional and that the stable and unstable manifolds  $W^s(p, \mu)$  and  $W^u(p, \mu)$  of  $p$  for  $\dot{u} = f(u, \mu)$  intersect transversally along  $h \times \{0\}$  in the product  $\mathbb{R}^n \times \mathbb{R}^3$  of state and parameter space.

Note that this requires  $n \geq 4$ . We have the following result.

**THEOREM 5.23 ([411]).** Suppose that  $h$  is a homoclinic orbit to the hyperbolic equilibrium  $p$  and assume that *Hypothesis 5.13* is met, then the set of parameter values for which a single-round homoclinic orbit exists, forms a Whitney umbrella in parameter space.

To visualize the geometry, consider a degenerate homoclinic orbit  $h(t)$  in  $\mathbb{R}^4$  to an equilibrium with stable and unstable manifolds that are two-dimensional. Take a cross section  $\Sigma$  transverse to  $h(t)$  at  $t = 0$ , then  $W^s(p, \mu)$  and  $W^u(p, \mu)$  intersect  $\Sigma$  along curves  $W_\Sigma^s(p, \mu)$  and  $W_\Sigma^u(p, \mu)$ . Suppose that these curves intersect in a single point when  $\mu = 0$  and have a quadratic tangency at this point then, in a generic three-parameter family, the set of parameter values for which  $W_\Sigma^s(p, \mu)$  and  $W_\Sigma^u(p, \mu)$  intersect forms a Whitney umbrella.

**5.1.10. Homoclinic orbits to nonhyperbolic equilibria.** In this section, we discuss homoclinic orbits to nonhyperbolic equilibria and restrict ourselves to bifurcations of overall codimension at most two. In particular, we consider only local bifurcations of codimension at most two and remark that homoclinic orbits to nonhyperbolic equilibria can be robust under perturbations that do not unfold the local bifurcation at the equilibrium: examples are provided by homoclinic orbits to saddle-node (codimension one) or Hopf/saddle-node (codimension two) equilibria. Additional degeneracies of the homoclinic orbit may, of course, further increase the codimension. We do not consider homoclinic orbits to pitchfork equilibria, which have been discussed in [106].

We start with homoclinic orbits to saddle-node equilibria. Consider a one-parameter family  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^n$  unfolding a saddle-node equilibrium  $p$  for  $\mu = 0$ . The geometry

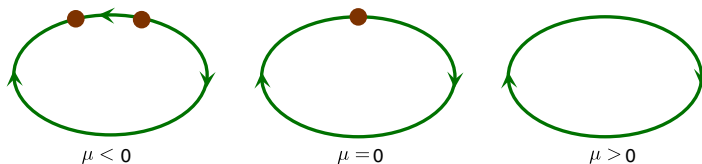


Fig. 5.9. The unfolding of a generic homoclinic orbit to a saddle-node equilibrium as described by Theorem 5.24.

of the flow near  $p$  is clarified in Section 2.2, and we assume that Hypothesis 2.5 is met. Thus, the dimensions of the stable and unstable sets  $\mathcal{M}^s(p)$  and  $\mathcal{M}^u(p)$  of  $p$  add up to  $n + 1$ : these sets can therefore intersect transversally to give an isolated homoclinic orbit.

**THEOREM 5.24** ([18,365]). *In the above setup, assume that Hypotheses 2.5 and 2.7 are met, then the homoclinic orbit and the equilibrium  $p$  form a normally hyperbolic set diffeomorphic to a circle, which therefore persists for  $\mu$  near zero.*

On the persistent invariant circle, the dynamics changes from two hyperbolic equilibria with heteroclinic connections between them for  $\mu < 0$ , say, to a single periodic orbit for  $\mu > 0$ , with the homoclinic loop to the saddle-node equilibrium forming the boundary at  $\mu = 0$ ; see Figure 5.9.

If the sets  $\mathcal{M}^s(p)$  and  $\mathcal{M}^u(p)$  intersect simultaneously and transversally at several distinct homoclinic orbits, or have a tangency at a single homoclinic orbit, then the resulting ODE can lie on the boundary of Morse–Smale systems; see Section 4.4 for details. In particular, coexisting transverse intersections of  $\mathcal{M}^s(p)$  and  $\mathcal{M}^u(p)$  lead to the creation of suspended horseshoes in an appropriate unfolding; see Theorem 4.5.

Next, we consider a codimension-two bifurcation of homoclinic orbits to saddle-node equilibria where Hypothesis 2.7 is violated. By reversing the direction of time if necessary, we may assume that Hypothesis 2.7(1) is violated. Thus, consider a two-parameter family  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^n$  with a saddle-node equilibrium  $p$  at  $\mu_1 = 0$ . We need the following unfolding condition for the homoclinic connection when varying  $\mu_2$ .

**HYPOTHESIS 5.14.** *Upon setting  $\mu_1 = 0$  and varying only  $\mu_2$ , the manifolds  $W^{cu}(p, \mu)$  and  $W^s(p, \mu)$  intersect transversally in  $\mathbb{R}^n \times \mathbb{R}$  along  $h \times \{0\}$ .*

**THEOREM 5.25** ([98,106,266,350]). *In the above two-parameter setting, assume that  $p$  is a saddle-node equilibrium when  $\mu_1 = 0$ . We assume that Hypothesis 2.5 is met if  $\mu_1$  is varied and that there is a homoclinic orbit  $h(t)$  at  $\mu = 0$  for which Hypothesis 2.7(2) is met but Hypothesis 2.7(1) is violated. Suppose further that the unfolding condition Hypothesis 5.14 is met. The bifurcation diagram then contains a one-sided curve that emerges from  $\mu = 0$  tangent to the  $\mu_2$ -axis along which a homoclinic orbit to a hyperbolic equilibrium exists. Heteroclinic connections and periodic orbits bifurcate also; see Figure 5.10.*

The codimension-three bifurcation of a homoclinic orbit to a Bogdanov–Takens equilibrium<sup>7</sup> has been treated in [128]. As the unfolding of a local Bogdanov–Takens

<sup>7</sup>The local bifurcation of codimension two with two non-semisimple eigenvalues at zero; see Section 5.4.1.

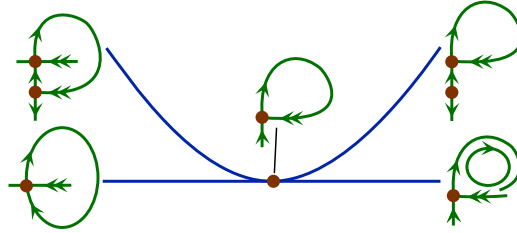


Fig. 5.10. The unfolding of a nongeneric homoclinic orbit to a saddle-node equilibrium as described by Theorem 5.25.

bifurcation gives rise to small homoclinic orbits, both small and large homoclinic orbits occur here. A codimension-three bifurcation of a homoclinic orbit to a saddle-node equilibrium in which both Hypotheses 2.7(1) and 2.7(2) are violated occurs in a model for solitary pulses in an excitable reaction-diffusion medium [433]: this bifurcation is dynamically more complicated and leads to inclination-flip homoclinic orbits and cascades of T-point bifurcations [198].

Finally, we consider homoclinic loops to an equilibrium that undergoes a supercritical Hopf bifurcation. This bifurcation, which is commonly referred to as a Shil'nikov–Hopf bifurcation, has codimension two, and we therefore treat two-parameter families  $\dot{u} = f(u, \mu)$ . Suppose that the Hopf bifurcation of the equilibrium  $p$  occurs at  $\mu_1 = 0$  and unfolds generically in the parameter  $\mu_1$  so that Hypothesis 2.8 is met. Assume also a generic unfolding that breaks the intersection of the centre-stable with the unstable manifold of  $p$  when varying  $\mu_2$  for  $\mu_1 = 0$ .

**HYPOTHESIS 5.15 (Generic Unfolding).** *For  $\mu_1 = 0$ , the intersection of  $W^{cs}(p, \mu)$  and  $W^u(p, \mu)$  unfolds generically with respect to the parameter  $\mu_2$ .*

Shil'nikov–Hopf bifurcations lead to complicated dynamics such as suspended horseshoes via two different mechanisms: first, their unfolding may contain homoclinic tangencies of the stable and unstable manifolds of the small periodic orbits that are created in the Hopf bifurcation; secondly, wild saddle-focus homoclinic orbits are created at such bifurcations. Hirschberg and Knobloch [182] considered three-dimensional flows and study first-return maps, while Deng and Sakamoto [111] derived bifurcation equations in higher-dimensional systems. In higher-dimensional systems, additional conditions, such as a nondegeneracy condition akin to Hypothesis 2.2(1), and the absence of inclination-flip and orbit-flip configurations akin to Hypothesis 2.4, are needed. We formulate here a bifurcation result in  $\mathbb{R}^3$  and refer to [111] for the higher-dimensional results.

**THEOREM 5.26 ([111,182]).** *Suppose  $\dot{u} = f(u, \mu)$  is a two-parameter family in  $\mathbb{R}^3$  for which a homoclinic orbit to a Hopf equilibrium exists when  $\mu = 0$ . If Hypotheses 2.8(3) and 5.15 are met, then there is a one-sided curve in parameter space that emerges from  $\mu = 0$  and is transverse to the curve of Hopf bifurcations along which a generic wild saddle-focus homoclinic orbit exists. Furthermore, there is a curve tangent to the curve of Hopf bifurcations along which homoclinic tangencies of the stable and*

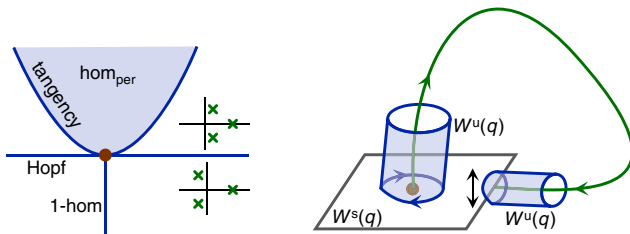


Fig. 5.11. The unfolding of a Shil'nikov–Hopf homoclinic orbit described in Theorem 5.26 is sketched in the left panel, where the insets show the eigenvalues at the equilibrium: a wild saddle-focus homoclinic orbit exists along the curve 1-hom, while the stable and unstable manifolds of the bifurcating small periodic orbits  $q$  have precisely two transverse intersections above the parabola, which disappear through a tangency along the parabola. The mechanism that creates and destroys transverse homoclinic orbits of the small periodic orbits  $q$  for fixed  $\mu_2 > 0$  as  $\mu_1$  varies, is illustrated in the right panel, where the arrows indicate how the unstable manifolds of equilibrium and periodic orbits move as  $\mu_1$  varies.

*unstable manifolds of the small periodic orbits created in the Hopf bifurcation occur; see Figure 5.11.*

We finish with a few remarks on the interaction of homoclinic and Hopf/saddle-node bifurcations. Local unfoldings of Hopf/saddle-node equilibria are discussed in Section 5.4.2, and we remark that small Shil'nikov homoclinic orbits occur in certain cases of the local unfolding. With a global homoclinic orbit present, global Shil'nikov homoclinic orbits will also occur. These bifurcations of a homoclinic orbit to an equilibrium at a Hopf/saddle-node bifurcation arise in models of semiconductor lasers with optical injection [238,434], and we refer also to [239] for a numerical investigation of such global Shil'nikov homoclinic orbits with one or more global excursions.

## 5.2. Heteroclinic cycles in generic systems

In this section, we discuss primarily bifurcations from heteroclinic cycles that connect different equilibria, though cycles that involve a periodic orbit instead of an equilibrium are considered briefly in Section 5.2.4. Thus, we consider the system  $\dot{u} = f(u)$  in  $\mathbb{R}^n$  with a heteroclinic cycle that consists of disjoint equilibria  $p_i$  and heteroclinic orbits  $h_i(t)$  with  $h_i(t) \in W^u(p_i) \cap W^s(p_{i+1})$  for each  $i$  with  $1 \leq i \leq \ell$ , where indices are taken modulo  $\ell$ ; see Section 2.3. We will mostly consider heteroclinic cycles with two heteroclinic orbits as the codimension gets too large otherwise. We recall that the Morse index of an equilibrium  $p$  is defined via  $\text{ind}(p) := \dim W^u(p)$ .

**5.2.1. Heteroclinic cycles with saddles of identical Morse index.** We discuss heteroclinic cycles built from codimension-one heteroclinic connections that occur when the equilibria involved have the same Morse index. One observation is that homoclinic orbits can appear in unfoldings of heteroclinic bifurcations. A prototype result is Theorem 5.27 below by Chow, Deng, and Terman. We start with results on heteroclinic cycles between two equilibria with real leading eigenvalues. Bifurcations that involve

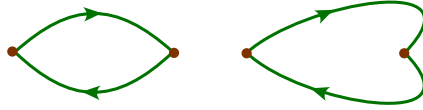


Fig. 5.12. Planar heteroclinic cycles can be nontwisted [left] or doubly twisted [right].

equilibria with complex conjugate leading eigenvalues exhibit more complicated features and will be discussed briefly afterwards.

Let  $\dot{u} = f(u, \mu)$  be a two-parameter family in  $\mathbb{R}^n$  that has a heteroclinic cycle for  $\mu = 0$  with heteroclinic orbits  $h_1$  and  $h_2$  that connect the hyperbolic equilibria  $p_1$  and  $p_2$  with identical Morse index: thus, we assume that [Hypothesis 2.10\(1\)](#) is met at  $\mu = 0$  with codimension  $d_i = 1$  for  $i = 1, 2$ . Furthermore, we assume that the leading eigenvalues at both equilibria are unique and real ([Hypothesis 2.11\(2\)](#)) and that [Hypothesis 2.12](#) is met along both heteroclinic orbits, so that neither is in either orbit- or inclination-flip configuration. As a consequence, there exists a two-dimensional continuous plane bundle  $E^c$  along the heteroclinic cycle  $h_1 \cup h_2 \cup p_1 \cup p_2$  that is invariant under the variational equations along  $h_1$  and  $h_2$ . The plane  $E_{p_i}^c$  is spanned by two eigenvectors  $e_i^s$  and  $e_i^u$  that belong to the leading eigenvalues and which we pick according to

$$e_i^s = \lim_{t \rightarrow \infty} \dot{h}_{i-1}(t) / \|\dot{h}_{i-1}(t)\|, \quad e_i^u = \lim_{t \rightarrow -\infty} \dot{h}_i(t) / \|\dot{h}_i(t)\|, \quad (5.3)$$

with indices taken modulo 2. We choose orientations on  $E_{p_i}^c$  so that the bases  $\{e_i^s, e_i^u\}$  are both positively oriented. By continuity, this induces an orientation on the plane bundle  $E_i^c(t)$  along  $h_i(t)$  by continuing the orientation from  $t = -\infty$ . We define the orientation index  $\mathcal{O}_i = \pm 1$  as follows:

$$\mathcal{O}_i = \begin{cases} 1 & \text{if the orientation } E_i^c(t) \text{ along } h_i(t) \\ & \text{matches the orientation of } E_{p_i}^c \text{ when } t \rightarrow \infty, \\ -1 & \text{otherwise} \end{cases} \quad (5.4)$$

and refer to [Figure 5.12](#) for an illustration.

**HYPOTHESIS 5.16 (Twist Conditions).** *Consider the following twist conditions along the heteroclinic cycle:*

- (1) *nontwisted:*  $\mathcal{O}_1, \mathcal{O}_2 = 1$ ,
- (2) *single twisted:*  $\mathcal{O}_1 \mathcal{O}_2 = -1$ ,
- (3) *doubly twisted:*  $\mathcal{O}_1, \mathcal{O}_2 = -1$ .

Note that, in the spirit of [Section 3.4](#), there exists a two-dimensional centre manifold near the heteroclinic cycle which will be orientable if the cycle is nontwisted or doubly twisted, and non-orientable otherwise.

**THEOREM 5.27 ([96,97,107,230,341]).** *Let  $\dot{u} = f(u, \mu)$  be a two-parameter family that admits a heteroclinic cycle with heteroclinic orbits  $h_1, h_2$  to two hyperbolic equilibria  $p_1, p_2$  with identical Morse index when  $\mu = 0$ .*

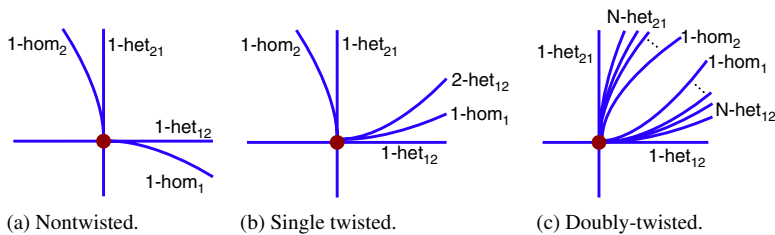


Fig. 5.13. Bifurcation diagrams for Theorem 5.27: subscripts denote to which equilibria the solution connects. In the right panel, there is a unique curve of  $N$ -heteroclinic orbits that connect  $p_1$  to  $p_2$  for each  $N \geq 2$ , and these curves accumulate onto the branch of homoclinic orbits to  $p_2$ .

- (1) Suppose that *Hypotheses* 2.10 with  $d_i = 1$ , 2.11(1) at  $p_i$ , and 2.12(2), (4) for  $h_i$  are met for  $i = 1, 2$ . In the parameter plane, there are then two curves of heteroclinic orbits that intersect at  $\mu = 0$ . Branching off  $\mu = 0$  and tangent to the curves of the heteroclinic orbit  $h_i$  are two one-sided curves of homoclinic orbits. Other solutions may bifurcate as well.
- (2) Suppose that *Hypotheses* 2.10 with  $d_i = 1$ , 2.11(2) with<sup>8</sup>  $-v_i^s/v_i^u < 1$ , and 2.12 for  $h_i$  hold for  $i = 1, 2$ . If *Hypothesis* 5.16(1) is also met, then the complete bifurcation diagram is given in Figure 5.13(a). If *Hypothesis* 5.16(2) is met, the complete bifurcation set given in Figure 5.13(b) which contains, in addition, a one-sided curve of 2-heteroclinic orbits that makes two passages near the twisted heteroclinic orbit. If *Hypothesis* 5.16(3) is met, then the complete bifurcation set, given by Figure 5.13(c), contains, for each  $N \geq 1$ , a unique one-sided curve of  $N$ -heteroclinic orbits between  $p_1$  and  $p_2$  that make  $N + 1$  passages near  $h_1$  and  $N$  near  $h_2$  as well as another branch of  $N$ -heteroclinic orbits between  $p_2$  and  $p_1$  that make  $N$  passages near  $h_1$  and  $N + 1$  near  $h_2$ .

The geometric picture is as follows. There is a two-dimensional normally hyperbolic centre manifold that contains the heteroclinic cycle for  $\mu = 0$  (see Section 3.4 and [341]). Take two cross sections  $\Sigma_1$  and  $\Sigma_2$  in the centre manifold transverse to the heteroclinic orbits  $h_1$  and  $h_2$ , respectively, then there are  $C^1$  coordinates on these cross sections in which the transition maps are of the form  $x \mapsto b_i(\mu) + a_i(\mu)x^{-v_i^s/v_i^u} + o(x^{-v_i^s/v_i^u})$  for  $x > 0$ ; see Section 3.6.3. If  $-v_i^s/v_i^u > 1$  for both  $i = 1, 2$ , then the return map restricted to the centre manifold is a contraction and, consequently, there can be at most one periodic orbit (if  $-v_i^s/v_i^u < 1$ , then the return map is an expansion, and the conclusion still holds).

The above results assumed, apart from real simple leading eigenvalues, that the saddle quantities are either both larger or both smaller than one: thus, we assumed that we are in the first of the two cases distinguished below:

**HYPOTHESIS 5.17 (Saddle Quantities).** We distinguish two cases for the saddle quantities  $\lambda_i = -v_i^s/v_i^u$ :

<sup>8</sup>We assume here that the saddle quantities are both smaller than one: if they are both larger than one, the result holds for the time-reversed system; note that this changes the definition of 1-het<sub>12</sub> to 1-het<sub>21</sub>, and vice versa, in Figure 5.13, but not that of 1-hom<sub>j</sub>.

- (1)  $\lambda_1, \lambda_2 > 1$  (the case  $\lambda_1, \lambda_2 < 1$  is brought to this case by changing the direction of time);
- (2)  $\lambda_1 \lambda_2 < 1$ .

Heteroclinic cycles under [Hypothesis 5.17\(2\)](#) have been considered by Shashkov [354–356], and we refer to these references for the bifurcation diagrams. Depending on the orientability of the primary heteroclinic orbits, the two-parameter bifurcation diagrams show, in addition to homoclinic and heteroclinic bifurcation curves, bifurcation curves of saddle-node or periodic-doubling bifurcations of periodic orbits.

Bifurcations from heteroclinic cycles where at least one of the equilibria has complex conjugate leading eigenvalues, lead to great complexity in bifurcation diagram and dynamics. Various cases have been studied in [70,71,359], and we refer to these references and also to [376] for an overview of the resulting dynamics. Apart from determining bifurcation curves in an unfolding, one can also look for prevalent dynamics as parameters are varied. We present a result in this direction due to San Martín, involving a heteroclinic cycle in  $\mathbb{R}^3$  between an equilibrium with real leading stable eigenvalues and a second equilibrium with complex conjugate stable eigenvalues.

Consider a two-parameter family  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^3$  that has a heteroclinic cycle consisting of connecting orbits  $h_1$  and  $h_2$  and hyperbolic equilibria  $p_1$  and  $p_2$  of index 1 when  $\mu = 0$ .

**HYPOTHESIS 5.18** (Heteroclinic Cycle with One Saddle Focus). *Suppose that  $p_1$  has real eigenvalues  $v_1^{ss} < \overline{v_1^s} < 0 < v_1^u$ , while  $p_2$  has a real eigenvalue  $v_2^u > 0$  and complex eigenvalues  $v_2^s, \overline{v_2^s}$  with  $\operatorname{Re} v_2^s < 0$ . For  $\mu = 0$ , there is a heteroclinic cycle with  $h_1 \in W^u(p_1) \cap W^s(p_2)$  and  $h_2 \in W^u(p_2) \cap W^s(p_1)$ .*

**HYPOTHESIS 5.19** (Expanding Heteroclinic Cycle). *The heteroclinic cycle is expanding:  $\operatorname{Re} v_2^s v_1^s < v_2^u v_1^u$ .*

The following result shows that hyperbolic dynamics is prevalent in an unfolding of an expanding heteroclinic cycle that involves a saddle focus.

**THEOREM 5.28** ([348]). *Consider a two-parameter family of ODEs on  $\mathbb{R}^3$  that satisfies [Hypotheses 5.18](#), [5.19](#) and [2.10\(2\)](#). Furthermore, assume that  $h_2$  satisfies [Hypothesis 2.12\(2\)](#), and  $h_1$  satisfies [Hypothesis 2.12\(1\)](#). We also assume the generic condition that, for  $\mu$  near 0, there are  $C^2$  linearizing coordinates near  $p_1$  and  $p_2$ . Last, we assume that the heteroclinic cycle is the locally maximal invariant set in a small open neighbourhood  $\mathcal{U}$  of itself when  $\mu = 0$ . Let  $P$  be the set of parameter values  $\mu$  near zero for which  $\dot{u} = f(u, \mu)$  has a chain recurrent set in  $\mathcal{U}$  that is equal to the two equilibria  $p_j(\mu)$  and at most one nontrivial hyperbolic basic set, then*

$$\lim_{\varepsilon \downarrow 0} \frac{|B_\varepsilon(0) \cap P|}{|B_\varepsilon(0)|} = 1,$$

where  $B_\varepsilon(0)$  is a disc of radius  $\varepsilon$  around zero, and  $|\cdot|$  denote two-dimensional Lebesgue measure.



Similarly hyperbolic dynamics is prevalent for contracting heteroclinic cycles where the inequality in [Hypothesis 5.19](#) is reversed: in this case, the single hyperbolic basic set in the definition of  $P$  is replaced by an attracting periodic orbit; see [\[348\]](#).

A central question in the study of heteroclinic cycles of planar ODEs has been to estimate its cyclicity, that is, to find bounds on the number of cycles that can appear in unfoldings.<sup>9</sup> In [Section 5.1.5](#), this problem is discussed for the unfolding of planar homoclinic loops near resonant eigenvalues; see, in particular, [Theorem 5.14](#). Here, we briefly discuss general heteroclinic cycles and begin with the flow near a hyperbolic equilibrium with eigenvalues that are not necessarily at resonance. Let  $\dot{u} = f(u, \mu)$  be a smooth family of ODEs on  $\mathbb{R}^2$  with a saddle type equilibrium at  $\mu = 0$ . For  $\mu$  small and fixed  $\ell \in \mathbb{N}$ , the vector field near the equilibrium is then  $C^\ell$ -equivalent to

$$\begin{aligned}\dot{x}_s &= v^s x_s + \sum_{i=0}^{N(\ell)} \alpha_i(\mu) (x_s x_u)^i x_u, \\ \dot{x}_u &= v^u x_u,\end{aligned}$$

where  $\alpha_i(\mu)$  are smooth functions of  $\mu$ . The  $C^\ell$ -equivalence is by rescaling time in an  $x$ -dependent  $C^\ell$ -smooth fashion and conjugating with diffeomorphisms of class  $C^\ell$  in  $(x_s, x_u, \mu)$ ; see [\[332\]](#) and use [\[53\]](#) for the parameter-dependent case. Recall that a Dulac map is a local transition map from a cross section transverse to the local stable manifold to a cross section transverse to the local unstable manifold. We may assume that these cross sections are given by  $\{x_s = 1\}$  and  $\{x_u = 1\}$ . Mourtada proves the following expansion for the Dulac map.

**THEOREM 5.29** ([\[290\]](#)). *For each fixed  $k \in \mathbb{N}$ , there is a neighbourhood  $U_k$  of  $\mu = 0$  in  $\mathbb{R}$  such that the Dulac map  $D(x, \mu)$  has the asymptotic expansion*

$$D(x, \mu) = x^{-v^s/v^u} [a(\mu) + \psi(x, \mu)]$$

for  $\mu \in U_k$ , where  $a$  is a smooth function of  $\mu$  and

$$\lim_{x \rightarrow 0} x^j \partial_x^j \psi(x, \mu) = 0$$

for each  $j \leq k$ .

A *hyperbolic polycycle* is a polycycle for a planar ODE that consists of hyperbolic equilibria and connecting heteroclinic orbits. The above expansion for the Dulac map near a hyperbolic equilibrium can now be used to bound the cyclicity of hyperbolic polycycles. Note that a return map on a cross section transverse to a hyperbolic polycycle is a composition of local diffeomorphisms and Dulac maps. Mourtada derived the following bound for the number of limit cycles that can appear in the unfolding of hyperbolic polycycles. We remark that Kaloshin [\[215\]](#) obtained an explicit bound for the cyclicity of more general polycycles, which we state as [Theorem 5.36](#) further below.

---

<sup>9</sup>This question has its origin in Hilbert's 16th problem; see [\[206\]](#) for a review.

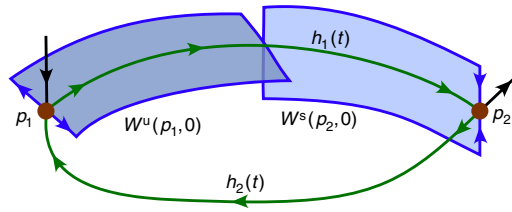


Fig. 5.14. The geometric configuration of heteroclinic orbits at a T-point is illustrated.

**THEOREM 5.30** ([290,291]). *There is a function  $C : \mathbb{N} \rightarrow \mathbb{N}$  such that the following holds. Let  $\dot{u} = f(u, \mu)$  be a generic family of planar ODEs depending on a parameter  $\mu \in \mathbb{R}^k$  that has a hyperbolic polycycle that contains  $k$  equilibria when  $\mu = 0$ , then the number of limit cycles that exist near the polycycle for small values of  $\mu$  is bounded by  $C(k)$ .*

We mention that  $C(2) = 2$ ,  $C(3) = 3$ ,  $C(4) = 5$ . The cyclicity of hyperbolic polycycles with two equilibria was studied in [87,130,292].

**5.2.2. T-points: Heteroclinic cycles with saddles of different index.** T-points are codimension-two bifurcations of heteroclinic cycles that involve equilibria with different Morse indices. We consider

$$\dot{u} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^3 \times \mathbb{R}^2 \quad (5.5)$$

and assume that  $p_i$  with  $i = 1, 2$  are hyperbolic equilibria of (5.5) for all  $\mu$ .

**HYPOTHESIS 5.20.** *The linearization  $f_u(p_1, 0)$  has simple eigenvalues  $v_1^s, v_1^u, \tilde{v}_1^u$  with  $v_1^s < 0 < \operatorname{Re} v_1^u \leq \operatorname{Re} \tilde{v}_1^u$ , while the linearization  $f_u(p_2, 0)$  has simple eigenvalues  $\tilde{v}_2^s, v_2^s, v_2^u$  with  $\operatorname{Re} \tilde{v}_2^s \leq \operatorname{Re} v_2^s < 0 < v_2^u$ . We assume that the saddle quantities*

$$\lambda_1 := -\frac{\operatorname{Re} v_1^u}{v_1^s}, \quad \lambda_2 := -\frac{v_2^u}{\operatorname{Re} v_2^s}$$

satisfy  $\lambda_j \neq 1$ .

Thus, the manifolds  $W^s(p_1, \mu)$  and  $W^u(p_2, \mu)$  are one-dimensional, while  $W^u(p_1, \mu)$  and  $W^s(p_2, \mu)$  are two-dimensional. We assume that (5.5) has a heteroclinic cycle for  $\mu = 0$  that consists of a transversely constructed heteroclinic orbit  $h_1(t) \in W^u(p_1, 0) \cap W^s(p_2, 0)$  and a codimension-two connection  $h_2(t) \in W^u(p_2, 0) \cap W^s(p_1, 0)$ , as indicated in Figure 5.14. We also need to make several genericity assumptions.

**HYPOTHESIS 5.21.** *We assume the following:*

- (1) *The heteroclinic orbit  $h_1(t) \in W^u(p_1, 0) \cap W^s(p_2, 0)$  satisfies Hypothesis 2.10 with  $d = 0$ , while the orbit  $h_2(t) \in W^u(p_2, 0) \cap W^s(p_1, 0)$  satisfies Hypothesis 2.10 with  $d = 2$ .*

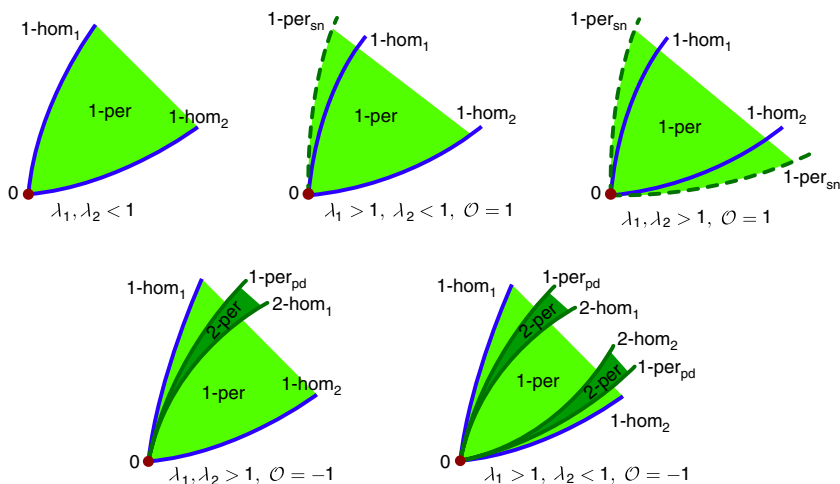


Fig. 5.15. Shown are the bifurcation diagrams in the parameter plane  $\mu \in \mathbb{R}^2$  near a generic T-point at  $\mu = 0$  that involves two hyperbolic equilibria with real and simple eigenvalues: the diagrams depend on the saddle quantities  $\lambda_j$  defined in Hypothesis 5.20 and the orientation index  $\mathcal{O}$  from Hypothesis 5.21(3).

- (2) Neither  $h_1$  nor  $h_2$  are in an orbit-flip configuration (that is,  $h_j$  obeys Hypothesis 2.12(2) and (4) for  $j = 1, 2$ ).
- (3) If the eigenvalues of  $p_1$  and  $p_2$  are all real, we assume, in addition, that the heteroclinic cycle is not in an inclination-flip configuration: more precisely, we assume that the closure of  $W^s(p_2, 0)$  is homeomorphic to a cylinder ( $\mathcal{O} := 1$ ) or to a Möbius band ( $\mathcal{O} := -1$ ); in this case, the closure of  $W^u(p_1, 0)$  is also homeomorphic to a cylinder if  $\mathcal{O} = 1$  and to a Möbius band if  $\mathcal{O} = -1$ .

The following theorem summarizes the different bifurcation diagrams at T-points when the eigenvalues of the equilibria  $p_j$  are real.

**THEOREM 5.31** ([67]). Assume that Hypotheses 5.20 and 5.21(1)–(3) are met. If the eigenvalues of both equilibria  $p_1$  and  $p_2$  are all real, then the bifurcation diagrams near  $\mu = 0$  are as shown in Figure 5.15.

Next, we consider the case where one of the equilibria (say  $p_1$ ) has complex eigenvalues, while the other equilibrium ( $p_2$ ) has only real eigenvalues.

**THEOREM 5.32** ([67, 155, 156]). Assume that Hypotheses 5.20 and 5.21(1)–(2) are met. Assume furthermore that the eigenvalues of  $p_2$  are real, while  $p_1$  is a focus so that  $\text{Im } v_1^u = -\text{Im } \tilde{v}_1^u > 0$ . The bifurcation diagram of heteroclinic orbits and 1-homoclinic orbits is shown in Figure 5.16(a). In addition, for each  $\mu$  close to zero, there are infinitely many hyperbolic periodic orbits near the heteroclinic cycle and the dynamics contains a shift of two symbols.

Finally, we consider the situation where both  $p_1$  and  $p_2$  are foci. In coordinate systems in which the vector fields near  $p_1$  and  $p_2$  are linearized, we pick a number  $T \gg 1$  and

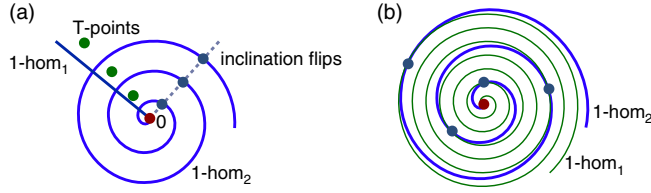


Fig. 5.16. Shown are the partial bifurcation diagrams in the parameter plane  $\mu \in \mathbb{R}^2$  near a generic T-point at  $\mu = 0$ . Panel (a) is for the case where the unstable eigenvalues of  $p_1$  are not real, while the eigenvalues of  $p_2$  are all real: There is a sequence of parameter values that converge to  $\mu = 0$  and correspond to T-points of (5.5); furthermore, the homoclinic orbits to the saddle  $p_1$  undergo a sequence of inclination-flips that accumulate at  $\mu = 0$ . Panel (b) shows the bifurcation diagram of 1-homoclinic orbits when both equilibria have complex leading eigenvalues.

place small sections at  $h_2(-T)$  near  $p_2$  and at  $h_2(T)$  near  $p_1$ . Without loss of generality, the return map  $\Pi$  along  $h_2$  that maps these section into each other satisfies

$$\Pi_u(h_2(-T)) = \begin{pmatrix} 1 & 0 \\ d & d \end{pmatrix}.$$

We set

$$\gamma := \frac{1}{2} \left( d^2 + \frac{1}{d^2} \right) \geq 1,$$

$$\Gamma := \left( \frac{\lambda_1}{\operatorname{Im} v_1^u} \right)^2 + \left( \frac{\lambda_2}{\operatorname{Im} v_2^s} \right)^2 - 2\gamma \frac{\lambda_1 \lambda_2}{|\operatorname{Im} v_1^u| |\operatorname{Im} v_2^s|} + 1 - \gamma^2.$$

We can now state the following theorem.

**THEOREM 5.33 ([70,71]).** *Assume that Hypotheses 5.20 and 5.21(1)–(2) are met. Assume furthermore that the unstable eigenvalues of  $p_1$  and the stable eigenvalues of  $p_2$  are complex and that  $\lambda_1 \operatorname{Im} v_2^s \neq \lambda_2 \operatorname{Im} v_1^u$  and  $\Gamma \neq 0$ . Under these assumptions, the following is true:*

*First, the manifolds  $W^u(p_1, 0)$  and  $W^s(p_2, 0)$  intersect transversely at infinitely many different heteroclinic orbits. If  $\Gamma < 0$ , the class of systems for which there is a heteroclinic orbit along which  $W^u(p_1, 0)$  and  $W^s(p_2, 0)$  intersect non-transversely is dense, is the class of systems with generic T-points.*

*Second, the bifurcation diagram of 1-homoclinic orbits is as shown in Figure 5.16(2): in particular, there exists an infinite sequence of parameter values that accumulate at  $\mu = 0$  so that (5.5) has a pair of coexisting homoclinic orbits to  $p_1$  and  $p_2$ : the two homoclinic bifurcation curves intersect transversely in the parameter plane when  $\Gamma > 0$ , while near each system for which  $\Gamma < 0$  there is a system that satisfies the assumptions stated above, for which one of these intersections is not transverse.*

Depending on the values of the saddle quantities  $\lambda_i$ , the dynamics of (5.5) can contain attractors and repellers near parameter values at which homoclinic orbits to the foci  $p_1$  and  $p_2$  coexist; see Section 5.1.2.

We remark that T-points arise in the Lorenz equation [155,156] as well as in systems that model the oxidation on platinum surfaces [198,432,433], the propagation of calcium waves [330,380], Josephson junctions [409] and laser systems [420].

**5.2.3. Heteroclinic cycles with nonhyperbolic equilibria.** In generic two-parameter families of ODEs, heteroclinic cycles that involve a saddle-node equilibrium and a hyperbolic equilibrium may appear. We remark that heteroclinic cycles with two saddle-node equilibria may also occur: such a cycle is of codimension two if it forms a normally hyperbolic invariant circle that contains the two saddle-node equilibria.

Planar systems with heteroclinic cycles that consist of two heteroclinic orbits, which connect a hyperbolic saddle to a saddle-node equilibrium and back, have been studied by Grozovskii [164]. Polycycles where both unstable separatrices of the hyperbolic saddle are heteroclinic orbits to the saddle-node equilibrium are also of codimension-two: they have also been treated in [164]. Homburg [189] studied general heteroclinic cycles between a hyperbolic equilibrium and a saddle-node equilibrium for ODEs in  $\mathbb{R}^n$ . We present here a single theorem for the case where the hyperbolic equilibrium has a pair of complex conjugate leading eigenvalues: for simplicity, this result is stated for three-dimensional flows; generalizations to higher dimensions involve geometric conditions akin to no-flip conditions.

**THEOREM 5.34 ([189]).** *Let  $\dot{u} = f(u, \mu)$  be a two-parameter family of ODEs on  $\mathbb{R}^3$  with  $\mu \in \mathbb{R}^2$ . Assume that a heteroclinic cycle exists for  $\mu = 0$  which connects a hyperbolic equilibrium  $p_1$  and a saddle-node equilibrium  $p_2$  so that the following assumptions are met. The eigenvalues of  $f_u(p_1, 0)$  at the hyperbolic equilibrium  $p_1$  satisfy Hypothesis 2.3(2) and are of Type B as explained in Section 5.1.6, while the saddle-node equilibrium  $p_2$  satisfies Hypotheses 2.5 and 2.6(2) with respect to  $\mu_1$ , so that  $\mu_1$  unfolds the saddle node. Furthermore, we assume that  $W^u(p_1, 0) \not\subset W^s(p_2, 0)$  and that the intersection of the stable manifold  $W^s(p_1, \mu)$  and the centre manifold  $W^c(p_2, \mu)$  is unfolded generically with respect to the parameter  $\mu_2$  when  $\mu_1 = 0$ . The bifurcation set then contains a sequence of inclination-flip homoclinic bifurcations that converge to  $\mu = 0$ . Branching from the inclination-flip bifurcation points are curves of saddle-node and period-doubling bifurcations of periodic orbits and curves corresponding to 2-homoclinic orbits to  $p_1$  as shown in Figure 5.17.*

The inclination-flip bifurcations occur as the stable manifold of  $p_1$  undergoes an arbitrary number of rotations in the vicinity of  $p_2$  for parameter values near  $\mu_1 = 0$  (recall that  $\mu_1$  unfolds the saddle-node bifurcation). The above result assumed eigenvalues of type B that lead to the homoclinic-doubling bifurcation. For type-C eigenvalues, one likewise finds a converging sequence of inclination-flip homoclinic orbits with more complicated bifurcation diagrams as in Section 5.1.6.

We conclude this section with a few remarks on heteroclinic bifurcations for planar ODEs. Kotova and Stanzo [236] compiled a list of codimension one, two and three bifurcations of heteroclinic cycles (the ‘Kotova zoo’). The interest lies in estimating the

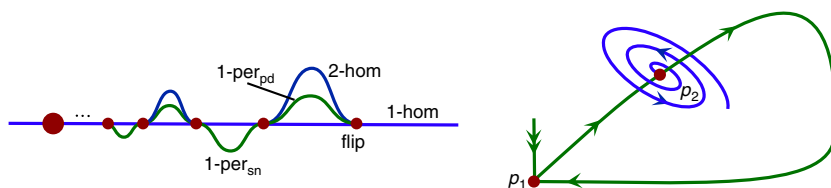


Fig. 5.17. Bifurcation set and phase-space dynamics at  $\mu = 0$  for Theorem 5.34.

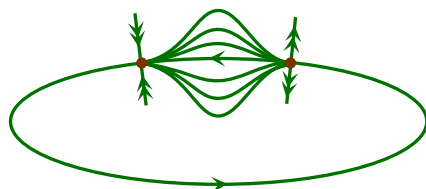


Fig. 5.18. An arbitrary number of limit cycles can occur in perturbations from a 'lips-shaped' heteroclinic cycle; see Theorem 5.35.

cyclicity, that is, the number of limit cycles that are born in an unfolding of these cycles. It turns out that there are heteroclinic cycles in three parameter families for which any number of limit cycles bifurcates.

**THEOREM 5.35 ([236]).** *For each  $N \in \mathbb{N}$ , there is a planar ODE  $\dot{u} = f(u)$  with a heteroclinic cycle that connects two saddle-node equilibria as shown in Figure 5.18 so that  $N$  limit cycles occur for an arbitrary small perturbation of  $\dot{u} = f(u)$ .*

Note that the heteroclinic cycle shown in Figure 5.18 contains a normally attracting saddle-node equilibrium  $p_-$  and a normally repelling saddle-node equilibrium  $p_+$ . To prove the preceding theorem, three-parameter families  $\dot{u} = f(u, \mu)$  are considered. The normal form near the saddle-node equilibria  $p_{\pm}$  are given by

$$\begin{aligned}\dot{x} &= \frac{x^2 + \delta_{\pm}(\mu)}{1 + a_{\pm}(\mu)} \\ \dot{y} &= \pm y,\end{aligned}$$

which can be integrated explicitly to construct local transition maps. Composing with global transition maps yields a return map that can be studied to prove the above result.

In contrast, generic  $k$ -parameter families that unfold elementary polycycles always lead to a bounded number of limit cycles: a polycycle (which, by definition, consists of finitely many heteroclinic orbits and equilibria) is called elementary if each of its equilibria has at least one nonzero eigenvalue.

**THEOREM 5.36 ([209,215]).** *There is a function  $E : \mathbb{N} \rightarrow \mathbb{N}$  such that the following holds. Let  $\dot{u} = f(u, \mu)$  be a generic family of planar ODEs that depends on a parameter  $\mu \in \mathbb{R}^k$  and has an elementary polycycle at  $\mu = 0$ . The number of limit cycles that exist*

near the polycycle for small values of  $\mu$  is then bounded by  $E(k)$ . Furthermore, the explicit estimate  $E(k) \leq 2^{25k^2}$  holds.

The existence of the function  $E(k)$  is due to Il'yashenko and Yakovenko, while Kaloshin proved the explicit estimate.

**5.2.4. Heteroclinic cycles containing periodic orbits.** In this section, we consider heteroclinic cycles that involve heteroclinic connections between a hyperbolic equilibrium and a hyperbolic periodic orbit. We focus on three-dimensional flows and refer to [323] for results in higher dimensions for cycles of codimension one and two.

Let  $\dot{u} = f(u, \mu)$  be a one-parameter family on  $\mathbb{R}^3$ . Suppose that the differential equation with  $\mu = 0$  has a hyperbolic periodic orbit  $q(x)$  of saddle type and a hyperbolic equilibrium  $p$  with eigenvalues  $\nu^{ss} < \nu^s < 0 < \nu^u$ . We assume that one separatrix in  $W^u(p)$  is contained in the stable manifold  $W^s(q)$  of  $q$ , and we denote this solution by  $h_1$ . The manifolds  $W^u(q)$  and  $W^s(p)$  are assumed to intersect transversally along a heteroclinic solution  $h_2$  from  $q(t)$  to  $p$ . Finally, the singular cycle  $\overline{h_1 \cup h_2}$  is assumed to be the maximal invariant set in some neighbourhood  $\mathcal{U}$  of itself.

**HYPOTHESIS 5.22 (Genericity and Unfolding Conditions).** *Consider the following genericity conditions:*

- (1)  $W^u(q)$  and  $W^s(p)$  intersect transversally along  $h_2$ .
- (2)  $h_2 \notin W^{ss}(p)$ .
- (3) The distance between  $W^u(p)$  and  $W^s(q)$ , measured in a cross section, varies with nonzero speed in  $\mu$ .

**HYPOTHESIS 5.23 (Expanding Versus Contracting Cycles).** *We distinguish the following two cases:*

- (1) *Expanding singular cycle:*  $-\nu^s/\nu^u < 1$ ;
- (2) *Contracting singular cycle:*  $-\nu^s/\nu^u > 1$ .

Bifurcations from expanding and contracting singular cycles are quite different in nature, and we illustrate this with the following bifurcation result.

**THEOREM 5.37 ([35]).** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs that unfolds a singular cycle that satisfies Hypothesis 5.22.*

- (1) *If Hypothesis 5.23(1) is met, then the bifurcation set has zero Lebesgue measure.*
- (2) *If Hypothesis 5.23(2) is met, then there are parameter values arbitrarily close to  $\mu = 0$  for which an attracting periodic orbit exists near the singular cycle.*

Further results on singular cycles for equilibria with real eigenvalues can be found in [250,293,305,349], while results for singular cycles that contain an equilibrium with complex conjugate leading eigenvalues appear in [43,68,69]. Codimension-two singular cycles involving tangencies of stable and unstable manifolds are considered in [81,283]. Papers in which singular cycles are treated using Lin's method include [227,240,324] and [36]; the latter paper uses Fenichel theory to describe trajectories near the periodic orbits.

### 5.3. Conservative and reversible systems

Systems that are reversible or conserve a certain energy-like quantity appear in many applications. In this section, we consider homoclinic bifurcations in such systems. We decided to review them together as they share several dynamical features and also since many models are both conservative and reversible: for instance, the energy  $H(q, p) = \frac{1}{2}\|p\|^2 + V(q)$  leads to a Hamiltonian system that is reversible and conservative.

In the introduction, we mentioned travelling waves in parabolic partial differential equations as an important source of ODE models where homoclinic orbits are of interest. One reason for the interest in reversible ODEs is the fact that standing waves  $u(x)$  of reaction-diffusion systems  $U_t = DU_{xx} + F(u)$  are captured by reversible systems. For surveys of reversible systems and their applications, we direct the reader to [75,76,253].

**5.3.1. Introduction and hypotheses.** In this section, we consider homoclinic orbits in systems

$$\dot{u} = f(u), \quad u \in \mathbb{R}^{2n} \quad (5.6)$$

that are conservative or reversible. A differential system is conservative if it preserves an appropriate real-valued quantity. More precisely, we assume the following:

**HYPOTHESIS 5.24 (Conservative Systems).** *There is a smooth function  $\mathcal{H} : \mathbb{R}^{2n} \rightarrow \mathbb{R}$  with  $\langle \nabla \mathcal{H}(u), f(u) \rangle = 0$  for all  $u \in \mathbb{R}^{2n}$ , and  $\nabla \mathcal{H}(u) = 0$  only at a discrete set of points in  $\mathbb{R}^{2n}$ .*

Thus, if [Hypothesis 5.24](#) is met, then  $\mathcal{H}(u(t)) = \mathcal{H}(u(0))$  along any solution  $u(t)$  of (5.6). A particular example of conservative systems are Hamiltonian systems given by

$$\dot{u} = J \nabla \mathcal{H}(u), \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad u \in \mathbb{R}^n \times \mathbb{R}^n.$$

If [Hypothesis 5.24](#) is met and  $h(t)$  is a homoclinic orbit, then

$$\psi(t) = \nabla \mathcal{H}(h(t)) \quad (5.7)$$

is a nontrivial bounded solution to the adjoint equation (2.3)

$$\dot{w} = -f_u(h(t))^* w. \quad (5.8)$$

Homoclinic orbits  $h(t)$  to hyperbolic equilibria  $p$  are found at the intersection of stable and unstable manifolds of  $p$  which all lie in the codimension-one surface  $\mathcal{H}^{-1}(p)$ :

**HYPOTHESIS 5.25 (Transversality).** *Assume that  $p$  is a hyperbolic equilibrium of (5.6) and that  $h(t)$  is a homoclinic orbit with  $h(0) \in W^s(p) \pitchfork W^u(p)$  in  $\mathcal{H}^{-1}(p)$  (which is equivalent to assuming [Hypothesis 2.2\(1\)](#)).*

A differential system is reversible if it admits a reverser  $\mathcal{R}$ :



**HYPOTHESIS 5.26** (Reversible Systems). *There exists a linear operator  $\mathcal{R} : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$  such that  $\mathcal{R}^2 = 1$ ,  $\dim \text{Fix}(\mathcal{R}) = n$ , and  $\mathcal{R}f(u) = -f(\mathcal{R}u)$  for all  $u \in \mathbb{R}^{2n}$ .*

If **Hypothesis 5.26** is met, then  $v(t) := \mathcal{R}u(-t)$  is a solution of (5.6) whenever  $u(t)$  is. In particular, we have  $\mathcal{R}W^s(p) = W^u(p)$  for any equilibrium  $p$ . Furthermore, if  $u(0) \in \text{Fix}(\mathcal{R})$ , then  $u(t) = \mathcal{R}u(-t)$  for all  $t \in \mathbb{R}$ , and we call  $u(t)$  reversible or symmetric:

**HYPOTHESIS 5.27** (Transversality). *We assume that*

- (1)  $h(t)$  is a symmetric homoclinic orbit of (5.6) to the hyperbolic equilibrium  $p$ .
- (2)  $W^s(p) \pitchfork \text{Fix}(\mathcal{R})$  in  $\mathbb{R}^{2n}$  at  $u = h(0)$ .

Homoclinic orbits in conservative or reversible systems are codimension-zero phenomena, that is, they persist under structure-preserving  $C^1$ -perturbations, and they are accompanied by periodic orbits.

**THEOREM 5.38** ([113,412]). *A homoclinic orbit  $h(t)$  to (5.6) that satisfies Hypotheses 5.24 and 5.25 (or Hypotheses 5.26 and 5.27) is accompanied by a family  $q_T(t)$  of unique 1-periodic orbits which is parameterized by their period  $T$  for all  $T$  sufficiently large. For symmetric homoclinic orbits in reversible systems, the accompanying periodic orbits are also symmetric. Furthermore, such homoclinic orbits persist under  $C^1$ -perturbations of the vector field  $f$  that preserve the conservative (reversible) structure.*

We remark that the spectra of the linearization of conservative or reversible systems about equilibria are symmetric:

**LEMMA 5.1.** *We have*

$$\nu \in \text{spec}(f_u(p)) \iff -\nu \in \text{spec}(f_u(p)),$$

*counted with multiplicity, if  $p$  is a hyperbolic equilibrium with  $\mathcal{H}_{uu}(p)$  invertible in a conservative system, or a symmetric equilibrium in a reversible system.*

**PROOF.** For reversible systems, we infer from **Hypothesis 5.26** and  $\mathcal{R}p = p$  that

$$\mathcal{R}f_u(p) = -f_u(\mathcal{R}p)\mathcal{R} = -f_u(p)\mathcal{R}$$

which shows that  $f_u(p)$  and  $-f_u(p)$  are similar. For conservative systems, **Hypothesis 5.24** implies that  $g(u) := \mathcal{H}_u(u)^* f(u) \equiv 0$ . Thus,  $0 = g_u(p) = \mathcal{H}_{uu}(p)f(p) + \mathcal{H}_u(p)^* f_u(p) = \mathcal{H}_u(p)^* f_u(p)$  which gives  $\mathcal{H}_u(p) = 0$  since  $p$  is hyperbolic. Computing the second derivative of  $g$ , we get

$$\begin{aligned} 0 = g_{uu}(p)[e_j, e_k] &= [\mathcal{H}_{uu}(p)e_j]^* [f_u(p)e_k] + [\mathcal{H}_{uu}(p)e_k]^* [f_u(p)e_j] \\ &= e_j^* \mathcal{H}_{uu}(p) f_u(p) e_k + e_k^* \mathcal{H}_{uu}(p) f_u(p) e_j \\ &= e_j^* \mathcal{H}_{uu}(p) f_u(p) e_k + e_j^* f_u(p)^* \mathcal{H}_{uu}(p) e_k \\ &= e_j^* [\mathcal{H}_{uu}(p) f_u(p) + f_u(p)^* \mathcal{H}_{uu}(p)] e_k \end{aligned}$$

for all  $j, k$  which shows that  $\mathcal{H}_{uu}(p) f_u(p) = -f_u(p)^* \mathcal{H}_{uu}(p)$ . Invertibility of  $\mathcal{H}_{uu}(p)$  then implies that  $f_u(p)$  and  $-f_u(p)^*$  are similar.  $\square$

**5.3.2. Bi-foci homoclinic orbits.** For conservative and reversible systems, homoclinic orbits to hyperbolic bi-foci are accompanied by infinitely many horseshoes. The result for conservative systems is as follows:

**THEOREM 5.39** ([112,340]). *Assume that [Hypotheses 5.24](#) and [5.25](#) are met. We also assume that there are precisely two leading stable eigenvalues  $v^s = -\alpha \pm i\beta$  which are simple with  $\beta > 0$ , and that*

$$\lim_{t \rightarrow \infty} e^{2\alpha t} \|h(t)\| \|h(-t)\| \neq 0$$

*(which is equivalent to assuming [Hypothesis 2.4\(1\)](#) and (3)). There are then infinitely many horseshoes close to the homoclinic orbit  $h(t)$ . Furthermore, for each  $N \geq 2$  and each sequence  $(k_1, \dots, k_N) \in \mathbb{N}^N$ , there are numbers  $k_*, T_* \geq 1$  and infinitely many  $N$ -homoclinic orbits near  $h(t)$  which are parameterized by  $k \geq k_*$  with return times given by*

$$T_j = \frac{2\pi(k_j + k)}{\beta} + T_* + O(e^{-\delta k}), \quad k \geq k_*, \quad j = 1, \dots, N.$$

*The  $N$ -pulses with these return times are unique.*

For Hamiltonian systems, the transversality condition can be relaxed significantly; see [61]. An application of the preceding theorem to fourth-order Hamiltonian systems can be found in [62]. Next, we state a similar result for reversible systems.

**THEOREM 5.40** ([74,174,340]). *Assume that [Hypotheses 5.26](#), [5.27\(1\)](#), and [2.2\(1\)](#) are met. We also assume that [Hypothesis 2.4\(1\)–\(2\)](#) hold and that there are precisely two leading stable eigenvalues  $v^s = -\alpha \pm i\beta$  which are simple with  $\beta > 0$ . There are then infinitely many horseshoes close to the homoclinic orbit  $h(t)$ , which consist of symmetric orbits. Furthermore, for each  $N \geq 2$  and each sequence  $(k_1, \dots, k_{\lceil N/2 \rceil}) \in \mathbb{N}^{\lceil N/2 \rceil}$ , there are numbers  $k_*, T_* \geq 1$  and infinitely many symmetric  $N$ -homoclinic orbits near  $h(t)$  which are parameterized by  $k \geq k_*$  with return times given by*

$$T_j = \frac{2\pi(k_j + k)}{\beta} + T_* + O(e^{-\delta k}), \quad k \geq k_*, \quad j = 1, \dots, \lceil N/2 \rceil.$$

*The  $N$ -pulses with these return times are unique.*

**5.3.3. Belyakov–Devaney transition.** For reversible and conservative systems that have a homoclinic orbit, we consider the codimension-one bifurcation where two real stable (and thus also the unstable) leading eigenvalues collide and become complex as the parameter is varied. This bifurcation was called Belyakov–Devaney in [75]; see also Section 5.1.4. For simplicity, we consider one-parameter families

$$\dot{u} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^4 \times \mathbb{R} \tag{5.9}$$

in  $\mathbb{R}^4$  that are reversible or conservative. For  $\mu = 0$ , we assume that  $h(t)$  is a homoclinic orbit to the hyperbolic equilibrium  $p$ . Recall that the imaginary parts of the two stable

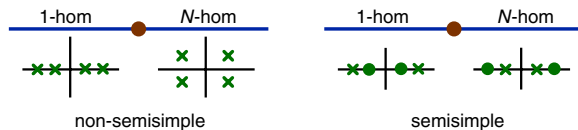


Fig. 5.19. In reversible or conservative systems, eigenvalues that collide on the real axis will typically be non-semisimple and therefore become complex: the resulting Belyakov–Devaney transition, described in Theorems 5.41 and 5.42, leads to  $N$ -homoclinic orbits for all  $N$  [left]. If an additional  $\mathbb{Z}_2$ -symmetry is present, semisimple eigenvalues may be enforced that cannot leave the real axis:  $N$ -homoclinic orbits may still bifurcate as shown in Theorem 5.43 [right].

eigenvalues are nonzero if the discriminant  $\Delta(\mu)$  of  $f_u(p, \mu)$  restricted to the stable directions is negative.

**HYPOTHESIS 5.28** (Non-semisimple Leading Eigenvalues). *At  $\mu = 0$ , the real leading stable eigenvalue  $v^s$  of  $f_u(p, 0)$  has geometric multiplicity one and algebraic multiplicity two with  $\partial_\mu \Delta(0) \neq 0$ , and we have*

$$\lim_{t \rightarrow \infty} \frac{1}{|t|} e^{|v^s t|} \|h(t)\| \neq 0, \quad \lim_{t \rightarrow -\infty} \frac{1}{|t|} e^{|v^s t|} \|\psi(t)\| \neq 0.$$

In the conservative case, we have the following result on the existence of  $N$ -homoclinic orbits.

**THEOREM 5.41.** *Assume that (5.9) satisfies Hypothesis 5.24 for all  $\mu$  and Hypotheses 5.25 and 5.28 at  $\mu = 0$ . Without loss of generality, assume that  $\partial_\mu \Delta(0) = -1$  so that the eigenvalues at  $p(\mu)$  are complex for  $\mu > 0$ . For each  $N \geq 2$  and each sequence  $(k_1, \dots, k_N) \in \mathbb{N}^N$ , there are numbers  $k_*$ ,  $T_* \geq 1$  and  $0 < \mu_N \ll 1$  so that (5.9) has infinitely many  $N$ -homoclinic orbits near  $h(t)$  for each  $0 < \mu < \mu_N$  which are parameterized by  $k \geq k_*$  with return times given by*

$$T_j = \frac{2\pi(k_j + k)}{\mu} + T_* + O(e^{-\delta k/\epsilon}), \quad k \geq k_*, \quad j = 1, \dots, N.$$

*The  $N$ -pulses with these return times are unique.*

An analogous theorem is true for reversible systems (see Figure 5.19).

**THEOREM 5.42.** *Assume that (5.9) satisfies Hypothesis 5.26 for all  $\mu$  and Hypotheses 5.27(1), 2.2(1) and 5.28 at  $\mu = 0$ . Without loss of generality, assume that  $\partial_\mu \Delta(0) = -1$  so that the eigenvalues at  $p(\mu)$  are complex for  $\mu > 0$ . For each  $N \geq 2$  and each sequence  $(k_1, \dots, k_{\lceil N/2 \rceil}) \in \mathbb{N}^{\lceil N/2 \rceil}$ , there are numbers  $k_*$ ,  $T_* \geq 1$  and  $0 < \mu_N \ll 1$  so that (5.9) has infinitely many symmetric  $N$ -homoclinic orbits near  $h(t)$  for each  $0 < \mu < \mu_N$  which are parameterized by  $k \geq k_*$  with return times given by*

$$T_j = \frac{2\pi(k_j + k)}{\beta} + T_* + O(e^{-\delta k}), \quad k \geq k_*, \quad j = 1, \dots, \lceil N/2 \rceil.$$

*The  $N$ -pulses with these return times are unique.*

**Theorem 5.41** was proved by Champneys and Toland [84] for a special class of Hamiltonian systems, while **Theorem 5.42** was established by Peroueme (unpublished notes) for  $N = 2$ . Alternatively, these theorems follow as in [340] from Lin's method upon using [338, Lemma 1.5] or [429, Lemma 2.1] to write down expansions in  $t$  of the quantities that appear in (3.13).

**5.3.4. Homoclinic orbits to equilibria with semisimple spectrum.** We consider reversible Hamiltonian systems of second-order ODEs of the form

$$\begin{aligned}\ddot{u}_1 &= u_1 + g_{u_1}(u_1, u_2) \\ \ddot{u}_2 &= (1 + \mu)u_2 + g_{u_2}(u_1, u_2),\end{aligned}\tag{5.10}$$

where the Hamiltonian is given  $H(u_1, u_2) = \dot{u}_1^2 + \dot{u}_2^2 + u_1^2 + u_2^2 + g(u_1, u_2)$ . Systems of the above type arise in coupled nonlinear Schrödinger equations [429] and in the study of three-dimensional water waves [163]. The key feature in (5.10) is the semisimple eigenvalue of multiplicity two that is unfolded by the parameter  $\mu$ : in contrast to the situation studied in Section 5.3.3, the unfolded eigenvalues are always real. We assume the following:

**HYPOTHESIS 5.29.** *We have  $g(0, 0) = g_u(0, 0) = 0$  and  $g(-u_1, u_2) = g(u_1, u_2)$  for all  $u = (u_1, u_2)$ . Furthermore, we assume that there is a constant  $\delta \notin \{0, 1\}$  so that  $\delta g_{u_1}(u_1, \delta u_1) = g_{u_2}(u_1, \delta u_1)$  for all  $u_1$ .*

The preceding hypothesis implies that the first-order system associated with (5.10) is equivariant with respect to the  $\mathbb{Z}_2$ -action  $\kappa : (u_1, u_2) \mapsto (-u_1, u_2)$ . It also implies that the subspace  $u_2 = \delta u_1$  is invariant under (5.10) when  $\mu = 0$ .

**HYPOTHESIS 5.30.** *For  $\mu = 0$ , Equation (5.10) has a pair  $u = (\pm h_1(t), h_2(t))$  of homoclinic orbits that lie in the invariant subspace  $u_2 = \delta u_1$  and satisfy Hypothesis 5.27.*

We refer to [163, 429] and [17] for conditions on  $g(u_1, u_2)$  that guarantee the existence of transverse homoclinic orbits. The next theorem shows that (5.10) admits  $N$ -homoclinic orbits near  $\mu = 0$ .

**THEOREM 5.43** ([163, 429]). *Assume that Hypotheses 5.29 and 5.30 are met for (5.10). For each  $N \geq 2$ , there is then an  $\mu_N > 0$  such that (5.10) has a unique pair of  $N$ -homoclinic orbits for each  $\mu$  with  $0 < |\mu| < \mu_N$  and  $\text{sign } \mu = \text{sign } \ln |\delta|$ : the  $N$  loops in each of the  $N$ -homoclinic orbits follow alternately  $(h_1, h_2)$  and  $(-h_1, h_2)$ . No other  $N$ -homoclinic orbits exist near  $\mu = 0$ .*

**5.3.5. Reversible systems with  $SO(2)$ -symmetry.** We consider the equation

$$\dot{u} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^4 \times \mathbb{R}\tag{5.11}$$

with  $f \in C^2$  and assume that this equation is reversible for all  $\mu$ ; see Hypothesis 5.26. In addition, we assume (5.11) is equivariant with respect to an  $\mathbb{S}^1$ -action for all values of  $\mu$ :

**HYPOTHESIS 5.31.** *There is a one-parameter group of orthogonal matrices  $T_\rho : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ , defined for  $\rho \in \mathbb{S}^1 := \mathbb{R}/2\pi\mathbb{Z}$ , such that  $T_{\rho_1}T_{\rho_2} = T_{\rho_1+\rho_2}$  for all  $\rho_1, \rho_2$  and  $\dim \text{Fix}(T_\rho) = \{0\}$  for  $\rho \neq 0$ . Furthermore, we assume that  $f(T_\rho u, \mu) = T_\rho f(u, \mu)$  and  $\mathcal{R}T_\rho = T_\rho \mathcal{R}$  for all  $(u, \mu)$  and all  $\rho$ .*

In particular,  $u = 0$  is an equilibrium for all  $\mu$  which we assume to be hyperbolic. The presence of the  $\mathbb{S}^1$ -action precludes that homoclinic orbits are transversely constructed in the sense of [Hypothesis 5.27\(2\)](#), since any homoclinic orbit  $h(t)$  yields an  $\mathbb{S}^1$ -orbit  $T_\rho h(t)$  of distinct symmetric homoclinic orbits for  $\rho \in \mathbb{S}^1$ . Thus, we assume that [Hypothesis 5.27\(1\)](#) is met with  $p = 0$  when  $\mu = 0$  and need then to assume that the parameter  $\mu$  unfolds the  $\mathbb{S}^1$ -orbit of homoclinic orbits. To make this precise, note that we have  $W^s(0, 0) = W^u(0, 0)$  due to the presence of the  $\mathbb{S}^1$ -action, which implies that the adjoint equation

$$w' = -f_u(h(t), 0, 0)^* w$$

has two bounded, linearly independent solutions  $\psi_1(t)$  and  $\psi_2(t)$ , which can be chosen so that

$$\psi_1(0) \in \text{Fix}(\mathcal{R}^*), \quad \psi_2(0) \in \text{Fix}(-\mathcal{R}^*).$$

Note that this implies  $\mathcal{R}^*\psi_1(t) = \psi_1(-t)$  and  $\mathcal{R}^*\psi_2(t) = -\psi_2(-t)$  for all  $t$ . We can now encode transversality with respect to  $\mu$  using a Melnikov integral.

**HYPOTHESIS 5.32.** *We assume that*

$$\int_{-\infty}^{\infty} \langle \psi_2(t), f_\mu(h(t), 0) \rangle dt \neq 0.$$

We are interested in  $N$ -homoclinic orbits near the  $\mathbb{S}^1$ -orbit  $\{T_\rho h(t) : \rho \in \mathbb{S}^1\}$ : in particular, to each  $N$ -homoclinic orbit near this group orbit, we can associate a sequence  $\{\rho_k\}_{k=1, \dots, N}$  so that the  $k$ th loop of the  $N$ -homoclinic orbit follows  $T_{\rho_k} h(t)$ . We can now state the following result:

**THEOREM 5.44** ([9,216,272,273]). *Assume that [Hypotheses 5.26](#), [5.27\(1\)](#), [5.31](#) and [5.32](#) are met. We also assume that the eigenvalues of  $f_u(0, 0)$  are  $\pm\alpha \pm i\beta$  with  $\alpha, \beta > 0$  and that*

$$\lim_{t \rightarrow \infty} \frac{\langle \psi_1(t), \psi_2(t) \rangle}{\|\psi_1(t)\| \|\psi_2(t)\|} \neq 0.$$

*For  $N = 2, 3$ , there are then sequences  $\mu_N^{j, \pm}$  with  $\mu_N^{j, \pm} \rightarrow 0$  as  $j \rightarrow \infty$  and  $\mu_N^{j, -} < 0 < \mu_N^{j, +}$  for all  $j$  so that [\(5.11\)](#) has an  $\mathbb{S}^1$ -orbit of symmetric  $N$ -homoclinic orbits for  $\mu = \mu_N^{k, \pm}$  with  $N = 2, 3$ . The 2-homoclinic orbits associated with  $\mu_2^{j, \pm}$  have  $\rho_1 \rho_2 = (-1)^j$ . Last, each of the  $N$ -homoclinic orbits constructed above for  $N = 2, 3$  satisfies the assumptions of this theorem.*

Of interest in applications to travelling waves, for instance in the complex Ginzburg–Landau equation, is the system

$$\dot{u} = f(u, \mu, c), \quad (u, \mu, c) \in \mathbb{R}^4 \times \mathbb{R} \times \mathbb{R}$$

where we assume that the equation for  $c = 0$  satisfies the hypotheses of [Theorem 5.44](#), while the parameter  $c$  breaks the reversibility and the homoclinic orbit, while preserving the  $\mathbb{S}^1$ -action. In this case, asymmetric  $N$ -homoclinic orbits can also be found near  $(\mu, c) = 0$ , and we refer to [\[272,273\]](#) for results.

**5.3.6. Reversible and Hamiltonian flip bifurcations.** The presence of reversers and conserved quantities has interesting implications for homoclinic flip bifurcations. Since the dynamics near reversible homoclinic orbits is symmetric with respect to time reflection, they undergo simultaneous orbit-flips in the stable and unstable direction, and the same is true for inclination-flips. For conservative systems,  $\nabla \mathcal{H}(h(t))$  is the unique nontrivial solution of the adjoint variational equation; see [\(5.7\)](#). Since we have  $\nabla \mathcal{H}(h(t)) \approx \mathcal{H}_{uu}(p)[h(t) - p]$  as  $|t| \rightarrow \infty$ , we find that if a homoclinic orbit in a conservative system undergoes an orbit-flip in the stable direction, say, it is at the same time in an inclination-flip configuration. Lastly, in reversible conservative systems, orbit-flips and inclination-flips of symmetric homoclinic orbits occur simultaneously and in both stable and unstable directions.

Throughout, we consider the equation

$$\dot{u} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^{2n} \times \mathbb{R} \quad (5.12)$$

with  $f \in C^4$  and assume that this equation is conservative or reversible for all  $\mu$ .

**HYPOTHESIS 5.33.** *Equation (5.12) has  $p = 0$  as equilibrium for  $\mu = 0$ . Furthermore, the linearization  $f_u(0, 0)$  has simple real eigenvalues  $-v^{\text{uu}} < -v^{\text{u}} < 0 < v^{\text{u}} < v^{\text{uu}}$ , and the real part of all other eigenvalues has modulus strictly larger than  $v^{\text{uu}}$ .*

We begin with the reversible non-conservative orbit-flip.

**THEOREM 5.45** ([\[344\]](#)). *Assume that [Hypothesis 5.26](#) is met for all  $\mu$ , and that [Hypotheses 2.2\(1\)](#), [5.27\(1\)](#) and [5.33](#) hold at  $\mu = 0$ . We assume that  $h(t)$  is in an orbit-flip configuration so that*

$$v^{\text{s}}(0) = 0, \quad \frac{d}{d\mu} v^{\text{s}}(0) \neq 0, \quad v_*^{\text{s}}(0) \neq 0, \quad S := \lim_{t \rightarrow \infty} e^{2v^{\text{uu}}t} \langle \psi(-t), h(t) \rangle \neq 0, \quad (5.13)$$

where we used the notation from [\(2.9\)](#) and [\(2.11\)](#).<sup>10</sup> We set  $\delta := -\text{sign } S \langle v_*^{\text{s}}(0), dv^{\text{s}}/d\mu(0) \rangle$ , then, for each  $N > 1$ , there is a  $\mu_N > 0$  so that [\(5.12\)](#) has a unique  $N$ -homoclinic orbit for each  $|\mu| < \mu_N$  with  $\text{sign } \mu = \delta$ .

<sup>10</sup>Reversibility and [\(5.13\)](#) imply that  $v^{\text{u}}(0) = 0$  and  $v_*^{\text{u}}(0) \neq 0$ .

Since conservative equations have (5.7) as the adjoint solution, they always violate assumption (5.13). To our knowledge, the reversible non-conservative inclination-flip has not been studied so far. Next, we state results on conservative non-reversible flip bifurcations.

**THEOREM 5.46** ([401]). *Assume that Hypothesis 5.24 is met for all  $\mu$ , and that Hypotheses 5.9, 5.25 and 5.33 hold at  $\mu = 0$ . We assume that  $h(t)$  is in a flip configuration in the stable direction so that*

$$v^s(0) = 0, \quad \frac{d}{d\mu} v^s(0) \neq 0, \quad v^u(0) \neq 0, \quad (5.14)$$

see (2.9) and (2.11) for the notation.<sup>11</sup> Under these assumptions, there is then a  $\delta \in \{\pm 1\}$  such that (5.12) has a unique  $N$ -homoclinic orbit for each  $N > 1$  and each sufficiently small  $\mu$  with sign  $\mu = \delta$ , and no  $N$ -homoclinic orbits for sign  $\mu = -\delta$ .

The bifurcation direction in the preceding theorem is made more explicit in [401], which contains also a complete characterization of the recurrent set and additional results about super-homoclinic orbits. Note again that the assumption (5.14) precludes reversibility of (5.12) since symmetric homoclinic orbits have  $v^u(0) = v^s(0)$ . It remains to consider the case where (5.12) is reversible and conservative.

**THEOREM 5.47** ([343]). *Assume that Hypotheses 5.24 and 5.26 are met for all  $\mu$ , and that Hypotheses 5.25 and 5.33 hold at  $\mu = 0$ . We assume that  $h(t)$  is in a flip configuration so that*

$$v^s(0) = 0, \quad \frac{d}{d\mu} v^s(0) \neq 0, \quad (5.15)$$

see (2.9) and (2.11) for the notation.<sup>12</sup> Under these assumptions, (5.12) has a unique  $N$ -homoclinic orbit for each  $N > 1$  and each sufficiently small nonzero  $\mu$ .

Thus, in contrast to the previous cases,  $N$ -homoclinic orbits bifurcate for either sign of the bifurcation parameter  $\mu$ .

**5.3.7. Coexisting homoclinic orbits.** As symmetric homoclinic orbits in reversible systems are typically persistent, one can find multiple coexisting homoclinic orbits. On the other hand, the action of the reverser  $\mathcal{R}$  forces nonsymmetric homoclinic orbits (which typically occur in one-parameter families of ODEs) to a symmetric hyperbolic equilibrium to come in pairs. The consequences of these two scenarios for the dynamics are reviewed in this section.

We consider homoclinic orbits to hyperbolic equilibria with unique real leading eigenvalues. Recall from Theorem 5.38 that nondegenerate symmetric homoclinic orbits in reversible systems, as well as nondegenerate homoclinic orbits in conservative systems, are accompanied by a sheet of periodic solutions. For two coexisting homoclinic orbits,

<sup>11</sup>Equations (5.7) and (5.14) imply then that  $v_*^u(0) = 0$  and  $v_*^s(0) \neq 0$ .

<sup>12</sup>Reversibility, (5.7) and (5.14) together imply that  $v^u(0) = v_*^u(0) = v_*^s(0) = 0$ .

one thus expects two sheets of periodic solutions. The sheets are normally hyperbolic and, depending on the geometry, transverse intersections of the stable and unstable manifolds of these sheets are feasible. This indeed occurs if both homoclinic orbits are in a bellows configuration, i.e. approach the equilibrium along the same direction for positive as well as negative time (see Section 5.1.8). For conservative systems, the transverse intersections of stable and unstable manifolds of single periodic orbits occur inside levels of the first integral and thus give a suspended horseshoe in each level. In other words, two homoclinic orbits in a bellows configuration for a conservative system give a one-parameter family of suspended horseshoes. The following theorem, contained in [193,404], can be obtained from standard constructions of invariant laminations [183].

**THEOREM 5.48 ([193,404]).** *Let  $\dot{u} = f(u)$  be a conservative, reversible ODE on  $\mathbb{R}^{2n}$  with first integral  $\mathcal{H}$  that satisfies [Hypotheses 5.24](#) and [5.26](#). Suppose that it has a hyperbolic symmetric equilibrium  $p$  and two symmetric homoclinic orbits  $h_1, h_2$  to  $p$ : we assume that both  $h_1$  and  $h_2$  satisfy [Hypotheses 2.1](#), [2.2\(1\)](#), and [2.4](#) and that they are in the bellows configuration so that [Hypothesis 5.12\(4\)](#) is met. In a small neighbourhood of  $\overline{h_1 \cup h_2}$ , there is then a one-parameter family of suspended horseshoes, parameterized by the energy  $\mathcal{H}$ , for values on one side of  $\mathcal{H}(p)$ .*

If the system is reversible but not conservative, then the dynamical picture is more complicated. The following two theorems summarize some of the features, and we refer to the cited papers for further properties, such as the existence of sheets of almost periodic orbits and of heterodimensional cycles.

**THEOREM 5.49 ([193,197]).** *Let  $\dot{u} = f(u)$  be a reversible ODE on  $\mathbb{R}^{2n}$  that satisfies [Hypothesis 5.26](#). Suppose that it has a hyperbolic symmetric equilibrium  $p$  and two symmetric homoclinic orbits  $h_1, h_2$  to  $p$ : we assume that both  $h_1$  and  $h_2$  satisfy [Hypotheses 2.1](#), [2.2\(1\)](#), and [2.4](#) and that they are in the bellows configuration so that [Hypothesis 5.12\(4\)](#) is met. There is then an invariant normally hyperbolic lamination that contains the nonwandering set near  $\overline{h_1 \cup h_2}$ . Furthermore, there are infinitely many sheets of symmetric periodic orbits, and arbitrarily small perturbations of  $f$  in the  $C^1$  topology create saddle-node bifurcations of periodic orbits.*

The following theorem covers the case of nonsymmetric homoclinic orbits in a one-parameter family of reversible ODEs.

**THEOREM 5.50 ([193,197]).** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of reversible ODEs on  $\mathbb{R}^{2n}$  with reverser  $\mathcal{R}$  as stated in [Hypothesis 5.26](#). For  $\mu = 0$ , we assume that this system has a hyperbolic symmetric equilibrium  $p$  and two homoclinic orbits  $h_1$  and  $h_2$  with  $h_1 = \mathcal{R}h_2$  to  $p$  that satisfy [Hypotheses 2.1](#), [2.2](#), [2.4](#) and [5.12\(4\)](#). For each  $\mu$  close to zero, there are then infinitely many sheets of symmetric periodic orbits. Furthermore, for  $\mu$  on one side of  $\mu = 0$ , there are infinitely many hyperbolic periodic orbits. Arbitrarily small perturbations of  $f$  in the  $C^1$  topology create saddle-node bifurcations of periodic orbits.*

**5.3.8. Degenerate homoclinic orbits.** Recall that a homoclinic orbit to a hyperbolic equilibrium is called degenerate if the tangent spaces of the stable and unstable manifolds



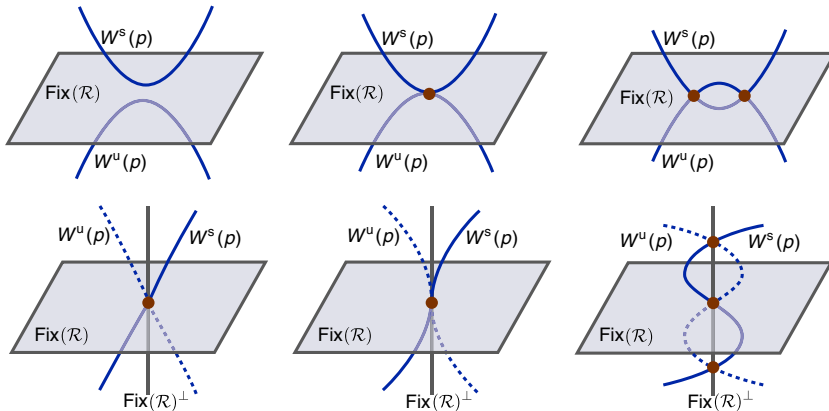


Fig. 5.20. The unfolding of degenerate reversible homoclinic orbits that satisfy Hypothesis 5.34(1) [top row] and 5.34(2) [bottom row] is shown: the phase diagram in the middle column is for  $\mu = 0$ .

along the homoclinic orbit have at least a plane in common. In generic systems, degenerate homoclinic orbits are of codimension three; see Section 5.1.9. In reversible systems, symmetric homoclinic orbits can be degenerate in two different ways which are of codimension one or codimension two. Such degenerate symmetric homoclinic orbits hence occur generically in one-parameter or two-parameter families. We review the situation near degenerate homoclinic orbits in reversible systems.

Thus, consider a family  $\dot{u} = f(u, \mu)$ ,  $\mu \in \mathbb{R}$  or  $\mu \in \mathbb{R}^2$ , of reversible ODEs on  $\mathbb{R}^{2n}$  with reverser  $\mathcal{R}$ . We may assume that  $\text{Fix}(-\mathcal{R})$  is perpendicular to  $\text{Fix}(\mathcal{R})$ . Assume that  $h(t)$  is a symmetric homoclinic orbit for  $\mu = 0$ , then we can choose a cross section  $\Sigma \subset \dot{h}(0)^\perp$  at  $h(0)$  transverse to  $h$ , so that  $\Sigma$  is symmetric and thus contains part of  $\text{Fix}(\mathcal{R})$ . A tangency of  $W^s(p, 0)$  and  $W^u(p, 0)$  can occur in two different ways, which we detail in the following hypothesis; see also Figure 5.20.

**HYPOTHESIS 5.34 (Generic Unfolding).** *Consider the following nondegeneracy and unfolding conditions:*

- (1) *The manifold  $W^u(p, 0) \cap \Sigma$  has a quadratic tangency with  $\text{Fix}(\mathcal{R})$  at  $h(0)$ , and  $W^u(p, \mu) \cap \Sigma \times \mathbb{R}$  intersects  $\text{Fix}(\mathcal{R}) \times \mathbb{R}$  transversally at  $(h(0), 0)$  in the extended phase space  $\Sigma \times \mathbb{R}$ .*
- (2) *The manifold  $W^u(p, 0) \cap \Sigma$  has a cubic tangency with  $[\text{Fix}(\mathcal{R}) \cap \Sigma]^\perp$  at  $h(0)$ , and  $W^u(p, \mu) \cap \Sigma \times \mathbb{R}^2$  intersects  $[\text{Fix}(\mathcal{R}) \cap \Sigma]^\perp \times \mathbb{R}^2$  transversally at  $(h(0), 0)$  in the extended phase space  $\Sigma \times \mathbb{R}^2$ .*

Homoclinic orbits near the degenerate homoclinic orbit can be found by studying the geometry of  $W^s(p, \mu)$  and  $W^u(p, \mu)$  in the vicinity of  $\text{Fix}(\mathcal{R}) \cap \Sigma$ . Since  $W^s(p) = \mathcal{R}W^u(p)$ , any intersection of  $W^u(p)$  with  $\text{Fix}(\mathcal{R})$  in  $\Sigma$  yields a homoclinic orbit. We begin by stating a result in  $\mathbb{R}^4$  due to Fiedler and Turaev that covers the case outlined in Hypothesis 5.34(1).

**HYPOTHESIS 5.35** (Transversality Conditions). Assume that  $W^{s,lu}(p, 0) \cap \Sigma$  is transverse to both  $\text{Fix}(\mathcal{R})$  and  $[\text{Fix}(\mathcal{R}) \cap \Sigma]^\perp$  in  $\Sigma$  at  $h(0)$ .

**THEOREM 5.51** ([136]). Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of reversible ODEs on  $\mathbb{R}^4$  with reverser  $\mathcal{R}$  as in [Hypothesis 5.26](#). Assume that  $W^s(p, 0)$  is tangent to  $\text{Fix}(\mathcal{R})$  at  $h(0)$  and that [Hypotheses 2.3\(2\)](#), [2.4\(2\)](#), (4) and [Hypotheses 5.34\(1\)](#) and [5.35](#) are all met. At  $\mu = 0$ , two symmetric homoclinic orbits that exist on one side of  $\mu = 0$  (say  $\mu > 0$ ) collide and disappear when  $\mu < 0$ . Furthermore, one of the following two alternatives holds:

- (1) For  $\mu > 0$ , there is a single surface of symmetric periodic solutions joining the homoclinic orbits. No periodic solutions exist for  $\mu < 0$ .
- (2) For  $\mu < 0$ , there is a surface of symmetric periodic solutions that breaks into two components each bounded by a homoclinic orbit when  $\mu > 0$ .

Interestingly, the surfaces of periodic solutions in the above theorem contain both hyperbolic and elliptic periodic solutions with real and complex conjugate Floquet multipliers.

The next theorem is for degenerate homoclinic orbits that satisfy [Hypothesis 5.34\(2\)](#): in this case, while the stable and unstable manifolds are tangent to each other, the tangent space of their intersection does not lie in  $\text{Fix}(\mathcal{R})$  but instead in  $\text{Fix}(-\mathcal{R})$ . Since this bifurcation is of codimension two, we consider two-parameter families  $\dot{u} = f(u, \mu)$  of reversible ODEs.

**THEOREM 5.52** ([225]). Assume that the system  $\dot{u} = f(u, \mu)$  with  $(u, \mu) \in \mathbb{R}^{2n} \times \mathbb{R}^2$  is reversible with reverser  $\mathcal{R}$  as stated in [Hypothesis 5.26](#). Suppose that  $W^u(p, 0)$  intersects  $\text{Fix}(\mathcal{R})$  transversally at  $h(0)$  and that  $T_{h(0)}W^u(p, 0) \cap \text{Fix}(\mathcal{R})^\perp$  is one-dimensional. If [Hypothesis 5.34\(2\)](#) is met, then a symmetric homoclinic orbit to  $p$  exists for all small  $\mu$ , and there is a one-sided curve in the parameter plane that terminates at  $\mu = 0$  along which the family has two nonsymmetric homoclinic orbits.

The coexisting homoclinic orbits that occur in the unfoldings described in the preceding two theorems come with complicated recurrent dynamics even when the leading eigenvalues are real: see [Theorems 5.49](#) and [5.50](#) in the previous section.

**5.3.9. Saddle-centre homoclinic orbits.** We begin by considering Hamiltonian systems

$$\dot{u} = J\nabla\mathcal{H}(u), \quad u \in \mathbb{R}^4, \quad (5.16)$$

where  $J$  is the skew-symmetric matrix

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which satisfy the following hypothesis:

**HYPOTHESIS 5.36.** We assume that  $u = 0$  is a saddle-centre equilibrium so that the linearization about 0 has two nonzero eigenvalues  $v^c = \pm i\omega$  on the imaginary axis in

addition to nonzero stable and unstable eigenvalues at  $\pm v^u$ . We also assume that  $h(t)$  is a homoclinic orbit to the origin.

We record that homoclinic orbits to saddle-centres in Hamiltonian systems in  $\mathbb{R}^4$  are a codimension-two phenomenon, which can be seen from the following observations. A cross section  $\Sigma$  is foliated by level sets of  $\mathcal{H}$ , and the centre-stable manifold  $W^{cs}(0) \cap \Sigma$  is tangent, typically at a single point, to the level set of  $\mathcal{H}(0)$ . The same is true for  $W^{cu}(0) \cap \Sigma$ , and a homoclinic orbit to the origin exists if these two tangencies occur at the same point. Since the level sets of  $\mathcal{H}$  intersected with  $\Sigma$  are two dimensional, two parameters are needed to find homoclinic orbits to saddle-centres.

We first investigate the dynamics of (5.16) itself and discuss the unfolding under parameter variations afterwards. In appropriate coordinates  $(q_1, q_2, p_1, p_2)$ , the Hamiltonian  $\mathcal{H} : \mathbb{R}^4 \rightarrow \mathbb{R}$  has the Taylor expansion

$$\mathcal{H}(q_1, q_2, p_1, p_2) = -q_1 p_1 + \frac{\omega}{2}(q_2^2 + p_2^2) + O(\|(q, p)\|^3) \quad (5.17)$$

where we may assume that  $\omega > 0$  to make the energy increasing in the  $(q_2, p_2)$  coordinates. In particular,  $\mathcal{H}(0) = 0$ , and the origin has a two-dimensional centre manifold which, by the Lyapunov-centre theorem, is filled with periodic orbits  $u_E(t)$ , which are parameterized by their positive energy  $E = \mathcal{H}(u_E(0))$  for  $0 < E < E_0$  for some  $E_0 > 0$ .

To introduce a crucial genericity assumption needed below, we choose two-dimensional transverse sections  $\Sigma_{\pm}$  to the homoclinic orbit in the energy level set  $\mathcal{H}^{-1}(0)$  that contain the points  $h(\pm\ell)$  for some sufficiently large times  $\ell \geq 1$ . Using the centre coordinates  $(q_2, p_2)$  as coordinates in both sections  $\Sigma_{\pm}$ , it can be shown that the symplectic Poincaré map  $\Pi$  along the homoclinic orbit from  $\Sigma_-$  to  $\Sigma_+$  is, after an appropriate adjustment of  $\ell$ , of the form

$$\Pi(q_2, p_2) = \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \times \text{rotation matrix} \times \begin{pmatrix} q_2 \\ p_2 \end{pmatrix} + O(\|(q_2, p_2)\|^2). \quad (5.18)$$

Thus, the linearization of the Poincaré map about the homoclinic orbit contains stretching and contracting directions when  $a \neq 0$ , while it is a pure rotation for  $a = 0$ . We shall also need a second geometric condition which we explain below after stating it:

**HYPOTHESIS 5.37.** *Consider the following geometric conditions:*

- (1) *The quantity  $a$  in (5.18) is nonzero.*
- (2) *The homoclinic orbit  $h(t)$  approaches the origin along the positive  $p_1$  and  $q_1$  axes as  $t \rightarrow \pm\infty$  or else along the negative  $p_1$  and  $q_1$  axes.*

The second assumption above implies that there are solutions in the level set  $\mathcal{H}^{-1}(0)$  that pass from  $\Sigma_+$  to  $\Sigma_-$  near the equilibrium  $p$ . If it is not met, then such solutions cannot exist.

**THEOREM 5.53** ([235,260,280]). *Assume that Hypotheses 5.36 and 5.37(1) are met, and that  $\mathcal{H} \in C^3$ , then there is an  $E_0 > 0$  such that the stable and unstable manifolds of each periodic orbit  $u_E(t)$  near the origin intersect transversally near  $h(t)$  for  $0 < E < E_0$ .*

In particular, the level sets  $\{u; \mathcal{H}(u) = E\}$  contain horseshoes for all  $0 < E \ll 1$ . This result is also valid in  $\mathbb{R}^{2n}$  provided the leading stable and unstable eigenvalues are unique and simple (but possibly complex),  $W^{cs}(0) \cap W^u(0)$  and  $W^{cu}(0) \cap W^s(0)$  at  $h(0)$  in  $\mathcal{H}^{-1}(0)$ , and the homoclinic orbit is not in an orbit-flip configuration (see [Hypothesis 2.4\(2\)](#) and (4)).

Under additional assumptions, horseshoes can also be shown to exist in the zero-energy level set.

**THEOREM 5.54 ([161]).** Assume that [Hypotheses 5.36](#) and [5.37](#) are met, that  $\mathcal{H}$  is analytic, and that

$$\frac{\omega|a - 1/a|}{\nu^u} > \frac{3}{2},$$

then the energy level sets  $\{u; \mathcal{H}(u) = E\}$  contain horseshoes for all  $E$  with  $|E|$  sufficiently small.

See [162] for a study of Lyapunov stability of two saddle-centre homoclinic orbits in Hamiltonian systems with an additional  $\mathbb{Z}_2$ -equivariance.

Next, we unfold the situation considered above for reversible Hamiltonian systems, and refer to [234,235] for unfoldings in nonreversible Hamiltonian systems. In reversible Hamiltonian systems, symmetric homoclinic orbits to a saddle-centre are of codimension one, so let

$$\dot{u} = J\nabla\mathcal{H}(u, \mu), \quad (u, \mu) \in \mathbb{R}^4 \times \mathbb{R} \quad (5.19)$$

be a one-parameter family of Hamiltonian systems that satisfies [Hypothesis 5.36](#) at  $\mu = 0$ . In this case, the Poincaré map (5.18) depends on  $\mu$  and has the expansion

$$\begin{aligned} \Pi(q_2, p_2, \mu) &= \mu\pi_0 + \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \times \text{rotation matrix} \times \begin{pmatrix} q_2 \\ p_2 \end{pmatrix} \\ &\quad + O((|q_2| + |p_2| + |\mu|)^2) \end{aligned} \quad (5.20)$$

for some  $\pi_0 \in \mathbb{R}^2$ . We assume that  $\pi_0 \neq 0$  and that (5.19) is reversible:

**HYPOTHESIS 5.38.** (1) The vector  $\pi_0$  in (5.20) is not zero.

(2) There is a linear map  $\mathcal{R} : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  with  $\mathcal{R}^2 = 1$  and  $\dim \text{Fix}(\mathcal{R}) = 2$  so that  $\mathcal{R}J = -J\mathcal{R}$  and  $\mathcal{H}$  is invariant under  $\mathcal{R}$  (that is,  $\mathcal{H}(\mathcal{R}u, \mu) = \mathcal{H}(u, \mu)$  for all  $(u, \mu)$ ). We assume that  $h(0) \in \text{Fix}(\mathcal{R})$ .

We are interested in  $N$ -homoclinic orbits near  $h(t)$  for parameter values  $\mu$  near zero and therefore define

$$\begin{aligned} \Lambda_N &:= \{\mu; \mu \approx 0 \text{ and (5.19) has an } N\text{-homoclinic orbit near } h(t)\}, \\ \Lambda &:= \bigcup_{N \geq 1} \Lambda_N. \end{aligned}$$

**THEOREM 5.55** ([161,280]). Assume that  $\mathcal{H}$  is analytic in  $(u, \mu)$  and that (5.19) satisfies Hypothesis 5.38 for all  $\mu$  and Hypotheses 5.36 and 5.37 at  $\mu = 0$ .

- (1) If  $\mu$  is in  $\Lambda_k$ , then both  $(\mu - \varepsilon, \mu)$  and  $(\mu, \mu + \varepsilon)$  contain infinitely many points of  $\Lambda_{km}$  for any  $m \geq 1$  and any  $\varepsilon > 0$ .
- (2) The set  $\Lambda$  is countable, and each point in  $\Lambda$  is an accumulation point of  $\Lambda$ .

In particular, for each  $N \geq 2$ , there are sequences  $\mu_N^\pm(k) \rightarrow 0$  as  $k \rightarrow \infty$  with  $\mu_N^-(k) < 0 < \mu_N^+(k)$  so that (5.19) has an  $N$ -homoclinic orbit for  $\mu = \mu_N^\pm(k)$ .

Lastly, we consider reversible systems

$$\dot{u} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^4 \times \mathbb{R} \quad (5.21)$$

with homoclinic orbits to a saddle-centre equilibrium that are not Hamiltonian.

**THEOREM 5.56** ([78]). Assume that (5.21) satisfies Hypothesis 5.26 for all  $\mu$  and Hypothesis 5.36 at  $\mu = 0$ . We also assume that  $h(t)$  is symmetric and that the vector field  $f(u, \mu)$  can be conjugated near the origin to a finite-order normal form by a  $C^1$ -diffeomorphism that commutes with  $\mathcal{R}$ . Lastly, we assume that

$$\frac{d}{d\mu} h^u(\mu)|_{\mu=0} \notin T_{h^u(0)} W^{\text{cs}}(0) \quad (5.22)$$

where  $h^u(\mu)$  is the unique intersection near  $h(t)$  of the one-dimensional global unstable manifold  $W^u(0)$  with the section  $\Sigma_+$ . Under these assumptions, 2-homoclinic orbits can exist either for  $\mu > 0$  or else for  $\mu < 0$ . Moreover, there is a sequence  $\mu_k \rightarrow 0$  as  $k \rightarrow \infty$  so that (5.21) has 2-homoclinic orbits near  $h(t)$  for  $\mu = \mu_k$  (and the  $\mu_k$  have the same sign independently of  $k$ ).

This theorem shows that reversible Hamiltonian and reversible non-Hamiltonian systems with homoclinic orbits to saddle-centre equilibria behave in a fundamentally different way. It is worthwhile to remark that it is the assumption (5.22) that discriminates between the two cases: indeed, (5.17) shows that the level set  $\mathcal{H}^{-1}(0)$  is tangent to  $W^{\text{cs}}(0)$  in Hamiltonian systems; see [78, Lemma 2]. We refer to [78] for a comprehensive discussion and unfolding results in the situation where the Hamiltonian structure is broken while reversibility is retained; see also [224,427] for results on homoclinic orbits to saddle-centres in reversible systems and to [362] for infinite-dimensional conservative systems.

**5.3.10. Homoclinic orbits to nonhyperbolic equilibria.** In Section 5.1.10, we discussed homoclinic orbits to equilibria which themselves undergo a local bifurcation. Obviously, analogous bifurcations are possible in ODEs that preserve a structure such as reversibility. Recall that homoclinic orbits to saddle-centre equilibria were discussed in Section 5.3.9. Thus, we report here on reversible transcritical and pitchfork bifurcations.

We first discuss bifurcations from a symmetric homoclinic orbit to an equilibrium that undergoes a reversible transcritical bifurcation. Rather than stating detailed bifurcation results, for which we refer to [419], we focus on the geometric arguments that lead to these results. Thus, consider a two-parameter family of reversible ODEs on  $\mathbb{R}^4$  that

satisfies [Hypothesis 5.26](#). Suppose that the origin has a double zero eigenvalue and two real eigenvalues  $\nu^s < 0$  and  $\nu^u = -\nu^s > 0$  when  $\mu = 0$ . Suppose that the unfolding is generic, so that the vector field on the two-dimensional centre manifold  $W^c(0, \mu)$  is given by

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= \mu_1 x + x^2.\end{aligned}$$

For  $\mu_1 \neq 0$ , the flow on the centre manifold has a saddle and a focus equilibrium, and there exists a small symmetric homoclinic orbit to the saddle equilibrium that encloses the focus equilibrium. At  $\mu = 0$ , we now also assume the existence of a large symmetric homoclinic orbit  $h$  to the origin. Take a symmetric three-dimensional cross section  $\Sigma$  transverse to  $h(0)$  for  $\mu = 0$ , then the fixed point space  $\text{Fix}(\mathcal{R})$  intersects  $\Sigma$  in a two-dimensional plane. The geometry of single-round homoclinic and heteroclinic orbits can now be understood from the following observations. First, we know that there are three-dimensional centre-stable and centre-unstable manifolds  $W^{cs}(0, \mu)$  and  $W^{cu}(0, \mu)$ , which are fibred by the stable and unstable manifolds of orbits in the centre manifolds. This yields copies of the dynamics on the centre manifold in the two-dimensional intersections  $W^{cs}(0, \mu) \cap \Sigma$  and  $W^{cu}(0, \mu) \cap \Sigma$ . Under the assumption that  $W^{cs}(0, 0) \cap \Sigma$  is transverse to  $W^{cu}(0, 0) \cap \Sigma$  at  $h(0)$ , these sets are also transverse to  $\text{Fix}(\mathcal{R})$  and intersect it along curves. The curves  $W^{cs}(0, \mu) \cap \Sigma \cap \text{Fix}(\mathcal{R})$  provide symmetric homoclinic orbits to recurrent orbits in the centre manifold.

Bifurcations in systems with an additional  $\mathbb{Z}_2$  symmetry, where the equilibrium undergoes a reversible pitchfork bifurcation, are discussed in [\[417–419\]](#). Depending on the group action and on normal-form coefficients, the bifurcating two equilibria may be connected by a heteroclinic cycle or be accompanied by symmetric or nonsymmetric figure-eight homoclinic orbits to the persisting equilibrium.

**5.3.11. Heteroclinic cycles and snaking.** We consider the equation

$$\dot{u} = f(u, \mu), \quad (u, \mu) \in \mathbb{R}^4 \times \mathbb{R} \quad (5.23)$$

with  $f \in C^2$  and assume that this equation is reversible for all  $\mu$ , see [Hypothesis 5.26](#), and that it has two symmetric hyperbolic equilibria,  $u = 0$  and  $u = p \neq 0$ , for all  $\mu$  near zero.

**HYPOTHESIS 5.39.** For  $\mu = 0$ , Equation (5.23) has a heteroclinic orbit  $h_1(t)$  that connects  $u = 0$  to  $u = p$  and is transversely constructed so that [Hypothesis 2.2](#) is met.

If [Hypothesis 5.39](#) is met, then  $h_2(t) := \mathcal{R}h_1(t)$  is a transversely constructed heteroclinic orbit between  $u = p$  and  $u = 0$ . Our goal is to describe homoclinic orbits  $h(t)$  that connect  $u = 0$  to itself and are obtained by gluing the heteroclinic orbits  $h_1(t)$  and  $h_2(t)$  together for  $\mu$  close to zero.

**THEOREM 5.57** ([228]). Assume that [Hypotheses 5.26](#) and 5.39 are met and that the eigenvalues of  $f_u(p, 0)$  are  $\pm\alpha \pm i\beta$  with  $\alpha, \beta > 0$ . Then there are constants  $a \neq 0$ ,  $b \in \mathbb{R}$ , and  $L_* > 0$  such that (5.23) has a homoclinic orbit  $h(t)$  to  $u = 0$  for  $\mu$  close to

zero that spends time  $L \geq L_*$  near  $u = p$  if and only if

$$\mu = a \sin(\beta L + b) e^{-\alpha L} + o(e^{-\alpha L}), \quad L \geq L_*.$$

In particular, there are infinitely many homoclinic orbits to  $u = 0$  when  $\mu = 0$ ; all but finitely many of them disappear for  $\mu \neq 0$ . In the same setting, the existence of multi-pulses was recently considered in [229] under the assumption that  $u = 0$  is also a bi-focus.

We now discuss the situation where the equilibrium  $p$  is replaced by a periodic orbit  $q(t)$ . It is easier to formulate the hypotheses in the conservative context:

**HYPOTHESIS 5.40.** *Equation (5.23) is conservative with an energy  $\mathcal{H}(u, \mu)$  that is invariant under the reverser  $\mathcal{R}$  for all  $\mu$ . The origin  $u = 0$  is a hyperbolic equilibrium of (5.23), and we may assume that  $\mathcal{H}(0, \mu) = 0$  for all  $\mu$ . Furthermore, for each  $\mu$ , (5.23) has a symmetric periodic orbit  $q(t, \mu)$  with  $q(0, \mu) \in \text{Fix}(\mathcal{R})$  and  $\mathcal{H}(q(t, \mu), \mu) = 0$  that depends smoothly on  $\mu$  and has two positive Floquet multipliers  $e^{\pm\alpha(\mu)}$  with  $\alpha(\mu) > 0$  (the other two Floquet multipliers are necessarily equal to one).*

Next, we define

$$\Gamma := \{(\varphi, \mu) \in \mathbb{S}^1 \times \mathbb{R}; W^u(0, \mu) \cap W^{ss}(q(\varphi, \mu), \mu) \neq \emptyset\}$$

which encodes and captures all heteroclinic orbits that connect  $u = 0$  to  $q(t, \mu)$ . We assume that  $\Gamma$  is a graph:

**HYPOTHESIS 5.41.** *The set  $\Gamma$  is the graph of a smooth function  $z : \mathbb{S}^1 \rightarrow \mathbb{R}$ , and we assume that  $z'(\varphi) = 0$  implies  $z''(\varphi) \neq 0$ .*

The next result shows that the heteroclinic orbits described in Hypothesis 5.41 and their symmetric counterparts can be glued together to construct homoclinic orbits that connect  $u = 0$  to itself and spend a long time near the periodic orbits  $q(t, \mu)$ .

**THEOREM 5.58** ([36,85,101,237,424]). *Assume that Hypotheses 5.40–5.41 and an additional technical condition (which can be found in [36]) are met, then there are constants  $L_* \gg 1$  and  $\eta > 0$  so that the following is true: for each  $L > L_*$ , (5.23) has a symmetric homoclinic orbit  $h(t)$  that spends time  $L$  near  $q(t, \mu)$  if and only if  $\mu = z(\varphi_0 + L) + O(e^{-\eta L})$  for an appropriate  $\varphi_0 \in \{0, \pi\}$ , and  $h(0)$  lies near  $q(\varphi_0, \mu)$  in  $\text{Fix}(\mathcal{R})$ .*

Geometric versions of the preceding theorem were first given in [101,424]. The theorem as stated was proved in [36], and the results established there are, in fact, valid for higher-dimensional systems that are only reversible and not necessarily conservative. In [85,237], the conclusions of Theorem 5.58 were shown to hold near degenerate Turing instabilities of the Swift–Hohenberg equation.

We remark that [64–66] gave numerical evidence for the existence of asymmetric homoclinic orbits of (5.23) whose existence was subsequently proved in [36] under assumptions similar to those stated above. Near degenerate Turing bifurcations, these results follow again from [85,237]. We refer to the recent review [226] for a list of open problems.

#### 5.4. Homoclinic orbits arising through local bifurcations

Homoclinic orbits can emerge in local bifurcations, and the study of local bifurcations therefore provides one way to obtain rigorous existence results for homoclinic orbits. We focus here on homoclinic bifurcations near nilpotent singularities, near Hopf/saddle-node bifurcations in generic and reversible systems, and at  $0^{2+i\omega}$  resonances in reversible systems. Other local bifurcations that lead to homoclinic orbits (in particular, 1:3 and 1:4 resonances, and codimension-three Hopf/Bogdanov–Takens singularities) will not be discussed here.

Recall that two differential equations on  $\mathbb{R}^n$  are said to be topologically equivalent if there exists a homeomorphism  $h$  that maps orbits of the first system to orbits of the second equation, while preserving the direction of time. For two parameter-dependent families of ODEs on  $\mathbb{R}^n$ , one can seek homeomorphisms  $\Phi(\cdot, \mu)$  of  $\mathbb{R}^n$  and  $\phi$  on the parameter space that provide an equivalence  $v = \Phi(u, \mu)$  between  $\dot{u} = f(u, \mu)$  and  $\dot{v} = g(v, \phi(\mu))$ . One speaks of a (fibre  $C^0, C^r$ )-equivalence if  $\Phi(\cdot, \mu)$  is  $C^0$  for each  $\mu$  and  $\phi$  is  $C^r$ . One speaks of a  $(C^0, C^r)$ -equivalence if  $(u, \mu) \mapsto \Phi(u, \mu)$  is  $C^0$  and  $\phi$  is  $C^r$ . In local bifurcation theory, one uses, of course, local versions of these notions.

**5.4.1. Nilpotent singularities.** A nilpotent singularity is an equilibrium of an ODE  $\dot{u} = f(u)$  for which the linearization about the equilibrium has multiple eigenvalues at zero and no other eigenvalues on the imaginary axis. The algebraic and geometric multiplicities of the zero eigenvalue distinguish different nilpotent singularities. By restricting to a centre manifold, we may assume that the eigenvalues of the linearization about the equilibrium are all zero. The most elementary example of a nilpotent singularity is then the Bogdanov–Takens bifurcation in  $\mathbb{R}^2$  where the linearization is a nontrivial  $2 \times 2$  Jordan block. It is well known that small homoclinic orbits with real leading eigenvalues at the equilibrium occur in the unfolding of the Bogdanov–Takens singularity, and we will recall this statement and review homoclinic dynamics in unfoldings of higher-codimension nilpotent singularities that may lead to other related dynamics such as small Lorenz-like attractors.

First, we review the Bogdanov–Takens bifurcation where the eigenvalue at zero has algebraic multiplicity two and geometric multiplicity one. This bifurcation was studied by Bogdanov [48] and Takens [389] (reproduced in [394]); a streamlined analysis can be found in [334]. The result is that a generic two-parameter family that unfolds a geometrically simple but algebraically double eigenvalue at the origin is (fibre  $C^0, C^\infty$ )-equivalent to

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= x^2 + \mu_1 + y(\mu_2 \pm x).\end{aligned}$$

It was shown in [125] that this equivalence is, in fact, a  $(C^0, C^\infty)$ -equivalence.

**THEOREM 5.59.** *The two-parameter unfolding of a nilpotent Bogdanov–Takens singularity at  $\mu = 0$  contains one-sided curves of Hopf bifurcations and of homoclinic loops that branch from a curve of saddle-node bifurcations at  $\mu = 0$ ; see Figure 5.21.*



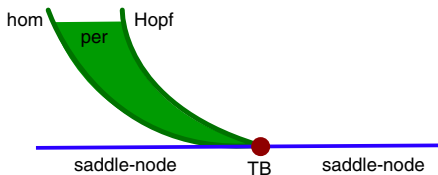


Fig. 5.21. The unfolding of a Bogdanov–Takens bifurcation (labelled TB) is illustrated.

Codimension-three bifurcations of nilpotent equilibria in the plane have been studied by Dumortier, Sotomayor and Roussarie [127,129]: intricate bifurcation diagrams, including homoclinic and heteroclinic bifurcations, arise, and we refer to these references for details.

Ibáñez and Rodríguez [204] considered nilpotent singularities with linearization

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (5.24)$$

in  $\mathbb{R}^3$  and proved that saddle-focus homoclinic orbits occur in the three-parameter unfolding, confirming a conjecture in [25]. This singularity appears in differential equations for coupled Brusselators [121]. In passing, we note that saddle-focus homoclinic orbits occur also near nilpotent singularities of higher codimension; see [203,387]. Returning to (5.24), a generic unfolding is given by the normal form

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= z, \\ \dot{z} &= \mu_1 + \mu_2 y + \mu_3 z + x^2 + bxy + cxz + dy^2 + eyz + O(\|(x, y, z, \mu)\|^3), \end{aligned} \quad (5.25)$$

which depends on the parameters  $\mu = (\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3$ .

**THEOREM 5.60** ([204]). *For any given neighbourhood  $\mathcal{U}$  of  $0 \in \mathbb{R}^3$ , there are parameter values  $\mu$  arbitrarily close to zero for which (5.25) has a wild saddle-focus homoclinic orbit in  $\mathcal{U}$ .*

The proof involves a singular re-scaling that reduces the problem to a study of perturbations of

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= z, \\ \dot{z} &= c^2 - \frac{1}{2}x^2 - y \end{aligned}$$

for  $c = 15\sqrt{22/19^3}$ . This system has a T-point heteroclinic cycle between the equilibria  $p_+ = (\sqrt{2}c, 0, 0)$  and  $p_- = (-\sqrt{2}c, 0, 0)$ : the two-dimensional stable manifold of  $p_-$  and the two-dimensional unstable manifold of  $p_+$  intersect along an isolated heteroclinic

orbit, while the codimension-two heteroclinic connection between  $p_-$  and  $p_+$  lies in the  $x$ -axis.

Next, we consider nilpotent singularities with rank-one linearization

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

The associated linear system is equivariant with respect to the  $\mathbb{Z}_2$ -action  $(x, y, z) \mapsto (-x, -y, z)$ . Within the class of  $\mathbb{Z}_2$ -equivariant perturbations, one finds miniature Lorenz-like attractors.

**THEOREM 5.61 ([124]).** *Consider a family of ODEs on  $\mathbb{R}^3$  with parameter  $\mu \in \mathbb{R}^5$  whose third-order truncation is given by*

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= \mu_1 x - x^3 + \mu_3 y + \mu_4 xz + \mu_5 yz, \\ \dot{z} &= \mu_2 z + x^2, \end{aligned}$$

*then there exist arbitrarily small values of  $\mu$  for which the differential equation has a small Lorenz-like attractor near the origin.*

To prove this result, Dumortier, Kokubu, and Oka showed that the unfolding of the singularity contains inclination-flip homoclinic orbits and that the eigenvalues at the equilibrium are such that Rychlik's theorem in [336] (see [Theorem 5.81](#)) applies which gives the existence of Lorenz-like attractors. In [373], further information on the connection between  $\mathbb{Z}_2$ -equivariant unfoldings of the nilpotent singularity with a triple zero eigenvalue of rank one and the Shimizu–Morioka system

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= x - \bar{\mu}_2 y - xz, \\ \dot{z} &= -\bar{\mu}_3 z + x^2 \end{aligned}$$

is obtained: Consider  $\mathbb{Z}_2$ -equivariant systems with third-order truncation

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= \mu_1 x - \mu_2 y + axz - a_1 x(x^2 + y^2) - a_2 yz + a_3 y(x^2 + y^2), \\ \dot{z} &= -\mu_3 + z^2 + b(x^2 + y^2) \end{aligned}$$

for  $ab > 0$  and  $\mu_3 > 0$ , and let  $\tau^2 = \mu + a\sqrt{\mu_3}$ . The time scaling  $t \rightarrow s/\tau$ , the phase-space scaling  $x \rightarrow x\sqrt{\tau^3/(ab)}$ ,  $y \rightarrow y\tau\sqrt{\tau^3/(ab)}$ ,  $z \rightarrow \sqrt{\mu_3} + z\tau^2/a$ , and the parameter scaling  $\mu_2 = \bar{\mu}_2\tau$ ,  $\mu_3 = (\bar{\mu}_3\tau/2)^2$  then produces the Shimizu–Morioka system in the limit  $\tau \rightarrow 0$ . We refer to [370] for a bifurcation study of the Shimizu–Morioka system.

**5.4.2. Hopf/saddle-node bifurcations in generic and reversible systems.** In the previous section, we reviewed local bifurcations that give rise to small homoclinic orbits; in

particular, wild saddle-focus homoclinic orbits appear in the unfolding of a codimension-three nilpotent singularity. We emphasize that the unfoldings considered so far were determined by finite jets.

Wild saddle-focus homoclinic orbits can also arise in Hopf/saddle-node bifurcations, which can occur in two-parameter families of ODEs. In this case, the normal form is invariant under rotations, and the perturbations that create homoclinic orbits are flat and therefore not captured by finite jets. It is an open problem to determine whether saddle-focus homoclinic orbits bifurcate for analytic two-parameter families that do not possess rotational symmetry.

To set the scene, let  $\dot{u} = f(u, \mu)$  be a two-parameter family of ODEs on  $\mathbb{R}^3$ . For  $\mu = 0$ , we assume that  $p$  is an equilibrium whose linearization has a simple eigenvalue at zero and a pair of purely imaginary eigenvalues. Under certain nondegeneracy conditions, the normal form near  $p$  is given by

$$\begin{aligned}\dot{x} &= v_1 + x^2 + s|z|^2 + O(\|(x, z, \bar{z})\|^4), \\ \dot{z} &= (v_2 + i\omega)z + (a + ib)xz + x^2z + O(\|(x, z, \bar{z})\|^4),\end{aligned}$$

where  $(x, z) \in \mathbb{R} \times \mathbb{C}$ , and  $v = (v_1, v_2) \in \mathbb{R}^2$  is the unfolding parameter; see, for instance, [247]. In the above expression, we have  $s = \pm 1$ , while  $\omega$ ,  $a$ , and  $b$  are smooth functions of  $v$  with  $\omega(0) \neq 0$  and  $a(0) \neq 0$ . The truncated normal form, in which only terms of order at most three are retained, is equivariant under rotations of  $z$ . The following proposition, proved in [58,390], shows that a normal form exists that is rotationally symmetric except possibly for flat terms.

**PROPOSITION 5.2.** *There exists a smooth coordinate change that transforms the above system into a differential equation of the form*

$$\begin{aligned}\dot{x} &= v_1 + x^2 + s|z|^2 + R_1(x, |z|, v) + S_1(x, z, \bar{z}, v) \\ \dot{z} &= (v_2 + i\omega)z + (a + ib)xz + x^2z + zR_2(x, |z|, v) + S_2(x, z, \bar{z}, v),\end{aligned}$$

where  $R_1(x, |z|, v)$  and  $zR_2(x, |z|, v)$  are  $O(\|(x, z)\|^4)$ , and  $S_j$  are smooth functions that are flat in  $(x, z, \bar{z}, v)$  at the origin in  $\mathbb{R} \times \mathbb{C} \times \mathbb{R}^2$ .

Writing the  $z$ -equation of the truncated system in polar coordinates  $(\rho, \phi)$ , we can ignore the equation for the angle  $\phi$  due to rotational symmetry and obtain the equation

$$\begin{cases} \dot{x} = v_1 + x^2 + s\rho^2, \\ \dot{\rho} = v_2\rho + ax\rho + \rho x^2 \end{cases} \quad (5.26)$$

for the amplitudes  $(x, \rho)$ . Depending on the signs of  $s$  and  $a$ , different unfoldings need to be considered. Here, we shall discuss only the case  $s = 1$ ,  $a < 0$ , where a heteroclinic cycle occurs in the unfolding of the amplitude system. Thus, we impose the condition:

**HYPOTHESIS 5.42 (Coefficient Condition).** *Assume that  $s = 1$  and  $a < 0$  in (5.26).*

An analysis of the amplitude equation gives the following picture; see also Figure 5.22. For  $v_1 < 0$ , there are two hyperbolic equilibria  $p_+ = (\sqrt{-v_1}, 0, 0)$  and  $p_- = (-\sqrt{-v_1}, 0, 0)$ .

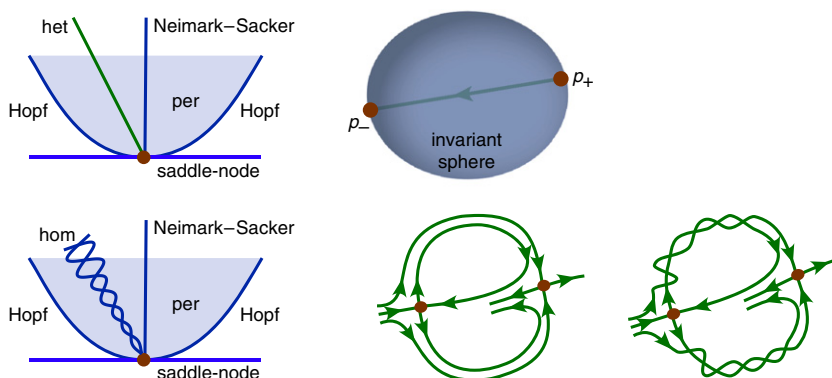


Fig. 5.22. The upper-left panel contains the bifurcation diagram of the truncated amplitude system in the  $(v_1, v_2)$ -parameter plane: the curve labelled 'het' corresponds to the existence of an invariant sphere in phase space, shown in the upper-right panel, whose surface is filled with heteroclinic connections from  $p_-$  to  $p_+$ , while the axis connecting these equilibria contains the indicated heteroclinic orbit from  $p_+$  back to  $p_-$ . Numerical computations for the non-truncated equation show entwined wiggling curves of saddle-focus homoclinic orbits of which the lower-left panel gives an impression. The centre and right panels in the bottom row indicate possible geometries of the invariant manifolds of  $p_-$  and  $p_+$  at parameter values where saddle-focus homoclinic orbits appear: homoclinic orbits to  $p_-$  can appear for parameter values where transverse intersections of  $W^u(p_-)$  and  $W^s(p_+)$  exist [right] but also for parameter values for which an invariant region exists [centre].

Inside the region bounded by the Hopf bifurcation curve, the equilibrium  $p_+$  has two complex conjugate stable eigenvalues, while  $p_-$  has two complex conjugate unstable eigenvalues. Along a curve  $\Gamma$  that emerges from the origin in the parameter plane, we find an invariant sphere of heteroclinic connections from  $p_-$  to  $p_+$ , while the  $x$ -axis contains a heteroclinic orbit from  $p_+$  to  $p_-$ . Generic perturbations from the truncated normal form will create transverse intersections of  $W^u(p_-)$  and  $W^s(p_+)$  [165]. Moreover, the  $x$ -axis may no longer be invariant, so that homoclinic orbits to  $p_-$  and to  $p_+$  could exist. A detailed analysis yields the existence of parameter values for which saddle-focus homoclinic orbits exists in generic families.

**THEOREM 5.62 ([59]).** *A generic two-parameter family on  $\mathbb{R}^3$  that unfolds a Hopf/saddle-node equilibrium at  $v = 0$  and satisfies Hypothesis 5.42 has saddle-focus homoclinic orbits for parameter values arbitrarily close to  $v = 0$ . These parameter values are contained in a wedge-shaped region whose width is flat in  $v$  as  $v \rightarrow 0$  and that is tangent at  $v = 0$  to the curve  $\Gamma$  for which a sphere of heteroclinic connections from  $p_+$  to  $p_-$  exists in the truncated normal form. Furthermore, the saddle-focus homoclinic orbits are wild for  $-2 < a < 0$ .*

Gaspard [146] noted the possible existence of a flow-invariant region for parameter values for which saddle-focus homoclinic orbits to  $p_-$  occur (the invariant region is bounded by the stable manifold of  $p_+$ ), which implies the existence of an attractor in this region; see Figure 5.22.

Much subsequent work has focused on the question as to whether specific families, which consist of the truncated normal form with small analytic terms added to it that break

the rotational symmetry, would unavoidably contain saddle-focus homoclinic bifurcations. When adding small third-order symmetry-breaking terms to the normal form, one finds two curves of saddle-focus homoclinic orbits (to  $p_+$  and to  $p_-$ ) that cross infinitely often on parameter space; see [80,146,219,220] for earlier results, [34] for more recent progress on this topic, and Figure 5.22 for an illustration of the bifurcation diagram.

Next, we discuss Hopf/saddle-node bifurcations in reversible systems in  $\mathbb{R}^3$ , where the fixed-point space of the reverser is a line of equilibria. In contrast to the situation for generic systems, where this bifurcation has codimension two (see above), Hopf/saddle-node bifurcations have codimension one in reversible systems. We remark that the conservative case is also of codimension one and refer to [59,104] for results on small homoclinic orbits that emerge in this setting.

The reversible Hopf/saddle-node bifurcation occurs in the Michelson system [279] given by

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= z \\ \dot{z} &= c^2 - \frac{1}{2}x^2 - y,\end{aligned}$$

which is reversible with respect to the involution  $\mathcal{R}(x, y, z) = (-x, y, -z)$ . Michelson's system is equivalent to the third-order system  $\ddot{x} + \dot{x} + \frac{1}{2}x^2 = c^2$ , which is the integrated travelling-wave equation of the Kuramoto–Sivashinsky equation  $u_t + uu_x + u_{xx} + u_{xxx} = 0$ . Analogous to the preceding analysis of generic Hopf/saddle-node bifurcation, the truncated normal form for reversible systems is invariant under rotations and reduces therefore to a planar amplitude system of the form

$$\begin{aligned}\dot{x} &= \nu + x^2 + \rho^2, \\ \dot{\rho} &= a\rho x\end{aligned}$$

with unfolding parameter  $\nu$ , which replaces the amplitude equation (5.26) for the generic Hopf/saddle-node bifurcation. Lamb, Teixeira and Webster showed in [254] that reversible perturbations of this normal form lead to the existence of infinitely many  $N$ -homoclinic orbits, where the number of rounds is defined with respect to a rescaled normal form.

**THEOREM 5.63 ([254]).** *A generic one-parameter family on  $\mathbb{R}^3$  that is reversible with respect to a reverser  $\mathcal{R}$  with one-dimensional fixed-point space and that unfolds a Hopf/saddle-node equilibrium admits infinitely many hyperbolic basic sets, infinitely many parameter values that correspond to saddle-focus  $N$ -homoclinic orbits, and infinitely many parameter values that correspond to symmetric  $N$ -heteroclinic cycles.*

In contrast to the situation for dissipative and conservative Hopf/saddle-node bifurcations, where heteroclinic cycles are of codimension two, symmetric heteroclinic cycles are codimension one in the reversible context. It was further established in [254] that the Michelson system has infinitely many bifurcations to  $N$ -homoclinic orbits and to symmetric  $N$ -heteroclinic cycles when varying the parameter  $c$ . For additional information on global bifurcations in the Michelson system, we refer to [123,255,421,422].

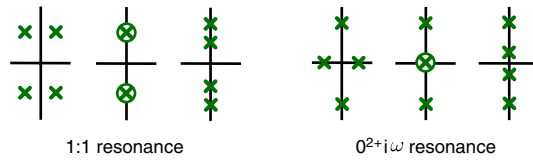


Fig. 5.23. The eigenvalues of an equilibrium in the unfolding of a 1:1 resonance [left] and a  $0^{2+}i\omega$  resonance are shown; see Sections 5.4.3 and 5.4.4.

**5.4.3. 1:1 resonances in reversible systems.** We now turn to local bifurcations in reversible systems and begin with the 1:1 resonance: this bifurcation occurs at equilibria where two purely imaginary pairs of eigenvalues collide on the imaginary axis and turn into a quadruplet of complex conjugate eigenvalues as a one-dimensional parameter is varied; see Figure 5.23. At the bifurcation point, the equilibrium therefore has a pair of two non-semisimple eigenvalues at  $\pm i\omega$  for some  $\omega > 0$ . Restricting to a centre manifold, we may therefore consider a one-parameter family  $\dot{u} = f(u, \mu)$  on  $\mathbb{R}^4$  that is reversible for all  $\mu$ . The generic unfolding of the 1:1 resonance was studied by Iooss and Peroueme [212], and we report first on their results.

Thus, suppose that we are given a reversible family of ODEs on  $\mathbb{R}^4$  so that the origin is an equilibrium for all  $\mu$ . For  $\mu = 0$ , we assume that  $i\omega$  is a non-semisimple eigenvalue of multiplicity two for some  $\omega > 0$  (and so is then  $-i\omega$ ). The normal form of a reversible family near the origin is, to any finite order, given by

$$\dot{A} = i\omega A + B + iAP\left(|A|^2, \frac{i}{2}(A\bar{B} - \bar{A}B); \mu\right) \quad (5.27)$$

$$\dot{B} = i\omega B + AQ\left(|A|^2, \frac{i}{2}(A\bar{B} - \bar{A}B); \mu\right) + iAP\left(|A|^2, \frac{i}{2}(A\bar{B} - \bar{A}B); \mu\right),$$

where  $(A, B) \in \mathbb{C}^2$  and the reverser acts as  $(A, B) \mapsto (\bar{A}, -\bar{B})$ . The functions  $P(u, v; \mu)$  and  $Q(u, v; \mu)$  are real polynomials in  $(u, v)$  that vanish at  $(u, v, \mu) = 0$ ; we write

$$Q(u, v; \mu) = q_1\mu + q_2u + q_3v + O((|u| + |v| + |\mu|)^2). \quad (5.28)$$

Note that the eigenvalues at the origin have nonzero real part for all  $\mu$  with  $q_1\mu > 0$ .

**THEOREM 5.64 ([212]).** *If the normal-form coefficients of a one-parameter family of reversible ODEs with a 1:1 resonance at  $\mu = 0$  satisfy  $q_1 \neq 0$  and  $q_2 < 0$ , then the system has a pair of small symmetric homoclinic orbits to the origin for each  $\mu$  close to zero for which  $q_1\mu > 0$ , and these homoclinic orbits satisfy Hypothesis 5.27.*

If  $q_2 > 0$ , then a pair of homoclinic orbits to periodic orbits bifurcates, and we refer to [212] for the precise results.

To prove the preceding theorem, Iooss and Peroueme first observed that the normal form (5.27) is integrable: two conserved quantities are given by  $v = \frac{i}{2}(A\bar{B} - \bar{A}B)$  and  $|B|^2 - \int_0^{|A|^2} Q(u, v; \mu) du$ . The normal form is furthermore equivariant under the

$\mathbb{S}^1$ -action  $(A, B) \mapsto e^{i\phi}(A, B)$ . If  $q_2 < 0$ , an  $\mathbb{S}^1$ -orbit of small homoclinic orbits to the origin bifurcates into  $q_1\mu > 0$ , and Iooss and Peroueme showed that the symmetric homoclinic orbits survive the perturbations by nonflat terms that are neglected in the normal form.

We remark that Bolle and Buffoni [49] constructed algebraically decaying homoclinic orbits for Hamiltonian systems that are exactly at a 1:1 resonance, provided a certain normal-form coefficient has the correct sign.

Next, we consider the situation where  $q_2$  changes sign as a second parameter is varied. This case is important in applications to snaking that we discussed in Section 5.3.11. When  $q_2 = 0$  and the normal-form coefficient  $q_4$  corresponding to the term  $q_4 u^2$  in the expansion of  $Q$  in (5.28) satisfies  $q_4 \neq 0$ , then the normal form can exhibit homoclinic orbits to the origin and heteroclinic cycles that connect the origin to small periodic orbits and back: this scenario was first studied by Woods and Champneys [424]. The persistence of these solutions for the full system involves asymptotics beyond all orders and has not yet been proved rigorously: we refer to the work by Chapman and Kozyreff [85, 237], who provided a detailed analysis of this bifurcation using formal methods.

**5.4.4.  $0^{2+}i\omega$  resonances in reversible systems.** In this section, we discuss  $0^{2+}i\omega$  resonances in reversible systems: these bifurcations occur at equilibria that have eigenvalues 0, 0, and  $\pm i\omega$  with  $\omega > 0$ ; see Figure 5.23. We outline the results contained in the monograph [264] by Lombardi who studied homoclinic loops near  $0^{2+}i\omega$  resonances in both four- and infinite-dimensional state space. Consider a one-parameter family  $\dot{u} = f(u, \mu)$  of reversible ODEs on  $\mathbb{R}^4$  with reverser  $\mathcal{R}$  that satisfies the following hypothesis.

**HYPOTHESIS 5.43 ( $0^{2+}i\omega$  Resonance).** *At  $\mu = 0$ , the origin is an equilibrium so that  $f_u(0, 0)$  has eigenvalues 0, 0,  $\pm i\omega$  for some  $\omega > 0$  with (generalized) eigenvectors  $v_0, v_1, v_{\pm}$  that satisfy*

$$f_u(0, 0)v_0 = 0, \quad f_u(0, 0)v_1 = v_0, \quad f_u(0, 0)v_{\pm} = \pm i\omega v_{\pm},$$

*and the reverser  $\mathcal{R}$  maps  $v_0$  to  $v_0$ .*

We remark that the case where  $\mathcal{R}v_0 = -v_0$  is referred to as the  $0^{2-}i\omega$  resonance. In the following, we write  $v_0^*, v_1^*, v_{\pm}^*$  for a basis that is dual to  $v_0, v_1, v_{\pm}$ .

**HYPOTHESIS 5.44 (Generic Unfolding of  $0^{2+}i\omega$  Resonances).** *Consider the following conditions on the quadratic terms and the dependence on the parameter:*

- (1)  $c_1 = \langle v_1^*, f_{u\mu}(0, 0)v_0 \rangle \neq 0$ ;
- (2)  $c_2 = \langle v_1^*, f_{uu}(0, 0)[v_0, v_0] \rangle \neq 0$ .

Without loss of generality, we may assume that  $c_1 > 0$  is positive. Under these assumptions, the equilibrium persists for  $\mu$  near zero, and the double zero eigenvalues of the linearization about it move from the imaginary axis for  $\mu < 0$  onto the real axis for  $\mu > 0$ . Thus, for  $\mu > 0$ , the linearization about the origin has a fast oscillatory part corresponding to the purely imaginary eigenvalues at  $\pm i\omega$  and a slow hyperbolic part with real eigenvalues  $\pm\sqrt{\mu[c_1 + O(\mu)]}$ . The question then is whether or not small-amplitude

homoclinic orbits can bifurcate for  $\mu > 0$ , and the following result shows that this will not happen for generic families due to exponentially small terms beyond all orders.

**THEOREM 5.65** ([264]). *Assume that the reversible one-parameter family  $\dot{u} = f(u, \mu)$  is analytic in  $(u, \mu) \in \mathbb{R}^4 \times \mathbb{R}$  and satisfies the [Hypotheses 5.43](#) and [5.44](#). For all  $c_1\mu > 0$  close to zero, the system then admits a sheet of small periodic solutions  $q_{\kappa, \mu}$  that are parameterized by their amplitude  $\kappa$ , and there are positive constants  $\kappa_1, \kappa_2, \sigma > 0$  so that the ODE has a pair of symmetric homoclinic orbits to the periodic solution  $q_{\kappa, \mu}$  for each  $\kappa$  with*

$$\kappa_1 c_1 \mu e^{-\omega[\pi - \sigma(c_1 \mu)^{3/10}]/\sqrt{c_1 \mu}} < \kappa < \kappa_2 c_1 \mu.$$

*Furthermore, there is a constant  $\kappa_0 \geq 0$ , which generically is positive, so that the ODE has no single-round symmetric homoclinic orbits to  $q_{\kappa, \mu}$  for*

$$0 \leq \kappa < \kappa_0 c_1 \mu e^{-\pi/[\omega\sqrt{c_1 \mu}]}.$$

The qualifier ‘single-round’ in the preceding theorem is with respect to a homoclinic orbit to the origin in a rescaled normal-form family of ODEs that we now describe. Using a singular rescaling and the rescaled parameter  $\nu = \sqrt{c_1 \mu}$ , the one-parameter family can be transformed into the normal form

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{v} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} y \\ x - \frac{3}{2}x^2 - c(\nu^2 + w^2) \\ -w(\omega/\nu + xv + bv x) \\ v(\omega/\nu + xv + bv x) \end{pmatrix} + R(x, y, v, w, \nu) \quad (5.29)$$

with higher-order terms  $R(x, y, v, w, \nu)$ ; see [264]. The truncated normal form is integrable and has a sheet of small symmetric periodic orbits  $q_{\kappa, \nu}$  for  $\nu > 0$  that is parameterized by the amplitude  $\kappa$ . Moreover, the truncated normal form admits a symmetric homoclinic orbit  $h(t)$  to the origin, and the stable and unstable manifolds of each periodic orbit  $q_{\kappa, \nu}$  coincide to form a two-parameter family of homoclinic orbits between symmetric periodic orbits. Among the circle of homoclinic orbits to a periodic orbit, two intersect the fixed point space of  $\mathcal{R}$  and therefore correspond to symmetric homoclinic orbits. Integrability holds not only for the normal form truncated at second-order terms but, in fact, for the normal forms truncated at any order.

The geometry becomes clearer in a symmetric three-dimensional cross section  $\Sigma$  placed at  $h(0)$ . The two-dimensional stable and unstable manifolds  $W^s(q_{\kappa, \nu})$  and  $W^u(q_{\kappa, \nu})$  of each individual periodic orbit  $q_{\kappa, \nu}$  are identical and, for  $\kappa > 0$ , their intersection with  $\Sigma$  are circles that intersect the plane  $\text{Fix}(\mathcal{R})$  transversally. The one-dimensional stable and unstable manifolds of the origin, on the other hand, each intersect  $\Sigma$  in a point, and these intersection points coincide and lie in  $\text{Fix}(\mathcal{R})$  for all truncated normal forms. We remark that the truncated system also admits elliptic periodic orbits and invariant tori.

The central issue is to study the effect of perturbations of the truncated normal form which will perturb the intersections  $W^s(q_{\kappa, \nu}) \cap \Sigma$  and  $W^u(q_{\kappa, \nu}) \cap \Sigma$ . In particular, we expect that the homoclinic orbit to the origin breaks as there is no reason why the



zero-dimensional intersections of the stable and unstable manifolds of the origin with  $\Sigma$  should coincide. The precise statement for analytic families that unfold the local bifurcation forms the content of the preceding theorem; we remark that this result does not address the possible existence of multi-round homoclinic orbits. The proof of the theorem involves the complexification of the system and the time variable which leads to a complex differential equation in  $\mathbb{C}^4$  in a complex time variable.

We remark that the geometry sketched above suggests that an additional parameter could be used to control the existence of symmetric homoclinic orbits to the origin: these homoclinic orbits should then occur along curves in the parameter plane. This idea is expounded in [77], which also contains careful numerical experiments.

The occurrence of exponentially small phenomena is illustrated in the following toy model which is taken from [264]. Consider the differential equations

$$\begin{aligned}\dot{x} &= 1 - x^2, \\ \dot{z} &= \frac{i\omega z}{\varepsilon} + i\varepsilon\rho(1 - x^2)\end{aligned}$$

on  $\mathbb{R} \times \mathbb{C}$ , which respect the reverser  $\mathcal{R}(x, z) = (-x, \bar{z})$ ; note that  $\text{Fix}(\mathcal{R})$  is the real axis in  $\{0\} \times \mathbb{C}$ . This system has two families of periodic solutions given by  $q_k^\pm(t) = (\pm 1, ke^{i\omega t/\varepsilon})$ . For  $\rho = 0$ , there is a sheet of heteroclinic solutions, given by  $h_{k,\phi}(t) = (\tanh t, ke^{i\omega t/\varepsilon + i\phi})$ , which connect  $q_k^-$  to  $q_k^+$ . The connections corresponding to  $\phi = 0, 1$  are symmetric. Thus, this toy problem has features reminiscent of those of the normal form of the  $0^{2+i\omega}$  resonance, where the parameter  $\rho$  can now be thought of as breaking the truncated normal form. To determine the fate of the stable manifold  $W^s(1, 0)$  of the equilibrium  $(x, z) = (1, 0)$  upon varying  $\rho$ , we note that it is given explicitly by

$$u(t) = \left( \tanh t, -i\rho\varepsilon \int_t^\infty \frac{e^{i\omega(t-s)/\varepsilon}}{\cosh^2 s} ds \right).$$

For  $t \rightarrow -\infty$ ,  $u(t)$  converges to the periodic solution  $q_{K(\varepsilon)}^-$  with

$$K(\varepsilon) = \rho\varepsilon \int_{-\infty}^\infty \frac{e^{-i\omega s/\varepsilon}}{\cosh^2 s} ds = \frac{\pi\omega\rho}{\sinh(\omega\pi/2\varepsilon)},$$

which is asymptotic to  $2\pi\omega\rho e^{-\omega\pi/2\varepsilon}$  as  $\varepsilon \rightarrow 0$ . Thus, the splitting distance between stable and unstable manifolds is exponentially small in  $\varepsilon$ . For more general problems such as the  $0^{2+i\omega}$  resonance, oscillatory integrals of the form  $\int_{-\infty}^\infty e^{i\omega s/\varepsilon} g(u(s)) ds$  with a small parameter  $\varepsilon$  for a solution  $u(t)$  and an analytic function  $g$  need to be studied. We refer to [210,211,263] for further details and results in this direction.

### 5.5. Equivariant systems

In flows that are equivariant under the action of a symmetry group, global bifurcations can differ substantially from those in generic flows. The restriction to equivariant perturbations

often reduces the codimension of bifurcations and may also constrain the dynamics found in an unfolding. One striking example is that of heteroclinic cycles which can appear robustly in equivariant systems.

Homoclinic and heteroclinic bifurcation theory for equivariant systems is a broad area in itself. To stay within the central theme of this survey, we limit ourselves to flows that are equivariant under the linear action of a finite group and discuss existence, stability, and bifurcations of heteroclinic cycles within this framework. Thus, continuous group actions will not be considered, and we will also not discuss how homoclinic and heteroclinic dynamics can emerge in local bifurcations. For general background on equivariant flows, we refer the reader to the books [158] and [93,138,139]; the latter references also contain sections on homoclinic and heteroclinic bifurcation theory.

To set the scene, let  $\Gamma$  be a finite group with a linear action  $x \mapsto \gamma x$  on  $\mathbb{R}^n$ . We may assume that  $\Gamma \subset O(n)$ , so that its action leaves the inner product  $\langle \cdot, \cdot \rangle$  invariant. We may also assume that the action of  $\Gamma$  is faithful.<sup>13</sup> By definition, a differential equation

$$\dot{u} = f(u), \quad u \in \mathbb{R}^n \quad (5.30)$$

is  $\Gamma$ -equivariant if the following statement holds:

$$\begin{aligned} u(t) \text{ is a solution to (5.30) if, and only if,} \\ \gamma u(t) \text{ is a solution to (5.30) for each } \gamma \in \Gamma. \end{aligned}$$

Equivalently,  $\Gamma$ -equivariance means that

$$f(\gamma u) = \gamma f(u), \quad \forall \gamma \in \Gamma, \quad \forall u \in \mathbb{R}^n. \quad (5.31)$$

For each  $u \in \mathbb{R}^n$ , we write  $\Gamma u = \{\gamma u \in \mathbb{R}^n; \gamma \in \Gamma\}$  for its group orbit and  $\Gamma_u = \{\gamma \in \Gamma; \gamma u = u\}$  for its isotropy group. If  $\Sigma \subset \Gamma$  is a subgroup of  $\Gamma$ , we write  $\text{Fix}(\Sigma) = \{u \in \mathbb{R}^n; \gamma u = u \forall \gamma \in \Sigma\}$  for the fixed-point space of  $\Sigma$ : note that (5.31) implies that  $\text{Fix}(\Sigma)$  is flow invariant. For any set  $\{\gamma_i\}$  of elements in  $\Gamma$ , we denote by  $\langle \{\gamma_i\} \rangle$  the group generated by  $\{\gamma_i\}$ : this is the smallest subgroup of  $\Gamma$  that contains the set  $\{\gamma_i\}$ . Finally, we recall that an isotypic component is a subspace of  $\mathbb{R}^n$  which is given as the sum of isomorphic irreducible subspaces of the action of  $\Gamma$ . The associated isotypic decomposition of  $\mathbb{R}^n$  is the decomposition of  $\mathbb{R}^n$  into isotypic components; see, for instance, [158].

To avoid any possible confusion between group and flow orbits, we shall use the terms ‘trajectory’ and ‘solution’ to refer to flow orbits and reserve, in this section, the term ‘orbit’ for group orbits.

**5.5.1. Robust heteroclinic cycles.** Equivariant flows may admit heteroclinic cycles between equilibria that persist under equivariant perturbations: this is because symmetry may force subspaces to be invariant, and a heteroclinic solution can now connect an equilibrium within an invariant subspace to a second equilibrium that is stable within this invariant subspace. We note that robust heteroclinic cycles can also occur in flows

<sup>13</sup>That is, for each  $\gamma \in \Gamma$ , there is an  $x \in \mathbb{R}^n$  with  $\gamma x \neq x$ .

that are not equivariant but have invariant subspaces for all parameter values: this is a common feature of models in population dynamics, where the coordinate subspaces  $\{u \in \mathbb{R}^n; u_i = 0\}$  are typically invariant (if species  $i$  is not present at  $t = 0$ , its population  $u_i(t)$  vanishes for all  $t$ ); see [184]. An extensive review of robust heteroclinic cycles is [241], which also includes a historical overview and descriptions of relevant experiments.

We ought to mention that the terminology of homoclinic and heteroclinic dynamics in equivariant systems varies widely in the literature: for consistency, we shall adhere to the notation introduced earlier in this paper, though this may not always agree with the prevalent terminology used in the literature.

Let  $\dot{u} = f(u)$  be a  $\Gamma$ -equivariant differential equation. A collection of different hyperbolic equilibria  $p_1, \dots, p_\ell$  and heteroclinic solutions  $h_j(t)$  from  $p_j$  to  $p_{j+1}$  is called a heteroclinic cycle.<sup>14</sup> Since the pointwise isotropy groups  $\Sigma_j = \Gamma_{h_j(t)}$  along each heteroclinic solution do not depend on  $t$ , we refer to it as the isotropy group of  $h_j$ . Define the fixed point spaces  $S_j = \text{Fix}(\Sigma_j)$  and recall that these spaces are flow invariant.

**HYPOTHESIS 5.45 (Robustness).** *We distinguish the following properties:*

- (1)  $W^u(p_j) \cap S_j$  and  $W^s(p_{j+1}) \cap S_j$  intersect transversally in  $S_j$ .
- (2)  $\dim W^u(p_j) = 1$ , and  $p_{j+1}$  is a sink in  $S_j$ .
- (3) Each fixed point space  $S_j$  is two-dimensional.

If **Hypothesis 5.45(1)** is met, then the manifold of heteroclinic connections from  $p_j$  to  $p_{j+1}$  is robust under  $\Gamma$ -equivariant perturbations, since the subspace  $S_j$  will continue to be invariant. Let  $\text{ind}_{S_j}(p_j)$  denote the Morse index  $\dim[W^u(p_j) \cap S_j]$  of  $p_j$  inside  $S_j$ , then the dimension of the manifold of heteroclinic connections in  $S_j$  is equal to  $\text{ind}_{S_j}(p_j) - \text{ind}_{S_j}(p_{j+1})$ . If this dimension is one, then the space  $S_j$  contains a robust isolated heteroclinic trajectory that connects  $p_j$  to  $p_{j+1}$ . We call the resulting heteroclinic cycle a *robust heteroclinic cycle*. Note that **Hypothesis 5.45(2)**, which is often found in the literature, is stronger: heteroclinic cycles that satisfy this assumption are often attracting, and we refer to Section 5.5.2 for stability results in this direction.

If **Hypothesis 5.45(2)–(3)** is met, we call the heteroclinic cycle a *simple heteroclinic cycle*. Any connected component in the image of a heteroclinic cycle under the group  $\Gamma$  is called a *heteroclinic network*. A *homoclinic cycle* is a polycycle<sup>15</sup> that is equal to the group orbit  $\langle \gamma \rangle \bar{h}$  of a heteroclinic trajectory  $h(t)$  that connects the equilibrium  $p$  to  $\gamma p$  for some  $\gamma \in \Gamma$ . The element  $\gamma \in \Gamma$  is called the twist of the homoclinic cycle: the twist is well defined modulo the isotropy group of  $p$ . A homoclinic network is a connected component of the group orbit  $\Gamma H$  of a homoclinic cycle  $H = \langle \gamma \rangle \bar{h}$ . We may now also define robust and simple homoclinic cycles in an analogous fashion. The following lemma characterizes homoclinic cycles:

**LEMMA 5.2 ([192]).** *Let  $h$  be a heteroclinic trajectory that connects the equilibria  $p$  and  $\gamma p$  for some  $\gamma \in \Gamma$ , then  $\Gamma h$  is connected, and thus a homoclinic cycle, if and only if  $\Gamma = \langle \gamma, \Gamma_p \rangle$ .*

<sup>14</sup>Throughout this section, all indices are taken modulo  $\ell$ .

<sup>15</sup>Recall that any connected invariant set that is the union of finitely many heteroclinic cycles is called a polycycle.

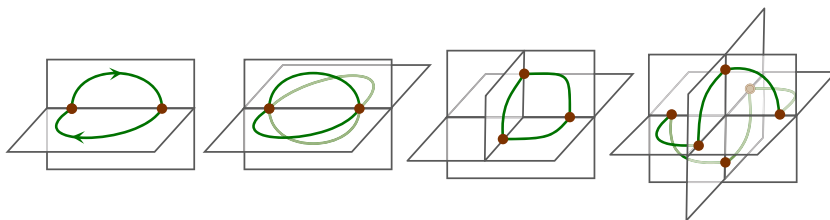


Fig. 5.24. Shown are the four different robust homoclinic cycles in  $\mathbb{R}^3$ : the symmetry groups are  $\Gamma = \mathbb{Z}_2 \ltimes \mathbb{Z}_2^2$  in the left two panel pictures and  $\Gamma = \mathbb{Z}_3 \ltimes \mathbb{Z}_2^3$  in the right two panels. For each  $\Gamma$ , the two different homoclinic cycles correspond to the same homoclinic network.

PROOF. Note that  $H_1 = \langle \gamma \rangle \bar{h}$  is trivially connected. Define inductively  $H_{i+1} = \bigcup_{q \in H_i} \Gamma_q H_i$ , where the union is over equilibria in  $H_i$ , and note that each  $H_i$  is connected. Since  $\Gamma$  is finite, this process terminates and yields the homoclinic cycle  $H$ . For Abelian groups  $\Gamma$ , the isotropy groups  $\Gamma_p$  are identical for all equilibria  $p$  in  $H$ , and thus  $H = H_2$ . For general  $\Gamma$ , the preceding construction shows that the isotropy groups of equilibria in  $H$  are conjugate to  $\Gamma_p$  via elements of  $\langle \gamma, \Gamma_p \rangle$ . Therefore,  $\Gamma = \langle \gamma, \Gamma_p \rangle$ .  $\square$

We will now discuss the classification of simple homoclinic cycles in  $\mathbb{R}^3$  and  $\mathbb{R}^4$ , the only dimensions for which a complete classification is known, and refer the reader to [383] for details and an historical overview. Homoclinic cycles for which  $\text{ind}_{S_j}(p_j) - \text{ind}_{S_j}(p_{j+1}) = 0$  will be discussed in Section 5.5.3: such cycles are typically of codimension one and will break under equivariant perturbations.

In  $\mathbb{R}^3$ , two different homoclinic networks exist that arise from four different homoclinic cycles:

THEOREM 5.66. *Figure 5.24 and the table*

| $\ell$ | $\Gamma$                              |
|--------|---------------------------------------|
| 2      | $\mathbb{Z}_2 \ltimes \mathbb{Z}_2^2$ |
| 4      | $\mathbb{Z}_2 \ltimes \mathbb{Z}_2^2$ |
| 3      | $\mathbb{Z}_3 \ltimes \mathbb{Z}_2^3$ |
| 6      | $\mathbb{Z}_3 \ltimes \mathbb{Z}_2^3$ |

contain the classification of simple homoclinic cycles in  $\mathbb{R}^3$ : listed are the number  $\ell$  of heteroclinic trajectories and the group  $\Gamma$  for which the cycle exists.

Figure 5.24 illustrates the four different homoclinic cycles in  $\mathbb{R}^3$ . Consider first the left two panels: the associated symmetry group is  $\Gamma = \mathbb{Z}_2 \ltimes \mathbb{Z}_2^2$ , where  $\mathbb{Z}_2$  and  $\mathbb{Z}_2^2$  are generated by  $(x, y, z) \mapsto (-x, z, y)$  and  $(x, y, z) \mapsto (x, \pm y, \pm z)$ , respectively. Two equilibria lie on the  $x$ -axis, and the homoclinic cycle consists of two heteroclinic trajectories when  $\gamma(x, y, z) = (-x, z, y)$  and of four trajectories when  $\gamma(x, y, z) = (-x, z, -y)$ . Both

cycles generate the same homoclinic network. Such homoclinic networks were studied, for instance, in [23,320]. An explicit system which contains these homoclinic networks is given by [345]

$$\begin{aligned}\dot{x} &= vx + z^2 - y^2 - x^3 + \beta x(y^2 + z^2), \\ \dot{y} &= y(\lambda + ay^2 + bz^2 + cx^2) + yx, \\ \dot{z} &= z(\lambda + az^2 + by^2 + cx^2) - zx\end{aligned}$$

where  $v > 0$  is small and  $\lambda \in (\lambda_H(v), \sqrt{v} + cv)$  for some  $\lambda_H(v) = -\frac{1}{2}v + O(v)$ ; the homoclinic cycle connects the equilibria  $(\pm\sqrt{v}, 0, 0)$ .

Next, consider the homoclinic cycles in the two rightmost panels in Figure 5.24, which again generate the same homoclinic network. The associated symmetry group is  $\Gamma = \mathbb{Z}_3 \ltimes \mathbb{Z}_2^3$  where  $\mathbb{Z}_3$  and  $\mathbb{Z}_2^3$  are generated by  $(x, y, z) \mapsto (y, z, x)$  and  $(x, y, z) \mapsto (\pm x, \pm y, \pm z)$ , respectively. An example of a  $\mathbb{Z}_3 \ltimes \mathbb{Z}_2^3$ -equivariant system that contains this network is given by

$$\begin{aligned}\dot{x} &= x(\lambda + ax^2 + by^2 + cz^2), \\ \dot{y} &= y(\lambda + ay^2 + bz^2 + cx^2), \\ \dot{z} &= z(\lambda + az^2 + bx^2 + cy^2)\end{aligned}$$

with  $a < 0$  and  $\lambda > 0$ : this ODE has a robust homoclinic cycle if, and only if,  $b < a < c$  or  $c < a < b$  [166]. To see how the robust homoclinic cycle arises, assume that  $b < a < c$ . The equilibria in the cycle are given by the group orbit of  $p_1 = (\sqrt{-\lambda/a}, 0, 0)$ : it is easy to check that  $(\pm\sqrt{-\lambda/a}, 0, 0)$  are sinks and  $(0, 0, \pm\sqrt{-\lambda/a})$  are saddles in the  $(x, z)$ -plane. Since  $b < 0$ , the region  $\{0 < z < \sqrt{-\lambda/a}\}$  is forward invariant, and it is also easy to verify that the unstable manifold of  $(0, 0, \sqrt{-\lambda/a})$  is bounded. This implies that  $W^u(p_1)$  goes to  $(\pm\sqrt{-\lambda/a}, 0, 0)$ . It follows from the results presented in Section 5.5.2 that the corresponding homoclinic cycle is asymptotically stable when  $2a > b + c$ .

We will now sketch the arguments that lead to Theorem 5.66 and refer to [383] for further details. Let  $\gamma \in \Gamma$  be the twist so that  $p_{j+1} = \gamma p_j$ . Choose a basis  $\{e_1, e_2, e_3\}$  of  $\mathbb{R}^3$  so that<sup>16</sup>  $S_1 = \langle e_1, e_2 \rangle$ ,  $S_2 = \gamma S_1 = \langle e_2, e_3 \rangle$  and  $S_3 = \gamma S_2 = \langle \cos(t)e_2 + \sin(t)e_3, e_1 \rangle$ . The angle  $t$  between consecutive equilibria  $p_j$  and  $p_{j+1}$  is called the *connecting angle*. The matrix  $A$  that represents  $\gamma$  in the basis  $\{e_j\}$  is

$$A = \begin{pmatrix} 0 & 0 & 1 \\ \alpha \sin(t) & \cos(t) & 0 \\ -\alpha \cos(t) & \sin(t) & 0 \end{pmatrix}$$

with  $\det A = \alpha$  and  $\alpha = 1$  or  $\alpha = -1$ . We claim that the connecting angle is either  $t = \pi/2$  or  $t = \pi$ . Indeed, let  $R = \text{diag}(1, 1, -1)$  be the matrix that fixes  $S_1$ , then  $ARA^{-1} = \text{diag}(-1, 1, 1)$  is a matrix that represents an element of  $\Gamma$ , and we conclude

<sup>16</sup>We use the notation  $\langle \{u_j\} \rangle$  for the vector space spanned by the vectors  $u_j$  in  $\mathbb{R}^3$ .

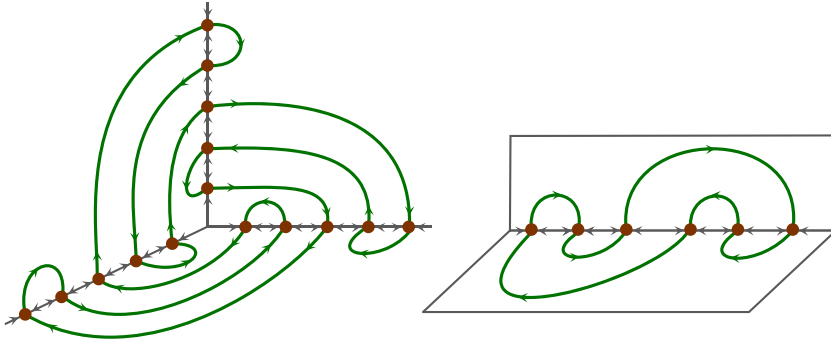


Fig. 5.25. Left panel: a heteroclinic cycle in a  $\mathbb{Z}_3 \times \mathbb{Z}_2^3$ -equivariant ODE in  $\mathbb{R}^3$ , where the action of the symmetry is defined by reflections in the coordinate planes and a permutation of the coordinates. Right panel: a heteroclinic cycle in a  $\mathbb{Z}_2 \times \mathbb{Z}_2^2$ -equivariant ODE in  $\mathbb{R}^3$ , where the action of the symmetry is defined by reflection in two coordinate planes and a permutation of two coordinates.

that the matrices

$$\begin{pmatrix} 0 & 0 & 1 \\ \sin(t) & \cos(t) & 0 \\ -\cos(t) & \sin(t) & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ -\sin(t) & \cos(t) & 0 \\ \cos(t) & \sin(t) & 0 \end{pmatrix}$$

both represent elements of  $\Gamma$ . The real parts of the complex eigenvalues of these matrices satisfy

$$\frac{1}{2} \cos(t) - \frac{1}{2} = \cos(\pi a), \quad \frac{1}{2} \cos(t) + \frac{1}{2} = \cos(\pi b)$$

for some rational numbers  $a, b$ . Hence  $\sin((a+b)\pi/2) \sin((a-b)\pi/2) = 1/2$ , which is only possible when  $a+b = a-b = 1/2$ , so that  $t = \pi/2$  or  $t = \pi$ . Finally, if  $t = \pi$ , the homoclinic cycle contains two equilibria. On the other hand, for  $t = \pi/2$ , it contains either three or six equilibria depending on the sign of  $\det A$ . This concludes the sketch of the arguments for [Theorem 5.66](#).

Breaking the  $\mathbb{Z}_2$ -symmetry in  $\Gamma = \mathbb{Z}_2 \times \mathbb{Z}_2^2$  creates a heteroclinic cycle with two equilibria from the homoclinic cycle. Likewise, breaking the  $\mathbb{Z}_3$ -symmetry in  $\Gamma = \mathbb{Z}_3 \times \mathbb{Z}_2^3$  creates a heteroclinic cycle with three or six equilibria near the homoclinic cycle. Hawker and Ashwin [177], see also [178], classify heteroclinic cycles in  $\mathbb{Z}_3$ -equivariant ODEs in  $\mathbb{R}^3$ ; see [Figure 5.25](#) for simple heteroclinic cycles in  $\mathbb{Z}_3 \times \mathbb{Z}_2^3$ - and  $\mathbb{Z}_2 \times \mathbb{Z}_2^2$ -equivariant ODEs in  $\mathbb{R}^3$  (clearly, many more heteroclinic cycles exist).

Simple homoclinic cycles in  $\mathbb{R}^4$  come in three types. Type B homoclinic cycles are contained in a three-dimensional invariant subspace.

**HYPOTHESIS 5.46** (Type A, B, C Homoclinic Cycles). *We distinguish the following configurations for simple heteroclinic trajectories in  $\mathbb{R}^4$ :*

- (1) *Type A:  $S_j + S_{j+1}$  is not a fixed point space.*

- (2) *Type B:  $S_j + S_{j+1}$  is a fixed point space that contains the homoclinic cycle.*  
 (3) *Type C:  $S_j + S_{j+1}$  is a fixed point space that does not contains the homoclinic cycle.*

Simple homoclinic cycles of type B and C in  $\mathbb{R}^4$  were investigated in [242], while Sottocornola [381–383] studied type A simple homoclinic cycles in  $\mathbb{R}^4$ . To state the classification result, we need the definition of structure angles that was introduced in [381–383]. Recall that the twist is the group element  $\gamma \in \Gamma$  for which  $p_{j+1} = \gamma p_j$ . As for homoclinic cycles in  $\mathbb{R}^3$ , we choose a basis  $\{e_1, e_2, e_3, e_4\}$  in which  $S_1, S_2 = \gamma S_1$ , and  $S_3 = \gamma S_2$  are given by

$$\begin{aligned} S_1 &= \langle e_1, e_2 \rangle, & S_2 &= \langle e_2, e_3 \rangle, \\ S_3 &= \langle \cos(t)e_2 + \sin(t)e_3, \cos(s)e_1 + \sin(s)e_4 \rangle \end{aligned}$$

for *structure angles*  $s, t$ , which can be thought of as the tilt and connecting angles of the hyperplane. The matrix that represents the twist  $\gamma$  is given by

$$A_{t,s}^\alpha = \begin{pmatrix} 0 & 0 & \cos(s) & -\sin(s) \\ \alpha \sin(t) & \cos(t) & 0 & 0 \\ -\alpha \cos(t) & \sin(t) & 0 & 0 \\ 0 & 0 & \sin(s) & \cos(s) \end{pmatrix}, \quad (5.32)$$

where  $\det(A) = \alpha$  with  $\alpha \in \{\pm 1\}$ . The symmetry group  $\Gamma$  contains the matrix  $R_1 = \text{diag}(1, 1, -1, -1)$  since the elements of  $\Sigma_1$  are the identity on  $S_1$  and  $-1$  on  $\langle e_3 \rangle$ . For type B and C homoclinic cycles,  $\Gamma$  also contains  $R_2 = \text{diag}(1, 1, 1, -1)$ .

**THEOREM 5.67** ([242, 381–383]). *The classification of simple homoclinic cycles in  $\mathbb{R}^4$  is in Table 1 where we list the homoclinic network, the generators of  $\Gamma$ , and the number of heteroclinic trajectories of the homoclinic cycle  $\langle \gamma \rangle h$  for possible twists  $\gamma$ , where  $h$  is a heteroclinic trajectory in  $S = \text{Fix}(R_1)$  that connects equilibria in  $\gamma^{-1}S \cap S$  to  $S \cap \gamma S$ .*

We briefly outline the strategy for proving the preceding theorem. First, the simple homoclinic cycles we found in  $\mathbb{R}^3$  occur in  $\mathbb{R}^4$  as robust type A or type B homoclinic cycles, depending on whether  $\mathbb{R}^3$  is a fixed point space or not. We note that they can also occur as homoclinic cycles that are not simple. To see this, consider the action of  $\mathbb{Z}_2 \times \mathbb{Z}_2^2$  on  $\mathbb{R}^4$  given by the linear maps  $(x, y, z, u) \mapsto (-x, z, y, -u)$ , which generates  $\mathbb{Z}_2$ , and  $(x, y, z, u) \mapsto (x, \pm y, \pm z, u)$ , which, in turn, generate  $\mathbb{Z}_2^2$ . The subspaces  $S_1 = \{y = 0\}$  and  $S_2 = \{z = 0\}$  are three-dimensional fixed-point spaces. Suppose that a differential equation with two hyperbolic equilibria  $p_1 = (x_*, 0, 0, u_*)$  and  $p_2 = (-x_*, 0, 0, -u_*)$  is given, and assume further, that the equilibria have different indices when restricting the flow to the fixed-point space  $S_1$ , so that  $\dim[W^u(p_1) \cap S_1] = 2$  and  $\dim[W^u(p_2) \cap S_1] = 1$ . If the manifolds  $W^u(p_1) \cap S_1$  and  $W^u(p_2) \cap S_1$  have a transverse intersection, its  $\Gamma$  image is a robust homoclinic network. The other three-dimensional simple homoclinic cycles can similarly provide models of robust homoclinic cycles in  $\mathbb{R}^4$  that are not simple. Next, consider any other simple homoclinic cycle of Type A in  $\mathbb{R}^4$ . If  $\Gamma \subset SO(4)$ , Sottocornola proved that the structure angles are multiples of  $\pi/4$ , and a direct analysis leads to the different simple homoclinic cycles listed in Table 1, where the homoclinic cycles of length

Table 1. The classification of simple homoclinic cycles in  $\mathbb{R}^4$  is shown: the columns contain the type (including, for type A cycles, a sign that indicates whether or not  $\Gamma \subset SO(4)$  and the number of equilibria it contains), the generators of  $\Gamma$ , and the number of heteroclinic trajectories of the homoclinic cycle  $\langle \gamma \rangle_{\bar{h}}$ .

| Type                              | Homoclinic network                   | Generators                                    | Twist                                                                              | Cycle length |
|-----------------------------------|--------------------------------------|-----------------------------------------------|------------------------------------------------------------------------------------|--------------|
| A with $\Gamma \subset SO(4)$     | $\mathcal{H}_2^{A,+}$                | $A_{\pi,0}^1, R_1$                            | $A_{\pi,0}^1$                                                                      | 2            |
|                                   |                                      |                                               | $R_1 A_{\pi,0}^1$                                                                  | 4            |
|                                   | $\mathcal{H}_6^{A,+}$                | $A_{\frac{\pi}{2},0}^1, R_1$                  | $A_{\frac{\pi}{2},0}^1$                                                            | 3            |
|                                   |                                      |                                               | $R_1 A_{\frac{\pi}{2},0}^1$                                                        | 6            |
|                                   | $\mathcal{H}_8^{A,+}$                | $A_{\frac{\pi}{2},\frac{\pi}{2}}^1, R_1$      | $A_{\frac{\pi}{2},\frac{\pi}{2}}^1, R_1 A_{\frac{\pi}{2},\frac{\pi}{2}}^1$         | 8            |
|                                   | $\mathcal{H}_{48}^{A,+}$             | $A_{\frac{\pi}{4},\frac{\pi}{4}}^1, R_1$      | $A_{\frac{\pi}{4},\frac{\pi}{4}}^1$                                                | 24           |
|                                   |                                      |                                               | $R_1 A_{\frac{\pi}{4},\frac{\pi}{4}}^1$                                            | 12           |
|                                   |                                      |                                               |                                                                                    |              |
| A with $\Gamma \not\subset SO(4)$ | $\mathcal{H}_{2k^2}^{A,-}, k \geq 1$ | $A_{\frac{\pi}{k},\frac{\pi}{k}}^{-1}, R_1$   | $A_{\frac{\pi}{k},\frac{\pi}{k}}^{-1}$                                             | $2k$         |
|                                   |                                      |                                               | $R_1 A_{\frac{\pi}{k},\frac{\pi}{k}}^{-1}$                                         | 4            |
| B                                 | $\mathcal{H}_2^B$                    | $A_{\pi,0}^1, R_1, R_2$                       | $A_{\pi,0}^1, R_2 A_{\pi,0}^1$                                                     | 2            |
|                                   |                                      |                                               | $R_1 A_{\pi,0}^1, R_1 R_2 A_{\pi,0}^1$                                             | 4            |
|                                   | $\mathcal{H}_6^B$                    | $A_{\frac{\pi}{2},0}^1, R_1, R_2$             | $A_{\frac{\pi}{2},0}^1, R_2 A_{\frac{\pi}{2},0}^1$                                 | 3            |
|                                   |                                      |                                               | $R_1 A_{\frac{\pi}{2},0}^1, R_1 R_2 A_{\frac{\pi}{2},0}^1$                         | 6            |
| C                                 | $\mathcal{H}_8^C$                    | $A_{\frac{\pi}{2},\frac{\pi}{2}}^1, R_1, R_2$ | $A_{\frac{\pi}{2},\frac{\pi}{2}}^1, R_1 A_{\frac{\pi}{2},\frac{\pi}{2}}^1$         | 8            |
|                                   |                                      |                                               | $R_2 A_{\frac{\pi}{2},\frac{\pi}{2}}^1, R_1 R_2 A_{\frac{\pi}{2},\frac{\pi}{2}}^1$ | 4            |

12 and 24 give the same homoclinic network. The structure angles for the remaining type A cycles that have  $\Gamma \not\subset SO(4)$  turn out to satisfy either  $t = s$  or  $t + s = \pi$ . In the second case, there is an infinite family of simple homoclinic cycles, all with four equilibria, that are parameterized by  $t = \pi/k$  with  $k > 1$ . In the first case, there is, likewise, an infinite family of simple homoclinic cycles with  $l = 2k$  equilibria for  $t = \pi/k$  with  $k > 1$ : each cycle gives the same network. For the existence of ODEs that admit these homoclinic cycles, Sottocornola relies on a result by Ashwin and Montaldi [33].

In passing, we note that the classification of simple homoclinic cycles of types B and C in  $\mathbb{R}^4$  extends to a classification of those simple heteroclinic cycles that intersect each connected component of  $[S_{j-1} \cap S_j] \setminus \{0\}$  in at most one point [242]. It turns out that there are four simple heteroclinic networks of type B and three of type C which satisfy the preceding condition: in [242], these networks are denoted by  $B_1^+, B_2^+, B_1^-, B_3^-$  and  $C_1^-, C_2^-, C_4^-$ , where the subscript  $m$  is the number of different group orbits of equilibria ( $m = 1$  corresponds to homoclinic networks) and the superscript  $\pm$  indicates whether or not  $-\text{id} \in \Gamma$  (the minus sign means that  $-\text{id} \in \Gamma$ ).



Finally, we briefly discuss the construction of homoclinic and heteroclinic cycles in  $\mathbb{R}^n$  with symmetry groups  $\Gamma = \mathbb{Z}_n \ltimes \mathbb{Z}_2^n$ ; details can be found in [138], and the reference [114] contains further information on their appearance in local bifurcations. A very useful tool is the invariant-sphere theorem which we explain first. Consider differential equations on  $\mathbb{R}^n$  of the form

$$\dot{u} = \lambda u + Q(u), \quad (5.33)$$

where  $\lambda > 0$  and  $Q$  is a homogeneous polynomial of degree  $d$  that satisfies  $\langle Q(x), x \rangle < 0$  for all  $x \in \mathbb{S}^{n-1}$ ; we refer to such maps  $Q$  as *contracting*. We can now split the vector field into its spherical and radial vector components so that

$$\begin{aligned} \dot{x} &= r^{d-1} (Q(x) - \langle Q(x), x \rangle x), \\ \dot{r} &= \lambda r + r^d \langle Q(x), x \rangle. \end{aligned}$$

The following invariant-sphere theorem due to Field shows that the radial part is, in a certain sense, irrelevant.

**THEOREM 5.68 ([137]).** *The differential equation (5.33) admits a unique invariant topological manifold  $\mathbb{S}^{n-1}$  which is homeomorphic to the  $(n-1)$ -sphere and attracts every point except the origin. The flow of (5.33) restricted to  $\mathbb{S}^{n-1}$  is topologically equivalent to the phase differential equation  $x' = Q(x) - \langle Q(x), x \rangle x$  on  $\mathbb{S}^{n-1}$ .*

To illustrate its use, consider the family

$$\begin{cases} \dot{x}_1 = x_1 + ax_1r^2 + bx_1x_2^2 + cx_1x_3^2 + dx_1x_4^2, \\ \dot{x}_2 = x_2 + ax_2r^2 + bx_2x_3^2 + cx_2x_4^2 + dx_2x_1^2, \\ \dot{x}_3 = x_3 + ax_3r^2 + bx_3x_4^2 + cx_3x_1^2 + dx_3x_2^2, \\ \dot{x}_4 = x_4 + ax_4r^2 + bx_4x_1^2 + cx_4x_2^2 + dx_4x_3^2, \end{cases} \quad (5.34)$$

of  $\mathbb{Z}_4 \ltimes \mathbb{Z}_2^4$ -equivariant ODEs, where  $r^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$ . The action of the symmetry group  $\mathbb{Z}_4 \ltimes \mathbb{Z}_2^4$  is generated by  $(x_1, x_2, x_3, x_4) \mapsto (-x_1, x_2, x_3, x_4)$  and  $(x_1, x_2, x_3, x_4) \mapsto (x_4, x_1, x_2, x_3)$ . **Theorem 5.68** can now be applied to yield an attracting invariant topological sphere provided  $a < 0$ ,  $2a + c < 0$ ,  $4a + b + d < 0$ , and  $4a + b + c + d < 0$  [425]. An analysis of the flow inside the invariant planes shows that a homoclinic cycle of type C exists when  $a < 0$  and  $bd < 0$ .

More generally, if the differential equation (5.33) is  $\mathbb{Z}_n \ltimes \mathbb{Z}_2^n$ -equivariant for an appropriate contracting homogeneous polynomial  $Q$ , then the invariant-sphere theorem shows that it suffices to consider the associated phase differential equation for  $x \in \mathbb{S}^{n-1}$ . Define  $\Lambda_{n-1} = \{(x_1, \dots, x_n) \in \mathbb{S}^{n-1}; x_1, \dots, x_n \geq 0\}$ , then  $\Lambda_{n-1}$  is left invariant by the action of  $\mathbb{Z}_n \subset \Gamma$ , we have  $\mathbb{S}^{n-1} = \mathbb{Z}_2^n \Lambda_{n-1}$ , and  $\gamma \text{int}(\Lambda_{n-1}) \cap \Lambda_{n-1} = \emptyset$  for each  $\gamma \in \mathbb{Z}_2^n$  with  $\gamma \neq \text{id}$ . A homoclinic network is thus given by a heteroclinic cycle for the phase differential equation on  $\Lambda_{n-1}$ . A given homoclinic cycle is called a  $k$ -face homoclinic cycle if the heteroclinic trajectories for the phase differential equation are on the  $k$ -faces of the simplex  $\Lambda_{n-1}$ .

We end this section with a brief discussion of more general heteroclinic cycles that contain infinitely many connecting trajectories. For instance, the definition of heteroclinic cycle given in [243], in which unstable manifolds of group orbits of equilibria connect to stable manifolds of group orbits of equilibria, allows for manifolds of heteroclinic trajectories. Ashwin and Field [32] gave a definition of a heteroclinic network which, in the context of trajectories that connect hyperbolic equilibria, consists of the following ingredients. A compact invariant set  $\Sigma$ , which consists of finitely many hyperbolic equilibria  $p_1, \dots, p_k$  and trajectories connecting them, is a heteroclinic network if

- (1)  $\Sigma$  is indecomposable: for every  $\varepsilon > 0$  and  $T > 0$ , any two points in  $\Sigma$  can be connected by an  $(\varepsilon, T)$ -pseudo trajectory<sup>17</sup>;
- (2)  $\Sigma$  has finite depth: if  $\Lambda_{i+1}$  denotes the recurrent set of the flow restricted to  $\Lambda_i$ , then there is a finite sequence  $\Lambda_0 = \Sigma, \dots, \Lambda_N = \{p_1, \dots, p_k\}$  (the minimal such  $N$  is called the depth of  $\Sigma$ ; note the similarity with the Birkhoff centre, which is defined analogously for non-wandering sets).

Ashwin and Field further generalize the notion of heteroclinic network to sets that contain periodic solutions or other recurrent invariant sets instead of equilibria. The sometimes intricate dynamics near invariant polycycles and heteroclinic networks is discussed in various papers: see, for instance, [14, 139] for heteroclinic networks in  $\mathbb{R}^4$  that contain equilibria with complex conjugate eigenvalues and produce suspended horseshoes. Some other key references are [15, 103, 318].

**5.5.2. Asymptotic stability of heteroclinic networks.** Consider a robust heteroclinic cycle for which Hypothesis 5.45(2) is met: the  $\omega$ -limit point  $p_{j+1}$  of the heteroclinic trajectory  $h_j$  is an attracting equilibrium inside the fixed-point space  $S_j = \text{Fix}(\Sigma_j)$  of the isotropy group  $\Sigma_j$  of the heteroclinic orbit  $h_j(t)$ . The geometry given by the subspaces  $S_j$  allows us to divide the spectrum of  $f_u(p_j)$  into four disjoint sets:

- (1) Radial eigenvalues, whose generalized eigenspaces lie in  $V_j^r = S_{j-1} \cap S_j$ ;
- (2) Contracting eigenvalues, whose generalized eigenspaces lie in  $V_j^c = S_{j-1} \ominus V_j^r$ ;
- (3) Transverse eigenvalues, whose generalized eigenspaces lie in  $V_j^t = (S_{j-1} + S_j)^\perp$ ;
- (4) Expanding eigenvalues, whose generalized eigenspaces lie in  $V_j^e = S_j \ominus V_j^r$ .

We remark that not all eigenvalues in  $V_j^e$  need to have positive real part. Define

$$\begin{aligned} r_j &= \min\{\text{Re } \lambda; \lambda \text{ is an eigenvalue of } f_u(p_j)|_{V_j^r}\}, \\ c_j &= \min\{\text{Re } \lambda; \lambda \text{ is an eigenvalue of } f_u(p_j)|_{V_j^c}\}, \\ t_j &= \max\{\text{Re } \lambda; \lambda \text{ is an eigenvalue of } f_u(p_j)|_{V_j^t}\}, \\ e_j &= \max\{\text{Re } \lambda; \lambda \text{ is an eigenvalue of } f_u(p_j)|_{V_j^e}\}. \end{aligned}$$

If  $S_{j-1} + S_j = \mathbb{R}^n$ , that is, when there are no transverse directions, we set  $t_j = -\infty$ . Note that the eigenspaces corresponding to  $c_j$ ,  $t_j$ ,  $e_{j+1}$ , and  $t_{j+1}$  all lie in  $S_j^\perp$ . The following

<sup>17</sup>An  $(\varepsilon, T)$ -pseudo trajectory from  $x$  to  $y$  is a finite set of points  $x_0 = x, \dots, x_n = y$  and times  $t_i > T$  with  $\|x_{i+1} - \varphi_{t_i}(x_i)\| < \varepsilon$ .

Table 2. Necessary and sufficient conditions for asymptotic stability of simple heteroclinic networks of type B and C in  $\mathbb{R}^4$ , where  $C_j = |c_j/e_j|$  and  $T_j = |t_j/e_j|$ .

| Type    | $\Gamma$                              | Stability condition                                                                                                                    |
|---------|---------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| $B_1^+$ | $\mathbb{Z}_2 \ltimes \mathbb{Z}_2^3$ | $C_1 > 1$                                                                                                                              |
| $B_2^+$ | $\mathbb{Z}_2^3$                      | $C_1 C_2 > 1$                                                                                                                          |
| $B_1^-$ | $\mathbb{Z}_3 \ltimes \mathbb{Z}_2^4$ | $C_1 > 1$                                                                                                                              |
| $B_3^-$ | $\mathbb{Z}_2^4$                      | $C_1 C_2 C_3 > 1$                                                                                                                      |
| $C_1^-$ | $\mathbb{Z}_4 \ltimes \mathbb{Z}_2^4$ | $C_1 - T_1 > 1$                                                                                                                        |
| $C_2^-$ | $\mathbb{Z}_2 \ltimes \mathbb{Z}_2^4$ | $C_1 + C_2 + T_1 T_2 > \min\{2, 1 + C_1 C_2\}$                                                                                         |
| $C_4^-$ | $\mathbb{Z}_2^4$                      | $C_1 C_3 + C_2 C_4 + T_1 T_2 C_3 + T_2 T_3 C_4 + T_3 T_4 C_1 + T_4 T_1 C_2 + T_1 T_2 T_3 T_4 + T_1 T_2 \min\{2, 1 + C_1 C_2 C_3 C_4\}$ |

condition is the analogue of type A, which we defined in [Hypothesis 5.46\(1\)](#) for systems in  $\mathbb{R}^4$ , in higher-dimensional systems.

**HYPOTHESIS 5.47.** *The eigenspaces corresponding to  $c_j, t_j, e_{j+1}, t_{j+1}$  lie in the same  $\Sigma_j$  isotopic component.*

**THEOREM 5.69** ([243]). *Suppose that  $\Gamma$  is a finite group that acts linearly, and assume that  $\dot{u} = f(u)$  is a  $\Gamma$ -equivariant equation on  $\mathbb{R}^n$  that admits a heteroclinic network  $H$  which satisfies [Hypothesis 5.45\(2\)](#). Assume, furthermore, that there are  $C^1$  linearizing coordinates near the equilibria in  $H$ . Write  $C_j = |c_j/e_j|$  and  $T_j = |t_j/e_j|$ , then the network  $H$  is asymptotically stable if  $\prod_{i=1}^k \min\{C_j, 1 - T_j\} > 1$ . If [Hypothesis 5.47](#) is met, this spectral condition is generically necessary and sufficient for asymptotic stability.*

Thus, radial eigenvalues play no role in the spectral conditions for asymptotic stability. Note that for homoclinic networks, where the spectral bounds are independent of  $j$ , the stability condition becomes  $-c > e$ . We shall now work out the asymptotic stability conditions for some simple heteroclinic networks of type B and C in  $\mathbb{R}^4$ .

**THEOREM 5.70** ([242]). *Let  $\dot{u} = f(u)$  be a  $\Gamma$ -equivariant equation on  $\mathbb{R}^4$  that admits a simple heteroclinic network of type B or C. Assume that the heteroclinic network intersects each connected component of  $V_j^r \setminus \{0\}$  in at most one point. Generically, the conditions listed in [Table 2](#) are then necessary and sufficient for asymptotic stability.*

We now attempt to make the conditions in [Table 2](#) more transparent. The proof of [Theorem 5.70](#) makes use of transition matrices [140, 184], which we now introduce. Near each equilibrium  $p_j$ , take coordinates  $x = (x^r, x^c, x^t, x^e)$  corresponding to the splitting  $\mathbb{R}^4 = V_j^r \oplus V_j^c \oplus V_j^t \oplus V_j^e$ . Pick the incoming and outgoing cross sections<sup>18</sup>

<sup>18</sup>We abuse notation by denoting these sections by  $\Sigma_j^{\text{in}, \text{out}}$  which are *not* the isotropy groups  $\Sigma_j$  used earlier.

$\Sigma_j^{\text{in}} = \{x; |x^t|, |x^e| \leq 1, |x^r|^2 + |x^c|^2 = 1\}$  and  $\Sigma_j^{\text{out}} = \{x; |x^r|, |x^c|, |x^t| \leq 1, x^e = 1\}$ , define transition maps  $\Pi_j^{\text{loc}} : \Sigma_j^{\text{in}} \rightarrow \Sigma_j^{\text{out}}$  and  $\Pi_j^{\text{far}} : \Sigma_j^{\text{out}} \rightarrow \Sigma_j^{\text{in}}$ , and write  $\Pi_j = \Pi_j^{\text{far}} \circ \Pi_j^{\text{loc}}$ . The first-return map is therefore given by the composition  $\Pi_k \circ \cdots \circ \Pi_1$ . In linearizing coordinates near  $p_j$ , we have

$$\Pi_j^{\text{loc}}(x^r, x^c, x^t, x^e) = (x^r(x^e)^{-r_j/e_j}, x^c(x^e)^{-c_j/e_j}, x^t(x^e)^{-t_j/e_j}, 1).$$

For  $x$  near  $h_{j-1} \cap \Sigma_j^{\text{in}}$ , the radii  $|x^r|^2$  and  $|x^c|^2$  will generically be nonzero, which suggests that the  $(x^t, x^e)$  components are more important. Expanding the transition maps  $\Pi_j^{\text{far}}$  in a Taylor series, the  $(x^t, x^e)$  components  $\Pi_j^t, \Pi_j^e$  of  $\Pi_j$  are, at lowest order, given by  $\Pi_j^t(x) = (a_j(x^e)^{-c_j/e_j} + b_j x^t(x^e)^{-t_j/e_j})$  and  $\Pi_j^e(x) = (c_j(x^e)^{-c_j/e_j} + d_j x^t(x^e)^{-t_j/e_j})$ , respectively. As the  $(x^r, x^c)$  coordinates are absent, we can consider the lowest-order terms of  $(\Pi_j^t, \Pi_j^e)$  as a map  $\pi_j : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  with

$$\pi_j(x^t, x^e) = (a_j(x^e)^{-c_j/e_j} + b_j x^t(x^e)^{-t_j/e_j}, c_j(x^e)^{-c_j/e_j} + d_j x^t(x^e)^{-t_j/e_j}).$$

For type B cycles, we have  $a = d = 0$ , while we have  $b = c = 0$  for type C cycles. Hence, in both cases, we can write  $\pi_j(x^t, x^e) = (E(x^t)^{\alpha_j}(x^e)^{\beta_j}, F(x^t)^{\gamma_j}(x^e)^{\delta_j})$  for suitable coefficients  $\alpha_j, \beta_j, \gamma_j, \delta_j$ . The transition matrix is now defined as the matrix  $M_j = \begin{pmatrix} \alpha_j & \beta_j \\ \gamma_j & \delta_j \end{pmatrix}$ , and asymptotic stability is deduced from the matrix  $M = M_k \cdots M_2 M_1$ : the network is asymptotically stable if the row sums of iterates  $M^l$  diverge to infinity as  $l \rightarrow \infty$ . For type C cycles, this conditions becomes  $\text{trace } M > \min\{2, 1 + \det M\}$ , which translates into the specific conditions stated in [Theorem 5.70](#).

An interesting phenomenon occurs for robust heteroclinic networks that contain equilibria of different indices: although the basin of attraction of such networks may not necessarily contain an open neighbourhood of the network, it may have full Lebesgue measure at points of the heteroclinic network. Such a heteroclinic network is called *essentially asymptotically stable* [[57,278](#)].

An example that exhibits essentially asymptotically stable networks can be constructed as follows. Consider a differential equation in  $\mathbb{R}^4$  that is  $\mathbb{Z}_2^2$ -equivariant under the action generated by  $(u_1, u_2, u_3, u_4) \mapsto (u_1, -u_2, -u_3, u_4)$  and  $(u_1, u_2, u_3, u_4) \mapsto (u_1, u_2, -u_3, -u_4)$ . Consider a heteroclinic cycle with equilibria  $p_1 = (-1, 0, 0, 0)$  and  $p_2 = (1, 0, 0, 0)$ , and assume that both  $f_u(p_1)$  and  $f_u(p_2)$  have four distinct real eigenvalues with  $\text{ind}(p_1) = 2$  and  $\text{ind}(p_2) = 1$ . Suppose that, inside the fixed-point plane  $\{(u_3, u_4) = 0\}$ , there are two heteroclinic trajectories that are related by symmetry and connect  $p_1$  to  $p_2$ : we assume that these trajectories form an attracting normally hyperbolic circle inside  $\{(u_3, u_4) = 0\}$ . Similarly, we assume that there are two symmetry-related heteroclinic trajectories inside  $\{(u_2, u_3) = 0\}$  that connect  $p_2$  to  $p_1$  and again form an attracting normally hyperbolic circle. If the eigenvalues are such that the strong unstable direction at  $p_1$  and the leading stable direction at  $p_2$  both lie in  $\{(u_3, u_4) = 0\}$ , the  $\mathbb{Z}_2^2$ -invariant heteroclinic network can be essentially asymptotically stable. Recall the definitions of radial, transversal, contracting and expanding eigenvalues from the beginning of this section.

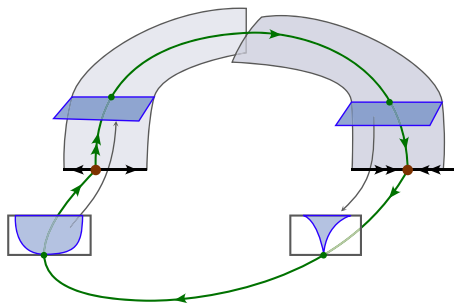


Fig. 5.26. The geometry behind a non-asymptotically stable heteroclinic network is illustrated. Only two heteroclinic trajectories are drawn. Starting near the upper heteroclinic trajectory, all points converge to the heteroclinic network, whereas a small number of points near the lower heteroclinic trajectory escape.

**THEOREM 5.71** ([278]). *Consider a  $\mathbb{Z}_2^2$ -equivariant equation on  $\mathbb{R}^4$  that admits a robust heteroclinic network as described above. If the open conditions*

$$c_1 c_2 > e_1 e_2, \quad -c_2(e_1 - t_1) > e_1 e_2, \quad t_2 < c_2$$

*on the transversal, contracting and expanding eigenvalues are satisfied, then the heteroclinic network is, generically, essentially asymptotically stable.*

Note that the heteroclinic network can be constructed so that it is contained in an attracting normally hyperbolic three-sphere  $\{\|u\| = 1\}$ . For this to hold, the additional eigenvalue conditions  $r_1 < c_1$  and  $r_2 < \min\{t_2, c_2\}$  must be satisfied. The flow near the heteroclinic network restricted to the three-sphere is illustrated in Figure 5.26. Further information on the stability properties of heteroclinic networks can be found in [244].

Lauterbach and Roberts [257] have shown how such essentially asymptotic robust heteroclinic networks appear in symmetry-breaking bifurcations of systems with spherical symmetry. The starting point in their work is a normally hyperbolic invariant three-sphere for an  $SO(3)$ -equivariant system for which the effects of forced symmetry-breaking are considered; see also Section 5.5.4.

Kirk and Silber gave another example of an essentially asymptotically stable heteroclinic network. It involves an equation on  $\mathbb{R}^4$  that is equivariant under the action of  $\mathbb{Z}_2^4$  generated by reflections of the coordinate axes. Assume that the  $i$ th coordinate axis contains the hyperbolic equilibrium  $p_i$  for each  $i = 1, \dots, 4$  and that there is a robust heteroclinic cycle inside the  $(u_1, u_2, u_3)$ -space that connects the equilibria  $p_1, p_2$  and  $p_3$ . We write  $H_{123}$  for the heteroclinic network in  $(u_1, u_2, u_3)$ -space that is the image under the action of the group. Likewise, assume the existence of a robust heteroclinic cycle that connects the equilibria  $p_1, p_2$  and  $p_4$  and of a corresponding heteroclinic network  $H_{124}$  in  $(u_1, u_2, u_4)$ -space. The connecting trajectories from  $p_i$  to  $p_j$  lie in the  $(u_i, u_j)$ -plane, and the equilibria  $p_1, p_3, p_4$  have one-dimensional unstable manifolds, while the equilibrium  $p_2$  has a two-dimensional unstable manifold. The following hypothesis captures the asymptotic stability of  $H_{123}$  and  $H_{124}$  restricted to the  $(u_1, u_2, u_3)$ -space and  $(u_1, u_2, u_4)$ -space, respectively; see Theorem 5.70. Denote by  $c_{ij}$  the contracting eigenvalue of  $f_u(p_j)$  that corresponds to the connecting trajectory from  $p_i$  to  $p_j$  and by  $e_{jk}$  the expanding

eigenvalue of  $f_u(p_j)$  restricted to the  $(u_j, u_k)$ -plane, which corresponds to the trajectory from  $p_j$  to  $p_k$ .

**HYPOTHESIS 5.48** (Asymptotic Stability of Three-dimensional Heteroclinic Networks). *The heteroclinic networks  $H_{123}$  and  $H_{124}$  are asymptotically stable within the enclosing three-dimensional spaces: we have*

$$-\frac{c_{12}c_{23}c_{31}}{e_{23}e_{31}e_{12}} > 1, \quad -\frac{c_{12}c_{24}c_{41}}{e_{24}e_{41}e_{12}} > 1.$$

**THEOREM 5.72** ([221]). *Let  $\dot{u} = f(u)$  be a  $\mathbb{Z}_2^4$ -equivariant ODE with a robust heteroclinic network as above, and assume Hypothesis 5.48 is met, then the robust heteroclinic network is, generically, essentially asymptotically stable.*

Postlethwaite and Dawes [318] considered heteroclinic networks in  $\mathbb{Z}_6 \ltimes \mathbb{Z}_2^6$ -equivariant differential equations in  $\mathbb{R}^6$  for which the equilibria lie on a single group orbit. They established the existence of trajectories that follow different heteroclinic trajectories in an irregular order, while converging to the heteroclinic network. Finally, we remark that homoclinic networks can also be essentially asymptotically stable [92, 119].

**5.5.3. Bifurcations from heteroclinic cycles.** In this section, we consider bifurcations from both robust and non-robust heteroclinic cycles.

First, we consider bifurcations from robust cycles: more specifically, we consider so-called *resonant bifurcations* where, by definition, the eigenvalue conditions for asymptotic stability become violated; see also Section 5.1.5. Certain resonant bifurcations from homoclinic cycles with real leading eigenvalues have been considered in [93, 352] and, in more generality, in [120]; a specific bifurcation that involves complex leading eigenvalues can be found in [319]. Here, we focus instead on *transverse bifurcations* where a transverse eigenvalue crosses the imaginary axis. Obviously, other types of bifurcations exist as well, but we are not aware of any systematic studies of bifurcations from robust heteroclinic cycles.

To outline the available results for transverse bifurcations of simple homoclinic cycles, we consider a  $\Gamma$ -equivariant system in  $\mathbb{R}^4$  and assume that  $H$  is a  $\Gamma$ -invariant simple homoclinic cycle consisting of equilibria  $p_1, \dots, p_\ell$ . Recall from Section 5.5.2 that the linearization about each of the equilibria in a simple homoclinic cycle has four real eigenvalues  $r, c, e, t$ , and we denote the corresponding eigenspaces at  $p_j$  by  $V_j^r, \dots, V_j^t$ . A transverse bifurcation occurs when the transverse eigenvalue  $t$  crosses through zero in  $V_j^t$ : the presence of symmetry implies that this is a pitchfork bifurcation.

**HYPOTHESIS 5.49** (Transverse Bifurcation). *Assume that an asymptotically stable homoclinic cycle loses its stability through a supercritical pitchfork bifurcation in the transverse directions:*

- (1) *The pitchfork bifurcation is supercritical: on a one-dimensional centre manifold, the differential equation has the normal form  $\dot{z} = \mu z + h(z, \mu)$  with  $h(0, \mu) = h_z(0, \mu) = 0$  and  $h_{zz}(0, \mu) < 0$ .*
- (2) *At  $\mu = 0$ , the contracting and expanding eigenvalue satisfy  $-c > e$ .*

We then have the following bifurcation result.

**THEOREM 5.73 ([91]).** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs on  $\mathbb{R}^4$ , which is equivariant under the linear action of a finite group  $\Gamma$ . We assume that the equation for  $\mu = 0$  admits a simple  $\Gamma$ -invariant homoclinic cycle  $H$  whose equilibria undergo a pitchfork bifurcation so that [Hypothesis 5.49](#) is met. Depending on the type of cycle (as defined in [Hypothesis 5.46](#)), we then have the following cases:*

- (1) *If  $H$  is of type A, then, generically, there is a unique branch of limit cycles that bifurcate from the homoclinic cycle. Furthermore, there is a constant  $d > 0$ , which depends only on the flow at  $\mu = 0$ , so that each periodic cycle is asymptotically stable for  $d < 1$  and unstable for  $d > 1$ .*
- (2) *If  $H$  is of type B, then there is a supercritical pitchfork bifurcation to two asymptotically stable homoclinic cycles.*
- (3) *If  $H$  is of type C, then there is a supercritical pitchfork bifurcation to four asymptotically stable homoclinic cycles.*

We will now briefly discuss the geometry for homoclinic cycles  $H$  of type B. In this case,  $H$  lies in a three-dimensional fixed-point space  $\text{Fix}(\tau)$  for some  $\tau \in \Gamma$ . If  $h_j$  is the heteroclinic trajectory of the homoclinic cycle inside the fixed-point space  $S_j = \text{Fix}(\xi_j)$  that connects the equilibrium  $p_j$  to  $p_{j+1}$ , then  $\text{Fix}(\xi_j \tau)$  is a three-dimensional fixed-point space that contains  $S_j$  and the transverse directions to  $S_j$ . Furthermore, the equilibria  $p'_j$ ,  $\tau p'_j$ ,  $p'_{j+1}$ , and  $\tau p'_{j+1}$  that are created in the pitchfork bifurcation are contained in  $\text{Fix}(\xi_j \tau)$ : in fact,  $p'_j$  and  $\tau p'_j$  are saddles, while  $p'_{j+1}$  and  $\tau p'_{j+1}$  are sinks in  $\text{Fix}(\xi_j \tau)$ . A continuity argument establishes the existence of heteroclinic trajectories from  $p'_j$  to  $p'_{j+1}$  and, through the action of  $\tau$ , from  $\tau p'_j$  to  $\tau p'_{j+1}$ , which are all robust homoclinic cycles. Checking the eigenvalue conditions at the equilibrium  $p'_j$ , where the eigenvalues are close to those about  $p_j$ , proves asymptotic stability.

Transverse bifurcations are among the bifurcation scenarios that occur in models of magnetic dynamos in rotating Bénard convection; see [92]. In [119], transverse bifurcations from an asymptotically stable homoclinic network in five-dimensional differential equations were considered, which lead to an essentially asymptotically stable homoclinic network (see [Section 5.5.2](#) for this notion) that is contained in an asymptotically stable heteroclinic network.

Our next example of bifurcations from robust cycles is an inclination-flip bifurcation of a simple type-A homoclinic cycle in  $\mathbb{R}^4$  that has been studied by Worfolk [425]. Consider the action of  $\mathbb{Z}_4 \ltimes \mathbb{Z}_2^4$  generated by  $(x_1, x_2, x_3, x_4) \mapsto (-x_1, x_2, x_3, x_4)$  and  $(x_1, x_2, x_3, x_4) \mapsto (x_4, x_1, x_2, x_3)$ . Let  $\Gamma$  be the subgroup of index two consisting of those elements that preserve orientation.

**THEOREM 5.74 ([425]).** *Let  $\dot{u} = f(u, \mu)$  with  $u = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$  be a one-parameter family on  $\mathbb{R}^4$  which is equivariant under the action of  $\Gamma$  as stated above. Suppose that the family has a robust homoclinic cycle and that, at  $\mu = 0$ , the stable manifold  $W^s(p_{i+1}, 0)$  is in an inclination-flip configuration so that [Hypothesis 2.12\(1\)](#) is violated. For a generic family of equivariant ODEs for which [Hypotheses 2.9, 2.10](#),*

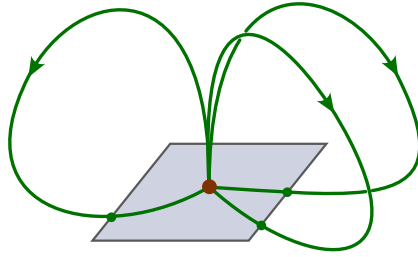


Fig. 5.27. Illustration of three homoclinic trajectories approaching  $p$  when  $\dim E_p^s = 2$ . Note that the state space has to be at least four dimensional.

2.11(2), and 2.12(2), (3), (4) are met, there exist values of  $\mu$  arbitrarily close to zero for which there exist hyperbolic invariant chaotic sets.

The genericity conditions referred to in the preceding theorem involve smooth linearizability assumptions. We remark that local bifurcations in  $\Gamma$ -equivariant ODEs have been studied in [168]: the starting point is the system

$$\begin{cases} \dot{x}_1 = x_1 + ax_1r^2 + bx_1x_2^2 + cx_1x_3^2 + dx_1x_4^2 + ex_2x_3x_4, \\ \dot{x}_2 = x_2 + ax_2r^2 + bx_2x_3^2 + cx_2x_4^2 + dx_2x_1^2 - ex_1x_3x_4, \\ \dot{x}_3 = x_3 + ax_3r^2 + bx_3x_4^2 + cx_3x_1^2 + dx_3x_2^2 + ex_1x_2x_4, \\ \dot{x}_4 = x_4 + ax_4r^2 + bx_4x_1^2 + cx_4x_2^2 + dx_4x_3^2 - ex_1x_2x_3, \end{cases} \quad (5.35)$$

where  $r^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$ , see also (5.34), and it is illustrated how complicated dynamics, including Shil'nikov homoclinic loops, arise in unfoldings. The inclination-flip homoclinic cycle from Theorem 5.74 may possibly occur in the preceding cubic normal form.

Having discussed bifurcations from robust homoclinic cycles, we focus now on bifurcations from codimension-one homoclinic cycles. An example of such a bifurcation is the Takens–Bogdanov bifurcation in  $\mathbb{R}^4$  with  $\mathbb{D}_3$ -symmetry that has been studied by Matthies [274]. The action of  $\mathbb{D}_3$  on  $\mathbb{R}^4 \sim \mathbb{C}^2$  is generated by  $(u_1, u_2) \mapsto (e^{2\pi i/3}u_1, e^{2\pi i/3}u_2)$  and  $(u_1, u_2) \mapsto (\overline{u_1}, \overline{u_2})$ , and we assume that the  $\mathbb{D}_3$  invariant equilibrium at the origin undergoes the equivariant analogue of the Takens–Bogdanov bifurcation, where the linearization at the origin is given by  $\dot{u}_1 = u_2, \dot{u}_2 = 0$ . The unfolding of this bifurcation contains a  $\mathbb{D}_3$ -symmetric configuration of three homoclinic loops to the origin (resembling a clover of homoclinic loops; see Figure 5.27) that occurs on a one-sided branch. The linearization of the ODE on the two-dimensional stable and unstable directions is semi-simple, and the unfolding of the homoclinic cycle near the homoclinic branch contains a suspended topological Markov chain.

Below, we discuss generalizations of this example in more detail. An important difference compared with homoclinic loops in generic flows is the possibility that multiple semisimple eigenvalues can be enforced by the symmetry, as is the case in the situation studied in [274]. Symmetry can also force heteroclinic trajectories to approach



an equilibrium along non-leading directions, thus enforcing orbit-flip configurations (or inclination-flips, if this occurs in the adjoint system). The bifurcation result we present below allows for multiple semisimple eigenvalues, but assumes that heteroclinic trajectories approach equilibria along the leading directions.

Thus, consider a  $\Gamma$ -equivariant system that admits a homoclinic cycle which consists of the heteroclinic trajectories  $h_1, \dots, h_\ell$ , where  $h_j$  connects hyperbolic equilibria  $p_j$  to  $p_{j+1}$  (taking indices modulo  $\ell$ ). We will now assume that this connection is of codimension one inside the fixed-point space  $S_j$  of  $\Gamma_{h_j}$ . We write  $\text{ind}_{S_j}(p) := \dim[W^u(p) \cap S_j]$  for the Morse index of an equilibrium  $p$  inside  $S_j$ .

**HYPOTHESIS 5.50.** *We have  $\text{ind}_{S_j}(p_j) = \text{ind}_{S_j}(p_{j+1})$  for all  $j$ .*

Recall that  $E_j^{\text{ls}}$  and  $E_j^{\text{lu}}$  denote the leading stable and unstable eigenspaces of the Jacobian  $f_u(p_j, 0)$ .

**HYPOTHESIS 5.51.** *The isotropy groups  $\Gamma_{p_j}$  act real absolutely irreducibly on  $E_j^{\text{ls}}$  (so that  $v_j^s$  is real), and  $0 < -v_j^s < \text{Re } v_j^u$ .*

Denoting by  $E_{j,*}^{\text{ls}}$  and  $E_{j,*}^{\text{lu}}$  the leading stable and unstable eigenspaces of the adjoint Jacobian  $f_u(p_j, 0)^*$ . Assuming the nondegeneracy condition [Hypothesis 2.10\(1\)](#), there is a unique bounded solution  $\psi_j$  of the adjoint variational equation  $\dot{w} = -f_u(h_j, 0)^* w$  along the solution  $h_j$ , and we record that  $\psi_j$  is contained in  $S_j$ . The following conditions, which require that  $h_j$  and  $\psi_j$  decay with rates coming from the leading directions, are analogues of the orbit-flip and inclination-flip conditions we introduced for generic systems. For equivariant systems, these conditions are open, but not necessarily generic.

**HYPOTHESIS 5.52.** *Consider the following conditions on the geometry of  $h_j$ :*

- (1) *The limit  $v_{j+1}^s = \lim_{t \rightarrow \infty} h_j(t)/\|h_j(t)\|$  lies in the leading stable eigenspace so that  $v_{j+1}^s \in E_{j+1}^{\text{ls}}$ ;*
- (2)  *$S_j \cap E_{j,*}^{\text{ls}} \neq \{0\}$ , and the limit  $v_{j,*}^s = \lim_{t \rightarrow -\infty} \psi_j(t)/\|\psi_j(t)\|$  satisfies  $v_{j,*}^s \in E_{j,*}^{\text{ls}}$ .*

Take cross sections  $\Sigma_j$ , transverse to the heteroclinic trajectory  $h_j$ , related by symmetry. For a given  $\ell \times \ell$  matrix  $M$  with entries  $M_{ij} \in \{0, 1\}$ , let  $\mathcal{B}_M$  be the topological Markov chain defined by  $M$ . For  $\kappa \in \mathcal{B}_M$ , we denote a solution by  $u(\mu, \kappa)(\cdot)$  if there is a monotonically increasing sequence  $(\tau_i)_{i \in \mathbb{Z}}$  such that

$$u(\mu, \kappa)(\tau_i) \in \Sigma_{\kappa_i}, \quad \text{but } u(\mu, \kappa)(t) \notin \bigcup_{j=1}^{\ell} \Sigma_j \text{ for } t \notin \{\tau_i, i \in \mathbb{Z}\}.$$

We call  $\kappa$  the itinerary of  $u(\mu, \kappa)(\cdot)$ . Let  $\Pi(\cdot, \mu)$  be the first-return map defined on  $\bigcup_{j=1}^{\ell} \Sigma_j$  (more precisely, the domain of  $\Pi(\cdot, \mu)$  will be a subset of  $\bigcup_{j=1}^{\ell} \Sigma_j$ ):

$$\Pi(\cdot, \mu) : \bigcup_{j=1}^{\ell} \Sigma_j \rightarrow \bigcup_{j=1}^{\ell} \Sigma_j, \quad \Pi(u(\mu, \kappa)(\tau_i), \mu) = u(\mu, \kappa)(\tau_{i+1}). \quad (5.36)$$

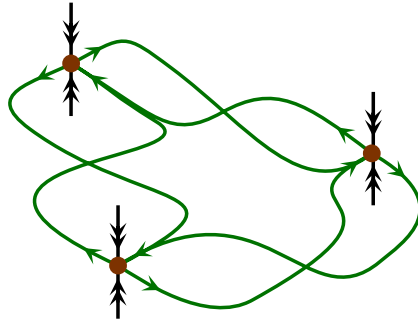


Fig. 5.28. An impression of a codimension-one  $\mathbb{D}_3$ -invariant homoclinic cycle for a  $\mathbb{D}_3$ -equivariant flow in  $\mathbb{R}^3$  with three  $\mathbb{Z}_2$ -symmetric equilibria. Note that a codimension-two bifurcation that involves resonance conditions, an inclination-flip, or an orbit-flip condition would lead to  $\mathbb{D}_3$ -symmetric singular hyperbolic attractors in an unfolding (akin to Lorenz-like attractors); see also Section 5.5.5.

The next theorem describes how recurrent sets change through the bifurcation. For a given matrix  $A$ , we write  $|A| = (|a_{ij}|)_{ij}$ .

**THEOREM 5.75 ([192]).** *Let  $\Gamma$  be a finite group and  $\dot{u} = f(u, \mu)$  be a one-parameter  $\Gamma$ -equivariant family of differential equations on  $\mathbb{R}^n$  that admits a homoclinic network for  $\mu = 0$ . Suppose that [Hypotheses 5.50–5.52](#) and [2.10](#) are met, then the system contains a recurrent set near the homoclinic network that is given as follows. Define the matrix  $A = (a_{ij})_{i,j \in \{1, \dots, m\}}$  by*

$$a_{ij} = \begin{cases} 0 & \text{if } \omega(h_i) \neq \alpha(h_j), \\ \text{sign}\langle v_i^s, v_{j,*}^s \rangle, & \text{if } \omega(h_i) = \alpha(h_j). \end{cases}$$

Write  $M_+ = \frac{1}{2}(A + |A|)$  and  $M_- = -\frac{1}{2}(A - |A|)$ . For  $\mu > 0$  small enough, there is an invariant set  $\mathcal{D}_\mu \subset \bigcup_{j=1}^\ell \Sigma_j$  of  $\Pi(\cdot, \mu)$  such that, for each  $\kappa \in \mathcal{B}_{M_+}$ , there exists a unique solution  $u(\mu, \kappa)(t)$  with  $u(\mu, \kappa)(0) \in \mathcal{D}_\mu$ . Moreover,  $\Pi(\cdot, \mu)|_{\mathcal{D}_\mu}$  is topologically conjugated to the left shift on  $\mathcal{B}_{M_+}$ . An analogous statement holds for  $\mu < 0$  with  $\mathcal{B}_{M_+}$  replaced by  $\mathcal{B}_{M_-}$ .

The recurrent trajectories described in the preceding theorem generate the entire recurrent set for  $\mu \neq 0$  only when the inner products  $\langle e_i^s, e_j^- \rangle$  are nonzero for all  $i, j$  with  $\omega(h_i) = \alpha(h_j)$ .

**5.5.4. Forced symmetry breaking.** In this section, we consider equivariant systems in which the some of the symmetries are broken upon changing parameters: in other words, we consider systems  $\dot{u} = f(u, \mu)$  that are equivariant under  $\Gamma$  for  $\mu = 0$  but respect only a proper subgroup  $\tilde{\Gamma}$  of  $\Gamma$  for  $\mu \neq 0$ . This situation is commonly referred to as *forced symmetry breaking*.<sup>19</sup>

<sup>19</sup>This contrasts so-called *spontaneous symmetry breaking* which involves changes in the isotropy group of a critical element (or a attractor or any transitive invariant set) for systems that are equivariant under a fixed symmetry group.

Forced symmetry breaking of relative equilibria in systems with continuous symmetries provides a mechanism for creating robust heteroclinic cycles. This research was initiated by Lauterbach and Roberts, who used forced symmetry breaking from  $SO(3)$  to the tetrahedral group  $\mathbb{T}$  as an example.

**THEOREM 5.76 ([257]).** *Let  $\dot{u} = f(u)$  be an  $SO(3)$ -equivariant system with an  $SO(3)$ -orbit  $\mathcal{E}$  of equilibria, whose isotropy subgroups are conjugate to  $O(2)$ , so that  $\mathcal{E}$  is a normally hyperbolic invariant manifold. Let  $\dot{u} = \tilde{f}(u)$  be a  $\mathbb{T}$ -equivariant system that is close to  $\dot{u} = f(u)$  in the  $C^1$  topology. Restricted to the perturbed invariant manifold  $\tilde{\mathcal{E}}$  near  $\mathcal{E}$ , the system  $\dot{u} = \tilde{f}(u)$  then admits equilibria with isotropy subgroups conjugate to  $\mathbb{D}_2$  and  $\mathbb{Z}_3$ ; in addition, it admits either equilibria with isotropy conjugate to  $\mathbb{Z}_2$  or robust heteroclinic cycles that connect the equilibria with  $\mathbb{D}_2$  symmetry. Moreover, there exist  $\mathbb{T}$ -equivariant equations  $\dot{u} = \tilde{f}(u)$  arbitrarily close to  $\dot{u} = f(u)$  in the  $C^1$  topology for which  $\tilde{\mathcal{E}}$  contains one of the following:*

- (1) *stable equilibria with  $\mathbb{Z}_3$  symmetry;*
- (2) *stable equilibria with  $\mathbb{D}_2$  symmetry and equilibria with  $\mathbb{Z}_2$  symmetry;*
- (3) *stable relative homoclinic cycles that connect the equilibria with  $\mathbb{D}_2$  symmetry.*

There is a dual statement on forced symmetry breaking from  $SO(3)$  to  $O(2)$ -symmetry for equilibria with isotropy  $\mathbb{T}$ , in which a circle of heteroclinic cycles occurs that connect equilibria with  $\mathbb{D}_2$  symmetry. If we break the symmetry from  $SO(3)$  to  $\mathbb{D}_n$ , then non-asymptotically stable heteroclinic cycles, see [Theorem 5.71](#), can bifurcate:

**THEOREM 5.77 ([257]).** *Let  $\dot{u} = f(u)$  be an  $SO(3)$ -equivariant system with an  $SO(3)$ -orbit  $\mathcal{E}$  of equilibrium points with isotropy subgroups conjugate to  $\mathbb{T}$  and assume that  $\mathcal{E}$  is a normally hyperbolic invariant manifold. Let  $\dot{u} = \tilde{f}(u)$  be a  $\mathbb{D}_n$ -equivariant system close to  $\dot{u} = f(u)$  in the  $C^1$  topology, then the perturbed invariant manifold  $\tilde{\mathcal{E}}$  near  $\mathcal{E}$  has the following properties:*

- (1) *If  $3|n$ , then  $\tilde{\mathcal{E}}$  contains equilibria or periodic solutions with  $\mathbb{Z}_3$  symmetry;*
- (2) *If  $n$  is odd, then  $\tilde{\mathcal{E}}$  contains equilibria or periodic solutions with  $\mathbb{Z}_2$  symmetry;*
- (3) *If  $n$  is even, then equilibria with  $\mathbb{D}_2$ -symmetry and heteroclinic cycles and/or equilibria with  $\mathbb{Z}_2$ -symmetry exist in  $\tilde{\mathcal{E}}$ .*

We refer the reader to [88,200,256,314] for further results of a similar flavour.

Next, we consider a few specific examples of the effect of forced symmetry breaking on robust relative homoclinic cycles. Recall the classification of robust relative homoclinic cycles in  $\mathbb{R}^3$  that we outlined in [Theorem 5.66](#). Let  $\Gamma = \mathbb{Z}_3 \ltimes \mathbb{Z}_2^3$  and suppose that  $\dot{u} = f_0(u)$  is a  $\Gamma$ -equivariant system on  $\mathbb{R}^3$  that admits a robust relative homoclinic cycle. We will consider perturbations  $\dot{u} = f(u, \mu)$  with  $f(u, 0) = f_0(u)$  that are only  $\mathbb{Z}_3$ -equivariant for all  $\mu$ . Generically, the two-dimensional fixed-point spaces that contain the homoclinic cycle for  $\mu = 0$  will no longer be invariant for  $\mu \neq 0$ , and we expect that the homoclinic cycle ceases to exist. The following result, adapted from [345], shows that the original symmetry  $\Gamma$  may force the homoclinic cycle to be in an inclination-flip configuration, which is then unfolded by small perturbations that retain  $\mathbb{Z}_3$  symmetry. We

remark that it is not hard to think of other examples where the symmetry enforces orbit-flips that can be unfolded by forced symmetry breaking. Let  $\mu \in \mathbb{R}^2$  and consider the following condition:

**HYPOTHESIS 5.53 (Inclination Flip).** *We assume that  $v_j^s, v_j^u, v_{j,*}^u \neq 0$  at  $\mu = 0$ , while  $v_{j,*}^s(0) = 0$  with  $\partial_{\mu_2} v_{j,*}^s(0) \neq 0$ .*

**THEOREM 5.78 ([345]).** *Let  $\dot{u} = f(u, \mu)$  be a  $\mathbb{Z}_3$ -equivariant two-parameter family on  $\mathbb{R}^3$  which, for  $\mu = 0$ , is  $\mathbb{Z}_3 \times \mathbb{Z}_2^3$ -equivariant and has a simple relative homoclinic cycle. Suppose that [Hypothesis 2.10\(1\)](#) with  $d_i = 0$  and [Hypothesis 2.12\(2\)](#) are met. If  $c < r$ , where  $c, r$  are the contracting and radial eigenvalues at the equilibria, then the relative homoclinic cycle is of inclination-flip type. Furthermore, if the unfolding condition [Hypothesis 5.53](#) is met, then the bifurcation diagram is as in [Theorems 5.1, 5.16](#) and [5.17](#), depending on the eigenvalue  $v^s = r$  and  $v^u = e$  (the expanding eigenvalue).*

**PROOF.** Symmetry forces the bounded solution  $\psi_j$  to the adjoint variational equation along the heteroclinic trajectory  $h_j$  to be perpendicular to the fixed point space of  $\Gamma_{h_j}$ . For  $c < r$ , the relative homoclinic cycle is therefore of inclination-flip type. An orbit space reduction [89] reduces the problem to a generic inclination-flip bifurcation of a homoclinic orbit.  $\square$

An analogous result holds for forced symmetry breaking of robust homoclinic cycles from  $\mathbb{Z}_2 \times \mathbb{Z}_2^2$  to  $\mathbb{Z}_2$  symmetry in systems in  $\mathbb{R}^3$ . A more detailed classification can be found in [345], where also four-dimensional representations are considered. The additional transverse eigenvalue of the robust homoclinic cycle could be a leading eigenvalue and must therefore be taken into account.

**5.5.5. Homoclinic orbits in systems with  $\mathbb{Z}_2$ -symmetry.** The unfolding of multiple homoclinic orbits to a hyperbolic equilibrium can lead to nontrivial dynamics, including suspended horseshoes. One reason to devote an individual section to homoclinic orbits in differential equations with reflection symmetries is that additional degeneracies may create Lorenz-like strange attractors. To see this, consider a  $\mathbb{Z}_2$ -equivariant differential equation with homoclinic orbits to a symmetric equilibrium  $p$ . We write  $\rho$  for the linear map that generates the action of  $\mathbb{Z}_2$ . If  $h$  is a homoclinic solution to  $p$ , then  $\rho h$  is a second homoclinic solution to  $p$ . Analogous to [Hypothesis 5.12](#), we distinguish different geometries of the closure of  $h \cup \rho h$ .

**HYPOTHESIS 5.54 (Geometric Configurations).** *Suppose that  $p$  has unique real leading eigenvalues  $v^s$  and  $v^u$  with eigenvectors  $v^s$  and  $v^u$ . We distinguish:*

- (1) *Figure-eight:*  $v^s = -\rho v^s$  and  $v^u = -\rho v^u$ ;
- (2) *Butterfly (expanding):*  $-v^s/v^u < 1$  and  $v^s = \rho v^s, v^u = -\rho v^u$ ;
- (3) *Butterfly (contracting):*  $-v^s/v^u > 1$  and  $v^s = \rho v^s, v^u = -\rho v^u$ ;
- (4) *Bellows:*  $v^s = \rho v^s$  and  $v^u = \rho v^u$ .

**THEOREM 5.79.** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of  $\mathbb{Z}_2$ -equivariant ODEs which admits two different homoclinic loops  $h$  and  $\rho h$  to the hyperbolic equilibrium  $p$  when  $\mu = 0$ . Suppose that [Hypotheses 2.1, 2.2, 2.3\(2\)](#), and [2.4](#) are met.*

- (1) If *Hypothesis 5.54*(1) or *5.54*(3) is met, then there are two nonsymmetric periodic solutions for  $\mu$  on one side of  $\mu = 0$  and a single symmetric periodic solution for  $\mu$  on the other side of  $\mu = 0$ .
- (2) If *Hypothesis 5.54*(2) or *Hypothesis 5.54*(4) holds, then there is no recurrent dynamics (apart from the equilibrium) for  $\mu$  on one side of  $\mu = 0$  and a suspended horseshoe for  $\mu$  on the other side of  $\mu = 0$ .

Shil'nikov [369,371] observed that Lorenz-like attractors could be created via three different codimension-two bifurcations in  $\mathbb{Z}_2$ -equivariant systems with two homoclinic loops: these are resonant leading eigenvalues, inclination-flips, and orbit-flips. The resonant bifurcation has subsequently been studied by Robinson, the inclination-flip by Rychlik, and the orbit-flip by Golmakani and Homburg. We now present the results of these analyses in the following three theorems.

First, we discuss  $\mathbb{Z}_2$  equivariant systems with homoclinic loops at resonance, which have been studied by Robinson; see also [287,288]. Although we only state a result on the existence of Lorenz-like attractors, unfoldings of homoclinic loops at resonance can also give rise to contracting Lorenz models [288,329]; see also Section 4.2.

**THEOREM 5.80** ([327–329]). *There exists an open set of two-parameter families of  $\mathbb{Z}_2$ -equivariant systems with the following properties. Each such family has two homoclinic loops for  $\mu = 0$  that satisfy *Hypotheses 2.1, 2.2, 2.3*(2), and *2.4*, and there is an open set in the parameter plane for which a Lorenz-like attractor exists.*

In fact, Robinson showed that Lorenz-like attractors occur in the cubic system

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= x - 2x^3 + \alpha y + \beta x^2 y \pm yz, \\ \dot{z} &= -\gamma z + \delta x^2.\end{aligned}$$

Recall that the Lorenz equations are quadratic. Rychlik proved a similar phenomenon in the unfolding of  $\mathbb{Z}_2$ -equivariant ODEs with inclination-flip homoclinic loops.

**THEOREM 5.81** ([336]). *Let  $\dot{u} = f(u, \mu)$  be a two-parameter family of  $\mathbb{Z}_2$ -equivariant ODEs in  $\mathbb{R}^3$  that admit a hyperbolic equilibrium  $p$  with one-dimensional unstable and two-dimensional stable manifolds and has two symmetry-related homoclinic orbits to  $p$  when  $\mu = 0$ . Suppose furthermore that *Hypotheses 2.1, 2.2* and *5.8* are met, and that this bifurcation is of Type B, as explained in Section 5.1.6. There are then parameter values arbitrarily close to  $\mu = 0$  for which the system has a Lorenz-like attractor.*

In fact, Rychlik showed that Lorenz attractors occur in the  $\mathbb{Z}_2$ -equivariant cubic system

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= x - 2x^3 + \alpha y + \beta x^2 y + \delta xz, \\ \dot{z} &= -\gamma z + x^2,\end{aligned}$$

which, up to a change of coordinates, coincides with the original Lorenz equations when

$\beta = 0$ . Finally, orbit-flip homoclinic bifurcations can give rise to Lorenz-like attractors in much the same way.

**THEOREM 5.82 ([157]).** *Let  $\dot{u} = f(u, \mu)$  be a two-parameter family of  $\mathbb{Z}_2$ -equivariant ODEs in  $\mathbb{R}^3$  that admit a hyperbolic equilibrium  $p$  with one-dimensional unstable and two-dimensional stable manifolds and has two symmetry-related homoclinic orbits to  $p$  when  $\mu = 0$ . Assume that the action of the symmetry has a one-dimensional fixed-point space, that [Hypotheses 2.1](#), [2.2](#) and [5.10](#) are met, and that this bifurcation is of Type B, as explained in [Section 5.1.7](#). There are then parameter values arbitrarily close to  $\mu = 0$  for which the system has a Lorenz-like attractor.*

Degenerate homoclinic loops of certain  $\mathbb{Z}_2$ -equivariant ODEs in  $\mathbb{R}^4$  have been investigated in [\[30,413\]](#). These bifurcation studies were motivated by the equations

$$\begin{aligned}\dot{s}_1 &= s_2, \\ \dot{s}_2 &= \frac{I - s_2 - \sin s_1 \cos r_1}{3 + \beta}, \\ \dot{r}_1 &= r_2, \\ \dot{r}_2 &= -\frac{r_2 + \sin r_1 \cos r_1}{\beta}\end{aligned}$$

which model two coupled Josephson junctions with a capacitive load [\[29\]](#). This system is equivariant under the action of  $\mathbb{Z}_2$  given by  $(s_1, s_2, r_1, r_2) \mapsto (s_1, s_2, -r_1, -r_2)$ , which fixes elements of the plane  $\{(s_1, s_2, r_1, r_2) \in \mathbb{R}^4; r_1 = r_2\}$ . There are values of  $\beta$  and  $I$  for which this system has a symmetric homoclinic loop in the fixed-point space along which the stable and unstable manifolds of the origin have a two-dimensional common tangent space. Among the phenomena found for nearby parameters is the existence of a one-sided curve in the parameter plane along which a pair of nonsymmetric homoclinic loops in a bellows configuration exists. Further analytic work revealed the existence of curves of pitchfork and periodic-doubling bifurcations of periodic solutions. Numerical work suggests a complicated bifurcation diagram including chaotic dynamics.

## 6. Related topics

In this section, a number of topics are reviewed that do not fit into the previous sections but are of interest to homoclinic bifurcation theory. Specifically, we discuss topological index theory for homoclinic orbits and moduli of stability associated to heteroclinic and homoclinic orbits in some detail. We also give an overview of existence theorems for homoclinic orbits, numerical techniques for continuing homoclinic orbits and detecting their bifurcations, variational methods for constructing homoclinic orbits in Hamiltonian systems, techniques for studying homoclinic orbits in singularly perturbed systems, and, finally, extensions to infinite-dimensional systems. However, for the latter topics, we will be brief and focus primarily on giving pointers to existing literature.

### 6.1. Topological indices

Homoclinic orbits occur along curves in two-dimensional parameter spaces. Of interest is then the fate of these curves for generic systems: the issue is whether or not we can list generic possibilities for start and end points of such curves and identify bifurcations along them that do not rely on unfolding and nondegeneracy conditions.

First, we discuss topological bifurcations that do not rely on unfolding conditions and present a prototypical result due to Nii. Recall that two-dimensional homoclinic centre manifolds are orientable annuli or non-orientable Möbius strips. Nii proved that, if we are given a path of homoclinic orbits in the parameter plane so that the orientations of the homoclinic centre manifold are different at the end points, then a 2-homoclinic orbit has branched somewhere along the path. We stress that there is no need to check unfolding conditions at the bifurcation, though we need nondegeneracy conditions at the end points of the path that guarantee the existence of two-dimensional homoclinic centre manifolds. The proof given by Nii uses Conley index theory, see [2,3] for an overview, to find the 2-homoclinic solutions; we remark that Lin's method can also be utilized to prove the following theorem.

**THEOREM 6.1** ([299]). *Let  $\dot{u} = f(u, \mu)$  be a two-parameter family of ODEs on  $\mathbb{R}^n$  so that a homoclinic orbit to a hyperbolic equilibrium  $p$  exists along a path  $s : [0, 1] \rightarrow \mathbb{R}^2$  in the parameter plane. Suppose that the equilibria along the path have one-dimensional unstable manifolds and that [Hypotheses 2.1](#), [2.2](#) and [2.3\(2\)](#) are met for all  $\mu \in s([0, 1])$ . Furthermore, we assume that the homoclinic orbit is not in a flip configuration so that [Hypothesis 2.4\(1\)–\(2\)](#) are met for  $\mu = s(0), s(1)$ . Finally, suppose that at  $\mu = s(0)$  the homoclinic centre manifold is orientable, while at  $\mu = s(1)$  the homoclinic centre manifold is non-orientable, then there exists a  $t \in (0, 1)$  so that a 2-homoclinic orbit is created in the homoclinic bifurcation at  $\mu = s(t)$ . The statement remains true if the homoclinic centre manifolds are non-orientable for  $\mu = s(0), s(1)$  but  $-v^s/v^u > 1$  at  $\mu = s(0)$ , and  $-v^s/v^u < 1$  at  $\mu = s(1)$ .*

Next, we discuss the fate of curves of homoclinic orbits in the bifurcation plane. In the spirit of the continuation theory for periodic orbits in [\[16\]](#), Fiedler [\[135\]](#) introduced an index for homoclinic orbits in the interior of the closure of the class of Morse–Smale flows (that is, for tame homoclinic orbits) and used this index to derive a path-following result for homoclinic orbits in generic families. The idea is to follow curves of homoclinic orbits up to the boundary (of the interior of the closure) of the class of Morse–Smale flows or to a collection of bifurcations of higher codimension.

Let us illustrate the outcome of the resulting homoclinic continuation theory by showing that certain bifurcation diagrams cannot occur in generic two-parameter families of three-dimensional flows. Here, we use the term ‘generic’ to refer to two-parameter systems that contain homoclinic bifurcations of codimension one and two that, furthermore, only unfold generically, that is, occur only along curves and at isolated points, respectively. Adopting the term ‘noose’ from the continuation theory of periodic orbits, we say that a bifurcation diagram of homoclinic orbits has a homoclinic noose if it contains the structures shown in [Figure 6.1](#). The following result can be deduced from the general homoclinic continuation theory that we describe further below.

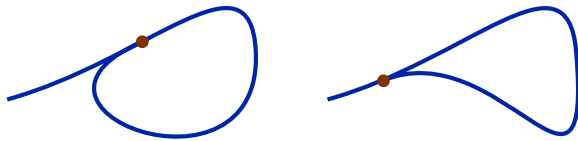


Fig. 6.1. Bifurcation curves of homoclinic orbits that contain a homoclinic noose.

**PROPOSITION 6.1.** *The bifurcation diagram of a generic two-parameter family of ODEs  $\dot{u} = f(u, \mu)$  in  $\mathbb{R}^3$  cannot contain homoclinic nooses.*

**PROOF.** Suppose that the bifurcation diagram of  $\dot{u} = f(u, \mu)$  contains a homoclinic noose. The branching point is then necessarily a homoclinic-doubling bifurcation, that is, either a resonant homoclinic bifurcation or an inclination-flip, or an orbit-flip (see Section 5), and the noose consists of the primary orbit that returns to itself as a 2-homoclinic orbit. Draw a smooth curve in the parameter plane close to the homoclinic noose and consider the bifurcations of periodic orbits along this curve. The bifurcation diagram for the periodic orbits in the parameter that parameterizes the chosen curve then contains a noose, which is impossible by [152,218].  $\square$

Other applications of a topological continuation theory for homoclinic orbits include the existence of homoclinic-doubling cascades, see Section 4.5, and of cascades of  $T$ -points in [198]. We remark that these results are similar in spirit and detail to the existence proofs of cascades of period-doubling bifurcations that use continuation theory for periodic orbits [430].

We now outline homoclinic continuation theory itself and follow the account given in [195] where a continuation result without genericity conditions is proved. Let  $\dot{u} = f(u, \mu)$  be a two-parameter family of ODEs in  $\mathbb{R}^3$ . We denote by  $\mathcal{P}$  the set of compact subsets of  $\mathbb{R}^3$ , equipped with the Hausdorff metric, and define

$$G = \left\{ (\mu, h) \in \mathbb{R}^2 \times \mathcal{P}; \begin{array}{l} h \text{ is the union of an equilibrium and a} \\ \text{homoclinic orbit of } \dot{u} = f(u, \mu) \end{array} \right\}. \quad (6.1)$$

For  $(\mu, h) \in G$ , let  $l(\mu, h)$  denote the arc length of  $h$ . For simplicity, we assume  $l(\mu, h)$  is finite, which is guaranteed, for instance, if the equilibrium in  $h$  is hyperbolic. We say that  $k$  is a virtual length of  $(\mu, h)$  if there exists a sequence of smooth perturbations  $f_i(\cdot, \nu)$  of the family  $f(\cdot, \nu)$  with  $f_i(\cdot, \nu) \rightarrow f(\cdot, \nu)$  as  $i \rightarrow \infty$  so that  $\dot{u} = f_i(u, \nu)$  has a homoclinic orbit  $h_i$  at parameter values  $\mu_i$  with  $\mu_i \rightarrow \mu$ ,  $h_i \rightarrow h$  in the Hausdorff topology, and  $l(\mu_i, h_i) \rightarrow k$  as  $i \rightarrow \infty$ . We write  $\tau(\mu, h)$  for the set of virtual lengths of  $(\mu, h) \in G$ .

To decide whether a homoclinic orbit can be continued globally, we associate an index with each such orbit. First, pick an  $(\mu, h) \in G$  and assume that it is a codimension-one homoclinic orbit that is unfolded generically upon varying  $\mu$  and satisfies  $\tau(\mu, h) = \{l(\mu, h)\}$ . There is then a sequence  $\mu_i$  of parameter values that converges to  $\mu$  and a sequence of periodic orbits  $q_i$  for these parameter values that converges to  $h$  in the Hausdorff topology as  $i \rightarrow \infty$ . Our assumptions imply furthermore that  $q_i$  is the unique periodic orbit for the parameter value equal to  $\mu_i$  for all sufficiently large  $i$  and that its



unstable manifold  $W^u(q_i)$  is either orientable or non-orientable. Define the index

$$\phi(\mu, h) = \begin{cases} 0 & \text{if } W^u(q_i) \text{ is non-orientable for large } i; \\ 1 & \text{if } W^u(q_i) \text{ is orientable for large } i. \end{cases} \quad (6.2)$$

Note that one- and three-dimensional unstable manifolds are always orientable. If  $W^u(q_i)$  is two-dimensional, then there exists a two-dimensional homoclinic centre manifold  $W^c(h)$  of  $h$ ; see Section 3.4: this manifold is orientable if  $\phi(\mu, h) = 1$  and non-orientable otherwise. In particular, it is possible to define  $\phi(\mu, h)$  using only the equation at the parameter value  $\mu$ .

Next, we extend this definition of  $\phi(\mu, h)$ , as follows, to the entire set  $G$ . For each  $(\mu, h) \in G$ , we set  $\phi(\mu, h) = 1$  if the virtual length of  $(\mu, h)$  is bounded and if there exists a sequence of families  $f_i(\cdot, v)$  with  $f_i(\cdot, v) \rightarrow f(\cdot, v)$  as  $i \rightarrow \infty$ , and  $f_i(\cdot, v)$  has a generically unfolded homoclinic orbit  $h_i$  of codimension one at parameter values  $\mu_i$  so that  $\mu_i \rightarrow \mu$ ,  $h_i \rightarrow h$  in the Hausdorff topology, and  $\phi(\mu_i, \tau_i) = 1$  as  $i \rightarrow \infty$ . For all other  $(\mu, h) \in G$ , we set  $\phi(\mu, h) = 0$ . Let

$$G_1 = \{(\mu, h) \in G; \phi(\mu, h) = 1\} \quad (6.3)$$

be the set of  $(\mu, h)$  of index one.

Finally, we can state precisely what the global continuation of homoclinic orbits  $(\mu, h)$  in  $G_1$  refers to. Let  $(\mu, h) \in G_1$  so that  $h$  is the union of a homoclinic orbit and a hyperbolic equilibrium, and write  $\Gamma_1$  for the connected component of  $G_1$  that contains  $(\mu, h)$ . We call  $(\mu, h)$  globally continuable if either

- $\Gamma_1 \setminus \{(\mu, h)\}$  is connected

or else each component  $C_1$  of  $\Gamma_1 \setminus \{(\mu, h)\}$  satisfies at least one of the following conditions:

- $C_1$  is unbounded;
- there exists a sequence  $(\mu_i, h_i) \in C_1$  so that  $\sup_i \tau(\mu_i, h_i) = \infty$  or so that  $(\mu_i, h_i) \rightarrow (\tilde{\mu}, \tilde{h}) \in G$  as  $i \rightarrow \infty$  and  $(\tilde{\mu}, \tilde{h})$  has unbounded virtual length;
- there exists a sequence  $(\mu_i, h_i) \in C_1$  so that  $\mu_i$  has a limit as  $i \rightarrow \infty$ , and  $h_i$  converges, in the Hausdorff topology, to a closed invariant set that contains either a nonhyperbolic equilibrium or more than two orbits.

Note that the closure of a homoclinic orbit always consists of two orbits: thus, if the closed invariant set consists of more than two orbits, it may contain two homoclinic orbits or a heteroclinic cycle.

**THEOREM 6.2 ([195]).** *A generically unfolded codimension-one homoclinic orbit in  $G_1$  is globally continuable.*

## 6.2. Moduli

Recall that two differential equations  $\dot{u} = f(u)$  and  $\dot{u} = g(u)$  on  $\mathbb{R}^n$  are topologically equivalent if there is a homeomorphism that maps orbits of  $\dot{u} = f(u)$  to orbits of  $\dot{u} = g(u)$ ,

Table 3. Moduli for heteroclinic orbits are listed, where  $\nu^s$  is the leading stable eigenvalue (or Floquet multiplier) at  $\alpha(h)$  and  $\nu^u$  is the leading unstable eigenvalue (or Floquet multiplier) at  $\omega(h)$ .

|             | $\alpha(h)$ | Equilibrium | Equilibrium                                                                                               | Periodic orbit                                                                  | Periodic orbit                                                                              |
|-------------|-------------|-------------|-----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| $\omega(h)$ | Leading     | Real        | Complex                                                                                                   | Real                                                                            | Complex                                                                                     |
| Equilibrium | Real        |             |                                                                                                           |                                                                                 |                                                                                             |
| Equilibrium | Complex     |             | $\frac{\operatorname{Re} \nu^s \operatorname{Im} \nu^u}{\operatorname{Im} \nu^s \operatorname{Re} \nu^u}$ | $\frac{\operatorname{Re} \nu^s}{\operatorname{Im} \nu^s} \frac{1}{\ln  \nu^u }$ | $\frac{\operatorname{Re} \nu^s}{\operatorname{Im} \nu^s} \frac{1}{\ln  \nu^u }, \arg \nu^u$ |
| Periodic    | Real        |             |                                                                                                           | $\frac{\ln  \nu^s }{\ln  \nu^u }$                                               | $\frac{\ln  \nu^s }{\ln  \nu^u }, \arg \nu^u$                                               |
| Periodic    | Complex     |             |                                                                                                           |                                                                                 | $\frac{\ln  \nu^s }{\ln  \nu^u }, \arg \nu^s, \arg \nu^u$                                   |

while preserving the direction of time. Any invariant of topological equivalence is called a modulus.

Palis [308] proved that heteroclinic orbits that involve a tangency between the stable and unstable manifolds of two periodic orbits give rise to a modulus that can be expressed as the quotient of the leading Floquet multipliers of the periodic orbits. Homoclinic tangencies of stable and unstable manifolds of periodic orbits can give rise to infinitely many moduli: see [159] for a survey of results. Homoclinic and heteroclinic orbits to equilibria can also give rise to moduli, and we remark that moduli near saddle-focus homoclinic orbits were treated in Theorem 5.5 in Section 5.1.2.

In this section, we review some other results in this direction. Moduli of stability that occur for systems with heteroclinic connections between critical elements (and for families that unfold these heteroclinic connections) have been studied by van Strien, extending and generalizing work by Belouqui [39] and Newhouse, Takens and Palis [295, 296, 391, 392]. The following result lists necessary and sufficient conditions for the existence of a topological equivalence near a heteroclinic orbit for two nearby vector fields.

**THEOREM 6.3 ([414]).** *Let  $\dot{u} = f(u)$  be an ODE on  $\mathbb{R}^n$  that has a heteroclinic orbit  $h(t)$  connecting hyperbolic critical elements (equilibria or periodic orbits)  $\alpha(h)$  to  $\omega(h)$  and suppose that additional generic conditions are met (for equilibria, these are Hypotheses 2.9, 2.10(1) with  $d = 1$ , 2.11(3), and 2.12). Table 3 contains the complete list of moduli of topological equivalence in this situation: if the moduli of two nearby vector fields with such heteroclinic orbits are equal, then there exists a topological equivalence by a near-identity homeomorphism in a neighbourhood of the closure of the heteroclinic orbit.*

Heteroclinic connections of codimension two are considered by Bonatti and Dufraine [51]. In [393], a complete set of three invariants of conjugacy are constructed for attracting planar heteroclinic cycles with two hyperbolic equilibria: the moduli arise because the time averages of continuous functions along orbits that converge to the heteroclinic cycle typically do not converge [148], and we refer to [393] for the relation between the moduli and these time averages.

Moduli are also relevant for the comparison of families of vector fields. Consider two families  $\dot{u} = f(u, \mu)$  and  $\dot{v} = g(v, \nu)$  on  $\mathbb{R}^n$ . A topological equivalence between these families is given by a family  $\Phi(\cdot, \mu)$  of homeomorphisms of  $\mathbb{R}^n$  and a homeomorphism  $\phi$

Table 4. Moduli for unfoldings of heteroclinic orbits. In the table,  $v^s$  is the leading stable eigenvalue (or Floquet multiplier) at  $\alpha(h)$  and  $v^u$  is the leading unstable eigenvalue (or Floquet multiplier) at  $\omega(h)$ .

|             | $\alpha(h)$ | Equilibrium | Equilibrium                                                                                                | Periodic orbit                                                   | Periodic orbit                                                             |
|-------------|-------------|-------------|------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|----------------------------------------------------------------------------|
| $\omega(h)$ | Leading     | Real        | Complex                                                                                                    | Real                                                             | Complex                                                                    |
| Equilibrium | Real        |             |                                                                                                            |                                                                  |                                                                            |
| Equilibrium | Complex     |             | $\frac{\operatorname{Re} v^s}{\operatorname{Im} v^s}, \frac{\operatorname{Re} v^u}{\operatorname{Im} v^u}$ | $\frac{\operatorname{Re} v^s}{\operatorname{Im} v^s}, \ln  v^u $ | $\frac{\operatorname{Re} v^s}{\operatorname{Im} v^s}, \ln  v^u , \arg v^u$ |
| Periodic    | Real        |             |                                                                                                            | $\ln  v^s , \ln  v^u $                                           | $\ln  v^s , \ln  v^u , \arg v^u$                                           |
| Periodic    | Complex     |             |                                                                                                            |                                                                  | $\ln  v^s , \ln  v^u , \arg v^s, \arg v^u$                                 |

on the parameter space so that  $(v, v) = (\Phi(u, \mu), \phi(\mu))$  relates their orbits. As in Section 5.4, one distinguishes different regularity properties of  $\Phi(\cdot, \mu)$  and  $\phi$ : the properties most relevant here are (fibre  $C^0, C^0$ )-equivalence (the above definition) and  $(C^0, C^0)$ -equivalence (where  $\Phi(u, \mu)$  is continuous in  $(u, \mu)$ ). We consider one-parameter families and assume the following:

**HYPOTHESIS 6.1 (Generic Unfolding).** *The unions of  $W^s(\omega(h))$  and  $W^u(\alpha(h))$  in the product  $\mathbb{R}^n \times \mathbb{R}$  of state and parameter space intersect transversally.*

The next theorem is the analogue of Theorem 6.3 for families that unfold a heteroclinic bifurcation of codimension one.

**THEOREM 6.4 ([414]).** *Let  $\dot{u} = f(u, \mu)$  be a one-parameter family of ODEs on  $\mathbb{R}^n$  that has a heteroclinic orbit  $h(t)$  for  $\mu = 0$  which connects hyperbolic critical elements (equilibria or periodic orbits)  $\alpha(h)$  to  $\omega(h)$ . Suppose that certain generic conditions are met (for equilibria, these are Hypotheses 2.9, 2.10(1) with  $d = 1$ , 2.11(3), and 2.12) and that the generic unfolding condition Hypothesis 6.1 holds. Table 4 then lists moduli of topological equivalence: If the moduli of two nearby families with such a heteroclinic orbit are equal, then there exists a  $(C^0, C^0)$ -equivalence by near-identity homeomorphisms in a neighbourhood of the closure of the heteroclinic orbit.*

A topological equivalence between two nearby families in  $\mathbb{R}^2$  that unfold a homoclinic bifurcation gives rise to a modulus if one requires that the parameter change is a diffeomorphism.

**THEOREM 6.5 ([125]).** *For  $j = 1, 2$ , let  $\dot{u} = f_j(u, \mu)$  be one-parameter families in  $\mathbb{R}^2$  that have a homoclinic orbit to hyperbolic equilibrium  $p_i$  at  $\mu = 0$ . Suppose that Hypotheses 2.2 and 2.3(2) are met for both vector fields and denote the stable and unstable leading eigenvalues at  $p_j$  by  $v_j^s$  and  $v_j^u$ . Suppose that there exists a topological equivalence  $(\Phi, \phi)$  between these two systems by a homeomorphism  $\Phi(\cdot, \mu)$  that depends continuously on  $\mu$  and a diffeomorphism  $\phi$ , then necessarily  $-v_1^s/v_1^u = -v_2^s/v_2^u$  at  $\mu = 0$ .*

The proof uses  $C^1$  linearizations near the equilibria to obtain the expressions  $x \mapsto \alpha_j(\mu) + x^{-v_j^s/v_j^u} (1 + \vartheta_j(x, \mu))$  for the first-return maps on curves that are transverse to the homoclinic orbits, where  $\alpha_j$  is differentiable, and  $\vartheta_j$  is continuous and vanishes along

$\{x = 0\}$  and  $\{\mu = 0\}$ . A conjugacy of the vector fields then implies a conjugacy of these maps that can be further analyzed to find moduli. [Theorem 6.5](#) can be applied to prove  $(C^0, C^\infty)$ -equivalence of generic families with Bogdanov–Takens bifurcations and their normal forms; see [Section 5.4.1](#).

We briefly comment on extensions to global stability where the conjugacy is not restricted to a neighbourhood of the heteroclinic connection. A family of vector fields is called structurally stable if every nearby family is  $(C^0, C^0)$ -equivalent. The following result by Labarca and Plaza characterizes structurally stable families that unfold heteroclinic bifurcations.

**THEOREM 6.6** ([\[249,252\]](#)). *A generic one-parameter family  $\dot{u} = f(u, \mu)$  on a compact three-dimensional manifold whose non-wandering set consists of finitely many hyperbolic critical elements and that has no heteroclinic cycles is structurally stable provided the following hold*

- (1) *Stable and unstable manifolds of periodic orbits intersect transversally;*
- (2) *If  $p$  is an equilibrium with one-dimensional unstable (stable) manifold and complex conjugate stable (unstable) eigenvalues, then  $W^u(p)$  is contained in the stable (unstable) manifold of an attracting (repelling) critical element.*

We refer to [\[416\]](#) for the global stability of families with non-trivial recurrent sets and unfoldings of heteroclinic bifurcations. An investigation of how the geometry of stable and unstable manifolds induces moduli can also be found in [\[188\]](#). An extension of [Theorem 6.6](#) by Plaza and Vera, incorporating local bifurcations, is contained in [\[316\]](#). One-parameter families of gradient vector fields on compact manifolds of any dimension turn out to be generically structurally stable [\[312\]](#), and the same is true for two-parameter families of gradient vector fields [\[72,415\]](#).

Finally, we consider Lorenz-like attractors. Guckenheimer and Williams established that geometric Lorenz models have two moduli.

**THEOREM 6.7** ([\[167\]](#)). *There is an open set  $\mathcal{U}$  in the space of smooth ODEs on  $\mathbb{R}^3$  and a continuous mapping  $k : \mathcal{U} \rightarrow \mathbb{R}^2$  so that the following holds. Each  $f \in \mathcal{U}$  has a Lorenz-like attractor, and  $f, \tilde{f} \in \mathcal{U}$  are topologically equivalent by a homeomorphism close to identity precisely if  $k(f) = k(\tilde{f})$ .*

The natural mapping  $k$  is obtained by considering the kneading sequences of the unstable manifolds [\[201,325\]](#). The above theorem becomes more precise in the language of interval maps; the statement below on expanding Lorenz maps applies to Lorenz-like vector fields by identifying points on leaves of the stable foliation on a cross section. We refer to [\[167,325,423\]](#) for the construction; see also [Section 3.5](#).

**HYPOTHESIS 6.2** (Expanding Lorenz Maps). *Consider  $f : [-1, 1] \rightarrow [-1, 1]$  that satisfy:*

- (1)  *$f$  is continuous and strictly increasing away from zero;*
- (2)  *$\lim_{x \uparrow 0} f(x) = 1$  and  $\lim_{x \downarrow 0} f(x) = -1$ ;*
- (3)  *$f$  is topologically expanding: there exists an  $\varepsilon > 0$  so that, for all  $x_0, y_0$  whose orbits do not contain zero,  $|f^i(x_0) - f^i(y_0)| > \varepsilon$  for some  $i \in \mathbb{N}$ .*

Given an expanding Lorenz map and a point  $x$  that is not a preimage of 0, define its kneading sequence  $k(x) \in \{-1, 1\}^{\mathbb{N}}$  by

$$k(x)(i) = \begin{cases} -1, & f^i(x) < 0, \\ 1, & f^i(x) > 0. \end{cases}$$

For general  $x \in [-1, 1]$ , we define its upper and lower kneading sequences by

$$k^+(x) = \lim_{y \downarrow x} k(y), \quad k^-(x) = \lim_{y \uparrow x} k(y),$$

where the limits run over all points  $y$  that are not preimages of 0, and define the *kneading invariant*

$$K(f) = (k^+(-1), k^-(1)).$$

We write  $\sigma : \{-1, 1\}^{\mathbb{N}} \rightarrow \{-1, 1\}^{\mathbb{N}}$  for the left shift operator defined by  $[\sigma\alpha](i) = \alpha(i+1)$  and take the lexicographical ordering on  $\{-1, 1\}^{\mathbb{N}}$ .

**THEOREM 6.8 ([201]).** *If  $f$  is an expanding Lorenz map, then the kneading invariant  $K(f) = (\alpha, \beta)$  satisfies*

$$\alpha \leq \sigma^n \alpha < \beta, \quad \alpha < \sigma^n \beta \leq \beta \tag{6.4}$$

for all  $n \in \mathbb{N}$ . Conversely, given a pair of sequences  $\alpha, \beta \in \{-1, 1\}^{\mathbb{N}}$  satisfying (6.4), there exists an expanding Lorenz map  $f$  with  $K(f) = (\alpha, \beta)$ , and  $f$  is unique up to conjugacy.

The combinatorial structure encoded by the kneading invariant is also apparent in the organization of heteroclinic bifurcation curves as presented in the numerical study of the Lorenz equations in [117].

### 6.3. Existence results

The existence of homoclinic orbits has been proved in many concrete models and applications. We give a few examples here and refer otherwise to some of the references listed in Section 5 for many other examples: in particular, we mention the review of local bifurcations in Section 5.4 that lead to small homoclinic orbits, and the cubic equivariant vector fields given in Section 5.5.5 that admit two homoclinic loops to the origin at a resonance or inclination-flip bifurcation, which lead to geometric Lorenz attractors in appropriate unfoldings.

Sandstede provided in [339] a general construction of vector fields which have homoclinic orbits that undergo various codimension-two bifurcations. The idea is to begin with a two-dimensional vector field that leaves a planar algebraic curve invariant and has a homoclinic solution which lies on this curve. A third coordinate is then added in such a way that the geometry near the algebraic curve can be changed, while the curve itself remains invariant. Further perturbations can now be added to break the homoclinic orbit.

This construction results in the system

$$\begin{aligned}\dot{x} &= ax + by - ax^2 + (\mu_2 - \alpha z)x(2 - 3x) + \delta z, \\ \dot{y} &= bx + ay - \frac{3}{2}bx^2 - \frac{3}{2}axy - (\mu_2 - \alpha z)2y - \delta z, \\ \dot{z} &= cz + \mu_1 x + \gamma xz + \alpha\beta(x^2(1 - x) - y^2)\end{aligned}$$

that involves the real parameters  $a, b, c, \alpha, \beta, \gamma, \mu_1, \mu_2$  and  $\delta \in \{0, 1\}$ .

**THEOREM 6.9 ([339]).** *For  $\mu = 0$ , the above system has a homoclinic orbit which is contained in the Cartesian leaf  $\Gamma = \{(x, y, z) \in \mathbb{R}^3; x^2(1 - x) - y^2 = 0, z = 0\}$ . First, suppose that  $\delta = 0$ , then the eigenvalues of the linearization at the origin are real, and the following codimension-two homoclinic bifurcations occur:*

- *A resonant bifurcation occurs for  $a = 0$  if  $c < -\sqrt{b^2 - 4\mu_2^2}$  and for  $c = a + \sqrt{b^2 - 4\mu_2^2}$  otherwise. This bifurcation is unfolded by  $\mu_1$  and  $a$ .*
- *An inclination-flip occurs for  $c < a - b$  and  $\beta = 1$ . This bifurcation is unfolded by  $\mu_1$  and  $\alpha - \alpha_0$  for a certain  $\alpha_0$  that depends on  $a, b, c$  and  $\gamma$ .*
- *An orbit-flip occurs for  $c > a - b$ ,  $\beta = 0$  and sufficiently small  $\alpha > 0$ . The unfolding parameters are  $\mu_1, \mu_2$ .*

*Next, suppose that  $\delta = 1$ , then the eigenvalues of the linearization at the origin consists of a complex conjugate pair and a real eigenvalue, and a saddle-focus homoclinic orbit occurs for  $c = a - b$ ,  $\gamma = 0$ ,  $\alpha = 0$ .*

Finally, we comment on the Lorenz system

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= \rho x - y - xz, \\ \dot{z} &= -\beta z + xy\end{aligned}$$

for which a number of results have been obtained that prove rigorously the existence of homoclinic orbits for some parameter values and which do not rely on numerical computations (rigorous or otherwise). In particular, a theoretical existence proof for homoclinic orbits has been given in [176], using an analytic implementation of the shooting method: the authors prove that, for  $\sigma$  near 10 and  $\beta$  near 1, there exists a  $\rho \in (1, 1000)$  for which the Lorenz equations have a homoclinic orbit. Leonov [258] states that, for  $\sigma > (2\beta + 1)/3$ , there is a  $\rho > 1$  for which the Lorenz equations have a homoclinic orbit; this refines earlier results from [42].

#### 6.4. Numerical techniques

Since homoclinic and heteroclinic orbits are genuinely global dynamical objects, it is typically very hard to prove their existence and nondegeneracy in a given explicit system of differential equations. In these situations, numerical computations are often the only way to get insight into the existence and bifurcation structure of connecting orbits.

The most efficient and accurate algorithms seek homoclinic and heteroclinic orbits as solutions to appropriate boundary-value problems on bounded time intervals  $(-T, T)$ . Path-following in systems parameters can then be used to continue connecting orbits and to locate parameter values where they undergo bifurcations. We refer to the surveys [1, Sections 6 and 8] and [6, Section 6.1] for details and references.

Algorithms of this kind can also be used for large systems, and specifically for discretizations of partial differential equations, and we refer to [83] for an overview. They have also been applied to delay differential equations [337] and even to equations with advanced and retarded terms [8].

These methods have been implemented via the driver HOMCONT [82] in the software package AUTO [118] by Doedel. AUTO also allows users to switch to  $N$ -homoclinic orbits at bifurcation points using a numerical implementation of Lin's method developed in [301]. DDE-BIFTOOL is a MATLAB package that implements a similar functionality for delay differential equations [131].

### 6.5. Variational methods

For Hamiltonian systems of the form

$$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \nabla H(p, q), \quad (p, q) \in \mathbb{R}^{2n}, \quad (6.5)$$

global methods from the calculus of variations can often be used to prove the existence of homoclinic, heteroclinic and periodic orbits. To find connecting orbits in this manner, an appropriate variational formulation needs to be set up whose critical points are the desired heteroclinic orbits. The difficulty lies in finding variational formulations to which minimization techniques, mountain-pass theorems or other global methods can be applied by verifying the necessary hypotheses.

This approach has been used successfully to construct  $N$ -pulses (multibump orbits) near given primary homoclinic orbits without having to impose any nondegeneracy conditions. In particular, multibump orbits have been constructed near homoclinic orbits to bi-foci in [61, 63, 214]. Such orbits have also been found in four-dimensional systems where the primary homoclinic orbit converges to a centre with non-semisimple eigenvalues on the imaginary axis [49] and in systems that have two primary homoclinic orbits to the same saddle equilibrium [46]. In [379, 407, 408], variational methods were used to find multibump orbits in the Swift–Hohenberg equation.

We refer to the survey [5] for more details and a comprehensive list of references.

### 6.6. Singularly perturbed systems

Many physical systems admit two or more natural time scales. In ODE models, multiple time scales typically manifest themselves via the presence of small parameters that multiply the time derivatives of some of the variables. We shall give a very brief outlook

of such singularly perturbed problems. Specifically, we consider systems of the form

$$\begin{aligned}\dot{u} &= f_1(u, v) \\ \dot{v} &= \epsilon f_2(u, v),\end{aligned}\tag{6.6}$$

where  $t$  is the time variable and  $(u, v) \in \mathbb{R}^{n_1 \times n_2}$ . The singular character of these equations is reflected in the assumption that  $0 < \epsilon \ll 1$ . That the perturbation caused by  $\epsilon$  is singular and becomes more visible if we use the slow time  $s = \epsilon t$  which yields the slow system

$$\begin{aligned}\epsilon u' &= f_1(u, v) \\ v' &= f_2(u, v)\end{aligned}\tag{6.7}$$

in the slow time variable  $s$ . For  $\epsilon > 0$ , Equations (6.6) and (6.7) are equivalent. However, for  $\epsilon = 0$ , we obtain the fast system

$$\dot{u} = f_1(u, v), \quad \dot{v} = 0,\tag{6.8}$$

where  $v$  plays the role of a parameter, and the slow system

$$v' = f_2(u, v), \quad f_1(u, v) = 0,\tag{6.9}$$

which is a differential-algebraic system, where  $v$  is constrained to the surface  $\mathcal{M}_0 = \{(u, v); f_1(u, v) = 0\}$ , which we refer to as the slow manifold. Note that the elements of  $\mathcal{M}_0$  correspond to the equilibria of (6.8). Since the systems (6.8) and (6.9) are equations with fewer dependent variables, we can exploit this reduction in dimension to understand the occurrence and bifurcations of homoclinic and heteroclinic orbits for  $\epsilon > 0$  by investigating the two systems (6.8) and (6.9) for  $\epsilon = 0$  separately and gluing their solutions together to get solutions that persist for  $\epsilon > 0$ .

Rigorous matched asymptotic expansion provides one possible avenue for gluing slow and fast solutions together, and we refer to [262] for further details. Geometric singular perturbation theory, originating in work by Fenichel, offers an alternative approach: if the slow manifold  $\mathcal{M}_0$  is normally hyperbolic for (6.8), then it persists as an invariant manifold  $\mathcal{M}_\epsilon$  of (6.6) for  $\epsilon > 0$  with dynamics that is close to the dynamics of (6.9) on  $\mathcal{M}_0$ . The so-called Exchange Lemma [60,213,245,351], due to Jones and Kopell, then describes the dynamics near the manifold  $\mathcal{M}_\epsilon$  and allows one to carry out the matching of slow and fast solutions, and we refer to [213] for a review of this approach; see Figure 6.2 for an illustration. If  $\mathcal{M}_0$  is not normally hyperbolic, then geometric blowup can often be used to analyze the dynamics, and [126,246] contain results in this direction.

Last, we mention that homoclinic orbits in near-integrable Hamiltonian systems are often the building blocks of complicated chaotic behaviour, and we refer to [173] for a comprehensive book on this topic.

## 6.7. Infinite-dimensional systems

Many of the results reviewed and summarized in this survey can be generalized to infinite-dimensional dynamical systems. We give a brief list of such systems and a few pointers to the relevant literature.



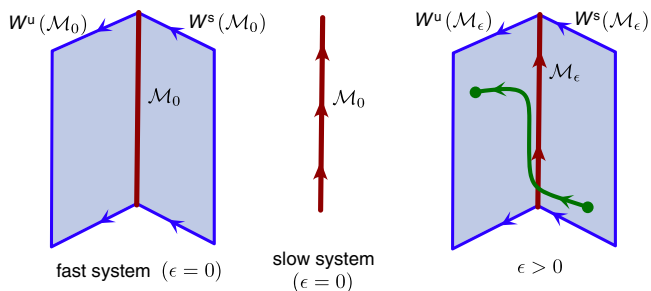


Fig. 6.2. Shown are the dynamics of the fast system [left] and the slow system [centre] for  $\epsilon = 0$ , and the anticipated dynamics for  $\epsilon > 0$  [right].

Delay differential equations are systems where the evolution of the solution  $u(t)$  depends not only on its state at time  $t$  but also on its history: they occur often as models in population dynamics, in laser systems, and in system with time-delayed feedback. A typical delay differential equation is of the form

$$\dot{u}(t) = f(u(t), u(t - \tau)), \quad u \in \mathbb{R}^n,$$

where  $\tau > 0$  is the temporal delay. Such equations generate dynamical systems on the function space  $C^0([-\tau, 0], \mathbb{R}^n)$ . Both geometric and analytical approaches can be used to study homoclinic and heteroclinic bifurcations in delay differential equations, and we refer to [171, 172] and [426] for results and further references.

Functional differential equations of mixed type (FDEs) are of the form

$$\dot{u}(t) = \sum_{j=-m}^m \alpha_j u(t + j) + f(u(t)), \quad u \in \mathbb{R}^n, \quad (6.10)$$

for constants  $\alpha_j \in \mathbb{R}$  and  $m \geq 1$ . Thus, in contrast to delay differential equations, the equation for the rate of change of  $u$  at time  $t$  depends here not only on the past but also on the future. FDEs are ill-posed in the sense that, given an initial condition  $u(t)$  defined on the interval  $[-m, m]$ , a solution to (6.10) may not exist. In particular, FDEs do not generate a flow on an appropriate function space, which prevents us from studying homoclinic bifurcations using geometric Poincaré-map based approaches. Now however, the existence of exponential dichotomies has been established for FDEs; see [175, 271]. This opens up the possibility of using Lin's method for studying homoclinic bifurcations for FDEs and work in this direction can be found in [150, 202].

Partial differential equations (PDEs) provide another important class of infinite-dimensional dynamical systems. Consider, for instance, systems of parabolic partial differential equations

$$u_t = D\Delta u + f(u), \quad x \in \Omega \subset \mathbb{R}^d, \quad u \in \mathbb{R}^n$$

with Dirichlet or Neumann conditions, where  $\Delta = \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2}$  denotes the Laplace operator

on a domain  $\Omega$  with a smooth boundary,  $D$  is a diagonal positive matrix, and  $f$  is a smooth nonlinearity. More generally, we may consider an abstract system of the form

$$\dot{u} = Au + f(u), \quad (6.11)$$

where  $A$  is a sectorial operator with dense domain defined on some Banach space  $X$ . These equations generate semiflows: solutions with prescribed initial data at  $t = 0$  exist for  $t > 0$  but not necessarily for  $t < 0$  in backward time. Having a semiflow available is sufficient to use many of the analytical and geometric techniques we discussed in Section 3, and we refer to [181] for results in this direction. In particular, the bifurcation of periodic orbits with large period from a homoclinic orbit, considered in Section 3.6, was investigated in [94]. We also mention that the homoclinic centre manifold theory developed in [342] is applicable to Equation (6.11) with  $A$  sectorial.

Homoclinic bifurcation theory for elliptic partial differential equations

$$\Delta u + f(u) = 0, \quad x \in \Omega \times \mathbb{R}, \quad u \in \mathbb{R}^n$$

on cylindrical domains is often of interest as elliptic PDEs on cylinders arise when studying travelling waves of parabolic PDEs on such domains. Similar to FDEs, elliptic PDEs are ill-posed as initial-value problems. Exponential dichotomy theory for such systems was developed in [315], and we refer to [315,346,347] for results on homoclinic and heteroclinic bifurcations for elliptic and pseudo-elliptic PDEs.

## References

### *Surveys in volumes 1–3*

- [1] W.-J. Beyn, A.R. Champneys, E. Doedel, W. Govaerts, Y.A. Kuznetsov and B. Sandstede, *Numerical continuation, and computation of normal forms*, Handbook of Dynamical Systems, Vol. 2, North-Holland, Amsterdam (2002), 149–219.
- [2] J. Franks and M. Misiurewicz, *Topological methods in dynamics*, Handbook of Dynamical Systems, Vol. 1A, North-Holland, Amsterdam (2002), 547–598.
- [3] K. Mischaikow and M. Mrozek, *Conley index*, Handbook of Dynamical Systems, Vol. 2, North-Holland, Amsterdam (2002), 393–460.
- [4] E.R. Pujals and M. Sambarino, *Homoclinic bifurcations, dominated splitting, and robust transitivity*, Handbook of Dynamical Systems, Vol. 1B, Elsevier B.V., Amsterdam (2006), 327–378.
- [5] P.H. Rabinowitz, *Variational methods for Hamiltonian systems*, Handbook of Dynamical Systems, Vol. 1A, North-Holland, Amsterdam (2002), 1091–1127.
- [6] B. Sandstede, *Stability of travelling waves*, Handbook of Dynamical Systems, Vol. 2, North-Holland, Amsterdam (2002), 983–1055.
- [7] F. Takens and A. Vanderbauwhede, *Local invariant manifolds and normal forms*, Handbook of Dynamical Systems, Vol. 3, North-Holland, Amsterdam (2010), 89–124.

### *Other sources*

- [8] K.A. Abell, C.E. Elmer, A.R. Humphries and E.S. Van Vleck, *Computation of mixed type functional differential boundary value problems*, SIAM J. Appl. Dyn. Syst. **4** (2005), 755–781.
- [9] A. Afendikov and A. Mielke, *Bifurcation of homoclinic orbits to a saddle-focus in reversible systems with  $SO(2)$ -symmetry*, J. Differential Equations **159** (1999), 370–402.
- [10] V.S. Afraïmovich, V.V. Bykov and L.P. Shil'nikov, *On structurally unstable attracting limit sets of Lorenz attractor type*, Trans. Mosc. Math. Soc. **1983** (1983), 153–216.

- [11] V.S. Afraïmovich, S.-N. Chow and W. Liu, *Lorenz-type attractors from codimension one bifurcation*, J. Dynam. Differential Equations **7** (1995), 375–407.
- [12] V.S. Afraïmovich and S.-B. Hsu, *Lectures on Chaotic Dynamical Systems*, American Mathematical Society, Providence, RI (2003).
- [13] V.S. Afraïmovich and Y.B. Pesin, *Dimension of Lorenz type attractors*, Mathematical Physics Reviews, Vol. 6, Harwood Academic Publishers, Chur (1987), 169–241.
- [14] M.A.D. Aguiar, S.B.S.D. Castro and I.S. Labouriau, *Dynamics near a heteroclinic network*, Nonlinearity **18** (2005), 391–414.
- [15] M.A.D. Aguiar, S.B.S.D. Castro and I.S. Labouriau, *Simple vector fields with complex behavior*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **16** (2006), 369–381.
- [16] K.T. Alligood and J.A. Yorke, *Families of periodic orbits: virtual periods and global continuability*, J. Differential Equations **55** (1984), 59–71.
- [17] C.J. Amick and J.F. Toland, *Global uniqueness of homoclinic orbits for a class of fourth order equations*, Z. Angew. Math. Phys. **43** (1992), 591–597.
- [18] A.A. Andronov, E.A. Leontovich, I.I. Gordon and A.G. Maïer, *Theory of Bifurcations of Dynamic Systems on a Plane*, Halsted Press (1973).
- [19] V. Araújo, *Large deviations bound for semiflows over a non-uniformly expanding base*, Bull. Braz. Math. Soc. (N.S.) **38** (2007), 335–376.
- [20] V. Araújo and M.J. Pacífico, *Three Dimensional Flows*, Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro (2007).
- [21] V. Araújo and M.J. Pacífico, *Physical measures for infinite-modal maps*, Fund. Math. **203** (2009), 211–262.
- [22] V. Araújo, M.J. Pacífico, E.R. Pujals and M. Viana, *Singular-hyperbolic attractors are chaotic*, Trans. Amer. Math. Soc. **361** (2009), 2431–2485.
- [23] D. Armbruster, J. Guckenheimer and P.J. Holmes, *Heteroclinic cycles and modulated travelling waves in systems with  $O(2)$  symmetry*, Physica D **29** (1988), 257–282.
- [24] D. Armbruster, E. Stone and V. Kirk, *Noisy heteroclinic networks*, Chaos **13** (2003), 71–86.
- [25] A. Arneodo, P. Couillet, E.A. Spiegel and C. Tresser, *Asymptotic chaos*, Physica D **14** (1985), 327–347.
- [26] A. Arneodo, P. Couillet and C. Tresser, *A possible new mechanism for the onset of turbulence*, Phys. Lett. A **81** (1981), 197–201.
- [27] A. Arneodo, P. Couillet and C. Tresser, *Possible new strange attractors with spiral structure*, Comm. Math. Phys. **79** (1981), 573–579.
- [28] V.I. Arnol'd, V.S. Afraïmovich, Y.S. Il'yashenko and L.P. Shil'nikov, *Bifurcation Theory and Catastrophe Theory*, Springer-Verlag, Berlin (1999).
- [29] D.G. Aronson, M. Golubitsky and M. Krupa, *Coupled arrays of Josephson junctions and bifurcation of maps with  $S_N$  symmetry*, Nonlinearity **4** (1991), 861–902.
- [30] D.G. Aronson, S.A. van Gils and M. Krupa, *Homoclinic twist bifurcations with  $Z_2$  symmetry*, J. Nonlinear Sci. **4** (1994), 195–219.
- [31] A. Arroyo and F. Rodríguez Hertz, *Homoclinic bifurcations and uniform hyperbolicity for three-dimensional flows*, Ann. Inst. H. Poincaré Anal. Non Linéaire **20** (2003), 805–841.
- [32] P. Ashwin and M.J. Field, *Heteroclinic networks in coupled cell systems*, Arch. Ration. Mech. Anal. **148** (1999), 107–143.
- [33] P. Ashwin and J. Montaldi, *Group theoretic conditions for existence of robust relative homoclinic trajectories*, Math. Proc. Cambridge Philos. Soc. **133** (2002), 125–141.
- [34] I. Baldomá and T.M. Seara, *Breakdown of heteroclinic orbits for some analytic unfoldings of the Hopf-zero singularity*, J. Nonlinear Sci. **16** (2006), 543–582.
- [35] R. Bamón, R. Labarca, R. Mañé and M.J. Pacífico, *The explosion of singular cycles*, Inst. Hautes Études Sci. Publ. Math. **78** (1993), 207–232. 1994.
- [36] M. Beck, J. Knobloch, D.J.B. Lloyd, B. Sandstede and T. Wagenknecht, *Snakes, ladders, and isolas of localized patterns*, SIAM J. Math. Anal. **41** (2009), 936–972.
- [37] G.R. Belitskiĭ, *Functional equations, and conjugacy of local diffeomorphisms of finite smoothness class*, Funkcional. Anal. i Priložen. **7** (1973), 17–28.
- [38] G.R. Belitskiĭ and A.Y. Kopanskii, *Equivariant Sternberg–Chen theorem*, J. Dynam. Differential Equations **14** (2002), 349–367.
- [39] J.A. Beloqui, *Modulus of stability for vector fields on 3-manifolds*, J. Differential Equations **65** (1986), 374–395.

- [40] L.A. Belyakov, *Bifurcation set in a system with homoclinic saddle curve*, Math. Notes **28** (1981), 910–916.
- [41] L.A. Belyakov, *Bifurcations of systems with a homoclinic curve of the saddle-focus with a zero saddle value*, Mat. Zametki **36** (1984), 681–689, 798.
- [42] V.N. Belykh, *Bifurcation of separatrices of a saddle point of the Lorenz system*, Differ. Equ. **20** (1984), 1184–1191.
- [43] V.N. Belykh and V.V. Bykov, *Bifurcations for heteroclinic orbits of a periodic motion and a saddle-focus and dynamical chaos*, Chaos Solitons Fractals **9** (1998), 1–18.
- [44] M.S. Berger, *Nonlinearity and Functional Analysis*, Academic Press, New York (1977).
- [45] P. Bergé, Y. Pomeau and C. Vidal, *Order within Chaos: Towards a Deterministic Approach to Turbulence*, John Wiley & Sons (1984).
- [46] M. Berti and P. Bolle, *Variational construction of homoclinics and chaos in presence of a saddle–saddle equilibrium*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **27** (1998), 331–377.
- [47] V.S. Biragov and L.P. Shil'nikov, *On the bifurcation of a saddle-focus separatrix loop in a three-dimensional conservative dynamical system*, Selecta Math. Soviet **11** (1992), 333–340.
- [48] R. Bogdanov, *Versal deformation of a singularity of a vector field on the plane in the case of zero eigenvalues*, Selecta Math. Soviet **1** (1976), 389–421.
- [49] P. Bolle and B. Buffoni, *Multibump homoclinic solutions to a centre equilibrium in a class of autonomous Hamiltonian systems*, Nonlinearity **12** (1999), 1699–1716.
- [50] C. Bonatti, L.J. Díaz and M. Viana, *Dynamics Beyond Uniform Hyperbolicity*, Springer-Verlag, Berlin (2005).
- [51] C. Bonatti and E. Dufraine, *Équivalence topologique de connexions de selles en dimension 3*, Ergodic Theory Dynam. Systems **23** (2003), 1347–1381.
- [52] C. Bonatti, A. Pumarino and M. Viana, *Lorenz attractors with arbitrary expanding dimension*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), 883–888.
- [53] P. Bonckaert, *Conjugacy of vector fields respecting additional properties*, J. Dynam. Control Systems **3** (1997), 419–432.
- [54] P. Bonckaert and P. De Maesschalck, *Gevrey and analytic local models for families of vector fields*, Discrete Contin. Dyn. Syst. Ser. B **10** (2008), 377–400.
- [55] A.R. Borisjuk, *Global bifurcations on the Klein bottle. The unimodal case*, Math. Notes **71** (2002), 316–329.
- [56] A.R. Borisjuk, *Global bifurcations on a Klein bottle. The general case*, Sb. Math. **196** (2005), 465–483.
- [57] W. Brannath, *Heteroclinic networks on the tetrahedron*, Nonlinearity **7** (1994), 1367–1384.
- [58] H.W. Broer, *Formal normal form theorems for vector fields and some consequences for bifurcations in the volume preserving case*, Dynamical Systems and Turbulence, Warwick 1980 (Coventry, 1979/1980), Lecture Notes in Math., Vol. 898, Springer, Berlin (1981), 54–74.
- [59] H.W. Broer and G. Vegter, *Subordinate Shil'nikov bifurcations near some singularities of vector fields having low codimension*, Ergodic Theory Dynam. Systems **4** (1984), 509–525.
- [60] P. Brunovský, *Tracking invariant manifolds without differential forms*, Acta Math. Univ. Comenian. (N.S.) **65** (1996), 23–32.
- [61] B. Buffoni, *Infinitely many large amplitude homoclinic orbits for a class of autonomous Hamiltonian systems*, J. Differential Equations **121** (1995), 109–120.
- [62] B. Buffoni, A.R. Champneys and J.F. Toland, *Bifurcation and coalescence of a plethora of homoclinic orbits for a Hamiltonian system*, J. Dynam. Differential Equations **8** (1996), 221–279.
- [63] B. Buffoni and E. Séré, *A global condition for quasi-random behavior in a class of conservative systems*, Comm. Pure Appl. Math. **49** (1996), 285–305.
- [64] J. Burke and E. Knobloch, *Localized states in the generalized Swift–Hohenberg equation*, Phys. Rev. E **73** (2006), 056211.
- [65] J. Burke and E. Knobloch, *Homoclinic snaking: structure and stability*, Chaos **17** (2007), 037102.
- [66] J. Burke and E. Knobloch, *Snakes and ladders: localized states in the Swift–Hohenberg equation*, Phys. Lett. A **360** (2007), 681–688.
- [67] V.V. Bykov, *The bifurcations of separatrix contours and chaos*, Physica D **62** (1993), 290–299.
- [68] V.V. Bykov, *One mechanism of appearance of a spiral quasitractor involving heteroclinic contours*, Comput. Math. Appl. **34** (1997), 229–244.
- [69] V.V. Bykov, *On bifurcations leading to chaos in Chua's circuit*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **8** (1998), 685–699.

- [70] V.V. Bykov, *On systems with separatrix contour containing two saddle-foci*, J. Math. Sci. (New York) **95** (1999), 2513–2522.
- [71] V.V. Bykov, *Orbit structure in a neighborhood of a separatrix cycle containing two saddle-foci*, Methods of Qualitative Theory of Differential Equations and Related Topics, Amer. Math. Soc, Providence, RI (2000), 87–97.
- [72] M.J.D. Carneiro and J. Palis, *Bifurcations and global stability of families of gradients*, Inst. Hautes Études Sci. Publ. Math. **70** (1989), 103–168. 1990.
- [73] J.C. Ceballos and R. Labarca, *A note on modulus of stability for cycles of the complex type*, Physica D **55** (1992), 37–44.
- [74] A.R. Champneys, *Subsidiary homoclinic orbits to a saddle-focus for reversible systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **4** (1994), 1447–1482.
- [75] A.R. Champneys, *Homoclinic orbits in reversible systems and their applications in mechanics, fluids and optics*, Physica D **112** (1998), 158–186.
- [76] A.R. Champneys, *Homoclinic orbits in reversible systems ii: Multi-bumps and saddle-centres*, CWI Quarterly **12** (1999), 185–212.
- [77] A.R. Champneys, *Codimension-one persistence beyond all orders of homoclinic orbits to singular saddle centres in reversible systems*, Nonlinearity **14** (2001), 87–112.
- [78] A.R. Champneys and J. Härterich, *Cascades of homoclinic orbits to a saddle-centre for reversible and perturbed Hamiltonian systems*, Dyn. Stab. Syst. **15** (2000), 231–252.
- [79] A.R. Champneys, J. Härterich and B. Sandstede, *A non-transverse homoclinic orbit to a saddle-node equilibrium*, Ergodic Theory Dynam. Systems **16** (1996), 431–450.
- [80] A.R. Champneys and V. Kirk, *The entwined wiggling of homoclinic curves emerging from saddle-node/Hopf instabilities*, Physica D **195** (2004), 77–105.
- [81] A.R. Champneys, V. Kirk, E. Knobloch, B.E. Oldeman and J.D.M. Rademacher, *Unfolding a tangent equilibrium-to-periodic heteroclinic cycle*, SIAM J. Appl. Dyn. Syst. **8** (2009), 1261–1304.
- [82] A.R. Champneys, Y.A. Kuznetsov and B. Sandstede, *A numerical toolbox for homoclinic bifurcation analysis*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 867–887.
- [83] A.R. Champneys and B. Sandstede, *Numerical computation of coherent structures*, Numerical Continuation Methods for Dynamical Systems, B. Krauskopf, H.M. Osinga and J. Galan-Vioque, eds, Springer (2007), 331–358.
- [84] A.R. Champneys and J.F. Toland, *Bifurcation of a plethora of multi-modal homoclinic orbits for autonomous Hamiltonian systems*, Nonlinearity **6** (1993), 665–721.
- [85] S.J. Chapman and G. Kozyreff, *Exponential asymptotics of localized patterns and snaking bifurcation diagrams*, Physica D **238** (2008), 319–354.
- [86] K.-T. Chen, *Equivalence and decomposition of vector fields about an elementary critical point*, Amer. J. Math. **85** (1963), 693–722.
- [87] L. Cherkas, *Structure of the sequence function in the neighborhood of a separatrix cycle during perturbation of an analytic autonomous system on the plane*, Differ. Equ. **17** (1981), 321–328.
- [88] D. Chillingworth and R. Lauterbach, *Dynamics and geometry in forced symmetry breaking: a tetrahedral example*, Math. Proc. Cambridge Philos. Soc. **137** (2004), 411–432.
- [89] P. Chossat, *The reduction of equivariant dynamics to the orbit space for compact group actions*, Acta Appl. Math. **70** (2002), 71–94.
- [90] P. Chossat and D. Armbruster, *Dynamics of polar reversals in spherical dynamos*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **459** (2003), 577–596.
- [91] P. Chossat, M. Krupa, I. Melbourne and A. Scheel, *Transverse bifurcations of homoclinic cycles*, Physica D **100** (1997), 85–100.
- [92] P. Chossat, M. Krupa, I. Melbourne and A. Scheel, *Magnetic dynamos in rotating convection—a dynamical systems approach*, Dynam. Contin. Discrete Impuls. Systems **5** (1999), 327–340.
- [93] P. Chossat and R. Lauterbach, *Methods in Equivariant Bifurcations and Dynamical Systems*, World Scientific Publishing, River Edge, NJ (2000).
- [94] S.-N. Chow and B. Deng, *Bifurcation of a unique stable periodic orbit from a homoclinic orbit in infinite-dimensional systems*, Trans. Amer. Math. Soc. **312** (1989), 539–587.
- [95] S.-N. Chow, B. Deng and B. Fiedler, *Homoclinic bifurcation at resonant eigenvalues*, J. Dynam. Differential Equations **2** (1990), 177–244.

- [96] S.-N. Chow, B. Deng and D. Terman, *The bifurcation of homoclinic and periodic orbits from two heteroclinic orbits*, SIAM J. Math. Anal. **21** (1990), 179–204.
- [97] S.-N. Chow, B. Deng and D. Terman, *The bifurcation of homoclinic orbits from two heteroclinic orbits - a topological approach*, Appl. Anal. **42** (1991), 275–299.
- [98] S.-N. Chow and X.-B. Lin, *Bifurcation of a homoclinic orbit with a saddle-node equilibrium*, Differential Integral Equations **3** (1990), 435–466.
- [99] E. Colli, *Infinitely many coexisting strange attractors*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), 539–579.
- [100] M.J. Costa, *Chaotic behaviour of one-dimensional saddle-node horseshoes*, Discrete Contin. Dyn. Syst. **9** (2003), 505–548.
- [101] P. Coullet, C. Riera and C. Tresser, *Stable static localized structures in one dimension*, Phys. Rev. Lett. **84** (2000), 3069–3072.
- [102] S. Crovisier, *Saddle-node bifurcations for hyperbolic sets*, Ergodic Theory Dynam. Systems **22** (2002), 1079–1115.
- [103] M. Dellnitz, M.J. Field, M. Golubitsky, J. Ma and A. Hohmann, *Cycling chaos*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **5** (1995), 1243–1247.
- [104] A. Delshams and T.M. Seara, *On the existence of homoclinic orbits in a family of vector fields near an equilibrium point*, European Conference on Iteration Theory (Calde de Malavella, 1987), World Sci. Publishing, Teaneck, NJ (1989), 167–177.
- [105] B. Deng, *The Shil'nikov problem, exponential expansion, strong  $\lambda$ -lemma,  $C^1$ -linearization, and homoclinic bifurcation*, J. Differential Equations **79** (1989), 189–231.
- [106] B. Deng, *Homoclinic bifurcations with nonhyperbolic equilibria*, SIAM J. Math. Anal. **21** (1990), 693–720.
- [107] B. Deng, *The bifurcations of countable connections from a twisted heteroclinic loop*, SIAM J. Math. Anal. **22** (1991), 653–679.
- [108] B. Deng, *Homoclinic twisting bifurcations and cusp horseshoe maps*, J. Dynam. Differential Equations **5** (1993), 417–467.
- [109] B. Deng, *On Shil'nikov's homoclinic-saddle-focus theorem*, J. Differential Equations **102** (1993), 305–329.
- [110] B. Deng, *Exponential expansion with principal eigenvalues. Nonlinear dynamics, bifurcations and chaotic behavior*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 1161–1167.
- [111] B. Deng and K. Sakamoto, *Shil'nikov–Hopf bifurcations*, J. Differential Equations **119** (1995), 1–23.
- [112] R.L. Devaney, *Homoclinic orbits in Hamiltonian systems*, J. Differential Equations **21** (1976), 431–438.
- [113] R.L. Devaney, *Blue sky catastrophes in reversible and Hamiltonian systems*, Indiana Univ. Math. J. **26** (1977), 247–263.
- [114] A.P.S. Dias, B. Dionne and I. Stewart, *Heteroclinic cycles and wreath product symmetries*, Dyn. Stab. Syst. **15** (2000), 353–385.
- [115] K. Díaz-Ordaz, M.P. Holland and S. Luzzatto, *Statistical properties of one-dimensional maps with critical points and singularities*, Stoch. Dyn. **6** (2006), 423–458.
- [116] L.J. Díaz and R. Ures, *Critical saddle-node cycles: Hausdorff dimension and persistence of tangencies*, Ergodic Theory Dynam. Systems **22** (2002), 1117–1140.
- [117] E.J. Doedel, B. Krauskopf and H.M. Osinga, *Global bifurcations of the Lorenz manifold*, Nonlinearity **19** (2006), 2947–2972.
- [118] E. Doedel, R.C. Paffenroth, A.R. Champneys, T.F. Fairgrieve, Y.A. Kuznetsov, B.E. Oldeman, B. Sandstede and X. Wang, *AUTO2000: Continuation and bifurcation software for ordinary differential equations (with HOMCONT)*, Technical report, Concordia University (2002).
- [119] R. Driesse and A.J. Homburg, *Essentially asymptotically stable homoclinic networks*, Dynamical Systems **24** (2009), 459–471.
- [120] R. Driesse and A.J. Homburg, *Resonance bifurcation from homoclinic cycles*, J. Differential Equations **246** (2009), 2681–2705.
- [121] F. Drubi, S. Ibáñez and J.A. Rodríguez, *Coupling leads to chaos*, J. Differential Equations **239** (2007), 371–385.
- [122] E. Dufraine, *Some topological invariants for three-dimensional flows*, Chaos **11** (2001), 443–448.
- [123] F. Dumortier, S. Ibáñez and H. Kokubu, *Cocoon bifurcation in three-dimensional reversible vector fields*, Nonlinearity **19** (2006), 305–328.

- [124] F. Dumortier, H. Kokubu and H. Oka, *A degenerate singularity generating geometric Lorenz attractors*, Ergodic Theory Dynam. Systems **15** (1995), 833–856.
- [125] F. Dumortier and R. Roussarie, *On the saddle-loop bifurcation*, Bifurcations of Planar Vector Fields, LNM, Vol. 1455, Springer-Verlag (1990).
- [126] F. Dumortier and R. Roussarie, *Geometric singular perturbation theory beyond normal hyperbolicity*, Multiple-time-scale Dynamical Systems (Minneapolis, MN, 1997), Springer, New York (2001), 29–63.
- [127] F. Dumortier, R. Roussarie and J. Sotomayor, *Generic 3-parameter families of vector fields on the plane, unfolding a singularity with nilpotent linear part. The cusp case of codimension 3*, Ergodic Theory Dynam. Systems **7** (1987), 375–413.
- [128] F. Dumortier, R. Roussarie and J. Sotomayor, *Bifurcations of cuspidal loops*, Nonlinearity **10** (1997), 1369–1408.
- [129] F. Dumortier, R. Roussarie, J. Sotomayor and H. Zolpolhkadek, *Bifurcations of planar vector fields*, Lecture Notes in Mathematics, Vol. 1480, Springer-Verlag, Berlin (1991).
- [130] M. El Morsalani and A. Mourtada, *Degenerate and non-trivial hyperbolic 2-polycycles: appearance of two independant Écalle-Roussarie compensators and Khovanskii's theory*, Nonlinearity **7** (1994), 1593–1604.
- [131] K. Engelborghs, T. Luzyanina and G. Samaey, DDE-BIFTOOL: A matlab package for bifurcation analysis of delay differential equations, Technical report, University of Leuven (2001).
- [132] J.W. Evans, N. Fenichel and J.A. Feroe, *Double impulse solutions in nerve axon equations*, SIAM J. Appl. Math. **42** (1982), 219–234.
- [133] J.A. Feroe, *Existence of traveling wave trains in nerve axon equations*, SIAM J. Appl. Math. **46** (1986), 1079–1097.
- [134] J.A. Feroe, *Homoclinic orbits in a parametrized saddle-focus system*, Physica D **62** (1993), 254–262. Homoclinic chaos (Brussels, 1991).
- [135] B. Fiedler, *Global pathfollowing of homoclinic orbits in two-parameter flows*, Pitman Res. Notes Math. **352** (1996), 79–146.
- [136] B. Fiedler and D.V. Turaev, *Coalescence of reversible homoclinic orbits causes elliptic resonance*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 1007–1027.
- [137] M.J. Field, *Equivariant bifurcation theory and symmetry breaking*, J. Dynam. Differential Equations **1** (1989), 369–421.
- [138] M.J. Field, *Lectures on bifurcations, dynamics and symmetry*, Pitman Research Notes in Mathematics Series, Vol. 356, Longman, Harlow (1996).
- [139] M.J. Field, *Dynamics and symmetry*, ICP Advanced Texts in Mathematics, Vol. 3, Imperial College Press, London (2007).
- [140] M.J. Field and J.W. Swift, *Stationary bifurcation to limit cycles and heteroclinic cycles*, Nonlinearity **4** (1991), 1001–1043.
- [141] A.C. Fowler and C. Sparrow, *Bifocal homoclinic orbits in four dimensions*, Nonlinearity **4** (1991), 1159–1182.
- [142] J.-M. Gambaudo, P. Glendinning and C. Tresser, *The gluing bifurcation. I. Symbolic dynamics of the closed curves*, Nonlinearity **1** (1988), 203–214.
- [143] J.-M. Gambaudo and C. Tresser, *On the dynamics of quasi-contractions*, Bol. Soc. Brasil. Mat. **19** (1988), 61–114.
- [144] P. Gaspard, *Generation of a countable set of homoclinic flows through bifurcation*, Phys. Lett. A **97** (1983), 1–4.
- [145] P. Gaspard, *Generation of a countable set of homoclinic flow through bifurcation in multidimensional systems*, Acad. Roy. Belg. Bull. Cl. Sci. (5) **70** (1984), 61–83.
- [146] P. Gaspard, *Local birth of homoclinic chaos*, Physica D **62** (1993), 94–122.
- [147] P. Gaspard, R. Kapral and G. Nicolis, *Bifurcation phenomena near homoclinic systems: a two-parameter analysis*, J. Statist. Phys. **35** (1984), 697–727.
- [148] A. Gaunersdorfer, *Time averages for heteroclinic attractors*, SIAM J. Appl. Math. **52** (1992), 1476–1489.
- [149] N.K. Gavrilov and A.L. Shil'nikov, *Example of a blue sky catastrophe*, Methods of Qualitative Theory of Differential Equations and Related Topics, Amer. Math. Soc. Transl. Ser. 2, Vol. 200, Amer. Math. Soc, Providence, RI (2000), 99–105.
- [150] M. Georgi, A. Jukes, J.S.W. Lamb and K. Webster, *Homoclinic period blow-up and the detection of multipulse solutions in lattice differential equations*, preprint (2008).

- [151] R.W. Ghrist, *Resonant gluing bifurcations*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **10** (2000), 2141–2160.
- [152] R.W. Ghrist, P.J. Holmes and M.C. Sullivan, *Knots and links in three-dimensional flows*, Lecture Notes in Mathematics, Vol. 1654, Springer-Verlag, Berlin (1997).
- [153] R.W. Ghrist and T. Young, *From Morse–Smale to all knots and links*, Nonlinearity **11** (1998), 1111–1125.
- [154] P. Glendinning, *Subsidiary bifurcations near bifocal homoclinic orbits*, Math. Proc. Cambridge Philos. Soc. **105** (1989), 597–605.
- [155] P. Glendinning and C. Sparrow, *Local and global behavior near homoclinic orbits*, J. Statist. Phys. **35** (1984), 645–696.
- [156] P. Glendinning and C. Sparrow, *T points: a codimension two heteroclinic bifurcation*, J. Statist. Phys. **43** (1986), 479–488.
- [157] A. Golmakani and A.J. Homburg, *Lorenz attractors in unfoldings of homoclinic flip bifurcations*, Dynamical Systems (in press).
- [158] M. Golubitsky, I. Stewart and D.G. Schaeffer, *Singularities and Groups in Bifurcation Theory II*, Springer-Verlag, New York (1988).
- [159] S.V. Gonchenko, L.P. Shil’nikov, O.V. Sten’kin and D.V. Turaev, *Bifurcations of systems with structurally unstable homoclinic orbits and moduli of  $\Omega$ -equivalence*, Comput. Math. Appl. **34** (1997), 111–142.
- [160] S.V. Gonchenko, D.V. Turaev, P. Gaspard and G. Nicolis, *Complexity in the bifurcation structure of homoclinic loops to a saddle-focus*, Nonlinearity **10** (1997), 409–423.
- [161] C. Grotta Ragazzo, *Irregular dynamics and homoclinic orbits to Hamiltonian saddle centres*, Comm. Pure Appl. Math. **50** (1997), 105–147.
- [162] C. Grotta Ragazzo, *On the stability of double homoclinic loops*, Comm. Math. Phys. **184** (1997), 251–272.
- [163] M.D. Groves and B. Sandstedt, *A plethora of three-dimensional periodic travelling gravity-capillary water waves with multipulse transverse profiles*, J. Nonlinear Sci. **14** (2004), 297–340.
- [164] T.M. Grozovskii, *Bifurcations of polycycles ‘apple’ and ‘half-apple’ in typical two-parameter families*, Differ. Equ. **32** (1996), 459–470.
- [165] J. Guckenheimer, *On a codimension two bifurcation*, Dynamical Systems and Turbulence, Warwick 1980 (Coventry, 1979/1980), Lecture Notes in Math., Vol. 898, Springer, Berlin (1981), 99–142.
- [166] J. Guckenheimer and P.J. Holmes, *Structurally stable heteroclinic cycles*, Math. Proc. Cambridge Philos. Soc. **103** (1988), 189–192.
- [167] J. Guckenheimer and R.F. Williams, *Structural stability of Lorenz attractors*, Inst. Hautes Études Sci. Publ. Math. **50** (1979), 59–72.
- [168] J. Guckenheimer and P.A. Worfolk, *Instant chaos*, Nonlinearity **5** (1992), 1211–1222.
- [169] L.-S. Guimond, *Homoclinic loop bifurcations on a Möbius band*, Nonlinearity **12** (1999), 59–78.
- [170] L.-S. Guimond and C. Rousseau, *Finite cyclicity of finite codimension nondegenerate homoclinic loops with real eigenvalues in  $\mathbb{R}^3$* , Qual. Theory Dyn. Syst. **2** (2001), 151–204.
- [171] J.K. Hale and X.-B. Lin, *Examples of transverse homoclinic orbits in delay equations*, Nonlinear Anal. **10** (1986), 693–709.
- [172] J.K. Hale and X.-B. Lin, *Heteroclinic orbits for retarded functional-differential equations*, J. Differential Equations **65** (1986), 175–202.
- [173] G. Haller, *Chaos near resonance*, Applied Mathematical Sciences, Vol. 138, Springer-Verlag, New York (1999).
- [174] J. Härterich, *Cascades of reversible homoclinic orbits to a saddle-focus equilibrium*, Physica D **112** (1998), 187–200.
- [175] J. Härterich, B. Sandstedt and A. Scheel, *Exponential dichotomies for linear non-autonomous functional differential equations of mixed type*, Indiana Univ. Math. J. **51** (2002), 1081–1109.
- [176] S.P. Hastings and W.C. Troy, *A proof that the Lorenz equations have a homoclinic orbit*, J. Differential Equations **113** (1994), 166–188.
- [177] D. Hawker and P. Ashwin, *Classification of robust heteroclinic cycles for vector fields in  $\mathbb{R}^3$  with symmetry*, J. Phys. A **38** (2005), 8319–8335.
- [178] D. Hawker and P. Ashwin, *Robust bursting to the origin: heteroclinic cycles with maximal symmetry equilibria*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **15** (2005), 2819–2832.
- [179] S. Hayashi, *Connecting invariant manifolds and the solution of the  $C^1$  stability and  $\Omega$ -stability conjectures for flows*, Ann. Math. (2) **145** (1997), 81–137.



- [180] S. Hayashi, *Correction to: 'Connecting invariant manifolds and the solution of the  $C^1$  stability and  $\Omega$ -stability conjectures for flows'* [Ann. Math. (2) 145 (1) (1997) 81–137], Ann. Math. (2) **150** (1999), 353–356.
- [181] D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Mathematics, Vol. 840, Springer-Verlag, Berlin (1981).
- [182] P. Hirschberg and E. Knobloch, *Shil'nikov–Hopf bifurcation*, Physica D **62** (1993), 202–216.
- [183] M.W. Hirsch, C.C. Pugh and M. Shub, *Invariant manifolds*, Lecture Notes in Mathematics, Vol. 583, Springer-Verlag, Berlin (1977).
- [184] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge (1998).
- [185] P.J. Holmes, *A strange family of three-dimensional vector fields near a degenerate singularity*, J. Differential Equations **37** (1980), 382–403.
- [186] P.J. Holmes, *Symmetries, heteroclinic cycles and intermittency in fluid flow*, Turbulence in Fluid Flows, Springer, New York (1993), 49–58.
- [187] A.J. Homburg, *Global aspects of homoclinic bifurcations of vector fields*, Mem. Amer. Math. Soc. **121** (1996), viii+128.
- [188] A.J. Homburg, *Heteroclinic bifurcations of  $\Omega$ -stable vector fields on 3-manifolds*, Discrete Contin. Dynam. Systems **4** (1998), 559–580.
- [189] A.J. Homburg, *Singular heteroclinic cycles*, J. Differential Equations **161** (2000), 358–402.
- [190] A.J. Homburg, *Periodic attractors, strange attractors and hyperbolic dynamics near homoclinic orbits to saddle-focus equilibria*, Nonlinearity **15** (2002), 1029–1050.
- [191] A.J. Homburg, A. Jukes, J. Knobloch and J.S.W. Lamb, *Saddle-nodes and period-doublings of smale horseshoes: a case study near resonant homoclinic bellows*, Bull. Belg. Math. Soc. Simon Stevin **15** (2008), 833–850.
- [192] A.J. Homburg, A.C. Jukes, J. Knobloch and J.S.W. Lamb, *Bifurcation from codimension one relative homoclinic cycles*, Transactions Amer. Math. Soc. (in press).
- [193] A.J. Homburg and J. Knobloch, *Multiple homoclinic orbits in conservative and reversible systems*, Trans. Amer. Math. Soc. **358** (2006), 1715–1740.
- [194] A.J. Homburg, H. Kokubu and M. Krupa, *The cusp horseshoe and its bifurcations in the unfolding of an inclination-flip homoclinic orbit*, Ergodic Theory Dynam. Systems **14** (1994), 667–693.
- [195] A.J. Homburg, H. Kokubu and V. Naudot, *Homoclinic-doubling cascades*, Arch. Rational Mech. Anal. **160** (2001), 195–243.
- [196] A.J. Homburg and B. Krauskopf, *Resonant homoclinic flip bifurcations*, J. Dynam. Differential Equations **12** (2000), 807–850.
- [197] A.J. Homburg and J.S.W. Lamb, *Symmetric homoclinic tangles in reversible systems*, Ergodic Theory Dynam. Systems **26** (2006), 1769–1789.
- [198] A.J. Homburg and M.A. Natiello, *Accumulations of T-points in a model for solitary pulses in an excitable reaction-diffusion medium*, Physica D **201** (2005), 212–229.
- [199] A.J. Homburg and T. Young, *Universal scalings in homoclinic doubling cascades*, Comm. Math. Phys. **222** (2001), 269–292.
- [200] C. Hou and M. Golubitsky, *An example of symmetry breaking to heteroclinic cycles*, J. Differential Equations **133** (1997), 30–48.
- [201] J.H. Hubbard and C. Sparrow, *The classification of topologically expansive Lorenz maps*, Comm. Pure Appl. Math. **43** (1990), 431–443.
- [202] H.J. Hupkes and S.M. Verduyn-Lunel, *Lin's method and homoclinic bifurcations for functional differential equations of mixed type*, Indiana Univ. Math. J. **58** (2009), 2433–2487.
- [203] S. Ibáñez and J.A. Rodríguez, *Shil'nikov bifurcations in generic 4-unfoldings of a codimension-4 singularity*, J. Differential Equations **120** (1995), 411–428.
- [204] S. Ibáñez and J.A. Rodríguez, *Shil'nikov configurations in any generic unfolding of the nilpotent singularity of codimension three on  $\mathbb{R}^3$* , J. Differential Equations **208** (2005), 147–175.
- [205] Y.S. Il'yashenko, *Limit cycles of polynomial vector fields with nondegenerate singular points on the real plane*, Funct. Anal. Appl. **18** (1984), 199–209.
- [206] Y.S. Il'yashenko, *Centennial history of Hilbert's 16th problem*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), 301–354.

- [207] Y.S. Il'yashenko and W. Li, *Nonlocal bifurcations*, Mathematical Surveys and Monographs, Vol. 66, American Mathematical Society, Providence, RI (1999).
- [208] Y.S. Il'yashenko and S.Y. Yakovenko, *Finitely-smooth normal forms of local families of diffeomorphisms and vector fields*, Russ. Math. Surv. **46** (1991), 1–43.
- [209] Y.S. Il'yashenko and S.Y. Yakovenko, *Finite cyclicity of elementary polycycles in generic families*, Concerning the Hilbert 16th problem, Amer. Math. Soc., Providence, RI (1995), 21–95.
- [210] G. Iooss and E. Lombardi, *Normal forms with exponentially small remainder: application to homoclinic connections for the reversible  $0^{2+}i\omega$  resonance*, C. R. Math. Acad. Sci. Paris **339** (2004), 831–838.
- [211] G. Iooss and E. Lombardi, *Polynomial normal forms with exponentially small remainder for analytic vector fields*, J. Differential Equations **212** (2005), 1–61.
- [212] G. Iooss and M.-C. Pérouème, *Perturbed homoclinic solutions in reversible 1:1 resonance vector fields*, J. Differential Equations **102** (1993), 62–88.
- [213] C.K.R.T. Jones, T.J. Kaper and N. Kopell, *Tracking invariant manifolds up to exponentially small errors*, SIAM J. Math. Anal. **27** (1996), 558–577.
- [214] W.D. Kalies, J. Kwapisz and R.C.A.M. Vandervorst, *Homotopy classes for stable connections between Hamiltonian saddle-focus equilibria*, Comm. Math. Phys. **193** (1998), 337–371.
- [215] V.Y. Kaloshin, *The existential Hilbert 16-th problem and an estimate for cyclicity of elementary polycycles*, Invent. Math. **151** (2003), 451–512.
- [216] T. Kapitula and S. Maier-Paape, *Spatial dynamics of time periodic solutions for the Ginzburg–Landau equation*, Z. Angew. Math. Phys. **47** (1996), 265–305.
- [217] J.P. Keener, *Chaotic behavior in piecewise continuous difference equations*, Trans. Amer. Math. Soc. **261** (1980), 589–604.
- [218] P. Kent and J. Elgin, *Noose bifurcation of periodic orbits*, Nonlinearity **4** (1991), 1045–1061.
- [219] V. Kirk, *Breaking of symmetry in the saddle-node Hopf bifurcation*, Phys. Lett. A **154** (1991), 243–248.
- [220] V. Kirk, *Merging of resonance tongues*, Physica D **66** (1993), 267–281.
- [221] V. Kirk and M. Silber, *A competition between heteroclinic cycles*, Nonlinearity **7** (1994), 1605–1621.
- [222] M. Kisaka, H. Kokubu and H. Oka, *Bifurcations to N-homoclinic orbits and N-periodic orbits in vector fields*, J. Dynam. Differential Equations **5** (1993), 305–357.
- [223] M. Kisaka, H. Kokubu and H. Oka, *Supplement to homoclinic doubling bifurcation in vector fields*, Dynamical Systems (Santiago, 1990), Longman Sci. Tech., Harlow (1993), 92–116.
- [224] J. Klaus and J. Knobloch, *Bifurcation of homoclinic orbits to a saddle-centre in reversible systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **13** (2003), 2603–2622.
- [225] J. Knobloch, *Bifurcation of degenerate homoclinics in reversible and conservative systems*, J. Dynam. Differential Equations **9** (1997), 427–444.
- [226] E. Knobloch, *Spatially localized structures in dissipative systems: open problems*, Nonlinearity **21** (2008), T45–T60.
- [227] J. Knobloch and T. Rieß, *Lin's method for heteroclinic chains involving periodic orbits*, Nonlinearity **23** (2010), 23–54.
- [228] J. Knobloch and T. Wagenknecht, *Homoclinic snaking near a heteroclinic cycle in reversible systems*, Physica D **206** (2005), 82–93.
- [229] J. Knobloch and T. Wagenknecht, *Snaking of multiple homoclinic orbits in reversible systems*, SIAM J. Appl. Dyn. Syst. **7** (2008), 1397–1420.
- [230] H. Kokubu, *Homoclinic and heteroclinic bifurcations of vector fields*, Japan J. Appl. Math. **5** (1988), 455–501.
- [231] H. Kokubu, M. Komuro and H. Oka, *Multiple homoclinic bifurcations from orbit flip. I. Successive homoclinic-doublings*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 833–850.
- [232] H. Kokubu and V. Naudot, *Existence of infinitely many homoclinic-doubling bifurcations from some codimension- three homoclinic orbits*, J. Dynam. Differential Equations **9** (1997), 445–462.
- [233] H. Kokubu and R. Roussarie, *Existence of a singularly degenerate heteroclinic cycle in the Lorenz system and its dynamical consequences. I*, J. Dynam. Differential Equations **16** (2004), 513–557.
- [234] O.Y. Koltsova, *Families of multi-round homoclinic and periodic orbits near a saddle-centre equilibrium*, Regul. Chaotic Dyn. **8** (2003), 191–200.
- [235] O.Y. Koltsova and L.M. Lerman, *Families of transverse Poincaré homoclinic orbits in 2N-dimensional Hamiltonian systems close the system with a loop to a saddle-center*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 991–1006.

- [236] A. Kotova and V. Stanzo, *On few-parameter generic families of vector fields on the two-dimensional sphere*, Concerning the Hilbert 16th Problem, Amer. Math. Soc, Providence, RI (1995), 155–201.
- [237] G. Kozyreff and S.J. Chapman, *Asymptotics of large bound states of localized structures*, Phys. Rev. Lett. **97** (2006), 044502.
- [238] B. Krauskopf and B.E. Oldeman, *A planar model system for the saddle-node Hopf bifurcation with global reinjection*, Nonlinearity **17** (2004), 1119–1151.
- [239] B. Krauskopf and B.E. Oldeman, *Bifurcations of global reinjection orbits near a saddle-node hopf bifurcation*, Nonlinearity **19** (2006), 2149–2167.
- [240] B. Krauskopf and T. Rieß, *A Lin's method approach to finding and continuing heteroclinic connections involving periodic orbits*, Nonlinearity **21** (2008), 1655–1690.
- [241] M. Krupa, *Robust heteroclinic cycles*, J. Nonlinear Sci. **7** (1997), 129–176.
- [242] M. Krupa and I. Melbourne, *Asymptotic stability of heteroclinic cycles in systems with symmetry. II*, Proc. Roy. Soc. Edinburgh Sect. A **134** (2004), 1177–1197.
- [243] M. Krupa and I. Melbourne, *Asymptotic stability of heteroclinic cycles in systems with symmetry*, Ergodic Theory Dynam. Systems **15** (1995), 121–147.
- [244] M. Krupa and I. Melbourne, *Nonasymptotically stable attractors in  $O(2)$  mode interactions*, Normal Forms and Homoclinic Chaos (Waterloo, ON, 1992), Amer. Math. Soc, Providence, RI (1995), 219–232.
- [245] M. Krupa, B. Sandstede and P. Szmolyan, *Fast and slow waves in the FitzHugh–Nagumo equation*, J. Differential Equations **133** (1997), 49–97.
- [246] M. Krupa and P. Szmolyan, *Extending slow manifolds near transcritical and pitchfork singularities*, Nonlinearity **14** (2001), 1473–1491.
- [247] Y.A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York (2004).
- [248] Y.A. Kuznetsov, O. De Feo and S. Rinaldi, *Belyakov homoclinic bifurcations in a tritrophic food chain model*, SIAM J. Appl. Math. **62** (2001), 462–487.
- [249] R. Labarca, *Stability of parametrized families of vector fields*, Dynamical Systems and Bifurcation Theory (Rio de Janeiro, 1985), Longman Sci. Tech., Harlow (1987), 121–213.
- [250] R. Labarca, *Bifurcation of contracting singular cycles*, Ann. Sci. École Norm. Sup. (4) **28** (1995), 705–745.
- [251] R. Labarca and M.J. Pacifico, *Stability of singularity horseshoes*, Topology **25** (1986), 337–352.
- [252] R. Labarca and S. Plaza, *Global stability of families of vector fields*, Ergodic Theory Dynam. Systems **13** (1993), 737–766.
- [253] J.S.W. Lamb and J.A.G. Roberts, *Time-reversal symmetry in dynamical systems: a survey*, Physica D **112** (1998), 1–39.
- [254] J.S.W. Lamb, M.A. Teixeira and K.N. Webster, *Heteroclinic bifurcations near Hopf-zero bifurcation in reversible vector fields in  $\mathbb{R}^3$* , J. Differential Equations **219** (2005), 78–115.
- [255] Y.-T. Lau, *The ‘cocoon’ bifurcations in three-dimensional systems with two fixed points*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **2** (1992), 543–558.
- [256] R. Lauterbach, S. Maier-Paape and E. Reissner, *A systematic study of heteroclinic cycles in dynamical systems with broken symmetries*, Proc. Roy. Soc. Edinburgh Sect. A **126** (1996), 885–909.
- [257] R. Lauterbach and M. Roberts, *Heteroclinic cycles in dynamical systems with broken spherical symmetry*, J. Differential Equations **100** (1992), 22–48.
- [258] G.A. Leonov, *On estimates of the bifurcation values of the parameters of a Lorenz system*, Russ. Math. Surv. **43** (1988), 216–217.
- [259] E.A. Leontovich, *On the generation of limit cycles from separatrices*, Doklady Akad. Nauk SSSR (N.S.) **78** (1951), 641–644.
- [260] L.M. Lerman, *Hamiltonian systems with a separatrix loop of a saddle-center*, Methods of the Qualitative Theory of Differential Equations, Gor’kov. Gos. Univ., Gorki (1987), 89–103 (Russian).
- [261] X.-B. Lin, *Using Melnikov’s method to solve Shil’nikov’s problems*, Proc. Roy. Soc. Edinburgh **116A** (1990), 295–325.
- [262] X.-B. Lin, *Construction and asymptotic stability of structurally stable internal layer solutions*, Trans. Amer. Math. Soc. **353** (2001), 2983–3043.
- [263] E. Lombardi, *Non-persistence of homoclinic connections for perturbed integrable reversible systems*, J. Dynam. Differential Equations **11** (1999), 129–208.
- [264] E. Lombardi, *Oscillatory integrals and phenomena beyond all algebraic orders*, Lecture Notes in Mathematics, Vol. 1741, Springer-Verlag, Berlin (2000).
- [265] E.N. Lorenz, *Deterministic non-periodic flow*, J. Atmos. Sci. **20** (1963), 130–141.

- [266] V.I. Lukyanov, *Bifurcations of dynamical systems with a saddle-point-separatrix loop*, Differ. Equ. **18** (1983), 1049–1059.
- [267] S. Luzzatto, I. Melbourne and F. Paccaut, *The Lorenz attractor is mixing*, Comm. Math. Phys. **260** (2005), 393–401.
- [268] S. Luzzatto and W. Tucker, *Non-uniformly expanding dynamics in maps with singularities and criticalities*, Inst. Hautes Études Sci. Publ. Math. **89** (1999), 179–226.
- [269] S. Luzzatto and M. Viana, *Positive Lyapunov exponents for Lorenz-like families with criticalities*, Astérisque **261** (2000), 201–237.
- [270] D.V. Lyubimov, A.S. Pikovsky and M.A. Zaks, *Universal scenarios of transitions to chaos via homoclinic bifurcations*, Soviet Scientific Reviews, Section C: Mathematical Physics Reviews, Vol. 8, Harwood Academic Publishers, Chur (1989).
- [271] J. Mallet-Paret and S.M. Verduyn-Lunel, *Exponential dichotomies and Wiener–Hopf factorizations for mixed-type functional differential equations*, J. Differential Equations (in press).
- [272] V. Manukian, *Existence and stability of multi-pulses with applications to nonlinear optics*. Ph.D. thesis, Ohio State University (2005).
- [273] V. Manukian and B. Sandstede, *Multi-hump pulses in systems with reflection and phase invariance*, J. Differential Equations **247** (2009), 1866–1898.
- [274] K. Matthies, *A subshift of finite type in the Takens–Bogdanov bifurcation with  $D_3$  symmetry*, Doc. Math. **4** (1999), 463–485.
- [275] V.S. Medvedev, *On a new type of bifurcations on manifolds*, Math. USSR, Sb. **41** (1982), 403–407.
- [276] V.S. Medvedev, *On the ‘blue sky catastrophe’ bifurcation on two-dimensional manifolds*, Mat. Zametki **51** (1992), 118–125.
- [277] I. Melbourne, *Intermittency as a codimension-three phenomenon*, J. Dynam. Differential Equations **1** (1989), 347–367.
- [278] I. Melbourne, *An example of a nonasymptotically stable attractor*, Nonlinearity **4** (1991), 835–844.
- [279] D. Michelson, *Steady solutions of the Kuramoto–Sivashinsky equation*, Physica D **19** (1986), 89–111.
- [280] A. Mielke, P.J. Holmes and O. O’Reilly, *Cascades of homoclinic orbits to, and chaos near, a Hamiltonian saddle-center*, J. Dynam. Differential Equations **4** (1992), 95–126.
- [281] C.A. Morales, *Lorenz attractor through saddle-node bifurcations*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), 589–617.
- [282] C.A. Morales, *On inclination-flip homoclinic orbit associated to a saddle-node singularity*, Bol. Soc. Brasil. Mat. (N.S.) **27** (1996), 145–160.
- [283] C.A. Morales and M.J. Pacifico, *Degenerated singular cycles of inclination-flip type*, Ann. Sci. École Norm. Sup. (4) **31** (1998), 1–16.
- [284] C.A. Morales and M.J. Pacifico, *Inclination-flip homoclinic orbits arising from orbit-flip*, Nonlinearity **14** (2001), 379–393.
- [285] C.A. Morales, M.J. Pacifico and E.R. Pujals, *Strange attractors across the boundary of hyperbolic systems*, Comm. Math. Phys. **211** (2000), 527–558.
- [286] C.A. Morales, M.J. Pacifico and E.R. Pujals, *Robust transitive singular sets for 3-flows are partially hyperbolic attractors or repellers*, Ann. Math. (2) **160** (2004), 375–432.
- [287] C.A. Morales, M.J. Pacifico and B. San Martín, *Expanding Lorenz attractors through resonant double homoclinic loops*, SIAM J. Math. Anal. **36** (2005), 1836–1861.
- [288] C.A. Morales, M.J. Pacifico and B. San Martín, *Contracting Lorenz attractors through resonant double homoclinic loops*, SIAM J. Math. Anal. **38** (2006), 309–332.
- [289] C.A. Morales and E.R. Pujals, *Singular strange attractors on the boundary of Morse–Smale systems*, Ann. Sci. École Norm. Sup. (4) **30** (1997), 693–717.
- [290] A. Mourtada, *Cyclicité finie des polycycles hyperboliques de champs de vecteurs du plan: mise sous forme normale*, Bifurcations of Planar Vector Fields (Luminy, 1989), Lecture Notes in Math., Vol. 1455, Springer, Berlin (1990), 272–314.
- [291] A. Mourtada, *Cyclicité finie des polycycles hyperboliques de champs de vecteurs du plan. Algorithme de finitude*, Ann. Inst. Fourier (Grenoble) **41** (1991), 719–753.
- [292] A. Mourtada, *Degenerate and non-trivial hyperbolic polycycles with two vertices*, J. Differential Equations **113** (1994), 68–83.
- [293] E.M. Muñoz Morales, B. San Martín and J.A. Vera Valenzuela, *Nonhyperbolic persistent attractors near the Morse–Smale boundary*, Ann. Inst. H. Poincaré Anal. Non Linéaire **20** (2003), 867–888.

- [294] V. Naudot, *Strange attractor in the unfolding of an inclination flip homoclinic orbit*, Ergodic Theory Dynam. Systems **16** (1996), 1071–1086.
- [295] S. Newhouse, J. Palis and F. Takens, *Stable arcs of diffeomorphisms*, Bull. Amer. Math. Soc. **82** (1976), 499–502.
- [296] S. Newhouse, J. Palis and F. Takens, *Bifurcations and stability of families of diffeomorphisms*, Inst. Hautes Études Sci. Publ. Math. **57** (1983), 5–71.
- [297] H.K. Nguyen and A.J. Homburg, *Resonant heteroclinic cycles and Lorenz type attractors in models for skewed varicose instability*, Nonlinearity **18** (2005), 155–173.
- [298] H.K. Nguyen and A.J. Homburg, *Global bifurcations to strange attractors in a model for skewed varicose instability in thermal convection*, Physica D **211** (2005), 235–262.
- [299] S. Nii,  *$N$ -homoclinic bifurcations for homoclinic orbits changing their twisting*, J. Dynam. Differential Equations **8** (1996), 549–572.
- [300] V.P. Nozdracheva, *Bifurcations of a structurally unstable separatrix loop*, Differ. Uravn. **18** (1982), 1551–1558, 1654.
- [301] B.E. Oldeman, A.R. Champneys and B. Krauskopf, *Homoclinic branch switching: a numerical implementation of Lin's method*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **13** (2003), 2977–2999.
- [302] B.E. Oldeman, B. Krauskopf and A.R. Champneys, *Death of period-doublings: locating the homoclinic-doubling cascade*, Physica D **146** (2000), 100–120.
- [303] I.M. Ovsyannikov and L.P. Shil'nikov, *On systems with saddle-focus homoclinic curve*, Math. USSR Sbornik **58** (1987), 557–574.
- [304] I.M. Ovsyannikov and L.P. Shil'nikov, *Systems with a homoclinic curve of multidimensional saddle-focus type, and spiral chaos*, Math. USSR Sbornik **73** (1992), 415–443.
- [305] M.J. Pacifico and A. Rovella, *Unfolding contracting singular cycles*, Ann. Sci. École Norm. Sup. (4) **26** (1993), 691–700.
- [306] M.J. Pacifico, A. Rovella and M. Viana, *Infinite-modal maps with global chaotic behavior*, Ann. Math. (2) **148** (1998), 1–44.
- [307] M.J. Pacifico, A. Rovella and M. Viana, *Corrigendum: 'Infinite-modal maps with global chaotic behavior' [Ann. Math. (2) 148 (2) (1998) 441–484]*, Ann. Math. (2) **149** (1999), 705.
- [308] J. Palis, *A differentiable invariant of topological conjugacies and moduli of stability*, Dynamical Systems, Vol. III—Warsaw, Soc. Math., France, Paris, Astérisque, No. 51 (1978), 335–346.
- [309] J. Palis, *A global perspective for non-conservative dynamics*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), 485–507.
- [310] J. Palis, *Open questions leading to a global perspective in dynamics*, Nonlinearity **21** (2008), T37–T43.
- [311] J. Palis and C.C. Pugh, *Fifty problems in dynamical systems*, Dynamical Systems—Warwick 1974 (Proc. Sympos. Appl. Topology and Dynamical Systems, Univ. Warwick, Coventry, 1973/1974; presented to E.C. Zeeman on his fiftieth birthday), Lecture Notes in Math., Vol. 468, Springer, Berlin (1975), 345–353.
- [312] J. Palis and F. Takens, *Stability of parametrized families of gradient vector fields*, Ann. Math. (2) **118** (1983), 383–421.
- [313] J. Palis and F. Takens, *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations*, Cambridge University Press (1993).
- [314] M.J. Parker, M.G.M. Gomes and I. Stewart, *Forced symmetry-breaking of square lattice planforms*, J. Dynam. Differential Equations **18** (2006), 223–255.
- [315] D. Peterhof, B. Sandstede and A. Scheel, *Exponential dichotomies for solitary-wave solutions of semilinear elliptic equations on infinite cylinders*, J. Differential Equations **140** (1997), 266–308.
- [316] S. Plaza and J. Vera, *Bifurcation and stability of families of hyperbolic vector fields in dimension three*, Ann. Inst. H. Poincaré Anal. Non Linéaire **14** (1997), 119–142.
- [317] J. Porter and E. Knobloch, *Dynamics in the 1:2 spatial resonance with broken reflection symmetry*, Physica D **201** (2005), 318–344.
- [318] C.M. Postlethwaite and J.H.P. Dawes, *Regular and irregular cycling near a heteroclinic network*, Nonlinearity **18** (2005), 1477–1509.
- [319] C.M. Postlethwaite and J.H.P. Dawes, *A codimension-two resonant bifurcation from a heteroclinic cycle with complex eigenvalues*, Dyn. Syst. **21** (2006), 313–336.
- [320] M.R.E. Proctor and C.A. Jones, *The interaction of two spatially resonant patterns in thermal convection. I. Exact 1:2 resonance*, J. Fluid Mech. **188** (1988), 301–335.

- [321] A. Pumariño and J.A. Rodríguez, *Coexistence and persistence of strange attractors*, Lecture Notes in Mathematics, Vol. 1658, Springer-Verlag, Berlin (1997).
- [322] A. Pumariño and J.A. Rodríguez, *Coexistence and persistence of infinitely many strange attractors*, Ergodic Theory Dynam. Systems **21** (2001), 1511–1523.
- [323] J.D.M. Rademacher, *Homoclinic orbits near heteroclinic cycles with one equilibrium and one periodic orbit*, J. Differential Equations **218** (2005), 390–443.
- [324] J.D.M. Rademacher, *Lyapunov–Schmidt reduction for unfolding heteroclinic networks of equilibria and periodic orbits with tangencies*, J. Differential Equations **249** (2010), 305–348.
- [325] D. Rand, *The topological classification of Lorenz attractors*, Math. Proc. Cambridge Philos. Soc. **83** (1978), 451–460.
- [326] R.C. Robinson, *Differentiability of the stable foliation for the model Lorenz equations*, Dynamical Systems and Turbulence, Warwick 1980 (Coventry, 1979/1980), Lecture Notes in Math., Vol. 898, Springer, Berlin (1981), 302–315.
- [327] R.C. Robinson, *Homoclinic bifurcation to a transitive attractor of Lorenz type*, Nonlinearity **2** (1989), 495–518.
- [328] R.C. Robinson, *Homoclinic bifurcation to a transitive attractor of Lorenz type. II*, SIAM J. Math. Anal. **23** (1992), 1255–1268.
- [329] R.C. Robinson, *Nonsymmetric Lorenz attractors from a homoclinic bifurcation*, SIAM J. Math. Anal. **32** (2000), 119–141.
- [330] M.M. Romeo and C.K.R.T. Jones, *The stability of traveling calcium pulses in a pancreatic acinar cell*, Physica D **177** (2003), 242–258.
- [331] R. Roussarie, *On the number of limit cycles which appear by perturbation of separatrix loop of planar vector fields*, Bol. Soc. Brasil. Mat. **17** (1986), 67–101.
- [332] R. Roussarie, *Bifurcation of planar vector fields and Hilbert’s sixteenth problem*, Progress in Mathematics, Vol. 164, Birkhäuser Verlag, Basel (1998).
- [333] R. Roussarie and C. Rousseau, *Almost planar homoclinic loops in  $\mathbb{R}^3$* , J. Differential Equations **126** (1996), 1–47.
- [334] R. Roussarie and F. Wagnen, *A study of the Bogdanov–Takens bifurcation*, Resenhas **2** (1995), 1–25.
- [335] A. Rovella, *The dynamics of perturbations of contracting Lorenz maps*, Bol. Soc. Brasil. Mat. (N.S.) **24** (1993), 233–259.
- [336] M.R. Rychlik, *Lorenz attractors through Shil’nikov-type bifurcation. Part I*, Ergodic Theory Dynam. Systems **10** (1990), 793–821.
- [337] G. Samaey, K. Engelborghs and D. Roose, *Numerical computation of connecting orbits in delay differential equations*, Numer. Algorithms **30** (2002), 335–352.
- [338] B. Sandstede, *Verzweigungstheorie homokliner Verdopplungen*. Ph.D. thesis, University of Stuttgart (1993).
- [339] B. Sandstede, *Constructing dynamical systems having homoclinic bifurcation points of codimension two*, J. Dynam. Differential Equations **9** (1997), 269–288.
- [340] B. Sandstede, *Instability of localized buckling modes in a one-dimensional strut model*, Philos. Trans. Roy. Soc. London Ser. A **355** (1997), 2083–2097.
- [341] B. Sandstede, *Stability of  $N$ -fronts bifurcating from a twisted heteroclinic loop and an application to the FitzHugh–Nagumo equation*, SIAM J. Math. Anal. **29** (1998), 183–207.
- [342] B. Sandstede, *Center manifolds for homoclinic solutions*, J. Dynam. Differential Equations **12** (2000), 449–510.
- [343] B. Sandstede, *Multi-pulses in reversible conservative vector fields* (in preparation).
- [344] B. Sandstede, C.K.R.T. Jones and J.C. Alexander, *Existence and stability of  $N$ -pulses on optical fibres with phase-sensitive amplifiers*, Physica D **106** (1997), 167–206.
- [345] B. Sandstede and A. Scheel, *Forced symmetry breaking of homoclinic cycles*, Nonlinearity **8** (1995), 333–365.
- [346] B. Sandstede and A. Scheel, *Essential instabilities of fronts: bifurcation, and bifurcation failure*, Dyn. Syst. **16** (2001), 1–28.
- [347] B. Sandstede and A. Scheel, *Defects in oscillatory media: toward a classification*, SIAM J. Appl. Dynam. Syst. **3** (2004), 1–68.
- [348] B. San Martín, *Saddle-focus singular cycles and prevalence of hyperbolicity*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), 623–649.

- [349] B. San Martín, *Contracting singular cycles*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), 651–659.
- [350] S. Schecter, *The saddle-node separatrix-loop bifurcation*, SIAM J. Math. Anal. **18** (1987), 1142–1156.
- [351] S. Schecter, *Exchange lemmas. II. General exchange lemma*, J. Differential Equations **245** (2008), 411–441.
- [352] A. Scheel and P. Chossat, *Bifurcation d'orbites périodiques à partir d'un cycle homocline symétrique*, C. R. Acad. Sci. Paris Sér. I Math. **314** (1992), 49–54.
- [353] G.R. Sell, *Smooth linearization near a fixed point*, Amer. J. Math. **107** (1985), 1035–1091.
- [354] M.V. Shashkov, *On bifurcations of separatrix contours with two saddles*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **2** (1992), 911–915.
- [355] M.V. Shashkov, *On the bifurcations of separatrix contours on two-dimensional surfaces. I*, Selecta Math. Soviet. **11** (1992), 341–353; Translation of *Methods in qualitative theory and bifurcation theory (Russian)*, 116–129, Gor'kov. Gos. Univ., Gorki, 1989; MR1102920 (92d:34086). Selected translations.
- [356] M.V. Shashkov, *On bifurcations of separatrix contours on two-dimensional surfaces. II*, Selecta Math. **13** (1994), 175–182; Translation of *Methods in qualitative theory and bifurcation theory (Russian)*, 41–49, Nizhegorod. Gos. Univ., NizhniNovgorod, 1990; MR1224008 (94f:34081). Selected translations.
- [357] M.V. Shashkov, *Impossibility of complete bifurcations description for some classes of systems with simple dynamics*, International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999), World Sci. Publ., River Edge, NJ (2000), 68–70.
- [358] M.V. Shashkov and L.P. Shil'nikov, *The existence of a smooth invariant foliation for Lorentz-type maps*, Differ. Equ. **30** (1994), 536–544.
- [359] M.V. Shashkov and D.V. Turaev, *On the complex bifurcation set for a system with simple dynamics*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 949–968.
- [360] M.V. Shashkov and D.V. Turaev, *An existence theorem of smooth nonlocal center manifolds for systems close to a system with a homoclinic loop*, J. Nonlinear Sci. **9** (1999), 525–573.
- [361] M.V. Shashkov and D.V. Turaev, *A proof of Shilnikov's theorem for  $C^1$ -smooth dynamical systems*, Methods of Qualitative Theory of Differential Equations and Related Topics, Amer. Math. Soc., Providence, RI (2000), 149–163.
- [362] J. Shatah and C. Zeng, *Orbits homoclinic to centre manifolds of conservative PDEs*, Nonlinearity **16** (2003), 591–614.
- [363] L.P. Shil'nikov, *Some cases of generation of period motions from singular trajectories*, Mat. Sb. (N.S.) **61** (103) (1963), 443–466.
- [364] L.P. Shil'nikov, *A case of the existence of a countable number of periodic motions*, Soviet Math. Dokl. **6** (1965), 163–166.
- [365] L.P. Shil'nikov, *On the generation of a periodic motion from a trajectory which leaves and re-enters a saddle-saddle state of equilibrium*, Sov. Math., Dokl. **7** (1966), 1155–1158.
- [366] L.P. Shil'nikov, *On the generation of a periodic motion from trajectories doubly asymptotic to an equilibrium state of saddle type*, Math. USSR Sbornik **6** (1968), 427–438.
- [367] L.P. Shil'nikov, *A certain new type of bifurcation of multidimensional dynamic systems*, Dokl. Akad. Nauk SSSR **189** (1969), 59–62.
- [368] L.P. Shil'nikov, *A contribution to the problem of the structure of an extended neighborhood of a rough state of saddle-focus type*, Math. USSR Sbornik **10** (1970), 91–102.
- [369] L.P. Shil'nikov, *The bifurcation theory and quasi-hyperbolic attractors*, Uspehi Mat. Nauk **36** (1981), 240–241.
- [370] A.L. Shil'nikov, *On bifurcations of the Lorenz attractor in the Shimizu–Morioka model*, Physica D **62** (1993), 338–346.
- [371] L.P. Shil'nikov, *Bifurcations and strange attractors*, Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002), Higher Education (2002), 349–372.
- [372] A.L. Shil'nikov and G. Cymbalyuk, *Homoclinic bifurcations of periodic orbits en a route from tonic-spiking to bursting in neuron models*, Regul. Chaotic Dyn. **9** (2004), 281–297.
- [373] A.L. Shil'nikov, L.P. Shil'nikov and D.V. Turaev, *Normal forms and Lorenz attractors*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **3** (1993), 1123–1139.
- [374] A.L. Shil'nikov, L.P. Shil'nikov and D.V. Turaev, *Blue-sky catastrophe in singularly perturbed systems*, Mosc. Math. J. **5** (2005), 269–282.
- [375] L.P. Shil'nikov, A.L. Shil'nikov, D.V. Turaev and L.O. Chua, *Methods of Qualitative Theory in Nonlinear Dynamics. Part I*, World Scientific Publishing Co. (1998).

- [376] L.P. Shil'nikov, A.L. Shil'nikov, D.V. Turaev and L.O. Chua, *Methods of Qualitative Theory in Nonlinear Dynamics. Part II*, World Scientific Publishing Co. (2001).
- [377] L.P. Shil'nikov and D.V. Turaev, *Super-homoclinic orbits and multi-pulse homoclinic loops in Hamiltonian systems with discrete symmetries*, Regul. Khaoticheskaya Din. **2** (1997), 126–138 [V.I. Arnol'd (on the occasion of his 60th birthday) (Russian)].
- [378] L.P. Shil'nikov and D.V. Turaev, *A new simple bifurcation of a periodic orbit of 'blue sky catastrophe' type*, Methods of Qualitative Theory of Differential Equations and Related Topics, Amer. Math. Soc, Providence, RI (2000), 165–188.
- [379] D. Smets and J.B. van den Berg, *Homoclinic solutions for Swift–Hohenberg and suspension bridge type equations*, J. Differential Equations **184** (2002), 78–96.
- [380] J. Sneyd, A. LeBeau and D. Yule, *Traveling waves of calcium in pancreatic acinar cells: model construction and bifurcation analysis*, Physica D **145** (2000), 158–179.
- [381] N. Sottocornola, *Complete classification of homoclinic cycles in  $\mathbb{R}^4$  in the case of a symmetry group  $G \subset SO(4)$* , C. R. Math. Acad. Sci. Paris **334** (2002), 859–864.
- [382] N. Sottocornola, *Robust homoclinic cycles in  $\mathbb{R}^4$* , Nonlinearity **16** (2003), 1–24.
- [383] N. Sottocornola, *Simple homoclinic cycles in low-dimensional spaces*, J. Differential Equations **210** (2005), 135–154.
- [384] C. Sparrow, *The Lorenz Equations: Bifurcation, Chaos and Strange Attractors*, Springer-Verlag (1982).
- [385] S. Sternberg, *On the structure of local homeomorphisms of euclidean  $n$ -space. II*, Amer. J. Math. **80** (1958), 623–631.
- [386] E. Stone and P.J. Holmes, *Random perturbations of heteroclinic attractors*, SIAM J. Appl. Math. **50** (1990), 726–743.
- [387] J.H. Sun, *Homoclinic bifurcations in symmetric unfoldings of a singularity with three-fold zero eigenvalue*, Acta Math. Sin. (Engl. Ser.) **21** (2005), 65–80.
- [388] F. Takens, *Partially hyperbolic fixed points*, Topology **10** (1971), 133–147.
- [389] F. Takens, *Forced oscillations and bifurcations*, Applications of Global Analysis, I (Sympos., Utrecht State Univ., Utrecht, 1973), Math. Inst. Rijksuniv. Utrecht, Utrecht (1974), 1–59 (Comm. Math. Inst. Rijksuniv. Utrecht, No. 3–1974).
- [390] F. Takens, *Singularities of vector fields*, Inst. Hautes Études Sci. Publ. Math. **43** (1974), 47–100.
- [391] F. Takens, *Global phenomena in bifurcations of dynamical systems with simple recurrence*, Jahresber. Deutsch. Math.-Verein., **81** (1978/79) 87–96.
- [392] F. Takens, *Moduli and bifurcations; nontransversal intersections of invariant manifolds of vectorfields*, Functional Differential Equations and Bifurcation (Proc. Conf., Inst. Ciênc. Mat. São Carlos, Univ. São Paulo, São Carlos, 1979), Lecture Notes in Math., Vol. 799, Springer, Berlin (1980), 368–384.
- [393] F. Takens, *Heteroclinic attractors: time averages and moduli of topological conjugacy*, Bol. Soc. Brasil. Mat. (N.S.) **25** (1994), 107–120.
- [394] F. Takens, *Forced oscillations and bifurcations*, Global Analysis of Dynamical Systems, Inst. Phys., Bristol (2001), 1–61.
- [395] Y. Togawa, *A modulus of 3-dimensional vector fields*, Ergodic Theory Dynam. Systems **7** (1987), 295–301.
- [396] C. Tresser, *About some theorems by L.P. Shil'nikov*, Ann. Inst. Henri Poincaré Phys. Théor. **40** (1984), 441–461.
- [397] W. Tucker, *The Lorenz attractor exists*, C. R. Acad. Sci. Paris Sér. I Math. **328** (1999), 1197–1202.
- [398] W. Tucker, *A rigorous ODE solver and Smale's 14th problem*, Found. Comput. Math. **2** (2002), 53–117.
- [399] D.V. Turaev, *Bifurcations of a homoclinic 'figure-eight' of a multidimensional saddle*, Russian Math. Surveys **43** (1988), 264–265.
- [400] D.V. Turaev, *On dimension of non-local bifurcational problems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **6** (1996), 919–948.
- [401] D.V. Turaev, *Multi-pulse homoclinic loops in systems with a smooth first integral*, Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems, Springer, Berlin (2001), 691–716.
- [402] D.V. Turaev, *An example of a resonant homoclinic loop of infinite cyclicity*, Mosc. Math. J. **5** (2005), 283–293.
- [403] D.V. Turaev and L.P. Shil'nikov, *On bifurcations of a homoclinic 'figure-eight' for a saddle with negative saddle value*, Soviet. Math. Dokl. **34** (1987), 397–401.
- [404] D.V. Turaev and L.P. Shil'nikov, *On Hamiltonian systems with homoclinic saddle curves*, Soviet Math. Dokl. **39** (1989), 165–168.



- [405] D.V. Turaev and L.P. Shil'nikov, *Blue sky catastrophes*, Dokl. Akad. Nauk **342** (1995), 596–599.
- [406] D.V. Turaev and L.P. Shil'nikov, *An example of a wild strange attractor*, Mat. Sb. **189** (1998), 137–160.
- [407] J.B. van den Berg, J. Hulshof and R.C.A.M. Vandervorst, *Traveling waves for fourth order parabolic equations*, SIAM J. Math. Anal. **32** (2001), 1342–1374.
- [408] J.B. van den Berg and R.C.A.M. Vandervorst, *Stable patterns for fourth-order parabolic equations*, Duke Math. J. **115** (2002), 513–558.
- [409] J.B. van den Berg, S.A. van Gils and T.P.P. Visser, *Parameter dependence of homoclinic solutions in a single long Josephson junction*, Nonlinearity **16** (2003), 707–717.
- [410] A. Vanderbauwhede, *Centre manifolds, normal forms and elementary bifurcations*, Dynamics Reported 2, Wiley, Chichester (1989), 89–169.
- [411] A. Vanderbauwhede, *Bifurcation of degenerate homoclinics*, Results Math. **21** (1992), 211–223.
- [412] A. Vanderbauwhede and B. Fiedler, *Homoclinic period blow-up in reversible and conservative systems*, Z. Angew. Math. Phys. **43** (1992), 292–318.
- [413] S.A. van Gils, M. Krupa and V. Tchistiakov, *Homoclinic twist bifurcation in a system of two coupled oscillators*, J. Dynam. Differential Equations **12** (2000), 733–806.
- [414] S.J. van Strien, *One parameter families of vector fields, bifurcations near saddle-connections*. Ph.D. thesis, Utrecht University, 1982.
- [415] G. Vegter, *Global stability of generic two-parameter families of gradients on three-manifolds*, Dynamical Systems and Bifurcations (Groningen, 1984), Lecture Notes in Math., Vol. 1125, Springer, Berlin (1985), 107–129.
- [416] J. Vera, *Stability of quasi-transversal bifurcation of vector fields on 3-manifolds*, Nonlinearity **9** (1996), 943–972.
- [417] T. Wagenknecht, *About a homoclinic pitchfork bifurcation in reversible systems with additional  $\mathbb{Z}_2$ -symmetry*, Nonlinearity **15** (2002), 2097–2119.
- [418] T. Wagenknecht, *Two-heteroclinic orbits emerging in the reversible homoclinic pitchfork bifurcation*, Nonlinearity **18** (2005), 527–542.
- [419] T. Wagenknecht and A.R. Champneys, *When gap solitons become embedded solitons: a generic unfolding*, Physica D **177** (2003), 50–70.
- [420] S. Wiczorek and B. Krauskopf, *Bifurcations of n-homoclinic orbits in optically injected lasers*, Nonlinearity **18** (2005), 1095–1120.
- [421] D. Wilczak, *Chaos in the Kuramoto–Sivashinsky equations—a computer-assisted proof*, J. Differential Equations **194** (2003), 433–459.
- [422] D. Wilczak, *Symmetric homoclinic solutions to the periodic orbits in the Michelson system*, Topol. Methods Nonlinear Anal. **28** (2006), 155–170.
- [423] R.F. Williams, *The structure of Lorenz attractors*, Inst. Hautes Études Sci. Publ. Math. **50** (1979), 73–99.
- [424] P.D. Woods and A.R. Champneys, *Heteroclinic tangles and homoclinic snaking in the unfolding of a degenerate reversible Hamiltonian–Hopf bifurcation*, Physica D **129** (1999), 147–170.
- [425] P.A. Worfolk, *An equivariant, inclination-flip, heteroclinic bifurcation*, Nonlinearity **9** (1996), 631–647.
- [426] J. Wu, *Theory and applications of partial functional-differential equations*, Applied Mathematical Sciences, Vol. 119, Springer-Verlag, New York (1996).
- [427] K. Yagasaki and T. Wagenknecht, *Detection of symmetric homoclinic orbits to saddle-centres in reversible systems*, Physica D **214** (2006), 169–181.
- [428] E. Yanagida, *Branching of double pulse solutions from single pulse solutions in nerve axon equations*, J. Differential Equations **66** (1987), 243–262.
- [429] A.C. Yew, *Multipulses of nonlinearly coupled Schrödinger equations*, J. Differential Equations **173** (2001), 92–137.
- [430] J.A. Yorke and K.T. Alligood, *Period doubling cascades of attractors: a prerequisite for horseshoes*, Comm. Math. Phys. **101** (1985), 305–321.
- [431] E.C. Zeeman, *Bifurcation, catastrophe, and turbulence*, New Directions in Applied Mathematics (Cleveland, Ohio, 1980), Springer, New York (1982), 109–153.
- [432] M.G. Zimmermann, S.O. Firlé, M.A. Natiello, M. Hildebrand, M. Eiswirth, M. Bär, A.K. Bangia and I.G. Kevrekidis, *Pulse bifurcation and transition to spatiotemporal chaos in an excitable reaction-diffusion model*, Physica D **110** (1997), 92–104.
- [433] M.G. Zimmermann and M.A. Natiello, *Homoclinic and heteroclinic bifurcations close to a twisted heteroclinic cycle*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **8** (1998), 359–375.
- [434] M.G. Zimmermann, M.A. Natiello and H.G. Solari, *Shil'nikov-saddle-node interaction near a codimension-2 bifurcation: laser with injected signal*, Physica D **109** (1997), 293–314.

# Author Index

Roman numbers refer to pages on which the author (or his/her work) is mentioned. Italic numbers refer to reference pages. Numbers between brackets are the reference numbers. No distinction is made between first and coauthor(s).

- Aarts, J. 170, 174, 221 [1]  
 Abdullah, K. 277, 325 [8]  
 Abell, K.A. 506, 509 [8]  
 Abraham, R. 7, 10, 40 [1]; 71, 78, 81 [1]; 81 [2]  
 Abraham, R.H. 252, 273, 319, 325 [9]  
 Aeyels, D. 349, 350, 352, 376 [1]  
 Afendikov, A. 456, 509 [9]  
 Afraïmovich, V.S. 383, 408, 409, 414, 424, 509 [10]; 510 [11]; 510 [12]; 510 [13]; 510 [28]  
 Aguiar, M.A.D. 485, 510 [14]; 510 [15]  
 Ahlfors, L. 143, 163, 221 [2]; 221 [3]  
 Albouy, A. 277, 325 [8]  
 Albrecht, J. 272, 277, 325 [10]  
 Alexander, J. 72, 81 [3]  
 Alexander, J.C. 457, 521 [344]  
 Alexiewicz, A. 65, 81 [4]  
 Alligood, K.T. 498, 499, 510 [16]; 524 [430]  
 Alpern, S. 75, 81 [5]  
 Amick, C.J. 455, 510 [17]  
 Anderson, R.M. 78–80, 81 [6]  
 Andronov, A.A. 421, 438, 510 [18]  
 Anosov, D.V. 52, 75, 78, 81 [7]; 81 [8]; 81 [9]  
 Araújo, V. 408, 409, 422, 510 [19]; 510 [20]; 510 [21]; 510 [22]  
 Araujo, A. 80, 81 [10]  
 Armbruster, D. 419, 480, 510 [23]; 510 [24]; 512 [90]  
 Arneodo, A. 409, 422, 468, 510 [25]; 510 [26]; 510 [27]  
 Arnol'd, V.I. 47, 82 [11]; 82 [12]; 424, 510 [28]  
 Arnold, V. 227, 244, 245, 245 [1]; 245 [2]; 245 [3]; 245 [4]  
 Arnold, V.I. 7, 8, 10, 11, 28, 37, 40 [2]; 40 [3]; 40 [4]; 41 [7]; 41 [8]; 41 [9]; 245, 245 [5]; 252–255, 257, 259–261, 263, 265, 270, 272–280, 283, 284, 287–290, 298, 302, 309, 313, 315, 317–319, 321, 322, 324, 325 [11]; 326 [12]; 326 [13]; 326 [14]; 326 [15]; 326 [16]; 326 [17]; 326 [18]; 326 [19]; 326 [20]; 326 [22]; 326 [23]; 326 [24]; 326 [25]; 326 [26]; 376 [2]  
 Aronson, D.G. 497, 510 [29]; 510 [30]  
 Aronszajn, N. 81, 82 [13]  
 Arroyo, A. 408, 510 [31]  
 Ashwin, P. 63, 82 [14]; 481, 483, 485, 510 [32]; 510 [33]; 515 [177]; 515 [178]  
 Atela, P. 148, 221 [4]  
 Aubry, S. 241, 245 [6]; 279, 337 [297]  
 Auerbach, H. 60, 82 [15]  
 Avez, A. 41 [7]; 252, 254, 263, 265, 273, 274, 277–279, 319, 326 [22]; 376 [2]  
 Avila, A. 49, 51, 65, 82 [16]; 82 [17]; 82 [18]  
 Avron, Y. 67, 75, 82 [19]  
 Badii, R. 366, 376 [3]; 376 [4]  
 Baesens, C. 293, 295, 326 [27]  
 Baker, I.N. 155, 220, 221 [5]; 221 [6]; 221 [7]  
 Baldomá, I. 472, 510 [34]  
 Bamón, R. 412, 450, 510 [35]  
 Banach, S. 60, 82 [15]; 82 [20]  
 Bangia, A.K. 448, 524 [432]  
 Bär, M. 448, 524 [432]  
 Barge, M. 177, 221 [8]  
 Barrow-Green, J. 252, 277, 326 [28]  
 Bates, L.M. 252, 265, 273, 274, 318, 319, 321, 322, 331 [138]  
 Beardon, A. 127, 221 [9]  
 Beck, M. 450, 466, 510 [36]  
 Belitskiĭ, G.R. 395, 396, 405, 423, 510 [37]; 510 [38]  
 Beloqui, J.A. 501, 510 [39]  
 Belyakov, L.A. 427, 428, 511 [40]; 511 [41]  
 Belykh, V.N. 450, 505, 511 [42]; 511 [43]  
 Benettin, G. 277, 280, 326 [29]; 332 [172]  
 Benyamini, Y. 82 [21]  
 Bergé, P. 418, 511 [45]  
 Berger, M.S. 404, 511 [44]  
 Bergweiler, W. 209, 221 [11]  
 Bernard, P. 68, 82 [22]; 245, 245 [7]  
 Bers, L. 143, 221 [3]  
 Berti, M. 245 [8]; 506, 511 [46]  
 Berweiler, W. 208, 221 [10]  
 Bessi, U. 245, 245 [9]; 245 [10]  
 Beyn, W.-J. 509 [1]  
 Bhattacharjee, R. 191, 195, 202, 221 [12]

- Bibikov, Yu.N. 252, 254, 277, 283, 287, 295, 313, 326 [30]; 326 [31]; 326 [32]
- Bin, L. 241, 245 [11]
- Biragov, V.S. 425, 511 [47]
- Birkhoff, G.D. 274, 278, 284, 326 [33]; 326 [34]
- Blanchard, P. 82 [23]; 127, 205, 221 [13]; 255, 326 [35]
- Bochi, J. 76, 82 [24]; 82 [25]
- Bodelón, C. 163, 189, 207, 221 [14]; 221 [15]
- Bogdanov, R. 467, 511 [48]
- Bohigas, O. 78, 82 [26]
- Bolle, P. 474, 506, 511 [46]; 511 [49]
- Bolle, Ph. 245 [8]
- Bolotin, S.V. 298, 302, 308, 313, 326 [36]
- Bolsinov, A.V. 318, 320, 327 [37]; 327 [38]
- Bonatti, C. 407, 408, 410, 501, 511 [50]; 511 [51]; 511 [52]
- Bonckaert, P. 395, 396, 444, 511 [53]; 511 [54]
- Bonetto, F. 261, 278, 324, 327 [39]; 332 [184]
- Borel, E. 111, 123 [2]
- Borisyuk, A.R. 421, 511 [55]; 511 [56]
- Borwein, J.M. 55, 82 [27]
- Bost, J.-B. 254, 277, 327 [40]
- Bourgain, J. 245, 245 [12]; 279, 313, 314, 327 [41]; 327 [42]
- Bowen, R. 10, 41 [10]
- Braaksma, B.L.J. 38, 41 [15]; 252–254, 259–261, 263, 268, 270–272, 274, 277, 283–296, 300, 306, 307, 313, 319, 324, 325, 327 [43]; 327 [44]; 328 [78]
- Brachet, M. 115, 123 [8]
- Brannath, W. 487, 511 [57]
- Branner, B. 142, 221 [16]
- Bricmont, J. 261, 278, 324, 327 [45]; 327 [46]
- Brillinger, D.R. 376 [5]
- Brin, M. 361, 376 [6]
- Bröcker, Th. 11, 15, 18, 28, 29, 41 [11]
- Broer, H.W. 3, 22, 37, 38, 41 [12]; 41 [13]; 41 [14]; 41 [15]; 41 [16]; 41 [17]; 50, 82 [28]; 91, 112, 114, 117, 118, 120, 123 [3]; 123 [4]; 123 [5]; 123 [6]; 252–254, 257, 259–263, 268–278, 280, 281, 283–296, 298–301, 303, 306, 307, 309, 312–317, 319–322, 324, 325, 325 [1]; 327 [43]; 327 [44]; 327 [47]; 327 [48]; 327 [49]; 327 [50]; 327 [51]; 327 [52]; 327 [53]; 327 [54]; 327 [55]; 327 [56]; 327 [57]; 327 [58]; 327 [59]; 328 [60]; 328 [61]; 328 [62]; 328 [63]; 328 [64]; 328 [65]; 328 [66]; 328 [67]; 328 [68]; 328 [69]; 328 [70]; 328 [71]; 328 [72]; 328 [73]; 328 [74]; 328 [75]; 328 [76]; 328 [77]; 328 [78]; 328 [79]; 328 [80]; 328 [81]; 328 [82]; 329 [83]; 329 [84]; 329 [85]; 329 [86]; 329 [87]; 329 [88]; 329 [89]; 329 [90]; 329 [91]; 329 [92]; 329 [93]; 329 [94]; 329 [95]; 329 [96]; 330 [132]; 347, 376 [7]; 470–472, 511 [58]; 511 [59]
- Bruno, A.D. 256, 270, 274, 283, 284, 313, 315, 329 [97]; 329 [98]; 329 [99]; 329 [100]
- Brunovský, P. 507, 511 [60]
- Buffoni, B. 453, 474, 506, 511 [49]; 511 [61]; 511 [62]; 511 [63]
- Burke, J. 511 [64]; 511 [65]; 511 [66]
- Buzzard, G. 50, 76, 82 [29]; 82 [30]; 82 [31]
- Bykov, V.V. 408, 450, 509 [10]; 511 [43]; 511 [67]; 511 [68]; 511 [69]; 512 [70]; 512 [71]
- Cabré, X. 272, 329 [101]; 329 [102]; 329 [103]
- Carleson, L. 127, 221 [17]
- Carneiro, M.J.D. 512 [72]
- Carr, J. 13, 41 [18]
- Casdagli, M. 49, 50, 53, 61, 86 [154]; 349, 352, 372, 376 [8]; 377 [40]
- Castro, S.B.S.D. 485, 510 [14]; 510 [15]
- Ceballos, J.C. 512 [73]
- Celletti, A. 277, 329 [104]; 329 [105]; 329 [106]
- Champneys, A.R. 416, 450, 453, 455, 464–466, 472, 474, 506, 509 [1]; 511 [62]; 512 [74]; 512 [75]; 512 [76]; 512 [77]; 512 [78]; 512 [79]; 512 [80]; 512 [81]; 512 [82]; 512 [83]; 512 [84]; 513 [118]; 520 [301]; 520 [302]; 524 [419]; 524 [424]
- Chapman, S.J. 466, 474, 512 [85]; 518 [237]
- Chen, K.-T. 395, 512 [86]
- Chenciner, A. 293, 299, 329 [107]; 329 [108]; 329 [109]; 330 [110]; 330 [111]
- Cheng, C.-Q. 245, 245 [13]
- Cheng, Ch.-Q. 301, 303, 330 [112]; 330 [113]; 330 [114]; 330 [115]; 330 [116]; 343 [447]; 343 [448]
- Cherkas, L. 445, 512 [87]
- Chierchia, L. 245, 245 [10]; 277, 329 [104]; 329 [105]; 329 [106]; 330 [117]; 330 [118]; 330 [119]; 330 [120]; 330 [121]; 330 [122]; 330 [123]; 330 [124]
- Chillingworth, D. 494, 512 [88]
- Chirikov, B.V. 251, 330 [125]; 335 [239]; 343 [469]
- Chossat, P. 383, 419, 477, 489, 490, 495, 512 [89]; 512 [90]; 512 [91]; 512 [92]; 512 [93]; 522 [352]
- Chow, S.-N. 9, 28, 30–32, 41 [19]; 330 [126]; 330 [127]; 330 [128]; 404, 414, 428, 429, 438, 441, 509, 510 [11]; 512 [94]; 512 [95]; 513 [96]; 513 [97]; 513 [98]
- Christensen, J.P.R. 55, 78, 82 [32]; 82 [33]; 82 [34]
- Chua, L.O. 383, 396, 399, 416, 421, 424, 432, 435, 443, 522 [375]; 523 [376]
- Cicogna, G. 330 [129]

- Ciliberto, S. 374, 376 [16]  
 Ciocci, M.-C. 252–254, 269, 271, 272, 281, 283, 285–288, 290, 294–296, 299, 314, 324, 327 [57]; 327 [58]; 330 [130]; 330 [131]; 330 [132]  
 Colli, E. 425, 513 [99]  
 Cong, F. 253, 301, 303, 313, 314, 316, 317, 330 [133]; 330 [134]; 331 [135]; 331 [136]; 334 [231]; 334 [232]  
 Contreras, G. 68, 82 [22]; 245, 245 [7]  
 Coppens, M.-O. 376, 377 [51]  
 Cornfeld, I.P. 8, 41 [20]; 51, 75, 82 [35]  
 Costa, M.J. 410, 513 [100]  
 Couillet, P. 115, 123 [8]; 409, 422, 466, 468, 510 [25]; 510 [26]; 510 [27]; 513 [101]  
 Cremer, H. 47, 82 [36]; 255, 331 [137]  
 Crovisier, S. 410, 513 [102]  
 Crutchfield, J.P. 347, 348, 376 [10]; 377 [32]  
 Csörnyei, M. 81, 82 [37]  
 Curry, J. 209, 221 [19]  
 Curry, S. 178, 180, 221 [18]  
 Cushman, R. 117, 123 [7]  
 Cushman, R.H. 252, 253, 265, 268, 273, 274, 298, 318–322, 324, 327 [59]; 328 [60]; 331 [138]; 331 [139]; 331 [140]; 331 [141]; 331 [142]; 331 [159]; 332 [161]  
 Cutler, C.D. 366, 376 [11]  
 Cymbalyuk, G. 422, 522 [372]  
  
 Damanik, D. 67, 82 [38]  
 Dani, J.S. 267, 331 [143]  
 Dani, S.G. 267, 331 [144]  
 Dawes, J.H.P. 485, 489, 520 [318]; 520 [319]  
 De Feo, O. 428, 518 [248]  
 de Jong, H.H. 279, 298, 331 [145]  
 de la Llave, R. 245, 246 [14]; 246 [15]; 254, 272, 274, 279, 301, 303, 324, 329 [101]; 329 [102]; 329 [103]; 331 [146]; 331 [147]; 331 [148]; 331 [149]; 331 [150]; 333 [193]; 333 [198]; 335 [241]  
 De Maesschalck, P. 395, 511 [54]  
 de Melo, W. 19, 20, 42 [49]; 49, 50, 82 [18]; 86 [142]  
 de Melo, W.C. 40, 41 [21]; 272, 287, 339 [339]  
 Debreu, G. 80, 82 [39]  
 DeGoede, J. 375, 376 [15]  
 del Rio, R. 75, 82 [41]  
 Dellnitz, M. 63, 82 [40]; 485, 513 [103]  
 Delnitz, M. 63, 85 [120]  
 Delshams, A. 245, 246 [15]; 252, 275, 276, 279, 280, 287, 314, 331 [150]; 331 [151]; 331 [152]; 331 [153]; 333 [195]; 472, 513 [104]  
 Deng, B. 396, 399, 404, 413, 424, 428, 429, 437–439, 441, 509, 512 [94]; 512 [95]; 513 [96]; 513 [97]; 513 [105]; 513 [106]; 513 [107]; 513 [108]; 513 [109]; 513 [110]; 513 [111]  
 Denker, M. 364, 376 [13]  
 Devaney, R.L. 8, 41 [22]; 146, 149, 151, 152, 155, 157, 163, 166, 169, 175, 180, 181, 183, 187, 189–191, 195, 202, 207, 215, 221, 221 [12]; 221 [14]; 221 [15]; 221 [20]; 221 [21]; 221 [22]; 221 [23]; 221 [24]; 221 [25]; 221 [26]; 221 [27]; 221 [28]; 221 [29]; 222 [30]; 222 [31]; 257, 261, 325 [2]; 331 [154]; 452, 453, 513 [112]; 513 [113]  
 Deville, R.E.L. 166, 222 [32]; 222 [33]  
 Diacu, F. 252, 277, 331 [155]  
 Dias, A.P.S. 484, 513 [114]  
 Díaz, L.J. 407, 408, 410, 511 [50]; 513 [116]  
 Díaz-Ordaz, K. 410, 513 [115]  
 Diks, C. 347, 375, 376, 376 [14]; 376 [15]  
 Dinaburg, E.I. 278, 331 [156]  
 Dionne, B. 484, 513 [114]  
 Doedel, E. 506, 509 [1]; 513 [118]  
 Doedel, E.J. 504, 513 [117]  
 Dolgopyat, D. 82 [42]; 82 [43]  
 Dolženko, E.P. 80, 83 [44]  
 Douady, A. 140, 143, 147–149, 153, 183, 187, 204, 205, 222 [34]; 222 [35]; 222 [36]; 222 [37]  
 Douady, R. 243–245, 246 [16]; 246 [17]  
 Dougherty, R. 58, 79, 83 [45]; 83 [46]  
 Driesse, R. 489, 490, 513 [119]; 513 [120]  
 Drubi, F. 468, 513 [121]  
 du Plessis, A.A. 18, 41 [25]  
 Dubey, P. 79, 83 [47]  
 Dufraine, E. 424, 501, 511 [51]; 513 [122]  
 Duistermaat, J.J. 318, 320–322, 324, 331 [139]; 331 [157]; 331 [158]  
 Dullin, H.R. 318, 324, 331 [140]; 331 [159]; 342 [441]  
 Dumortier, F. 3, 41 [13]; 263, 328 [61]; 438, 467–469, 472, 502, 507, 513 [123]; 514 [124]; 514 [125]; 514 [126]; 514 [127]; 514 [128]; 514 [129]  
 Dunford, N. 55, 57, 58, 83 [48]  
 Durkin, M.B. 155, 222 [38]  
  
 Eckmann, J.-P. 83 [49]; 366, 373, 374, 376 [16]; 377 [17]  
 Efstathiou, K. 321, 324, 331 [160]; 332 [161]; 332 [162]  
 Eiswirth, M. 448, 524 [432]  
 El Morsalani, M. 445, 514 [130]  
 Elgin, J. 499, 517 [218]  
 Eliasson, L.H. 254, 261, 278, 288, 303, 313, 324, 332 [163]; 332 [164]; 332 [165]; 332 [166]; 332 [167]; 332 [168]  
 Elmer, C.E. 506, 509 [8]

- Elphick, C. 115, 123 [8]  
 Enciso, G.A. 64, 83 [50]  
 Engelborghs, K. 506, 514 [131]; 521 [337]  
 Erëmenko, A.È. 257, 332 [169]  
 Eremenko, A. 139, 155, 222 [39]  
 Evans, J.W. 423, 514 [132]
- Fabbri, R. 278, 332 [170]  
 Fabes, E. 4, 41 [23]  
 Fagella, N. 175, 202, 208, 209, 221 [23]; 222 [40]; 222 [41]; 222 [42]  
 Fairgrieve, T.F. 506, 513 [118]  
 Falcolini, C. 330 [119]; 330 [120]; 330 [121]  
 Falconer, K.J. 354, 377 [18]  
 Falkoner, K. 354, 355, 377 [19]  
 Farey, J. 153, 222 [43]  
 Farmer, J.D. 347, 348, 376 [10]; 377 [32]  
 Fassò, F. 253, 268, 280, 319–322, 324, 327 [59]; 328 [60]; 332 [171]; 332 [172]  
 Féjóz, J. 246 [18]; 277, 312, 313, 315, 316, 332 [173]  
 Feldman, J. 261, 278, 332 [174]  
 Fenichel, N. 263, 270, 291, 298, 302, 308, 317, 332 [175]; 423, 514 [132]  
 Feroe, J.A. 423, 424, 514 [132]; 514 [133]; 514 [134]  
 Fiedler, B. 296, 332 [176]; 404, 428, 429, 452, 461, 498, 512 [95]; 514 [135]; 514 [136]; 524 [412]  
 Field, M. 9, 41 [24]; 63, 68, 83 [51]; 83 [52]  
 Field, M.J. 477, 484–486, 510 [32]; 513 [103]; 514 [137]; 514 [138]; 514 [139]; 514 [140]  
 Firle, S.O. 448, 524 [432]  
 Fomenko, A.T. 318, 320, 327 [37]; 327 [38]  
 Fomin, S.V. 8, 41 [20]; 51, 75, 82 [35]  
 Fontich, E. 252, 272, 280, 329 [101]; 329 [102]; 329 [103]; 333 [195]  
 Fornaess, J. 76, 83 [53]  
 Forni, G. 51, 82 [16]  
 Fowler, A.C. 426, 514 [141]  
 Franks, J. 498, 509 [2]  
 Frayssé, A. 61, 63, 83 [54]; 83 [55]; 83 [56]  
 Froeschlé, C. 301, 332 [177]; 332 [178]; 336 [269]  
 Fröhlich, J. 278, 332 [179]
- Gaeta, G. 330 [129]  
 Galgani, L. 252, 277, 280, 326 [29]; 333 [195]  
 Galin, D.M. 287, 332 [180]  
 Galka, A. 347, 377 [20]  
 Gallavotti, G. 261, 278, 303, 324, 327 [39]; 330 [122]; 330 [123]; 332 [181]; 332 [182]; 332 [183]; 332 [184]; 332 [185]; 333 [186]; 333 [187]; 333 [190]
- Gambaudo, J.-M. 293, 295, 313, 333 [188]; 434, 514 [142]; 514 [143]  
 Gamelin, T. 127, 221 [17]  
 Garíjo, A. 209, 222 [41]  
 Garnett, L. 209, 221 [19]  
 Gaspard, P. 422, 423, 425, 471, 472, 514 [144]; 514 [145]; 514 [146]; 514 [147]; 515 [160]  
 Gaunersdorfer, A. 501, 514 [148]  
 Gavosto, E. 76, 83 [53]  
 Gavrilov, N.K. 422, 514 [149]  
 Gawędzki, K. 261, 278, 327 [46]  
 Gentile, G. 261, 269, 270, 278, 303, 324, 327 [39]; 332 [184]; 332 [185]; 333 [186]; 333 [187]; 333 [189]; 333 [190]; 333 [191]  
 Georgi, M. 508, 514 [150]  
 Ghrist, R.W. 414, 422, 429, 499, 515 [151]; 515 [152]; 515 [153]  
 Giacobbe, A. 318, 324, 331 [140]; 331 [159]  
 Giannoni, M.J. 78, 82 [26]  
 Gibson, C.G. 18, 41 [25]; 284, 287, 298, 333 [192]  
 Gidea, M. 279, 333 [193]  
 Giorgilli, A. 252, 261, 277, 280, 300, 301, 313, 326 [29]; 332 [177]; 333 [194]; 333 [195]; 333 [196]; 333 [197]; 338 [314]; 338 [315]  
 Girsanov, I.V. 52, 71, 83 [57]  
 Giuliani, A. 261, 303, 333 [186]  
 Glendinning, P. 425, 426, 434, 446, 448, 514 [142]; 515 [154]; 515 [155]; 515 [156]  
 Goldberg, L. 153, 163, 175, 183, 187, 189, 207, 221 [14]; 221 [15]; 221 [24]; 222 [34]; 222 [45]  
 Goldberg, L.R. 139, 222 [44]  
 Golmakani, A. 497, 515 [157]  
 Golubitsky, M. 9, 41 [26]; 63, 82 [40]; 85 [120]; 261, 328 [62]; 477, 485, 494, 497, 510 [29]; 513 [103]; 515 [158]; 516 [200]  
 Gomes, M.G.M. 494, 520 [314]  
 Gonchenko, S. 74, 77, 83 [58]  
 Gonchenko, S.V. 422, 425, 501, 515 [159]; 515 [160]  
 González, A. 274, 331 [148]  
 González-Enríquez, A. 272, 333 [198]  
 Gordon, I.I. 421, 438, 510 [18]  
 Gorodetski, A. 77, 83 [59]  
 Gorodnik, A. 267, 333 [199]  
 Goryunov, V.V. 253, 289, 298, 326 [23]; 326 [24]  
 Govaerts, W. 509 [1]  
 Graczyk, J. 49, 83 [60]  
 Graff, S.M. 313, 314, 333 [200]  
 Grassberger, P. 356, 366, 377 [21]; 377 [22]; 377 [23]  
 Greene, J.M. 252, 333 [201]  
 Grobman, D. 108, 123 [9]  
 Grobman, D.M. 21, 39, 41 [27]  
 Grotta Ragazzo, C. 463, 464, 515 [161]; 515 [162]

- Groves, M.D. 455, 515 [163]  
 Grozovskiĭ, T.M. 448, 515 [164]  
 Guckenheimer, J. 9, 28, 41 [28]; 49, 83 [61]; 283, 288, 290, 293, 295, 326 [27]; 333 [202]; 408, 471, 480, 491, 503, 510 [23]; 515 [165]; 515 [166]; 515 [167]; 515 [168]  
 Guimond, L.-S. 431, 515 [169]; 515 [170]  
 Gusein-Zade, S.M. 11, 28, 41 [8]  
 Gustavson, F. 274, 284, 333 [203]  
 Gutiérrez, P. 275, 276, 280, 287, 314, 331 [151]; 331 [152]; 331 [153]  
 Gutierrez, C. 27, 41 [29]  
 Gutkin, E. 50, 83 [62]  
 Guzzo, M. 280, 332 [172]  
  
 Hadamard, J. 91, 123 [10]  
 Hale, J.K. 9, 28, 30–32, 41 [19]; 263, 330 [126]; 333 [204]; 508, 515 [171]; 515 [172]  
 Haller, G. 507, 515 [173]  
 Halmos, P. 75, 83 [63]; 83 [64]  
 Halsey, T.C. 366, 377 [24]  
 Han, Y. 303, 313, 318, 333 [205]; 336 [276]  
 Hanßmann, H. 253, 254, 277, 283, 286, 287, 294, 295, 297–299, 301, 313, 314, 324, 325, 327 [57]; 327 [58]; 328 [63]; 328 [64]; 328 [65]; 328 [66]; 328 [67]; 328 [68]; 328 [69]; 328 [70]; 333 [206]; 333 [207]; 333 [208]; 333 [209]; 333 [210]; 333 [211]; 334 [212]  
 Haro, À. 254, 334 [213]  
 Härterich, J. 416, 453, 464, 508, 512 [78]; 512 [79]; 515 [174]; 515 [175]  
 Hartman, P. 21, 39, 41 [30]; 108, 123 [11]  
 Hasselblatt, B. 8, 10, 41 [34]; 83 [65]; 84 [87]; 91, 123 [1]; 261, 335 [246]  
 Hastings, S.P. 505, 515 [176]  
 Hausdorff, F. 259, 269, 334 [214]  
 Hawker, D. 481, 515 [177]; 515 [178]  
 Hayashi, S. 408, 515 [179]; 516 [180]  
 Hayes, M. 163, 189, 207, 221 [14]; 221 [15]  
 Hénon, M. 76, 83 [66]  
 Henriksen, C. 209, 222 [42]  
 Henry, D. 509, 516 [181]  
 Hentschel, H.G.E. 366, 377 [25]  
 Herman, M. 27, 41 [31]  
 Herman, M.R. 47, 83 [67]; 237, 243, 246 [19]; 252, 254, 257, 260, 261, 264, 272, 288, 289, 316, 334 [215]; 334 [216]; 334 [217]; 334 [218]; 334 [219]; 334 [220]; 334 [221]; 334 [222]; 334 [223]; 334 [224]  
 Heunis, A.J. 65, 83 [68]  
 Hildebrand, M. 448, 524 [432]  
 Hille, E. 210, 222 [46]  
 Hirsch, M. 91, 95, 99, 108, 123 [12]  
 Hirsch, M.W. 13–15, 17, 20, 31, 41 [32]; 41 [33]; 64, 83 [50]; 83 [69]; 252, 253, 257, 263, 268, 270–273, 285, 291, 298, 302, 306, 308, 317, 319, 320, 334 [225]; 334 [226]; 459, 516 [183]  
 Hirschberg, P. 439, 516 [182]  
 Hof, A. 75, 83 [70]  
 Hofbauer, J. 478, 486, 516 [184]  
 Hoffmann-Jørgensen, J. 78, 87 [175]  
 Hohmann, A. 485, 513 [103]  
 Holland, M.P. 410, 513 [115]  
 Holm, D.D. 318, 324, 331 [140]  
 Holmes, P. 9, 28, 41 [28]; 252, 277, 283, 288, 290, 331 [155]; 333 [202]  
 Holmes, P.J. 419, 422, 436, 462, 464, 480, 499, 510 [23]; 515 [152]; 515 [166]; 516 [185]; 516 [186]; 519 [280]; 523 [386]  
 Homburg, A.J. 401, 402, 404, 405, 409–412, 416–420, 422, 425, 426, 429, 432, 433, 439, 448, 459, 478, 489, 490, 493, 497, 499, 500, 503, 513 [119]; 513 [120]; 515 [157]; 516 [187]; 516 [188]; 516 [189]; 516 [190]; 516 [191]; 516 [192]; 516 [193]; 516 [194]; 516 [195]; 516 [196]; 516 [197]; 516 [198]; 516 [199]; 520 [297]; 520 [298]  
 Hoo, J. 253, 283, 285–287, 298, 306, 307, 313, 314, 324, 328 [63]; 328 [64]; 328 [71]; 334 [227]  
 Hopf, E. 253, 263, 288, 296, 334 [228]; 334 [229]  
 Horozov, E. 322, 334 [230]  
 Hou, C. 494, 516 [200]  
 Hoveijn, I. 296, 298, 328 [72]; 328 [73]  
 Hsu, S.-B. 383, 510 [12]  
 Huang, D. 296, 313, 314, 334 [233]  
 Huang, Q. 253, 312, 317, 318, 334 [231]; 334 [232]; 336 [275]  
 Hubbard, J. 140, 143, 147–149, 153, 163, 189, 204, 205, 207, 221 [14]; 221 [15]; 222 [35]; 222 [36]; 222 [37]  
 Hubbard, J.H. 254, 261, 334 [234]; 334 [235]; 503, 504, 516 [201]  
 Huitema, G.B. 37, 38, 41 [14]; 41 [15]; 252–254, 259–261, 263, 268, 270–277, 280, 281, 283–290, 292–296, 300, 301, 303, 306, 307, 309, 312–317, 319, 324, 325, 328 [74]; 328 [75]; 328 [76]; 328 [77]; 328 [78]; 334 [236]  
 Hulshof, J. 506, 524 [407]  
 Humke, P.D. 61, 83 [71]  
 Humphreys, J. 113, 117, 123 [13]  
 Humphries, A.R. 506, 509 [8]  
 Hunt, B. 74, 84 [83]; 84 [84]  
 Hunt, B.R. 45, 46, 51, 55, 60, 62, 78, 83 [72]; 83 [73]; 83 [74]; 85 [134]  
 Hupkes, H.J. 508, 516 [202]  
 Huygens, Chr. 321, 334 [237]

- Ibáñez, S. 468, 472, 513 [121]; 513 [123]; 516 [203]; 516 [204]
- Il'yashenko, Y.S. 383, 395, 421, 424, 430, 444, 449, 510 [28]; 516 [205]; 516 [206]; 517 [207]; 517 [208]; 517 [209]
- Il'yashenko, Yu.S. 261, 334 [235]
- Ilyashenko, Y. 50, 82 [31]
- Ilyashenko, Yu. 84 [75]
- Iooss, G. 115, 123 [8]; 295, 299, 330 [110]; 330 [111]; 335 [238]; 473, 476, 517 [210]; 517 [211]; 517 [212]
- Izrailev, F.M. 251, 335 [239]
- Jaffard, S. 61, 63, 83 [55]; 83 [56]; 84 [76]
- Jakobson, M.V. 49, 84 [77]
- Jarque, X. 157, 183, 189, 190, 202, 221 [23]; 221 [25]; 221 [26]; 221 [27]
- Jensen, M.H. 366, 377 [24]
- Jin, D. 313, 314, 331 [136]
- Jitomirskaya, S. 67, 75, 82 [41]; 84 [78]
- Johnson, R. 278, 332 [170]; 335 [240]
- Jones, C.A. 480, 520 [320]
- Jones, C.K.R.T. 448, 457, 507, 517 [213]; 521 [330]; 521 [344]
- Jorba, À. 253, 274, 295, 296, 298, 303, 313–315, 328 [65]; 328 [66]; 331 [148]; 335 [241]; 335 [242]; 335 [243]; 335 [244]
- Joyeux, M. 318, 324, 331 [140]
- Jukes, A. 429, 478, 493, 508, 514 [150]; 516 [191]; 516 [192]
- Junge, A. 324, 342 [441]
- Jürgens, H. 354, 377 [33]
- Kadanoff, L.P. 366, 377 [24]
- Kahane, J.-P. 61, 63, 83 [56]; 84 [79]
- Kalies, W.D. 506, 517 [214]
- Kaloshin, V. 49, 74, 77, 83 [59]; 84 [82]; 84 [83]; 84 [84]; 84 [85]; 245, 245 [12]; 246 [20]
- Kaloshin, V.Y. 444, 449, 517 [215]
- Kaloshin, V.Yu. 62, 69, 72–74, 83 [73]; 84 [80]; 84 [81]; 85 [134]
- Kantz, H. 347, 366, 375, 377 [26]; 377 [27]
- Kaper, T.J. 507, 517 [213]
- Kapitula, T. 456, 517 [216]
- Kaplan, J.L. 374, 377 [28]
- Kappeler, T. 267, 279, 335 [245]
- Kapral, R. 425, 514 [147]
- Karpinska, B. 176, 222 [47]
- Katok, A. 8, 10, 41 [34]; 261, 335 [246]; 361, 376 [6]
- Katok, A.B. 51, 75, 84 [86]; 84 [87]; 84 [88]; 84 [89]
- Kaufmann, R. 62, 63, 84 [90]; 84 [91]
- Keen, L. 139, 215, 221, 221 [28]; 222 [44]; 222 [48]
- Keener, J.P. 420, 517 [217]
- Keller, G. 364, 376 [13]
- Kelley, A. 91, 123 [14]
- Kelley, J.L. 13, 41 [35]
- Kent, P. 499, 517 [218]
- Kerckhoff, S. 50, 51, 84 [92]
- Kevrekidis, I.G. 448, 524 [432]
- Khinchin, A.Ya. 256, 335 [247]
- Kierst, S. 61, 84 [93]
- Killip, R. 67, 82 [38]
- Kim, S. 293, 295, 326 [27]
- Kirk, V. 419, 450, 472, 489, 510 [24]; 512 [80]; 512 [81]; 517 [219]; 517 [220]; 517 [221]
- Kisaka, M. 432, 517 [222]; 517 [223]
- Klaus, J. 464, 517 [224]
- Klee, V.L. 57, 84 [94]
- Klingenberg, W. 78, 84 [95]
- Knauf, A. 254, 335 [248]
- Knill, O. 75, 83 [70]
- Knobloch, E. 433, 439, 450, 466, 511 [64]; 511 [65]; 511 [66]; 512 [81]; 516 [182]; 517 [226]; 520 [317]
- Knobloch, J. 401, 404, 429, 450, 459, 461, 464–466, 478, 493, 510 [36]; 516 [191]; 516 [192]; 516 [193]; 517 [224]; 517 [225]; 517 [227]; 517 [228]; 517 [229]
- Koçak, H. 287, 335 [249]
- Kokubu, H. 410, 416, 417, 432, 441, 469, 472, 499, 500, 513 [123]; 514 [124]; 516 [194]; 516 [195]; 517 [222]; 517 [223]; 517 [230]; 517 [231]; 517 [232]; 517 [233]
- Kolář, J. 60, 81, 84 [96]
- Kolmogorov, A.N. 45, 84 [97]; 251, 253, 260, 261, 274, 275, 277, 278, 303, 335 [250]; 335 [251]
- Koltsova, O.Y. 462, 463, 517 [234]; 517 [235]
- Komuro, M. 416, 517 [231]
- Kopanskii, A.Y. 396, 510 [38]
- Kopell, N. 507, 517 [213]
- Kotova, A. 448, 449, 518 [236]
- Kotus, J. 155, 209, 221 [6]; 222 [48]; 222 [49]
- Kozlov, V.V. 245, 245 [5]; 252, 254, 263, 265, 273–277, 284, 287, 302, 309, 313, 315, 317, 319, 321, 322, 324, 326 [25]
- Kozlovski, O. 49, 84 [98]
- Kozyreff, G. 466, 474, 512 [85]; 518 [237]
- Krauskopf, B. 222 [50]; 416, 417, 440, 448, 450, 504, 506, 513 [117]; 516 [196]; 518 [238]; 518 [239]; 518 [240]; 520 [301]; 520 [302]; 524 [420]
- Kriete, H. 222 [50]
- Krikorian, R. 65, 82 [17]; 288, 335 [252]; 335 [253]



- Krupa, M. 419, 432, 433, 478, 482, 483, 485, 486, 488–490, 497, 507, 510 [29]; 510 [30]; 512 [91]; 512 [92]; 516 [194]; 518 [241]; 518 [242]; 518 [243]; 518 [244]; 518 [245]; 518 [246]; 524 [413]
- Krych, M. 169, 180, 181, 221 [29]
- Kubichka, A.A. 312, 316, 318, 335 [254]; 335 [255]
- Kuksin, S. 66, 84 [99]
- Kuksin, S.B. 53, 87 [184]; 254, 279, 324, 325 [6]; 332 [168]; 335 [256]; 335 [257]; 335 [258]
- Kupiainen, A. 261, 278, 327 [46]
- Kupka, I. 19, 41 [36]; 41 [37]
- Kuratowski, K. 177, 222 [51]
- Kuznetsov, Y.A. 9, 28, 37, 38, 41 [38]; 91, 123 [15]; 383, 428, 470, 506, 509 [1]; 512 [82]; 513 [118]; 518 [247]; 518 [248]
- Kuznetsov, Yu.A. 283, 288, 290, 293, 294, 336 [260]; 336 [261]
- Kwapisz, J. 506, 517 [214]
- Küpper, T. 301, 303, 330 [133]
- Labarca, R. 412, 413, 450, 503, 510 [35]; 512 [73]; 518 [249]; 518 [250]; 518 [251]; 518 [252]
- Labouriau, I.S. 485, 510 [14]; 510 [15]
- Laederich, S. 246 [21]
- Lamb, J.S.W. 254, 336 [262]; 429, 451, 459, 472, 478, 493, 508, 514 [150]; 516 [191]; 516 [192]; 516 [197]; 518 [253]; 518 [254]
- Landau, L.D. 263, 296, 336 [263]; 336 [264]
- Laskar, J. 277, 336 [265]; 336 [266]
- Lasota, A. 65, 84 [100]
- Lau, E. 148, 222 [52]
- Lau, Y.-T. 472, 518 [255]
- Lauterbach, R. 383, 477, 488, 489, 494, 512 [88]; 512 [93]; 518 [256]; 518 [257]
- LaVaurs, P. 148, 222 [53]
- Lazutkin, V.F. 243, 246 [22]; 252, 272, 278, 324, 336 [267]; 336 [268]
- LeBeau, A. 448, 523 [380]
- LeDaeron, P.Y. 241, 245 [6]
- Ledrappier, F. 374, 377 [29]
- Lee, E. 153, 222 [54]
- Lega, E. 301, 332 [177]; 332 [178]; 336 [269]
- Lenz, D. 67, 82 [38]
- Leonov, G.A. 505, 518 [258]
- Leontovich, E.A. 421, 428, 430, 438, 510 [18]; 518 [259]
- Lerman, L.M. 462, 463, 517 [235]; 518 [260]
- Leslie, J.A. 50, 84 [101]
- Levi, M. 238, 239, 241, 242, 245, 246 [20]; 246 [21]; 246 [23]; 246 [24]; 246 [25]; 246 [26]; 278, 324, 325 [7]; 328 [79]
- Li, C.Z. 330 [127]
- Li, W. 383, 421, 517 [207]
- Li, Weigu 84 [75]
- Li, Y. 253, 280, 301, 303, 312–314, 316–318, 330 [128]; 330 [133]; 330 [134]; 331 [135]; 331 [136]; 333 [205]; 334 [231]; 334 [232]; 336 [277]; 336 [278]; 336 [279]; 336 [280]; 336 [281]; 336 [282]
- Lifshitz, E.M. 336 [264]
- Lin, X.-B. 399, 400, 438, 507, 508, 513 [98]; 515 [171]; 515 [172]; 518 [261]; 518 [262]
- Lindenstrauss, J. 82 [21]
- Littlewood, J.E. 242, 246 [27]
- Litvak-Hinenzon, A. 252, 269, 271, 272, 281, 285, 288, 290, 294–296, 324, 330 [132]; 336 [270]; 336 [271]; 336 [272]
- Liu, B. 303, 314, 336 [273]
- Liu, B.-f. 318, 336 [276]
- Liu, W. 414, 510 [11]
- Liu, Z. 296, 313, 314, 334 [233]
- Liu, Zh. 312, 313, 318, 336 [274]; 336 [275]
- Lloyd, D.J.B. 450, 466, 510 [36]
- Locatelli, U. 277, 333 [196]
- Lochak, P. 245, 246 [28]; 246 [29]; 277, 279, 280, 300, 336 [283]; 336 [284]; 336 [285]; 337 [286]; 337 [287]; 337 [288]
- Lombardi, E. 474–476, 517 [210]; 517 [211]; 518 [263]; 518 [264]
- Long, Y. 242, 246 [30]
- Looijenga, E.J.N. 18, 41 [25]
- Lorenz, E.N. 408, 518 [265]
- Los, J.E. 295, 299, 335 [238]; 337 [289]
- Loveĭkin, Yu.V. 312, 316, 318, 337 [290]; 337 [291]; 337 [292]; 337 [293]
- Lukyanov, V.I. 438, 519 [266]
- Lunter, G.A. 298, 328 [72]; 337 [294]
- Luskin, M. 4, 41 [23]
- Luzyanina, T. 514 [131]
- Luzzatto, S. 409, 410, 513 [115]; 519 [267]; 519 [268]; 519 [269]
- Lyashko, O.V. 253, 289, 298, 326 [23]; 326 [24]
- Lynch Hruska, S. 50, 82 [31]
- Lynch, P. 318, 324, 331 [140]
- Lyubich, M. 49, 82 [18]; 85 [105]; 182, 222 [55]
- Lyubich, M.Yu. 48, 49, 84 [102]; 84 [103]; 84 [104]; 139, 155, 222 [39]; 257, 332 [169]; 337 [295]
- Lyubimov, D.V. 420, 519 [270]
- Lü, Y. 155, 221 [6]
- Ma, J. 485, 513 [103]
- MacKay, R.S. 254, 279, 293, 295, 326 [27]; 337 [296]; 337 [297]; 337 [298]; 337 [299]
- Maĭer, A.G. 421, 438, 510 [18]



- Maier-Paape, S. 456, 494, 517 [216]; 518 [256]  
Makarov, N. 75, 82 [41]  
Malgrange, B. 28, 41 [39]; 319, 337 [300]  
Mallet-Paret, J. 508, 519 [271]  
Mañé, R. 8, 39, 41 [40]; 41 [41]; 61, 68, 76, 85 [106]; 85 [107]; 85 [108]; 293, 295, 299, 337 [301]; 412, 450, 510 [35]  
Mankiewicz, P. 81, 85 [109]  
Manukian, V. 456, 457, 519 [272]; 519 [273]  
Marco, J.-P. 245, 246 [29]; 246 [31]; 279, 280, 293–295, 299, 337 [286]; 337 [302]; 337 [303]  
Marmi, S. 254, 294, 324, 332 [168]; 337 [304]  
Marsden, J.E. 7, 10, 40 [1]; 41 [42]; 252, 273, 303, 319, 325 [9]; 337 [305]  
Marstrand, J.M. 62, 85 [110]  
Mas-Colell, A. 80, 85 [111]  
Mastropietro, V. 261, 327 [39]; 333 [187]; 333 [191]  
Masur, H. 50, 51, 84 [92]  
Mather, J. 28, 41 [43]; 68, 85 [112]; 241, 245, 246 [32]; 246 [33]; 246 [34]; 246 [35]; 246 [36]; 246 [37]; 246 [38]  
Mather, J.N. 275, 303, 337 [306]; 337 [307]  
Matoušková, E. 58, 59, 85 [113]; 85 [114]  
Matthies, K. 491, 519 [274]  
Mattila, P. 62, 85 [115]  
Matveev, V.S. 301, 303, 337 [308]  
Mayer, J. 171, 176, 222 [56]; 222 [57]  
Mazur, H. 51, 85 [116]  
Mazurkiewicz, S. 60, 85 [117]  
McClure, M. 61, 85 [118]  
McCracken, M.F. 41 [42]  
McMullen, C. 127, 164, 175, 176, 222 [58]; 222 [59]  
Medvedev, V.S. 421, 519 [275]; 519 [276]  
Mehta, M. 78, 85 [119]  
Meijer, H.G.E. 293, 336 [261]  
Meiss, J.D. 254, 337 [298]  
Melbourne, I. 63, 64, 68, 82 [14]; 82 [40]; 83 [51]; 83 [52]; 85 [120]; 85 [121]; 409, 419, 482, 483, 485–490, 512 [91]; 512 [92]; 518 [242]; 518 [243]; 518 [244]; 519 [267]; 519 [277]; 519 [278]  
Melnikov, V.K. 261, 279, 337 [309]; 337 [310]  
Meyer, K.R. 25, 37, 42 [44]  
Michelson, D. 472, 519 [279]  
Mielke, A. 456, 462, 464, 509 [9]; 519 [280]  
Milnor, J. 49, 68, 85 [122]; 85 [123]; 127, 129, 222 [60]  
Milnor, J.W. 272, 277, 324, 337 [311]  
Mischaiow, K. 498, 509 [3]  
Misiurewicz, M. 48, 85 [124]; 498, 509 [2]  
Mityagin, B.S. 52, 71, 83 [57]  
Montaldi, J. 483, 510 [33]  
Monteiro, P.K. 80, 81 [10]  
Montgomery, R. 261, 277, 338 [312]  
Moors, W.B. 55, 82 [27]  
Morales, C.A. 409, 410, 415, 416, 450, 496, 519 [281]; 519 [282]; 519 [283]; 519 [284]; 519 [285]; 519 [286]; 519 [287]; 519 [288]; 519 [289]  
Morbiddelli, A. 252, 261, 280, 300, 301, 313, 332 [177]; 333 [197]; 338 [313]; 338 [314]; 338 [315]  
Moreira, C.G. 77, 78, 85 [125]; 85 [126]  
Moreno Rocha, M. 146, 149, 151, 152, 183, 221 [27]; 222 [30]  
Morris, G.R. 240, 246 [39]  
Moser, J. 100, 123 [16]; 214, 215, 222 [61]; 237–239, 246 [25]; 246 [40]; 246 [41]; 247 [42]  
Moser, J.K. 7, 42 [61]; 252, 254, 263, 264, 270, 272, 274–279, 283, 284, 286–288, 290–292, 294, 298, 299, 302, 308, 312–314, 317, 335 [240]; 338 [316]; 338 [317]; 338 [318]; 338 [319]; 338 [320]; 338 [321]; 338 [322]; 338 [323]; 338 [324]; 341 [414]  
Mourtada, A. 444, 445, 514 [130]; 519 [290]; 519 [291]; 519 [292]  
Mrozek, M. 498, 509 [3]  
Munkres, J.R. 101, 123 [17]; 252, 253, 257, 268, 271–273, 285, 306, 319, 320, 338 [325]  
Muñoz Morales, E.M. 450, 519 [293]  
Mycielski, J. 48, 79, 83 [46]; 85 [127]  
Nakamura, K. 244, 247 [43]  
Narasimhan, R. 105, 123 [18]  
Natiello, M.A. 439, 440, 448, 499, 516 [198]; 524 [432]; 524 [433]; 524 [434]  
Naudot, V. 253, 283, 285–287, 293, 295, 298, 306, 307, 313, 314, 328 [64]; 328 [71]; 328 [80]; 416, 417, 432, 433, 499, 500, 516 [195]; 517 [232]; 520 [294]  
Neĭmark, Yu.I. 37, 42 [45]  
Neishtadt, A.I. 252, 254, 263, 265, 273–277, 280, 284, 287, 300, 302, 309, 313, 315, 317, 319, 321, 322, 324, 326 [25]; 337 [287]; 337 [288]; 338 [330]; 338 [331]  
Neishtadt, A.I. 245, 245 [5]  
Nekhoroshev, N. 245, 247 [44]  
Nekhoroshev, N.N. 64, 85 [128]; 277, 279, 280, 300, 338 [326]; 338 [327]; 338 [328]; 338 [329]  
Nevanlinna, R. 209, 210, 212, 222 [62]  
Newhouse, S. 501, 520 [295]; 520 [296]  
Newhouse, S.E. 28, 32, 34, 35, 40, 42 [46]; 76, 85 [129]; 85 [130]; 252, 263, 272, 288, 294, 338 [332]; 338 [333]  
Nguyen, H.K. 409, 410, 433, 520 [297]; 520 [298]  
Nicol, M. 63, 83 [51]

- Nicolis, G. 422, 425, 514 [147]; 515 [160]  
 Niederman, L. 64, 85 [131]; 245, 247 [45]; 280, 303, 337 [288]; 338 [334]; 338 [335]; 338 [336]  
 Nii, S. 498, 520 [299]  
 Nogueira, A. 51, 85 [132]  
 Nozdracheva, V.P. 428, 520 [300]  
  
 Oka, H. 416, 432, 469, 514 [124]; 517 [222]; 517 [223]; 517 [231]  
 Oldeman, B.E. 416, 440, 450, 506, 512 [81]; 513 [118]; 518 [238]; 518 [239]; 520 [301]; 520 [302]  
 Oliffson Kamphorst, S. 374, 376 [16]  
 Ollé, M. 253, 298, 313, 324, 338 [337]  
 O'Reilly, O. 462, 464, 519 [280]  
 Orlicz, W. 65, 81 [4]; 85 [133]  
 Oseledets, V.I. 374, 377 [30]  
 Osinga, H.M. 272, 340 [388]; 504, 513 [117]  
 Ott, W. 45, 62, 85 [134]; 85 [135]; 86 [136]  
 Oversteegen, L. 170, 174, 221 [1]  
 Ovsyannikov, I.M. 397, 422, 425, 427, 520 [303]; 520 [304]  
 Oxtoby, J.C. 13, 14, 42 [47]; 46, 53, 58, 60, 75, 79, 86 [137]; 86 [138]; 86 [139]; 252, 255, 256, 259, 338 [338]  
  
 Paccaut, F. 409, 519 [267]  
 Pacha, J.R. 253, 298, 313, 324, 338 [337]  
 Pacífico, M.J. 408–410, 412, 413, 415, 422, 450, 496, 510 [20]; 510 [21]; 510 [22]; 510 [35]; 518 [251]; 519 [283]; 519 [284]; 519 [285]; 519 [286]; 519 [287]; 519 [288]; 520 [305]; 520 [306]; 520 [307]  
 Packard, N.H. 347, 348, 376 [10]; 377 [32]  
 Paffenroth, R.C. 506, 513 [118]  
 Palis, J. 19, 20, 28, 31, 32, 34, 35, 38, 40, 42 [46]; 42 [48]; 42 [49]; 42 [50]; 49, 50, 76, 77, 86 [140]; 86 [141]; 86 [142]; 86 [143]; 91, 99, 101, 104, 107, 109, 111, 123 [19]; 123 [20]; 123 [21]; 123 [22]; 123 [23]; 272, 287, 288, 338 [332]; 339 [339]; 407, 410, 421, 501, 503, 512 [72]; 520 [295]; 520 [296]; 520 [308]; 520 [309]; 520 [310]; 520 [311]; 520 [312]; 520 [313]  
 Parasyuk, I.O. 253, 276, 312, 316, 318, 335 [254]; 335 [255]; 337 [290]; 337 [291]; 337 [292]; 337 [293]; 339 [340]; 339 [341]; 339 [342]; 339 [343]; 339 [344]; 339 [345]; 339 [346]; 339 [347]; 339 [348]  
 Parker, M.J. 494, 520 [314]  
 Patera, J. 287, 339 [349]  
 Pavani, R. 278, 332 [170]  
 Peitgen, H.-O. 354, 377 [33]  
 Percival, I.C. 254, 337 [299]  
 Pérouème, M.-C. 473, 517 [212]  
  
 Perron, O. 91, 123 [24]; 123 [25]; 123 [26]  
 Pesin, Y.B. 354, 361, 377 [34]; 409, 510 [13]  
 Peterhof, D. 509, 520 [315]  
 Petersen, C. 205, 223 [63]  
 Petruska, G. 61, 83 [71]  
 Phelps, R.R. 81, 86 [144]  
 Pikovsky, A.S. 420, 519 [270]  
 Piranian, G. 175, 223 [64]  
 Plaza, S. 503, 518 [252]; 520 [316]  
 Poláčik, P. 296, 332 [176]  
 Politi, A. 366, 376 [3]; 376 [4]  
 Pomeau, Y. 418, 511 [45]  
 Popov, G. 272, 339 [350]; 339 [351]; 339 [352]; 339 [353]  
 Porter, J. 433, 520 [317]  
 Pöschel, J. 247 [46]; 252, 254, 260, 264, 267, 269, 270, 272, 274, 275, 277–281, 287, 313, 324, 335 [245]; 338 [324]; 339 [354]; 339 [355]; 339 [356]; 339 [357]  
 Postlethwaite, C.M. 485, 489, 520 [318]; 520 [319]  
 Prasad, V. 75, 81 [5]  
 Preiss, D. 80, 86 [145]  
 Priesley, M.B. 366, 377 [35]  
 Procaccia, I. 356, 366, 377 [21]; 377 [22]; 377 [23]; 377 [24]; 377 [25]  
 Proctor, M.R.E. 480, 520 [320]  
 Pugh, C. 26, 27, 42 [52]; 42 [53]; 91, 95, 99, 104, 108, 110, 123 [12]; 123 [27]  
 Pugh, C.C. 13, 20, 31, 41 [33]; 263, 270, 291, 298, 302, 308, 317, 334 [226]; 421, 459, 516 [183]; 520 [311]  
 Puig, J. 278, 328 [81]; 339 [358]  
 Pujals, E. 76, 86 [146]  
 Pujals, E.R. 408, 409, 415, 509 [4]; 519 [285]; 519 [286]; 519 [289]  
 Pumariño, A. 410, 425, 511 [52]; 521 [321]; 521 [322]  
 Pustyl'nikov, L.D. 243, 247 [47]  
  
 Qian, D. 330 [124]  
 Qiu, Q. 276, 343 [461]  
 Quinn, F. 53, 69, 86 [147]  
 Quispel, G.R.W. 254, 314, 315, 339 [359]; 339 [364]  
  
 Rabinowitz, P.H. 506, 509 [5]  
 Rademacher, J.D.M. 450, 512 [81]; 521 [323]; 521 [324]  
 Raghavan, S. 267, 331 [144]  
 Rand, D. 503, 521 [325]  
 Ratiu, T.S. 273, 303, 319, 337 [305]; 339 [360]  
 Rees, M. 48, 86 [148]  
 Reinhardt, W.P. 339 [361]

- Reissner, E. 494, 518 [256]  
 Renyi, A. 356, 377 [36]  
 Rieß, T. 450, 517 [227]; 518 [240]  
 Riedel, F. 80, 86 [149]  
 Riera, C. 466, 513 [101]  
 Rinaldi, S. 428, 518 [248]  
 Rink, B.W. 279, 319, 321, 324, 339 [362]; 339 [363]  
 Rippon, P. 155, 221 [7]  
 Robbin, J.W. 39, 42 [54]  
 Roberts, G. 163, 189, 207, 221 [14]; 221 [15]  
 Roberts, J.A.G. 254, 336 [262]; 339 [364]; 451, 518 [253]  
 Roberts, M. 488, 494, 518 [257]  
 Robinson, C. 19, 20, 42 [55]; 91, 123 [28]  
 Robinson, R.C. 22, 39, 42 [56]; 42 [57]; 402, 410, 429, 496, 521 [326]; 521 [327]; 521 [328]; 521 [329]  
 Robutel, P. 277, 336 [266]; 340 [365]  
 Rockett, A.M. 256, 340 [366]  
 Rodnianski, I. 49, 84 [85]  
 Rodríguez, J.A. 425, 468, 513 [121]; 516 [203]; 516 [204]; 521 [321]; 521 [322]  
 Rodríguez Hertz, F. 408, 510 [31]  
 Rogers, J.T. 176, 222 [57]  
 Rokhlin, V.A. 75, 86 [150]  
 Rom-Kedar, V. 295, 336 [270]; 336 [271]; 336 [272]  
 Romeo, M.M. 448, 521 [330]  
 Roose, D. 506, 521 [337]  
 Roussarie, R. 293, 295, 298, 300, 313, 328 [80]; 328 [82]; 329 [83]; 401, 410, 430, 431, 438, 444, 467, 468, 502, 507, 514 [125]; 514 [126]; 514 [127]; 514 [128]; 514 [129]; 517 [233]; 521 [331]; 521 [332]; 521 [333]; 521 [334]  
 Rousseau, C. 287, 339 [349]; 401, 431, 515 [170]; 521 [333]  
 Rovella, A. 409, 422, 450, 520 [305]; 520 [306]; 520 [307]; 521 [335]  
 Rudin, W. 358, 377 [37]  
 Rudnev, M. 303, 313, 340 [367]; 340 [368]  
 Rudolph, D. 51, 85 [132]  
 Ruelle, D. 69, 83 [49]; 86 [151]; 252, 262, 263, 272, 294, 296, 338 [333]; 340 [369]; 340 [370]; 340 [371]; 359, 366, 373, 374, 376 [16]; 377 [17]; 377 [38]; 377 [39]  
 Rychlik, M.R. 402, 469, 496, 521 [336]  
 Ryd, G. 205, 223 [63]  
 Rüssmann, H. 252, 253, 256, 264, 269, 270, 272, 274–278, 284, 289, 313, 340 [372]; 340 [373]; 340 [374]; 340 [375]; 340 [376]; 340 [377]; 340 [378]; 340 [379]; 340 [380]; 340 [381]  
 Sacker, R. 37, 42 [58]  
 Sadovskii, D.A. 318, 324, 331 [140]; 331 [141]; 332 [161]; 332 [162]  
 Sakamoto, K. 439, 513 [111]  
 Salamon, D. 236, 238, 239, 247 [48]  
 Salamon, D.A. 272, 274, 277, 280, 340 [382]; 340 [383]  
 Saleh, K. 293, 295, 298, 313, 328 [80]; 340 [384]  
 Samaey, G. 506, 514 [131]; 521 [337]  
 Sambarino, M. 76, 86 [146]; 408, 509 [4]  
 San Martín, B. 410, 443, 444, 450, 496, 519 [287]; 519 [288]; 519 [293]; 521 [348]; 522 [349]  
 San, Vũ Ngọc 324, 340 [385]  
 Sanders, J. 117, 123 [7]  
 Sandstede, B. 382, 399–401, 416, 423, 429, 433, 434, 441, 442, 450, 453, 455–458, 466, 480, 494, 495, 504–509, 509 [1]; 509 [6]; 510 [36]; 512 [79]; 512 [82]; 512 [83]; 513 [118]; 515 [163]; 515 [175]; 518 [245]; 519 [273]; 520 [315]; 521 [338]; 521 [339]; 521 [340]; 521 [341]; 521 [342]; 521 [343]; 521 [344]; 521 [345]; 521 [346]; 521 [347]  
 Sard, A. 53, 69, 86 [147]  
 Sarnak, P. 78, 86 [152]  
 Sattinger, D.H. 9, 42 [59]  
 Sauer, T. 45, 46, 49–51, 53, 55, 61, 62, 78, 83 [74]; 86 [153]; 86 [154]; 349, 352, 377 [40]  
 Saupé, D. 354, 377 [33]  
 Sauzin, D. 245, 246 [31]; 279, 280, 293–295, 299, 301, 337 [302]; 337 [303]; 340 [386]  
 Sbano, L. 273, 319, 339 [360]  
 Schaefer, H.H. 86 [155]  
 Schaeffer, D.G. 9, 41 [26]; 477, 515 [158]  
 Schecter, S. 438, 507, 522 [350]; 522 [351]  
 Scheel, A. 419, 480, 489, 490, 494, 495, 508, 509, 512 [91]; 512 [92]; 515 [175]; 520 [315]; 521 [345]; 521 [346]; 521 [347]; 522 [352]  
 Scheurle, J. 252, 314, 340 [387]  
 Schilder, F. 272, 340 [388]  
 Schleicher, D. 148, 222 [52]  
 Schlomiuk, D. 287, 339 [349]  
 Schmit, C. 78, 82 [26]  
 Schouten, J.C. 376, 377 [51]  
 Schreiber, T. 347, 366, 377 [27]  
 Schwartz, J.T. 55, 57, 58, 83 [48]  
 Seara, T. 245, 246 [15]  
 Seara, T.M. 279, 331 [150]; 472, 510 [34]; 513 [104]  
 Sell, G.R. 4, 41 [23]; 395, 522 [353]  
 Séré, E. 511 [63]  
 Sevryuk, M.B. 8, 37, 41 [9]; 41 [14]; 42 [60]; 252, 254, 259–261, 263, 267, 268, 270–277, 280, 281, 283–288, 290, 294–296, 300, 301, 303, 306, 307, 309, 312–319, 324, 325, 326 [26]; 328 [76]; 328 [77]; 339 [359]; 340 [389]; 340 [390];

- 341 [391]; 341 [392]; 341 [393]; 341 [394];  
341 [395]; 341 [396]; 341 [397]; 341 [398];  
341 [399]; 341 [400]; 341 [401]; 341 [402];  
341 [403]; 341 [404]; 341 [405]; 341 [406];  
341 [407]; 341 [408]; 341 [409]; 341 [410];  
341 [411]; 341 [412]
- Shannon, C. 80, 86 [156]; 86 [157]
- Shashkov, M.V. 401, 402, 423, 424, 436, 443,  
522 [354]; 522 [355]; 522 [356]; 522 [357];  
522 [358]; 522 [359]; 522 [360]; 522 [361]
- Shatah, J. 464, 522 [362]
- Shaw, R.S. 347, 348, 376 [10]; 377 [32]
- Shi, H. 60, 79, 86 [158]; 86 [159]
- Shil'nikov, A.L. 383, 396, 399, 416, 421, 422,  
424, 432, 435, 443, 469, 514 [149]; 522 [370];  
522 [372]; 522 [373]; 522 [374]; 522 [375];  
523 [376]
- Shil'nikov, L.P. 383, 396, 397, 399, 402, 408, 410,  
411, 414, 416, 421–427, 432, 435, 438, 443, 459,  
469, 496, 501, 509 [10]; 510 [28]; 511 [47];  
515 [159]; 520 [303]; 520 [304]; 522 [358];  
522 [363]; 522 [364]; 522 [365]; 522 [366];  
522 [367]; 522 [368]; 522 [369]; 522 [371];  
522 [373]; 522 [374]; 522 [375]; 523 [376];  
523 [377]; 523 [378]; 523 [403]; 523 [404];  
524 [405]; 524 [406]
- Shilnikov, L. 74, 77, 83 [58]
- Shraiman, B.I. 366, 377 [24]
- Shub, M. 13, 20, 31, 41 [33]; 50, 69, 86 [160];  
86 [161]; 91, 95, 99, 104, 108, 110, 123 [12];  
123 [27]; 263, 270, 291, 298, 302, 308, 317,  
334 [226]; 459, 516 [183]
- Siegel, C.L. 7, 42 [61]; 48, 86 [162]; 252, 254,  
255, 257, 264, 274–277, 284, 341 [413];  
341 [414]
- Sigmund, K. 66, 86 [163]; 86 [164]; 478, 486,  
516 [184]
- Silber, M. 489, 517 [221]
- Simó, C. 252, 261, 262, 272, 278, 280, 295, 296,  
299, 300, 328 [73]; 328 [81]; 329 [83]; 329 [84];  
329 [85]; 329 [86]; 329 [87]; 329 [94]; 333 [195];  
341 [415]; 341 [416]; 341 [417]
- Simon, B. 66, 67, 75, 82 [19]; 82 [41]; 83 [70];  
84 [78]; 86 [165]; 86 [166]
- Sinaĭ, Ya.G. 8, 41 [20]; 278, 331 [156]
- Sinai, Ya.G. 51, 75, 82 [35]
- Singer, D. 209, 223 [65]
- Smale, S. 13, 19, 42 [62]; 42 [63]; 74, 86 [167];  
102, 107, 109, 123 [20]; 123 [29]; 177, 223 [66]
- Smets, D. 506, 523 [379]
- Smillie, J. 50, 51, 84 [92]
- Smith, H.L. 64, 83 [50]; 87 [168]
- Smith, R.L. 366, 370, 372, 377 [50]
- Sneyd, J. 448, 523 [380]
- Solari, H.G. 440, 524 [434]
- Solecki, S. 58, 87 [169]
- Solomyak, B. 72, 87 [170]
- Sotomayor, J. 438, 468, 514 [127]; 514 [128];  
514 [129]
- Sottocornola, N. 479, 480, 482, 523 [381];  
523 [382]; 523 [383]
- Sousa Dias, E. 273, 319, 339 [360]
- Sparrow, C. 408, 425, 426, 446, 448, 503, 504,  
514 [141]; 515 [155]; 515 [156]; 516 [201];  
523 [384]
- Spencer, T. 278, 332 [179]
- Spiegel, E.A. 468, 510 [25]
- Spivak, M. 253, 268, 271, 273, 285, 306, 319,  
342 [418]
- Stanzo, V. 448, 449, 518 [236]
- Stark, J. 254, 337 [298]
- Stefan, P. 320, 322, 342 [419]
- Steinmetz, N. 127, 223 [67]
- Sten'kin, O.V. 501, 515 [159]
- Stepin, A.M. 51, 75, 84 [88]; 84 [89]
- Sternberg, S. 21, 42 [64]; 42 [65]; 111, 124 [30];  
124 [31]; 395, 523 [385]
- Stewart, I. 9, 41 [26]; 63, 64, 85 [121]; 324,  
342 [420]; 477, 484, 494, 513 [114]; 515 [158];  
520 [314]
- Stoker, J.J. 291, 295, 342 [421]
- Stone, E. 419, 510 [24]; 523 [386]
- Strelcyn, J.-M. 277, 326 [29]
- Sudakov, V.N. 52, 71, 87 [171]; 87 [172]
- Sullivan, D. 139, 209, 221 [19]; 223 [68]
- Sullivan, M.C. 422, 499, 515 [152]
- Sun, J.H. 468, 523 [387]
- Sun, Y.-S. 330 [114]
- Sussmann, H.J. 320, 322, 342 [422]
- Svanidze, N.V. 252, 272, 342 [423]
- Świątek, G. 47, 49, 83 [60]; 87 [173]
- Swift, J.W. 486, 514 [140]
- Szmolyan, P. 433, 507, 518 [245]; 518 [246]
- Szpilrajn, E. 61, 84 [93]
- Szűsz, P. 256, 340 [366]
- Takens, F. 3, 25, 28, 31, 32, 34, 35, 38, 40,  
41 [13]; 41 [15]; 41 [16]; 42 [46]; 42 [50];  
42 [66]; 42 [67]; 42 [68]; 61, 76–78, 84 [95];  
86 [143]; 87 [174]; 91, 101, 104, 111, 114, 117,  
120, 123 [5]; 123 [6]; 123 [21]; 123 [22];  
123 [23]; 124 [32]; 124 [33]; 124 [34]; 252–254,  
259–261, 263, 264, 268, 270–274, 277, 280, 281,  
283–290, 292–296, 300, 306, 307, 313, 319–322,  
324, 325, 325 [1]; 328 [60]; 328 [61]; 328 [78];  
329 [88]; 329 [89]; 329 [90]; 329 [91]; 329 [92];  
338 [332]; 338 [333]; 340 [370]; 340 [371];  
342 [424]; 342 [425]; 342 [426]; 342 [427]; 347,

- 349–351, 357, 359–362, 366, 375, 376 [7];  
376 [15]; 377 [41]; 377 [42]; 377 [43]; 377 [44];  
377 [45]; 377 [46]; 377 [47]; 384, 387, 389, 390,  
395, 407, 410, 467, 470, 501, 503, 509 [7];  
520 [295]; 520 [296]; 520 [312]; 520 [313];  
523 [388]; 523 [389]; 523 [390]; 523 [391];  
523 [392]; 523 [393]; 523 [394]
- Tangerman, F. 175, 215, 222 [31]
- Tangerman, F.M. 22, 41 [17]; 50, 82 [28]; 252,  
257, 329 [93]
- Tatjer, J.C. 261, 262, 272, 329 [86]
- Tchistiakov, V. 497, 524 [413]
- Teixeira, M.A. 472, 518 [254]
- Terman, D. 441, 513 [96]; 513 [97]
- Terra, G. 273, 319, 339 [360]
- Theiler, J. 366, 377 [48]
- Thieme, H.R. 64, 87 [168]
- Thom, R. 11, 13, 15, 42 [69]; 42 [70]; 253, 289,  
297, 298, 342 [428]
- Thompson, B. 79, 86 [159]
- Thurston, W. 49, 85 [123]
- Tirapegui, E. 115, 123 [8]
- Tišer, J. 80, 86 [145]
- Togawa, Y. 424, 523 [395]
- Toland, J.F. 453, 455, 510 [17]; 511 [62]; 512 [84]
- Tong, H. 366, 370, 372, 377 [50]
- Topsøe, F. 78, 87 [175]
- Török, A. 68, 83 [52]
- Treschev, D. 245, 247 [49]
- Treshchëv, D.V. 253, 254, 275, 279, 287, 298,  
301–303, 308, 313, 326 [36]; 342 [429];  
342 [430]; 342 [431]; 342 [432]
- Tresser, C. 153, 222 [45]; 409, 422, 423, 434, 466,  
468, 510 [25]; 510 [26]; 510 [27]; 513 [101];  
514 [142]; 514 [143]; 523 [396]
- Troy, W.C. 505, 515 [176]
- Trubowitz, E. 261, 278, 332 [174]
- Tsujii, M. 69, 70, 72, 73, 87 [176]; 87 [177];  
87 [178]; 87 [179]
- Tucker, W. 408, 410, 519 [268]; 523 [397];  
523 [398]
- Tudoran, R. 273, 319, 339 [360]
- Turaev, D. 74, 77, 83 [58]
- Turaev, D.V. 383, 396, 399, 401, 410, 411, 416,  
420–425, 431, 432, 435, 436, 443, 458, 459, 461,  
469, 501, 514 [136]; 515 [159]; 515 [160];  
522 [359]; 522 [360]; 522 [361]; 522 [373];  
522 [374]; 522 [375]; 523 [376]; 523 [377];  
523 [378]; 523 [399]; 523 [400]; 523 [401];  
523 [402]; 523 [403]; 523 [404]; 524 [405];  
524 [406]
- Ulam, S.M. 58, 75, 79, 86 [138]; 86 [139]
- Umemura, Y. 52, 71, 87 [180]
- Urbanski, M. 209, 222 [49]
- Ures, R. 410, 513 [116]
- Valdinoci, E. 245, 245 [10]; 303, 342 [433]
- van den Berg, J.B. 448, 506, 523 [379]; 524 [407];  
524 [408]; 524 [409]
- van den Bleek, C.M. 376, 377 [51]
- van der Meer, J.-C. 122, 124 [39]; 298, 331 [142];  
342 [435]
- van Gils, S.A. 448, 497, 510 [30]; 524 [409];  
524 [413]
- van Noort, M. 296, 328 [73]; 329 [94]; 342 [436]
- van Ommen, J.R. 376, 377 [51]
- van Strien, S. 76, 87 [181]; 100, 124 [40]
- van Strien, S.J. 3, 40, 41 [13]; 42 [71]; 263, 270,  
328 [61]; 342 [437]; 501, 502, 524 [414]
- van Veen, L. 299, 342 [438]
- Van Vleck, E.S. 506, 509 [8]
- van Zwet, W.R. 375, 376 [15]
- Vanderbauwhede, A. 91, 117, 119, 120, 122,  
124 [35]; 124 [36]; 124 [37]; 124 [38]; 124 [39];  
253, 254, 263, 270, 283, 286, 287, 291, 298, 299,  
302, 308, 314, 317, 324, 327 [58]; 342 [434];  
384, 387, 389, 390, 437, 452, 509 [7]; 524 [410];  
524 [411]; 524 [412]
- Vandervorst, R.C.A.M. 506, 517 [214]; 524 [407];  
524 [408]
- Varchenko, A.N. 11, 28, 41 [8]
- Vasil'ev, V.A. 253, 289, 298, 326 [23]; 326 [24]
- Veech, W.A. 51, 87 [182]; 87 [183]
- Vegter, G. 261, 296, 298, 328 [62]; 328 [72];  
328 [73]; 329 [95]; 471, 472, 503, 511 [59];  
524 [415]
- Vera Valenzuela, J.A. 450, 519 [293]
- Vera, J. 503, 520 [316]; 524 [416]
- Verbitski, E. 361, 366, 377 [46]; 377 [47]
- Verbitski, E.A. 366, 377 [52]
- Verduyn-Lunel, S.M. 508, 516 [202]; 519 [271]
- Viana, M. 76, 82 [25]; 175, 223 [69]; 407, 408,  
410, 422, 511 [50]; 511 [52]; 519 [269];  
520 [306]; 520 [307]
- Vidal, C. 418, 511 [45]
- Villanueva, J. 253, 261, 274, 277, 295, 296, 298,  
313–315, 324, 328 [65]; 328 [66]; 331 [148];  
335 [242]; 335 [243]; 335 [244]; 338 [337];  
342 [439]; 342 [440]
- Vishik, M.I. 53, 87 [184]
- Visser, T.P.P. 448, 524 [409]
- Vitolo, R. 299, 329 [87]
- Vogt, W. 272, 340 [388]
- Waalkens, H. 324, 342 [441]
- Wagener, F. 467, 521 [334]

- Wagener, F.O.O. 253, 272, 287–289, 293, 295, 296, 298, 313, 325, 328 [65]; 328 [66]; 328 [67]; 328 [80]; 329 [90]; 329 [96]; 340 [384]; 342 [427]; 342 [442]; 342 [443]; 342 [444]; 342 [445]; 342 [446]
- Wagenknecht, T. 450, 464–466, 510 [36]; 517 [228]; 517 [229]; 524 [417]; 524 [418]; 524 [419]; 524 [427]
- Wang, D. 330 [127]
- Wang, Sh. 301, 303, 313, 330 [115]; 343 [447]; 343 [448]; 343 [449]
- Wang, X. 506, 513 [118]
- Wayne, C.E. 254, 301, 303, 331 [149]; 343 [450]
- Webster, K. 508, 514 [150]
- Webster, K.N. 472, 518 [254]
- Wei, B. 303, 343 [451]
- Whitney, H. 61, 87 [185]; 260, 319, 343 [452]; 343 [453]; 349, 377 [54]
- Wieczorek, S. 448, 524 [420]
- Wiggins, S. 283, 288, 290, 303, 313, 340 [367]; 340 [368]; 343 [454]
- Wilczak, D. 472, 524 [421]; 524 [422]
- Wilkinson, A. 69, 83 [65]; 86 [151]; 86 [161]
- Williams, R.F. 408, 503, 515 [167]; 524 [423]
- Wirthmüller, K. 18, 41 [25]
- Woods, P.D. 466, 474, 524 [424]
- Worfolk, P.A. 484, 490, 491, 515 [168]; 524 [425]
- Wu, J. 508, 524 [426]
- Xia, J. 245, 247 [50]
- Xu, J. 272, 276, 313, 314, 343 [449]; 343 [455]; 343 [456]; 343 [457]; 343 [458]; 343 [459]; 343 [460]; 343 [461]; 344 [473]; 344 [474]; 344 [475]
- Yagasaki, K. 464, 524 [427]
- Yakovenko, S.Y. 395, 449, 517 [208]; 517 [209]
- Yan, J. 245, 245 [13]; 330 [116]
- Yanagida, E. 431, 524 [428]
- Yew, A.C. 455, 524 [429]
- Yi, Y. 280, 301, 303, 312–314, 318, 330 [128]; 333 [205]; 336 [277]; 336 [278]; 336 [279]; 336 [280]; 336 [281]; 336 [282]
- Yihe, D. 312, 318, 336 [275]
- Yoccoz, J.-C. 78, 85 [126]; 254, 256, 260, 270, 289, 312, 316, 317, 324, 332 [168]; 343 [462]; 343 [463]; 343 [464]; 343 [465]; 343 [466]
- Yorke, J. 45, 62, 72, 81 [3]; 85 [135]; 86 [136]
- Yorke, J.A. 45, 46, 49–51, 53, 55, 61, 62, 65, 78, 83 [74]; 84 [100]; 86 [153]; 86 [154]; 349, 352, 374, 377 [28]; 377 [40]; 498, 499, 510 [16]; 524 [430]
- You, J. 253, 272, 276, 298, 301, 303, 313, 328 [68]; 328 [69]; 328 [70]; 330 [133]; 343 [458]; 343 [459]; 343 [460]; 343 [461]; 343 [467]; 343 [468]
- Young, L.S. 374, 377 [29]
- Young, T. 414, 418, 515 [153]; 516 [199]
- Yule, D. 448, 523 [380]
- Zajíček, L. 60, 80, 87 [186]; 87 [187]; 87 [188]
- Zaks, M.A. 420, 519 [270]
- Zame, W.R. 78–80, 81 [6]; 86 [157]
- Zaslavskii, G.M. 251, 343 [469]
- Zeeman, E.C. 11, 28, 42 [72]; 410, 524 [431]
- Zehnder, E. 236, 238, 239, 242, 246 [26]; 247 [42]; 247 [48]; 252, 260, 261, 272, 274, 313, 314, 316, 340 [382]; 343 [470]; 343 [471]; 344 [472]
- Zeng, C. 464, 522 [362]
- Zhang, D. 272, 313, 343 [449]; 344 [473]; 344 [474]; 344 [475]
- Zharnitsky, V. 242, 243, 247 [51]; 247 [52]; 247 [53]; 247 [54]
- Zhilinskiĭ, B.I. 318, 324, 331 [140]
- Zhu, W.-zh. 318, 336 [276]
- Zimmermann, M.G. 439, 440, 448, 524 [432]; 524 [433]; 524 [434]
- Zolpolhkadek, H. 468, 514 [129]
- Zou, M. 303, 335 [241]
- Zung, Nguyen Tien 318, 322, 324, 344 [476]; 344 [477]; 344 [478]



# Subject Index

- Action-angle variables, 227, 229, 230
- adiabatic invariants, 244, 245
- admissible measure, 54, 55
- almost every, 45, 49, 51, 54, 55, 60, 62, 65, 76
- almost everywhere, 49
- almost sure, 51
- angle doubling, 153
- approximation function, 269
- Arnold determinant, 275
- Arnold diffusion, 240, 245
- Arnold family of circle maps, 261
- Arnold tongues, 122, 261
- asymptotic value, 134, 215
- attractor, 351
  - Hénon, *see* Hénon attractor
  - Lorenz, *see* Lorenz attractor
- augmented itinerary, 197
- auto covariance, 366
- automorphism, 6
  
- Baire category, 53, 60
- Baire generic, 73, 74, 76, 77
- Baire property, 13
- Baker domain, 138
- basin of attraction, 135
- bifurcation
  - accessible, 414
  - homoclinic, *see* homoclinic bifurcation
  - local, *see* local bifurcation
- bifurcation diagram, 155
- bifurcations, 28
- Billiard, 243
- blue sky catastrophe, 403, 421
- Borel measure, 54
- Borel probability measures, 54, 55, 57, 58, 70
- Borel set, 71
- Borel subsets, 71
- bounded distortion, 350
- box counting dimension, 354
- Brin-Katok, 361
- Brjuno number, 131
- Bruno condition, 256, 270
  
- Cantor bouquet, 163, 166, 169, 172, 219
- Cantor manifold, 321
- Cantor set, 259, 269
  
- Cantor stratification, 299
- catastrophe theory, 11, 28
- category, 52
- centre manifolds, 30, 99
- centre stable manifold, 99
- centre unstable manifold, 97
- codimension, 382
- completely integrable, 227
- complex exponential map, 155
- complex standard family, 208
- conditionally periodic flow, 266
- conjugacy, 108
- conjugate, 38
- connection lemma, 408
- conservative, 451
- correlation dimension, 360
- correlation integral, 360
- critical elements, 406, 408
- critical value, 134
- cross section, 392, 393
- crown, 174
- cyclicity, 444, 448
  
- deaugmentation map, 197
- deformation, 15
- dense, 55
- density point, 280
- diffeomorphism, 6, 349
- differential equation, 5, 349
- Diophantine conditions, 236, 237, 241, 255, 259, 264, 268
  - mild, 269
  - normal-internal, 285, 291, 293, 306
- Diophantine stability, 309
- distance function, 385
  
- eigenvalues
  - leading, 386
- elementary, 23
- embedding, 349
- endomorphism, 6, 349
- endpoint, 176
- entropy, 361
- equilibrium, 381
  - elliptic, 302
  - hyperbolic, 302, 383
  - nonhyperbolic, 389



- of mixed type, 302
- relative, 289
- equivariant, 477
- Ergodic Hypothesis, 278
- essential singularity, 134
- essentially asymptotically stable, 487
- evolution map, 3
- evolutions, 3, 349
- explosion, 218
- exponential condensation, 300
- exponential tract, 212
- exterior Riemann map, 143
- external ray, 143, 207
  
- families, 49, 52, 69–72, 78
- family, 47, 49, 52, 53, 69–71
- Farey tree, 153, 170
- Fatou set, 138
- fattening, 292
- Fermi–Ulam’s ping-pong, 242
- filled Julia set, 140
- fingers, 193
- first-return map, 393, 394
- fixed point, 128
- forced symmetry breaking, 493
- formal linearization, 111
- formal normalization, 113
- fraying, 294
- frequency map, 271, 273
- local, 319
- frequency module, 267
- frequency-ratio map, 275
- full Lebesgue measure, 62, 73, 74
- full measure, 50, 70
- fundamental dichotomy, 141, 205
- fundamental domain, 106, 108
- fundamental set, 194
  
- general solution, 5
- generating function, 230, 238
- generic, 45, 46, 49–51, 53, 58, 60, 61, 63–68, 72, 74–76, 78, 81, 349
- generic set, 47, 50
- generically, 49, 60, 62, 63, 79
- genericity, 13, 52, 67, 74, 76, 78, 80
- gradient systems, 10
- graph transform, 91
- Grobman and Hartman, 21
- group orbit, 477
  
- Haar measure, 56, 62
- Haar null set, 55
- Haar zero set, 55
- hairs, 164, 188, 190
  
- Hamiltonian, 7, 23, 25
- Hamiltonian matrix, 305
- Hamiltonian systems, 7
- Hamiltonian vector field, 273
- Liouville integrable, 273
- locally, 318
- Hartman Grobman linearization, 39
- Hausdorff dimension, 175, 176
- heat equation, 6
- Hénon-like attractor, 415, 425, 431, 433
- Herman method, 312
- heteroclinic cycle, 381, 391
- intermittent time series, 418
- robust, 478
- simple, 478
- T-points, 445
- heteroclinic network, 478
- heteroclinic orbit, 381
- inclination-flip, 393
- nondegenerate, 391
- orbit-flip, 393
- homoclinic and heteroclinic bifurcations, 38
- homoclinic bifurcation
- gluing bifurcation, 419, 434
- homoclinic-doubling cascades, 416
- Shil’nikov–Hopf bifurcation, 390, 439
- homoclinic centre manifold, 388, 400, 420
- homoclinic cycle, 478
- robust, 478
- simple, 478
- homoclinic network, 478
- homoclinic orbit, 381
- almost homoclinic point, 408
- bellows, 435, 497
- bellows, reversible, 459
- bi-focus, 426
- butterfly, 435
- coexisting, 434
- coexisting, reversible, 458
- degenerate, 436, 497
- degenerate, reversible, 459
- figure-eight, 435
- inclination-flip, 388, 417, 431, 447, 448, 457, 469, 490, 494, 496
- multi-round, 382, 406
- non-orientable, 388
- nondegenerate, 386
- orbit-flip, 388, 390, 433, 457, 497
- orientable, 388, 500
- resonant, 428
- saddle-focus, 422, 471
- saddle-node homoclinic orbit, 414
- superhomoclinic, 411, 419
- tame, 422, 436, 498

- wild, 422
- homoclinic tangency, 407
- saddle-node homoclinic orbit, 415
- homological equation, 258, 281
- Hopf bifurcation, 33
- quasi-periodic, 290
- skew, 296
- subcritical, 290
- supercritical, 290
- Hopf Neïmark Sacker bifurcation, 36
- hyperbolic, 18
- hyperbolic component, 158
- hyperbolic fixed points, 108
- immediate basin of attraction, 135
- incompressible fluids, 8
- indecomposable continuum, 176
- infinitesimally symplectic matrix, 305
- initial state, 3
- integrability condition, 105
- integrable approximation, 114
- integrable vector field, 268, 271, 284, 289
- invariant foliations, 104
- invariant manifolds
  - centre manifolds, 389
  - centre-stable and -unstable manifolds, 390
  - homoclinic centre manifold, 400
  - leading manifolds  $W^{ls,u}$  and  $W^{s,lu}$ , 388
  - stable and unstable manifolds, 383
  - strong stable and unstable manifolds, 387
- invariant sets
  - stable and unstable set, 390
- invariant torus, 266
- atropic, 317
- coisotropic, 315
- elliptic, 302, 308
- elliptic normal modes of, 314
- Floquet, 288
- frequencies of, 267
- frequency vector of, 267
- higher dimensional, 315
- hyperbolic, 302, 308
- – whiskers of, 303, 314
- internal frequencies of, 305
- internal frequency vector of, 284
- isotropic, 303
- KAM, 274
- Lagrangean, 273
- lower dimensional, 303
- non-resonant, 267
- normal frequencies of, 285, 305
- ‘normal’ matrix of, 284, 305
- of mixed type, 302, 308
- reducible, 314
- resonant, 301
- superexponentially sticky, 300
- with parallel dynamics, 267
- with quasi-periodic dynamics, 267
- inverse graph transform, 100
- isotropy group, 477
- isotypic
  - component, 477
  - decomposition, 477
- itinerary, 196
- Jet extensions, 16
- Julia ray, 212
- Julia set, 133, 136, 155
- Kantz-Diks test, 375
- Knaster continuum, 176, 178
- Knaster-like continuum, 183
- kneading sequence, 190
- Kolmogorov–Arnold theorem, 237
- Kupka-Smale theorem, 19
- lambda-lemma, 411
- lattice of frequencies, 267
- leaves, 104
- Lebesgue almost every, 45, 48–53, 59, 62, 72
- Lebesgue measure, 47–53, 56, 58, 59, 62, 63, 69, 70, 72
- Lebesgue measure zero, 45–47, 50, 52–54, 56, 59, 62, 63, 73
- Liapunov exponents, 373
- Liapunov-Schmidt reduction, 120
- $\omega$ -limit, 351
- limit cycle, 420
  - bounds, 430
- Lin’s method, 399
- linear prediction, 372
- linear prevalence, 60, 73
- Linearization theorem, 129
- linearizations, 108
- linearizing coordinates, 394
- Liouville–Arnold theorem, 227
- local bifurcation
  - $0^{2+}i\omega$  resonance, 474
  - 1:1 resonance, 473
  - Bogdanov–Takens, 467
  - Hopf bifurcation, 390
  - saddle-node bifurcation, 389
- local bifurcations, 30
- local dynamical system, 5, 31
- Localization, 98
- localization procedure, 287, 307
- locally invariant manifolds, 98
- logarithmic singularity, 211

- Lorenz attractor, 408
- geometric or Lorenz-like, 408, 414, 469, 495, 503
- Lorenz equations, 408
- Lorenz models
  - contracting, 409
  - hooked, 410
- Malgrange, Thom, Mather, 29
- Mandelbrot set, 142
- meagre, 53, 80
- meagreness, 54
- measure of relative frequencies, 358
- measure zero, 51, 52, 54, 56, 70, 81
- measure-theoretically, 58
- Melnikov integral, 385
- miniversal unfolding, 284
- Misiurewicz point, 157
- moduli, 39, 501
- saddle-focus homoclinic orbit, 424, 436
- Montel's Theorem, 134
- Morse index, 391
- Morse lemma, 29
- Morse–Smale flow, 414
- Moser's twist theorem, 236
- multiplier map, 162
- negligibility, 81
- negligible, 52, 69
- Nekhoroshev estimate, 280
- Nekhoroshev theorem, 280
- non-shy, 58, 59
- nondegeneracy condition, 271
- BHT, 285, 305
- iso-energetic, 275
- Kolmogorov, 273
- global, 319
- Moser, 284
- Rüssmann, 275
- nonlinear, 74
- nonlinear prevalence, 68–70, 72, 74
- normal family, 133
- normal forms, 112, 389
- normal hyperbolicity, 110
- normal linear stability, 286, 307
- normally attracting, 103
- not shy, 58
- nowhere, 60
- nowhere dense, 60
- open  $C^r$ -dense, 68
- open and dense, 51, 68, 73
- open dense, 49, 50, 66, 68, 69, 76, 78
- open dense set, 46
- orbit, 127
  - symmetric, 452
- orbits, 3
- orientation index, 388, 401
- Outer billiard, 244
- Paley–Wiener estimates, 282
- parabolic equations, 6
- parallel flow, 266
- parameter plane, 183, 202
- partial differential equations, 6
- partial linearizations, 109
- partially hyperbolic set, 408
- period doubling bifurcations, 35, 158, 191
- period maps, 6
- periodic orbit
  - multi-round, 382, 406
- periodic point, 128
- persistent, 14
- perturbing family, 15
- Poincaré map, 6, 11, 393
- Poincaré metric, 133
- Poincaré return, 10
- Poisson manifold, 318
- polycycle, 391
  - hyperbolic, 444
- porosity, 81
- porous, 80
- power spectrum, 366
- predictability, 351, 370
- prevalence, 45, 52–55, 58, 59, 62, 64, 65, 67–70, 72, 78, 80
- prevalent, 50, 53–55, 58–64, 66, 69–74, 79–81
- prevalent transversality theorem, 45
- prevalently, 61, 62
- probability measure, 54, 57, 71, 74
- probability zero, 46
- Pugh's Closing Lemma, 26
- quasi-convex Hamiltonian, 299
- Quasi-Ergodic Hypothesis, 317
- quasi-invariant, 70, 71
- quasi-periodic attractor, 262
- quasi-periodic diffeomorphism, 261
- quasi-periodic flow, 267
- quasi-periodic stability, 261, 268, 272, 274, 286, 307
  - global, 320, 323
  - weak, 287
- quasiperiodic forcing, 242
- read out function, 349
- reconstruction
  - map, 349
  - vectors, 349
- reconstruction measures, 359

- recurrent, 10
- recurrent orbit, 382
- regularity properties, 322
- relative prevalence, 80
- renormalization operator, 418
- residual, 13, 54, 65, 78
- resonance horns, 122
- resonance tongue, 261
- resonant zone, 301
- reverser, 451
- reversible, 451
- reversible system, 254
  - weakly, 254
- reversible systems, 8
- Riemann mapping theorem, 132
- rotation number, 434
- saddle quantity, 386
- saddle-node bifurcation, 32, 156
- Sard's theorem, 14
- Schwarz lemma, 133, 166
- Schwarzian derivative, 210
- second category, 52
- (semi-)algebraic set, 17
- (semi-)group property, 3
- shift map, 10, 195
- Shil'nikov variables, 396
- shy, 53–60, 71, 79, 80
- shyness, 53–55, 58, 59, 79, 81
- singular
  - cycle, 408, 410, 414, 450
  - horseshoe, 413, 432
  - hyperbolic attractor, 409, 493
  - hyperbolic set, 409
- singularity, 18
- spherical pendulum, 321
- Splitting lemma, 29
- stable fingers, 195
- stable foliation, 107, 401
- stable manifold, 20
- stable set, 138
- state space, 3
- stationary, 358
- stationary point, 18
- stem, 176
- straight brush, 170
- stroboscopic maps, 6
- strong shyness, 79
- strong unstable manifold, 97
- structural stability, 38
- structurally stable, 187
- structurally unstable, 183
- sub-shift, 10
- suspended horseshoe, 407, 410
  - creation or disappearance, 410
- suspension, 11
- symbolic dynamics, 9
- symplectic systems, 7
- systems depending on parameters, 9
- systems with symmetry, 9
- taking restrictions, 13
- tangent family, 213
- The inclination lemma, 99
- theorem of Grobman and Hartman, 108
- Thom's transversality lemma, 14
- time  $t$  map, 12
- time series, 349
  - deterministic, 352
  - intermittent, 418
  - random, 352
- time set, 3, 349
- topological entropy, 356
- topological genericity, 51, 53, 60, 78, 80
- topologically generic, 45–47, 49, 52, 66, 73, 79, 80
- total measure, 52
- transition matrix, 201
- transitive sets
  - robust, 409
- transversality theorem, 72
- twist, 233, 478
- twist map, 264
  - theorem on, 264
- typical/prevalent, 63
- unfolding, 15
- unstable foliation, 105
- unstable manifold, 21, 97
- Van der Pol oscillator, 262
- variational equation, 384
  - adjoint, 384, 391
- vector field, 5, 349
- versal unfolding, 284
- volume preserving, 23, 25
- volume preserving systems, 7
- Yoccoz' Theorem, 132