

Applied
Mathematical
Sciences
41

Colin Sparrow

The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors



Springer-Verlag

Applied Mathematical Sciences | Volume 41

Applied Mathematical Sciences

1. *John*: Partial Differential Equations, 4th ed.
2. *Sirovich*: Techniques of Asymptotic Analysis.
3. *Hale*: Theory of Functional Differential Equations, 2nd ed.
4. *Percus*: Combinatorial Methods.
5. *von Mises/Friedrichs*: Fluid Dynamics.
6. *Freiberger/Grenander*: A Short Course in Computational Probability and Statistics.
7. *Pipkin*: Lectures on Viscoelasticity Theory.
9. *Friedrichs*: Spectral Theory of Operators in Hilbert Space.
11. *Wolovich*: Linear Multivariable Systems.
12. *Berkovitz*: Optimal Control Theory.
13. *Bluman/Cole*: Similarity Methods for Differential Equations.
14. *Yoshizawa*: Stability Theory and the Existence of Periodic Solution and Almost Periodic Solutions.
15. *Braun*: Differential Equations and Their Applications, 3rd ed.
16. *Lefschetz*: Applications of Algebraic Topology.
17. *Collatz/Wetterling*: Optimization Problems.
18. *Grenander*: Pattern Synthesis: Lectures in Pattern Theory, Vol I.
20. *Driver*: Ordinary and Delay Differential Equations.
21. *Courant/Friedrichs*: Supersonic Flow and Shock Waves.
22. *Rouche/Habets/Laloy*: Stability Theory by Liapunov's Direct Method.
23. *Lamperti*: Stochastic Processes: A Survey of the Mathematical Theory.
24. *Grenander*: Pattern Analysis: Lectures in Pattern Theory, Vol. II.
25. *Davies*: Integral Transforms and Their Applications, 2nd ed.
26. *Kushner/Clark*: Stochastic Approximation Methods for Constrained and Unconstrained Systems
27. *de Boor*: A Practical Guide to Splines.
28. *Keilson*: Markov Chain Models—Rarity and Exponentiality.
29. *de Veubeke*: A Course in Elasticity.
30. *Sniatycki*: Geometric Quantization and Quantum Mechanics.
31. *Reid*: Sturmian Theory for Ordinary Differential Equations.
32. *Meis/Markowitz*: Numerical Solution of Partial Differential Equations.
33. *Grenander*: Regular Structures: Lectures in Pattern Theory, Vol. III.
34. *Kevorkian/Cole*: Perturbation methods in Applied Mathematics.
35. *Carr*: Applications of Centre Manifold Theory.
36. *Bengtsson/Ghil/Källén*: Dynamic Meteorology: Data Assimilation Methods.
37. *Saperstone*: Semidynamical Systems in Infinite Dimensional Spaces.
38. *Lichtenberg/Lieberman*: Regular and Stochastic Motion.
39. *Piccinini/Stampacchia/Vidossich*: Ordinary Differential Equations in \mathbb{R}^n .
40. *Naylor/Sell*: Linear Operator Theory in Engineering and Science.
41. *Sparrow*: The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors.
42. *Guckenheimer/Holmes*: Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.
43. *Ockendon/Taylor*: Inviscid Fluid Flows.
44. *Pazy*: Semigroups of Linear Operators and Applications to Partial Differential Equations.
45. *Glashoff/Gustafson*: Linear Optimization and Approximation: An Introduction to the Theoretical Analysis and Numerical Treatment of Semi-Infinite Programs.
46. *Wilcox*: Scattering Theory for Diffraction Gratings.
47. *Hale et al.*: An Introduction to Infinite Dimensional Dynamical Systems—Geometric Theory.
48. *Murray*: Asymptotic Analysis.
49. *Ladyzhenskaya*: The Boundary-Value Problems of Mathematical Physics.
50. *Wilcox*: Sound Propagation in Stratified Fluids.
51. *Golubitsky/Schaeffer*: Bifurcation and Groups in Bifurcation Theory, Vol. I.
52. *Chipot*: Variational Inequalities and Flow in Porous Media.
53. *Majda*: Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables.
54. *Wasow*: Linear Turning Point Theory.

Colin Sparrow

The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors

With 91 Illustrations



Springer-Verlag
New York Heidelberg Berlin

Colin Sparrow
King's College
Cambridge
U.K.

Mathematical Subject Classification Codes: 35B32, 25Q20, 86A10

Library of Congress Cataloging in Publication Data
Sparrow, Colin.

The Lorenz equations.

(Applied mathematical sciences; v. 41)

Bibliography: p.

Includes index.

I. Lorenz equations. 2. Bifurcation theory.

I. Title. II. Series: Applied mathematical sciences
(Springer-Verlag New York Inc.); v. 41.

QA1.A647 vol. 41 [QA372] 510s [515.3'52] 82-19435

© 1982 by Springer-Verlag New York Inc.

All rights reserved. No part of this book may be translated or reproduced in any form without written permission from Springer-Verlag, 175 Fifth Avenue, New York, New York 10010, U.S.A.

9 8 7 6 5 4 3 2

ISBN-13: 978-0-387-90775-8 e-ISBN-13: 978-1-4612-5767-7
DOI: 10.1007/978-1-4612-5767-7

Preface

The equations which we are going to study in these notes were first presented in 1963 by E. N. Lorenz. They define a three-dimensional system of ordinary differential equations that depends on three real positive parameters. As we vary the parameters, we change the behaviour of the flow determined by the equations. For some parameter values, numerically computed solutions of the equations oscillate, apparently forever, in the pseudo-random way we now call "chaotic"; this is the main reason for the immense amount of interest generated by the equations in the eighteen years since Lorenz first presented them. In addition, there are some parameter values for which we see "preturbulence", a phenomenon in which trajectories oscillate chaotically for long periods of time before finally settling down to stable stationary or stable periodic behaviour, others in which we see "intermittent chaos", where trajectories alternate between chaotic and apparently stable periodic behaviours, and yet others in which we see "noisy periodicity", where trajectories appear chaotic though they stay very close to a non-stable periodic orbit.

Though the Lorenz equations were not much studied in the years between 1963 and 1975, the number of man, woman, and computer hours spent on them in recent years - since they came to the general attention of mathematicians and other researchers - must be truly immense. Besides this mainstream of "Lorenz" research, countless authors have quoted the equations as an example or referred to them in the course of some debate or other. Despite all this interest, there has never been an attempt to tie together all the various different parameter ranges. The first purpose of these notes is to attempt to fill that gap.

For some parts of these notes I can claim no great originality. I have, wherever possible, referenced other authors, but should warn that these notes are not intended to be a complete review of the literature. I have been guided by what I have seen, and apologize to those authors whom I have omitted or unintentionally misinterpreted. As often, the most notable omissions are likely to be Russian authors. Nowhere should my references be taken to indicate any firm belief that the author references wrote either the first, or the best, paper on a subject. In attempting to give as thorough a description as possible of the various bifurcations which lead from one well studied parameter range to another, I hope that I have added to, rather than subtracted from, the various separate contributions so far.

These notes are also intended to illustrate an approach. It cannot be claimed that the Lorenz equations show all the different behaviours of a general set of chaotic ordinary differential equations. Indeed, they possess certain special properties (such as a symmetry) that indicate that this cannot be so. Nonetheless, many other systems behave in ways which seem to be very similar to one or more of the behaviours shown by the Lorenz equations, and an understanding of the Lorenz equations can be expected to increase our understanding of these other systems. The way in which we obtain this understanding is to move constantly back and forth between theory, models which demonstrate those properties which we deem to be important at any particular time, and numerical experiments on the equations themselves. Thus, we avoid some of the pitfalls of more single-minded approaches. Those who seek to know, with mathematical certainty, what the Lorenz equations "do" will be disappointed. Most often we proceed only to the point where we know "beyond all reasonable doubt", and readers should always bear in mind that much of these notes is dependent on computer generated numerical output which can always be misleading for reasons outside our control or even outside our comprehension.

It is my hope that these notes will be comprehensible to those readers with no previous knowledge of the Lorenz equations (or any other chaotic differential equations), as well as being informative and interesting to the "experts". It is assumed that readers know a little about differential equations and the various simpler bifurcations which can occur as parameters change.

Readers may notice that there is very little discussion of the problem of "real world" turbulence. This is a deliberate policy. It is not that I believe that the study of chaotic ordinary differential equations

can never be helpful in understanding real world phenomena; rather, I believe that until we know more about the behaviour of the finite-dimensional approximations that model the partial differential equations that model the world, I do not think I have anything very useful to say on the subject.

These notes are divided into nine chapters and eleven appendices. The first four chapters review what is known about the Lorenz equations in the most widely studied parameter ranges. Chapter 1 contains some general remarks and a description of those simple properties of the equations that can be deduced mathematically. In Chapter 2, we study the bifurcation associated with a homoclinic orbit. This study is more general than usual, since we shall see that there are many important homoclinic bifurcations in the Lorenz system. In Chapter 3, we describe the parameter range where it is believed that we have a well understood strange attractor in a whole interval of parameter values. Chapter 4 contains a description of the results of some simple numerical experiments in a parameter range where period doubling is observed. In Chapter 5, we attempt to reconcile Chapters 3 and 4. Using a combination of different numerical techniques and a careful theoretical analysis of the changes in the behaviour of the unstable manifold of the origin (which is dependent on our general knowledge of homoclinic bifurcations), we show how the behaviour changes from strange attractor to period doublings. In the process, we uncover various new properties of the Lorenz equations, including indications that our simple numerical experiments from Chapter 4 are misleading in various ways. Chapter 5 is probably not comprehensible without at least a quick reading of Chapters 1 through 4.

Chapter 6 is an attempt to justify the methods we have used earlier in the text to describe periodic orbits and other trajectories using sequences of symbols. In the process of this justification, we study the behaviour of the stable manifolds of the stationary points other than the origin. Chapter 6 could be omitted.

Chapter 7 contains an outline of an analysis of the behaviour when one of the parameters becomes large. This analysis goes considerably beyond earlier analyses and suggests that the "large r " behaviour may be qualitatively more complicated when the other two parameters are allowed to take values other than the usual ones. Chapter 7 could be read in isolation.

Chapter 8 contains a study of the Lorenz equations for parameter values suggested by Chapter 7. The expectations of qualitatively more

complicated behaviour are confirmed. Nonetheless, the general theory and approach of Chapters 1 through 7 shows us that we can, with very few numerical experiments, make a large number of interesting statements about this more complicated behaviour. In particular, we discover a kind of bifurcation not previously observed in the Lorenz equations, and can explain exactly how this bifurcation fits into the more general picture.

Chapter 9 contains a brief summary, a quick examination of some of the approaches used by other authors on the Lorenz equations, and a brief discussion of some Lorenz-like equations at present under investigation.

Throughout Chapters 1-9, I have attempted to confine my attention to those things with direct relevance to the actual Lorenz equations (as opposed to models or simplifications of the equations). The Appendices are of several types. Some contain little bits of mathematics which, though of direct relevance to the discussion in the main body of the text, have been relegated to an appendix so as to avoid breaking the flow of the description; often we can proceed just by quoting the results from the relevant appendix. Some appendices are self-contained and describe results with application to more than the Lorenz equations. Examples are the appendices on homoclinic bifurcations, and on numerical techniques for the location and following (with changing parameter) of non-stable periodic orbits. These appendices may be of separate use to some readers; in any case it is convenient to have these results in one place since they are referred to several times in the main body of the text. One appendix contains a review of work on a geometric model of the Lorenz equations in a parameter range where the strange attractor is believed to exist.

The numerical integration of differential equations, on which these notes depend, was done using standard integrating packages. Most of the simulations have been done on two different machines, using different packages. These were variously based on Merson's method, a variable order Adam's method, and a variable order Runge-Kutta method, all of which produced similar results.

Colin Sparrow
Berkeley, California
April 1982

Acknowledgements

These notes grew out of the latter half of my Ph.D. thesis, completed in 1980 for the University of Cambridge. I owe a great deal to my research supervisor, Alistair Mees, who encouraged, informed and cajoled, and to Peter Swinnerton-Dyer who first persuaded me that the Lorenz equations were still worth studying. The latter also helped with various pieces of analysis; in particular, the large r results (Chapter 7 and Appendix K) would not have been possible without his help.

I have had the opportunity to talk to many mathematicians and others about the Lorenz equations; to all I say thank you. Bob Williams deserves a special mention, not only did he talk with me at great length in the early stages of my research, but he also had me to stay in his house in Evanston for a week. Various people have looked through different versions of these notes; I am grateful to Jack Hale and John Guckenheimer for some general remarks, and to Peter Swinnerton-Dyer, Andrew Fowler, Edgar Knobloch and Sian Stumbles for detailed criticisms. I would also like to thank Kate MacDougall who did a wonderful job of typing the final version, Eleanor Addison who drew all the figures that were not drawn by the computer, David Abrahamson who gave editorial assistance, and the staff at Springer-Verlag who guided me through the whole painful process. As usual, however, I retain sole responsibility for all the errors, mistakes, and other deviations from the truth that may be contained within these notes.

I would like to conclude by thanking the staff of the two computer centers where I did most of my numerical work (Cambridge and Brown). Without them these notes would never have existed, and they deserve far more praise than they ever seem to receive.

Contents

	Page
Preface	v
Chapter 1. INTRODUCTION AND SIMPLE PROPERTIES	1
1.1. Introduction	1
1.2. Chaotic Ordinary Differential Equations	4
1.3. Our Approach to the Lorenz Equations	6
1.4. Simple Properties of the Lorenz Equations	8
Chapter 2. HOMOCLINIC EXPLOSIONS: THE FIRST HOMOCLINIC EXPLOSION	13
2.1. Existence of a Homoclinic Orbit	13
2.2. The Bifurcation Associated with a Homoclinic Orbit	16
2.3. Summary and Some General Definitions	22
Chapter 3. PRETURBULENCE, STRANGE ATTRACTORS AND GEOMETRIC MODELS	26
3.1. Periodic Orbits for the Hopf Bifurcation	26
3.2. Preturbulence and Return Maps	28
3.3. Strange Attractor and Homoclinic Explosions	32
3.4. Geometric Models of the Lorenz Equations	43
3.5. Summary	49
Chapter 4. PERIOD DOUBLING AND STABLE ORBITS	51
4.1. Three Bifurcations Involving Periodic Orbits	51
4.2. $99.524 < r < 100.795$. The x^2y Period Doubling Window	56
4.3. $145 < r < 166$. The x^2y^2 Period Doubling Window	59
4.4. Intermittent Chaos	62
4.5. $214.364 < r < \infty$. The Final xy Period Doubling Window	66
4.6. Noisy Periodicity	69
4.7. Summary	74
Chapter 5. FROM STRANGE ATTRACTOR TO PERIOD DOUBLING	76
5.1. Hooked Return Maps	76
5.2. Numerical Experiments	79
5.3. Development of Return Maps as r Increases: Homoclinic Explosions and Period Doubling	86
5.4. Numerical Experiments on Periodic Orbits	92
5.5. Period Doubling and One-Dimensional Maps	100
5.6. Global Approach and Some Conjectures	103
5.7. Summary	113
Chapter 6. SYMBOLIC DESCRIPTION OF ORBITS: THE STABLE MANIFOLDS OF C_1 AND C_2	115
6.1. The Maxima-in-z Method	116
6.2. Symbolic Descriptions from the Stable Manifolds of C_1 and C_2	119
6.3. Summary	130

	Page
Chapter 7. LARGE r	132
7.1. The Averaged Equations	132
7.2. Analysis and Interpretation of the Averaged Equations	136
7.3. Anomalous Periodic Orbits for Small b and Large r	140
7.4. Summary	149
Chapter 8. SMALL b	151
8.1. Twisting Around the z-Axis	151
8.2. Homoclinic Explosions with Extra Twists	153
8.3. Periodic Orbits Without Extra Twisting Around the z-Axis	158
8.4. Heteroclinic Orbits Between C_1 and C_2	164
8.5. Heteroclinic Bifurcations	169
8.6. General Behaviour When $b = 0.25$	171
8.7. Summary	176
Chapter 9. OTHER APPROACHES, OTHER SYSTEMS, SUMMARY AND AFTERWORD	179
9.1. Summary of Predicted Bifurcations for Varying Parameters σ , b and r	179
9.2. Other Approaches	184
9.3. Extensions of the Lorenz System	186
9.4. Afterword - A Personal View	189
Appendix A. DEFINITIONS	192
Appendix B. DERIVATION OF THE LORENZ EQUATIONS FROM THE MOTION OF A LABORATORY WATER WHEEL	194
Appendix C. BOUNDEDNESS OF THE LORENZ EQUATIONS	196
Appendix D. HOMOCLINIC EXPLOSIONS	199
Appendix E. NUMERICAL METHODS FOR STUDYING RETURN MAPS AND FOR LOCATING PERIODIC ORBITS	211
Appendix F. COMPUTATIONAL DIFFICULTIES INVOLVED IN CALCULATING TRAJECTORIES WHICH PASS CLOSE TO THE ORIGIN	221
Appendix G. GEOMETRIC MODELS OF THE LORENZ EQUATIONS	223
Appendix H. ONE-DIMENSIONAL MAPS FROM SUCCESSIVE LOCAL MAXIMA IN z	234
Appendix I. NUMERICALLY COMPUTED VALUES OF $k(r)$ FOR $\sigma = 10$ and $b = 8/3$	239
Appendix J. SEQUENCES OF HOMOCLINIC EXPLOSIONS	244
Appendix K. LARGE r; THE FORMULAE	257
Bibliography	262
Index	268

Chapter 1

Introduction and Simple Properties

(Mathematical definitions are given in Appendix A.)

1.1. INTRODUCTION

In 1963 E. N. Lorenz wrote a remarkable paper. In it he described a three parameter family of three-dimensional ordinary differential equations which appeared, when integrated numerically on a computer, to have extremely complicated solutions. These equations, now known as the Lorenz equations, have been studied by many authors in the years since 1963 and one of our aims, in these notes, is to contribute to this study. It is necessary, therefore, to explain some of the reasons why the equations generated so much interest initially and why they warrant further study.

Lorenz's search for a three-dimensional set of ordinary differential equations which would model some of the unpredictable behaviour which we normally associate with the weather (Lorenz is a meteorologist as well as a mathematician) is described in Lorenz (1979). The equations which he eventually hit upon were derived from a model of fluid convection. They are

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y-x) \\ \frac{dy}{dt} &= rx - y - xz \\ \frac{dz}{dt} &= xy - bz \end{aligned} \tag{*}$$

where σ , r and b are three real positive parameters.

Briefly, the original derivation (Lorenz, 1963) can be described as follows. A two-dimensional fluid cell is warmed from below and cooled from above and the resulting convective motion is modelled by a partial differential equation. The variables in this partial differential equa-

tion are expanded into an infinite number of modes, all but three of which are then set identically to zero. The three remaining modes give the equations (*). Roughly speaking, the variable x measures the rate of convective overturning, the variable y measures the horizontal temperature variation, and the variable z measures the vertical temperature variation. The three parameters σ , r and b are respectively proportional to the Prandtl number, the Rayleigh number, and some physical proportions of the region under consideration; consequently, all three are taken to be positive.

For wide ranges of values of the parameters, approximate solutions to the equations (*), calculated on a computer, look extremely complicated. Fig. 1.1 shows the projection onto the x, z plane ($y = \text{constant}$) of one such solution calculated when $\sigma = 10$, $b = 8/3$ and $r = 28$. Note that the trajectory shown does not intersect itself if we consider the full three-dimensional picture. The crossings in Fig. 1.1 are the result of projection onto two dimensions.

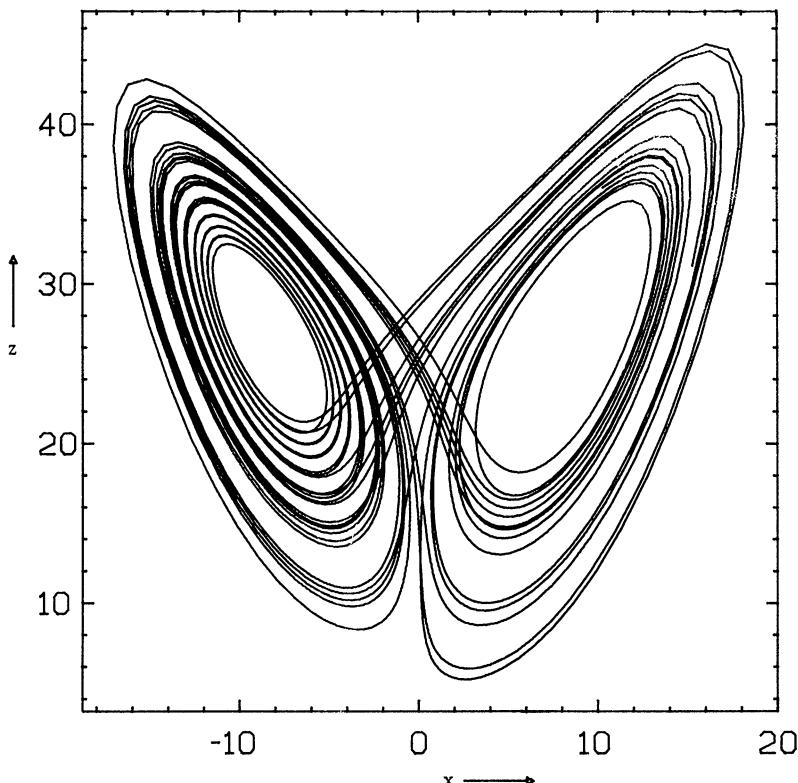


Figure 1.1. A numerically computed solution to the Lorenz equations projected onto the x, z plane ($\sigma = 10$, $b = 8/3$, $r = 28.0$).

In the next section (1.2) we will consider the extent to which figures such as 1.1 are artifacts of the computer programs which generate them and the extent to which they reflect the true dynamics of the equations (*). For the moment we should note that the figure appears to have various 'turbulent' properties as follows:

1. The trajectory shown in Fig. 1.1 is not periodic.
2. The figure does not appear to show a transient phenomenon. However long we continue the numerical integration the trajectory continues to wind around and around, first on one side, then on the other, without ever settling down to either periodic or stationary behaviour.
3. The general form of the figure does not depend at all on our choice of initial conditions (providing we ignore initial transient sections of a trajectory) or on our choice of integrating routine. Anything from a first-order Euler method with sufficiently small step size to a sophisticated multi-order integrating technique will produce a similar picture.
4. The details of the figure depend crucially on both the factors mentioned in 3) above. The exact sequence of loops which the trajectory makes is extremely sensitive to both changes in initial conditions and changes in the integrating routine. As a consequence of this, it is not possible to predict the details of how the trajectory will develop over anything other than a very short time interval.

Lorenz's original paper was titled "Deterministic Non-periodic Flows". Notice that the Lorenz equations are deterministic. They contain no random, noisy or stochastic terms and we know that they determine a unique flow which is valid for all time. The suggestion, motivated partly by Lorenz's work, that complicated 'turbulent' behaviour in systems with an infinite number of degrees of freedom (such as the atmosphere) might be modelled by simple deterministic finite-dimensional systems (such as the Lorenz equations) is one of the reasons why the Lorenz equations have attracted so much attention.

This question of the relationship between turbulent behaviour in the real world and 'turbulent' behaviour in finite-dimensional systems is still unresolved. Consequently we should state at once that Lorenz himself admits (1979) that the equations (*) are not a realistic model of his original fluid dynamical problem if the parameter r is far from unity. In addition, if we examine higher-dimensional truncations of Lorenz's original infinite-dimensional problem then it seems that the type

of behaviour observed in Fig. 1.1 does not always occur in the same way (see Chapter 9). However, these two observations do not imply that there is no relationship between infinite-dimensional and finite-dimensional 'turbulence', nor that the Lorenz equations are irrelevant to the debate about this relationship. Largely because of the amount of interest generated by the Lorenz equations, other authors have sought to discover, or stumble upon, other real world problems for which the equations (*) are an accurate model when r is much larger than one. They have had some success. Haken (1975) derives the Lorenz equations from a problem of irregular spiking in lasers, Malkus (1972) and Yorke & Yorke (1978) both studied a problem of convection in a toroidal region, and Knobloch (1981) discusses a derivation from a disc dynamo. Prof. Malkus has constructed a laboratory water wheel which behaves in a similar fashion to the equations (*) and from whose equations of motion (*) can be derived (Lorenz, 1979). This derivation is described in Appendix B. Using asymptotic methods, Pedlosky (1972) and Pedlosky & Frenzen (1980) have derived the Lorenz equations from a study of the dynamics of a weakly unstable, finite amplitude, baroclinic wave (two-layer model), Brindley & Moroz (1980) obtain the equations in a similar problem (continuously stratified model), and Gibbon & McGuinness (1980) discuss both the two-layer baroclinic model and a laser problem. See, also, Gibbon & McGuinness (1981) and Gibbon (1981).

Despite this great number of possible applications for the Lorenz equations, it is my intention to study them as differential equations without entering the sometimes heated debate about the relationship between the equations and the problems they are supposed to model. This debate is mentioned again, briefly, in Chapter 9 (after we know something about the Lorenz equations), and readers are referred to Ruelle & Takens (1971) for one of the earlier and more important papers about the problem. See, also, Ruelle (1976, 1977, 1979).

1.2. CHAOTIC ORDINARY DIFFERENTIAL EQUATIONS

We now know of a great number of sets of ordinary differential equations derived, convincingly or not, from an even greater number of real world problems, which have solutions that look, in some respect or other, like Fig. 1.1. These equations are generally called "chaotic", as are their numerically calculated (and approximate) solutions. We shall only be interested in dissipative chaotic systems in these notes; in Section 1.4 we shall see that the Lorenz equations are dissipative. Readers

interested in chaos in Hamiltonian systems should consult (for an introduction) either Helleman (1980) or Lichtenberg & Lieberman (1982). Both these references also discuss dissipative systems and the first reference contains an extensive bibliography.

When a system is bounded, as well as dissipative (which most, including the Lorenz system, are), we can deduce that all trajectories eventually tend towards some bounded set of zero volume lying in the phase space. (See Section 1.4). Though there are technical distinctions which allow us to define this set in various different ways (see Appendix C), we can state that we are especially interested in the bounded set of zero volume called the non-wandering set. This set contains all the recurrent behaviour of the flow and we expect that all true trajectories will tend towards it. The non-wandering set may have several components. A component might be a stationary point, a periodic orbit, or some more complicated set of zero volume. If we know the structure of the non-wandering set, the way that the flow behaves on the non-wandering set, and the parts of the non-wandering set which are attracting (since some parts of it may be unstable), then we can sensibly claim that we know all the important things about our differential equation. If, for some practical application, we need to know how a particular trajectory with a particular set of initial conditions behaves, we can attempt to discover this by experiment. Eventually, though, the trajectory will move close to the non-wandering set and its behaviour will be governed by the motion of the flow on this set.

In these notes we shall use the term "chaotic" to describe the numerical phenomenon seen in figures such as Fig. 1.1. We shall not use expressions like "chaotic attractor" or "strange attractor" unless we have a special reason to believe that we really are seeing an approximate trajectory which is lying close to a single (strange) attracting piece of the non-wandering set. In this way we will avoid confusing what we see with what we understand. But since much of these notes depends on numerical integrations, it is important that we ask what we can learn about the non-wandering set by looking at figures such as Fig. 1.1.

Essentially, all we can say about Fig. 1.1 is that there is something to be explained. The figure is not the result of inaccurate, inadequate or inappropriate numerical method. The Lorenz equations (and most other chaotic systems of equations) are not stiff or otherwise difficult to integrate numerically, and whilst the details of the figure do depend on the details of the computer program that produces it, the chaotic nature of

the solution does not. We can be certain, for example, that we are not merely having numerical difficulties following a trajectory that ought to be tending towards some low period stable periodic orbit with a large basin of attraction which is distant from other parts of the non-wandering set. That is about all that we can be certain of, though. Perhaps a large collection of low period periodic orbits, each with a basin of attraction smaller than the resolution of our computation, would cause our approximate solution to jump "randomly" from one of these orbits to another. Or perhaps we are seeing one small portion of a stable periodic orbit of very high period. (Indeed, if we are being pedantic and using a digital computer, we know we are seeing a portion of a very high period orbit; the machine has only a finite number of states to cycle through before it must repeat itself.) By merely looking at a chaotic trajectory we cannot decide between these and other possibilities. The point we are making here is not purely rhetorical; we shall see that the chaotic solutions shown in Figs. 1.1 and 5.6 are probably lying near non-wandering sets with quite different structures, despite their apparent similarity. We shall find the term "chaotic" useful only if we remember that our definition is loose, numerical, and non-rigorous.

1.3. OUR APPROACH TO THE LORENZ EQUATIONS

The actual techniques, numerical and mathematical, which we will use to study the Lorenz equations will be introduced as we proceed. Our general aim will be to discover as much as possible about the behaviour of the system for a wide range of parameter values. In particular, we will hope to understand the many different kinds of chaotic behaviour that have been observed by other authors. The emphasis will be on global, geometric and intuitive understanding, and we will be especially interested in the way the different chaotic regimes fit together in our global picture.

The notion of bifurcation is central to this approach. As we change the parameters, the behaviour of the flow will only change in an important way when the topology of the non-wandering set changes. Each time this occurs, we say there is a bifurcation. Many bifurcations can be dealt with theoretically at a local level; our problem will be to fit them all together into a global picture. If we can build a global picture, however tentative, that allows us to explain the observed changes in behaviour via a theoretically acceptable sequence of bifurcations, our understanding of the behaviour at particular parameter values will be

enhanced. If we cannot, then the problems involved in building the global picture may suggest to us where we should look to find as yet unobserved bifurcations.

It must be remembered that any global statements we make will be tentative. Non-linear systems of ordinary differential equations are not well understood, and the Lorenz equations are no exception. There are two levels (at least) at which any global picture will be uncertain. The first is the numerical and observational level. We can only ever do a finite number of numerical experiments at a finite number of parameter values, each with only finite accuracy. It is always possible that, for some parameter values we have not examined, the behaviour is completely different, or that there are strange things going on in some region of phase space that we have not investigated. Whilst we will take as much care with our numerical experiments as seems appropriate, we can never completely dispel doubts of this kind. However, until new numerical experiments indicate where we have gone wrong, the picture we have will be the best available description of the Lorenz equations. Providing the picture is self-consistent, it may remain of interest even if it eventually transpires that it is not an accurate description of the Lorenz system.

The second level of uncertainty is the theoretical one. Even if we assume that our numerical experiments are not misleading, our theoretical knowledge may not be adequate for us to be able to answer all the questions about the flow that we would like. I hope that I have not glossed over these uncertainties whenever they occur in the discussion.

Besides providing a new and more complete picture of the Lorenz equations, these notes are intended to fulfill another major function. That is to illustrate how one can use a combination of local theoretical results and numerical experiments to build up a global intuitive understanding of a system of chaotic differential equations. In this context we could ask, "Why the Lorenz equations?". The extent to which the Lorenz equations are typical of other chaotic systems will be discussed again in Chapter 9 after we know a little more about them. For the moment it will do to state that the Lorenz equations, for various different parameter values, seem to display most of the kinds of chaotic behaviour observed in other three-dimensional systems of chaotic differential equations. In fact, the Lorenz system appears to show a greater range of different behaviours than most systems. Of course, this may only be because the Lorenz system has been studied more extensively. But this extensive study

is another reason to choose the Lorenz equations; they are familiar and the available results will help us on our way. This consideration also motivates our choice of parameters. In the first six chapters we will concentrate on the parameter range $\sigma = 10$, $b = 8/3$, and $0 < r < \infty$. This parameter range includes the values first studied by Lorenz (1963) and provides a maximal intersection with other published work. In addition, the behaviour in this parameter range is qualitatively simpler than the behaviour observed for the different parameter values studied in Chapters 7 and 8.

We now commence our study of the Lorenz system by examining those results that are available using conventional, rigorous mathematics.

1.4. SIMPLE PROPERTIES OF THE LORENZ EQUATIONS

Much of the information in this section can be found in many papers. See, for example, Lorenz (1963) or Marsden (1977).

i) Symmetry

The Lorenz equations, (*), have a natural symmetry $(x, y, z) \rightarrow (-x, -y, z)$. This symmetry persists for all values of the parameters. I shall use the adjective "symmetric" to refer to those objects which are taken by the symmetry into themselves. Other objects will be referred to as "non-symmetric". Non-symmetric objects occur in pairs, of course, each one of the pair being taken into the other by the symmetry. Thus, if I sometimes use expressions like "for such and such a parameter value there is a non-symmetric, stable, periodic orbit," this should be taken as shorthand for the rather more cumbersome, "for such and such a parameter value there is a pair of non-symmetric, stable, periodic orbits, each of which is the image of the other under the natural symmetry."

ii) The z-axis

The z-axis, $x = y = 0$, is invariant. All trajectories which start on the z-axis remain on it and tend towards the origin $(0, 0, 0)$. Furthermore, all trajectories which rotate around the z-axis do so in a clockwise direction when viewed from above the plane $z = 0$. This follows from the fact that if $x = 0$ then $dx/dt > 0$ when $y > 0$ and $dx/dt < 0$ when $y < 0$. We can give a partial description of periodic orbits in the system by counting the number of times they wind around the z-axis. This description will not change as we alter the parameters, providing the same periodic orbit continues in existence.

iii) Existence of a bounded globally attracting set of zero volume

The divergence of the flow, $\partial\dot{x}/\partial x + \partial\dot{y}/\partial y + \partial\dot{z}/\partial z$, is $-(\sigma+b+1)$. Thus a volume element, V , is contracted by the flow into a volume element $V e^{-(\sigma+b+1)t}$ in a time t . (Notice that this simple observation already restricts the type of objects we expect to see in the Lorenz system. In Appendix A we define the stable objects ("sinks") to be those towards which all nearby trajectories tend, unstable objects ("sources") to be those from which all nearby trajectories diverge, and non-stable objects ("saddles") to be bounded invariant objects which are neither stable nor unstable. The dissipative nature of the flow implies that we cannot have unstable periodic orbits or unstable stationary points. Orbits and stationary points which are not stable must be non-stable; the existence of unstable objects would imply that there is, near the object, a volume that is expanded by the flow. In addition, we cannot have objects such as invariant tori in the flow; the volume within an invariant torus would not be contracting as required.)

We can show, as Lorenz (1963) did, that there is a bounded ellipsoid E in \mathbb{R}^3 which all trajectories eventually enter. The details of this calculation are in Appendix C. When the parameter r lies between 0 and 1 we can do even better. Consideration of the Liapunov function $V = x^2 + \sigma y^2 + \sigma z^2$ shows that the origin, $(0,0,0)$, is globally stable. For all parameter values, the existence of the bounded ellipsoid E and the negative divergence, taken together, show that there is a bounded set of zero volume within E towards which all trajectories tend. It is probably true, though not yet proved, that this set lies completely in $z \geq 0$ for all values of the parameters. In Appendix C we prove the result for $b < \sigma+1$; this certainly covers all the interesting cases considered in these notes.

iv) Stationary points

The origin is a stationary point for all parameter values. If $0 < r < 1$ we know that it is stable and globally attracting (see above). At $r = 1$ there is a bifurcation of a simple kind, and for $r > 1$ there are two other stationary points which we shall call C_1 and C_2 . The stationary points are $(\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$ and C_2 is the one lying in $x > 0$. The flow near the origin for r near one is shown schematically in Fig. 1.2.

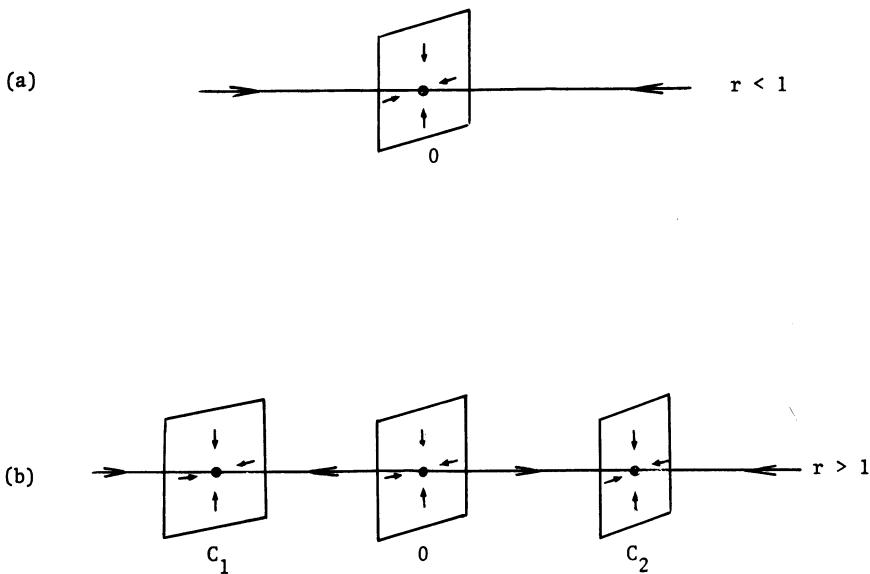


Figure 1.2. A schematic view of the flow near the origin for r near one.
(a) $r < 1$, (b) $r > 1$.

For $r > 1$ the origin is non-stable. The linearized flow near the origin has three real eigenvalues,

$$\lambda_1, \lambda_2 = \frac{1}{2}\{-\sigma-1 \pm ((\sigma-1)^2 + 4\sigma r)^{\frac{1}{2}}\} \quad \text{and} \quad \lambda_3 = -b.$$

λ_2 and λ_3 are negative, λ_1 is positive. The relative sizes of the eigenvalues will be important later. We have $-\lambda_2 > \lambda_1 > -\lambda_3$ providing that $r > 1 + b(\sigma+1+b)/\sigma$. For $\sigma = 10$ and $b = 8/3$ this condition is approximately $r > 4.644$.

The eigenvalues of the flow linearized near C_1 and C_2 are the roots of the equation $a^3 + a^2(\sigma+b+1) + ab(\sigma+r) + 2\sigma b(r-1) = 0$. The condition for all three roots to be real is complicated and not important. Suffice it to say that all three roots are real when r is very close to one but, when $\sigma = 10$ and $b = 8/3$, we have one real root and a complex conjugate pair of roots when $r > 1.346$. What is important is that if $r < \frac{\sigma(\sigma+b+3)}{(\sigma-b-1)}$, all three roots have negative real part. This implies that when $\sigma = 10$ and $b = 8/3$, the stationary points C_1 and C_2 are stable in the parameter range $1 < r < \frac{470}{19} \approx 24.74$.

Let us call the critical r -value (24.74..), r_H . When $r > r_H$ the complex roots of the equation above have positive real part and C_1 and C_2 are non-stable. The real root is negative for all r .

At $r = r_H$, as the complex eigenvalues cross the imaginary axis, there is a Hopf bifurcation (Marsden & McCracken, 1976) in which the points C_1 and C_2 lose their stability. The theory of Hopf bifurcation is now well advanced, and we know that there are two types. The bifurcation is "supercritical" if each point loses its stability by expelling a stable periodic orbit. It is "subcritical" if they lose their stability by absorbing a non-stable periodic orbit. In our case, we can show (Marsden & McCracken, 1976) that the bifurcation is subcritical. It appears that the bifurcation is likely to be subcritical for all σ and b values for which the bifurcation occurs in $r > 0$ (Hassard et al., 1981). See, also, Van Gib (1979). Marsden reports (private communication) that he and McCracken (1976) made numerical errors in their calculations which indicated supercritical bifurcation for some σ and b . The flow near C_1 for r near r_H is shown schematically in Fig. 1.3.

v) Summary

Let us summarize the results of this last section.

$0 < r < 1$ The origin is globally stable.

$1 < r$ The origin is non-stable. The flow linearized around the origin has two negative, and one positive, real eigenvalues.

$1 < r < 24.74..$ C_1 and C_2 are stable. All three eigenvalues of the flow, linearized about C_1 and C_2 , have negative real part. Providing $r > 1.346$ ($\sigma = 10$, $b = 8/3$) there is a complex conjugate pair of eigenvalues.

$24.74.. < r$ C_1 and C_2 are non-stable. The flow, linearized around C_1 and C_2 , has one negative real eigenvalue and a complex conjugate pair of eigenvalues with positive real part.

Notice that for $r > 24.74$ all three stationary points are non-stable. We are left with two immediate questions. These are:

1. Where do the two non-stable periodic orbits which are needed for the Hopf bifurcation at $r \approx 24.74$ come from? They must come into existence somewhere in the parameter range $r < 24.74$.

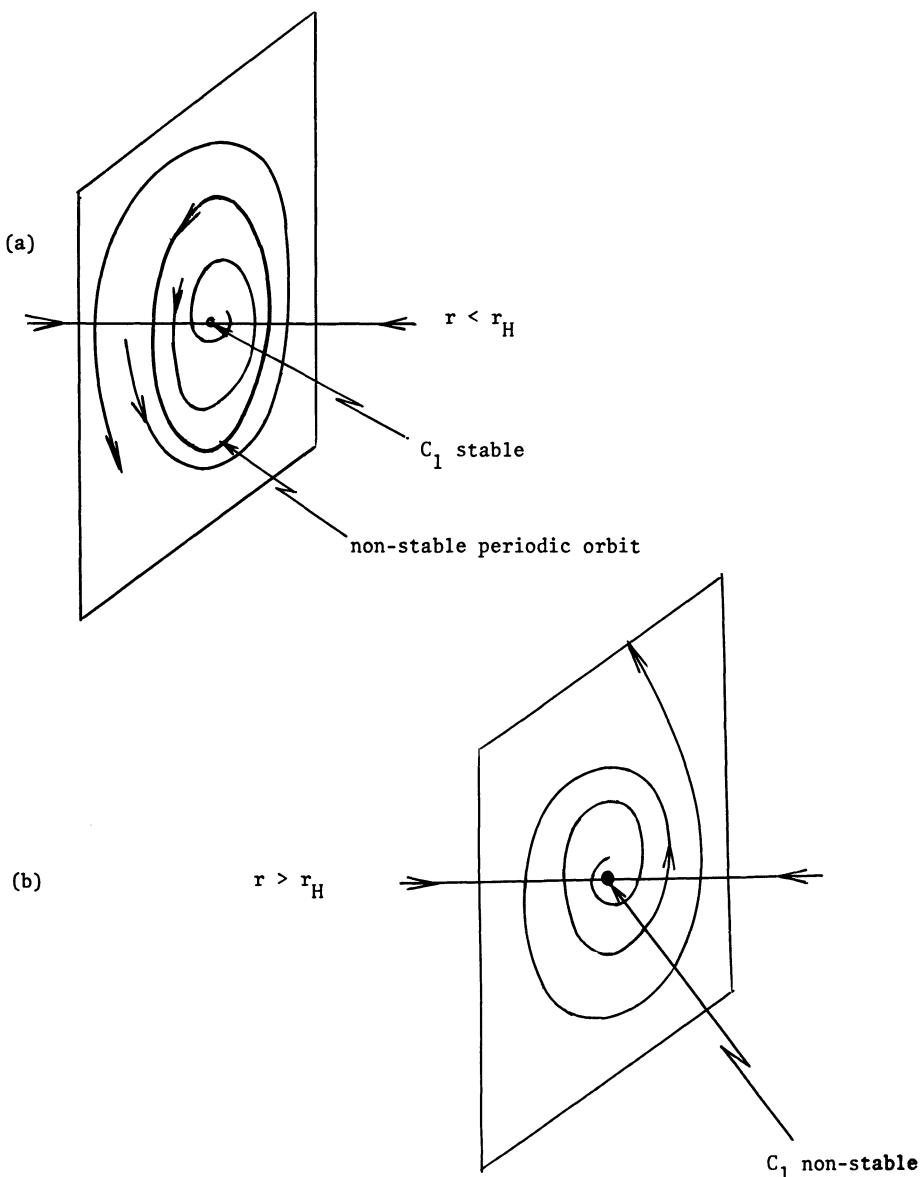


Figure 1.3. A schematic view of the flow near C_1 for r near r_H .
 (a) $r < r_H$, (b) $r > r_H$.

2. Where do most trajectories go when $r > 24.74$ and all three stationary points are non-stable?

We need the computer to help us answer these questions.

Chapter 2

Homoclinic Explosions: The First Homoclinic Explosion

2.1. EXISTENCE OF A HOMOCLINIC ORBIT

When $r > 1$ there is a two-dimensional sheet of initial values in \mathbb{R}^3 from which trajectories tend towards the origin. This two-dimensional sheet is called the stable manifold of the origin. Near the origin we know that this sheet looks like a plane (the plane associated with the two negative eigenvalues of the flow linearized near the origin) and if we wished we could approximate its position elsewhere by integrating the equations in backwards time with initial conditions lying on this plane and near to the origin. It appears that when r is only moderately larger than one, the stable manifold of the origin divides \mathbb{R}^3 into two halves in a fairly simple way. Trajectories started in one half-space tend towards C_1 and trajectories started in the other half-space tend towards C_2 . Trajectories started on the stable manifold of the origin tend, of course, towards the origin. (See Fig. 2.1.)

For a larger r -value, r' , approximately equal to 13.926..., it seems that the behaviour of the flow changes in an important way. As r increases towards r' , the spirals formed by the trajectories starting on the unstable manifold of the origin (which is a one-dimensional manifold associated with the positive eigenvalue at the origin, and on which trajectories tend in backwards time towards the origin) grow larger and larger. For $r > r'$ they "cross over" and are attracted to the "other" stationary point. (See Fig. 2.2.)

It is already difficult to imagine how the stable manifold of the origin must look. Trajectories cannot pierce this two-dimensional manifold, so it must now be twisted in a very strange way. We know that it

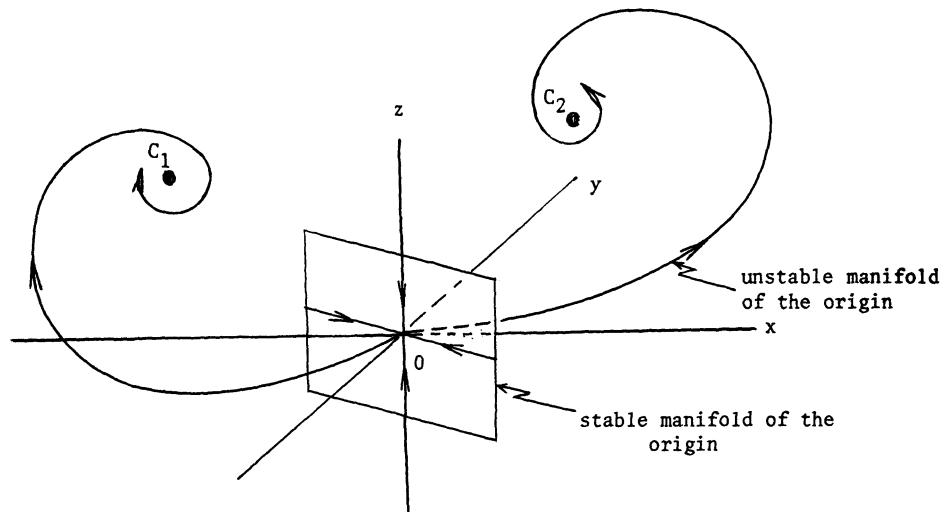


Figure 2.1. A schematic view of the flow in $1 < r < 13.926$.

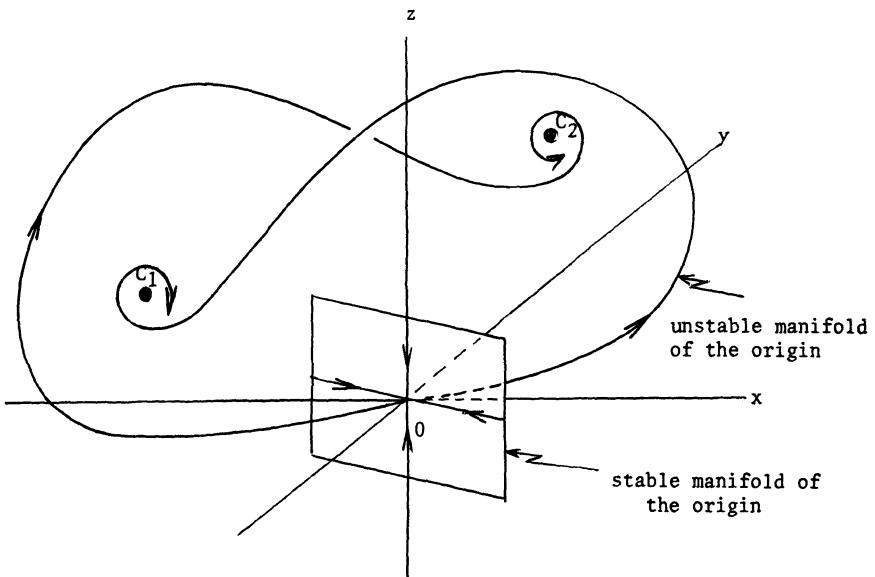


Figure 2.2. Behaviour of the unstable manifold of the origin for $r > 13.926$.

still looks flat near the origin, and that it includes the whole z -axis, but that is about all. We can understand this change in behaviour if we assume that at the critical r -value, r' , the stable manifold of the origin includes the unstable manifold of the origin; i.e., at $r = r'$, trajectories started on the unstable manifold of the origin will also lie in the stable manifold of the origin and will therefore tend, in both forwards and backwards time, towards the origin. In this situation we say that there is a homoclinic orbit associated with the stationary point at the origin. See the schematic Fig. 2.3(a).

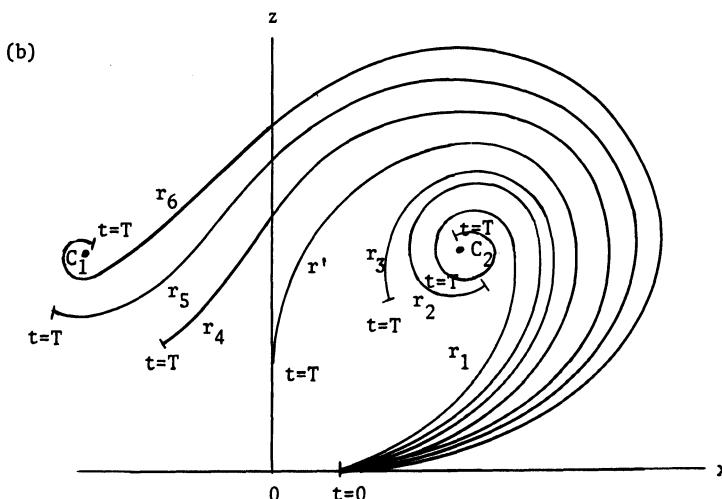
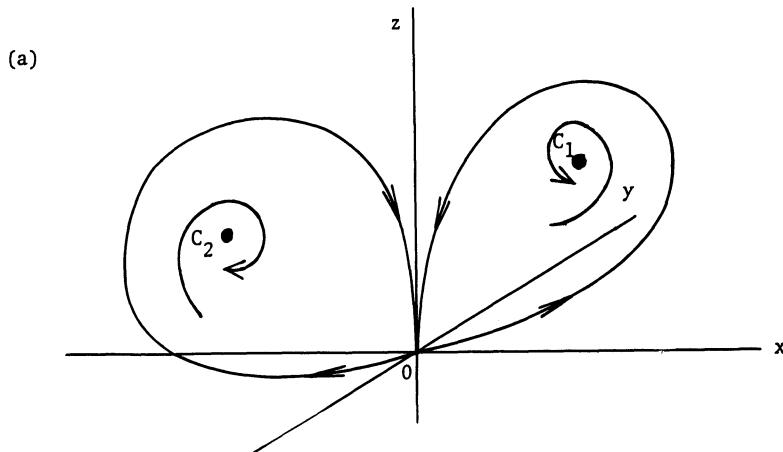


Figure 2.3. (a) Homoclinic orbits at $r = r'$. (b) Trajectories with the same initial conditions for different values of r near r' .

We could use some fairly sophisticated numerical techniques to attempt to confirm the existence of this homoclinic orbit (Hassard, 1980). We would need sophisticated techniques because trajectories calculated normally are only approximate, will inevitably jump off the stable manifold of the origin, and will then move away from it quite quickly. Fortunately, we do not need to do this since we can argue that there must be a homoclinic orbit for some r' in the following way. The behaviour of trajectories started at the same initial points will alter continuously with changing parameter. If a trajectory started at \underline{x}_1 reached \underline{x}_2 in a time t for one value of the parameter r , then trajectories started at \underline{x}_1 for nearby r -values will reach points near \underline{x}_2 in the same time t . Now, for $r < 13.926\dots$, trajectories started near the origin on its unstable manifold spiral quite rapidly towards one stationary point. For $r > 13.926\dots$, the "same" trajectory spirals quite rapidly towards the other stationary point. If we assume that there is a homoclinic orbit for some $r' \approx 13.926\dots$, we can see that this change in behaviour does not contradict the continuity property described above; for r -values very near r' , the trajectories will slow up near the origin and spend a lot of time there. Fig. 2.3(b) shows, schematically, the behaviour of the "same" trajectory, followed for the same lengths of time, for several different r -values satisfying $r_1 < r_2 < r_3 < r' < r_4 < r_5 < r_6$.

2.2. THE BIFURCATION ASSOCIATED WITH A HOMOCLINIC ORBIT

Numerical simulations for r just larger than $13.926\dots$ show nothing that we have not already described; all trajectories still seem to tend to either of the stationary points C_1 or C_2 rather rapidly. However, with the benefit of hindsight, we should look more carefully at the homoclinic orbit which occurs at $r = r'$. Our approach is quite general and will be relevant for any homoclinic orbits occurring in the Lorenz system.

What we do is look in a region of \mathbb{R}^3 very close to the homoclinic orbit, for values of r very near to r' , and ask whether there are any trajectories which remain forever within this region. There is of course the origin, and the homoclinic trajectory itself when $r = r'$, but we want to know if there are any more. There is a (mathematical) study of this question in Appendix D. See also the references given in Appendix D. The results depend crucially on the symmetry and on the relative magnitudes of the eigenvalues of the linearized flow near the origin. Other than that they are completely general.

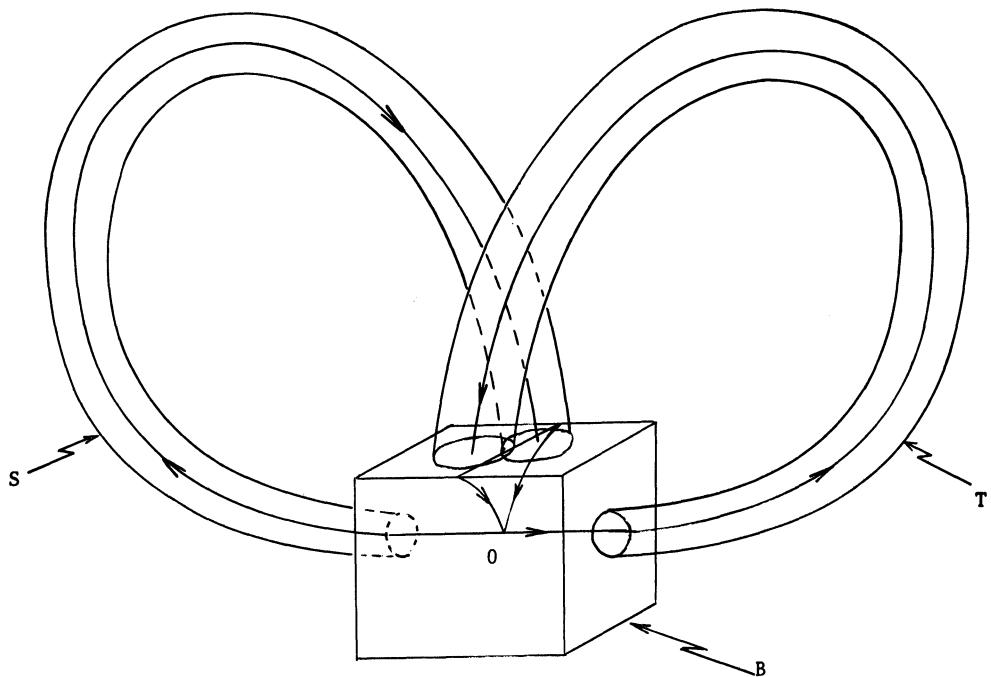


Figure 2.4. A small box B around the origin and two tubes T and S around the two branches of the unstable manifold of the origin.

What we do in Appendix D is to consider a small box B around the origin, together with two tubes, T and S , around the two branches of the unstable manifold of the origin. See Fig. 2.4. The box B is oriented so that the stable manifold of the origin intersects its top face in a line which divides this top face into two equal rectangles. We then calculate a return map on the top face of the box B ; i.e., we ask, for trajectories starting somewhere on the top face of the box, where do they next hit the top face of the box? This depends partly on what happens within the box B , which we can calculate using a linear analysis, and partly on what happens within the tubes. Fig. 2.5 illustrates what happens within the box B . Trajectories starting within the rectangle $BCEF$ on the top face of B will leave the box B through one of two thin sectors. Trajectories started to the right of AD (in $ABCD$) will leave through a thin section on the right-hand face of the box (shaded in Fig. 2.5), and

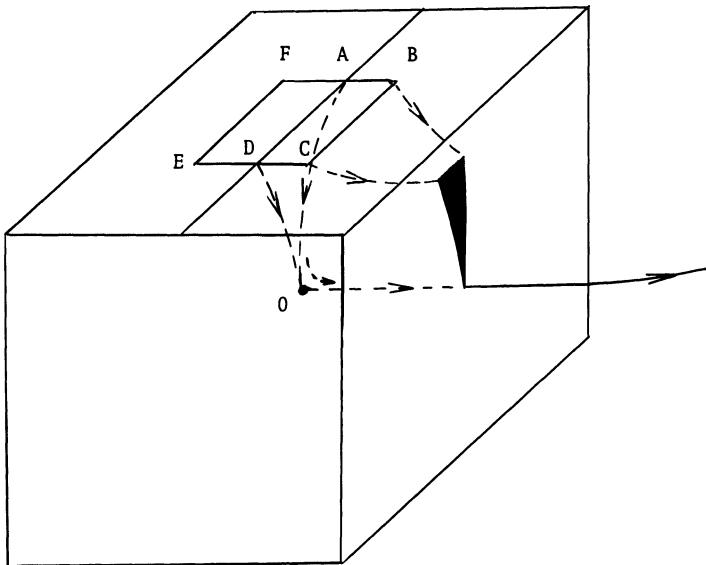


Figure 2.5. The flow within B is approximately linear. Trajectories started on the top face of B within ABCD leave B through the thin shaded region on the right-hand face of B. Trajectories started within ADEF leave B through a similar region on the left-hand face of B.

trajectories started to the left of AD (in ADEF) will leave through a thin sector on the left-hand face of the box. (Trajectories started on AD tend towards the origin.) The relative sizes of the eigenvalues determine the shape of these thin sectors; they will be longer and much thinner than the rectangles ABCD and ADEF that we started with. Each of the sectors has a vertex formed by the point at which the relevant branch of the unstable manifold of the origin leaves B.

We now need to know the effect the tubes T and S have on these thin sectors as we continue to follow trajectories after they leave the box B. Providing we have chosen B small enough, and chosen r close enough to r' , we can be sure that the shape of these sectors will be altered very little by what happens within T or S; the strong effects near the origin will more than compensate for any stretching or contracting that occurs elsewhere. (Though the tubes are "bigger" than the box, trajectories will actually spend most time in the box.) We only need to know where, and with what orientation, the tubes return the thin sectors to the top of the box B. Fig. 2.6 shows the six essentially different things that can happen.

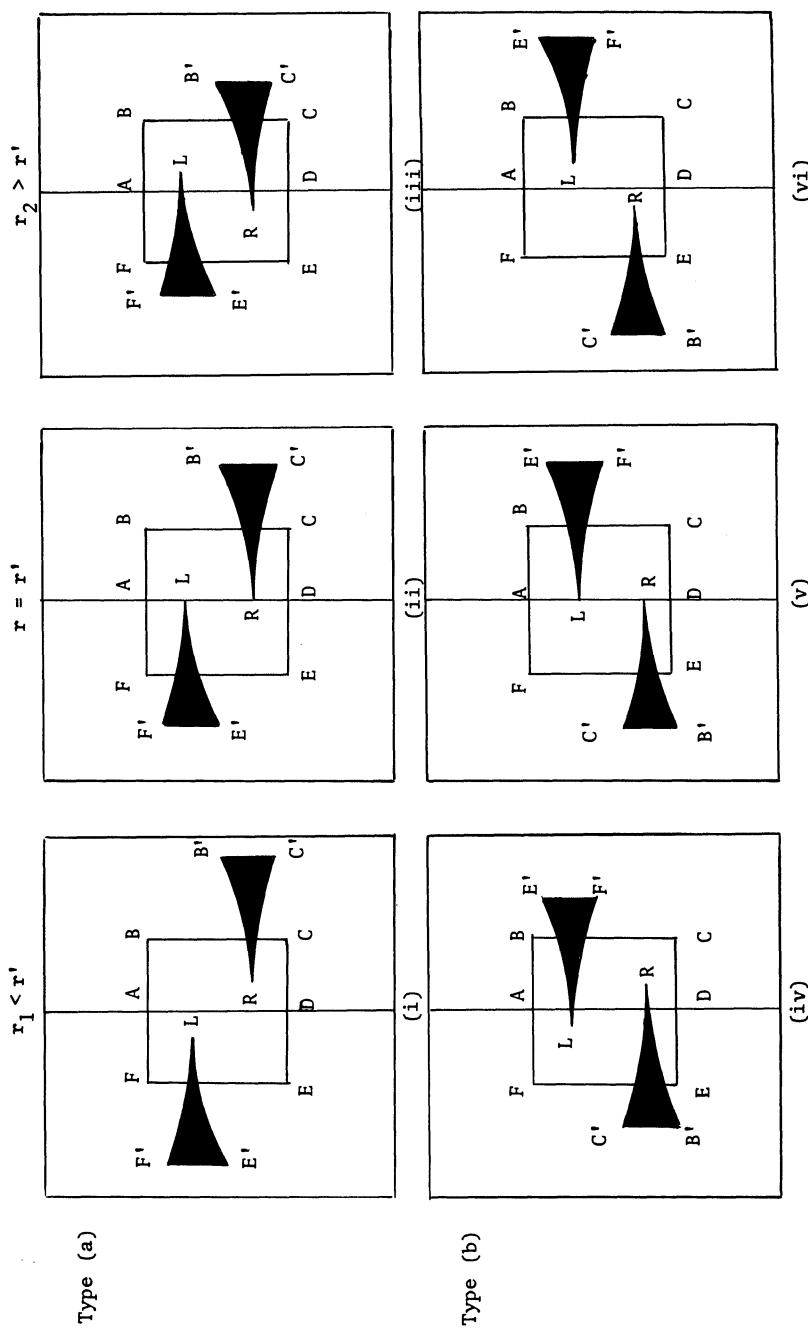


Figure 2.6. First return maps for the top face of B. Trajectories started within ABCD next strike the top face of B within the shaded area $RB'C'$; trajectories started within ADEF return to the shaded region $LE'F'$.

Each of Fig. 2.6 (i)-(vi) shows the top face of the box B and the first returns of a section of the face (the rectangle BCEF) under the flow. Trajectories started somewhere in BCEF next strike the top face of B somewhere within the shaded areas (Fig. 2.6). Trajectories started very near but to the right of AD (the stable manifold of the origin) return to points near R (which is the point where the right-hand branch of the unstable manifold of the origin first strikes the top face of B). Similarly, trajectories started very near but to the left of AD return to points near L. In all six diagrams we see that the images of the two rectangles ABCD and ADEF under the return map are strongly contracted in one direction (the direction parallel to AD) and stretched in another (the direction perpendicular to AD), and that they are similar in shape to the thin sections on the sides of B in Fig. 2.5.

The shaded areas have been drawn perpendicular to AD in all six diagrams. In fact, the shaded regions can be oriented at any angle (providing symmetry is preserved). The point is that the image of either ABCD or ADEF is a very thin sector (whose vertex is the point L or R); because it is very thin it can, be abuse of language, be said to stretch either towards or away from AD. This dichotomy matters, but the angle at which it stretches does not.

If, for r -values $r_1 < r'$, $r = r'$, $r_2 > r'$, we have the situation illustrated in Fig. 2.6(i)-(iii) we say the homoclinic orbit is type (a). If, on the other hand, we have the situation illustrated in Fig. 2.6(iv)-(vi) we say the homoclinic orbit is type (b). (Note that we have selected labellings $r_1 < r'$, $r = r'$ and $r_2 > r'$ in Figs. 2.6 as a result of a numerical experiment. When $r < 13.926\dots$, the right-hand branch of the unstable manifold of the origin spirals into C_2 and the point R lies to the right of AD. When $r > 13.926\dots$, it spirals in to C_1 and the point R lies to the left of AD. When we consider other homoclinic orbits we must remember that the bifurcations associated with moving from Fig. 2.6(i) to Fig. 2.6(iii) or from Fig. 2.6(iv) to Fig. 2.6(vi) can actually occur in either direction with increasing r .)

We must decide the type of the homoclinic orbit at $r' \approx 13.926\dots$. A simple numerical experiment should serve to convince us that it is type (a). For increasing values of r near the critical value r' we numerically calculate two trajectories. One, the right-hand branch of the unstable manifold of the origin, is calculated by starting a trajectory very close to the origin and displaced from it in the direction of the eigenvector associated with the positive eigenvalue of the linearized

flow around the origin. The other is another trajectory started just above (i.e., with $z > 0$) the first. Such a trajectory can be imagined to be starting on the top face of a small box B somewhere to the right of AD. For r somewhat less than r' both trajectories spiral in towards C_2 . For r somewhat greater than r' both spiral in towards C_1 . But, for some r -values between these two values (often located only with patience, particularly if the two trajectories are started very close together), we can see the right-hand branch of the unstable manifold of the origin tending towards C_1 , while the other trajectory tends towards C_2 . Thus the homoclinic orbit is type (a), since to explain our numerical result we must assume that there are r -values for which the point R lies to the left of AD, whilst some of the points in the thin sector stretching out from R lie to the right of AD; this only occurs in type (a). The numerical result could, of course, be misleading. But we shall see that there are other reasons to believe that this first homoclinic orbit is type (a) and so we shall concentrate on Fig. 2.6(i)-(iii).

It is fairly clear that in (i) any trajectory which starts to the right (say) of AD returns even further to the right. Thus, all trajectories (except those started on AD) eventually leave the small region we were considering (box plus tubes) and we have nothing interesting. In (ii) the same applies with the exception of the homoclinic orbit. Fig. 2.6(iii) is more complicated and has been analysed by several authors. See the references and the analysis in Appendix D.

The results of this analysis are that if we have a return map like Fig. 2.6(iii) or (iv), then there exists a "strange invariant set." This set consists of a countable infinity of periodic orbits, an uncountable infinity of aperiodic orbits, and an uncountable infinity of trajectories which terminate in the origin. All the orbits and trajectories remain forever within the small region we were studying (in both forwards and backwards time), and all intersect the top face of the box B within the shaded areas of Fig. 2.6. The set of points in which the strange invariant set intersects the top face of B looks like a Cantor set crossed with a Cantor set; the return map takes each of these points into another of these points, which is why we call the set "invariant". All of these orbits and trajectories are individually non-stable, as is the strange invariant set; almost all (in the sense of Lebesgue measure) trajectories still leave our small region of interest, and presumably spiral in towards either C_1 or C_2 . This explains why we do not see this "strange" addition to the non-wandering set when we do numerical experiments. On

the strange invariant set we have a phenomenon called "sensitive dependence on initial conditions." This means that almost all pairs of trajectories which are started close together (but both within the strange invariant set) eventually move far apart. See Appendix D for more details. For r -values very close to the homoclinic r -value, r' , trajectories and orbits in the strange invariant set pass very close to the origin. In particular, the periodic orbits in the strange invariant set have periods which tend to infinity as r tends to r' .

The fact that they never leave the tubes and box suggests a method for labelling the various trajectories which make up the strange invariant set. We can follow a trajectory (in both forwards and backwards time) and record the order in which it travels through the two tubes T and S . In fact (see Appendix D), there is a one-to-one correspondence between sequences, infinite to the left, of T 's and S 's, and trajectories in the invariant set. Repeating sequences correspond to periodic orbits, sequences which terminate correspond to trajectories which terminate in the origin, and aperiodic sequences correspond to trajectories which oscillate aperiodically through the tubes T and S forever. It should now be clear why we refer to the bifurcation associated with a homoclinic orbit as an "explosion". It seems worth distinguishing these bifurcations from the more mundane varieties, if only because it will make it easier to refer to them in the course of these notes.

2.3. SUMMARY AND SOME GENERAL DEFINITIONS

Since we shall see other homoclinic explosions as we proceed, let us gather together the information from this chapter which has general applicability and make some definitions that will make it easier to discuss them.

- 1) Note that the whole analysis of Appendix D, and the conclusions of this chapter, depend not at all on the shape of the tube T and its symmetric image S .
- 2) In the next chapter we will describe a global method for describing the "shape" of the tubes T and S . For the moment we need local definitions. When there is a homoclinic r -value at which the two branches of the unstable manifold of the origin can be surrounded by tubes of "shape" T and S (where S is the image under the symmetry of T), we say there is a T -homoclinic orbit. (Our choice of T rather than S in

this definition, and in the definitions which follow, is arbitrary. The symmetry allows us to specify the homoclinic orbit with only one tube, and the notation we have chosen is rather simpler than other choices which would mention both T and S .) Associated with this orbit there is a bifurcation which we call a T -homoclinic explosion. This explosion results in the production (or destruction, depending on the direction in which you are going) of a T -generated strange invariant set. This set contains one trajectory corresponding to every sequence, infinite on the left, of symbols T and S .

3) In particular, a T -generated invariant set will contain periodic orbits. These can be represented, if we wish, by the finite part of the sequence which repeats. The simplest periodic orbits will be T , S , TS , T^2S , S^2T , etc.. It is important to remember that these orbits actually correspond to infinite sequences; hence sequences TS and ST represent the same orbit, as do TS^2 , STS , and S^2T .

4) Because the Lorenz equations possess a symmetry, some of these periodic orbits will be symmetric. An orbit will be symmetric if some cyclic permutation of its finite symbolic representation replaces all the T 's with S 's, and vice versa. For example, ST and $S^2T^2ST^2S^2T$ are symmetric, whilst S^2T and S^2T^2ST are not.

5) There are two types of homoclinic explosion, type (a) and type (b). It may not seem that it was really worth distinguishing between the two types. After all, both give rise to the same T -generated strange invariant set. Notice, though, that the strange invariant set exists on different sides of the bifurcation, depending on the type. We shall find that it is crucial to have made this distinction.

6) There is another difference between the two types of explosion. This will turn out to be important later. For any particular r near r' , a trajectory started near the origin on the unstable manifold will, after a finite number of returns to the top face of our box B , leave the small region box plus tubes. The unstable manifold of the origin is not part of the T -generated strange invariant set. Nonetheless, we can, having selected the tubes T and S , move the parameter r nearer and nearer to r' , and force the unstable manifold of the origin to remain within the box plus tubes for as long as we like. In this sense, we can write down an infinite sequence of T 's and S 's which represents the behaviour of the unstable manifold of the origin as r approaches r' from above and below. Because of the symmetry, we can look at either branch

of the manifold; we choose to examine the right-hand branch (the one which leaves the origin into $x > 0$ and which strikes the top face of B at R), and define a sequence $k(r)$ of T's and S's which describes the behaviour of this branch for the period during which it remains within our small region of interest. It is possible to deduce the way in which $k(r)$ changes as we pass through either a type (a) or a type (b) explosion by looking at Figs. 2.6. However, it is a little easier to see these changes using the technique described at the end of Appendix D; we give the result below.

7) We can illustrate the difference between type (a) and type (b) explosions as follows:

$$\begin{array}{ccc} \text{Type (a)} & k(r) = \text{TTTT...} \\ & \text{No invariant set} \end{array} \quad \iff \quad \begin{array}{c} \left\{ \begin{array}{l} k(r) = \text{TSSSS...} \\ \text{Invariant set} \end{array} \right. \end{array}$$

$$\begin{array}{ccc} \text{Type (b)} & k(r) = \text{T TS TS TS...} \\ & \text{Invariant set} \end{array} \quad \iff \quad \begin{array}{c} \left\{ \begin{array}{l} k(r) = \text{TS TS TS TS...} \\ \text{No invariant set} \end{array} \right. \end{array}$$

Notice that to tell type (a) from type (b) on the basis of $k(r)$ alone, we need to follow the unstable manifold of the origin around the tubes at least three times. If we follow it only twice then the changes appear identical ($\text{TT} \iff \text{TS}$).

8) We have already said that numerical tests indicate that the homoclinic explosion at $r' \approx 13.926\dots$ is type (a). The change in behaviour of the unstable manifold of the origin on either side of this explosion (see Figs. 2.1 and 2.2) corresponds intuitively with the change in $k(r)$ that we expect for type (a). Furthermore, we had hoped that the flow was simple to understand when $r < 13.926\dots$; this too corresponds with what we expect with a type (a) explosion. If the explosion had been type (b), we would have expected it to destroy, rather than create, a strange invariant set as r increased, and would have had to explain how this set appeared somewhere in $r < r'$.

9) Convinced at last that the homoclinic explosion at $r' \approx 13.926$ really is type (a), we believe that for r -values just larger than r' there will be a strange invariant set, as described above, added to our non-wandering set. We shall call this set the *original strange invariant set* (because it is the first one), and the periodic orbits in it we shall call *original periodic orbits*. Providing that nothing much happens to this original invariant set as r increases (and we should remember our

analysis is only local), we have possible answers to the two questions posed at the end of the last chapter.

1. The simplest orbits, T and S, of the strange invariant set may be the orbits which are involved in the Hopf bifurcation at $r \approx 24.74$. These two are the only candidates; they do not wind around the z-axis. All other periodic orbits in the original invariant set do wind around the z-axis which, being invariant, does not allow them to unlink themselves from around it.
2. We have the original strange invariant set as a candidate for an attractor when all three stationary points are non-stable.

We now need to know whether these answers can be supported by numerical evidence.

Chapter 3

Preturbulence, Strange Attractors and Geometric Models

3.1. PERIODIC ORBITS FOR THE HOPF BIFURCATION

We commence our study of the parameter region $r > 13.926$ by attempting to confirm the first of the two suggestions made at the end of the last chapter. We can use the computer to locate stable and non-stable periodic orbits. The techniques used are based on Newton's method, are well known, and are described in Appendix E. To use the techniques it is necessary to have a reasonably good guess for the position and period of the orbit we wish to locate. Appendix E contains practical hints on how to come by such initial guesses, and on how to use the techniques. Appendix E also contains the description of a technique for following a periodic orbit with changing parameter.

Fig. 3.1 shows a non-stable periodic orbit at r -values 14.5, 20.0 and 24.5. At each r -value we have only followed the orbit around once; if the numerical integration is allowed to proceed, the trajectory wanders away from the periodic orbit very rapidly. We can be sure (numerically speaking) that it is the "same" periodic orbit which is shown in all three pictures.

The orbit shown in Fig. 3.1 is one of the two simplest periodic orbits belonging to the strange invariant set (and the first numerical evidence that we have for the existence of the strange invariant set). At $r = 14.5$ it passes quite close to the origin and its period is relatively large. In fact, it is impractical to follow the orbit into the parameter range $r < 14.5$, since the period increases very rapidly as r decreases. This is what we would expect if it is true, as we suppose, that the orbit was born in a homoclinic explosion (at which its period is infinite) for some $r' \approx 13.926$. At $r = 24.5$, the period of the orbit is relatively

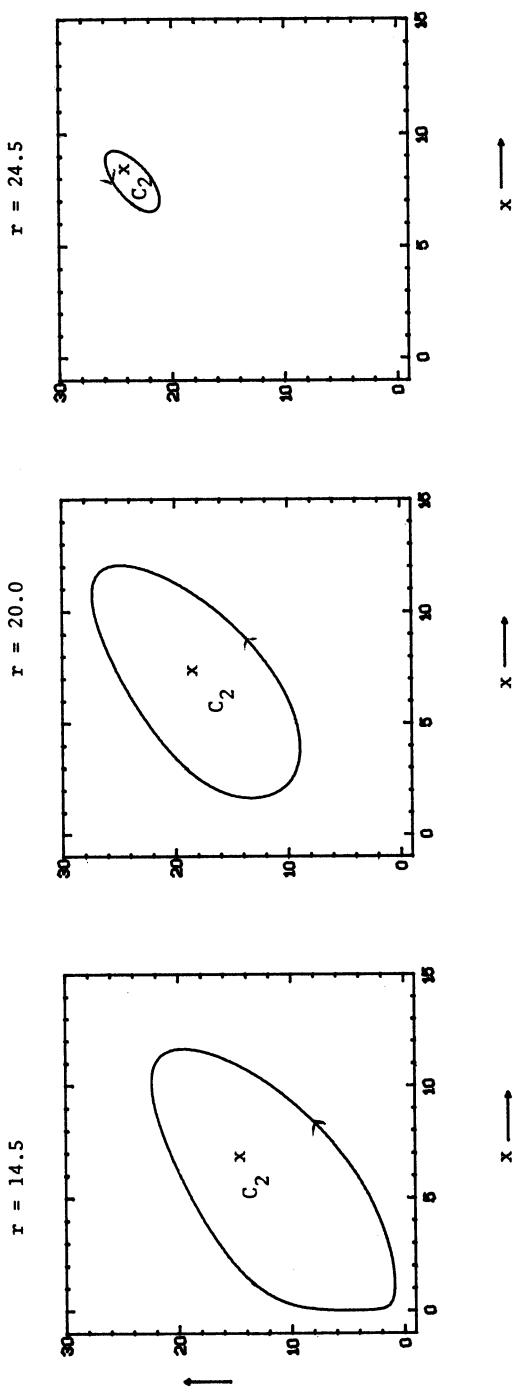


Figure 3.1. The same non-stable periodic orbit at three r -values in $13.926 < r < 24.74$.

short, it passes nowhere near the origin, and it seems extremely likely that this is the very orbit which is destined to be involved in the Hopf bifurcation at $r \approx 24.74$. If we continue to follow the orbit as r increases to 24.74, it continues to shrink towards the stationary point C_2 (and becomes too small to draw). We henceforth assume that we are correct in asserting that the two simplest orbits (T and S) born in the first homoclinic explosion ($r' \approx 13.926$) are the orbits involved in the Hopf bifurcation.

We can use the techniques of Appendix E to locate many of the shorter periodic orbits which we expect to find in the invariant set. However, since the presentation of any finite number of periodic orbits will still not convince us that there is actually an infinite number, and since we will have a lot to say about periodic orbits later in these notes, we will not describe these experiments here.

3.2. PRETURBULENCE AND RETURN MAPS

If we are to confirm the second suggestion at the end of the last chapter, we must first convince ourselves that the strange invariant set continues to exist in $r > 13.926$. The analysis in which we demonstrated its existence was only local and only valid for r -values near to r' (≈ 13.926). In Chapter 2 we argued the existence of the strange invariant set by calculating (analytically) a return map on a plane (the top face of the box B) very near to the origin. When r is larger than 13.926 we will have to compute return maps numerically, and therefore approximately. We will also have to compute them on planes far from the origin. The latter is no hardship. As r increases from 13.926, the strange invariant set may continue to exist topologically unchanged, but we certainly expect it to move in phase space; for r -values larger than 13.926 we no longer expect all interesting trajectories to pass close to the origin (see, for example, Fig. 3.1). The plane on which we choose to compute return maps is the plane $z = r - 1$. This plane contains the stationary points C_1 and C_2 , and it seems that all interesting trajectories intersect the plane for a wide range of r -values. Before proceeding, we should consider our attitude to numerically computed return maps.

When we compute a return map on a suitable return surface, we are no longer computing the long term behaviour of a solution. In general, return times - the times taken for a trajectory to return to the return surface - are short, and the return of each point can be computed with a

great deal of accuracy. (When return times are long, it is generally because the point we have selected lies close to the intersection of the return surface with the stable manifold of the origin; in this case the trajectory passes very close to the origin and analytical methods, like those used in Chapter 2 and Appendix D, may be used.) We cannot, of course, compute complete return maps. At most we can compute the returns of a finite number of points, and each of these computations will only be definitely accurate. After computing a sufficient number of returns, we will be able to make conjectures about the properties of the return map; e.g., we will be able to make conjectural statements like, "It appears that the return maps computed on the return plane $z = r-1$, for parameter values r in $r > r'$, are topologically identical to the return map of Fig. 2.6(iii)." Such statements will always be subject to numerical disproof; the calculation of just one more return may upset our picture completely. However, once we have made conjectures about the return map we may be able to say something interesting about the behaviour of the flow which we could not have said after just observing the long term behaviour of one or two trajectories. Notice that, even if we assume our conjectures about the return map are correct, we will not necessarily understand the behaviour of the flow; though return maps like Fig. 2.6(iii) are well understood, others are not.

Fig. 3.2 shows, schematically, the return map on the plane $z = r-1$ for an r -value near 22. As before, the rectangle ABCD is mapped into the thin sector RB'C', and the rectangle ADEF is mapped into the thin sector LF'E'. R and L are the points where the right and left-hand branches of the unstable manifold of the origin first intersect the return plane with $\dot{z} < 0$. (As in Chapter 2, we only consider intersections with the return plane for trajectories travelling downwards.) We choose to show a schematic return map in Fig. 3.2 because the numerically calculated return maps (which just consist of a collection of points) are hard to interpret. The major difference between Fig. 3.2 and a numerically computed map appears to be the width of the return areas. We have exaggerated the width in Fig. 3.2; numerically speaking, no width is discernable. Providing this return map still stretches distances "perpendicular" to AD, and providing it satisfies some special conditions that we will discuss later, it is topologically equivalent to Fig. 2.6(iii) and the same arguments used in Chapter 2 and Appendix D can be employed to argue the existence of the same strange invariant set. We cannot check the stretching condition everywhere, but it does appear to be true at any point we choose to examine.

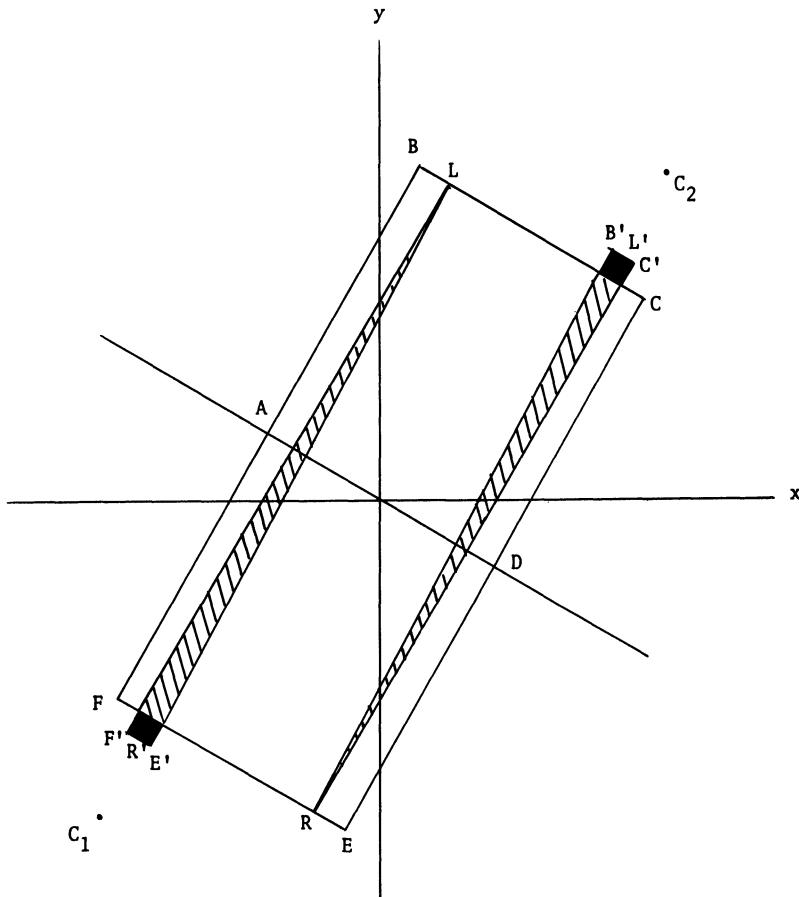


Figure 3.2. A schematic version of the return map on the plane $z = r-1$ for an r -value near 22.

We still expect almost all trajectories to leave the area BCEF eventually, since the strange invariant set is still non-stable. However, they will only do so after they have fallen into one of the small regions lying between $F'E'$ and FE or between $B'C'$ and BC . Almost all trajectories will fall into one of these regions eventually because of the mixing nature of the flow near the strange invariant set (see Appendix

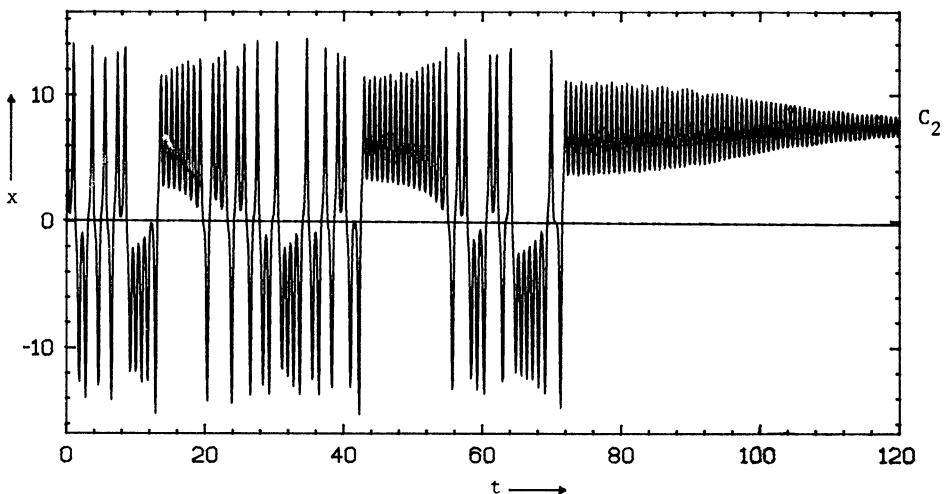


Figure 3.3. A typical "preturbulent" trajectory wanders chaotically near the strange invariant set before spiralling into C_2 . ($r = 22.4$.)

D), but, if the regions are very small, some trajectories may take a long time to escape from BCEF. As r increases, these "escape" regions seem to get smaller and smaller, and we can observe numerically calculated trajectories which wander near the strange invariant set for long periods of time. Fig. 3.3 shows a typical trajectory (the variable x is plotted against time) for $r = 22.4$. Though the trajectory eventually spirals into C_2 , it wanders chaotically near the strange invariant set for a long time.

The phenomenon seen in Fig. 3.3 is known as "preturbulence" (after Kaplan & Yorke, 1979a) or as "meta-stable chaos" (after Yorke & Yorke, 1979). These authors show that the average time spent wandering, by a trajectory which wanders at all, tends to infinity as r increases towards an r -value, r_A , which is approximately equal to 24.06. This corresponds to the observation that the small "escape" regions of Fig. 3.2 eventually shrink to nothing at $r = r_A$. At this r -value, the area BCEF is mapped into itself and if you start within BCEF you must remain there forever; the strange invariant set has become a strange attractor. (Note: We have some choice when selecting the region BCEF to consider. For $r < r_A$ return maps will look like Fig. 3.2 whatever choice, within reason, we make. When $r = r_A$ there will be some choices of region BCEF,

containing the whole invariant set, for which it is mapped into itself.)

Not all trajectories will show preturbulent behaviour in $r < r_A$. Trajectories started near enough to C_1 or C_2 will tend to these stationary points without wandering near the strange invariant set. (Notice that C_1 and C_2 are still stable at $r = r_A$. r_A is a lower r -value than r_H - the Hopf bifurcation value - and there will still be some trajectories attracted to C_1 and C_2 even after we have the strange attractor.) Trajectories started near the origin do not show preturbulent behaviour either. These trajectories strike the return plane near points R and L (Fig. 3.2) and then, on the next return, enter the small "escape" regions at once. This observation tells us how to calculate the critical r -value r_A (since trying to decide whether the mean duration of preturbulent behaviour has become infinite is not a satisfactory experimental technique). If we compute as close an approximation as possible to one of the branches of the unstable manifold of the origin, the critical r -value is the one for which this trajectory stops spiralling immediately into the appropriate stationary point.

We must now examine what happens in $r > r_A$.

3.3. STRANGE ATTRACTOR AND HOMOCLINIC EXPLOSIONS

We now consider the development of the strange attractor in $r > r_A$. Though it has been pointed out many times by many authors (e.g., Kaplan & Yorke, 1979b), it is worth emphasizing once more that r_A (≈ 24.06) is less than r_H (≈ 24.74). McLaughlin & Martin (1975) were wrong to assert that it is the subcritical Hopf bifurcation which results in an immediate transition to chaos. We begin our study of the strange attractor by considering the parameter value $r = 28.0$. This parameter value is a favorite of other authors and is the value first studied by Lorenz (1963). There is, however, nothing special about it.

Fig. 1.1 showed a typical "chaotic" trajectory for $r = 28.0$. Fig. 3.4(a) shows the downwards intersections of a rather longer trajectory (of which Fig. 1.1 was part) with the plane $z = r-1$. This longer trajectory was started near the origin on the unstable manifold of the origin (approximately) and we have labelled the point R where this trajectory first passed downwards through the plane. The point R' , the first return of R, is also labelled. Half the points in Fig. 3.4(a), including L and L' , were obtained by taking the symmetric images of each of the points of intersection already computed. To help us locate

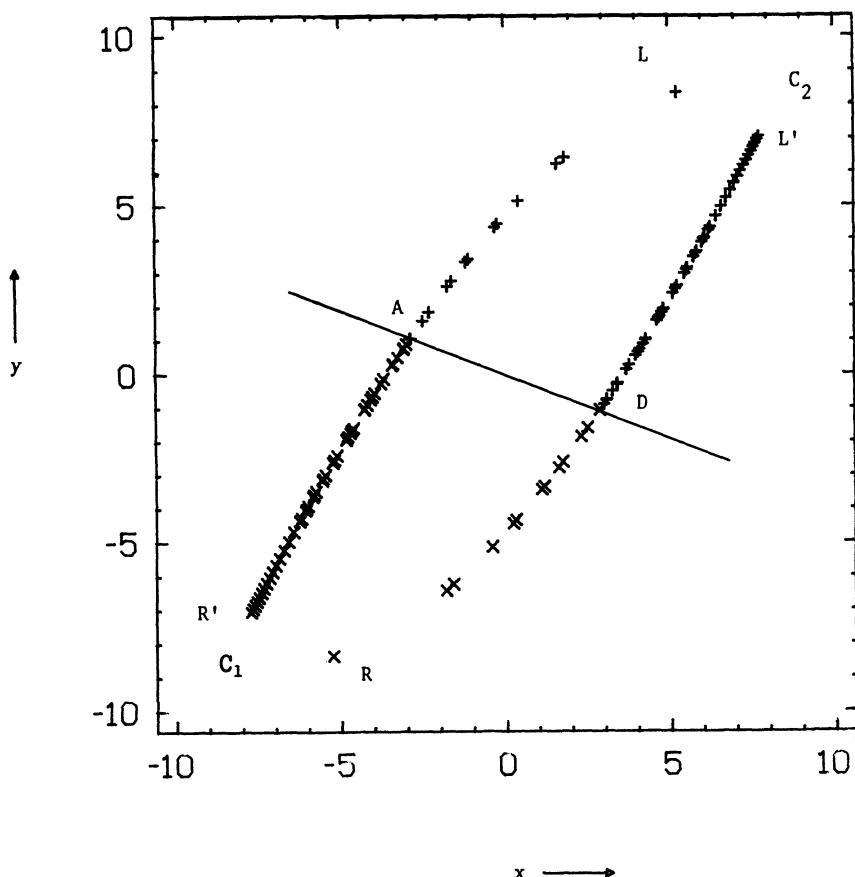


Figure 3.4. (a) Intersections of a numerically computed trajectory with the plane $z = r-1$. ($r = 28$.)

the stable manifold of the origin we have marked points with either an "x" or a "+", depending on whether the trajectory, after passing through the point, spirals to the left or to the right. The intersection of the stable manifold of the origin with the return plane, $z = r-1$, must include a line segment such as AD which separates the x's and +'s. (This line segment need not be straight, but that is more or less irrelevant in what follows.)

Notice that we do not (in Fig. 3.4(a)) see many points of intersection near R or L. This is not surprising since, though we actually expect the attractor to intersect the plane in points stretching all the

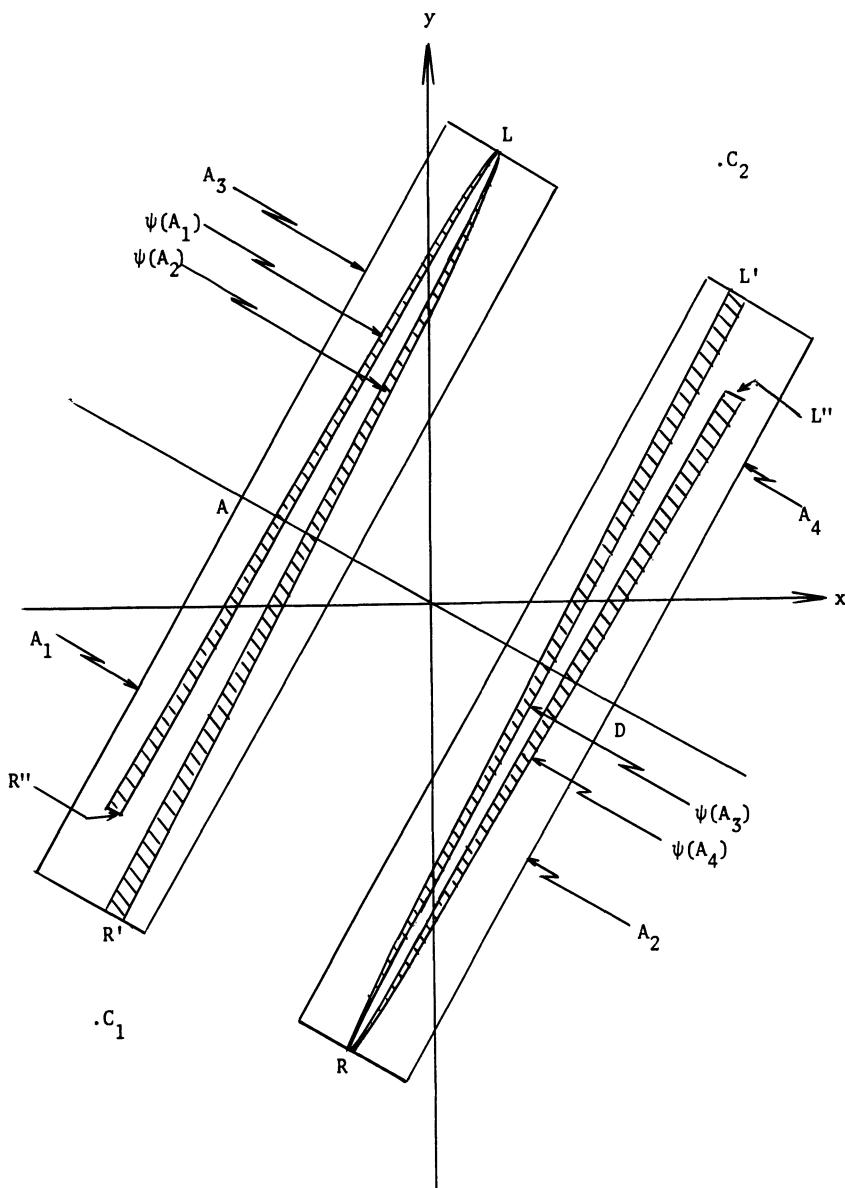


Figure 3.4. (b) Schematic view of the return map on the plane $z = r - 1$. ($r = 28.$)

way from R to L' and from L to R' , trajectories will only intersect the plane near R or L after passing very close to the origin. This is a very unlikely event as we explain in Appendix F. We would have to run a trajectory for a very long time before seeing even a few points of intersection near to R or L , and even if we can afford to run a trajectory indefinitely, there are still quite severe computational limitations on how close to the origin the trajectory can go. (Note: this same consideration implies that in the preturbulent parameter range, numerically calculated trajectories may actually wander near the non-stable strange invariant set longer than true trajectories.)

It is striking that the points of intersection of the trajectory with the plane seem to lie on two arcs, RL' and LR' , with no noticeable thickness. We will return to this point later; for the moment we should remember that there must be some thickness, and Fig. 3.4(b) shows a schematic return map in which this thickness is exaggerated. The four rectangles labelled A_1 to A_4 are mapped into longer and thinner areas labelled $\psi(A_1)$ to $\psi(A_4)$.

One way to visualize the return map is to imagine that points lying to the left of AD are under the influence of the stationary point C_1 , and that points lying to the right of AD are under the influence of C_2 . Each of these stationary points has a two-dimensional unstable manifold on which trajectories spiral outwards. The unstable manifold of C_1 intersects the plane near the arc $C_1R'L$ and the unstable manifold of C_2 intersects the plane near the arc $C_2L'R$. A trajectory starting at a point on the return plane will spiral outwards away from the relevant stationary point, at the same time moving rapidly towards the unstable manifold of that stationary point (recall that there is a negative real eigenvalue - of relatively large magnitude - associated with the linearized flow near each stationary point). Thus, for example, trajectories started to the left of AD will next strike the plane near to $C_1R'L$ and further from C_1 . This visualization of the return map is slightly misleading because it ignores the important effects near the origin. These can be incorporated into our understanding if we note that without the strong contracting effect near the origin, trajectories started very near to AD (and roughly equidistant from C_1 and C_2) would not know which of the two unstable manifolds to be attracted towards, and these same trajectories would not be expected to arrive back at the return plane quite so far from the relevant stationary point as they do. See Lanford (1976) for an excellent description of the attractor at $r = 28.0$.

Fig. 3.4(b) is different from the return maps that we have studied so far. Most importantly, we now have an attractor within the four rectangles, since the union of the four rectangles is mapped into itself. We will start by describing some aspects of this attractor in an intuitive way; in Section 3.4 we will discuss the special conditions that must hold if the attractor is to be as we are about to describe it.

We assume that the return map still stretches all distances "perpendicular" to AD. This implies that we have "sensitive dependence on initial conditions", since almost all pairs of trajectories started close together must diverge. (The exceptional pairs will be those trajectories started at points "the same distance" from AD.) This explains the chaotic appearance of Fig. 1.1, and implies that we have no stable periodic orbits within the attractor.

The attractor will have a Cantor set-like structure. In Fig. 3.4(b), the two long rectangles A_1 plus A_3 and A_2 plus A_4 are mapped into four long areas. The next application of the return map takes these four areas into eight long areas. In the limit, as we apply the return map over and over again, we will find that the attracting set intersects the return plane in a Cantor set of arcs. Thinking in terms of the flow (rather than in terms of the return map), the attractor will consist of a Cantor set of two-dimensional sheets (each of which intersects the return plane in an arc). We call it a "strange attractor". If we consider a general point on the attractor, it will lie on a two-dimensional sheet of points, all of which lie on the attractor, and nearby there will be other similar sheets, "parallel" to the one we are considering. Some points, such as the point R on Fig. 3.4(b), are special. At these points the attractor looks like a Cantor "book" or "fan". The point R lies on the spine of the book, and a Cantor set of sheets ("pages") radiates out from this spine. The special points all lie on the unstable manifold of the origin (as R does), and the special structure is due to the contraction which occurs as trajectories pass close to the origin. Notice that the origin is actually included in the attractor; Fig. 3.5 shows how the sheets come together at the origin.

In some sense the special points lying on the unstable manifold of the origin form the edge of the attractor. Though this is obvious in the case of the point R or in Fig. 3.5, it is not so obvious for all points lying on the unstable manifold of the origin; the unstable manifold of the origin intersects the return plane in many (possibly infinitely many) points in the interior of the four rectangles A_1 - A_4 .

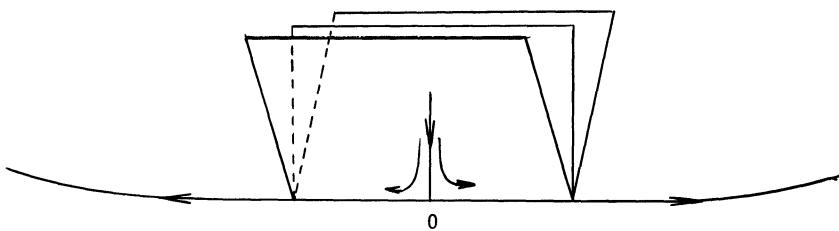


Figure 3.5. The sheets of the attractor are "pinched" together into a Cantor "book" near the origin.

Arguments similar to those used in studying the non-stable strange invariant set which existed in $r < r_A$ will show that the attractor contains, as before, a countable infinity of non-stable periodic orbits, an uncountable infinity of aperiodic trajectories, and an uncountable infinity of trajectories which terminate in the origin. We can still describe orbits and trajectories within the strange attractor with sequences of two symbols. We record the order in which a trajectory intersects the two halves of the return plane generated by the arc AD, and write an "x" every time it intersects the return plane to the right of AD, and a "y" every time it intersects the return plane to the left of AD. Thus we obtain an "x" every time a trajectory spirals around C_2 and a "y" every time it spirals around C_1 . This method of describing trajectories is global, unlike the local descriptions (T and S) used near a homoclinic explosion. We can, of course, use this global method to describe the tubes T and S of our local analysis. We emphasized, in Chapter 2, that our analysis of homoclinic explosions did not depend on the shape of the tubes T and S. The first homoclinic explosion (at $r \approx 13.926$) had rather simple tubes T and S which have global descriptions "x" and "y" respectively. Thus, we may say that the first homoclinic explosion was an x-homoclinic explosion, and that the original strange invariant set which it generated was an x-generated strange invariant set. (Once again, we emphasize that our use of "x" rather than "y" to describe the explosion and set is arbitrary. The symmetry of the system permits us to describe the explosion and invariant set accurately by specifying only one branch of the homoclinic orbit with which both are associated.)

The difference between the original invariant set and the strange attractor at $r = 28.0$ is not only a question of stability. We see this as follows:

1. The simplest orbits, x and y , are not part of the attracting set when $r = 28.0$. We expect this because they have been removed in the Hopf bifurcation, but the argument applies for any $r > r_A$ as well. If the x and y orbits were part of the attractor, they would have to intersect the return plane in the rectangles A_4 and A_1 respectively (Fig. 3.4(b)). These are the only two rectangles which map, at least partly, into themselves. But a glance at Fig. 3.4(b) shows that points in these two rectangles are always taken by the return map to points closer to AD or to points on the other side of AD . Thus, there can be no fixed points of the return map within these areas, and the x and y orbits are not part of the attractor. (At $r = r_A$, these orbits are, momentarily, part of the attractor. At this r -value, the region $\psi(A_1)$ stretches all the way to the end of area A_1 . The point R' is mapped neither towards C_1 - as in $r < r_A$ - nor towards AD - as in $r > r_A$ - but tends instead, under successive applications of the return map, towards a point near R' where the simple y orbit intersects the return plane.)
2. Though an analysis of Fig. 3.4(b) will, as usual, show that there is at most one orbit or trajectory corresponding to each sequence, infinite on the left, of symbols "x" and "y", we will no longer be able to conclude that there is a one-to-one correspondence. The original invariant set contained one trajectory corresponding to every sequence. But now, the number of consecutive x 's or y 's is limited. Consider the trajectory which starts near the origin and intersects the return plane first at R , then at R' ($= \psi(R)$), then at $\psi^2(R)$, etc.. As the integer i increases, $\psi^i(R)$ will get closer and closer to AD , and for some finite i^* , $\psi^{i^*}(R)$ will lie in rectangle A_3 on the other side of AD (with subsequent iterates wandering, where they will, over the attractor). This trajectory only managed i^* straight y 's in a row before generating an x ; it is easy to see that no other trajectory (contained within the invariant set) can manage more. When $r = 28.0$, $i^* = 24$, and so we know that all the periodic orbits and trajectories with 25 or more consecutive x 's or y 's in their symbolic representations are missing from the invariant set. Where have they gone?

Periodic orbits, and some aperiodic trajectories, are removed from the attractor in homoclinic explosions. Let us look, as an example, at the periodic orbit xy^{26} . This certainly existed in the original invariant set, and no longer exists when $r = 28.0$.

As r increases from r_A to 28, the integer i^* which determines the maximum number of revolutions which can occur consecutively on one side decreases. The point R' moves closer and closer to AD and the number of iterates required to map R' onto the right-hand side of AD decreases. (At $r = r_A$, i^* is infinite and R' is never mapped to the right of AD .) At the parameter value, below 28.0, where the integer i^* changes from 26 to 25, $\psi^{26}(R)$ lies exactly on AD , the stable manifold of the origin. This implies we have a homoclinic orbit. The bifurcation associated with this homoclinic orbit can be analyzed in exactly the same way as we analyzed the first homoclinic explosion at $r \approx 13.926$, except now the two tubes T and S are much more complicated. (They surround an xy^{26} and a yx^{26} homoclinic orbit respectively.) We can, nonetheless, describe all the periodic orbits and trajectories removed (or produced) in this homoclinic explosion with sequences of the two symbols T and S . Once we have done this, we can substitute $T = xy^{26}$ and $S = yx^{26}$ to obtain the global descriptions of the orbits and trajectories. A complete xy^{26} -generated invariant set will be involved; the simplest periodic orbits in this set will be $xy^{26} = T$, $yx^{26} = S$, $x^{27}y^{27} = xy^{26}yx^{26} = TS$, etc.. (Note that while we can quite properly describe periodic orbits with any cyclic permutation of some finite symbolic description, there is only one proper finite symbolic description of the homoclinic orbit. This is because the former are actually represented by doubly infinite repeating symbolic sequences, and the latter are actually represented by finite symbolic sequences.)

We can check numerically that the xy^{26} -homoclinic explosion is a type (a) explosion proceeding in the direction which removes an invariant set. (We do this the same way that we checked that the first homoclinic explosion was type (a) - see Chapter 2.) Actually, all homoclinic explosions in this parameter range will be type (a). Recall (from Chapter 2) that a homoclinic explosion is type (a) if a trajectory near the right-hand branch of the unstable manifold of the origin (which trajectory can be assumed to be starting on the return plane just to the right of AD), after following the unstable manifold around on its homoclinic course, finally returns to the return plane to the right of AD once more. Since our return map preserves orientation in the direction "perpendicular" to AD , this will always be the case.

The integer i^* gives only a very crude measure of the size of the attractor. It changes each time we have an xy^n -explosion for some n . However, there will be many other homoclinic explosions besides the xy^n

ones. It is the sequence $k(r)$ - which we defined in Chapter 2 to be the symbolic sequence which describes the behaviour of the right-hand branch of the unstable manifold of the origin - which determines the size of the attractor accurately (corresponding with our earlier observation that the unstable manifold of the origin formed the "edge" of the attractor). In considering the integer i^* , we were considering only the first few symbols in this sequence. In fact, if the rigorous conditions which we will describe in Section 3.4 and Appendix G apply, there will be homoclinic explosions at a dense set of r -values. Each of these explosions will remove a whole T -generated strange invariant set for the appropriate T . Consequently, the topological type of the attractor will change in every neighborhood of every r -value. We can state (Williams, 1980) that there will be an uncountable number of topologically distinct attractors in every neighborhood of every r -value. (The fact that homoclinic explosions occur at only a countable number of r -values does not contradict this result. The situation is similar to asking how many different sets $\{\text{rationals in } [0,\alpha]\}$ there are. Clearly there is an uncountable number of such sets, since the set is different for every value of α . However, each rational in the set leaves it as α passes through a rational value, and there is only a countable number of α -values at which a rational leaves the set.) It is worth emphasizing that this conclusion does nothing to disallow our local analysis of each homoclinic explosion. This analysis only involves trajectories which remain forever within the tubes around the particular homoclinic orbit we are considering. The fact that the unstable manifold of the origin will, within any small range of parameters that we consider, form an infinite number of other homoclinic orbits, is irrelevant. All such homoclinic orbits are formed only after the unstable manifold of the origin has left the small region under consideration.

The structure of the strange attractor, then, is different from the structure of the original strange invariant set because orbits and trajectories are actually disappearing; they are not just leaving the attractor and wandering off to some independent fate as the x and y orbits did. All the periodic orbits which disappear (except the x and y orbits) do so in homoclinic explosions. This is not true for all the aperiodic trajectories which disappear. In these notes, we will not consider aperiodic trajectories very often. They do not close up and do not, therefore, have topological difficulties coming into existence or disappearing; it is only the periodic orbits which require well-defined

bifurcations to appear or disappear. Take, for example, the aperiodic trajectory corresponding to the doubly infinite sequence $(xy^{26})^\infty x(xy^{26})^\infty$. This trajectory certainly exists in the original invariant set, and certainly exists no longer when $r = 28.0$. But there is no suitable T for which this trajectory belongs to a T -generated invariant set and so this trajectory was not removed in a homoclinic explosion. Or take the case of the aperiodic trajectory $(x)^\infty(y)^\infty$. This, too, exists in the original invariant set, and is part of the attractor for $r = r_A$. It is not, however, part of the attractor for any $r > r_A$. In general, if we know which periodic orbits exist, we will also know which aperiodic trajectories exist. However, aperiodic trajectories represent a non-trivial problem if we are to understand everything about the behaviour of the flow.

We can ask whether the strange attractor just continues to "collapse" as r increases, losing more and more of its orbits and trajectories in homoclinic explosions. Numerical experiments indicate that the integer i^* (which gives us some measure of the size of the attractor) decreases monotonically in $r_A < r < 28.0$. However, numerical experiments also indicate that at some r -value just larger than 28.0, the development of the attractor begins to reverse itself. When $r = 30.2$, i^* has increased to 33, and the attractor contains orbits and trajectories with up to 33 consecutive x's or y's in their symbolic representations. In $r > 30.2$, the integer i^* decreases steadily, reaching 2 for an r -value near 47.5 at which we have an xy^2 -homoclinic explosion, and reaching 1 for an r -value near 54.6 at which we have an xy -homoclinic explosion. For $r > 54.6$, the point R' lies to the right of AD . The first reversal of the development of the attractor should not require us to alter our view of the system very much; we just assume that the sequence of homoclinic explosions in $r < 28$ reverses itself and that in $28 < r < 30.2$ (approximately) homoclinic explosions replace some of the periodic orbits and trajectories which they earlier removed. We shall see, in Chapter 5, that the second reversal (at $r \approx 30.2$) is more complicated and that the return map changes in a qualitative way long before we reach r -values like 54.6. We can actually argue that this must be the case with a combinatorial argument, though readers may like to omit this argument on a first reading.

If we assume that the development of the attractor is monotonic in some interval, then each periodic orbit can only be removed from the attractor once. This implies that the set of all periodic orbits in the original invariant set can be divided up into disjoint subsets as follows:

(a) The x and y orbits involved in the Hopf bifurcation, (b) whatever is left in the strange attractor, and (c) a disjoint union of sets, each of which contains all the periodic orbits from some T -generated invariant set (these being the orbits which have been removed in various T -homoclinic explosions for different T 's). Combinatorially, it will only be possible to divide up the original periodic orbits into disjoint sets like this for certain limited choices of T 's. This implies that only certain homoclinic orbits can occur in the Lorenz system if our present picture of its development remains correct. (This very simple consideration allows us to correct a previous result for the maps discussed in Appendix G.) The xy -homoclinic orbit is one that cannot occur unless something has changed. If we try to imagine that the xy -homoclinic explosion is just a part of a sequence of homoclinic explosions which is removing orbits and trajectories from the strange invariant set, we will have to explain where two of these orbits came from. The two simplest orbits in the xy -generated invariant set removed in this bifurcation would have to be two distinct, non-symmetric orbits with symbolic descriptions $xy = T$ and $yx = S$. (Notice that we must look at the T and S descriptions of orbits involved in homoclinic explosions if we wish to know whether they are symmetric or not. There is no reason to suppose that we can tell the symmetry properties of an orbit from its global x and y description, unless the orbit in question happened to be born in the first x -homoclinic explosion; in this case the x/y and T/S descriptions of orbits are equivalent.) The original invariant set (and the strange attractor at $r = 28.0$) does not contain such orbits. The only orbit of an xy type which exists in the earlier sets is the symmetric orbit which is properly represented by the doubly infinite sequence $xyxyxy....$, and whose finite description is either xy or yx . (In Chapter 5 we will see that we can imagine a simple bifurcation which would produce the two "missing orbits" for the xy -explosion; the non-stable symmetric orbit can become stable by casting off two non-stable non-symmetric orbits as r increases. The appearance of the stable symmetric xy orbit would, of course, mean that our picture of the return map had changed in a qualitative way. In fact, the qualitative change which occurs for $\sigma = 10$ and $b = 8/3$ is much more complicated than this. It is possible, though, that the simpler situation just discussed is relevant for other values of the parameters σ and b . See Chapter 5.)

3.4. GEOMETRIC MODELS OF THE LORENZ EQUATIONS

The arguments of the last section were largely intuitive. There is now a considerable rigorous mathematical literature on "Lorenz attractors" (Williams 1977, 1979, 1980; Guckenheimer, 1976; Williams and Guckenheimer, 1980; Rand, 1978), but this literature is, unfortunately, not necessarily of direct relevance to the Lorenz flow. The "Lorenz attractors" that are well understood (and which have the properties described in the last section) occur in model flows that are constructed to have certain properties. Appendix G contains a brief description of some of the work on these geometric models of the Lorenz system. Williams (1980) explains the steps necessary to show that the attractor in the Lorenz flow is topologically equivalent to one of the model "Lorenz attractors". The most important step would be to demonstrate the existence of a contracting foliation on the flow.

Consider Fig. 3.6, which shows a schematic version of the return map. In this figure, motivated by Fig. 3.4(a), we have not indicated the width of the attractor. We merely remark that the attractor intersects the plane near the two arcs RL' and $L'R$. One application of the return map will map two points such as X and Y (Fig. 3.6) close to a point such as Z . If we have chosen the points X and Y correctly, the distance between successive iterates, $\psi^n(X)$ and $\psi^n(Y)$, will tend to zero as n increases. (Though most nearby trajectories diverge, we can start trajectories at points "equidistant" from AD and these trajectories will converge.) Since we are dealing with a continuous system, there must be a whole arc of points between X and Y such that trajectories started at any two points on this arc get closer together as they are followed under the flow. If we can fill the return plane with a continuum of such arcs (in such a way that each arc is taken by the return map into another of the arcs) then we say we have a contracting foliation. More rigorously, we require that there be constants c and λ , $0 < \lambda < 1$, such that if x_1 and x_2 lie on the same arc, then $\psi^n(x_1)$ and $\psi^n(x_2)$ also lie on the same arc (as one another) and the distance between these two n^{th} iterates is less than $c\lambda^n$ for all positive integers n . This is not a totally unfamiliar notion. Our analysis near the origin, in Appendix D, depended crucially on the local existence of a contracting foliation which allowed us to extract a return map in one dimension. In Appendix D, this contraction was given by the negative real eigenvalue of largest magnitude. Globally, we cannot appeal to properties of the

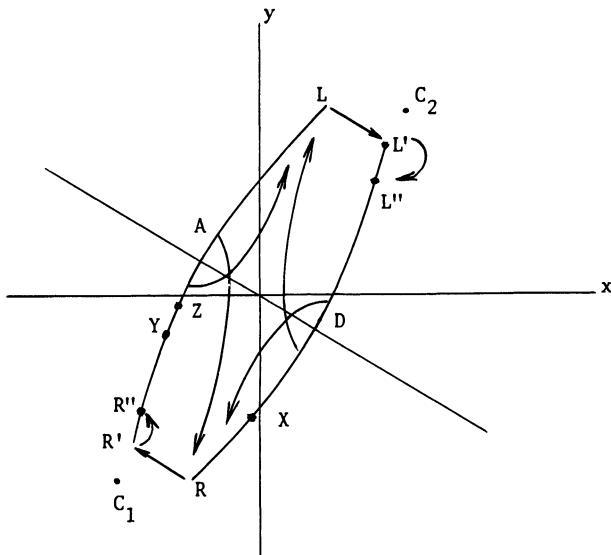


Figure 3.6. Simplified return map at $r = 28.0$.

linearized flow near the origin, and it does not seem likely that the existence (or non-existence) of a contracting foliation for the Lorenz flow will be proved in the near future. If there is no contracting foliation for the Lorenz flow, then we can expect that many of the detailed remarks about the structure of the strange attractor (Section 3.3) will be false. However, the global contracting foliation is not needed for the study of the individual homoclinic explosions which may occur; whatever the properties of the Lorenz flow, we can expect that many of the remarks about homoclinic explosions removing periodic orbits from the non-wandering set (we are no longer sure that we have a strange attractor) will remain true.

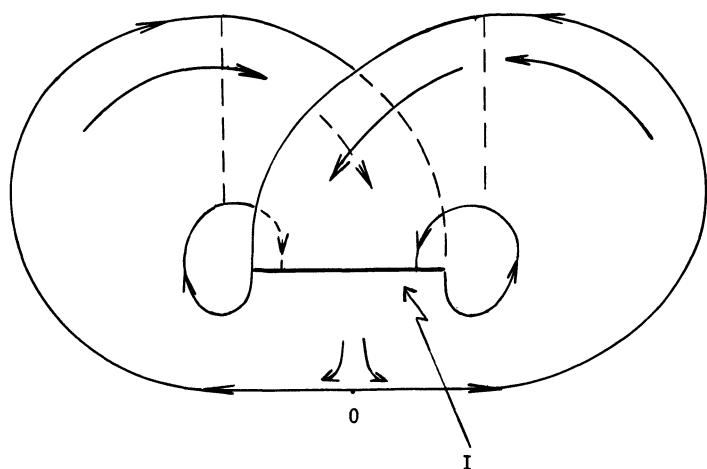
If we do assume the existence of a contracting foliation, or if we build a model flow that does have a foliation, we can reduce the dimension of the problem by one by identifying each contracting arc with a single point. The price that we have to pay for this simplification is the loss of unique histories for trajectories; different points on the same contracting arc have convergent futures, but the history of those points (for example, X and Y) may be quite different. Thus, the two-dimensional flow that we obtain will be a semi-flow, and it will exist on a "branched manifold". This is a two-dimensional sheet on which the

forward flow is uniquely determined, but which has branch points where two forward trajectories join up; we cannot follow the flow backwards, because each time we come to one of these points we cannot decide which branch of the manifold to take. Fig. 3.7(a) shows the appropriate branched manifold for the Lorenz flow. Trajectories started below the line I swing round on either the left or right branch of the manifold (in Fig. 3.7(a) we have drawn the right-hand branch in front of the left-hand branch) depending on whether they lie to the left or right of O. The left and right branches of the manifold join up again at the line I. Fig. 3.7(b) shows the first return map of the line I to itself. The study of one-dimensional maps of an interval to itself, like that shown in Fig. 3.7(b), is an essential part of the study of geometric models of the Lorenz flow. (It is relatively difficult to extract the appropriate one-dimensional maps from simulations of the Lorenz equations; however, we can extract rather different one-dimensional maps rather easily. Lorenz (1963, 1979, 1980a,b) has consistently looked at the approximately one-dimensional map which you get by plotting successive local maxima of the variable z against one another for trajectories computed numerically. Other authors have taken the same course. Studying these alternative one-dimensional maps - which are discussed in Appendix H - does not avoid the necessity of assuming some simplification equivalent to the one we have just discussed. We can derive the max-of-z maps from our simplification by considering the first return of the two dotted vertical lines on the branched manifold in Fig. 3.7(a). Since the system is symmetrical, we can identify the two lines together and obtain a single one-dimensional map of an interval to itself.)

It is possible to understand the whole development of the Lorenz system in $1 < r < 28$ by looking at one-dimensional maps like that shown in Fig. 3.7(b). This is probably worthwhile. (See, also, Guckenheimer's contribution to Guckenheimer et al (1980).) Fig. 3.8 shows five maps, one for each interesting parameter interval. A "trajectory" on one of these maps is a sequence of points obtained by repeated application of the map to some initial point.

For $r < 13.926$, the appropriate map is shown in Fig. 3.8(a). All trajectories move towards one or other of the stable points C_1 or C_2 . For $13.926 < r < 24.06$, the appropriate map is shown in Fig. 3.8(b). C_1 and C_2 are still stable and attract all trajectories starting between C_1 and Y, or between C_2 and X, directly. The points X and Y

(a)



(b)

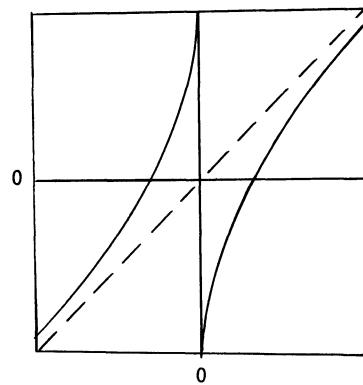


Figure 3.7. (a) The two-dimensional branched manifold for the Lorenz flow.
(b) The one-dimensional return map of the line I in Fig. 3.7(a).

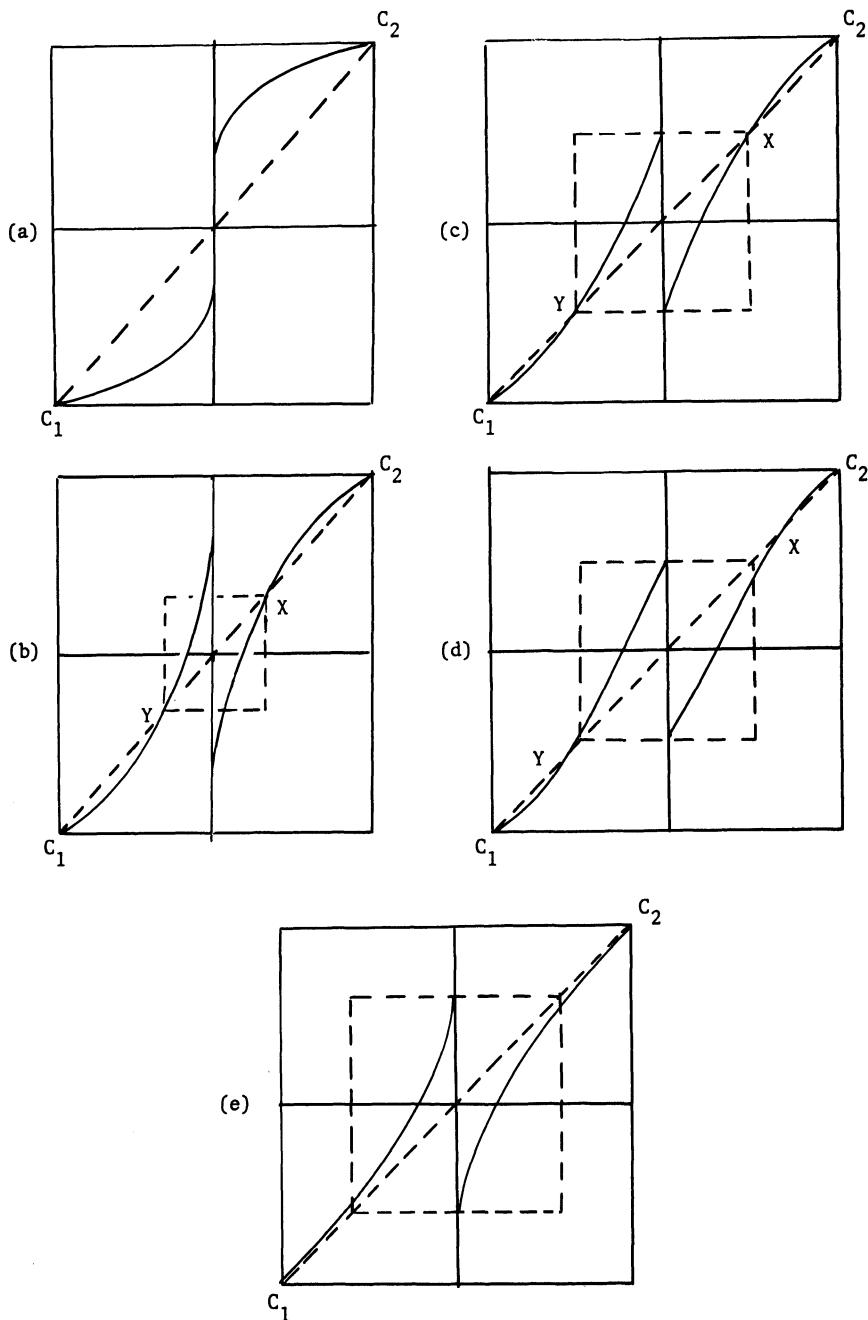


Figure 3.8. One-dimensional return maps, extracted as in Fig. 3.7, for the Lorenz flow. (a) $r < 13.926$; (b) $13.926 < r < 24.06$; (c) $r = r_A \approx 24.06$; (d) $r_A < r < r_H \approx 24.74$; (e) $r > r_H$.

represent the two simplest members of the strange invariant set (the orbits x and y). Between X and Y , i.e., within the dotted box, is the strange invariant set which can be analyzed in the same way as the one-dimensional map in Appendix D. Most trajectories that start within the dotted box will eventually fall on the sections of the return curves near to zero and outside the dotted box. Thereafter they tend directly to C_1 or C_2 . As r approaches 24.06, the sections of the return curves which lie outside the dotted box get smaller and smaller. The probability of a given trajectory leaving the dotted box within any given number of steps decreases. Here we see preturbulence for those trajectories which start within the dotted box. At $r = r_A \approx 24.06$, the appropriate map is shown in Fig. 3.8(c). Here, zero is mapped onto X or Y . For the Lorenz system, this implies that the two branches of the unstable manifold of the origin are included in the stable manifolds of the simple periodic orbits x and y . (The stable manifold of a periodic orbit is the union of those trajectories which tend towards it. All periodic orbits in the Lorenz system have a non-trivial stable manifold because of the volume contracting nature of the flow - see Chapter 1. The stable manifold of a non-stable periodic orbit will be two-dimensional; in the case of the x and y orbits these manifolds look, near to the orbits, like cylinders. Further from the orbits, these cylinders may open out and twist round in complicated ways. As the x and y orbits approach the Hopf bifurcation, the cylinders close in around the trajectory that tends towards the appropriate stationary point at a rate governed by the eigenvalue of the linearized flow near to C_1 and C_2 that is real and negative for all r - see Chapter 1. At $r = r_A$, the cylinders have closed in just enough to cut off the origin from the stationary points C_1 and C_2 . Trajectories started near the strange invariant set can never escape to the stationary points, and the set becomes an attractor.) On the one-dimensional map, if you start within the dotted box you remain there forever; again, the invariant set has become attracting. In $24.06 < r < 24.74$, the appropriate map is shown in Fig. 3.8(d). Trajectories started between X and C_2 will tend towards C_2 , and trajectories started between Y and C_1 will tend toward C_1 . Trajectories started between X and Y will tend to the dotted box, and once within the box will remain there forever. The box contains the strange attractor, and the study of that section of the return map which lies within the box is discussed in Appendix G. Notice that we have three attracting sets. The points X and Y divide their respective basins of attraction. In the real system this

means that the stable manifolds of the simple Hopf orbits divide \mathbb{R}^3 into three basins of attraction. Notice that X and Y are not part of the attracting set for $r > r_A$. Finally, in $r > 24.74$, we have Fig. 3.8(e). The points X and C_2 , and Y and C_1 , have coalesced (the Hopf bifurcation), and all trajectories now tend towards the dotted box which contains the strange attractor. The section of the return map within the dotted box is similar to that in the previous figure, and as far as the strange attractor is concerned we do not distinguish between cases (d) and (e). Once again we emphasize that the Hopf bifurcation is irrelevant to the development of the strange attractor.

Notice that in describing the development of the behaviour using these one-dimensional maps, we have not only had to assume that we are justified in extracting such maps, but we have had to assume that the maps look as we have drawn them. For instance, in all the pictures 3.8 (b)-(e), we have drawn the slope of the map greater than one everywhere between points X and Y. This is equivalent to assuming that our numerically computed return maps stretch all distances in one direction (the direction "perpendicular" to the contracting foliation). This is something we cannot ever check for the real Lorenz equations. Indeed, much of the analysis in Appendix G depends on the slope of the relevant part of the map being greater than $\sqrt{2}$ (though the results are not necessarily all false if the actual slope is not everywhere greater than $\sqrt{2}$). It is clear that we can only claim to understand the strange attractor if we make a number of assumptions; without these assumptions we cannot even claim with certainty that we have a strange attractor.

3.5. SUMMARY

If we are prepared to accept the assumptions which lead up to the derivation of the geometric model of the Lorenz equations, we can understand the development of the behaviour in the whole parameter interval $r < 28.0$ (and probably in a slightly longer interval also). This understanding comes either from the rigour of Appendix G, or from the more intuitive arguments in Section 3.3.

The advantage of the intuitive arguments is that many of them will apply whether or not the geometric model is a good one. If the model eventually turns out to be wrong, we will no longer be able to say with any certainty how the equations behave, or how the detailed structure of the non-wandering set changes. However, only very crude numerical experi-

ments are required to show the existence of homoclinic orbits, and we would still be able to conclude that a sequence of homoclinic explosions (perhaps no longer occurring at a dense set of r -values) removes periodic orbits and trajectories from the non-wandering set. There might, of course, be other bifurcations going on, about which we would know very little. Nonetheless, the kind of combinatorial arguments that we used at the end of Section 3.3 would still be useful; if we found homoclinic explosions removing periodic orbits which were not part of the original strange invariant set (whose existence for r just greater than 13.926 is known almost rigorously), we would know we had to look for bifurcations that explained the existence of these extra orbits.

We shall see, in Chapter 5, how this kind of approach can be generalized to cope with situations in which we know that the geometric model of the Lorenz flow is not appropriate. First, though, we examine some of the numerically observed behaviour which occurs for r -values considerably larger than 28.0.

Chapter 4

Period Doubling and Stable Orbits

Several authors have noticed that numerical simulations of the Lorenz equations indicate the existence of stable periodic orbits in some intervals of r -values. This behaviour is quite different from the behaviour discussed in Chapter 3, since for r -values near 28.0 we saw no stable periodic orbits, and had strong arguments that none could exist. We will attempt to reconcile the two phenomena in Chapter 5.

Three intervals of stable periodic behaviour can be found easily. The first is rather short but includes the r -value $r = 100$ (which is presumably an r -value likely to be included in any "random" selection of parameter values chosen for examination). The second is longer and includes $r = 150.0$. The third is approximately $200 < r < \infty$. In each of these r -intervals we can observe the now well-known phenomenon of a "period doubling cascade." Numerical simulations for "randomly" selected parameter values lying between these intervals (or between $r = 28$ and the first interval) show only chaotic behaviour. We will see, in Chapter 5, that there are actually many more parameter intervals in which periodic orbits are stable; we shall describe a technique to locate a few of them. However, we will study the three largest intervals now, and these will act as models for the others.

4.1. THREE BIFURCATIONS INVOLVING PERIODIC ORBITS

Readers should be familiar with three bifurcations that can happen to periodic orbits as the parameters change. The three kinds of bifurcation are:

1) The saddle-node bifurcation. Suppose the bifurcation happens at $r = r^*$. On one side of r^* we have a non-stable periodic and a stable periodic orbit. As the parameter increases (or decreases) towards r^* , the two orbits move very close together and their periods tend towards some common value. On the other side of r^* neither orbit exists. Fig. 4.1 shows the bifurcation diagram and a schematic picture of the flow either side of r^* . This bifurcation can occur between two symmetric, or between two non-symmetric orbits. (In some systems this bifurcation can occur with a non-stable and an unstable orbit "annihilating" one another (see Appendix A for definitions). In the Lorenz system we cannot have unstable orbits because the flow is volume contracting.)

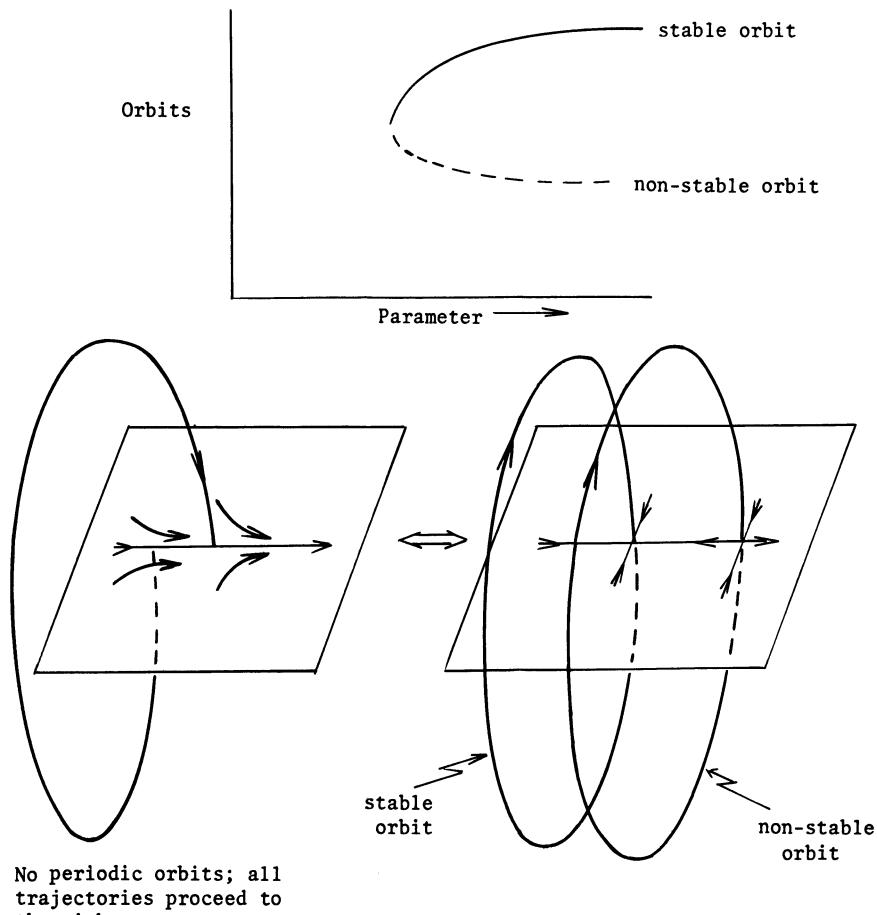


Figure 4.1. The saddle-node bifurcation. (a) Bifurcation diagram, (b) Schematic view of the flow; the periodic orbits can be both non-symmetric or both symmetric.

2) The symmetric saddle-node bifurcation (or symmetry breaking bifurcation). This bifurcation is the periodic orbit analogue of the bifurcation involving stationary points which occurred when $r = 1.0$. On one side of the bifurcation value, r^* , we have two non-symmetric orbits and a symmetric orbit. As r increases (or decreases) towards r^* , these three orbits move close together and their periods tend towards some common value. On the other side of r^* there is only one symmetric orbit. Fig. 4.2 shows the bifurcation diagram and a schematic view of the flow either side of the bifurcation. In the Lorenz system this bifurcation has only been observed when it involves two stable non-symmetric orbits and a non-stable symmetric orbit coalescing to form a stable symmetric orbit. (The other possibility is two non-stable non-symmetric orbits coalescing with a stable symmetric orbit to form a non-stable symmetric orbit.) It is called the symmetric saddle-node bifurcation because we could disturb the symmetry slightly in such a way that we are left with a saddle-node bifurcation. (It is sometimes called the symmetry breaking bifurcation because stability is transferred from a symmetric orbit to non-symmetric orbits.)

3) The period doubling bifurcation (or flip). On one side of the bifurcation value, r^* , we have a periodic orbit. As r increases (or decreases) towards r^* , its period approaches some value T . On the other side of r^* this periodic orbit continues to exist, but near it there is a periodic orbit of period $2T$. See Fig. 4.3. In the Lorenz system this bifurcation has only been observed when a stable periodic orbit becomes non-stable by throwing off a stable periodic orbit of twice the period. The other possibility is for a non-stable periodic orbit to become stable by throwing off a non-stable orbit of twice the period. The period doubling bifurcation can only occur in non-symmetric orbits. For it to work, trajectories moving near the orbit of period T must "flip" half-way round it on each revolution - thus allowing the possibility of a periodic orbit of period $2T$ after two revolutions; this cannot occur near a symmetric orbit.

(In Appendix E there is a more mathematical presentation of these bifurcations; there we explain how the computer program which locates periodic orbits and follows them with changing r can inform us if we are approaching a bifurcation, and if so, how to determine the type of the bifurcation. See also, for example, Guckenheimer et al (1980).)

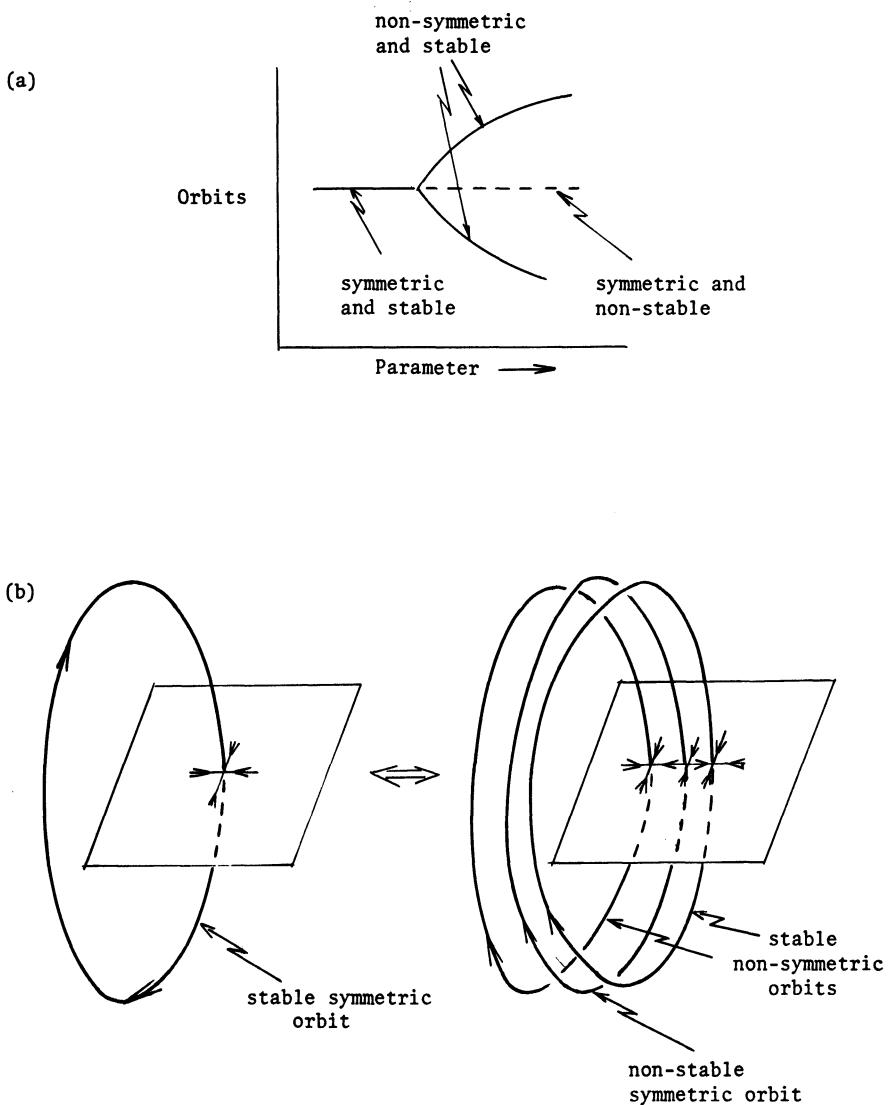


Figure 4.2. The symmetric saddle-node bifurcation. (a) Bifurcation diagram, (b) Schematic view of the flow.

Note: The three bifurcations described above are not the only bifurcations involving periodic orbits. It is possible, in some systems, that a periodic orbit will bifurcate to an invariant torus in a bifurcation analogous to the Hopf bifurcation for stationary points. This bifurcation is impossible in the Lorenz system since the volume within an invariant torus would not be contracting as required. There are also the homoclinic bifurcations and Hopf bifurcations that we have already studied. There are still other types of bifurcation which involve periodic orbits (see Chapter 8 for an example) but we shall not consider them now.

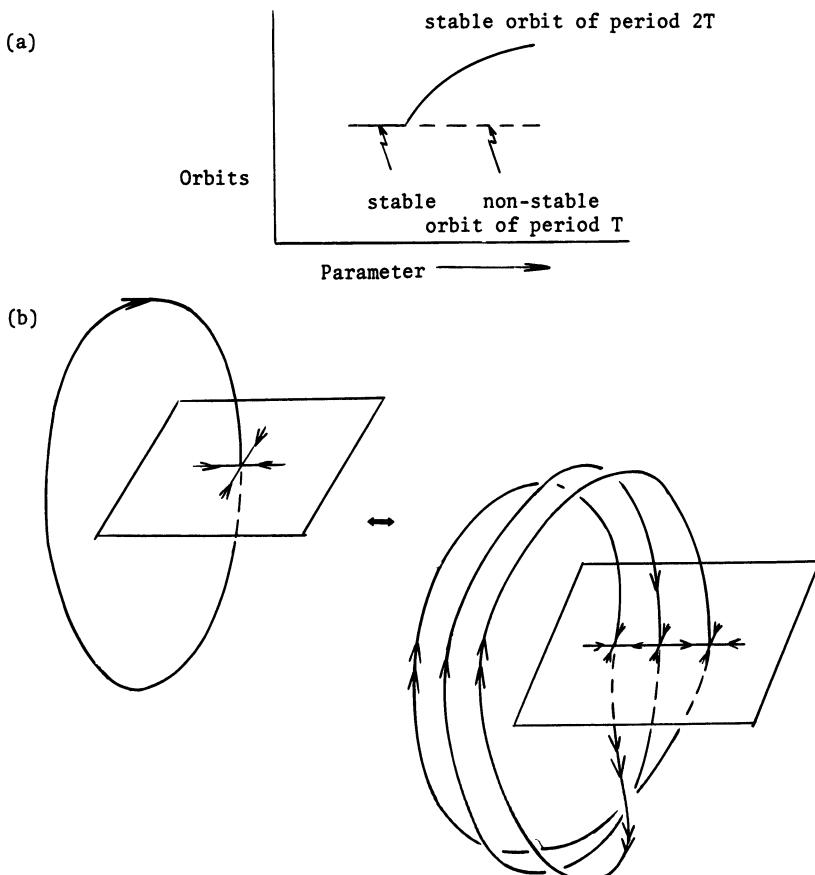


Figure 4.3. The period doubling bifurcation. (a) Bifurcation diagram, (b) Schematic view of the flow; all orbits are non-symmetric.

4.2, $99.524 < r < 100.795$. THE x^2y PERIOD DOUBLING WINDOW

This parameter interval has been examined by Franceschini (1980), from whose paper I have extracted the parameter values at which the various bifurcations occur. What we see is as follows:

i) Between $r \approx 99.98$ and $r \approx 100.795$ there is a stable periodic x^2y orbit. It is shown, for $r = 100.5$, in Fig. 4.4(a). (In this chapter we will continue to make use of the symbolic descriptions of periodic orbits that we used in Chapter 3; i.e., we will write an "x" everytime the orbit spirals round in $x > 0$ and a "y" everytime it spirals round in $x < 0$. This usage is very convenient since it is easy to name an orbit once we see it. We shall see in Chapter 6 that we are probably justified - in a topological way - in using these descriptions.) Of course, we still have symmetry so there is also a stable y^2x orbit in the same r -interval. All numerically calculated trajectories appear to be attracted to one or other of these stable orbits.

ii) Between $r \approx 99.629$ and $r \approx 99.98$ there is a stable x^2yx^2y periodic orbit (and its symmetric image, a stable y^2xy^2x orbit). The stable y^2xy^2x orbit is shown in Fig. 4.4(b). It is only just possible to distinguish the two loops of the orbit. As r increases towards 99.98 the two loops of this orbit get closer together and eventually merge. This is an example of the period doubling bifurcation. For $r > 99.98$ we have a stable x^2y orbit. For $r < 99.98$ we have a stable x^2yx^2y orbit and, though we cannot observe it directly, a non-stable x^2y orbit. The period doubling occurs as r decreases.

iii) In the interval $99.547 < r < 99.629$ we see a stable $x^2yx^2yx^2yx^2y$ ($= (x^2y)^4$) orbit. The period has doubled again.

iv) As r decreases further we see (using techniques such as those described in Appendix E) more doubling bifurcations. According to Franceschini (1980), the next bifurcations occur at $r \approx 99.529$ and $r \approx 99.5255$. Notice that the interval between bifurcations is decreasing rather rapidly as r decreases.

v) It seems that as r decreases we get an infinite number of period doubling bifurcations which occur at r -values which tend to some limit. We cannot verify this directly but there is a very good reason to believe it. It is well known that infinite sequences of period doublings occur in discrete one-dimensional maps of an interval to itself. (See, for example, May (1976) for an introduction. Collet & Eckmann (1980) or Guckenheimer (1979a, 1979b, 1980) have more recent results.) The sequences of period doublings occur in precisely the situation where the

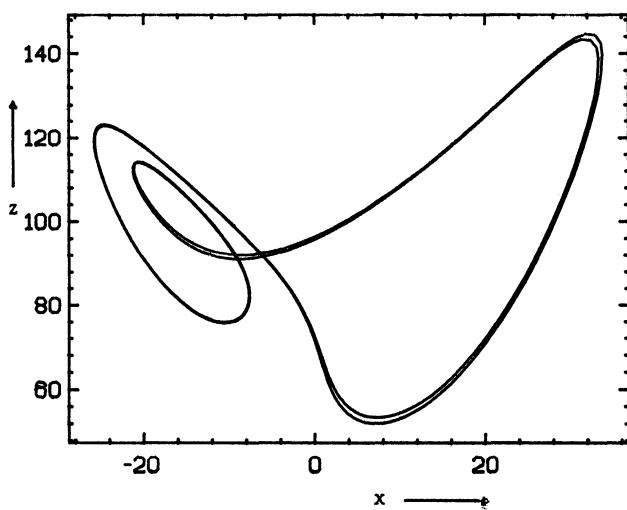
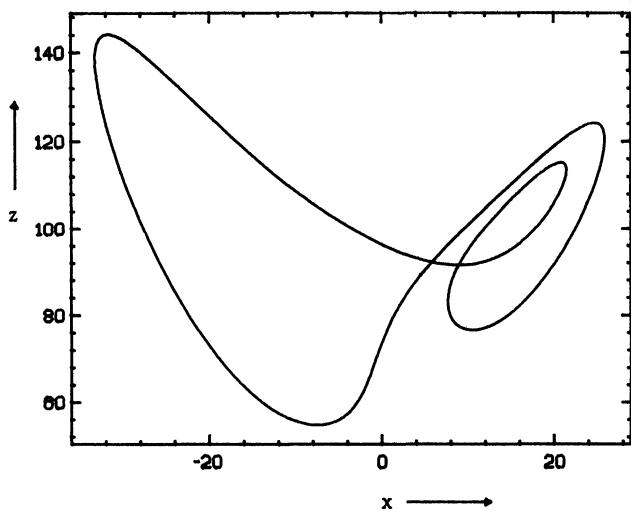


Figure 4.4. (a) Stable x^2y orbit. ($r = 100.5.$)
(b) Stable y^2xy^2x orbit. ($r = 99.65.$)

maps show alternate parameter intervals of stable periodic and chaotic behaviour. Furthermore, following conjectures by Feigenbaum (1978) and others, Collet, Eckmann & Koch (1980) and Collet & Eckmann (1980) have proved that there are some "universal" properties associated with these infinite sequences of bifurcations, whether they occur in one-dimensional maps or in flows. In particular, we can expect there to be an asymptotic

limit of the ratio $\frac{r_{n-1} - r_n}{r_n - r_{n+1}}$, where r_n is the parameter value at which the n^{th} doubling bifurcation occurs, which will be some "universal" constant, δ . δ has been calculated numerically for the one-dimensional discrete maps and equals (approximately) 4.6692016091029909... .

Franceschini (1980) looked at this ratio of the distances between successive bifurcations and found, for the sequence discussed above, a sequence of values tending to some limit near 4.67.

We henceforth assume that we do actually have an infinite sequence of doubling bifurcations, occurring as r decreases, which is over by some r -value, $r_\infty \approx 99.524$ (calculated by Franceschini using the constant δ).

We will call intervals such as $99.524 < r < 100.795$ *period doubling windows*. This particular interval we call, for obvious reasons, the x^2y period doubling window.

We leave consideration of what happens just below and just above period doubling windows (i.e., $r < 99.524$ and $r > 100.795$) until we have looked at another window. For the moment, notice that for r just less than 100.795 we have a symmetric pair of non-symmetric stable periodic orbits, x^2y and y^2x , and that for $r > 100.795$ we do not see them anymore. It would be reasonable to guess (correctly as it will turn out) that they are disappearing in simultaneous saddle-node bifurcations that involve non-stable x^2y and y^2x orbits which exist, unobserved, in the parameter range $r < 100.795$. When $r < 99.524$ we expect to have an infinite collection of non-stable periodic orbits "left over" from the infinite sequence of period doubling bifurcations; we do not lose any orbits in a period doubling bifurcation - they just lose their stability. We should have ($r < 99.524$) one each of $(x^2y)^{2^i}$ and $(y^2x)^{2^i}$ for $i = 0, 1, 2, 3, \dots$, as well as an extra (x^2y) and (y^2x) which are the non-stable orbits involved in the saddle-node bifurcation at $r \approx 100.795$. See Fig. 4.5 for the bifurcation diagram where some of the orbits are labelled for future reference.

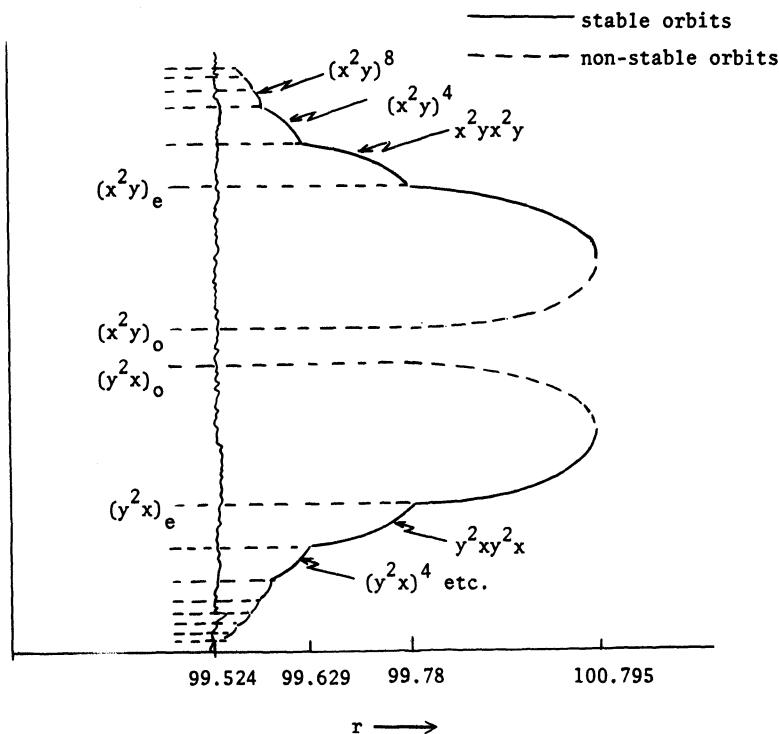
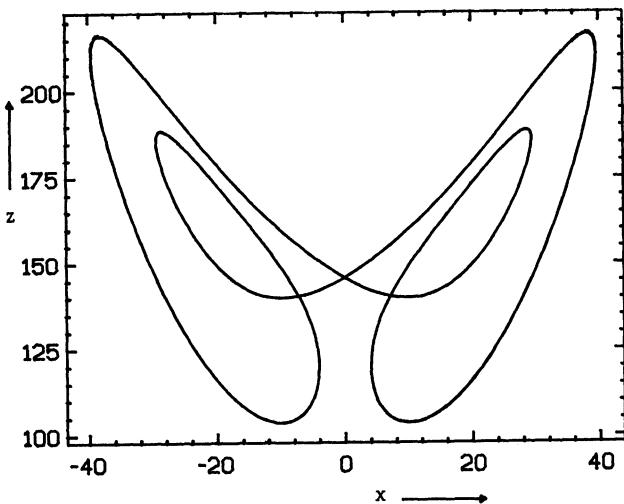


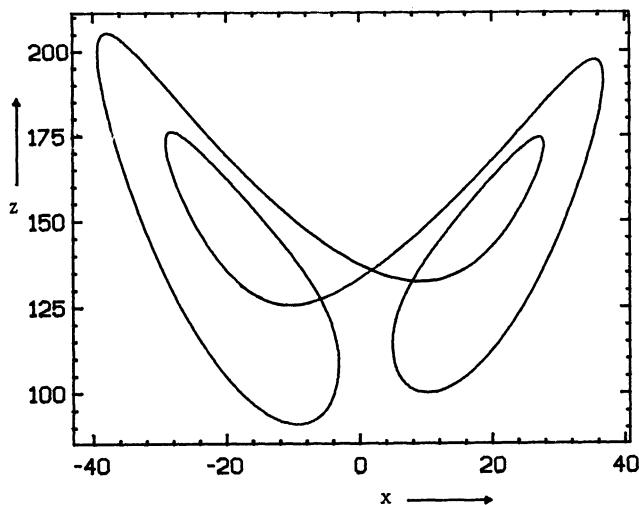
Figure 4.5. Bifurcation diagram for the x^2y^2 period doubling window.

4.3. $145 < r < 166$. THE x^2y^2 PERIOD DOUBLING WINDOW

This second period doubling window has been studied by Manneville & Pomeau (1979, 1980) and involves periodic orbits such as x^2y^2 . This window is very similar, in most respects, to the x^2y window studied above. The main difference is that it involves a stable *symmetric* periodic orbit shown, in Fig. 4.6(a) when $r = 160.0$. This orbit remains stable in the interval $154.4 < r < 166.07$ (approximately). Manneville & Pomeau (1979) suggest, misleadingly, that this orbit gradually loses its symmetry as r decreases. In fact, at $r \approx 154.4$, there is a *symmetric saddle-node bifurcation* (*symmetry breaking bifurcation*) and the stable symmetric orbit loses its stability to a pair of non-symmetric x^2y^2 periodic orbits which are stable in the interval $148.2 < r < 154.4$. One of them is

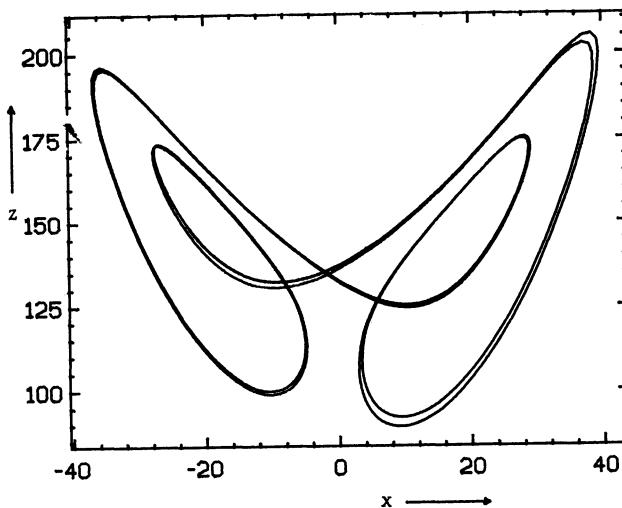


(a)



(b)

Figure 4.6. (a) Stable symmetric x^2y^2 orbit, $r = 160$. (b) Stable non-symmetric x^2y^2 orbit, $r = 148.5$. (c) Stable non-symmetric $x^2y^2x^2y^2$ orbit, $r = 147.5$.



(c)

Figure 4.6 (continued)

shown, for $r = 148.5$, in Fig. 4.6(b). The symmetric x^2y^2 orbit continues to exist below $r = 154.4$ though it is, of course, non-stable.

At $r \approx 148.2$ these non-symmetric x^2y^2 orbits undergo simultaneous period doubling bifurcations. A doubled, stable orbit is shown in Fig. 4.6(c). It seems that these doubled orbits double again, and then again, etc., exactly as described above for the x^2y window. There seems to be a limit point for this sequence of doubling bifurcations somewhere above $r = 145.0$. It is probably not worth the trouble to calculate the various bifurcation values exactly. See Fig. 4.7 for a bifurcation diagram where some of the orbits are labelled for future reference.

Other authors, e.g., Shimizu & Morioka (1978a,b,c) and Shimada & Nagashima (1978) have studied this and other period doubling windows, though their choice of parameter values ($\sigma = 16$, $b = 4$) is a little unusual. The authors of these papers are mistaken in their various assertions a) that they have located all the period doubling windows, b) that the symmetric stable orbits lose their symmetry gradually, and c) that the bifurcation at the top of the period doubling window is a very complicated one (see below).

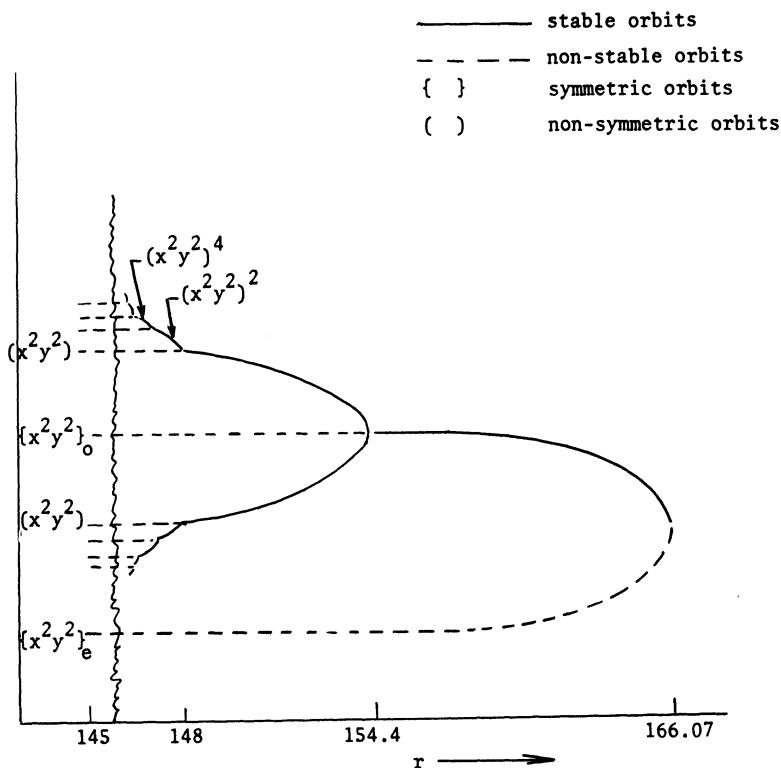


Figure 4.7. Bifurcation diagram for the x^2y^2 period doubling window.

4.4. INTERMITTENT CHAOS

We have to ask what happens at the top end (i.e., r large) of a period doubling window. Manneville & Pomeau (1979, 1980) have studied a phenomenon which they call "intermittent chaos" seen at r -values greater than $r = 166.07$, just outside the x^2y^2 period doubling window. Though the stable symmetric x^2y^2 orbit exists no longer, trajectories seem to move near it, or near where we would expect it to be if it did continue to exist, for long periods of time. They then wander off and behave chaotically for a while before returning to the almost periodic or "laminar" behaviour. Fig. 4.8 shows the variable x plotted against time for a typical trajectory calculated when $r = 166.3$. At this r -value the

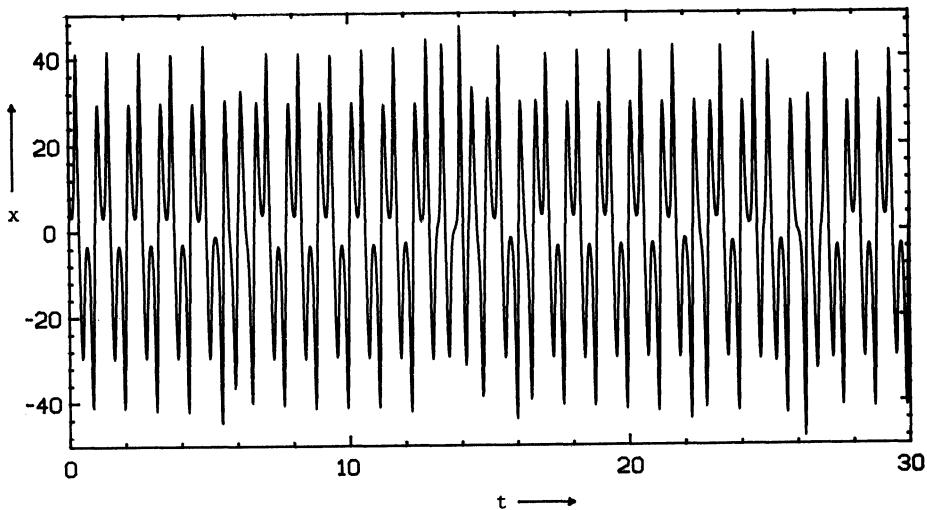


Figure 4.8. Intermittent behaviour at $r = 166.3$. There are periods of almost periodic behaviour interspersed with short "chaotic" bursts (*).

intervals of chaotic behaviour are very short. Most of the behaviour is "laminar".

As r increases, the chaotic periods of behaviour increase in length and the laminar periods decrease in length, until eventually the motion appears entirely chaotic. Near the "intermittency threshold" (the point where the stable x^2y^2 orbits ceases to exist), the mean length of the laminar intervals appears to vary at a rate proportional to $\frac{1}{(r - r_c)^{\frac{1}{2}}}$, where $r_c \approx 166.07$ (Manneville & Pomeau, 1979). The description of this process as a "continuous" transition to chaos is suggestive but misleading. In fact, there is a single bifurcation at the intermittency threshold of the saddle-node type.

Both the intermittency, and the bifurcation, can be understood intuitively if we look at some suitable part of some suitable return plane. Manneville & Pomeau (1979) looked at the plane $x = 0$ and recorded points of intersection, (y, z) , of trajectories with this plane. They then extracted a graph, y_n against y_{n+1} , of successive y -values. We do not necessarily expect this one-dimensional approach to yield a curve; rather we expect some kind of scatter diagram. But, as in Chapter 3, the points do all seem to lie on a curve with no noticeable thickness (though, un-

like Chapter 3, we have no hope of rigorously justifying our one-dimensional approach in the present situation). A small section of this graph is shown schematically in Figs. 4.9. Fig. 4.9(a) shows the situation when $r < r_c$. In this picture we expect there to be two periodic orbits represented by the two intersections of the graph with the line $y_{n+1} = y_n$. One of them will be stable, the other will be non-stable. In Fig. 4.9(b), which shows the situation for $r > r_c$, these two periodic orbits have disappeared in a saddle-node bifurcation which occurs as the graph becomes tangential to the line $y_{n+1} = y_n$. However, trajectories which intersect the return plane in points such as the point represented by X in Fig. 4.9(c), can be expected to follow paths determined by the path that the point X in Fig. 4.9(c) makes under successive iterations of the return map shown there. Such a path is illustrated on Fig. 4.9(c). Consequently, we expect that if a trajectory (which is wandering around near some non-wandering set whose structure we do not know) arrives at a place close to where the stable periodic orbit used to exist, then it will stay there for some time before escaping to other parts of the non-wandering set (not represented in Fig. 4.9). Providing the behaviour is chaotic (numerically speaking) at this parameter value, the trajectory will eventually fall once more into the region where it shows the almost periodic or laminar behaviour. The length of time spent on laminar behaviour will vary with the number of steps it takes points such as X on Fig. 4.9(c) to escape from the area of the graph shown on that figure. This will vary with the distance between the graph and the line $y_{n+1} = y_n$. If we assume that the curve shown in Fig. 4.9 is quadratic, and that the curve moves relative to the line $y_{n+1} = y_n$ at a constant speed as the parameter changes, then we can calculate (Manneville & Pomeau, 1979) that the number of these steps will vary at a rate proportional to $\frac{1}{(r - r_c)^{\frac{1}{2}}}$.

Intermittent chaos can also be observed at the top end of the x^2y period doubling window. In this case we had simultaneous saddle-node bifurcations involving x^2y and y^2x periodic orbits. Fig. 4.10 shows the variable x plotted against time for $r = 100.8$, just above the bifurcation value ≈ 100.795 . Most of the behaviour is laminar, and there are now two kinds of laminar behaviour - x^2y or y^2x . The short chaotic bursts connect periods of either x^2y or y^2x laminar behaviour together in a "random" sequence.

Notice that "intermittent chaos", like "chaos" itself, is a phenomenon observed numerically. To say that we see intermittent chaos is not to

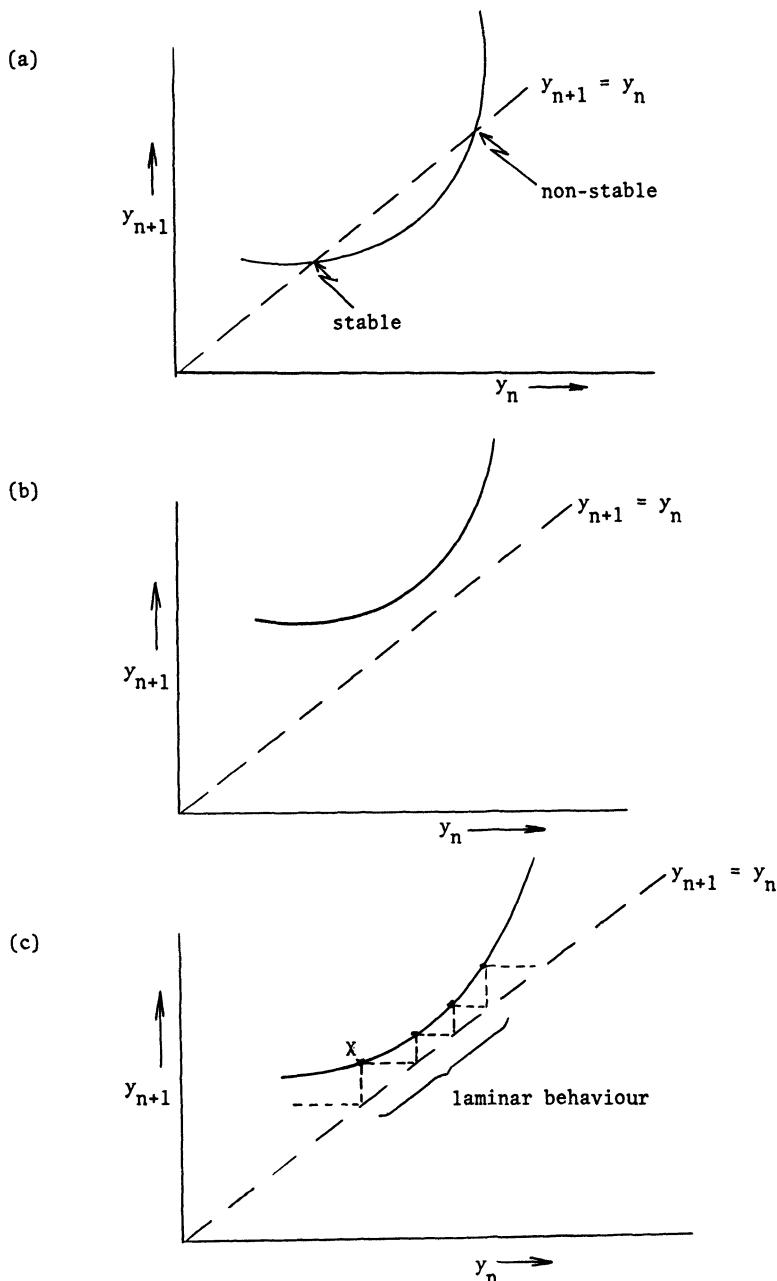


Figure 4.9. Schematic versions of a section of an approximately one-dimensional return map. (a) $r < r_c$. One stable and one non-stable periodic orbit. (b) $r > r_c$. The two orbits have disappeared in a saddle-node bifurcation. (c) Laminar periods occur in the flow as trajectories intersect the return plane several times near to where the periodic orbits used to exist.

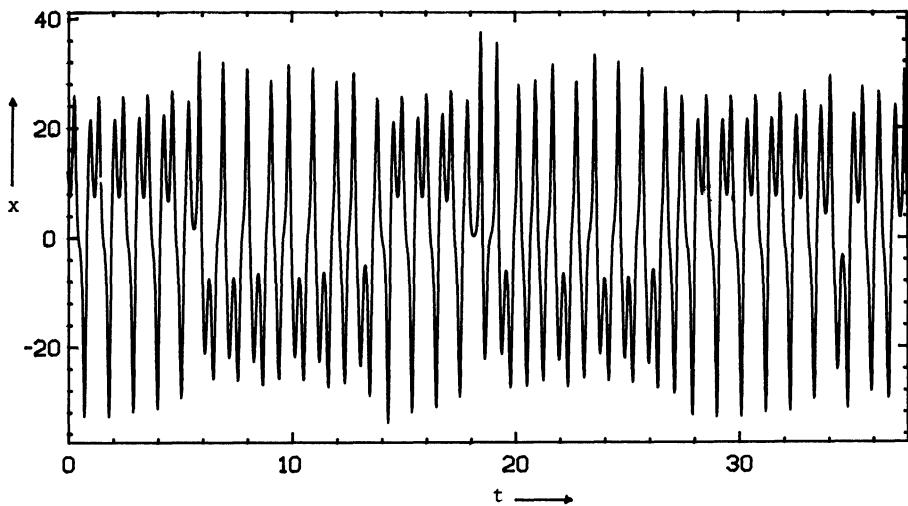


Figure 4.10. Intermittent behaviour above the x^2y window. Laminar bursts of x^2y and y^2x behaviour occur in "random" sequence, separated by short chaotic bursts.

claim that we know the structure of the non-wandering set, the behaviour of the flow near the non-wandering set, or which parts of the non-wandering set it is that are actually attracting.

4.5. $215.364 < r < \infty$. THE FINAL xy PERIOD DOUBLING WINDOW

We will examine one more period doubling window. This has been studied by Robbins (1979) and Lorenz (1980a) amongst others. We refer to it as the final xy period doubling window because it involves periodic orbits such as xy , and because it appears to be the last period doubling window. The window is exactly like the x^2y^2 window except that the final stable symmetric xy orbit which it produces does not suffer annihilation in a saddle-node bifurcation, but continues to exist for all $r > 313$. (We shall see in Chapter 7 that this is to be expected on theoretical grounds.) Fig. 4.11 shows some of the periodic orbits involved in this window. It appears that the infinite sequence of period doubling bifurcations accumulates at an r -value near 214.364 (Lorenz, 1980a).

Fig. 4.12 shows the bifurcation diagram for this window, and various of the orbits involved are labelled for future reference.

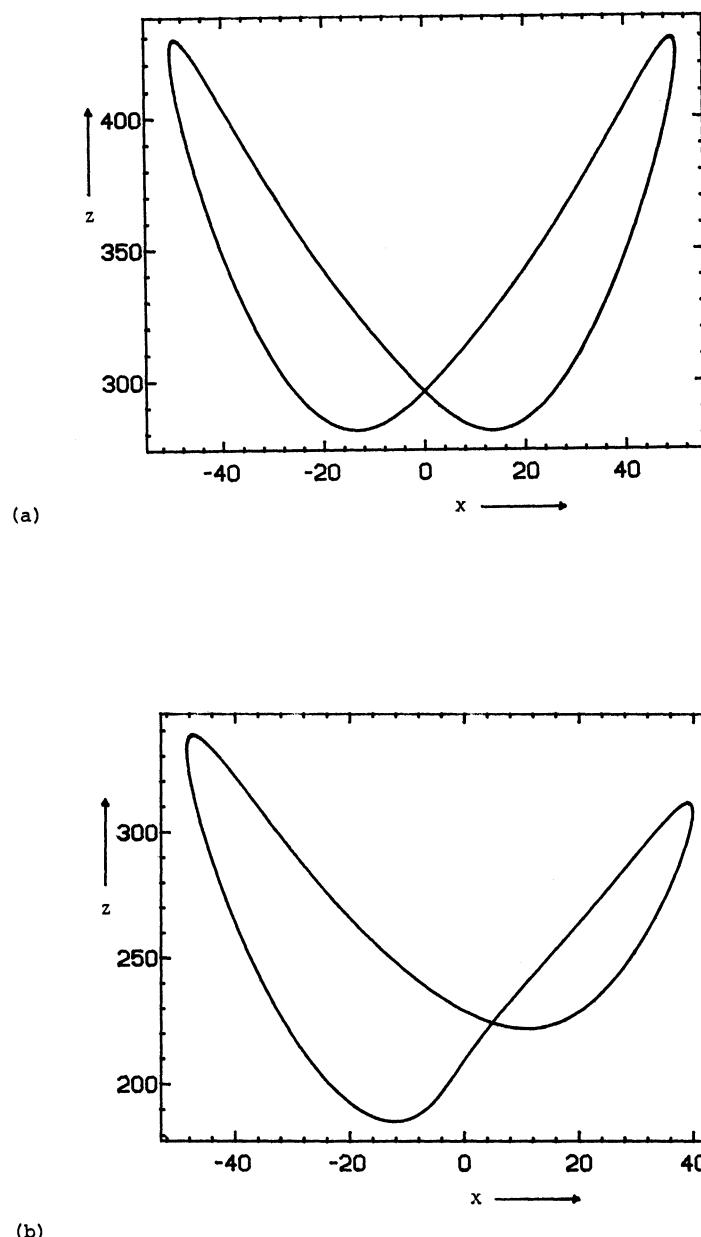


Figure 4.11. (a) Stable symmetric xy orbit. ($r = 350.$)
(b) Stable non-symmetric xy orbit. ($r = 260.$)

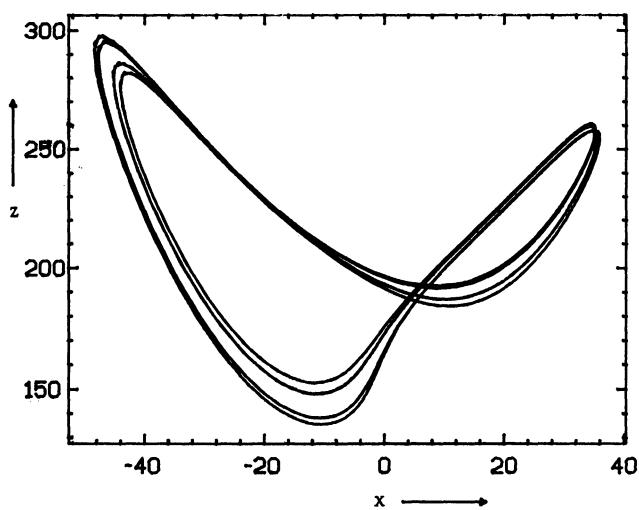
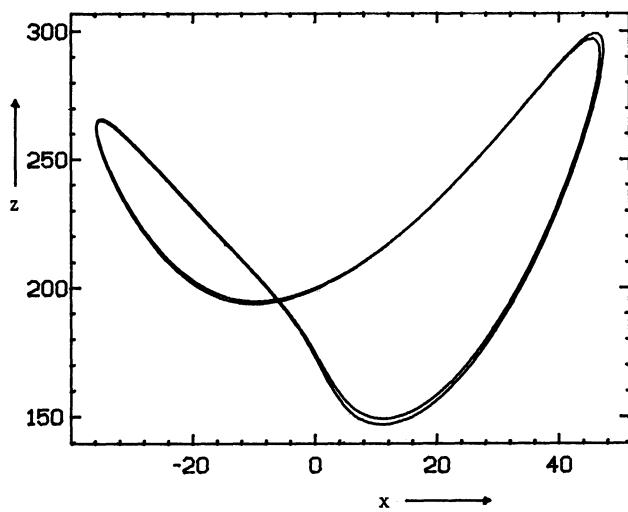


Figure 4.11. (c) Stable non-symmetric $(xy)^2$ orbit. ($r = 222.$)
(d) Stable non-symmetric $(xy)^4$ orbit. ($r = 216.2.$)

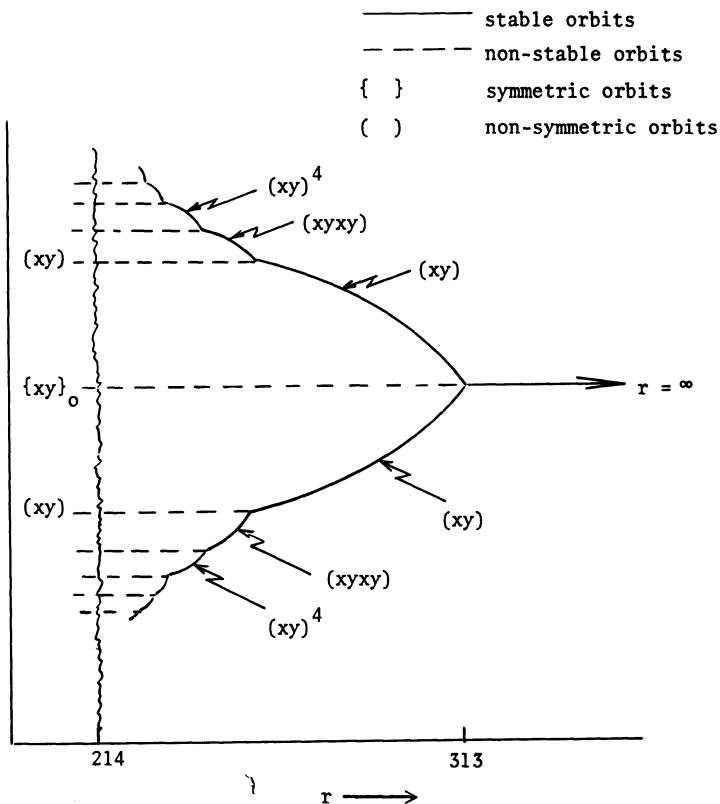


Figure 4.12. Bifurcation diagram for the final xy period doubling window.

Having no upper limit, there is no intermittent chaos associated with this window. But we still want to know what happens at the lower end of the various period doubling windows and shall use this one as an example.

4.6. NOISY PERIODICITY

We would like to know the behaviour of trajectories at, and below, the limit points of the infinite sequences of doubling bifurcations we have observed. Not very much is known about this region; Collet & Eckmann (1980) and Guckenheimer (1979a, 1979b, 1980) describe the situation for one-dimensional maps, but it is not known how many of the results

will carry over to differential equations.

At the accumulation point of the period doubling bifurcation value, r_∞ , we know that there are an infinite number of non-stable periodic orbits "left over" from all the doubling bifurcations. These continue to exist as we decrease r past r_∞ . For r less than r_∞ numerical simulation of the equations produces results such as those shown in Fig. 4.13. The motion lies quite close to the non-stable periodic orbits and Hénon & Pomeau (1976) and Ibañez & Pomeau (1978) both suggest that what we have is a strange attractor which is in some sense built up of "small instabilities" about the non-stable periodic orbit. This description is interesting, but possibly misleading, since we do not actually know the details of the behaviour in this region. In the first paper mentioned above (Hénon & Pomeau, 1976), a study of a local return map near to the non-stable periodic orbits motivated the study of a model return map; this map, now known as the Hénon map, is little understood despite being much studied.

We can describe some properties of the flow in $r < r_\infty$. Our description follows that of Lorenz (1980a) who has studied the parameter range $197.6 < r < 215.364$ in some detail. As we decrease r past r_∞ there is a sequence of "bifurcations" which mirrors the sequence that occurred in $r > r_\infty$. To understand this we need to define *semi-periodicity*. Lorenz (1979 & 1980a) defines semi-periodicity in the context of the Lorenz equations, but he does so from a one-dimensional map which he extracts, as an approximation, from the equations. We can presumably extend this definition to return maps in two dimensions without difficulty. Suppose that for some suitable return plane we can locate n non-overlapping connected regions $E_1, E_2, E_3, \dots, E_n$ on the return plane such that all trajectories eventually pass through these regions in cyclic order. Then we say the system is semi-periodic with period n . (Here n is an integer; we are using period in the sense of "number of intersections with a return plane", rather than "time spent going round the orbit".)

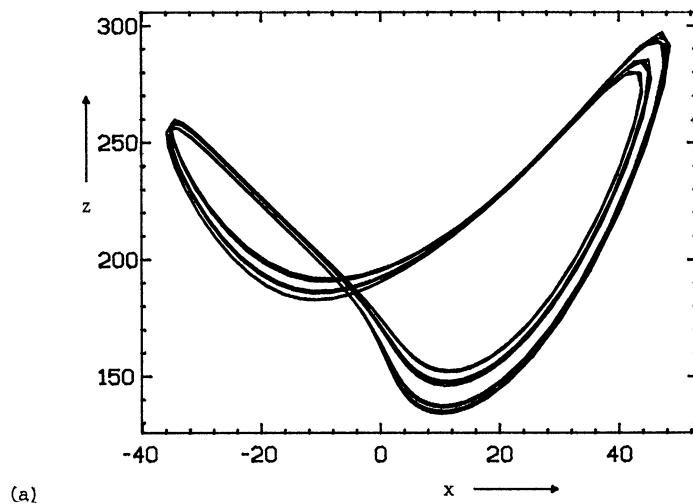
As r increases towards r_∞ , the system becomes semi-periodic with periods $2, 4, 8, \dots, 2^n, 2^{n+1}, \dots$ etc. at parameter values which show the same limiting behaviour we saw for the sequence of period doubling bifurcations which occurred as r decreased to r_∞ . So, for r -values just less than r_∞ the system is semi-periodic with period 2^n for some large n ; i.e., the attracting set is contained within a tube that surrounds the non-stable periodic orbit of period 2^n . (The symmetry implies there will actually be two tubes, one around each of the non-symmetric

orbits of period 2^n . Trajectories will go to one or other of these two tubes eventually.) The tube intersects some suitable return plane in 2^n non-overlapping regions. As r decreases the regions increase in diameter and begin to overlap in pairs. We are then left, for this lower- r -value, with semi-periodicity of period 2^{n-1} , and all trajectories eventually enter a slightly fatter tube around the non-stable periodic orbit of period 2^{n-1} .

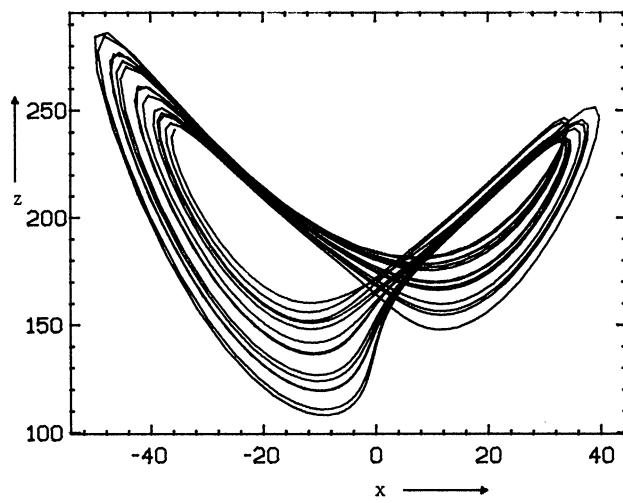
Fig. 4.13(a) shows the behaviour when $r = 212.0$. It looks almost like stable $(xy)^4$ periodic behaviour, though we know the periodic orbit in question lost its stability in a period doubling bifurcation at $r \approx 215.97$. The tubes here are very thin. Fig. 4.13(b) shows the behaviour at $r = 198.0$, just before the final change. The behaviour is xy semi-periodic, and we have two separated non-symmetric regions (only one of which is shown) in each of which we see non-symmetric chaotic behaviour. Fig. 4.13(c) shows the behaviour when $r = 197.4$. The two non-symmetric tubes have expanded sufficiently to overlap and we now have symmetric chaos. This final transition, as the two tubes grow large enough to include the non-stable symmetric xy orbit (left over from the symmetric saddle-node bifurcation at $r \approx 313$) marks the end of the semi-periodic parameter range associated with the final xy period doubling window. Lorenz (1980a) gives a table containing the parameter values at which the various changes, both above and below r_∞ , take place.

The definition of semi-periodicity we have given is one dependent on return maps. Lorenz (1980a) defines another kind of semi-periodicity which depends on the behaviour of the flow with time. He says a flow is *spectrally semi-periodic* if the spectrum of the output is a broad band with superimposed lines. It seems that the two definitions match up fairly well in the situation we have just described though there are problems. See Lorenz (1980a) for a discussion. The expression "noisy periodicity" comes from the title of Lorenz's (1980a) paper.

Semi- or noisy periodicity can also be observed in intervals below the other period doubling windows described in this chapter. Fig. 4.14 shows the behaviour when $r = 145.0$, just below the x^2y^2 window.



(a)



(b)

Figure 4.13. (a) Noisy periodicity. ($r = 212.$)
(b) Noisy periodicity. ($r = 198.$)

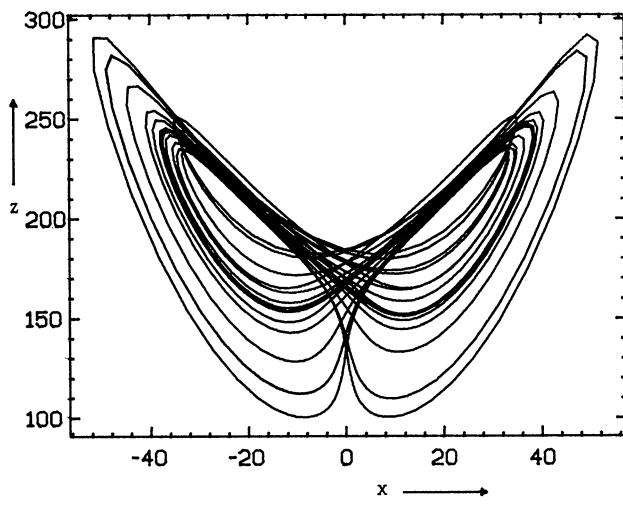


Figure 4.13. (c) Chaos. ($r = 197.4.$)

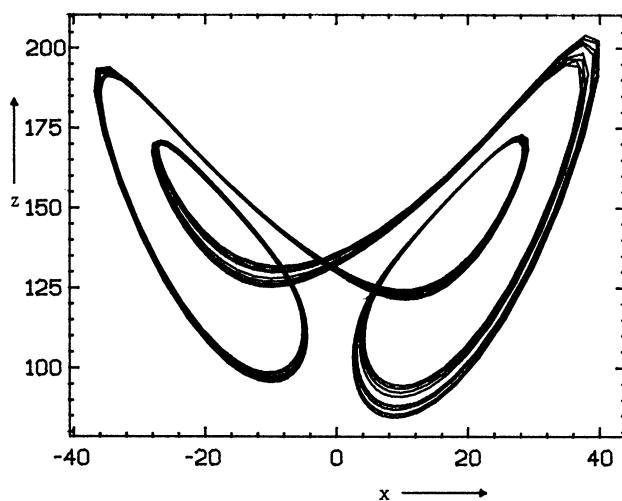


Figure 4.14. Noisy periodicity below the x^2y^2 window. ($r = 145.0.$)

4.7. SUMMARY

We have studied the three largest period doubling windows which occur in the Lorenz system for $\sigma = 10$ and $b = 8/3$. Conveniently, these three provide one example each of the three types of period doubling window we are ever likely to see. The x^2y window is an example of a window involving non-symmetric orbits, two of which eventually disappear in simultaneous saddle-node bifurcations. The x^2y^2 window is an example of a window involving a symmetric orbit which eventually disappears in a saddle-node bifurcation. The final xy window is unique in that the stable symmetric orbit, once it appears, persists for all r however large.

We have described some of the qualitative features of the chaotic behaviour which occurs between the period doubling windows. Intermittent chaos is to be found at the top of period doubling windows (r large), noisy periodicity at the bottom (r small). However, we do not know much about the chaotic regions. We have already stated that we will, in Chapter 5, show how to locate some additional stable periodic orbits (and their period doubling windows) lying in the apparently chaotic parameter intervals. But consider the following questions:

1. Are there parameter values for which we have no stable periodic orbits? If so, are these parameter values isolated or are there whole intervals of such values? In either case, what proportion of parameter space is taken up by parameter values at which we have no stable periodic orbit?
2. Are we sure that at any one parameter value there is at most one attracting set (or two if it is a pair of non-symmetric attractors)? Perhaps two or more period doubling windows overlap. Or perhaps stable periodic orbits and strange attractors can exist at the same parameter values?

These questions cannot be answered for the Lorenz equations (at these parameter values), nor for most other systems of chaotic ordinary differential equations. (The behaviour studied in this chapter - alternate parameter intervals of stable periodic and chaotic behaviour - is the kind of behaviour most often observed in systems of chaotic differential equations other than the Lorenz equations.) We will have a little more to say about these questions in Chapter 5, but the study of a set of differential equations is not really the place to try and answer them. (Though answers to these questions are available, or are becoming avail-

able, for many one-dimensional discrete maps of an interval to itself, they are not available for the Hénon map. This is a return map in two dimensions that is given analytically. We will have even less chance of finding answers when studying a system where numerical uncertainty is involved in the computation of a return map.) However, there are some interesting questions we can attempt to answer.

1. What is the qualitative change in the Lorenz equations which permits stable periodic orbits at these larger r -values when there were none at $r = 28.0$?
2. Where do all the periodic orbits involved in the period doubling windows come from as r increases? (Recall that at $r = 28.0$ we know symbolic descriptions for all the periodic orbits in the non-wandering set. In particular, there was at most one periodic orbit corresponding to each doubly infinite, infinitely repeating, sequence of symbols "x" and "y". Providing we assume that there is some topological justification for the symbolic sequences we have used in this chapter - implying that periodic orbits do not change their symbolic sequences as parameters change - we have to explain where the extra orbits we have seen in the period doubling windows come from. Notice, for example, that we see two x^2y orbits, and two symmetric x^2y^2 orbits, when at $r = 28.0$ we expect only one of each. In addition, we have seen non-symmetric x^2y^2 and xy orbits which do not exist at $r = 28.0$. Finally, all of the doubled orbits - x^2yx^2y or $xyxy$ for example - do not exist at $r = 28.0$; the infinite symbol sequences which would represent these orbits are the same sequences which represent the non-doubled x^2y or xy orbit, and, when $r = 28.0$, each sequence represents at most one orbit.)

We will address these two questions in the next chapter. In fact, they will motivate our search for some global understanding of the whole range of behaviours observed in the Lorenz equations for parameter values $\sigma = 10$ and $b = 8/3$.

Chapter 5

From Strange Attractor to Period Doubling

In Chapter 3 we argued that there was a whole range of r -values (near $r = 28.0$) for which the Lorenz equations possessed a strange attractor. We did not expect to find stable periodic orbits for any r -values in this range. In Chapter 4 we studied a very different range of r -values. In this range we did find stable periodic orbits. The purpose of this chapter is to show how the behaviour changes from strange attractor to period doubling windows as r increases. We will first examine the problem by studying return maps. Then we will ask how well we can model the Chapter 4 type behaviour with a one-dimensional discrete map of an interval to itself, and discuss the difficulties of this approach. Finally we shall work towards a global understanding of the Lorenz equations which will be useful when we want to know how the Lorenz equations behave for parameter values other than $\sigma = 10$ and $b = 8/3$, and which shows how strange attractor and period doubling fit together in a more general context.

5.1. HOOKED RETURN MAPS

In Chapter 3 we examined return maps on the plane $z = r-1$ for $r = 28.0$. After making some (apparently reasonable) conjectures about the properties of this return map we were able to conclude that there were no stable periodic orbits at this r -value, or at nearby r -values. The important element in the argument which led to this conclusion was the observation that the return map appeared to stretch all distances in one direction; thus, nearly all trajectories started at close together points would eventually diverge. For there to be stable periodic orbits, the

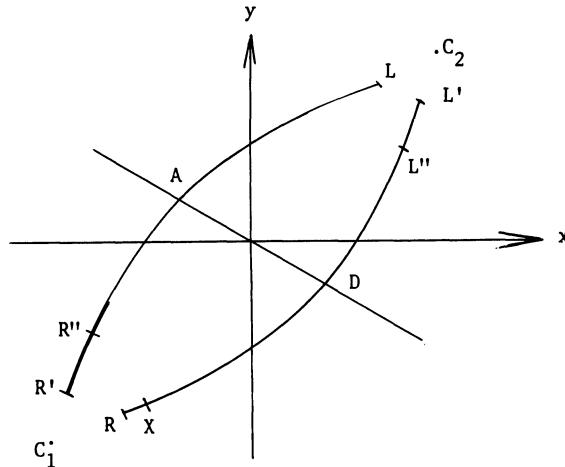


Figure 5.1. Schematic return map for r near 28.0 on the plane $z = r - 1$.

return map, near at least one of the points where the orbit intersects the return plane, must contract all distances. Furthermore, the contraction which occurs near this point must more than compensate for any expansion which occurs at the other points where the orbit intersects the return plane. We shall re-examine the return map at $r = 28.0$ and see how it changes as r increases. Consider Fig. 5.1 which shows the return map schematically.

Fig. 5.1 is virtually identical to the figures of Chapter 3. The labelling of points on the return map is as before; the point R is the point of first intersection of the right-hand branch of the unstable manifold of the origin with the return plane (with $z < 0$), the point R' is the first return of R , etc. The arc AD is part of the intersection of the stable manifold of the origin with the return plane. The return map takes long thin areas surrounding arcs RD , $R'A$, LA and $L'D$ into longer and much thinner areas near arcs $R'L$, $R''L$, $L'R$ and $L''R$ respectively. We expect the return map to develop "smoothly" as r increases, but wish to look for an area in which the stretching (which occurs in the direction "perpendicular" to AD) becomes a contraction; we will

then have a possible area in which stable periodic orbits might intersect the plane. The stretching action of the return map will be felt most strongly for areas of the return plane near AD; the nature of the flow near the origin ensures infinite stretching as we approach AD. Consequently, we shall concentrate on the region furthest from AD. We will look at the region near the point R.

Rather than actually considering points X near R (see Fig. 5.1), we can, by integrating some partial derivatives along with the Lorenz equations, compute the derivative of the map at R. (See Appendix E.) We will be particularly interested in what happens to a small vector \underline{dx} , chosen in the direction \vec{RX} , under the return map. The techniques of Appendix E allow us to compute a small vector $\underline{dy} = \vec{R'X'}$ in terms of \underline{dx} . The results we get are as follows:

1. For $24.06 < r < 30.1$ (approximately) \underline{dy} is in the direction $\vec{R'L}$ as expected. As r increases towards 30.1, the ratio $|\underline{dy}|/|\underline{dx}|$ becomes very small (and may actually tend to zero).
2. Between $r = 30.1$ and $r = 30.2$ (approximately), \underline{dy} swings rapidly but smoothly through 180° in an anticlockwise direction. For $r > 30.2$, \underline{dy} is approximately in the direction $\vec{LR'}$, the opposite of the direction expected.
3. For $r > 30.2$, \underline{dy} retains this direction and the ratio of vector lengths, $|\underline{dy}|/|\underline{dx}|$, increases again.

These results suggest that for $r > 30.2$ we should see a return map like that shown schematically in Fig. 5.2. A hook has been introduced into each of the arcs RL' and LR'.

It is necessary that we confirm Fig. 5.2 with some numerical experiments; these will also allow us to see how the return map develops as r continues to increase beyond 30.2. First, though, notice that the hook gives us an area of the return plane that is contracted very strongly (in all directions) by the return map. We can see this as follows.

Though the nature of the return map has changed near to R, there is no reason to expect that it has also changed near to R'. Thus, we still expect a region near R' to be mapped by the return map into an elongated and much thinner region near R''. Consequently, a small region near PQS (Fig. 5.2) on RL' will be mapped first to a small region near P'Q'S' on the bend in the hook (and the arc length P'Q'S' may well be longer than PQS) and then, by another application of the return map, to a very much smaller region near Q''. (Remember that the contraction in the

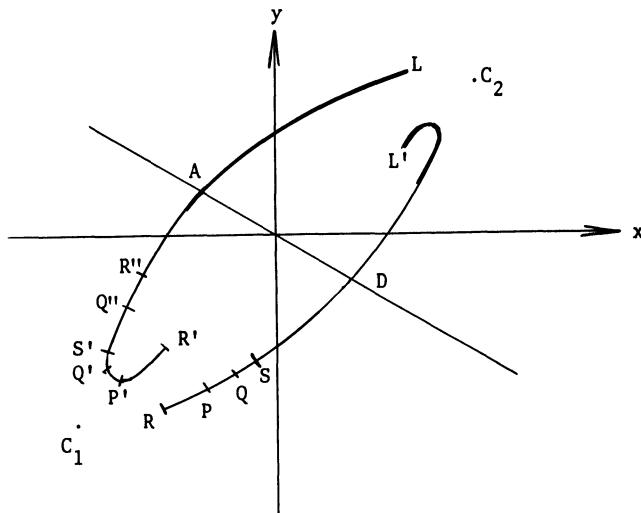


Figure 5.2. Schematic return map for $r > 30.2$ on the plane $z = r-1$.

direction "parallel" to AD is so strong that when we perform numerical experiments we cannot distinguish, after one return, the images of arcs RD and $R'A$; points such as P' , Q' and S' will be mapped extraordinarily close together.)

The contraction associated with the bend in the hook may lead to one (or more) of the periodic orbits becoming stable. This is not certain, for any periodic orbit which intersects the return plane near the bend in the hook also intersects the return plane in several other places; it is likely that the return map continues to stretch distances in one direction at these other intersections and the orbit will only be stable if the contraction (in both directions) which occurs near the bend in the hook more than compensates for the accumulated stretching elsewhere.

5.2. NUMERICAL EXPERIMENTS

It is not easy to observe the hooks directly when $r = 30.2$ and for r a little larger than that something goes wrong with our return plane $z = r-1$. As r increases through a value near 30.5, the unstable manifold of the origin becomes tangential to the plane $z = r-1$ on its first swing around the stationary point C_1 . Consequently, the first point of

intersection that we see with the plane is the point we have been calling R' . This problem does not require us to alter our view of the Lorenz system. We should remember that the system is three-dimensional and that slightly tilted return planes are as likely to be useful as horizontal ones. To stick with the plane $z = r-1$ would introduce a confusing discontinuity into the sequence of pictures of the return map as r varies; consequently we consider an alternative return surface.

There are several possible choices of return surface. One of the best might be the surface S given by $z = xy/b$; this is the surface on which trajectories momentarily move horizontally ($\dot{z} = 0$) and we can be reasonably sure that no interesting trajectories will become tangential to this surface. However, in order to obtain return maps that look as much like the ones we are used to as possible, we will consider a different return surface. We take a surface which looks like an old-fashioned tent. The ridge of the tent is the line $x = -y$, $z = 2(r-1)$, and the two sides of the tent are planes which slope downwards from this line to pass through the stationary points C_1 and C_2 . It seems that no interesting trajectories become tangential to this surface. The situation is illustrated in Fig. 5.3.

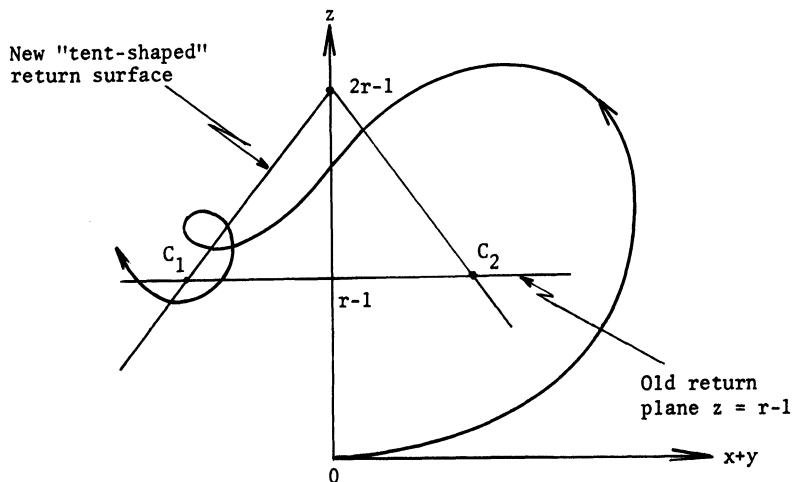


Figure 5.3. $r > 30.5$. The trajectory out from the origin "misses" the plane $z = r-1$ on its first swing around C_1 ; we consider a new return surface.

We restrict our attention (as usual) to points of intersection where trajectories are passing downward through the return surface (into the tent). We might expect the ridge of the tent to introduce some confusing corners into straight lines. Only one interesting line passes over this ridge and that is the section of the stable manifold, AD, of most interest to us. In drawing this line on previous return maps we have only been guessing at its position anyway; providing we know the position of the points A and D (where the stable manifold is near to the non-wandering set) the exact shape of the continuous line which joins them is not important.

We introduce ourselves to the new return surface by checking it at $r = 28.0$. Fig. 5.4(a) shows the intersections with the new surface of the same trajectories that were used to draw Fig. 3.4(a). All the essential features are identical and the ridge of the tent does not intersect anything interesting. As usual, the figure is plotted in the x,y plane; the z -coordinates of the various points can be calculated from the x and y values if required.

We still need to surmount another difficulty before we can see the hooks. As mentioned in Chapter 3, we do not see many points X near R on the return surface by just running a trajectory (since the trajectory has to pass very close to the origin to give us such points) and consequently, we do not see many points on the hook either. To get over this problem we actually calculate the first returns of a number of equally spaced points lying on the straight lines RC_2 and LC_1 . At least for r -values near 30, we expect most of the non-wandering set to intersect the return surface near these lines. In addition, if these initial points lie slightly away from the non-wandering set in the direction "parallel" to AD, the strong contraction in this direction will ensure that the first returns are very near the non-wandering set. Figs. 5.4(b) and (c) show the first returns of these points for r -values of 40 and 60. (It is necessary to start by calculating the first returns of points near R and L. These returns give us the bends in the hooks. When we see how far these hooks are from C_2 and C_1 , we know how far along the lines RC_2 and LC_1 to continue calculating returns. We wish to avoid calculating the first returns of points which lie closer to C_1 and C_2 than any part of the non-wandering set.)

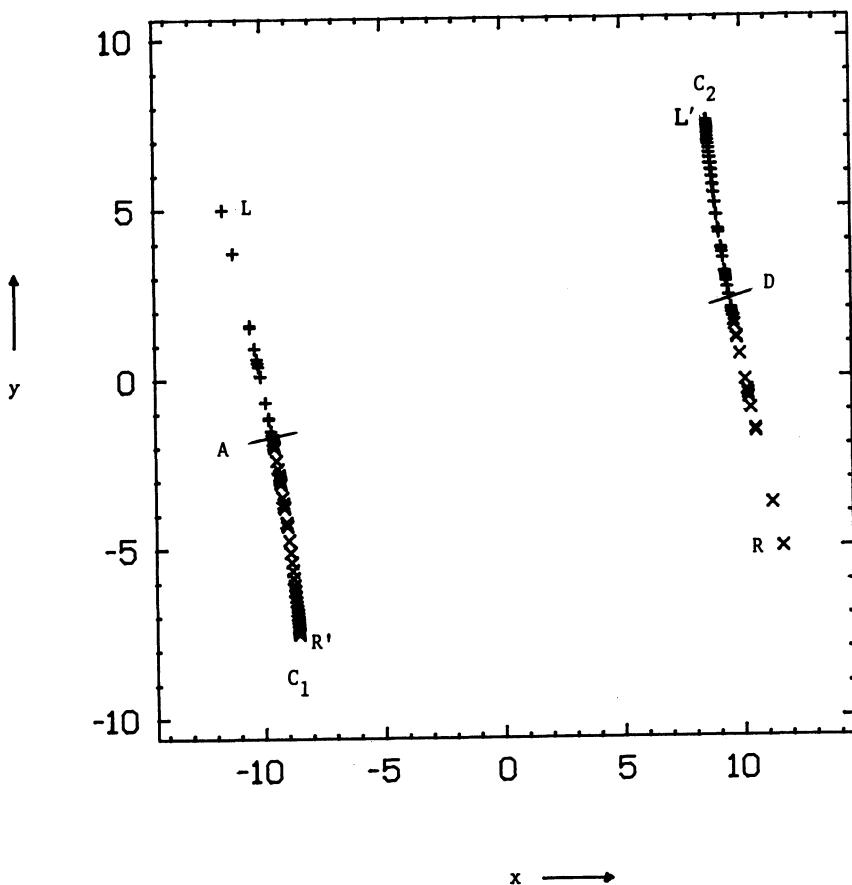
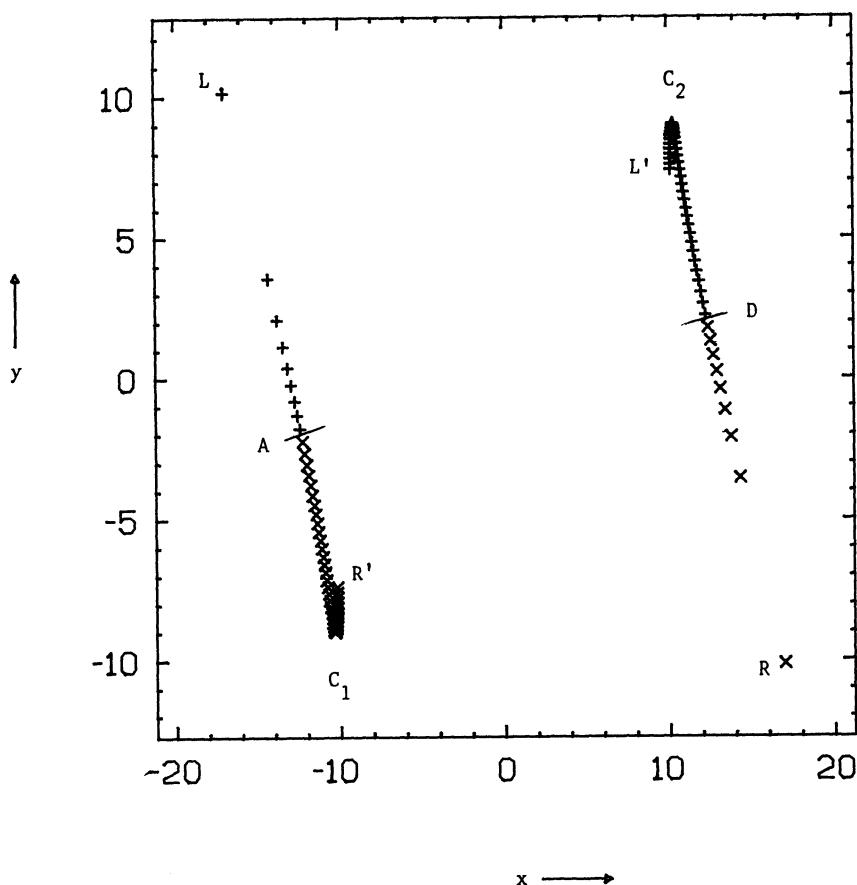


Figure 5.4. Numerically calculated returns on the tent shaped return surface. (a) $r = 28.0$.

Figure 5.4 shows the hooks very clearly. We can also see how the return map develops as r increases. (We are abusing language when we call these figures return maps; as stated above, we are actually calculating the first returns of a series of points on two straight lines. The figures will be useful as long as the returns we calculate continue to lie near the non-wandering set. Now that we are admitting the possibility of stable periodic orbits, we must consider that the existence of a stable periodic orbit would mean there will be "gaps" in the non-wandering set; certainly there will be a neighborhood of each point where the stable periodic orbit intersects the return surface in which no other

Figure 5.4. (b) $r = 40.0$.

parts of the non-wandering set can be found. However, for the r -values used to calculate Figs. 5.4, no stable periodic orbits can be observed - whether or not they exist! - and it seems likely that the points we have calculated all lie near to the non-wandering set, and, conversely, that there are no parts of the non-wandering set which intersect the return surface in points far from the points we have calculated.) We will discuss the development of the return maps in the next section. For the moment, notice the following two points. First, the chaotic behaviour that is observed at r -values of 40 and 60 looks very similar to the behaviour seen when $r = 28.0$, despite the difference in the return maps.

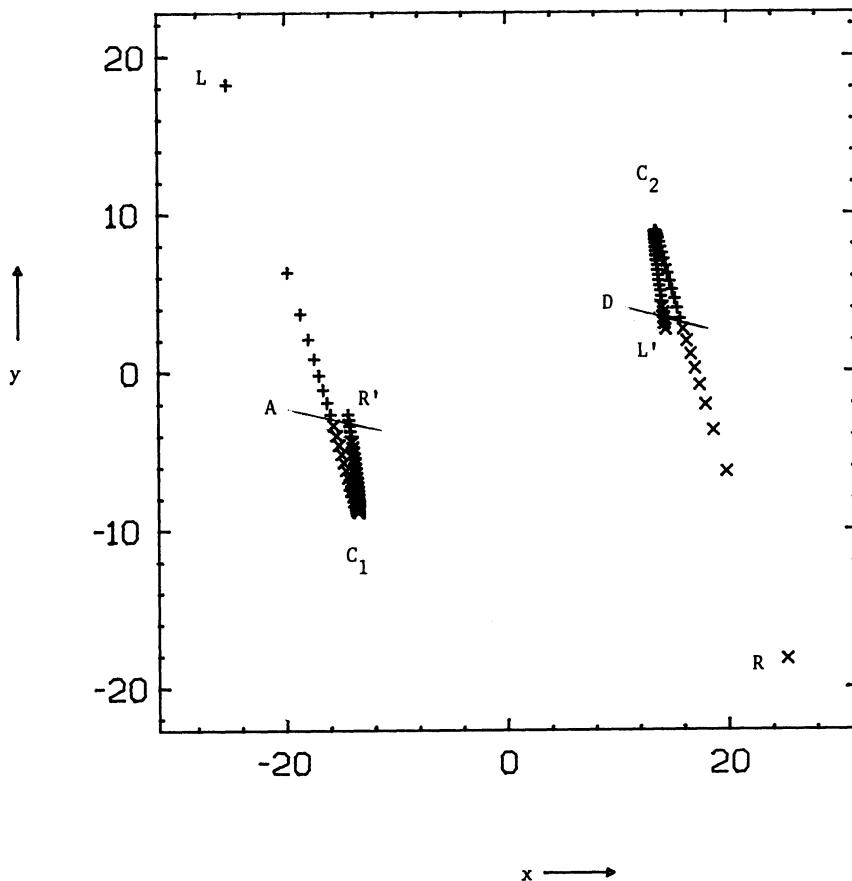


Figure 5.4. (c) $r = 60.0$.

Fig. 5.5 shows chaotic behaviour at $r = 60.0$ (cf. Fig. 1.1). If we plot the points of intersection of this trajectory with our return surface, all the points lie close to points on Fig. 5.4(c), though only one or two of them lie on the hook. Second, we can see that the bend in the hook is connected with the stability of periodic orbits in the following way. Fig. 5.6 shows return points calculated in the same way as Figs. 5.4(b-c), for an r -value of 100.5. The same figure has marked on it the three points of intersection of the x^2y periodic orbit which is stable at this parameter value. We can see that one of the points of intersection does lie on the bend in the hook.

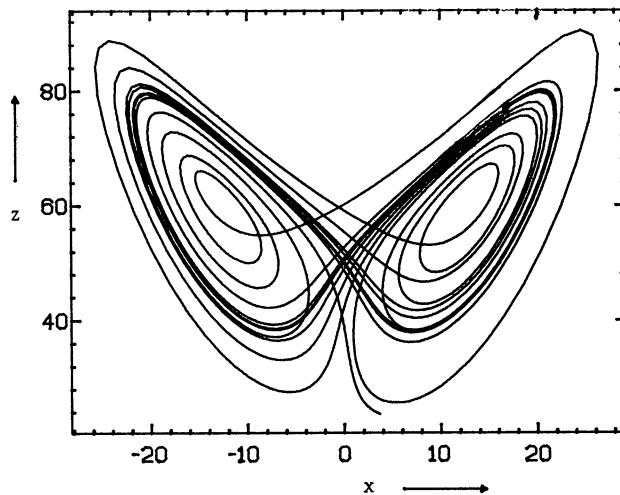


Figure 5.5. Chaotic behaviour for $r = 60.0$.

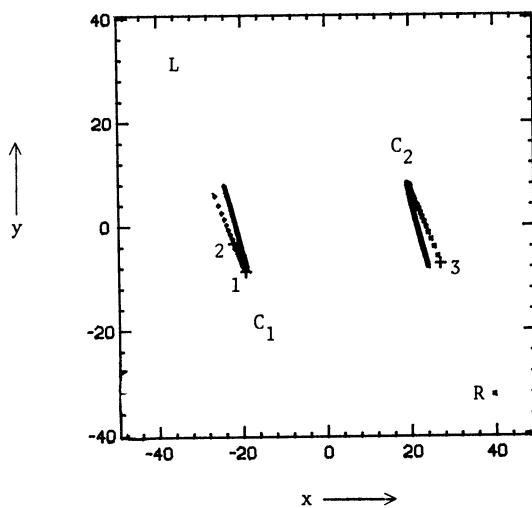


Figure 5.6. The three intersections of the stable x^2y orbit with the return surface, superimposed on the first returns of points lying on C_2R and C_1L . ($r = 100.5$.)

5.3. DEVELOPMENT OF RETURN MAPS AS r INCREASES: HOMOCLINIC EXPLOSIONS AND PERIOD DOUBLING

The numerical results of the last section suggest that as r increases from r -values near 30.2 to larger r -values (we are not sure how much larger yet), we have return maps which develop like those shown schematically in Fig. 5.7.

In each diagram of Fig. 5.7, the four rectangles (labelled A_1-A_4) are taken by the return map into the four shaded areas (labelled $\psi(A_1)$, $\psi(A_2)$, etc.). The thickness of the shaded areas has been exaggerated, and in numerical experiments we cannot see the space between return areas $\psi(A_1)$ and $\psi(A_2)$ or between $\psi(A_3)$ and $\psi(A_4)$. The union of the four rectangles is mapped into itself and so we have some non-wandering set intersecting the return surface within the four rectangles (and hence within the four shaded areas), all or some of which is attracting. The points B and G, the "bends in the hooks", are not well-defined; slightly different choices of the regions A_1-A_4 (such that they still map into themselves) will give slightly different points B and G. We call them points only for convenience.

We can still describe trajectories and orbits in the non-wandering set with sequences of two symbols "x" and "y". We write an "x" every time the trajectory intersects the return surface above AD (on the same side as C_2), and a "y" every time the trajectory intersects the return surface below AD (on the same side as C_1). These symbolic descriptions appear to agree with the symbolic descriptions used in Chapter 4, where we assigned symbolic sequences to orbits according to whether the local maxima in z occurred in $x > 0$ (an "x") or in $x < 0$ (a "y"). This correspondence between the two methods of assigning symbolic sequences is fortunate but not necessarily obvious; the two sequences will only correspond if we have chosen a return surface such that every interesting trajectory intersects the return surface between each local maximum in z , and, conversely, if trajectories reach a local maximum in z between every two intersections with the return surface. We will study symbolic descriptions more rigorously in Chapter 6. Now that we have hooks in the return maps, two or more periodic orbits may have the same symbolic descriptions. We will see this in a moment.

We can describe the development of the return maps with increasing r as follows:

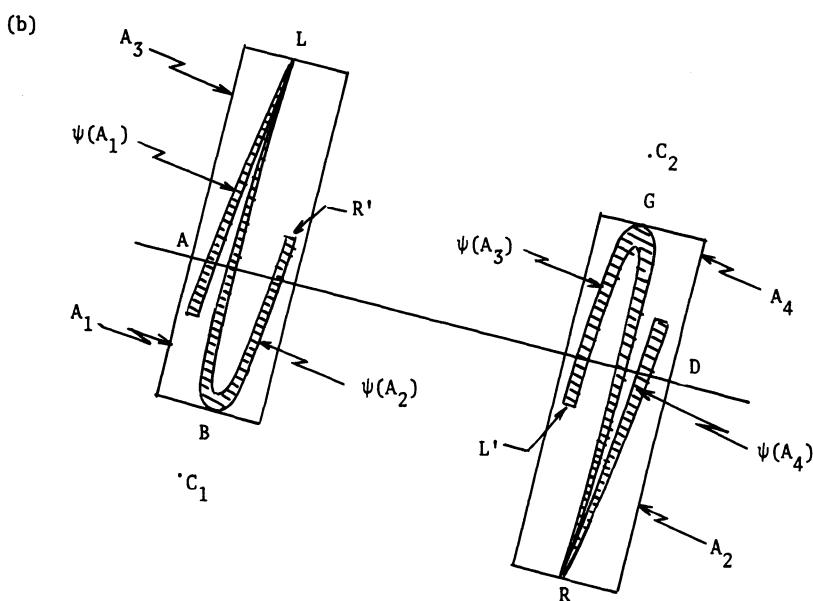
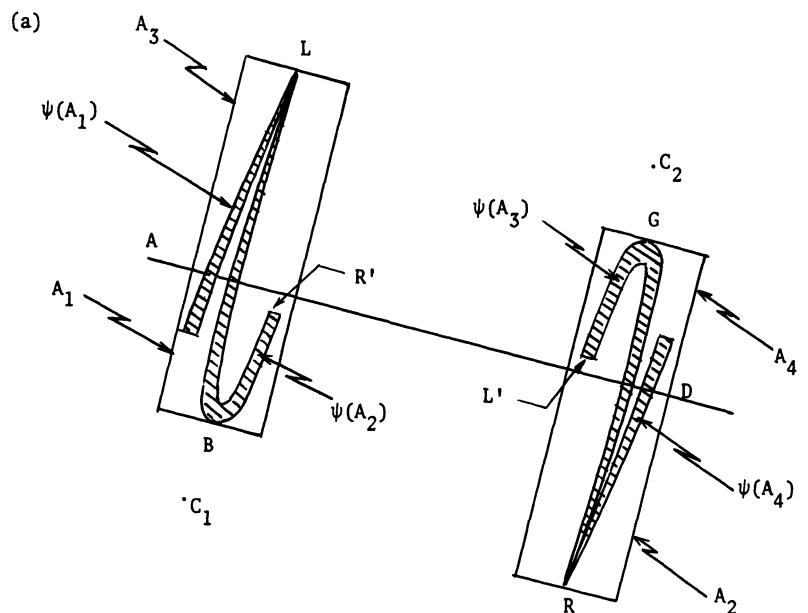


Figure 5.7. Schematic return maps: (a) $30.2 < r < 54.6$. R' is on the same side of AD as C_1 . (b) $54.6 < r$. R' is on the other side of AD from C_1 .

1. As r increases, the point R' (the first return of R) moves towards AD . At $r \approx 30.2$, $\psi^{33}(R)$ is the first iterate of R to lie above AD ; for an r -value near 47.5, $\psi^2(R)$ lies on AD (and we have an xy^2 -homoclinic orbit) at $r \approx 54.6$, R' itself lies on AD (and we have an xy -homoclinic orbit); and, for $r > 54.6$, R' lies above AD . In all of $r > 54.6$, the sequence $k(r)$ - which describes the behaviour of the right-hand branch of the unstable manifold of the origin - starts $xyx\dots$; we get the first x as the manifold swings around C_2 , crosses over the stable manifold of the origin, and intersects the return surface at R ; we get the y as the trajectory swings around C_1 , crosses over the stable manifold of the origin again, and intersects the return surface at R' ; then we get another x , etc.
2. As r increases, the bend in the hook, B , also moves towards AD . At $r \approx 30.2$, $\psi^{32}(B)$ is the first iterate of B to lie above AD (at this r -value, R' and B are very close together); for $r = 50$, $\psi^5(B)$ is the first iterate of B to lie above AD ; and, for some $r > 100$, $\psi(B)$ lies above AD . We can see that B moves towards AD more slowly than R' . This is what we would expect.

The two points R' and B (and their symmetric images, L' and G) are the points which determine the "size" of the non-wandering set. Let us illustrate this by considering periodic orbits of the type xy^2 .

When r was 28.0, there was one and only one xy^2 orbit in the non-wandering set. There is nothing about the appearance of the hooks to make us think this situation should change immediately. For r greater than 30.2 we expect the xy^2 orbit to continue in existence, and if r is not much greater than 30.2 we expect it to be unique. Fig. 5.8(a) shows, schematically, the three points of intersection of this orbit with the return surface when r is a little larger than 30.2. For larger values of r , we may be able to locate an additional xy^2 orbit. This new orbit will have one intersection with the return surface which is close to the "new", bent round part of the return map. See Fig. 5.8(b).

Notice that the arrangement of X 's and O 's on Fig. 5.8(b) is the only one possible. Starting with the points subscripted one, the point X_1 lies closer to R than the point O_1 . Consequently, X_3 must be nearer to AD than O_3 . For this to be true, X_2 must be further from AD than O_2 . Also, because of the relative positions of X_1 and O_1 , X_2 must be closer to R' (measured along the arc $R'BL$) than O_2 .

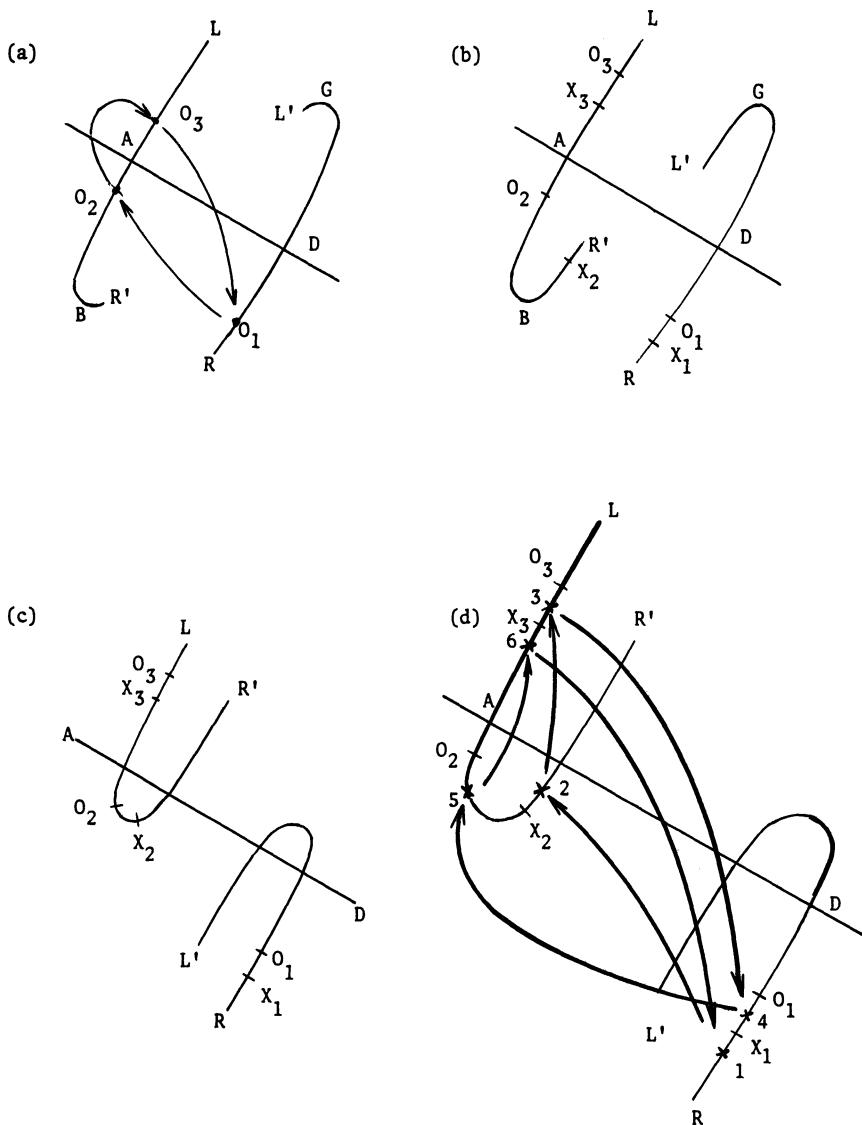


Figure 5.8. (a) The original x^2y orbit intersects the return surface in three points, none of which lie on the bent round part of the return map (c.f. Fig. 5.7). (b) A new x^2y orbit may intersect the return surface with one point on the bent round part of the return map. (c) As the bend in the hook moves towards AD the two x^2y orbits get closer together and the new one may become stable. (d) A doubled orbit may exist near the new x^2y orbit.

We can ask at what parameter value the new "X" orbit arrives. Notice that if $\psi(R')$ lies below AD, then $\psi(X_2)$ also lies below AD; this is not compatible with Fig. 5.8(b). We can argue that we can only have Fig. 5.8(b) for r-values above 47.5; this is approximately the r-value at which $\psi(R')$ crosses AD and we have an xy^2 -homoclinic orbit. Hence, we can argue that the new xy^2 orbit is born in the xy^2 -homoclinic explosion at $r \approx 47.5$. We certainly expect the xy^2 -explosion to either produce or destroy an xy^2 -generated strange invariant set. This set will contain an xy^2 periodic orbit as one of its simplest members.

We can check that the xy^2 -explosion produces a strange invariant set without reference to our schematic (and possibly misleading) return maps. Our analysis of homoclinic explosions did not depend on a knowledge of return maps on surfaces far from the origin. We can first check that there really is a homoclinic explosion. The sequence $k(r)$ changes from $xxxx\dots$ to $xyyx\dots$ at an r-value near 47.5, so we can argue the existence of a homoclinic orbit in the same way that we argued the existence of the first homoclinic orbit in Chapter 2. We can then check the type of the explosion using the techniques described in Chapter 2. It is type (b) and it is proceeding in the direction that produces a strange invariant set.

Remember that in Chapter 3 we had arguments that all homoclinic explosions would be type (a). The existence of the hooked return maps now invalidates those arguments. Recall (Chapter 2) that a homoclinic explosion is type (a) or (b) depending on whether a trajectory near the right-hand branch of the unstable manifold or the origin (which trajectory can be assumed to be starting on the return plane just to the right of AD), after following the unstable manifold of the origin around on its homoclinic course, finally returns to the right or to the left of AD. Before the bend in the hook appeared, the return map preserved orientation in the direction perpendicular to AD; hence, all explosions were type (a). With the bend in the hook, each time the two trajectories we are considering (the unstable manifold of the origin and a nearby trajectory) intersect the return surface on the new bent round part of the return map, their relative orientation perpendicular to AD is reversed. Hence, explosions will be type (b) if the number of times that the unstable manifold of the origin intersects the new bent round parts of the return maps in the course of its homoclinic orbit is odd. Since R' lies on the bent round part of the return map, all homoclinic trajectories in $r > 30.1$ have at least one intersection with the return surface that lies

on the new bent round part of the return map. Most of the shorter homoclinic orbits which we observe will have only this one "orientation reversing" intersection and, hence, will be type (b). However, there will be type (a) explosions whenever a homoclinic orbit intersects the return surface two, or a larger even number of times, in the bent round sections. There is also the possibility that the homoclinic orbit will actually pass through the bend in the hook; in this case, the explosion may be neither type (a) nor type (b). This event is extremely unlikely and we shall not discuss it now. The problem is mentioned again (in passing) in Appendix J.

We have, then, illustrated the fact that as R' moves with increasing r , we obtain a series of homoclinic explosions. These may be either type (a) or type (b). Appendix I contains a list of many of the shorter homoclinic explosions which occur as r increases, together with some extra information needed later in this chapter.

The other "point" which plays a big part in determining the size of the non-wandering set is the ill-defined point B . Referring again to Fig. 5.8 and to the x^2y orbits, as B moves towards AD there will be less and less room for the two points, X_2 and O_2 , in Fig. 5.8(b). Eventually we will have the situation shown in Fig. 5.8(c). The two orbits move close together, the "X" orbit moves onto the bend in the hook and becomes stable, and for a slightly larger r -value the two orbits annihilate one another in a saddle-node bifurcation. For larger r -values the non-wandering set is too small to contain either orbit and neither exists. (This illustrates the point that not all orbits on the bend in the hook are stable; as the X and O orbits move together on the bend in the hook, the X orbit becomes stable but the O orbit does not.)

Using a similar argument, we can actually understand the period doubling as well. Fig. 5.8(d) shows an xy^2xy^2 orbit near the xy^2 orbit. Note three points.

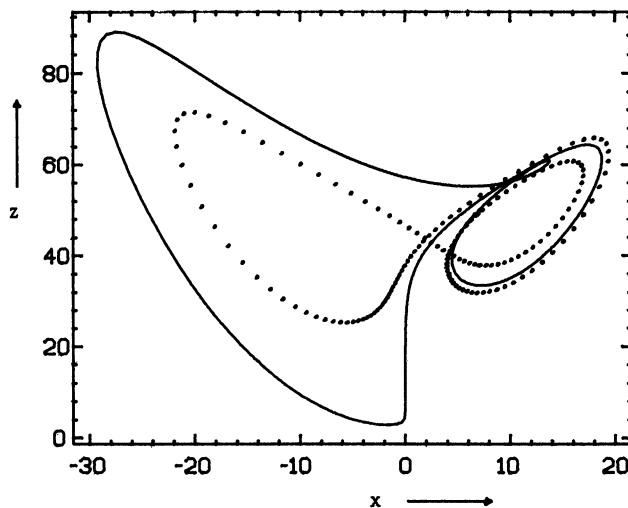
1. The xy^2xy^2 orbit must lie near the extra "X" xy^2 orbit (which has a point of intersection on the new bent round part of the return map) in order to obtain the "flip" necessary for a doubled orbit.
2. The xy^2xy^2 orbit has a point of intersection nearer to B than X_2 . On the assumption that both become stable at some point as r increases, the doubled orbit will become stable first.
3. The xy^2xy^2 orbit has a point of intersection nearer to R' than X_2 . Hence, it will appear (in a homoclinic explosion) at a larger r -value than the "X" orbit.

These arguments, though only intuitive, indicate how homoclinic explosions add the orbits we need for period doubling windows to the non-wandering set, and how the movement of the hook forces the period doubling windows to occur at larger r -values. We can check some of these observations for orbits in the real Lorenz system using the orbit locating and following techniques described in Appendix E.

5.4. NUMERICAL EXPERIMENTS ON PERIODIC ORBITS

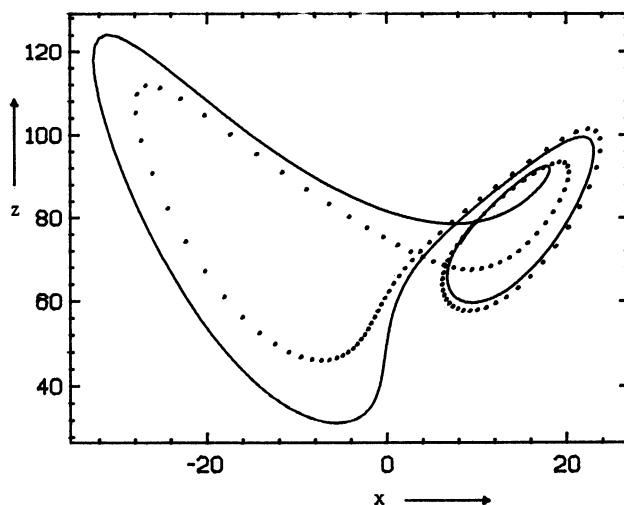
The arguments in the last section suggest that the "original" x^2y orbit (the one born in the first homoclinic explosion at $r \approx 13.926$) is the non-stable partner in the saddle-node bifurcation which marks the upper end (as we increase r) of the x^2y period doubling window that we discussed in Chapter 4. We can locate the orbit at $r = 28.0$ fairly easily. It can then be followed all the way up to $r \approx 100.795$. The orbit following program does indicate that a saddle-node bifurcation is approaching as r nears 100.795. Careful numerical experiments indicate that this orbit, $(x^2y)_0$, that we are following, is not the same as the stable $(x^2y)_e$ orbit which exists just below 100.795. (Subscripts correspond to the labellings of orbits in the bifurcation diagrams of Chapter 4.) The fact that the other orbit is stable means that it is easily located; it can be followed with decreasing r down to an r -value of 50, even though it loses its stability for r just less than 100 (see Appendix E). Fig. 5.9(a) and (b) show both x^2y orbits at $r = 50$ and $r = 80$. The extra orbit is shown in solid lines, the original orbit is shown dotted. Our conjecture that $(x^2y)_e$ was born in a homoclinic explosion for an r -value near 47.5 appears justified. Fig. 5.9(c) shows the y^2xy^2x orbit followed down from its region of stability in the x^2y period doubling window to an r -value of 61.5. Investigation of the behaviour of the unstable manifold of the origin indicates that there is an xy^2xy^2 -homoclinic explosion near $r = 48$. It is not easy to follow the y^2xy^2x orbit into the parameter range $48 < r < 61.5$ for numerical reasons which are explained in Appendix E, but if we could we would expect the orbit to pass closer and closer to the origin as we approached $r \approx 48$.

We can perform similar experiments for other of the periodic orbits observed in the period doubling windows in Chapter 4. The xy -homoclinic explosion (at $r \approx 54.6$, when R' lies on AD) is a type (b) explosion which produces an xy -generated strange invariant set. The two simplest



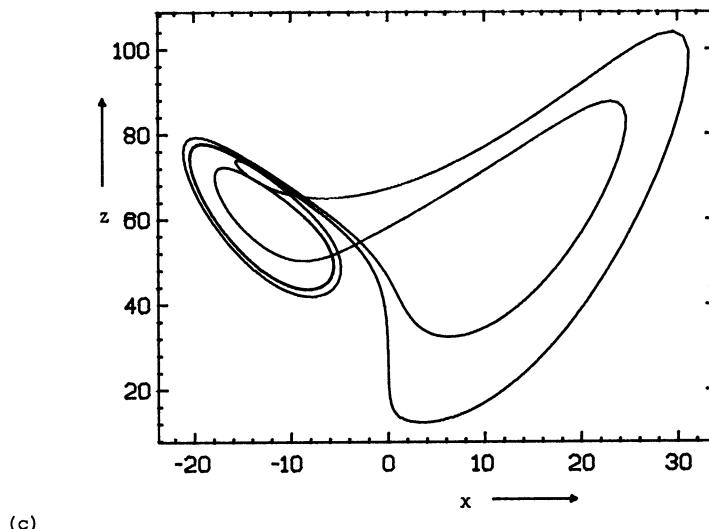
(a)

Figure 5.9. (a) $r = 50$. The two x^2y orbits. The new orbit (unbroken line) passes very close to the origin, as we would expect if it was produced in a homoclinic explosion at $r \approx 47.5$.



(b)

Figure 5.9. (b) $r = 80$. The two x^2y orbits, both still non-stable, move closer together.

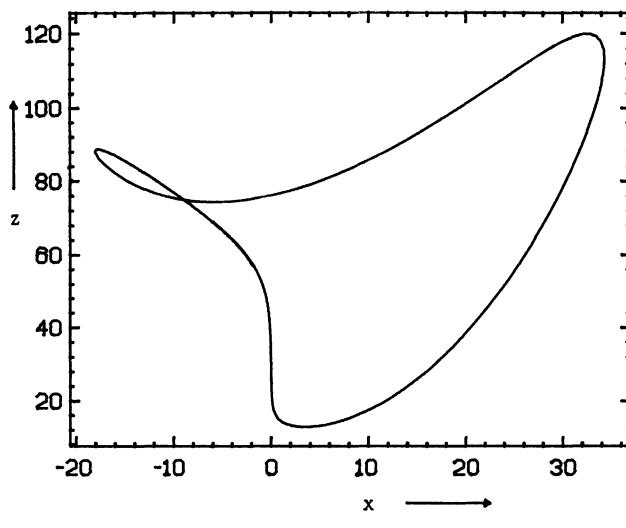


(c)

Figure 5.9. (c) $r = 61.5$. The (non-stable) y^2xyx^2 orbit.

members of this set are a pair of non-symmetric xy and yx ($= xy$) orbits. These are the orbits needed for the final xy period doubling window, and one of them is shown in Fig. 5.10(a) at $r = 70.0$. This same explosion produces, as its simplest symmetric member, an $xyyx$ ($= x^2y^2$) orbit. This is shown in Fig. 5.10(b) at $r = 65.0$. This orbit turns out to be the non-stable symmetric $(x^2y^2)_e$ orbit involved in the saddle-node bifurcation at the top of the x^2y^2 window. The $(x^2y^2)_o$ orbit which is stable at the top of the x^2y^2 window can be followed all the way down to small r -values; this is the original symmetric x^2y^2 orbit. Fig. 5.10(c) and (d) show a non-symmetric $xyxy$ orbit (needed for the xy window and produced in an $xyxy$ -explosion at $r \approx 59.5$) and a non-symmetric x^2y^2 orbit (needed for the x^2y^2 window and produced in an $xyyx$ -explosion at $r \approx 50$). All the orbits shown in Fig. 5.10 pass relatively close to the origin and all are "very non-stable".

Note the complicated way in which the period doubling windows and homoclinic explosions complement one another. Each homoclinic explosion may produce orbits for several different period doubling windows (e.g., the xy -explosion) and each period doubling window involves periodic orbits produced in several different explosions. Also notice that, though each window involves one "original" orbit, it is not always the original



(a)

Figure 5.10. (a) $r = 70.0$. A non-symmetric xy orbit produced in the xy -explosion at $r \approx 54.6$.

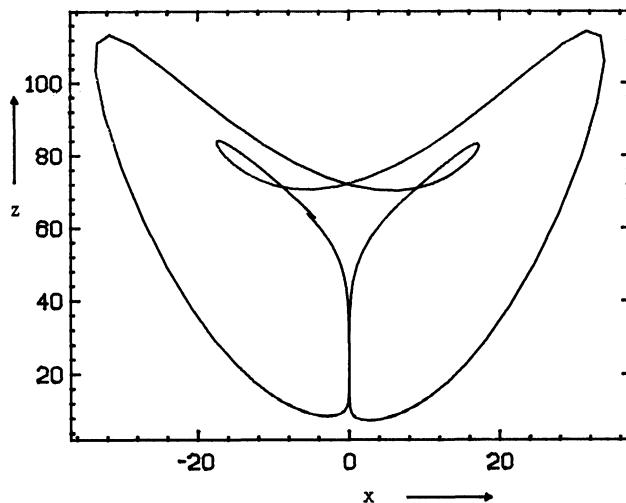
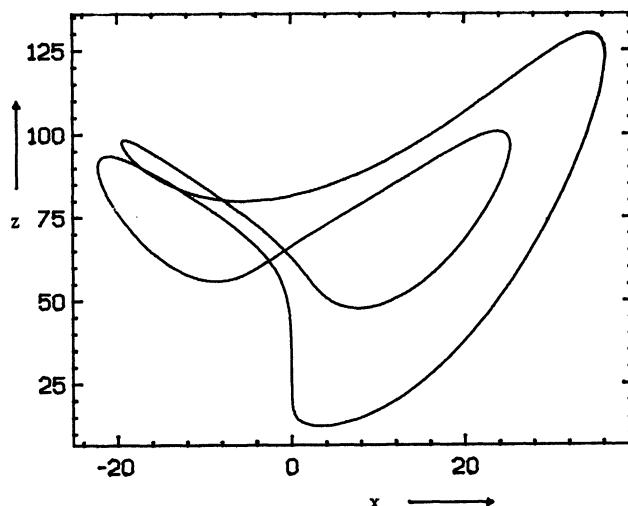
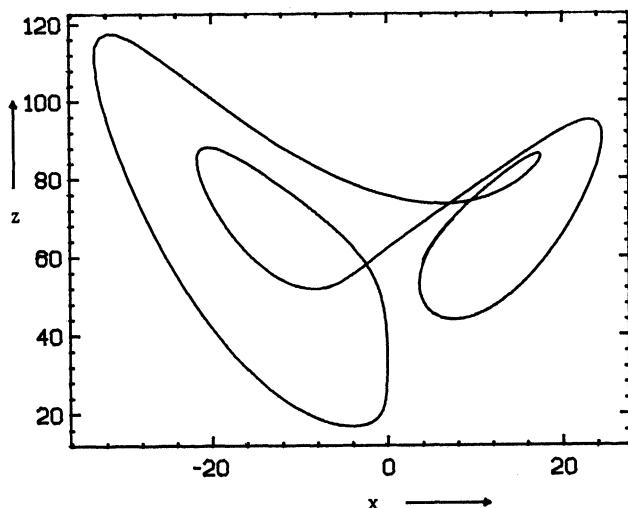


Figure 5.10. (b) $r = 65.0$. A symmetric x^2y^2 orbit produced in the xy -explosion at $r \approx 54.6$.



(c)

Figure 5.10. (c) $r = 75.1$. A non-symmetric $xyxy$ orbit produced in a $xyxy$ explosion at $r \approx 59.5$.



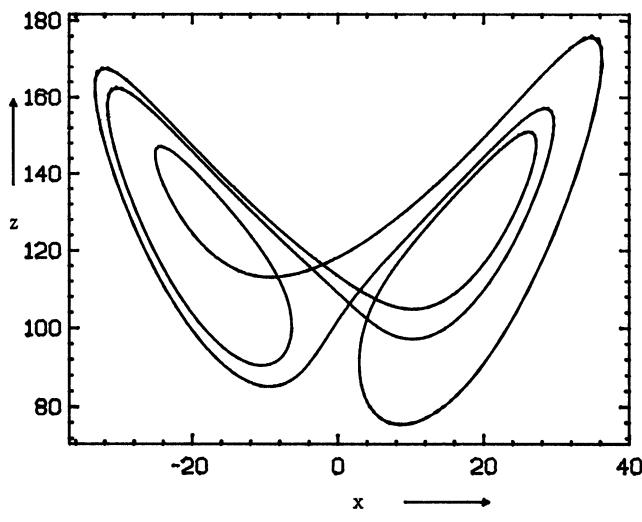
(d)

Figure 5.10. (d) $r = 70.5$. A non-symmetric x^2y^2 orbit produced in a xy^2x -explosion at $r \approx 50$.

orbit which becomes stable. In the x^2y window it is the extra orbit which becomes stable. In the x^2y^2 window it is the original orbit which becomes stable. Attempts to draw a figure like Fig. 5.8 for x^2y^2 orbits will indicate why this should be so.

So far, all the period doubling windows we have seen have involved one (and only one) original periodic orbit. We shall see (in Section 5.6) that there will be some period doubling windows which involve no original orbits; however, our observations show us how to locate other period doubling windows which do involve original orbits. Original orbits can be located at small r values and followed with increasing r . Just before they disappear, either they or a nearby orbit will be stable. Fig. 5.11 shows a stable x^2y^2xy orbit at $r = 126.52$ (the whole of this orbit's window appears to lie in $126.4 < r < 126.53$), a stable x^2yxy orbit at $r = 114.0$ (also a very short window) and a stable symmetric x^2yxy^2xy orbit at $r = 132.5$.

Other period doubling windows can doubtless be found using the same technique. Attempts to locate windows in $r < 99$, where we expect the orbits involved to have more than two consecutive x 's or y 's in their symbolic descriptions, have all failed. The original orbits (such as



(a)

Figure 5.11. (a) A stable x^2y^2xy orbit. ($r = 126.52$.)

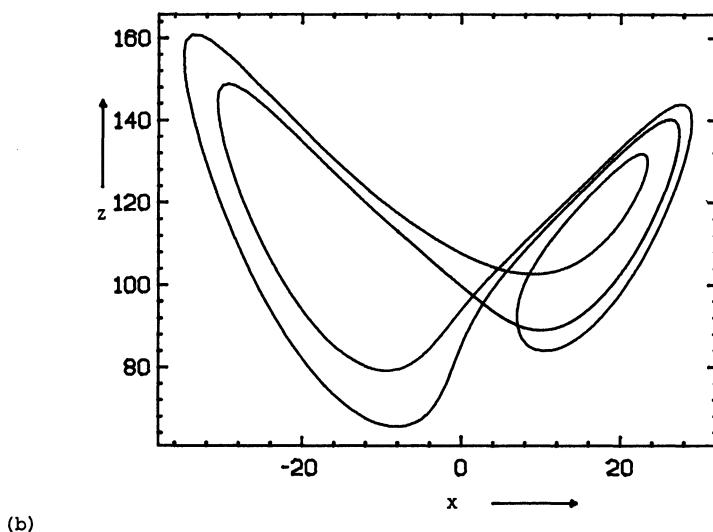
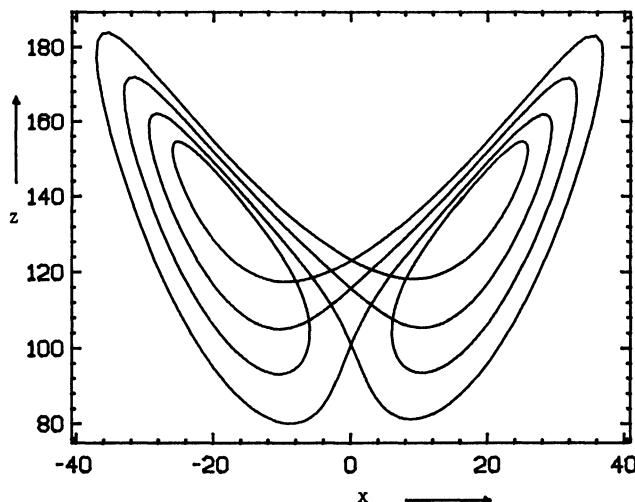


Figure 5.11. (b) A stable x^2yxy orbit. ($r = 114.0.$)

(c)

Figure 5.11. (c) A stable x^2yxy^2xy orbit. ($r = 132.5.$)



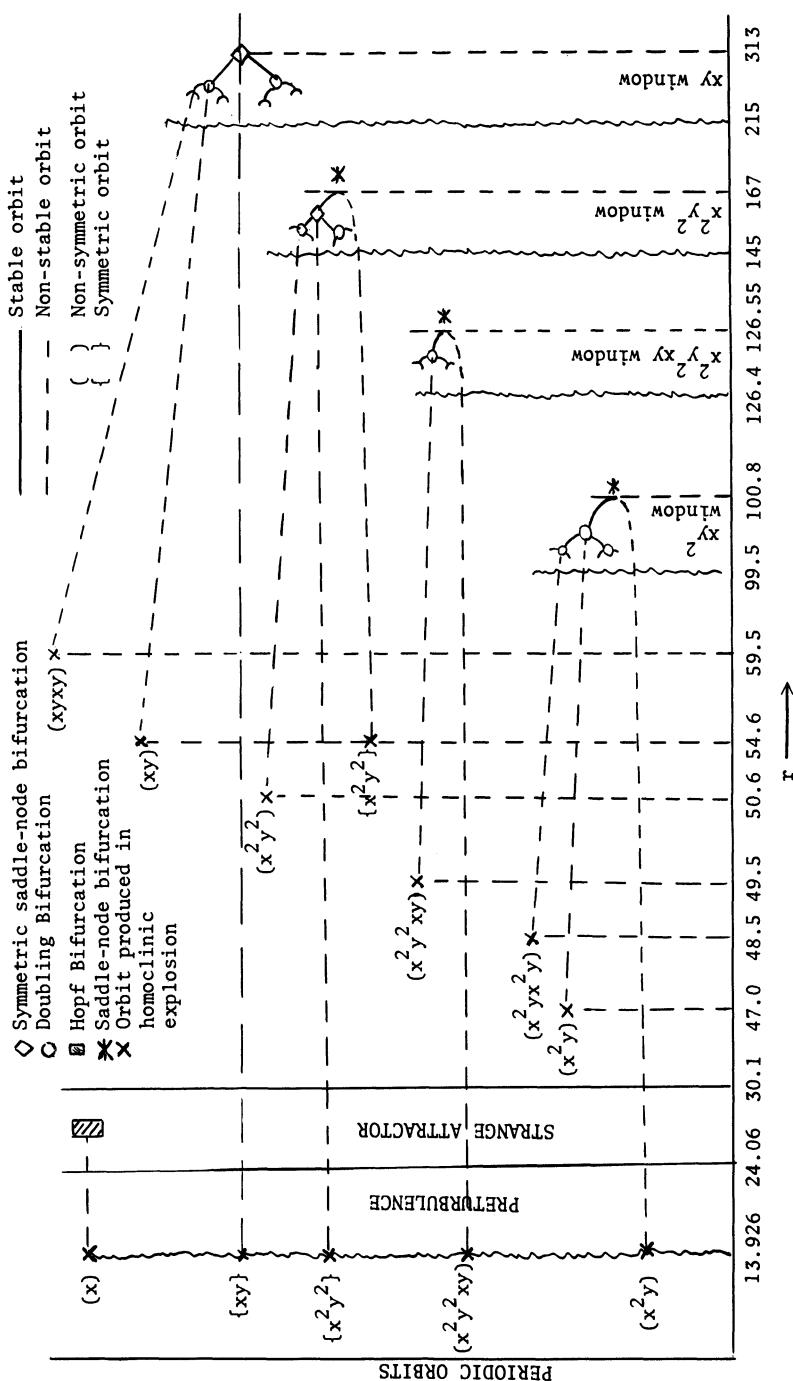


Figure 5.12. Bifurcation diagram showing a few of the shorter periodic orbits.

xy^3 , x^3y^2 , etc.) can be followed with increasing r up to the r -values where we expect them to disappear in period doubling windows. In all cases the orbit locating program indicates that we are approaching a saddle-node bifurcation just before the orbit disappears. However, the period doubling windows must be very short, and no stable periodic orbits have been located.

Fig. 5.12 summarizes some of what we know so far. The happy interplay of homoclinic explosions and period doubling windows motivates our global approach in Section 5.6.

5.5. PERIOD DOUBLING AND ONE-DIMENSIONAL MAPS

The existence of the hooks in the return maps will come as no surprise to students of other systems of chaotic differential equations. Most systems of chaotic ordinary differential equations show period doubling and chaotic behaviour like that described in Chapter 4, and most give rise to return maps like Fig. 5.13.

The return map of Fig. 5.13 is similar, in some ways, to our Lorenz return maps (when we take into account the symmetry of the Lorenz equations). One normal way to analyze these return maps (particularly if the system is very dissipative) is to consider a limit as the height of the rectangle (Fig. 5.13) decreases to zero. In this limit, the return map shrinks to a non-invertible one-dimensional map of an interval to itself. The one-dimensional map derived from Fig. 5.13 will look like Fig. 5.14.

The advantage of this approach is that one-dimensional maps like Fig. 5.14 are becoming very well understood. (See, for example, May (1976), Collet & Eckmann (1980), or Guckenheimer (1979a,b, 1980).) Typical results for maps like Fig. 5.14, providing the map satisfies certain not very restrictive conditions, are:

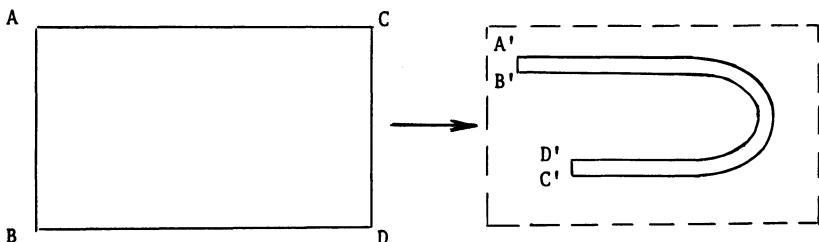


Figure 5.13. A return map frequently occurring in other systems of chaotic differential equations. The rectangle ABCD is mapped, by the flow, into the hooked area A'B'C'D'.

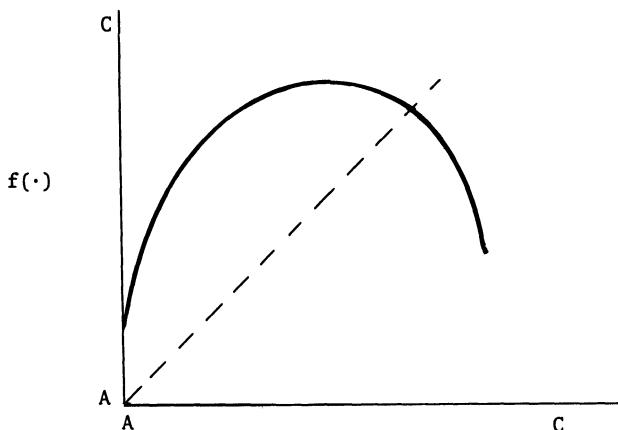


Figure 5.14. A one-dimensional return map derived from Fig. 5.13 as the height of the rectangle decreases to zero.

- i) There is at most one stable periodic orbit at any parameter value.
- ii) We can predict the order of period doubling windows exactly.
- iii) We can make progress with the question, "What is the measure of the set of parameter values for which there is no stable periodic orbit?" It seems likely that in many cases the answer is, "Greater than zero", though we also know that there are no parameter intervals for which no stable periodic orbit exists. (A set containing no intervals can still have measure greater than zero.)
- iv) People are beginning to understand the structure of the attracting set in the case where there is no stable periodic orbit.

One of the main tools used in studying these maps, and producing these results, is the study of symbolic sequences (kneading sequences) generated by following the successive iterates of the maximum of the map, and noting whether each iterate falls to the right or to the left of the maximum. It transpires that the behaviour of this "trajectory", and hence the value of the kneading sequence, determines most aspects of the behaviour of the map. The maximum of the map is a well-defined point - otherwise the analysis would not work. When we try to consider Fig. 5.14 as a perturbation of Fig. 5.14, we run into trouble. The "maximum of the map" is now "the bend in the hook", and this latter is not well-defined. Consequently, we do not know whether any of our detailed results carry

over to the two-dimensional situation or not. As we mentioned in Chapter 4, some aspects of the behaviour near the accumulation point of a series of period doubling bifurcations can be dealt with in several dimensions as well as in one dimension. But that is about all.

Consequently, the study of one-dimensional approximations to our system is not likely to be very useful in situations where we have hooked return maps. Most of the "gross" features of the behaviour which we do expect to carry over from the one-dimensional to the two-dimensional case can be observed nearly as easily in the full system as in the approximation. If anything, what is interesting is to look for differences between the behaviour of the two systems. Regrettably, such differences will be very hard to spot.

It is important to realize the difference between the situation with hooked return maps and the situation described in Chapter 3. There, we had at least the possibility that the extraction of the one-dimensional geometric model of the Lorenz equations could be justified via the existence of a contracting foliation. Attempts to construct contracting foliations for Fig. 5.13, or for our hooked return maps in the Lorenz system, fail. Ignoring the contraction, what is required for a foliation is a continuum of arcs on the return plane, each of which is taken by the return map into another of the arcs. Attempts to fill the rectangle in Fig. 5.13 with a continuum of such arcs come unstuck near the bend in the hook. However we draw them, continued applications of the return map introduce sharper corners into the arcs, and it is fairly easy to see that the conditions for a foliation cannot be fulfilled.

All this said, in Appendix J we do discuss some one-dimensional maps which model most of the observed behaviour of the Lorenz equations very well. Analysis of these maps (which we do not attempt) is more complicated than the analysis of maps like Fig. 5.14, and more complicated than the analysis of the one-dimensional maps from Appendix G. This is because the maps have both turning points (like Fig. 5.14) and a discontinuity (like Appendix G maps), these features modelling the bend in the hook and the special behaviour near the origin, respectively. We use the maps in the Appendix for purposes of emphasis and comparison.

5.6. GLOBAL APPROACH AND SOME CONJECTURES

Our knowledge of the Lorenz equations when $\sigma = 10$ and $b = 8/3$ is now quite detailed. We can describe the picture we have so far as follows:

1. There is a first homoclinic explosion which produces the original strange invariant set. This set is initially non-stable.
2. At an r -value that is approximately 24.06, the original invariant set becomes attracting. The x and y orbits go off to the Hopf bifurcations. Meanwhile an infinite sequence of homoclinic explosions begins.
3. There is an initial phase ($24.06 < r < 30.1$) in which homoclinic explosions remove original periodic orbits from the non-wandering set. These homoclinic explosions are all type (a). In this phase of the development we have a well understood strange attractor (we hope).
4. At some r -value near 30.1, the hooks appear in the return maps. Homoclinic explosions may now be either type (a) or type (b). At least some of the homoclinic explosions which occur add new periodic orbits to the non-wandering set. We have observed that some of the new orbits produced in these explosions are those required for the period doubling windows which occur in $r > 30.1$. We may, if we like, say that in this phase of the development, homoclinic explosions remove original periodic orbits from the non-wandering set in an indirect way; they do so by providing all the periodic orbits needed for a period doubling window which ends with an original periodic orbit being annihilated in a saddle-node bifurcation. In addition, homoclinic explosions produce all the periodic orbits we need for the final xy period doubling window which ends with the original symmetric xy orbit becoming stable and remaining that way forever.

It is clear that this view of the development of the Lorenz equations does not give us a complete understanding. We do not know, for example, anything much about the attracting set in $r > 30.1$ if it is not a periodic orbit. However, we can ask, "Is the scheme outlined above sufficient to explain the appearance and disappearance of all the periodic orbits which exist in the non-wandering set at different parameter values?" (It is certainly necessary that there be some explanation for the disappearance of all the periodic orbits except the final xy orbit, since one of the large r results - see Chapter 7 - is that for large enough r , the xy periodic orbit and the three non-stable stationary points

make up the whole of the non-wandering set.) It seems that the answer to the question is a general "yes", but that we have missed some important aspects of the behaviour. Both these points are argued at some length in Appendix J. We can summarize the arguments here.

1) Homoclinic Explosions and Period Doubling Windows are Enough

We observed, in Chapter 2, that there were rules about the way in which the symbol sequence $k(r)$ - which describes the behaviour of the right-hand branch of the unstable manifold of the origin - changes as r moves through a value at which there is either a type (a) or type (b) homoclinic orbit. Without knowing anything about the dynamics of the Lorenz equations, these rules alone will place combinatorial restrictions on which sequences of homoclinic explosions can occur as r increases. The restrictions will not prevent any particular homoclinic explosion occurring, but they will tell us that the existence of one homoclinic explosion at one r -value implies the existence of certain other homoclinic explosions at other r -values. We know the value of $k(r)$ at low r -values; it is $xxxxx\dots$ with the right-hand branch of the unstable manifold of the origin spiralling into C_2 . We can perform an experiment to determine the value of $k(r)$ at large r -values. (This will require that we believe that our method of symbolically describing trajectories continues to be valid at larger r -values - this matter is investigated in Chapter 6.) We will only need to do one experiment; we know that $k(r)$ must eventually remain the same for all large enough r -values since the large r results tell us that there are no periodic orbits - except xy - at these r -values, and continuing changes in $k(r)$ would imply continuing homoclinic explosions, which, in turn, would imply that periodic orbits continue to pop into and out of existence. Fig. 5.15 shows the unstable manifold of the origin for $r = 10000$. It seems that $k(r) = xyxyxyxy\dots$ with the trajectory eventually being attracted to the symmetric xy orbit and hence continuing to oscillate $\dots xyxyxyx\dots$ forever.

At this point, we can ask a purely combinatorial question. It is, "What is the cumulative effect on the number and type of periodic orbits in the system of any reasonable, combinatorially possible sequence of homoclinic explosions which changes $k(r)$ from $xxxx\dots$ to $xyxyxy\dots$?" The answer will clearly depend, in detail, on the exact sequence of explosions we consider. However, it seems that the following general answer is available.

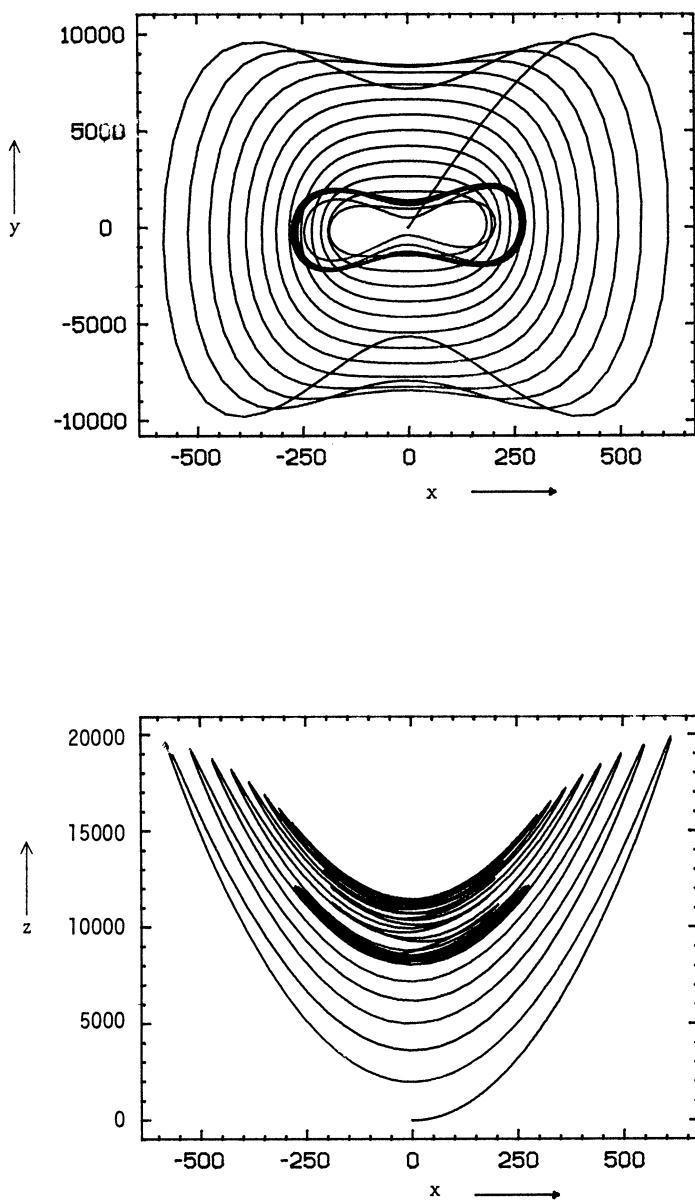


Figure 5.15. The right-hand branch of the unstable manifold of the origin.
 $r = 10000$. Two projections are shown: x against y and x against z .

Conjecture 1. Any non-destructive sequence of homoclinic explosions which changes $k(r)$ from $xxxxx\dots$ to $xyxy\dots$, and which is combinatorially possible, adds the following periodic orbits to the system.

- i) An x and a y orbit.
- ii) Exactly the right number and type of periodic orbits for an infinite number of period doubling windows which terminate in a saddle-node bifurcation. These period doubling windows will be modelled on the xy^2 -window (non-symmetric orbits) or on the x^2y^2 -window (symmetric orbits).
- iii) Exactly the right number and type of periodic orbits for a period doubling window which leaves a stable symmetric xy periodic orbit at the end.
- iv) No other periodic orbits.

Conjecture 1 is discussed in Appendix J. "Non-destructive" sequences are defined to be those which do not attempt to destroy an infinity of periodic orbits not previously produced by other homoclinic explosions in the sequence. The restriction to non-destructive sequences seems dynamically reasonable; though we can imagine a series of non-homoclinic bifurcations producing a finite number of periodic orbits for destruction in a homoclinic explosion, it is not easy to imagine a series of non-homoclinic bifurcations which could produce an infinite number of periodic orbits and then put them together into a strange invariant set suitable for destruction in a homoclinic explosion. It is important to note that this conjecture can be phrased purely combinatorially - without any reference to the Lorenz equations - and that the only aspects of the dynamical behaviour needed to motivate the statement of the conjecture (other than assumptions about the validity of our symbol sequences) is the behaviour near each homoclinic explosion. We should stress, once again, that the analysis of each homoclinic explosion is entirely rigorous and depends on no "model" assumptions about the flow, etc.

Conjecture 1 tells us that once we know the small and large r behaviour, we can already argue, in an intuitive way, that everything which we have seen for $\sigma = 10$ and $b = 8/3$ could have been predicted with very few (if any) numerical experiments. We could not, of course, argue that this is all we will see, though anything else which occurred could not involve any homoclinic behaviour (we have already considered the effect of all homoclinic explosions) and would have to have zero net effect on the number of periodic orbits in the system. What is more important,

however, is that we can argue that what we have seen is the "minimum" that we will see in any symmetrical system with the following properties:

1. A non-stable stationary point lying on the axis of symmetry with eigenvalues which have the same relative sizes that we have for the Lorenz system with $\sigma = 10$ and $b = 8/3$.
2. Some mechanism for describing with two symbols both periodic orbits (in such a way that the description of a periodic orbit does not change with changing parameter) and the behaviour of the one-dimensional unstable manifold of the stationary point (in such a way that the symbol sequence, $k(\cdot)$, so derived only changes at a homoclinic bifurcation).
3. $k(\cdot) = \text{xxxxxx...}$ at some parameter value for which there are no periodic orbits in the system, and $k(\cdot) = \text{xyxyxxyy....}$ at some other parameter value.

The Lorenz system, for a very wide range of parameter values (which almost certainly includes all of $3\sigma > 2b+1$ - see Chapter 7), will have these properties. For a large part of $3\sigma > 2b+1$, the large r behaviour will be the same as for $\sigma = 10$ and $b = 8/3$ and we can expect the whole development of behaviour for $1 < r < \infty$ to be "very similar" to the behaviour at $\sigma = 10$ and $b = 8/3$. In some of $3\sigma > 2b+1$ (loosely speaking, $\frac{\sigma}{b}$ large enough - see Chapter 7), the large r behaviour will be qualitatively more complicated, but there will be some finite r -value for which $k(r) = \text{xyxyxxyy....}$. In Chapter 8 we see that in this situation we can use Conjecture 1 to predict that there is very similar behaviour to that occurring for $\sigma = 10$ and $b = 8/3$, but that it all occurs in a parameter range where the stationary points C_1 and C_2 are still stable, and to predict the existence of a completely new type of bifurcation not seen when $\sigma = 10$ and $b = 8/3$.

Furthermore, there will be many systems which are "not quite the Lorenz system" for which our understanding will be useful, and finally, though we do not attempt it here or in Appendix J, we will be able to extend our understanding to systems without a symmetry.

The foregoing is the sense in which the description we gave at the beginning of this section is sufficient to explain all the arrivals and departures of periodic orbits. We do not need to look (when $\sigma = 10$ and $b = 8/3$) for any bifurcations other than homoclinic explosions, Hopf bifurcation, and period doubling windows to produce a consistent bifurcation picture of all the changes which occur in $1 < r < \infty$. We shall now explain

how our approach leads us to believe that there are some dynamical details of the behaviour which we have not yet observed.

2) Extra Period Doubling Windows

Appendix I contains a list of those homoclinic explosions which occur in the Lorenz system ($\sigma = 10$, $b = 8/3$) that can be detected by following the unstable manifold of the origin until it has generated seven symbols in the sequence $k(r)$. Hence, all the homoclinic orbits involved can be described with six or less symbols. As explained in Appendix I, this list may not be complete. However, of the more than sixty explosions in the list, only fifteen produce periodic orbits which are needed to dispose of original periodic orbits in period doubling windows (remember, we are only considering orbits of six symbols or less). These fifteen explosions are indicated in Appendix I, and all occur in $r < 61$. We can conjecture (see Appendix J):

Conjecture 2. There is some r -value, $r_{xy} < 61$, for which the unstable manifold of the origin is included in the stable manifold of the non-stable symmetric xy orbit (the original and only one). At this r -value, $k(r_{xy}) = xyxyxxyxy\dots$. All the homoclinic explosions required to produce the periodic orbits that are involved in period doubling windows with original orbits occur in $r < r_{xy}$ and all are type (b).

This conjecture is very reasonable in the light of Conjecture 1, and of the fact that arguments on diagrams like Fig. 5.8 indicate that we do expect there to be an r_{xy} -value as described (see Appendix J for more details). However, it leaves us with the question, "What is the effect of all the other homoclinic explosions?" Notice that the explosions continue up to r -values of nearly 500, while the final stable symmetric xy orbit gains its stability at an r -value near 313. Notice, also, that there are "extra" homoclinic explosions in the parameter range $r < 61$. It is unfortunate that it is only really practical to follow $k(r)$ for seven symbols; this only allows us to distinguish between type (a) and type (b) explosions when the homoclinic orbit has three or less symbols. However, we can test each explosion individually (as explained in Chapter 2) and can conjecture:

Conjecture 3. Every homoclinic explosion in the list in Appendix I creates a strange invariant set.

For every explosion tested, this conjecture appears to be correct. Some of the explosions are type (a) and some are type (b), but each is

proceeding in the direction which creates a strange invariant set. In Appendix J we argue that we expect this conjecture to be true. Furthermore, we can also argue that the apparently bizarre order of explosions - which causes $k(r)$ to jump back and forth between the same values - can be predicted. If Conjecture 3 is true (or partially true) we must explain how all these extra orbits disappear as r increases. Conjecture 1 still applies, and tells us that if all the explosions listed in Appendix I create invariant sets, we must also have an infinite number of other explosions (associated with homoclinic orbits which need more than six symbols to describe them) which produce all the orbits necessary for period doubling windows. This leads us to:

Conjecture 4. There is an infinite number of period doubling windows which do not involve original periodic orbits. These start at r -values near 30.1 (as do the windows which do involve original periodic orbits) and continue up to r -values near 500. These extra period doubling windows will not fit into the general sequence of period doubling windows which do involve original orbits; for example, an "extra" x^2y -window can be expected for an r -value greater than 120. One or more "extra" windows may occur concurrently with one another, or with a period doubling window which involves an original orbit.

This last conjecture can be rephrased in a way which may be more familiar to some readers. What it says is that in the period doubling window range ($r > 30.1$), the Lorenz system may have a large number of "horseshoes". The term "horseshoe" comes from Smale's (1967) paper, and is now often used (rather loosely) to describe any situation in which we have a hooked return map like Fig. 5.13. Our conjecture says that the Lorenz system does not just consist of a double horseshoe (which we would expect because of the symmetry) but that the behaviour is qualitatively more complicated than hitherto predicted. Roughly speaking, there is one horseshoe for each n up to some maximum which will change with r ; the n^{th} horseshoe contains orbits intersecting the return surface n times in the bent round part of the return map. For some parameter values, trajectories may wander between horseshoes; for others the behaviour in each horseshoe will be "independent". This latter situation is the one which pertains in $r > 197.6$, for example. In this parameter range the "innermost" horseshoe has crossed the stable manifold of the origin, and all non-transient numerically observed trajectories oscillate $xyxyxy.....$ forever. (See Figs. 4.11 and 4.13.) However, other horseshoes still intersect the stable manifold of the origin, and sufficiently

careful numerical experiments should be capable of locating chaotic and stable periodic behaviour not associated with the final xy period doubling window for r -values all the way up to 500. Fig. 5.16 shows an attempt to clarify this argument. Here, we show a rather more detailed picture of the return map for r -values greater than 54.6 (when R' crosses AD). We have marked six regions (A_1 - A_6) and their first returns. Regions A_2 and A_4 contain the bent round sections of the return map that have crossed AD. These regions are mapped into long thin hooked regions, adjacent to the first returns of regions A_1 and A_5 . The whole non-wandering set should intersect the return surface within the shaded regions of Fig. 5.16, but now we can distinguish between orbits which intersect the return surface at points lying in the different hooked return regions. Original periodic orbits (those which existed before the bends appeared in the return map) all intersect the return surface in the regions between B and L and between G and R. They never intersect the return surface in points in regions A_2 and A_4 and, therefore, none of their points of intersection with the return plane lie in the outer bent return regions. As r increases, and the bend in the hooks moves towards AD, the inner hooked return region will cross the line AD before the outer one does; the final xy period doubling window begins, but there are still periodic orbits and other parts of the non-wandering set intersecting the return plane in the outer hooked region, and these will be involved in period doubling windows, homoclinic explosions, etc., independent of the xy period doubling window which is all that is observed numerically. We can refine this argument by considering two or more applications of the return map; this will show us that there are actually several hooked return areas (lying within the two we have drawn), each of which may be responsible, as the hooks move towards AD, for independent period doubling windows. This is what we meant when we said the "bend in the hook" was not a well-defined point. The phenomenon we are discussing now is a product of the three-dimensional nature of the Lorenz flow; one-dimensional models of the behaviour in this parameter range will not model this phenomenon.

It is interesting that despite the large number of hours that have been spent studying (and computing with) the Lorenz equations, none of this extra behaviour has ever been observed. Some indication of its existence can be seen in Fig. 5.17 which shows the intersections of a chaotic trajectory with our return surface at $r = 180$. Most of the points of intersection lie near a hooked arc like those we saw earlier in

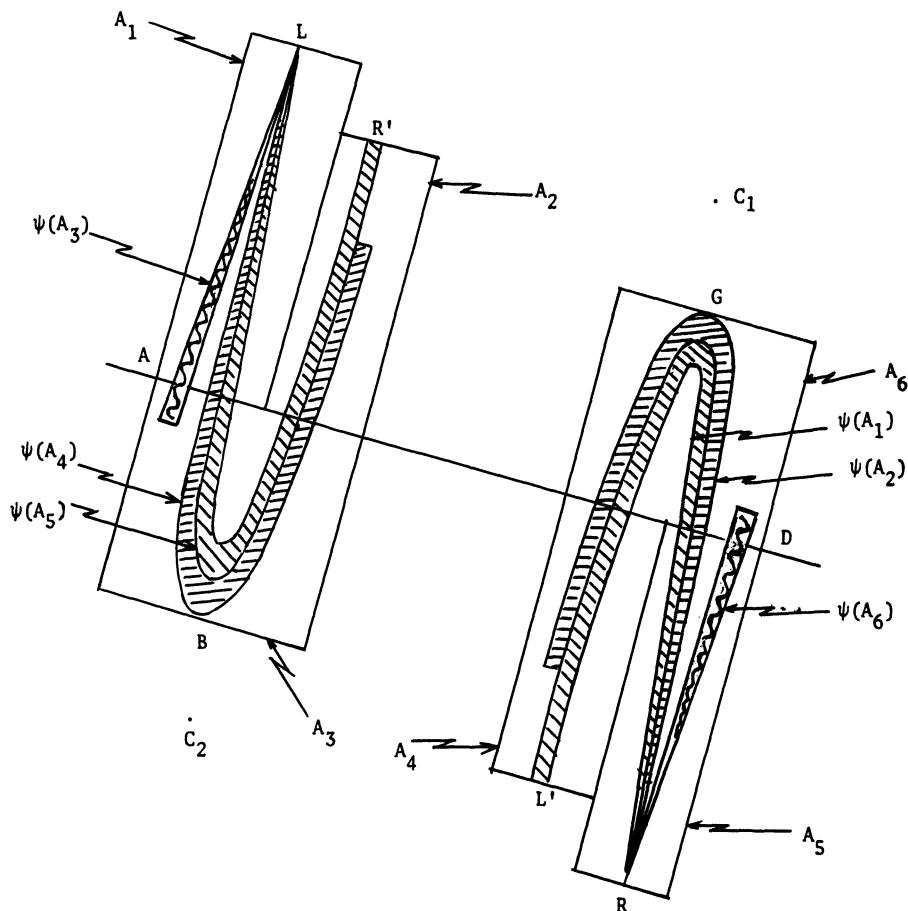


Figure 5.16. A more detailed schematic return map for $r > 54.6$.

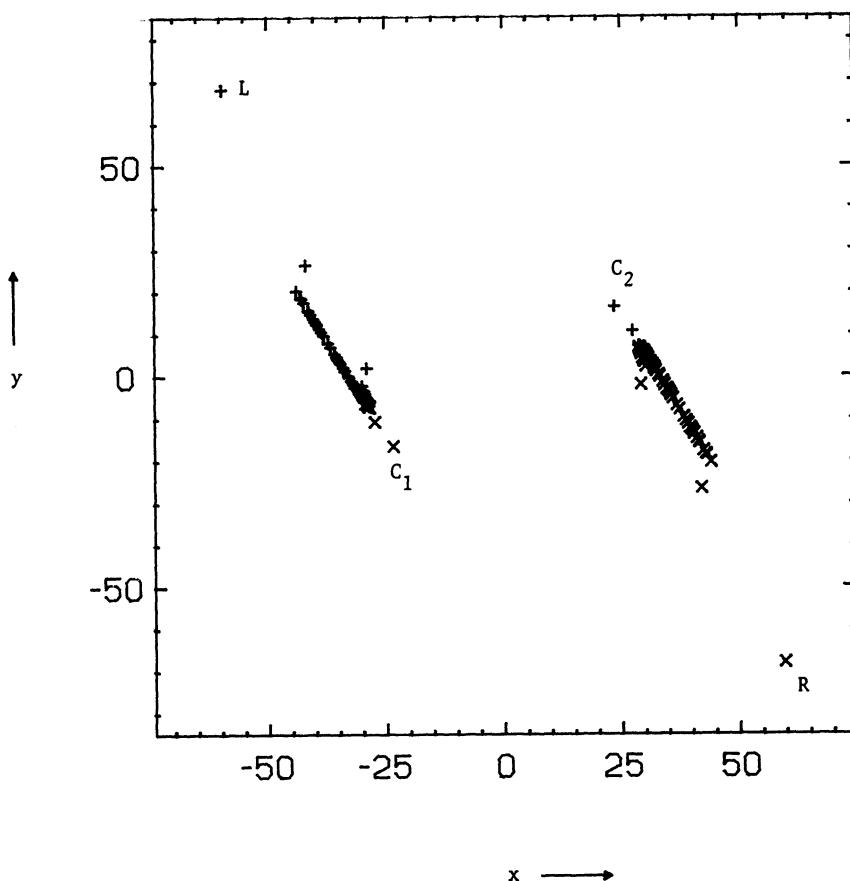


Figure 5.17. Intersections of a trajectory started near the origin with the return surface. $r = 180$.

the chapter. This arc intersects the stable manifold of the origin. However, there are a few points scattered around; these occur after the trajectory passes close to the origin (the trajectory was actually started near the origin which is appropriate since the arc crosses the stable manifold) and lie in some larger, outer "horseshoe".

Attempts to locate some extra behaviour should probably begin with the extra x^2y window. We can see, from Appendix I, that there is a type (b) xyx -explosion at an r -value near 61 and a type (a) yx -explosion at an r -value near 120. Both explosions will produce an x^2y orbit, and at some r -value larger than 120 the two will annihilate one another after an

x^2y period doubling window. It is the orbit produced near $r = 120$ that will become stable.

(Note: Whether or not the conjectures in this section are true, the existence of homoclinic explosions for r -values up to nearly 500 - confirmed fairly easily by numerical experiment - indicate that there is some previously unobserved behaviour to be studied and explained.)

5.7. SUMMARY

We may summarize the results of this chapter as follows.

1. There is a first homoclinic explosion which creates the original strange invariant set.
2. At some larger r -value the unstable manifold of the origin is attracted to this set, and a sequence of homoclinic explosions begins.
3. Initially, all homoclinic explosions are type (a) and all remove periodic orbits from the non-wandering set. In this phase of the development we probably have a strange attractor which is well understood.
4. At some point the hooks appear in the return maps and homoclinic explosions, which may now be either type (a) or type (b), produce periodic orbits which are added to the non-wandering set.
5. It seems that it is necessarily true that exactly the right number and type of periodic orbits are produced so that everything which is left can disappear in period doubling windows, excepting a stable symmetric xy orbit which persists for all r however large, and an x and a y orbit which are involved in the Hopf bifurcations.
(Conjecture 1)
6. Period doubling windows can be divided into two kinds; those that involve periodic orbits that existed in the original invariant set and those that don't.
7. There is one period doubling window for every original periodic orbit which still exists at the r -value where the hooks first appear. These windows occur in a predictable order (the ones involving orbits with the largest number of consecutive x 's or y 's occur first), and the final xy window is over by some r -value near 313.
8. There is an infinite number of period doubling windows which involve no original orbits. These windows may occur at the same time as windows involving original orbits, and one or more of these windows may overlap. In any case, these windows will not occur at the r -values which we would predict from the order of the windows involving ori-

ginal orbits, and they continue up to r -values near 500. None of these windows has ever been observed. They may involve periodic orbits with the same symbolic descriptions as original orbits (we will, for example, have an extra x^2y window in addition to the one which involves the original x^2y orbit), or they may not.

9. We can expect at least the behaviour described in 1 - 8 above for all parameter values σ and b satisfying $3\sigma > 2b+1$. We may, of course, see additional behaviour for some of these parameter values, and the behaviour may be difficult to observe. (For example, in Chapter 8 we see the 1 - 8 behaviour occurring in an r -interval in which C_1 and C_2 are still stable; this makes it difficult to observe!)
10. We can expect variations in the details of the behaviour as σ and b vary. For instance, we can expect the hooks to appear at different points in the development for different parameter values. Orbits which are removed in the type (a) only phase at one value of b , may be removed in a period doubling window at another value of b . Appendix J (Section 2) contains a description of the "homoclinic explosion explosion" which is associated with this change.
11. We have developed a powerful tool with which to approach the Lorenz equations. There are many unanswered questions, but we can now see the whole development in $0 < r < \infty$, for a wide range of parameters σ and b , as part of a single picture.

This almost concludes our study of the Lorenz equations for $\sigma = 10$, $b = 8/3$. In Chapter 6 we discuss symbolic descriptions of orbits - a necessary step if we are to apply our understanding to general values of the parameters σ and b . In Chapter 7 we prove the large r results which we have already used, as well as showing that we can expect qualitatively more complicated behaviour (in addition to the behaviour we have already studied) for some parameter values σ and b . In Chapter 8, we study some of this more complicated behaviour. The arguments from this chapter (Chapter 5) will be crucial to understanding the more complicated behaviour, as well as leading us to a type of bifurcation not so far observed in the Lorenz system.

Chapter 6

Symbolic Description of Orbits: The Stable Manifolds of C_1 and C_2

At various points in the last four chapters we have found it useful to be able to describe periodic orbits and trajectories with sequences of two symbols. When we studied homoclinic explosions in Chapter 2, we saw that we were rigorously justified in describing the periodic orbits and trajectories born (or destroyed) in these bifurcations with sequences of two symbols; each symbol corresponded to one of the two tubes surrounding (at the critical r -value) the two branches of the homoclinic orbit, and trajectories were assigned symbolic sequences according to the order in which they journeyed through these tubes. These descriptions were only local. In Chapters 3 through 5, we found we had need for a global method of describing orbits and trajectories. In Chapters 3 and 5, we defined symbol sequences according to the order in which trajectories intersected two halves of some suitable return surface. In Chapter 4, we assigned symbol sequences to periodic orbits according to whether their successive local maxima in the z -variable lay in $x > 0$ or $x < 0$. These two methods appeared to be equivalent. We chose "suitable" return surfaces by asking that all interesting trajectories should intersect the surface once and only once between local maxima in the variable z , and had to abandon the flat return plane $z = r-1$ when this condition ceased to hold for that return surface; thus, the symbol sequences generated by the two methods had the same length. Furthermore, the two different symbol sequences which we obtained always seemed to agree. We shall see that there are reasons to expect this correspondence when we examine the maxima-in- z method below. However, for the arguments of Chapter 5 to make sense, we require stronger conditions on our symbol sequences. These are:

1. Periodic orbits must retain the same symbol sequence as parameters change.
2. The sequence $k(r)$, which describes the behaviour of the unstable manifold of the origin, must change only at homoclinic explosions.

We have not observed any problems with either of the two points above; however, if we are to make the kind of arguments contained at the end of Chapter 5 for general values of the parameters σ and b , we should examine our symbol descriptions to see if there is any topological or mathematical reason to expect conditions (1) and (2) to hold when the parameters r , σ and b all vary. We examine the maxima-in-z method; similar but more complicated arguments will apply to various return surfaces.

6.1. THE MAXIMA-IN-z METHOD

Every trajectory (except those which terminate in a stationary point) necessarily reaches a series of local maxima in the variable z . These maxima occur either in the half-space $x > 0$ or in the half-space $x < 0$. We described each trajectory with a unique sequence of two symbols by following along the trajectory and writing an 'x' every time the trajectory had a maximum value of z in the half-space $x > 0$ and a 'y' every time it had a maximum value of z in the half-space $x < 0$.

The most obvious advantage of this method of describing orbits and trajectories is that the symbolic sequence assigned to an orbit has an obvious relationship to the way the orbit looks when we project it onto the x,z plane (which is the plane onto which we have projected all our numerically calculated trajectories in these notes). It might be objected that the choice of half-spaces $x > 0$ and $x < 0$ (and of the x,z plane for projections) is an arbitrary one, and that a different choice would have given different symbolic sequences which would have been equally useful. While it is true that different choices might have given different sequences, it is not true that these sequences would have been equally useful. We can see this as follows.

Consider the surface S given by $z = xy/b$. We expect trajectories to intersect this surface whenever $\dot{z} = 0$. The second derivative of z is given by

$$\ddot{z} = b^2 z + rx^2 - (\sigma+b+1)xy + \sigma y^2 - x^2 z. \quad (1)$$

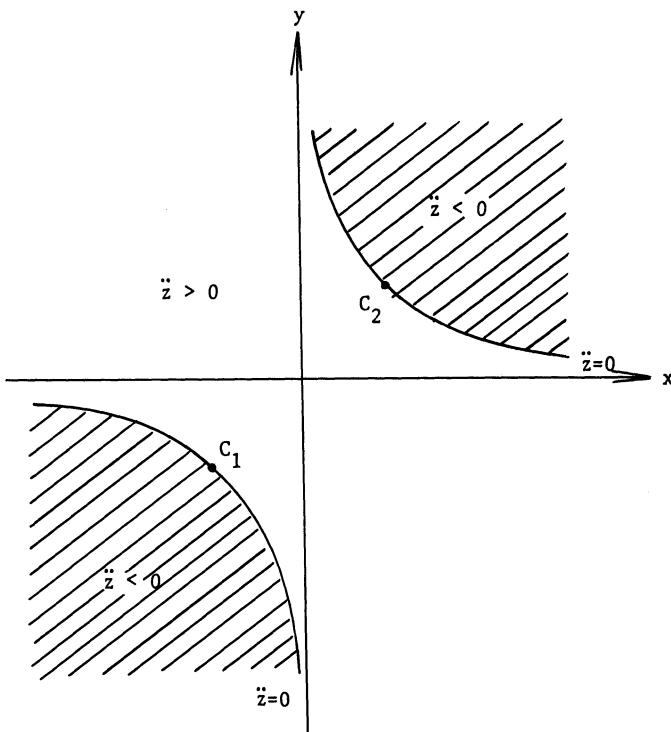


Figure 6.1. The surface $z = xy/b$ plotted in the x,y plane. Trajectories intersect the surface in local maxima within the shaded regions, and in local minima within the unshaded region.

On the surface S ($z = xy/b$) this becomes

$$\ddot{z} = rx^2 - (\sigma+1)xy + \sigma y^2 - b^{-1}x^3y. \quad (2)$$

Providing we assume that $4\sigma r > (\sigma+1)^2$ (which is true for all interesting values of the parameters), the quadratic part of expression (2) is positive definite. Hence, trajectories intersect the surface S at local maxima if and only if the intersections occur in one of the two shaded regions of Fig. 6.1 (where $\ddot{z} < 0$). If a trajectory intersects S in the unshaded region ($\ddot{z} > 0$), we have a local minimum. The two shaded regions are disjoint, and it is natural to match each region with one of the two symbols used to describe trajectories symbolically. Thus, our choice of the plane $x = 0$ for determining the symbolic sequences is one of the natural choices, since this plane separates the two shaded regions of Fig. 6.1. If we follow a periodic orbit as we change a parameter, we do not expect one of its local maxima to wander from the region $x > 0$ to the region $x < 0$, and, in this sense, our symbolic sequences will be robust.

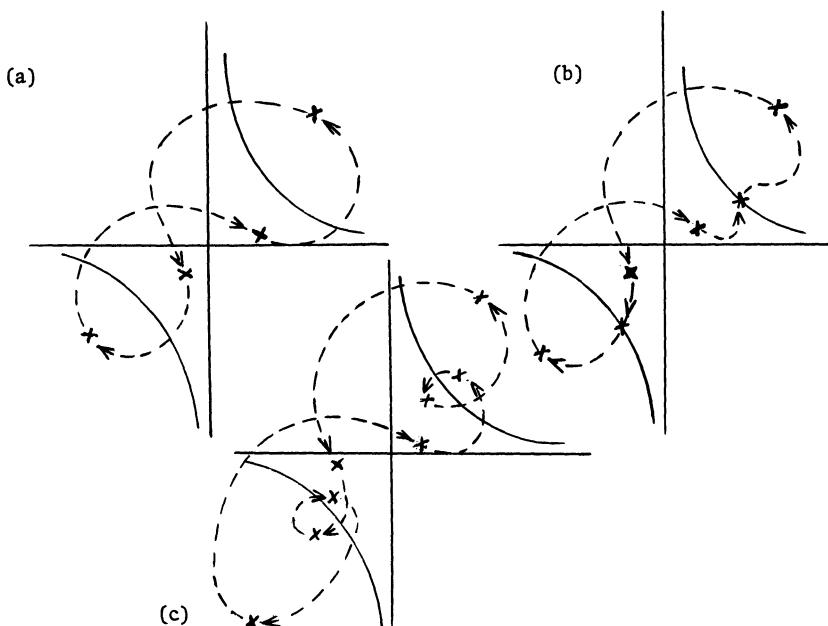


Figure 6.2. An xy orbit which intersects S in 4 points (a) may become tangential to the surface at two additional points lying on the boundaries between the regions (b) and thereafter intersect S in 8 points (c). The orbit will then have description x^2y^2 .

Regrettably, though, if we generate symbolic sequences using the "maxima-in-z method" described above, it is possible that the sequences assigned to any periodic orbit will change as we follow the orbits with changing parameter. The two arcs (Fig. 6.1) which separate the shaded and unshaded areas of the surface S are arcs of points at which trajectories intersect S with $\ddot{z} = 0$; they are points of inflection. It is possible, for instance, that an xy orbit (with two maxima and two minima) intersects S as shown in Fig. 6.2(a). For a slightly different parameter value the same orbit may intersect S as shown in Fig. 6.2(b), and, for yet a different parameter value, it may intersect S as shown in Fig. 6.2(c). Though the periodic orbit has been involved in no bifurcation, its symbolic description is now x^2y^2 .

This type of difficulty, illustrated in Fig. 6.2 and described above, is not hypothetical. We shall see (in Chapter 8) that for smaller values of the parameter b , the very sequence of changes illustrated in Fig. 6.2 actually occurs. Consequently, the maxima-in-z method of describing trajectories with sequences of two symbols is (generally speaking) not very

useful, despite the fact that it worked perfectly well in the parameter range $0 < r < \infty$, $\sigma = 10$, $b = 8/3$.

Notice that the maxima-in-z method is really a "which half of a return surface" method; our return surface here was the surface S on which $\dot{z} = 0$. It is likely that other return surfaces will also contain two disjoint regions in which trajectories pass upward through the return surface, thus providing a natural way to assign symbol sequences to the trajectories. In these notes, we have most often been concerned with points where trajectories pass downward through the return surface; in this case, it seems that a section of the stable manifold of the origin divides the return surface into two regions in a natural way. Nonetheless, we can expect similar problems with all return surfaces; we have already seen that trajectories (to be precise, the unstable manifold of the origin) become tangential to the return surface $z = r-1$ as r increases when $\sigma = 10$ and $b = 8/3$. We must approach the problem of symbolic description in a more topological way.

6.2. SYMBOLIC DESCRIPTIONS FROM THE STABLE MANIFOLDS OF C_1 AND C_2

We saw, in Chapter 1, that for all $r > 1$ there is a negative real eigenvalue associated with the linearized flow near the stationary points C_1 and C_2 . Though it involves a slight abuse of language (since when C_1 and C_2 are stable their stable manifolds are three-dimensional), it should be clear what we mean if we say that the existence of this eigenvalue implies that for all $r > 1$, C_1 and C_2 each have a one-dimensional stable manifold. Trajectories started on these manifolds will tend asymptotically towards C_1 or C_2 at a rate governed by the eigenvalue which is real and negative for all $r > 1$. Each of these manifolds will consist of two branches which meet at the stationary point. Trajectories started on one branch will tend towards the stationary point from one side, trajectories started on the other branch will tend towards the stationary point from the other side. The stable manifold of C_1 will be the image under the symmetry of the stable manifold of C_2 .

Trajectories not started on the stable manifolds of C_1 and C_2 will not intersect them; hence, we can attempt to devise a system of symbolic description of orbits and trajectories in which symbol sequences are assigned according to the order in which orbits and trajectories wind around these two manifolds. Providing the manifolds do not wind around

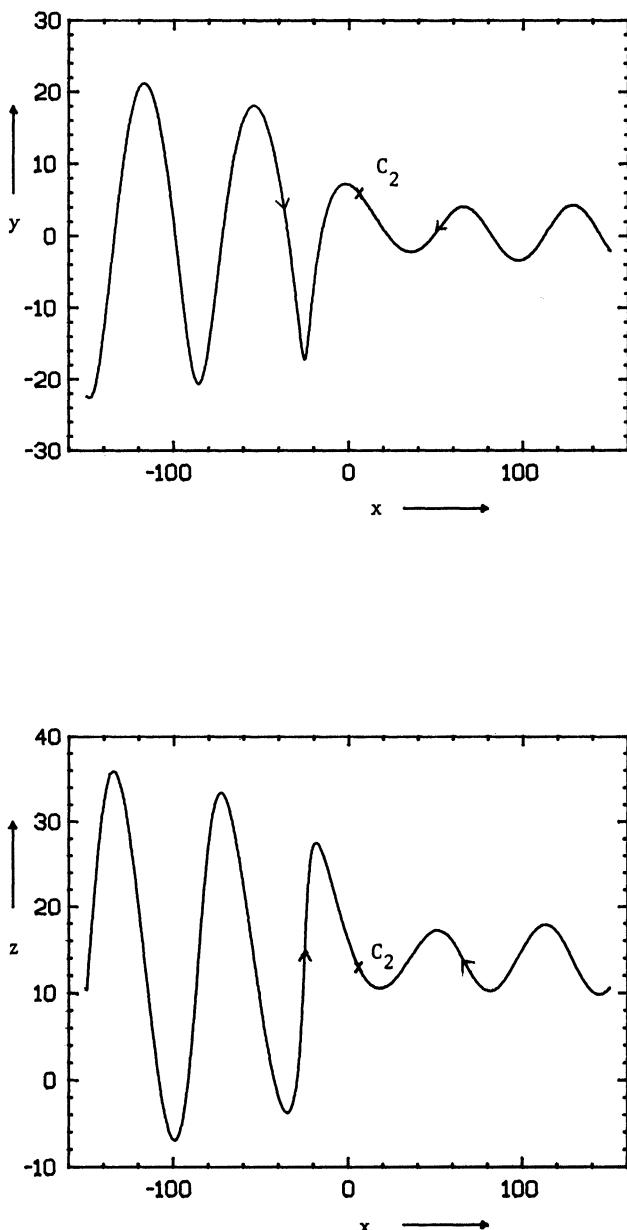


Figure 6.3. The Stable Manifold of C_2 . $r = 14.0$.

in a complicated way, which is the case when r is relatively small, this will be easy to do. Fig. 6.3 shows two projections of the stable manifold of C_2 when $r = 14.0$. The figure shows a relatively large range of x -values in which one branch of the manifold oscillates quite violently; however, in the region of dynamical interest (relatively close to C_2), the manifold behaves quite simply. If we compare pictures of the unstable manifold of the origin (at $r = 14$, just after the first homoclinic explosion) with Fig. 6.3, it is easy to see that the periodic orbits and aperiodic trajectories born in the first homoclinic explosion must wind around the stable manifolds of C_1 and C_2 in the "obvious" way. If we write an " x " everytime a trajectory winds around the stable manifold of C_2 and a " y " everytime it winds around the stable manifold of C_1 , we obtain the same symbolic descriptions that we have used throughout these notes.

As r increases, the behaviour of the manifolds becomes more complicated. Fig. 6.4 shows the stable manifold of C_2 for $r = 29.719$ and $r = 29.720$ ($\sigma = 10$, $b = 8/3$). For these r -values, it seems that one branch of the manifold spends a considerable time near the origin. Somewhere between the two r -values (29.719 and 29.720) one branch of the manifold has swung around the z -axis and, for $r > 29.72$, both branches of the manifold approach C_2 from x large and positive. It seems that this change must occur smoothly (though suddenly); the only other choice would be for the stable manifold to actually reach the origin at some r -value - in this case one branch of the unstable manifold of the origin and one branch of the stable manifold of C_2 would be the same trajectory, and we would expect to have seen some evidence of this in our numerical investigation of the behaviour of the unstable manifold of the origin.

As r increases beyond 29.72, more "half-swings" around the z -axis occur in the stable manifolds of C_1 and C_2 . Each new half-swing appears to be presaged by the manifold passing very close to (and below) the origin. This phenomenon is not well understood. Fig. 6.5 shows the stable manifold of C_2 at $r = 100$, when one branch of the manifold makes one and a half swings around the z -axis.

Notice that it is always the same branch of the stable manifold that changes its behaviour. The other branch seems to approach C_2 "smoothly and directly" for all r -values. It is not known in any detail how the behaviour of the manifolds of C_1 and C_2 changes as r continues to increase. The large r results of Chapter 7 suggest that one branch will always approach the stationary point without twisting around the z -axis

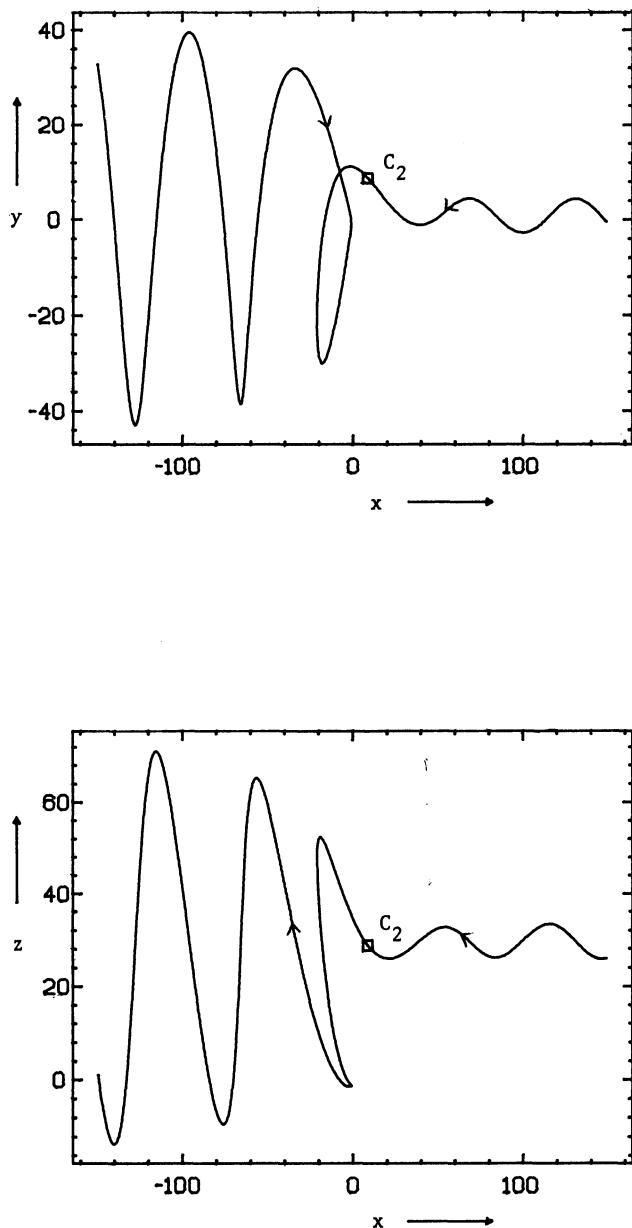
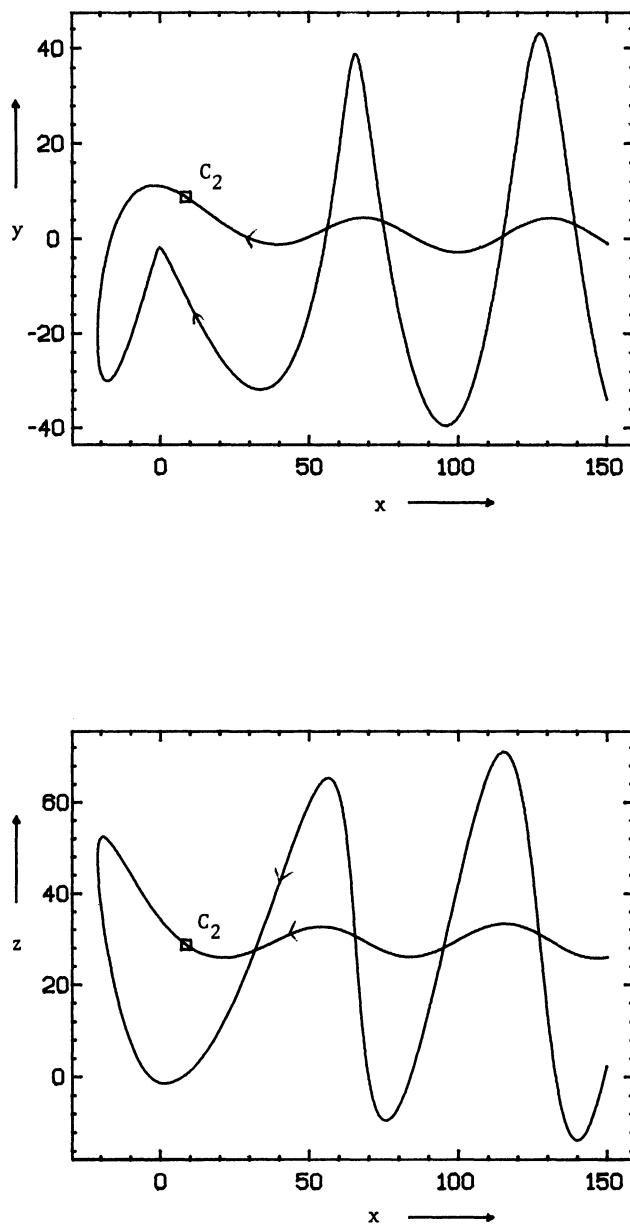


Figure 6.4. The stable manifold of C_2 . (a) $r = 29.719$.

Figure 6.4. (b) $r = 29.720$.

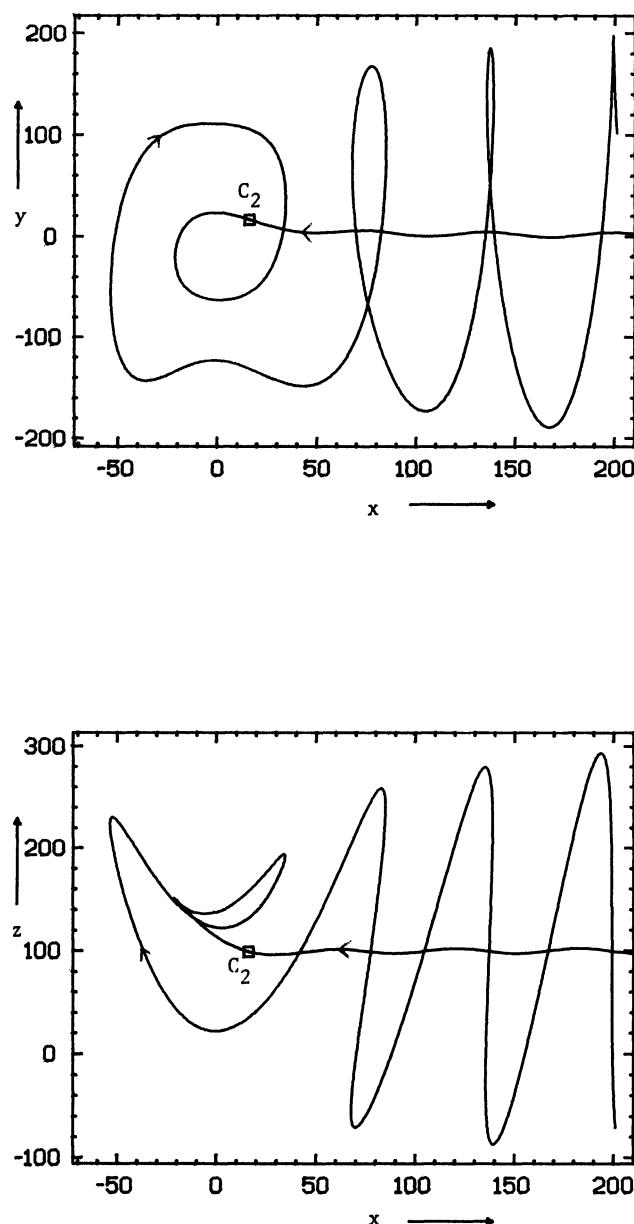


Figure 6.5. The stable manifold of C_2 . $r = 100$.

and with a z -coordinate near r and a y -coordinate near 0 (cf. the figures in this chapter). The large r results also suggest that, as r increases, the other branch will wind more and more around the z -axis; for large r we expect this branch of the manifold to approach the stationary point from above (z very large), twisting around the z -axis as it comes. The large r results give no details of how we expect the extra twists to appear as r increases.

When the manifolds look complicated (as in Fig. 6.5) it becomes quite hard to calculate the symbolic descriptions of orbits and trajectories according to the way in which they wind around the two manifolds. For the few orbits I have tried, the symbolic descriptions obtained are the usual ones; however, if we are to use the stable manifolds as a rigorous topological justification for our symbol sequences, we must adopt a more general approach.

We know, from Chapter 1 and Appendix C, that there is, for each r , a bounded ellipsoid E which all trajectories eventually enter. We also know that trajectories never leave this ellipsoid E . We shall concentrate on those sections of the stable manifolds of C_1 and C_2 that lie within E .

Each manifold intersects the boundary of E in at most two points; if there were more than two such points (one for each branch) then at least one of them would have to be a point at which a trajectory left the ellipsoid E . That there are at most two intersections then implies that the sections of each stable manifold which lie within E are connected. For most r -values, we expect each stable manifold to intersect E in exactly two points. However, for isolated r -values, one or both branches of a stable manifold may be part of the unstable manifold of some stationary point or periodic orbit (or more complicated attractor), in which case one or both branches of the stable manifold will lie completely within E . For most r -values, the manifolds of the two stationary points will not meet, cross or touch one another, either within E or on the boundary. This implies that (for most r -values) we can describe closed curves within E with respect to the two stable manifolds of the points C_1 and C_2 ; we will be topologically justified in specifying the direction and order in which the closed curve winds around the two manifolds. Essentially, we wish to find the first homotopy group of a sphere minus two non-intersecting lines. This is equivalent to finding the first homotopy group of a disc minus two distinct points, and that group is the free group on two symbols.

indication that either of these events occurs at any r -value. However, if either event did occur, it is extremely likely that the symbolic descriptions of many periodic orbits would change as parameters moved through the critical r -value.

(iii) For some r -values, one or both branches of each stable manifold may be part of the unstable manifold of some non-stable periodic orbit (or other, more complicated, non-stable part of the non-wandering set). Again, we have no numerical evidence that this ever occurs, but again, if it does, we can expect it to upset our symbolic descriptions.

It is clear that we cannot obtain a completely rigorous justification of our symbol sequences by using the stable manifolds of C_1 and C_2 . Furthermore, the other methods used in these notes worked perfectly well in practice. But before summarizing the relative merits of the different methods (and there are reasons why the topological method is to be preferred for use in theoretical arguments), one more aspect of the behaviour of the stable manifolds should be discussed and one more general point about symbolic descriptions should be made.

It has been suggested that the complicated behaviour of the stable manifolds of C_1 and C_2 is somehow necessary for the behaviour we have observed in the Lorenz equations. This suggestion is not convincing.

Fig. 6.6 shows the behaviour of the stable manifold of C_2 for parameter values $\sigma = 10$, $b = 0.25$ and $r = 480$ and 500. The behaviour seen for $r = 480$ ($b = 0.25$) is very simple; however, as we shall see in Chapter 8, for this smaller value of the parameter b we expect to see essentially the same range of behaviours already observed (Chapters 1-5, $b = 8/3$), all occurring for r -values much less than 480. For $r = 500$ ($b = 0.25$), we see that the first half-swing around the z -axis has occurred. We shall argue, in Chapter 8, that this change in behaviour is connected with heteroclinic trajectories between C_1 and C_2 (unlike the change illustrated in Fig. 6.4). These heteroclinic trajectories seriously upset the symbolic descriptions of some periodic orbits which exist at these r and b -values. Apparently (Chapter 8), further swings around the z -axis (which still occur as r increases for $b = 0.25$) are not connected with heteroclinic trajectories between C_1 and C_2 . None of this is very well understood, at least insofar as the behaviour of the stable manifolds of C_1 and C_2 is concerned.

Finally, it is worth discussing the z -axis. This is invariant for all values of the parameters σ , b and r , and, as we stated in Chapter 1, we could use this line to produce a particularly simple topological

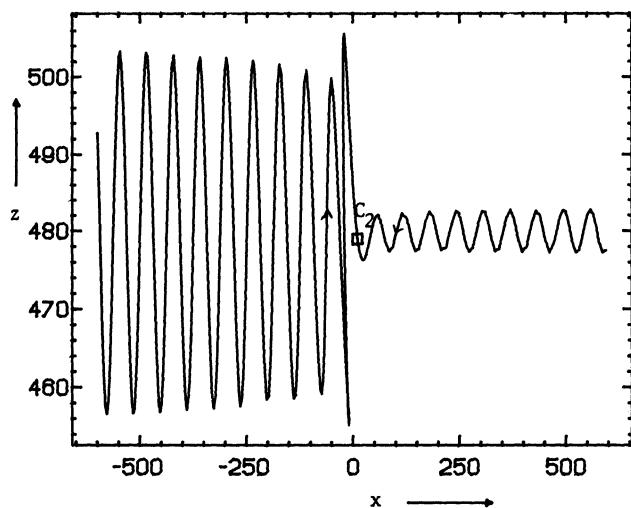
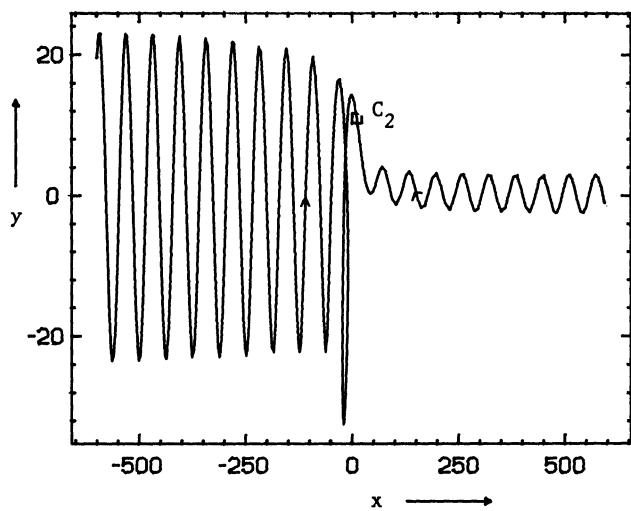


Figure 6.6. The stable manifold of C_2 for small b .
(a) $r = 480$, $b = 0.25$.

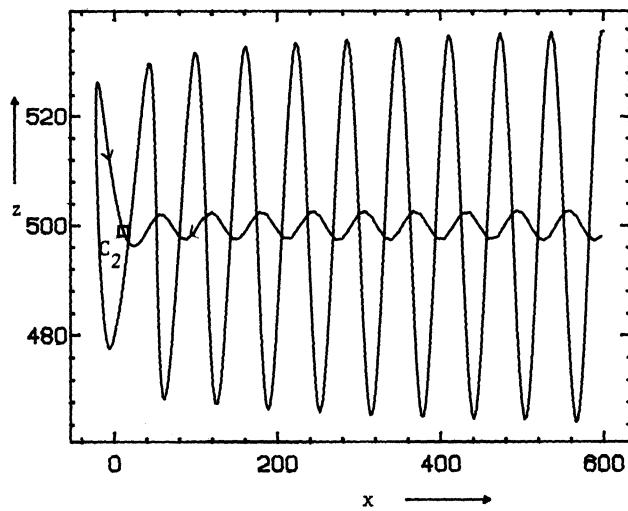
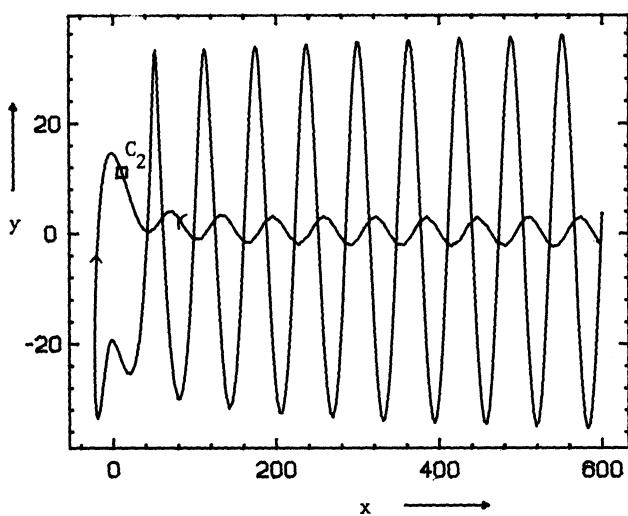


Figure 6.6. (b) $r = 500$, $b = 0.25$.

invariant for periodic orbits; orbits can only rotate around the z-axis in one direction (see Chapter 1), and we can be absolutely certain that the number of windings around the z-axis does not change as we follow an orbit with changing parameter. We have not used this topological invariant very often; by itself it does not distinguish between a sufficient number of orbits, and once we introduce the stable manifolds of C_1 and C_2 it becomes redundant. This redundancy is dynamically determined (it is not topological); all interesting trajectories appear to wind at least once around each of the two stable manifolds during each complete rotation around the z-axis. In Chapter 8 we will see that this is not always the case; nonetheless, the Chapter 5 arguments apply in an interesting parameter range where this redundancy still persists.

6.3. SUMMARY

We set out to find a global method of describing periodic orbits and trajectories with sequences of two symbols. We desired that the method would assign symbolic sequences to trajectories and orbits in such a way that the description of the trajectory or orbit would not change with changing parameter. We have not completely succeeded, but we have made some progress. We examined two methods of generating symbolic sequences and they can be contrasted as follows:

1. The maxima-in-z method is easy to use. It can be used to generate symbolic sequences for both closed orbits and aperiodic trajectories. However, it may fail (in the sense that periodic orbits may change their symbolic descriptions with changing parameter). We have no way of knowing when it might fail; it is likely that it will fail at different parameter values for different periodic orbits, and the fact that it does not appear to fail for the periodic orbits we have examined cannot be taken to imply that it does not fail for other periodic orbits that we have not examined.
2. The topological method is not so easy to use at some r-values. It can only be used to describe closed orbits. It may fail (in the sense that periodic orbits may change their symbolic descriptions as parameters change) but, if it does so, we know why it does so and what kind of numerical experiments to make to illustrate the problem. If it fails it will do so at isolated parameter values for which one or both branches of each stable manifold tends, in reverse time, towards some part of the non-wandering set. We can argue that the fact that

the method does not fail for some periodic orbits probably indicates that it does not fail at all ($\sigma = 10$, $b = 8/3$).

It should be clear why the second method (the topological one) is to be preferred for use in theoretical arguments. Though we may not be rigorously justified in assigning sequences to the unstable manifold of the origin when it is aperiodic, we can describe it with sequences of two symbols at each homoclinic explosion; this is probably sufficient for the purposes of the arguments in Chapters 3 through 5. In fact, the method is probably adequate to justify using the various arguments which depend on symbolic sequences over a wide range of parameter values. For $\sigma = 10$ and $b = 8/3$ it appears that the maxima-in-z method gives us equivalent descriptions; there is no particular reason to expect this, but it is very convenient.

Chapter 7

Large r

Numerical experiments indicate that for $\sigma = 10$, $b = 8/3$ and $r > 313$, there is a stable symmetric xy periodic orbit. Furthermore, we suggested in Chapter 5 that this periodic orbit and the three stationary points would, for large enough r , make up the whole of the non-wandering set. In this chapter, we show that there are theoretical reasons to expect both of these results. At the same time, we show that qualitatively more complicated large r behaviour may be expected for some values of the parameters σ and b .

7.1. THE AVERAGED EQUATIONS

Robbins (1979) shows the existence and uniqueness of a symmetric periodic orbit as r tends to infinity (see also Shimizu (1979) who gives an approximate analytic form for the orbit). Robbins' arguments have to be described as formal rather than rigorous, but the calculations are equivalent to the method of averaging and can in that way be rigorously justified. Robbins studied only local phenomena; our use of the method of averaging (see, for example, Swinnerton-Dyer, 1980) allows us to make a global analysis of the behaviour which turns up some interesting predictions. What follows is not a complete proof but an outline. For the details, readers are referred to Appendix K and to Sparrow & Swinnerton-Dyer (1982). A recent paper by Fowler & McGuinness (1982) is also relevant.

First, we notice that inside the bounded region in which trajectories eventually lie (see Appendix C) $x = O(r^{\frac{1}{2}})$, $y = O(r)$ and $z = O(r)$. The coordinates x , y , and z are not particularly convenient at large r

since they obscure the relative magnitudes of the variables. We therefore change coordinates so that everything is order one. We write

$$\varepsilon = r^{-\frac{1}{2}}; \quad x = \varepsilon^{-1}\xi; \quad y = \varepsilon^{-2}\sigma^{-1}\eta; \quad z = \varepsilon^{-2}(\sigma^{-1}z + 1) \quad \text{and} \quad t = \varepsilon\tau$$

where τ is our new independent variable. This transforms the Lorenz equations into:

$$\begin{aligned} \frac{d\xi}{d\tau} &= \eta - \varepsilon\sigma\xi \\ \frac{d\eta}{d\tau} &= -\xi z - \varepsilon\eta \\ \frac{dz}{d\tau} &= \xi\eta - \varepsilon b(z+\sigma) \end{aligned} \tag{1}$$

We now have a small parameter problem with $\varepsilon \rightarrow 0$ as $r \rightarrow \infty$. In the limit $\varepsilon = 0$ we obtain, from Equations (1),

$$\begin{aligned} \xi' &= \eta \\ \eta' &= -\xi z \\ z' &= \xi\eta \end{aligned} \tag{2}$$

which has the two obvious integrals;

$$\xi^2 - 2z = 2A; \quad \eta^2 + z^2 = B^2 \tag{3}$$

where A and B are constants. We choose the sign of B to be positive and then need to consider separately the two regions $A > B > 0$ and $B > A > -B$. (There are no real solutions in $A < -B$.) The existence of the integrals (3) suggests that the general trajectory of the approximate equations (2) is a closed orbit; the exceptions are the two lines of stationary points $\xi = \eta = 0$ and $\eta = z = 0$ and some homoclinic orbits;

$$\xi^2 - 2z = 2B; \quad \eta^2 + z^2 = B^2$$

associated with the stationary point $(0,0,-B)$. The situation is illustrated in Fig. 7.1. (See also Robbins, 1979.) The solutions lie on some cylinder, $\eta^2 + z^2 = B^2$. When $A > B$ there are two non-symmetric solutions, and when $|A| < B$ there is one symmetric solution. Notice that the symmetry $(x,y,z) \mapsto (-x,-y,z)$ carries over into a symmetry $(\xi,\eta,z) \mapsto (-\xi,-\eta,z)$ of equations (1) and (2); we expect remarks about symmetric and non-symmetric orbits in the new coordinates to have the obvious relevance to the old coordinates.

It is in proceeding from this point that the method of Robbins and the method of averaging formally diverge. Robbins expands solutions

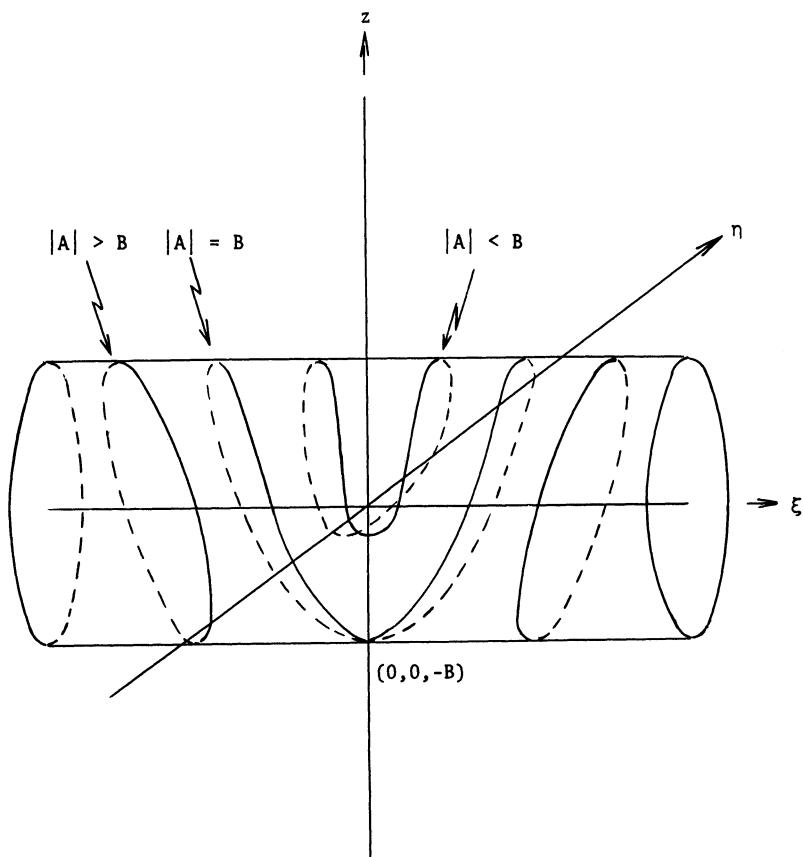


Figure 7.1. Solutions to the equations (2) are periodic orbits or homoclinic orbits ($|A| = B$) on some cylinder $\eta^2 + z^2 = B^2$.

$w(\epsilon, \tau)$ of the full equations (1) in powers of ϵ and discards terms of order greater than one. We proceed as follows. First, we complete the integration of the approximate, "limiting" equations, (2). This can be done in terms of elliptic integrals. Then we use the exact equations (1) to calculate the rate of change of variables A and B defined by equations (3). We should remember that equations (3) are not integrals of equations (1), and so we expect that A and B will vary slowly as the variables ξ , η and z move under the flow defined by (1). Simple substitution gives us:

$$\begin{aligned} A' &= \epsilon(-\sigma\xi^2 + bz + b\sigma) \\ BB' &= -\epsilon(\eta^2 + bz^2 + b\sigma z) \end{aligned} \tag{4}$$

and so A and B are slowly varying.

What we want is to have differential equations for A and B in terms of A and B . We must eliminate the variables ξ , η and z . We do this by substituting the solutions of equations (2) into (4) and then integrating over one complete period of the orbits which are solutions to (2). We end up with two equations which tell us the average rate of change of A and B . These first order equations are called the "averaged equations". The averaged equations are different in the two regions of the A, B plane that interest us ($A > B > 0$ and $|A| < B$). The method of averaging is, in general, only applicable where the orbits of (2) have bounded period (i.e., in some region away from the line $A = B$ which divides the two interesting parts of the A, B plane). However, the long period orbits of (2) spend all but a bounded part of the period close to a particular point, and the method therefore holds with some modification (Swinnerton-Dyer, 1980). To move $O(\epsilon)$ along a trajectory of the averaged equations involves going once around a large loop in the original variables; the loop almost closes up and goes once around the z -axis if we are in $|A| < B$ (but not otherwise). A stationary point of the averaged equations corresponds to a loop in the original variables which does close up (i.e., a periodic orbit). If the stationary point lies in $|A| < B$ it corresponds to a single symmetric orbit; if it lies in $A > B > 0$ it corresponds to a symmetric pair of non-symmetric orbits. It is shown in Swinnerton-Dyer (1980), that the solutions of the averaged equations differ only by $O(\epsilon)$, in a time $O(\epsilon^{-1})$, from the solutions to the full equations (1) (provided the relevant elliptic function is bounded - i.e., away from $A = B$). The averaged equations are extremely ugly. They are given in Appendix K, along with the calculations needed to reach the conclusions which follow.

7.2. ANALYSIS AND INTERPRETATION OF THE AVERAGED EQUATIONS

It is convenient to define a parameter $\lambda = \frac{(\sigma+1)}{(b+2)}$. Fig. 7.2 shows the behaviour of the averaged equations (which are actually two-dimensional) in the three parameter ranges, $\lambda < \frac{2}{3}$, $\frac{2}{3} < \lambda < 1$ and $\lambda > 1$. This behaviour can be described as follows:

1) The Line $A = -B$

Trajectories always move along the line $A = -B$ towards the point $(0,0)$. Near to $A = -B$, trajectories move towards the line unless B (and A) are very small. By examining equations (1), (3) and the change of coordinates, we can see that this boundary represents, in the original Lorenz equations, that part of the z -axis which lies above the plane $z = r$.

2) The Region $|A| < B$

Trajectories move in towards the region displayed in Figs. 7.2 from far away (i.e., the averaged equations are bounded). There is a unique stationary point in this region if $\lambda > \frac{2}{3}$ and none otherwise. When the stationary point exists, it is stable. This stationary point represents a symmetric periodic orbit in the original Lorenz equations. Fig. 7.1 indicates that the symmetric orbits represented by points in this region go once around the z -axis. Consequently, the stable stationary point of the averaged equations represents the stable symmetric xy orbit which exists for large enough r in $\lambda > \frac{2}{3}$.

3) The Line $A = B$

There is a non-stable stationary point on the line $A = B$ at (σ, σ) . This point represents the origin in the original Lorenz coordinates, and the piece of $A = B$ lying between $(0,0)$ and (σ, σ) represents that part of the z -axis which lies below the plane $z = r$ and above the plane $z = 0$. Trajectories of the averaged equations can move along $A = B$ towards the stationary point. However, there is a complication associated with this boundary. Trajectories which approach $A = B$ will reach it in finite time. They will then travel along $A = B$ towards (σ, σ) , but may, at any stage, leave the line $A = B$ before reaching the stationary point. Trajectories leaving the line $A = B$ will move into the region $|A| < B$ unless $B < \frac{3b\sigma}{(6\sigma - b - 2)}$ (which is approximately $B < \frac{b}{2}$ if σ is large compared with b). There is a genuine non-uniqueness in the behaviour on this line, associated with the fact that the method of averaging breaks down here. We will have more to say about this in Section 7.3.

4) The Region $A > B > 0$

Once again, trajectories move in from far away (boundedness of the averaged equations). There are no stationary points in this region if $\lambda < \frac{2}{3}$ or if $\lambda > 1$. In $\frac{2}{3} < \lambda < 1$ there is a unique stationary point which is non-stable. This stationary point represents, for the Lorenz equations, a pair of non-symmetric periodic orbits which do not wind around the z-axis.

5) The Line $B = 0$

There is a stationary point on the line $B = 0$ at $A = b/2$. This stationary point represents the two stationary points (C_1 and C_2) of the Lorenz equations. Trajectories on the line $B = 0$ move towards the stationary point and trajectories near $B = 0$ move toward the line unless $A < \frac{b\sigma}{2(b+1)}$. Since $\frac{b}{2} = \frac{b\sigma}{2(b+1)}$ when $\lambda = 1$, this stationary point is stable for $\lambda < 1$ and non-stable for $\lambda > 1$. (Trajectories moving towards the line $B = 0$ take an infinite time to reach it; we do not have the same problem here that we have on the line $A = B$.) Notice that the stability of the stationary point $(\frac{b}{2}, 0)$ corresponds to what we would expect from consideration of the Hopf bifurcation in the Lorenz system. The Hopf bifurcation occurs at $r_H = \frac{\sigma(\sigma+b+3)}{(\sigma-b-1)}$ and hence only occurs for a finite positive r-value if $\lambda > 1$. If $\lambda < 1$ the stationary points C_1 and C_2 remain stable for all r however large. In $\lambda > 1$, we may interpret the section of the line $B = 0$ lying in $A > \frac{b}{2}$ as representing one branch of the one-dimensional stable manifold of the stationary point C_1 and C_2 . The other branch will be represented by the union of the line $A = -B$ and that section of $B = 0$ lying between $(0,0)$ and $(\frac{b}{2}, 0)$. Notice how this observation corresponds with our numerical experiments in Chapter 6; there we saw that one branch of the stable manifold of C_1 always seemed to approach C_1 in a straightforward way without twisting around the z-axis. The other branch approaches C_1 after twisting around the z-axis an increasing number of times as r increased. These results suggest that this pattern continues as r increases; furthermore, $B = 0$ implies we will have $y \sim 0$ and $z \sim r$ (original variables) on the non-twisting branch (cf. the figures of Chapter 6).

6) Periodic Orbits of the Averaged Equations

One may show, using a simple index argument, that there can be no periodic solutions of the averaged equations lying completely within $A > B > 0$. Similarly, when $\lambda < \frac{2}{3}$ and there is no stationary point in

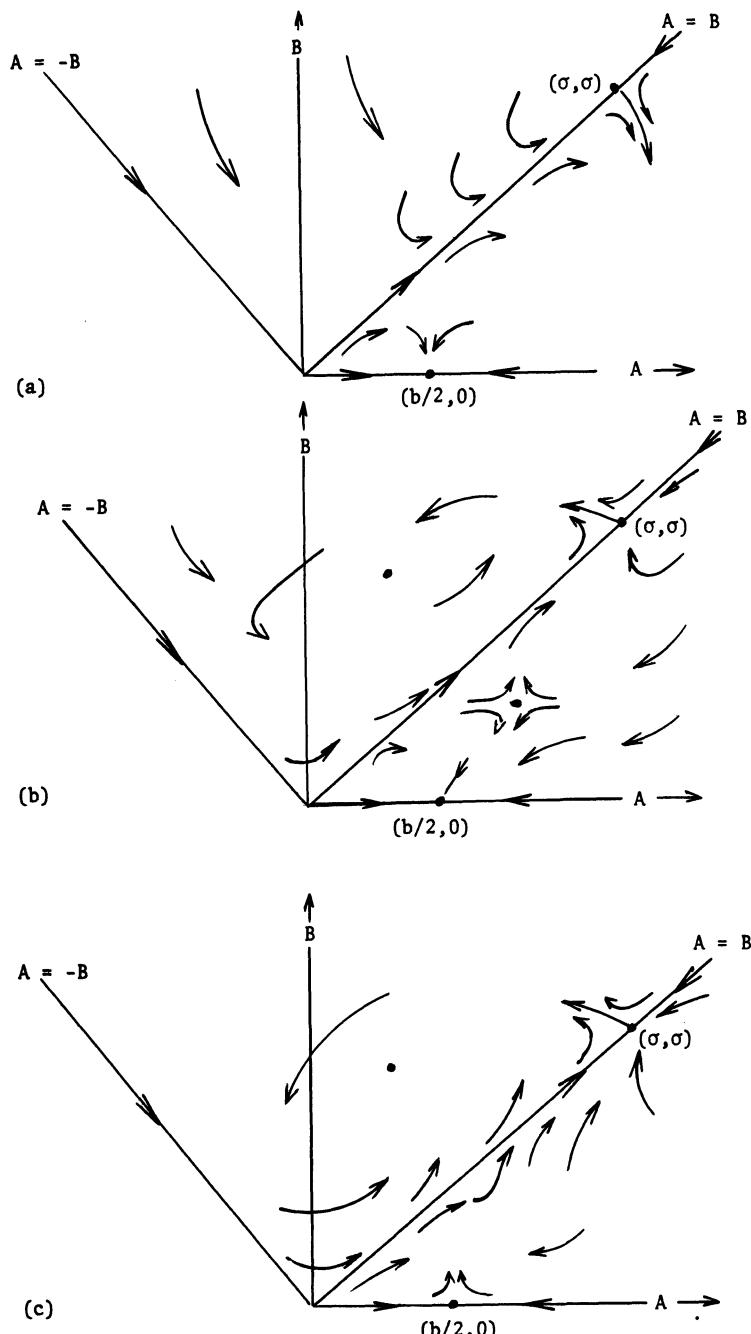


Figure 7.2. Phase portraits for the averaged equations. (a) $\lambda < \frac{2}{3}$; (b) $\frac{2}{3} < \lambda < 1$; (c) $1 < \lambda$.

$|A| < B$, we can show that there is no periodic orbit contained completely within $|A| < B$. We can also show (Appendix K) that the flow of the averaged equations is contracting in $|A| < B$, and hence that there are no periodic orbits of the averaged equations contained completely in $|A| < B$ for any values of the parameter λ .

These six sets of observations indicate that the proper phase portraits for the averaged equations are those shown in Fig. 7.2.

We can summarize our knowledge of the large r behaviour of the Lorenz equations as follows.

a) $\lambda < \frac{2}{3}$.

In this parameter range C_1 and C_2 remain stable for all r however large and there is no other interesting large r behaviour. This suggests that there may be no interesting bifurcations in $r > 1$ when $\lambda < \frac{2}{3}$. There may, of course, be some bifurcations in this parameter range, but the large r analysis suggests that any sequence of bifurcations which does occur reverses itself at large enough r -values and we end up at large r with the same situation that we have for r just greater than one.

b) $\frac{2}{3} < \lambda < 1$.

In this parameter range we have a stable symmetric orbit and a pair of non-stable non-symmetric orbits at large r . This suggests that we have the first homoclinic explosion (to produce these orbits) and that we have the subsequent behaviour changes observed when $\sigma = 10$ and $b = 8/3$ (to rid us of all the rest of the original strange invariant set born in the first homoclinic explosion), but that the Hopf bifurcation does not occur for any finite r -value. In this situation the two simple Hopf orbits (x and y) continue to exist for all r -values and the stationary points, C_1 and C_2 , remain stable. This may mean that it is difficult to observe any of the interesting behaviour numerically in this parameter range. Notice also that for λ just greater than $\frac{2}{3}$, the two stationary points of the averaged equations which represent periodic orbits lie very close to (σ, σ) (see Fig. 7.3); this leads us to expect that the large r periodic orbits (and probably the preceding behaviour changes also) lie very close to the z -axis and pass very close to the origin. This is an additional reason why the behaviour may be difficult to observe.

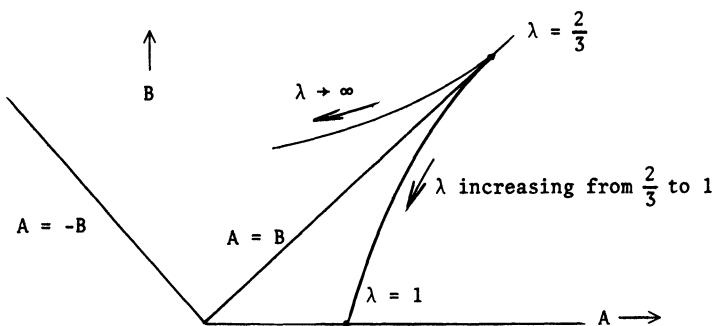


Figure 7.3. Position of the stationary points as λ changes (with appropriate rescaling to keep the stationary points on $A = B$ and $B = 0$ in the same place as λ changes).

c) $\lambda > 1$

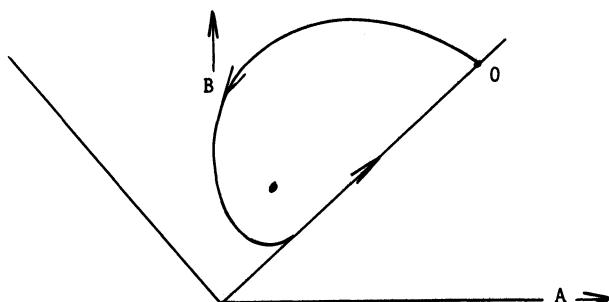
In this parameter range (which includes $\sigma = 10$ and $b = 8/3$) we have only the stable symmetric periodic orbit at large r -values. Again we can suggest that this implies that we have both the first homoclinic explosion and subsequent behaviour changes, but now the Hopf bifurcation occurs at a finite r -value and the points C_1 and C_2 are non-stable if r is large enough. $\lambda = 1$ is the parameter value at which the Hopf bifurcation occurs "at infinity". See Fig. 7.3.

7.3. ANOMALOUS PERIODIC ORBITS FOR SMALL b AND LARGE r

In the last section, we reached the conclusions which we needed for our analysis of the Lorenz equations when $\sigma = 10$ and $b = 8/3$. In this section we look more closely at some problems which may arise because of the indeterminacy of the averaged equations on the line $A = B$.

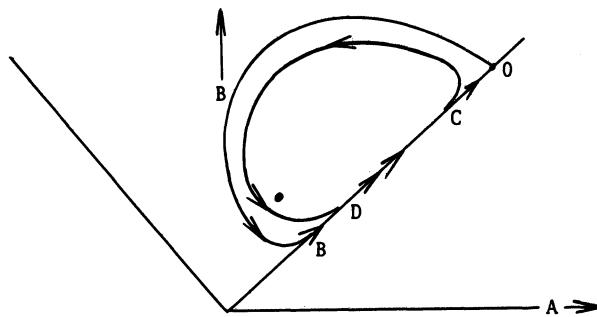
We know that there can be no periodic orbits of the averaged equations contained entirely within $|A| < B$. However, we may have a periodic orbit like that shown in Fig. 7.4(a). If this trajectory is possible, there may be many more complicated periodic orbits involving different portions of the line $A = B$ (because of the non-uniqueness of the flow on $A = B$). See, for example, Figs. 7.4(b) and (c).

(a)



The simplest periodic orbit of the averaged equations

(b)



A more complicated periodic orbit (OBCDO...) of the averaged equations

(c)

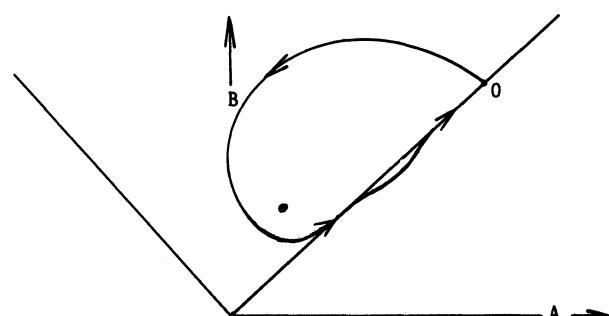
A periodic orbit of the averaged equations which enters $A > B > 0$

Figure 7.4. Schematic pictures of some possible periodic orbits of the averaged equations. (a) The simplest periodic orbit of the averaged equations; (b) A more complicated periodic orbit (OBCDO...); (c) A periodic orbit which enters $A > B > 0$.

The critical trajectory will be the one started at (σ, σ) . If this trajectory (which represents the unstable manifold of the origin) does not strike the line $A = B$, then no periodic orbits of the kind illustrated in Fig. 7.4 will be possible. This appears to be the case when $\sigma = 10$ and $b = 8/3$ and so the discussion which follows should not change our view of the Lorenz equations for these parameter values. However, numerical simulations of the averaged equations indicate that when $\sigma = 10$ and b is small enough, the more complicated situation does arise. Fig. 7.5 shows simulation of the averaged equations for several b values. These figures were produced in a fairly crude way and should be considered approximate; there are problems involved with integrating the equations near all three lines $A = -B$, $A = B$ and $B = 0$. In particular, numerically calculated trajectories do not move along the line $A = B$.

Fig. 7.5 indicates that we have the complicated situation for b less than some b^* , $2 < b^* < 8/3$, when $\sigma = 10$. Quite different considerations of other authors (Fowler & McGuinness, 1981a, 1981b) suggest that the appropriate more general condition might be $\frac{\sigma}{b}$ large enough or σ/b^2 large enough. This also seems likely on examination of the averaged equations. We must ask what the existence of these anomalous periodic orbits of the averaged equations implies for the Lorenz equations. As yet, the phenomenon is not entirely understood, but we can make quite a lot of sense of it.

If we have the anomalous periodic orbit of Fig. 7.4(a), we can draw an uncountable infinity of other anomalous periodic orbits. Fig. 7.4(b) and (c) show just two examples out of many. Some of these anomalous orbits will lie entirely within $|A| < B$ and some will have sections lying in $A > B > 0$ as well. All will spend some time on the line $A = B$. We wish to know, for large but finite r , which of these anomalous orbits corresponds to anything recognizable in the original variables. It appears that some of the anomalous orbits do correspond to periodic orbits of the Lorenz equations (and there will also be anomalous aperiodic trajectories which correspond to aperiodic trajectories in the Lorenz equations). Exactly which anomalous orbits correspond to Lorenz orbits will depend on the value of ϵ (and hence r) which we consider.

Though the averaged equations look the same for every ϵ -value (ϵ determines the rate at which trajectories move but not their shape - see equation (4)) the correspondence between averaged trajectories and Lorenz trajectories will alter with changing ϵ . Let us consider a particular averaged trajectory which reaches the line $A = B$ in finite time. At

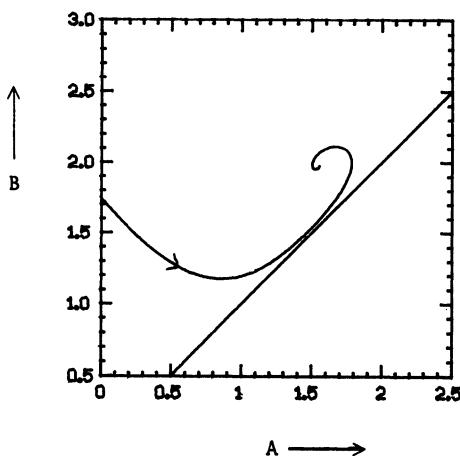
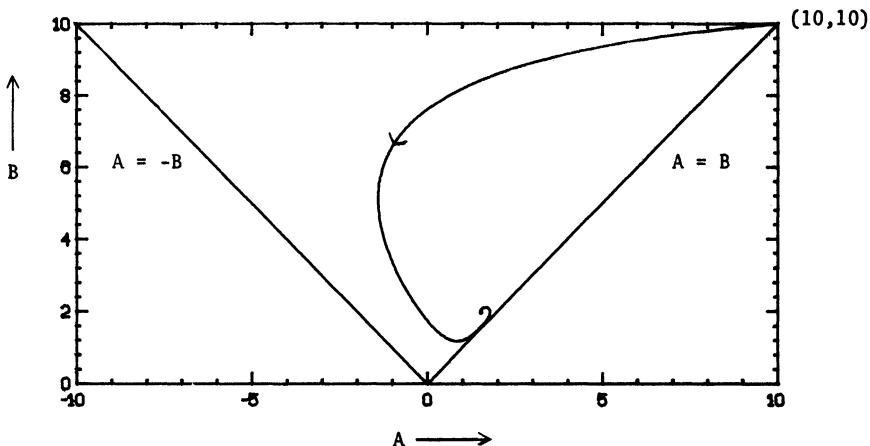


Figure 7.5. Numerical simulations of the averaged equations ($\sigma = 10$). Anomalous periodic orbits may occur if the trajectory started at (σ, σ) strikes $A = B$ ((b), (c) & (d)).

(a) $b = 8/3$.

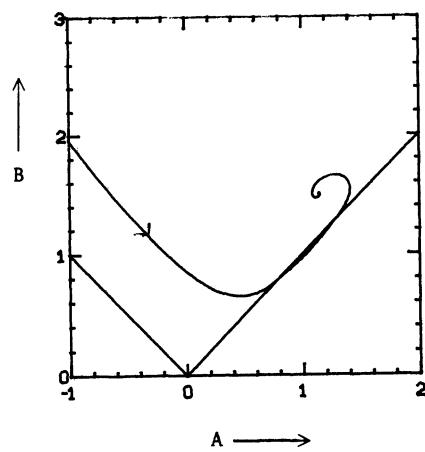
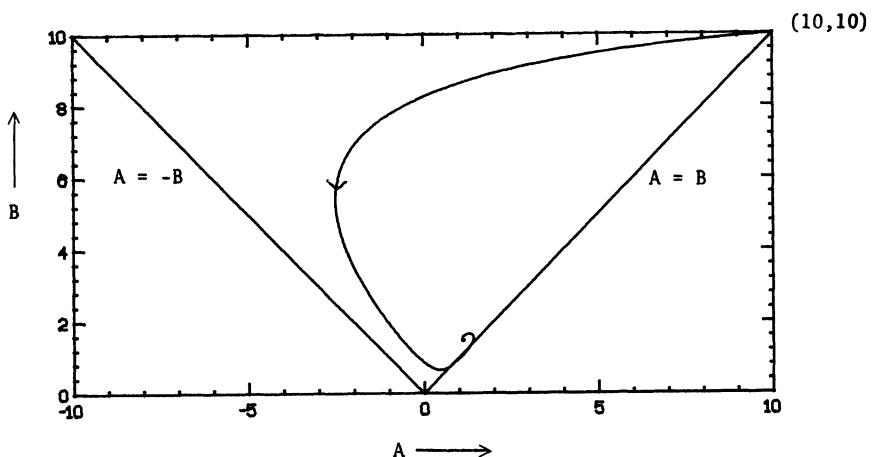


Figure 7.5. (b) $b = 2$.

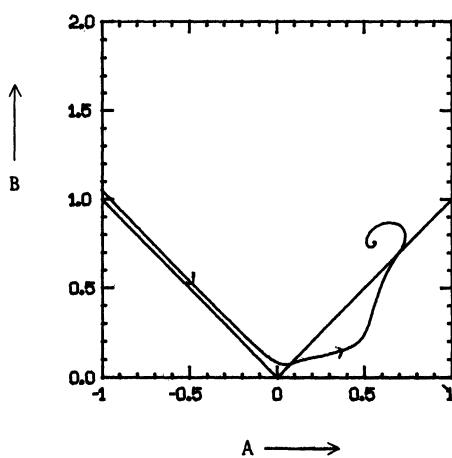
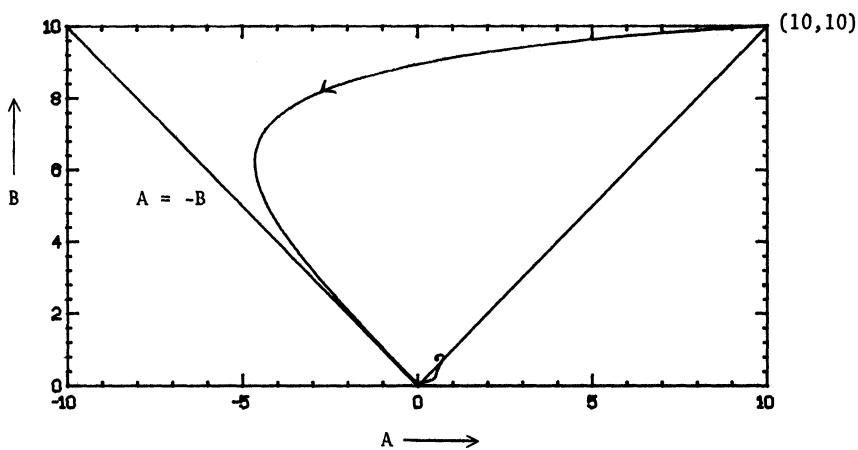


Figure 7.5. (c) $b = 1$.

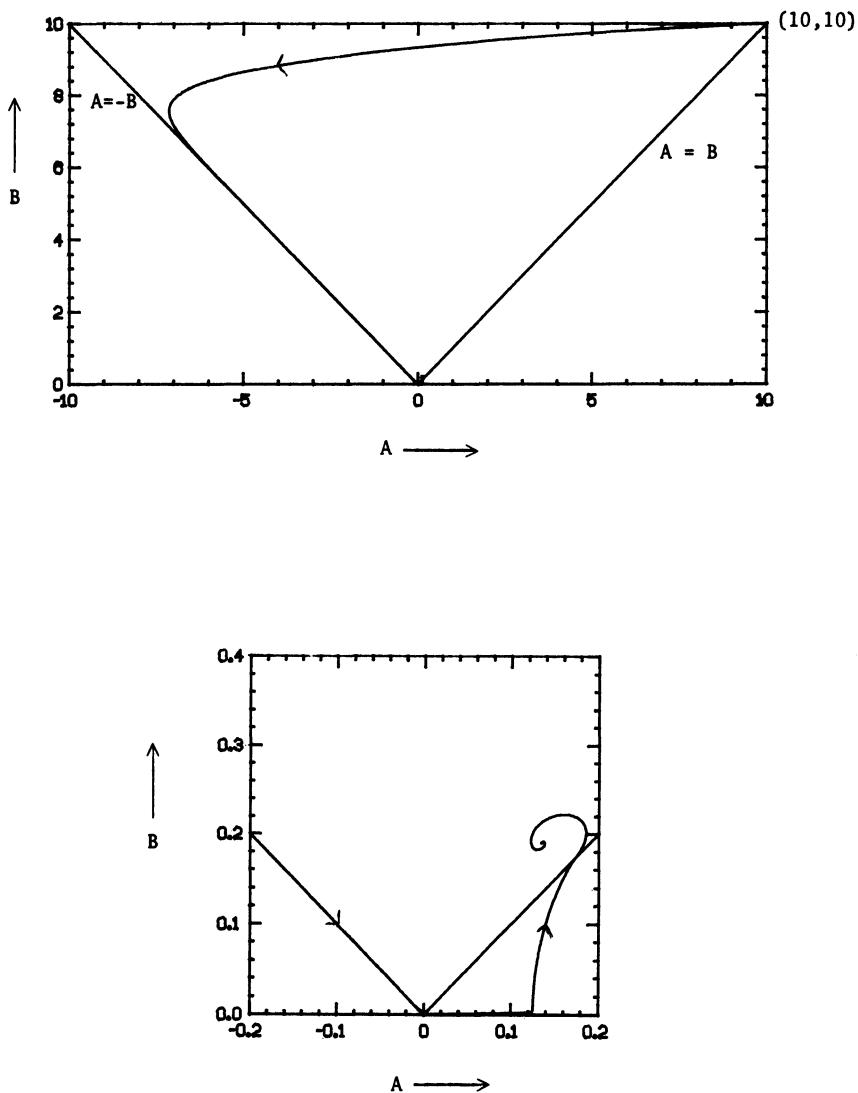


Figure 7.5. (d) $b = 0.25$.

the point where this trajectory reaches the line $A = B$, the values of the variables ξ and η for the corresponding Lorenz trajectory will be exponentially small but crucially non-zero (a trajectory started off the z -axis will never get onto the z -axis whatever our averaging technique tells us). The values of ξ and η at this point will depend on ϵ . The averaged trajectory now moves along $A = B$ towards (σ, σ) . The corresponding Lorenz trajectory moves (very slowly and exponentially close to the z -axis) towards the origin. The Lorenz trajectory will only continue to move towards the origin if the particular values of ξ and η we obtained happened to be such that the trajectory lay on the stable manifold of the origin. Otherwise, at some later time ($\sim O(\epsilon^{-1})$), the variables ξ and η increase again and the corresponding averaged trajectory moves off the line $A = B$. We might explain the indeterminacy on the line $A = B$ in the following non-rigorous way. If we wish to consider the averaged equations as representing the Lorenz equations for some particular large value of r , they are deterministic. Every averaged trajectory which reaches the line $A = B$ leaves it again (or not) at some determined point, but the averaged equations do not contain the information for us to decide which point this is. For another large r -value, the averaged equations will again be deterministic, but the same averaged trajectory arriving at the line $A = B$ will leave it at a different point. Consequently, if we just examine the averaged equations without any particular r -value in mind, we must have genuine non-uniqueness on the line $A = B$.

For any particular r -value, then, there will be many anomalous periodic orbits of the averaged equations which do not correspond to Lorenz orbits. However, there will be some (a countable infinity?) which do. We can say something about the way these orbits look. Remember that to progress $O(\epsilon)$ along an averaged trajectory (away from the line $A = B$) corresponds to going once around a large loop in the original variables; this loop will almost close up and will wind once around the z -axis if and only if the corresponding part of the averaged trajectory lies in $|A| < B$. Thus, the sections of anomalous periodic orbits which lie in $|A| < B$ correspond to parts of real periodic orbits which wind many times around the z -axis. The sections of the averaged trajectories which lie on the line $A = B$ correspond to parts of real periodic orbits which lie close to the z -axis and below the plane $z = r$. We can deduce that a periodic orbit of the Lorenz equations which corresponds to an anomalous periodic orbit of the averaged equations winds many times around the z -axis

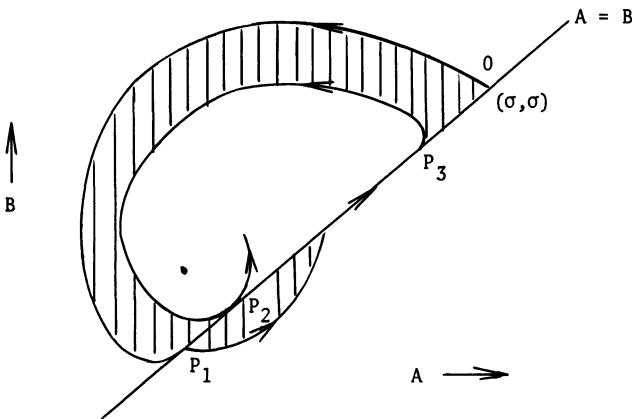


Figure 7.6. Anomalous periodic orbits of the averaged equations must lie in the shaded region. The largest possible arc lying entirely within $|A| < B$ is OP_1 . The shortest such arc is P_3P_2 .

between each passage into the region $z < r$. We can actually do better than this. Fig. 7.6 illustrates that there is a maximum and minimum length for the section of an anomalous trajectory which lies completely in $|A| < B$. The trajectory out from O strikes the line $A = B$ at P_1 and the arc OP_1 is the longest possible such arc that can exist in an anomalous periodic orbit. There is a point P_3 on $A = B$ from which the trajectory just touches the line $A = B$ at P_2 ; the arc P_3P_2 is the shortest arc that can exist in an anomalous periodic orbit. Consequently, at any particular r -value, there will be a maximum and a minimum number of twists around the z -axis that any anomalous Lorenz periodic orbit can make between passages into the region $z < r$. Furthermore, as r increases and ϵ decreases, both this maximum and minimum will increase (because the number of steps of length $O(\epsilon)$ on OP_1 and on P_3P_2 increases as ϵ decreases). Any particular anomalous Lorenz periodic orbit which exists for some large r value has a fixed number of windings around the z -axis. We deduce that any particular anomalous orbit exists only for a finite range of r -values, though we expect to find some anomalous orbits at any large enough r -value.

We can even predict how anomalous periodic orbits of the Lorenz equations will come into existence as r increases, and one way in which they disappear again at larger r -values. Consider some particular orbit which winds around the z -axis a maximum of n times between passages into the

region $z < r$. This orbit will first appear at some r -value where the orbit is represented by an averaged trajectory that passes, at some point, very close to the arc OP_1 of Fig. 7.6. In other words, it will appear in a homoclinic explosion. (Remember that the arc OP_1 represents the initial behaviour of the unstable manifold of the origin.) As r increases, the same Lorenz orbit will be represented by shorter and shorter anomalous periodic orbits of the averaged equations. At some r -value the orbit will be represented by an anomalous orbit which passes very close to the arc P_2P_3 . We may suggest that at this r -value there will be two close together anomalous periodic orbits of the averaged equations, both representing anomalous periodic orbits of the Lorenz equations, one of which is stable. As r increases the two Lorenz orbits will be represented by anomalous averaged orbits which come closer together and which approach P_2P_3 . Eventually there will not be any anomalous averaged orbits short enough to represent the two Lorenz orbits and we may be able to deduce that they have annihilated one another in a saddle-node bifurcation.

There is some hope that a careful study will allow us to describe more completely the various bifurcations which occur as r increases. We might, for example, expect to find some evidence of period doubling.

Finally, Figs. 7.5 suggest some of the quantitative differences we may expect for different parameter values which show anomalous behaviour. When we are just into the anomalous parameter range (Fig. 7.5(b)), we may expect relatively few anomalous orbits to exist at any one r -value. Here the arcs OP_1 and P_2P_3 will be close together. When we are well into the anomalous range (Fig. 7.5(c) and (d)), we can expect a wider range of anomalous orbits at any one r -value. In addition, some of these orbits will lie very close to the z -axis above the plane $z = r$ (because the trajectory out from (σ, σ) passes very close to the line $A = -B$) as well as passing close to the z -axis below the plane $z = r$.

7.4. SUMMARY

If the parameter $\lambda = \frac{(\sigma+1)}{(b+2)}$ is less than $\frac{2}{3}$, there is no interesting large r behaviour. If $\lambda > \frac{2}{3}$ there is a stable symmetric periodic orbit which exists for all r however large and which winds once around the z -axis. In $\lambda > 1$ this is the only orbit which persists for all r . In $\frac{2}{3} < \lambda < 1$ we have, in addition, a pair of non-symmetric, non-stable periodic orbits which do not wind around the z -axis. In this parameter

range the Hopf bifurcation does not occur and C_1 and C_2 remain stable for all r however large.

If λ is large enough (though exactly what is large enough probably depends on the values of σ and b) there are anomalous periodic orbits of the averaged equations and corresponding anomalous periodic orbits in the Lorenz system. These orbits will have long periods and will wind many times around the z -axis between passages into the region $z < r$. At any large r -value there will be some anomalous periodic orbits but any particular orbit will only exist for a finite range of r -values. At any large r -value there will be both an upper and lower limit on the number of times that a periodic orbit of the Lorenz equations, corresponding to an anomalous periodic orbit of the averaged equations, can wind around the z -axis between passages through the region $z < r$. As r increases, both of these limits increase; new periodic orbits with additional windings around the z -axis will appear in homoclinic explosions, and old periodic orbits (with an insufficient number of windings around the z -axis) will disappear, sometimes by pairwise annihilation in saddle-node bifurcations.

These conclusions agree with our previous study of the Lorenz equations at a non-anomalous parameter value ($\sigma = 10$ and $b = 8/3$). We will now look at the Lorenz equations for parameter values where the anomalous behaviour can be expected.

Chapter 8

Small b

Our investigation of the behaviour of the Lorenz equations for large r (Chapter 7) suggested that we would see qualitatively more complicated behaviour when the parameter b was small. We now investigate some aspects of this behaviour.

8.1. TWISTING AROUND THE z -AXIS

The Lorenz equations are:

$$\frac{dx}{dt} = \sigma(y - x)$$

$$\frac{dy}{dt} = rx - y - xz$$

$$\frac{dz}{dt} = xy - bz.$$

We can do some approximate analysis on the equations. Suppose that x and y are small. Then the product xy is small. If, in addition, b is small, then z changes only slowly. If we take z to be constant, the equations are reduced to a two-dimensional linear system in x and y . This system has one stationary point, $(0,0)$, and the eigenvalues of the system are:

$$\frac{1}{2}\{-\sigma - 1 \pm \sqrt{(\sigma+1)^2 + 4\sigma(r-z-1)}\}.$$

Providing $z > r - 1 + \frac{(\sigma+1)^2}{4\sigma}$, the eigenvalues are complex with negative real part. This implies that trajectories spiral in towards the z -axis. If $r - 1 + \frac{(\sigma+1)^2}{4\sigma} > z > r-1$, the eigenvalues are real and nega-

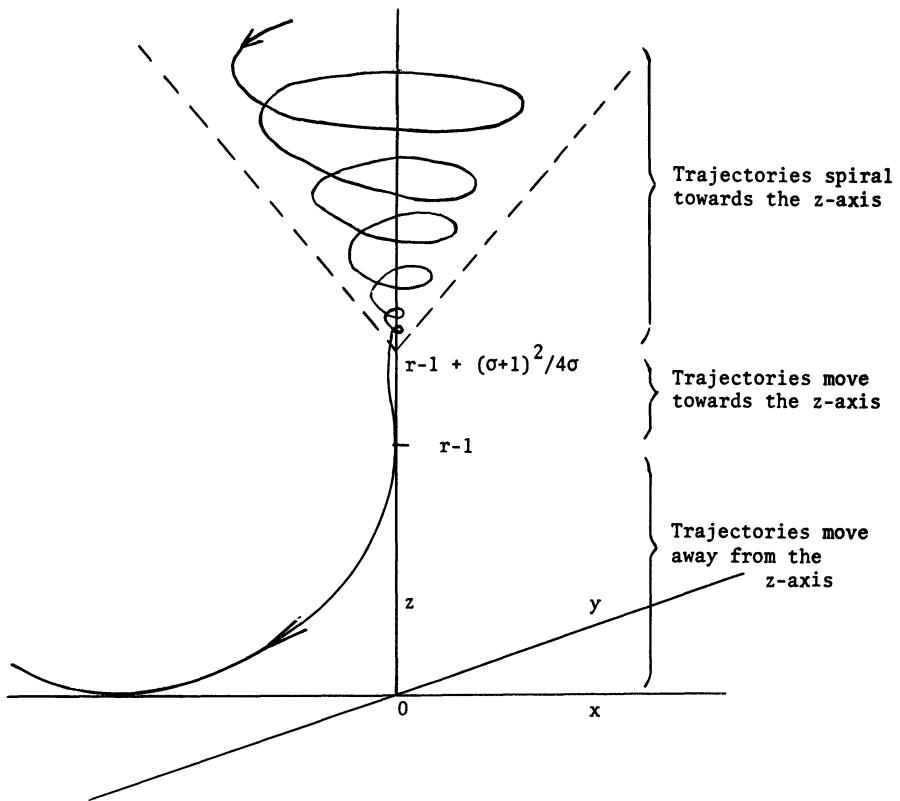


Figure 8.1. Trajectories started with large z and small x and y will spiral into the z -axis as z slowly decreases.

tive. Here, trajectories move towards the z -axis without spiralling. When $z < r-1$, one of the eigenvalues is positive and almost all trajectories move away from the z -axis.

Though this analysis is not rigorous when we allow z to vary, it is highly suggestive. We can expect that a trajectory, started with large enough z and small enough x and y , will behave roughly as shown in Fig. 8.1; i.e., we can expect there to be an inverted cone-shaped region inside which trajectories spiral in towards the z -axis while z slowly decreases.

The analysis above is relevant even for $b = 8/3$; trajectories started with large enough z and small enough x and y do behave in

the way suggested by Fig. 8.1. However, it seems that no interesting, non-transient trajectories enter the "cone" when b is as large as $8/3$. We could, if we had wished, have interpreted the fact that trajectories cross from side to side by saying that they entered the cone just enough to do one half-swing around the z -axis. This would not have added much to our general understanding.

When b is small, it seems that nearly all interesting trajectories enter the cone. This makes analysis of the behaviour more difficult and more interesting. We shall study the system when $b = 0.25$, keeping σ at 10.0. There is nothing special about these particular parameter values.

8.2. HOMOCLINIC EXPLOSIONS WITH EXTRA TWISTS

The results of Chapter 1 (simple properties of the system) and of Chapter 2 (analysis of homoclinic explosions and the first homoclinic explosion) are still relevant for $b = 0.25$. The Hopf bifurcation occurs for an r -value near 15.143 and the stationary points C_1 and C_2 are, therefore, stable in the interval $1 < r < 15.143$. There is a first homoclinic explosion at $r \approx 5.6$ which produces an "original", non-stable, strange invariant set, identical in all important respects to the one obtained in the first homoclinic explosion when b was $8/3$.

At $r \approx 7.7$. we see the first obvious "new" behaviour. The right-hand branch of the unstable manifold of the origin is shown for $r = 7.8$ in Fig. 8.2. The behaviour shown is "new" because the trajectory winds around the z -axis without "doing anything interesting" in $x < 0$. (We know that the trajectory actually goes round the z -axis because when $x = 0$, $\dot{x} > 0$ only if $y > 0$ and $\dot{x} < 0$ only if $y < 0$.) When b was $8/3$, the unstable manifold of the origin always looped around in $x < 0$ before crossing back into $x > 0$ (as did all other interesting trajectories). We can understand this as follows.

(Symbolic descriptions revisited.) In Chapter 6, we saw how periodic orbits and other trajectories might be described topologically with respect to the one-dimensional stable manifolds of the stationary points C_1 and C_2 (where by "stable manifold" we mean the one-dimensional manifold associated with the eigenvalue of the linearized flow near C_1 and C_2 that is real and negative for all $r > 1$). When $b = 0.25$, this method is even easier to justify than when b was $8/3$, since it appears (see Fig. 6.6) that these manifolds behave in a very simple way for all r -values up to approximately 500. Thus, periodic orbits may be described with

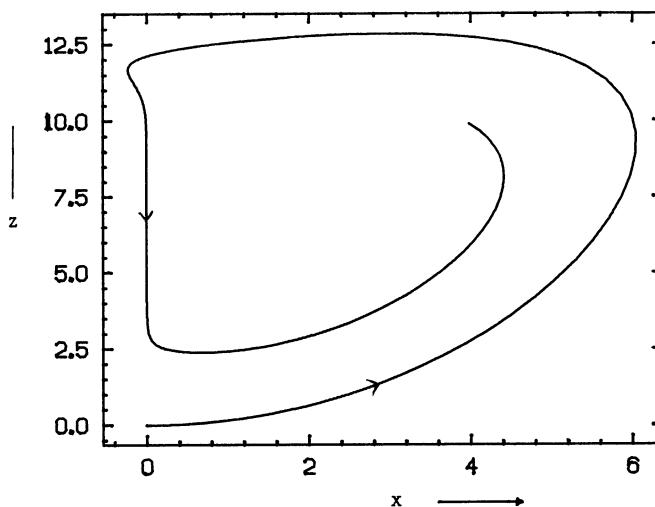


Figure 8.2. One branch of the unstable manifold of the origin. $b = 0.25$ and $r = 7.8$.

sequences of symbols corresponding to the successive loops which the orbit makes around the stable manifolds of C_1 and C_2 . The dynamics of the Lorenz equations at $b = 0.25$ (as at $b = 8/3$) seem to imply that we need only two symbols, since it appears that trajectories can only rotate in one direction around each stable manifold. As we mentioned in Chapter 6, we could have chosen to describe orbits with respect to three invariant manifolds, rather than two; the z -axis is an obvious choice for the third manifold. We would then have obtained symbolic descriptions containing three different symbols. However, when b was $8/3$ the z -axis symbol was redundant, because the dynamics of the equations appeared to determine that no interesting trajectories spiralled around the z -axis without looping at least once around the stable manifold of C_1 in $x < 0$ and at least once around the stable manifold of C_2 in $x > 0$. It is this dynamically determined simplification which no longer applies when $b = 0.25$.

Though we can still describe trajectories with only two symbols, it is no longer very useful to do so. A periodic orbit which winds once around the stable manifold of C_1 , and a periodic orbit which winds once around the z -axis and once around the stable manifold of C_1 , cannot be involved together in bifurcations; it is necessary that we distinguish between them. It is slightly more convenient to do this by adding sub-

scripts to the symbols we already have than by adding a third symbol. We will write x_n for a single loop around the stable manifold of C_2 (in $x > 0$) followed by n "half-twists" around the z -axis, and will write y_n for a single loop around the stable manifold of C_1 (in $x < 0$) followed by n "half-twists" around the z -axis. (Topologically speaking, a "half-twist" is not a well-defined concept. Dynamically speaking, we can get away with it because of the nice way trajectories behave on the plane $x = 0$.)

We can describe the periodic orbits from Chapters 2-5 with this new notation. Each subscript is either 0 or 1, and it is zero if the next symbol is the same, one if the next symbol is different. (To go twice around the same stable manifold we must not twist around the z -axis at all in between; to go around one stable manifold and then around the other we must go halfway around the z -axis in between.) Thus, our old xy orbit is now an x_1y_1 orbit, our old x orbit is now an x_0 orbit, and our old x^3y^2 orbit is now an $x_0x_0x_1y_0y_1$ orbit. The argument by which we determine subscripts for these old orbits suggests a rule for possible subscripts in more general orbits; if the subscript is odd, the next symbol must be different; if the subscript is even, the next symbol must be the same. So, x_2y_2 is not a possible orbit, but x_3y_3 is quite feasible.

Periodic orbits like the ones discussed when b was $8/3$, that require only 0 and 1 subscripts in their descriptions, we call *periodic orbits without extra twisting around the z -axis*. All other orbits do have extra twisting around the z -axis. Readers may like to look ahead to Fig. 8.12 which shows chaotic behaviour involving extra twisting around the z -axis.

Fig. 8.2 indicates that, for some r -value less than 7.8, we have an x_2 -homoclinic orbit. Numerical experiments indicate that as r increases, the unstable manifold of the origin winds more and more around the z -axis. By the time r is 12.0 there are four half-twists around the z -axis; trajectories started on the right-hand branch of the unstable manifold of the origin swing once around C_2 and then twist twice around the z -axis, emerging once more into $x > 0$ without swinging around C_1 . Somewhere in $7.8 < r < 12.0$, there is first an x_3 -homoclinic orbit and then an x_4 -homoclinic orbit. Fig. 8.3 shows the right-hand branch of the unstable manifold of the origin when $r = 5000$. Here, there is an immense number of half-twists around the z -axis, and when the trajectory first passes downwards through the plane $z = r-1$ it is very close to the

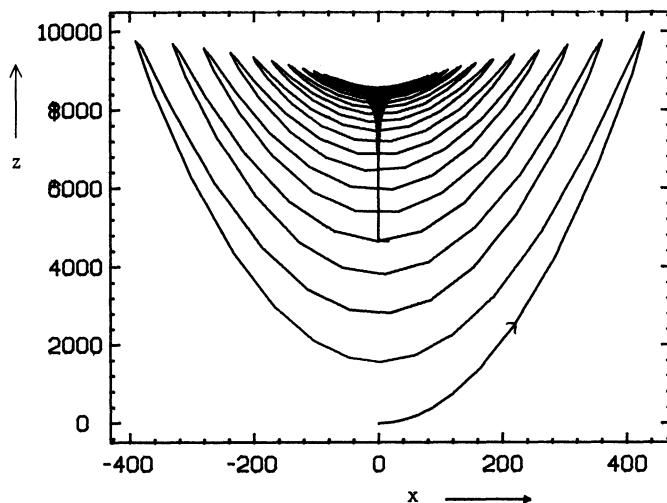


Figure 8.3. Unstable manifold of the origin for $r = 5000$, $b = 0.25$.

z -axis. This figure is a graphic illustration of the type of behaviour predicted in Section 8.1.

As r continues to increase, it seems that there are x_n -homoclinic orbits for ever-increasing values of n . (Though it is difficult to check that the trajectory in Fig. 8.3 doesn't wind around the stable manifold of C_1 while it is winding around the z -axis - because the stable manifold of C_1 behaves in a quite complicated way at $r = 5000$ - we can easily check this at r -values all the way up to 480; here there is already a very large number of twists, and the trajectory definitely does not twist around the stable manifold of C_1 in the process.) Each new x_n -homoclinic orbit will be associated with a homoclinic explosion. There is nothing about the extra twisting around the z -axis which changes our analysis of homoclinic explosions in Chapter 2 and Appendix D. We need to be a little more careful in deciding what the symbolic descriptions of the various orbits and trajectories in the associated strange invariant set actually are, but we can still be sure that each homoclinic explosion either produces or destroys some strange invariant set. The x_n -homoclinic explosions are all likely to produce strange invariant sets (as r increases); there are unlikely to be any orbits or trajectories with a large enough number of extra twists around the z -axis available to be destroyed. Simple arguments, in the spirit of Chapter 2, show that the x_2 -homoclinic

explosion is a type (b) birth and that the x_3 -homoclinic explosion is a type (a) birth. In fact, the types of the successive x_n -homoclinic explosions continue to alternate as n increases (if we assume that they are all births).

If we accept the foregoing, then we have evidence for the first of the predictions from Chapter 7. However large r becomes, we expect that new homoclinic explosions will continue to produce new periodic orbits. Furthermore, these orbits will twist many times around the z -axis between each passage through the region $z < r-1$ (as predicted) and will spend much of their periods near to the z -axis (as predicted). The maximum number of twists around the z -axis (occurring between passages through the region $z < r-1$) will increase as r increases.

The x_n -homoclinic explosions are not the only ones to occur as r increases. There must actually be an infinite number of homoclinic explosions between each pair of successive x_n -explosions. We can see this using Appendix D style arguments about the behaviour of the unstable manifold of the origin on either side of each homoclinic explosion. Just after the first homoclinic explosion at $r \approx 5.6$ (the x_1 -explosion), $k(r)$, the symbolic sequence which represents the right-hand branch of the unstable manifold of the origin, equals $x_1y_0y_0y_0\dots$ (as when $b = 8/3$). As we approach the x_2 -explosion $k(r) = x_1y_1x_1y_1x_1\dots$. If we approach the x_2 -explosion from above, $k(r) = x_2x_1y_1x_1y_1\dots$, and as we approach the x_3 -explosion, $k(r) = x_2x_2x_2x_2x_2\dots$. We can work out similar results for the values of $k(r)$ before and after all of the x_n -homoclinic explosions; they fall in an obvious way into one of the four patterns given above, depending on whether n is odd or even. We always find that the value of $k(r)$ immediately after the x_{n-1} -explosion is not the same as the value of $k(r)$ immediately before the x_n -explosion. To get from one to the other we need an infinite sequence of homoclinic explosions.

These other homoclinic explosions are not easy to detect numerically. When r is small (for instance between the x_1 and x_2 explosions) the stationary points C_1 and C_2 are still stable. Most numerically calculated trajectories tend rapidly to one or other of these points. For all values of r (particularly large ones) homoclinic orbits other than the x_n ones will only occur after the unstable manifold of the origin has passed near the origin one or more times. These trajectories are difficult to calculate for the reasons given in Appendix F.

8.3. PERIODIC ORBITS WITHOUT EXTRA TWISTING AROUND THE z-AXIS

Between the x_1 -homoclinic explosion and the x_2 -homoclinic explosion, the sequence $k(r)$ changes from $xyyyyyy\dots$ to $xyxyxyx\dots$. This is the whole change that the Lorenz system managed in the parameter range $b = 8/3$, $13.926 < r < \infty$. It is also the change we studied in Section 5.6 (and Appendix J), when we looked at arbitrary sequences of homoclinic explosions. If the conjectures of Chapter 5 are true, then there is some sequence of homoclinic explosions in $r < 7.8$ which produces:

- (i) The Hopf, x and y , orbits.
- (ii) Exactly the right number and type of periodic orbits for an infinite number of period doubling windows which end in saddle-node bifurcations.
- (iii) Exactly the right number and type of periodic orbits for a period doubling window which leaves us with a stable symmetric xy orbit.
- (iv) No other orbits.

All these orbits will be orbits without extra twisting around the z -axis.

This suggests that when $r = 7.8$ ($b = 0.25$), we already have enough periodic orbits in the system to produce some kind of replica of the entire sequence of different behaviours seen in the Lorenz system for $b = 8/3$, $1 < r < \infty$. We can see, from the results of Chapter 7 and from the considerations in the last section, that such a sequence of behaviours will more or less inevitably happen. The Chapter 7 results suggest that no one periodic orbit (except the stable symmetric xy) can continue to exist for all large r . The discussion in the last section suggests that all homoclinic explosions occurring after the x_2 -explosion will involve only periodic orbits with extra twisting around the z -axis. (If the homoclinic orbit winds many times around the z -axis, so will all the periodic orbits produced or destroyed in the associated homoclinic explosion.) Consequently, the orbits listed above must disappear somehow as r increases, and they cannot disappear in homoclinic explosions. It seems likely that they will disappear in the same way they did when b was $8/3$.

Can we find any numerical evidence for the suggestions above? We can, if we look carefully, establish that there are some (two or three) homoclinic explosions between the x_1 -explosion and the x_2 -explosion. As mentioned before, these experiments are not easy to do because the stationary points C_1 and C_2 are stable. Period doubling windows are

difficult to locate for the same reason. One technique for locating period doubling windows is to adapt the orbit following techniques described in Appendix E so that orbits can be followed when the parameters r and b both change. Then, an orbit can be located at $b = 8/3$ and some suitable r -value, and followed down into $b = 0.25$. (This technique also serves to convince us that we are looking at the "same" orbits when $b = 0.25$ that we studied when b was $8/3$.) This technique was used to locate the symmetric xy orbit at $b = 0.25$ and we can establish that this orbit has gained its stability (the event which marked the end of numerically observable interesting behaviour when $b = 8/3$) by an r -value approximately equal to 12.8. At this r -value C_1 and C_2 are still stable! (Notice that, while we expect all the orbits for the various period doubling windows to have been produced by the time the x_2 -explosion occurs at $r \approx 7.7$, we do not expect all the period doubling windows to have occurred at this r -value. The fact that the symmetric xy orbit is stable by $r = 12.8$ suggests that all the period doubling windows involving original periodic orbits will be over by this larger r -value.) As far as we can see, the xy orbit gains its stability in the usual way. Fig. 8.4 shows the stable symmetric xy orbit at $r = 12.8$ and a stable non-symmetric xy orbit at $r = 12.0$ (cf. Figs. 4.11). The same figure shows stable (non-transient) chaotic behaviour at $r = 10$.

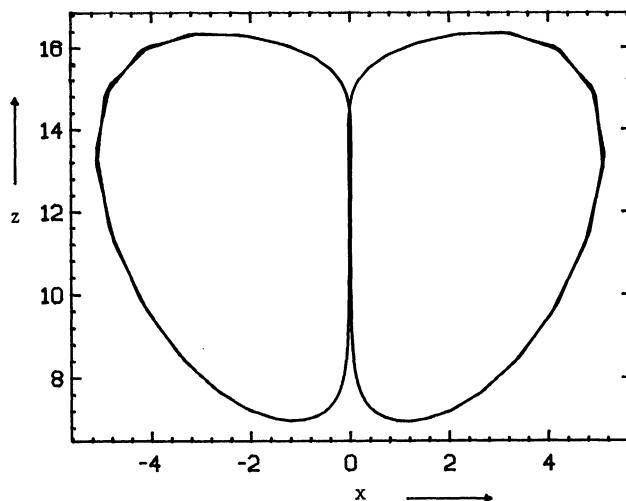


Figure 8.4. Some stable behaviours observed in a parameter range in which C_1 and C_2 are still stable. (a) Stable symmetric xy orbit at $r = 12.8$.

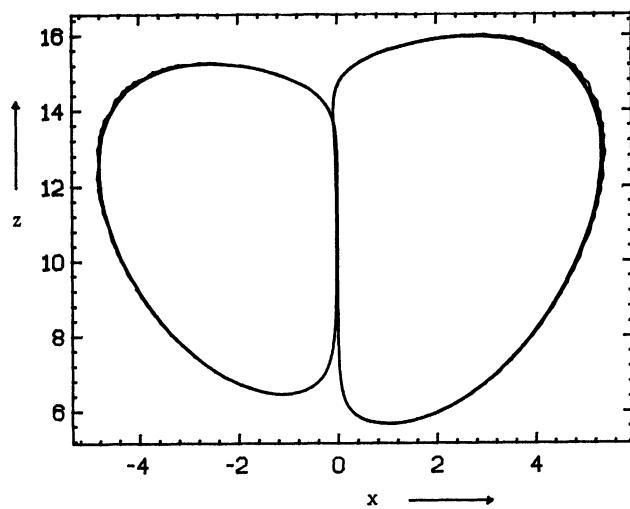


Figure 8.4. (b) Stable non-symmetric xy orbit at $r = 12.0$.

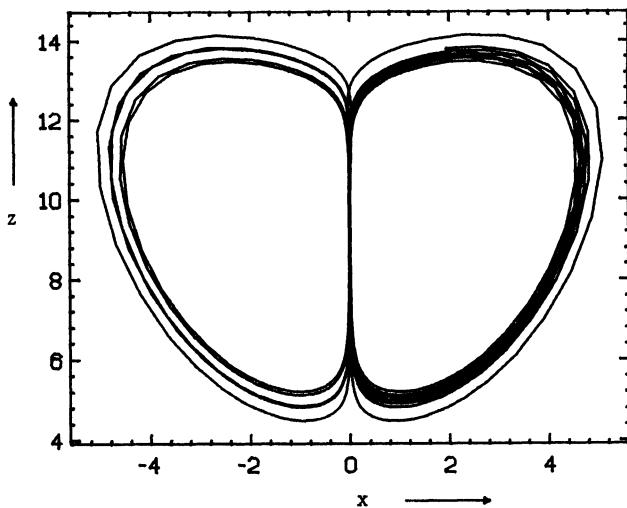


Figure 8.4. (c) Chaotic behaviour at $r = 10.5$.

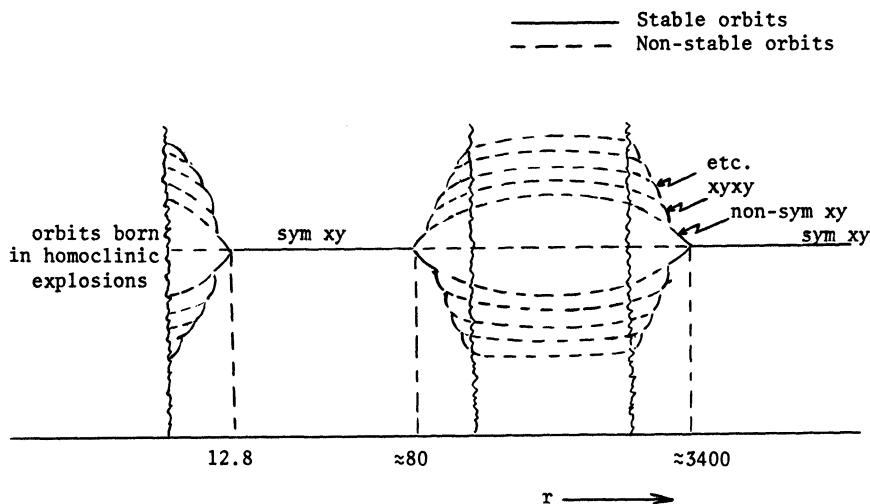


Figure 8.5. Bifurcation diagram for the xy orbit. $b = 0.25$.

When b was $8/3$, the final stable symmetric xy orbit seemed to retain its stability forever once it had become stable. When $b = 0.25$ this is not the case. The orbit remains stable up to $r \approx 80$, when it loses its stability in a period doubling window. (Unlike the other period doubling windows studied in these notes, the period doubling here occurs as r increases.) It seems reasonable to assume that the periodic orbits produced in this window remain in existence up to $r \approx 3500$ when there is another period doubling window (proceeding in the usual direction) which returns stability to the symmetric xy orbit. This then seems to retain its stability for all larger r -values. The bifurcation diagram for this xy orbit is shown in Fig. 8.5.

Before asking any general questions about periodic orbits with extra twisting around the z -axis, and before investigating the theoretical and numerical behaviour of the system in the parameter ranges which we have not investigated, it is interesting to ask, "Do we know, more or less, the fate of all the orbits without extra twisting around the z -axis?" At first sight it appears that in considering all the homoclinic explosions up to the x_2 -explosion we have considered all the explosions which involve such orbits. We also have, from the conjectures in Chapter 5, an explanation for the fate of each of the orbits produced in these explosions. Also, though it is apparent from Fig. 8.5 that we can expect there to be sequences of bifurcations involving periodic orbits without extra twisting at larger r -values, it seems reasonable to assume (without asking for details) that any such sequence reverses itself at yet larger r -values (as with Fig. 8.5). (It seems reasonable to assume this because of the result in Chapter 7 which says that as r increases there will be an increasing minimum number of extra twists in orbits which remain in existence; consequently, all orbits without extra twists - except the stable xy orbit - must have disappeared by some large r -value.) However, there is (at least) one thing which we have overlooked.

We know that the x_2 -homoclinic explosion produces a strange invariant set, and can work out the symbolic representations of the periodic orbits contained in this set. The simplest orbits in the set are x_2 and y_2 as we might expect. The simplest symmetric orbit is an x_1y_1 orbit ($= xy$). A moment's thought should convince us that this is true, even without checking rigorously; the simplest symmetric orbit has to be something, and we have already said that the other obvious choice, x_2y_2 , is not a possible orbit.

This leaves us with a problem. This "extra" xy orbit is the only member of the x_2 -generated invariant set without extra twisting around the z -axis. It cannot be involved in any subsequent homoclinic bifurcations (because these all involve orbits with extra twisting around the z -axis), it cannot continue to exist for large r as this would contradict the results of Chapter 7, it cannot be involved in any bifurcations with orbits with extra twisting around the z -axis (for topological reasons), and we think we have already explained the fate of all the other orbits without extra twisting, so it cannot be involved in a bifurcation with any of these orbits. If our (admittedly conjectural) arguments have any force, we must look for some "new" bifurcation, not previously observed in the Lorenz system, if we are to explain what happens to this

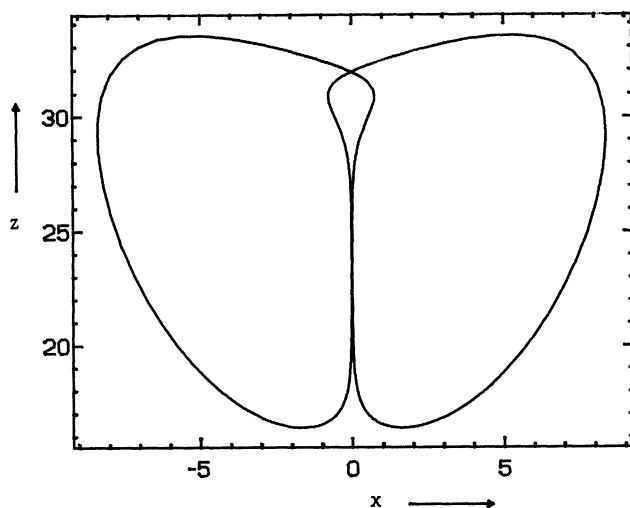


Figure 8.6. (a) The "extra" symmetric xy orbit born in the x_2 -homoclinic explosion. $r = 26.0$.

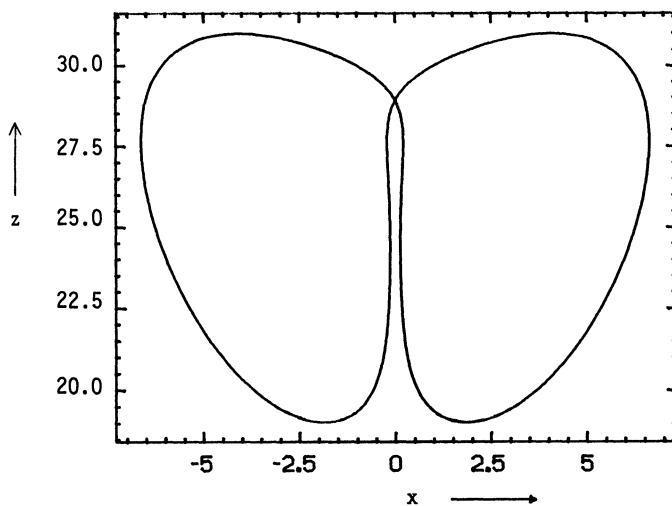


Figure 8.6. (b) The final symmetric xy orbit which is stable at this parameter value. $r = 26.0$.

orbit. Before proceeding we check, numerically, that the "extra" xy orbit actually exists. Fig. 8.6 shows the orbit (non-stable) at $r = 26.0$. For comparison, Fig. 8.6 also shows the other (stable) xy orbit at the same r-value.

8.4. HETEROCLINIC ORBITS BETWEEN C_1 AND C_2

The extra xy orbit, shown in Fig. 8.6, is non-stable. Using the orbit following techniques described in Appendix E, we can follow it all the way up to $r \approx 1353$. At this r-value it is annihilated, in a saddle-node bifurcation, by a stable periodic orbit. Our problem now is to explain where this orbit came from. We can follow the annihilating orbit back down to $r \approx 211$ where it is involved in another saddle-node bifurcation with a non-stable orbit. This process continues, and Fig. 8.7 shows a bifurcation diagram calculated by following these periodic orbits back and forth. The curve could have been followed further but the calculations become both predictable and tedious.

Each saddle-node bifurcation on this curve (Fig. 8.7) involves one non-stable periodic orbit and one stable periodic orbit, so every alternate "horizontal" section of the curve represents a "stable branch". It seems that the periodic orbits represented by these "stable branches" do not retain their stability for the whole distance between the two saddle-node bifurcations. In fact, they only remain stable in fairly short r-intervals near to the bifurcations. In between, we can assume that there are mirror-image period doubling windows, like the one shown in Fig. 8.5.

In many respects, this bifurcation curve should not be a surprise. Bifurcation curves cannot just stop, and once we have assumed that the curve is bounded away from large r-values (because there are no extra non-twisted orbits at large r), and that it is bounded away from small r-values (because we think we have discovered all the xy orbits which exist at small r), and that it cannot meet up with any other bifurcation curves (because we do not believe there are any other orbits around that it can bifurcate with), there are only two possible things that it can do. Either the orbit represented by the bifurcation curve must eventually become unbounded in amplitude, or it must become unbounded in period. (In four-dimensional systems, bifurcation curves can loop round and join up with themselves (Alligood, Mallet-Paret & Yorke, 1981). This requires the bifurcation where the loop joins up to be a period doubling bifurcation. It is reported that Alexander & Yorke have shown that this cannot

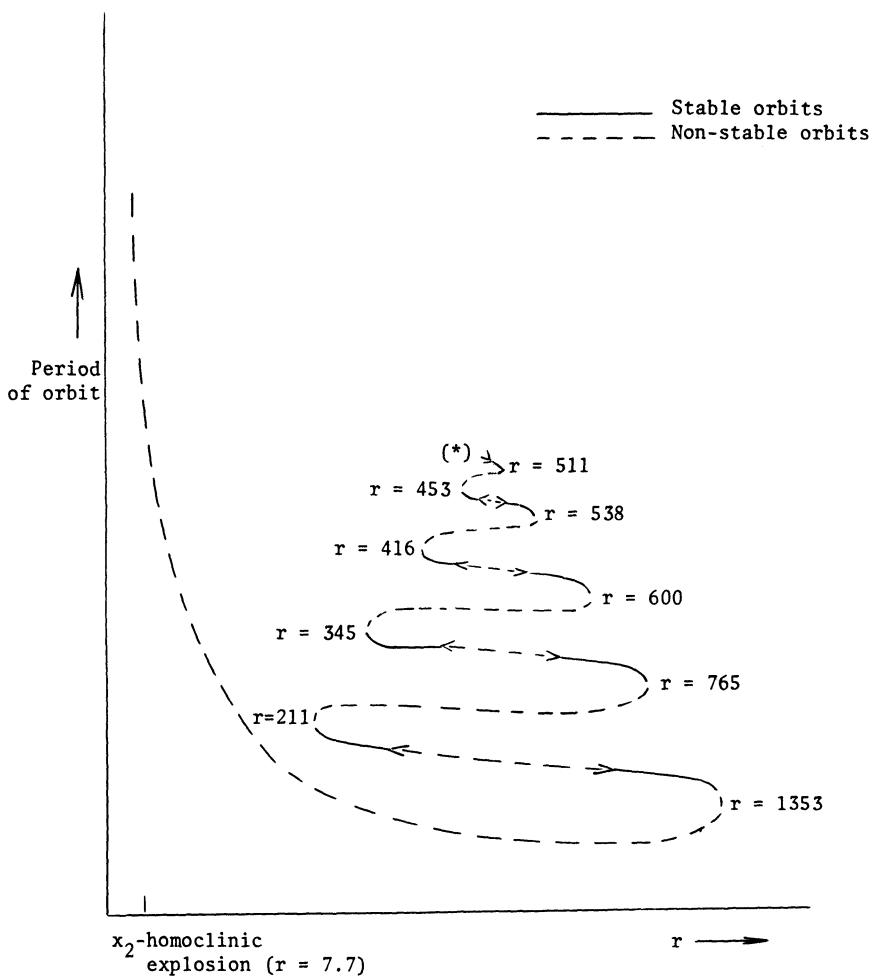


Figure 8.7. Bifurcation diagram for the "extra" xy orbit.

occur in a three-dimensional system. We also know that it cannot happen in the Lorenz system by counting the number of windings around the z -axis.) Periodic orbits cannot have unbounded amplitude in the Lorenz system. Consequently, the orbit represented by the curve in Fig. 8.7 should have period which tends to infinity as we follow the curve along. It is certainly true that the period of the orbit does increase as we follow Fig. 8.7; presumably it would continue to increase if we followed the curve further.

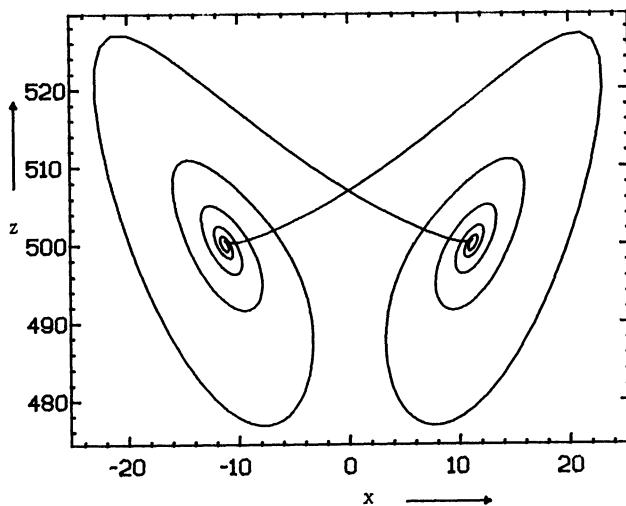


Figure 8.8. The orbit represented by the point (*) on Fig. 8.7. $r = 501$.

The only simple way for the period of an orbit to become infinite is for the orbit to pass very close to one or more of the stationary points. That is, after all, how the orbit came into existence - with infinite period at a homoclinic explosion. Our investigation of the unstable manifold of the origin makes it seem very unlikely that an orbit without extra twisting around the z-axis will pass close to the origin at the r-values we are considering. This leaves the points C_1 and C_2 . Fig. 8.8 shows the periodic orbit at the point where we stopped following the bifurcation curve of Fig. 8.7. The orbit passes moderately close to both C_1 and C_2 ; it seems to approach the points close to their one-dimensional stable manifolds, and then to spiral away from the points close to their two-dimensional unstable manifolds.

The first thing to notice about Fig. 8.8 is that the orbit shown does not look like an xy orbit; rather, it looks like either an x^5y^5 orbit or an x^6y^6 orbit. (Notice, though, that it is still an orbit without extra twisting around the z-axis; if this were not the case, we would have to conclude that we had made a mistake in following the orbit.) We can see that something has "gone wrong" with our symbolic descriptions. It is here that our discussion of symbolic descriptions (Chapter 6) is immediately relevant. If we assume that for some critical value, r^* , of r (somewhere in $480 < r < 500$), there are heteroclinic orbits between C_1

and C_2 , then we have no difficulty explaining how our symbolic descriptions have failed (see Chapter 6). This assumption also serves to explain Fig. 8.7. We can assume that the complete version of Fig. 8.7 looks something like Fig. 8.9, the bifurcation curve doing a kind of "damped harmonic motion" about the critical r -value as the period tends to infinity. As we follow the curve further and further, the periodic orbit which it represents will pass closer and closer to C_1 and C_2 ; we can imagine that the orbit is winding itself up around the stable manifolds of C_1 and C_2 . As the period of the orbit increases it will move closer to the heteroclinic orbits which exist at $r = r^*$. The heteroclinic orbits are trajectories which tend towards one of the stationary points in forwards time (along the stable manifold) and towards the other stationary point in reverse time (spiralling in on the two-dimensional unstable manifold of the relevant stationary point).

There are many details of the behaviour which we have not studied for r -values near r^* . However, we have explained the fate of the extra xy orbit. Notice that nearly all aspects of our understanding of the Lorenz equations have entered into the discovery of this heteroclinic bifurcation (which we will discuss further below). We can attempt to confirm the existence of the bifurcation using more direct numerical methods, but all the experiments fail to be convincing in one way or another. Numerical approximation of the stable manifolds involves integrating the equations in reverse time (when they are very non-dissipative) and we cannot hope to calculate a good approximation of a heteroclinic orbit this way. Attempts to calculate approximations to the heteroclinic orbits by approximating the unstable manifolds of C_1 or C_2 (which involves integration in forward time) would fail because these manifolds are two-dimensional and we have no way of deciding which of an infinite number of trajectories to select. We can see that the behaviour of the stable manifolds does change as r passes through r^* . Fig. 6.6 showed the stable manifold of C_2 at $r = 480$ and $r = 500$. However, we cannot conclude from that figure that there is a heteroclinic bifurcation for some r^* in $480 < r < 500$. In our discussion of the behaviour of the stable manifolds for $b = 8/3$ (Chapter 6) we saw that the behaviour could change very abruptly, but argued there that this change did not imply the existence of a bifurcation.

Direct numerical simulation of the equations at r -values near r^* , for various initial conditions, does not (so far) produce a convincing indication of the behaviour we have been discussing. This is true despite

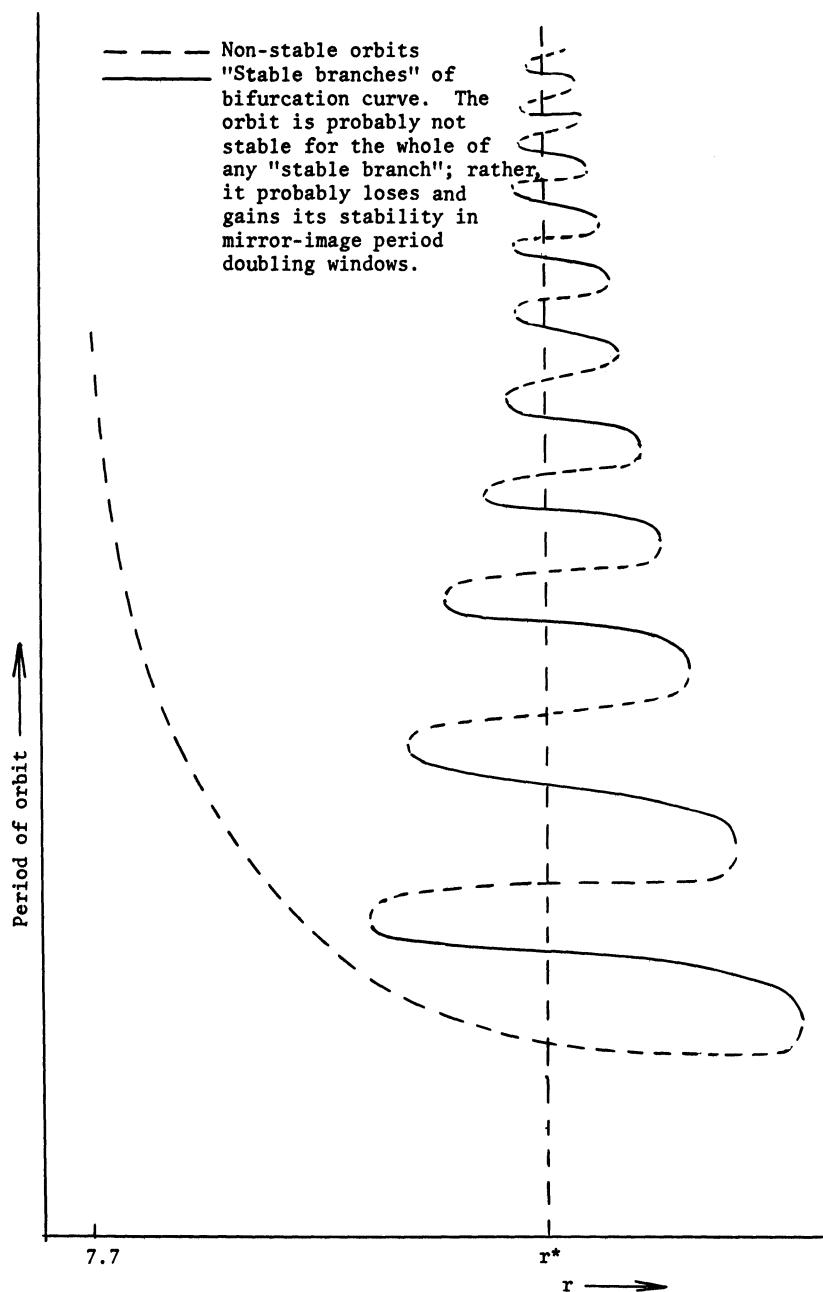


Figure 8.9. Probable bifurcation diagram for the "extra" orbit.

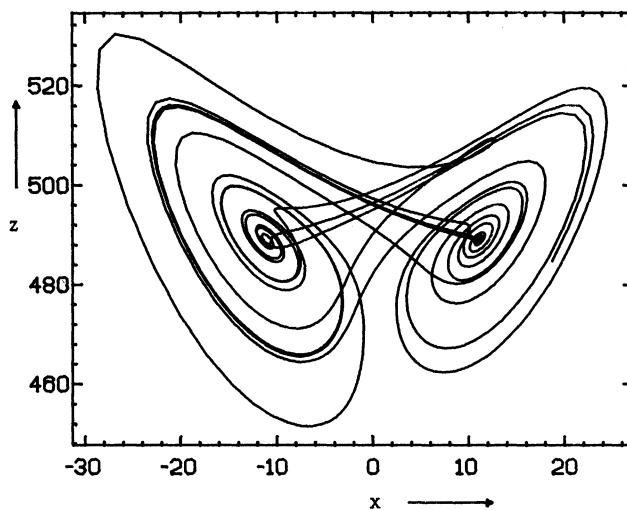


Figure 8.10. Chaotic behaviour at $r = 490$, $b = 0.25$.

the expected existence of many relevant stable periodic orbits (see Fig. 8.7) at r -values near r^* . Fig. 8.10 shows chaotic behaviour at $r = 490$.

The only real evidence we have for the heteroclinic bifurcations is the fact that we can locate the orbit represented by the bifurcation curves in Fig. 8.7 and 8.9, and that the symbolic description of this orbit with respect to the stable manifolds of C_1 and C_2 changes as r moves back and forth through r^* . More detailed numerical study of the parameter range $480 < r < 500$ may turn up more convincing numerical evidence but we would have been unlikely to select the parameter range for special attention without having gone through all the arguments in this chapter.

8.5. HETEROCLINIC BIFURCATIONS

Complete results for these bifurcations are not available, though some features of the behaviour for r near r^* are understood. The studies by Silnikov (1965, 1970) of homoclinic orbits to a single saddle/focus type stationary point, can be extended, with the help of the symmetry, to cope with heteroclinic trajectories between two such points. Recent papers by Arnéodo et al. (e.g., 1981b) also contain relevant information.

Tresser (private communication) reports that it is unlikely that there is just a single r -value for which we have heteroclinic trajectories;

this seems reasonable if we consider how the behaviour of the stable manifolds of C_1 and C_2 must change as we approach the critical r -value r^* . The heteroclinic trajectories at $r = r^*$ will be special because they are the "simplest" heteroclinic trajectories, and because the bifurcation associated with these trajectories will be the only one to produce a periodic orbit (the extra xy orbit) which has a bifurcation curve which exists for r -values very far from r^* and which appears and disappears in different kinds of bifurcation. The other heteroclinic bifurcations, occurring for r -values near r^* and on both sides of r^* , will involve heteroclinic trajectories which wander back and forth several times between C_1 and C_2 before settling down to one of the stationary points. There are also likely to be homoclinic orbits at r -values on both sides of r^* , formed by trajectories which wander back and forth between C_1 and C_2 before finally settling down to the stationary point at which they originated. All of these heteroclinic and homoclinic bifurcations are likely to occur in pairs, one on either side of r^* , and the periodic orbits which appear in one bifurcation of the pair are likely to disappear in the other bifurcation of the pair. (These statements are conjectural, and rely on our assumption that we have, in a sense, already accounted for all periodic orbits without extra twisting around the z -axis; the periodic orbits involved in all the heteroclinic and homoclinic bifurcations we are discussing will all be orbits without extra twisting around the z -axis, and, with the exception of the orbit involved in the heteroclinic bifurcation at r^* , it seems reasonable to assume that they must all appear and disappear in similar bifurcations.)

It is difficult to predict the behaviour of "most" trajectories (either theoretically or numerically) for r -values near r^* . Besides the orbits directly involved in heteroclinic or homoclinic bifurcations, there are likely to be many more orbits which appear in saddle-node bifurcations and subsequent period doublings, though these orbits are, again, likely to disappear in the same way that they appear. For instance, between the mirror image period doubling windows on each of the "stable branches" of the bifurcation curve shown in Figs. 8.7 and 8.9, we can expect a whole sequence of period doubling windows (similar to those studied in Chapter 4) which develop as r increases and then reverse themselves again. In addition to this behaviour existing near to C_1 and C_2 , we will have whatever behaviour we can expect from sequences of period doubling windows occurring because the final xy symmetric orbit is non-stable at these parameter values (Fig. 8.5), the behaviour from the lower

"stable branches" of bifurcation curves shown in Fig 8.7 or 8.9 (which involve orbits which are not close to C_1 or C_2), and behaviour due to the multitude of different sequences of bifurcations involving periodic orbits with extra twisting around the z-axis, which we have not even started to consider. There is much to be done!

8.6. GENERAL BEHAVIOUR WHEN $b = 0.25$

We have already seen that the behaviour of the Lorenz equations is rather more complicated when $b = 0.25$ than it was for $b = 8/3$. In addition, numerical experiments are more difficult to perform when $b = 0.25$, and there are numerous reasons to believe that important aspects of the behaviour will not show up however many experiments we do. Rather than attempting to give a detailed description of everything that happens, let us discuss each of the general approaches that are available and indicate the kind of general statements they will allow us to make.

1. Periodic Orbits From Homoclinic Explosions (cf. Section 5.6)

We have used the arguments from Section 5.6 to lead us to the heteroclinic bifurcation. So far, we have only applied these arguments to orbits without extra twisting around the z-axis, and have only dealt with homoclinic explosions occurring in the parameter range $r < 7.8$. We can speculate about the likely result of similar arguments applied to the infinite and never ending sequence of twisted homoclinic explosions which occurs as r increases.

It makes sense to break the sequence of homoclinic explosions into blocks, each block starting with an x_n -homoclinic explosion and finishing with an x_{n+1} -explosion. All the periodic orbits produced in such a block will have the property that the maximum number of half-twists they make around the z-axis between loops around the stable manifolds of either C_1 or C_2 is n . Furthermore, each such orbit will make n half-twists at least once; n is a maximum which is attained by each orbit. Hence, these orbits can only bifurcate (for topological reasons) with other orbits produced in the same block of homoclinic explosions. The large r results (Chapter 7) suggest that there is some large enough r_n such that all the periodic orbits produced in the n^{th} block of homoclinic explosions will have disappeared by $r = r_n$. This is because there is an increasing minimum for the number of half-twists in periodic orbits which remain as r increases. These observations suggest that combinatorial arguments (in the style of Chapter 5 or Appendix J) will lead us to some

conclusion which is almost, "Each block of homoclinic explosions leaves us with the right number and type of periodic orbits that they can all disappear in period doubling windows which terminate in saddle-node bifurcations." The extent to which we cannot draw this conclusion will indicate the extent to which we must look for new types of bifurcation. Two points should be made:

i) A Chapter 5-type counting analysis for the n^{th} block of explosions ($n > 1$) will be more complicated than the analysis for the first block. This is because, combinatorially speaking, homoclinic explosions may occur in the n^{th} block which require symbolic descriptions where the subscripts range over all values between zero and n . An $x_4x_3y_2y_1x_0$ -homoclinic explosion is quite possible, and may well occur for some value of r . Dynamically speaking, it appears that for large enough n (and large enough r), there will be a lower limit to the subscripts which can occur, though the range of possible subscripts is likely to be large. Once again, this is because the large r results tell us there is an increasing minimum to the number of half-twists allowed in periodic orbits, and hence an increasing minimum to the number of half-twists which can occur in a homoclinic orbit. We can understand this intuitively if we argue that trajectories which twist an insufficient number of times around the z -axis arrive at the plane $z = r-1$ too far from the z -axis to re-enter the inverted twisting cone (Section 8.1) on their next swing away from the z -axis; they will, presumably, be attracted instead to the stable symmetric xy periodic orbit. In considering periodic orbits with extra twists around the z -axis, we must also consider the possibility that the heteroclinic bifurcation we have described changes the symbolic descriptions of some of these orbits. It is to be hoped that it will either change all the symbolic descriptions in some obvious way, or that it will change none of them.

ii) We cannot claim, *a priori*, that a Chapter 5-type analysis of the n^{th} block is likely to suggest we should look for "new bifurcations". Indeed, the contrary is true. The first block, which we have already studied, is special because it is the first block. The respects in which this block failed to produce only period doubling windows terminating in saddle-node bifurcations were (a) it produced the simple orbits, x and y , for the Hopf bifurcations, and (b) it produced the stable xy symmetric orbit which persists for all r . If we look at the orbits which correspond to these "special" orbits in the $(2n)^{\text{th}}$ and $(2n+1)^{\text{th}}$ blocks we see:

(a) the x_{2n} -explosion produces two simple orbits x_{2n} and y_{2n} . However, the x_{2n+1} -explosion also produces x_{2n} and y_{2n} orbits (x_{2n+1} and y_{2n+1} are impossible) and providing we get enough other orbits for a complete period doubling window we can assume that these two pairs of orbits eventually annihilate one another in a saddle-node bifurcation. Hence, we need not look for an equivalent of the Hopf bifurcation for orbits with extra twisting around the z -axis; (b) the x_{2n+1} -explosion produces an $x_{2n+1}y_{2n+1}$ orbit as the simplest symmetric member of its strange invariant set. But the x_{2n+2} -explosion also produces a symmetric $x_{2n+1}y_{2n+1}$ orbit ($x_{2n+2}y_{2n+2}$ is impossible) and these two orbits will probably annihilate one another in a saddle-node bifurcation following a period doubling window. Consequently, we have no immediate reason to look for more heteroclinic bifurcations (perhaps involving extra twists around the z -axis), and we have no need to look for bifurcations to account for the non-appearance of additional stable symmetric orbits which persist for all r however large. These combinatorial arguments are supported by rather different arguments which follow. (Note: if it were not for the necessity of keeping one symmetric xy orbit for all r however large, we would have had no problem with our xy orbits. The extra xy orbit and the final xy orbit could just have disappeared in a saddle-node bifurcation. In this sense, the existence of the unique periodic orbit which remains stable for all large enough r implies the existence of the (possibly unique) collection of heteroclinic orbits for r -values near r^* .)

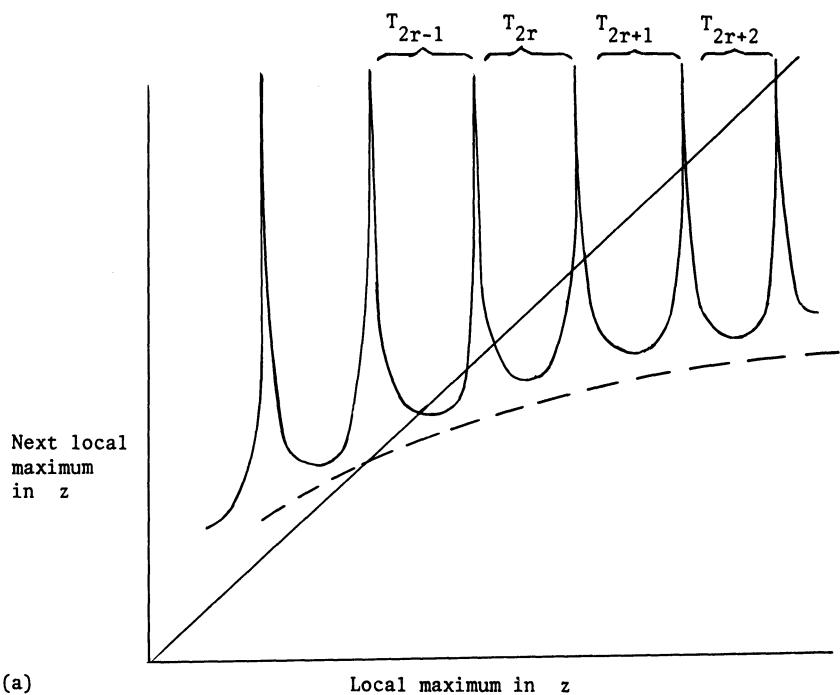
2. Analytic Calculation of One-Dimensional Maps

Fowler & McGuinness (1981a,b) studied the Lorenz equations in a limit, $\sigma \sim r \gg 1$, $b \sim 1$, in which trajectories show the twisting behaviour discussed in Section 8.1. Fowler & McGuinness analyzed this behaviour by separating the slow (twisting) decay of the variable z near the z -axis and the fast pulse which increases z again after the trajectory passes close to the origin. Using the ideas of matched asymptotic expansions, they managed to calculate a multiply-cusped one-dimensional return map for successive local maxima in z (cf. Appendix H). These results suggest that the appropriate title for this chapter might be "Small $\frac{b}{\sigma}$ ", as mentioned also in Chapter 7. Though their analysis is only valid for parameters $\sigma \sim r$, the way in which their multiply-cusped maps develop as r increases (at fixed σ) is very suggestive. Fig. 8.11 shows one-dimensional maps similar to the ones calculated by Fowler & McGuinness

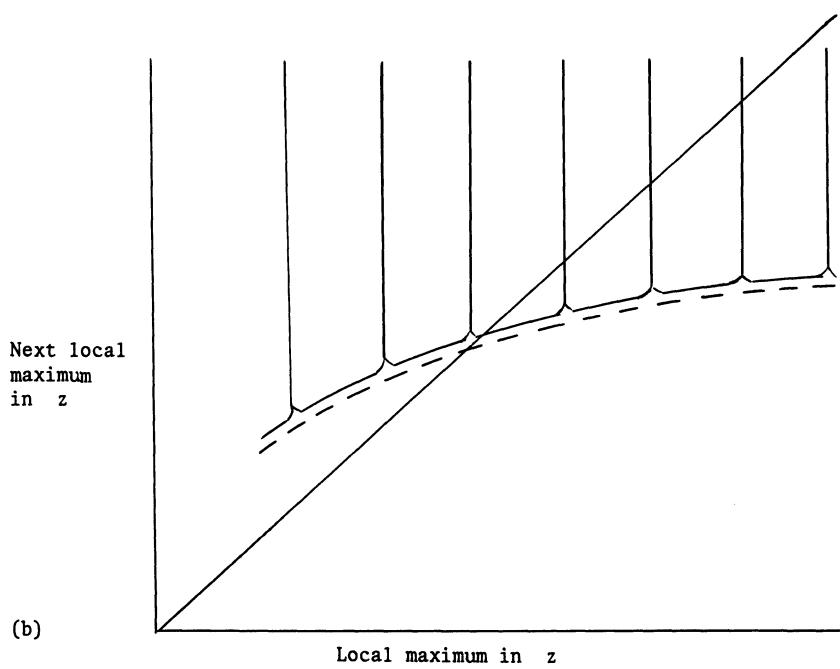
(1981a): the formulae for these maps are very complicated. The maps can be thought of as consisting of an envelope (shown dotted) with a series of cusps superimposed. In Fig. 8.11(a), we can see various fixed points of the map. These require careful interpretation since, by considering only the variable z , we have lost the symmetry of the Lorenz equations. If we suppose that the fixed points in the region we have marked T_{2r-1} correspond to $x_{2r-1}y_{2r-1}$ symmetric orbits, then the fixed points in T_{2r} each correspond to a symmetric pair, x_{2r} and y_{2r} , of non-symmetric orbits. As we move to the right, alternate troughs will contain fixed points corresponding to single symmetric orbits; the other troughs contain fixed points which correspond to symmetric pairs of non-symmetric orbits. The number of twists in these orbits increases as we move to the right.

If we fix σ and increase r , the cusps move to the left (Fowler & McGuinness, 1981a). As a new cusp appears on the right, we obtain two new fixed points; the appearance of the cusp corresponds to an x_n -homoclinic explosion for some n . Also, as the cusps move left, the two fixed points in T_{2r-1} (Fig. 8.11(a)) move closer together; eventually, the leftmost one becomes stable and then, for a larger r -value, they both disappear in a saddle-node bifurcation. Notice how this series of observations corresponds to our combinatorial arguments above (this section); the pattern of production and destruction of x_{2n+1} , y_{2n+1} , and $x_{2n}y_{2n}$ type orbits is exactly as we described it.

If b/σ is very small, the cusps become exceedingly narrow (Fowler & McGuinness, 1981a). See Fig. 8.11(b). In this case, we can expect one of the fixed points to be stable for nearly all r -values; the exceptional r -values will lie in short intervals for which a cusp lies over the intersection of the envelope with the 45° line. Fowler & McGuinness (1981b) show that there can be two different kinds of transition from one stable point (corresponding to a periodic orbit) to another as the cusps move to the left. Which of these two types of transition occurs will depend on the slope of the envelope and the form of the cusps. In one case, one expects to see intermittency immediately following the saddle-node bifurcation which destroyed one fixed point (see Chapter 4); there will be a short r -interval of intermittent-type chaotic behaviour before the fixed point in the next trough becomes stable (through period doubling). In the other case, there will already be attracting chaotic or periodic behaviour existing in the next trough before the saddle-node bifurcation occurs, and once the stable fixed point in one trough ceases



(a)

Local maximum in z 

(b)

Local maximum in z

Figure 8.11. Local maxima in the variable z plotted as a function of the previous local maximum in the variable z . (a) $\sigma \sim r \gg 1$, $b \sim 1$.
 (b) As $\sigma/b \rightarrow \infty$, the cusps become very narrow.

to exist, trajectories tend towards the other attracting set which already exists. Fowler & McGuinness report (1981b) that the second situation is more likely to occur when σ is large (in addition to σ/b being large).

For some parameter values, the Fowler & McGuinness approach has the advantage that it predicts the behaviour we are likely to see when we perform numerical experiments. It is not clear that the parameter values $\sigma = 10$, $b = 0.25$ fall into this category, though it is possible that at some r -values we will see the kinds of behaviour predicted in the last paragraph. In any case, the arguments from the paragraph before that (about the development as r increased with fixed σ) will be theoretically relevant. Note the particular disadvantages of the one-dimensional approach.

(i) It will not locate behaviour which is not of the fast/slow pulse variety. Hence, even if the method can be extended to cope with large r/σ , it will not detect the large r symmetric stable orbit.

(ii) It will not predict the heteroclinic bifurcation (which may occur in the region of validity of the analysis, $\sigma \sim r$) or deal well with the behaviour which leads up to the heteroclinic bifurcation. As we followed the extra xy orbit back and forth (Fig. 8.7), it continually acquired new local maxima in z as it wound around the stable manifolds of C_1 and C_2 . Each time this occurs, we expect the one-dimensional map to show a bifurcation (assuming that the analysis could locate the orbit we are discussing, which actually seems unlikely) though, topologically speaking, this event does not represent a bifurcation in the Lorenz system.

8.7. SUMMARY

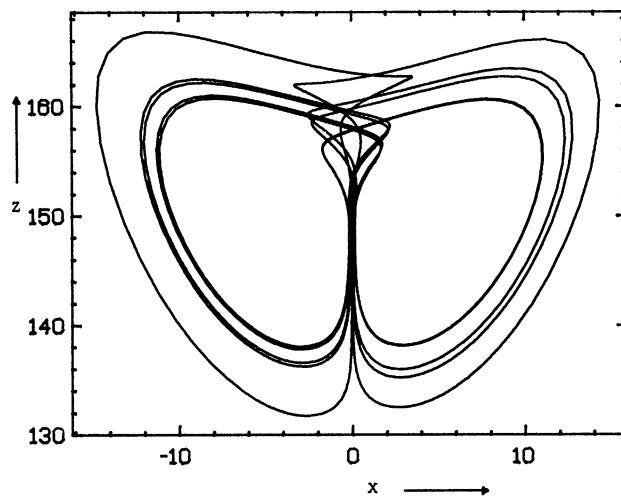
With an absolute minimum of numerical experiments, we have been able to describe many of the important features of the behaviour of the Lorenz equations when $b = 0.25$. Three of these features - twisting around the z -axis, homoclinic explosions for all r however large, and the heteroclinic bifurcation - can be expected to appear whenever the large r analysis predicts "anomalous" behaviour. In general, we expect such behaviour when $\frac{\sigma}{b}$ (or possibly σ/b^2) is large enough. In particular, we expect such behaviour in $0 < b < b^*$ for some $2 < b^* < 8/3$ when $\sigma = 10$. Furthermore, we expect none of this behaviour whenever the large r analysis does not predict anomalous behaviour. If one

or more of these features is missing (in a situation where we have the others) then something is wrong with our understanding of the Lorenz system.

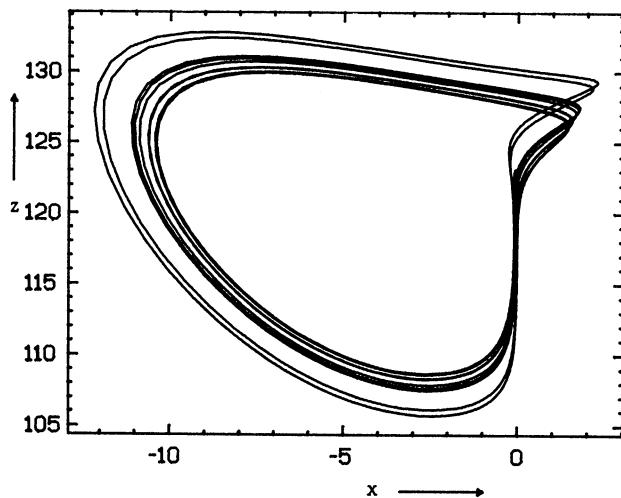
There are many details of the behaviour, some of which may be very important for various types of understanding, that we have not studied. Even if we exclude from consideration detailed theoretical questions which we cannot hope to answer (such as the detailed structure of the non-wandering set at almost any parameter value), and even if we exclude questions which present insurmountable numerical difficulties (such as observing the periodic orbits which exist at very large r -values), there are many aspects of the behaviour which could be investigated. We might, for example, be able to explain why the final stable symmetric xy orbit loses its stability in $80 < r < 3500$ (see Fig. 8.5). We might also be able to make some interesting statements about the behaviour of numerically calculated trajectories in this and other parameter ranges.

It may not be worth investigating these questions; they would certainly consume a lot of time and effort. In addition, answers to these questions might be expected to change with every change in σ and b (though if they did not, this would be interesting). The Lorenz system for small b is one of the most complicated three-dimensional chaotic systems of ordinary differential equations ever observed (if not the most) and attempts at complete understanding are almost certainly doomed.

Chapter 9 contains a brief summary of the types of behaviour that can be expected for a wide range of parameters σ , b and r . We end here by displaying two figures showing different kinds of chaotic behaviour for an even smaller value of b . See Fig. 8.12. A systematic search might be expected to yield similar "chaos with extra-twisting around the z -axis"-type figures with $b = 0.25$.



(a)



(b)

Figure 8.12. Chaos with extra twisting around the z -axis. $b = 0.1$.
(a) $r = 150$. At this parameter value there is also a stable symmetric
 xy orbit; (b) $r = 120$.

Chapter 9

Other Approaches, Other Systems, Summary and Afterword

9.1. SUMMARY OF PREDICTED BIFURCATIONS FOR VARYING PARAMETERS σ , b AND r

In this section we will not summarize all the results discussed in these notes (readers are referred to the relevant chapters for such summaries) but will attempt to draw together all the material that is useful in predicting bifurcations for general values of the parameters σ , b and r . We shall concentrate first on the two-dimensional parameter plane given by $\sigma = 10$ on which we have the most experience.

Fig. 9.1(a) shows the bifurcations we know analytically. The large r behaviour, calculated in Chapter 7, is marked along the top of the figure. Curve H shows the Hopf bifurcation value. The critical value, b^* , below which anomalous periodic orbits occur for all large r , lies somewhere between 2 and $8/3$. Fig. 9.1(b) uses Fig. 9.1(a) as a background and all the new curves on Fig. 9.1(b) are conjectural. The curve 0 represents the parameter values at which the first (as r increases) homoclinic explosion occurs. Since it seems likely that we cannot have the final stable symmetric orbit unless we have the first homoclinic explosion, we conjecture that this curve must exist for at least one r -value for every $b < 14.5$. ($b < 14.5$ is the parameter range in which the large r -symmetric orbit exists when $\sigma = 10$.) Numerical experiments may show that the curve 0 is more complicated than shown. It is possible, for example, that the curve enters the region $b > 14.5$ at some finite r -value. In any case, it seems likely that the curve 0 terminates at $b = 14.5$ and $r = \infty$ (which parameter values mark the boundary between no large- r orbits and some large- r orbits when $\sigma = 10$). If the curve 0 does enter $b > 14.5$ then it must turn back towards $b = 14.5$ at some point; in this case we would have some b -values (greater than 14.5) for

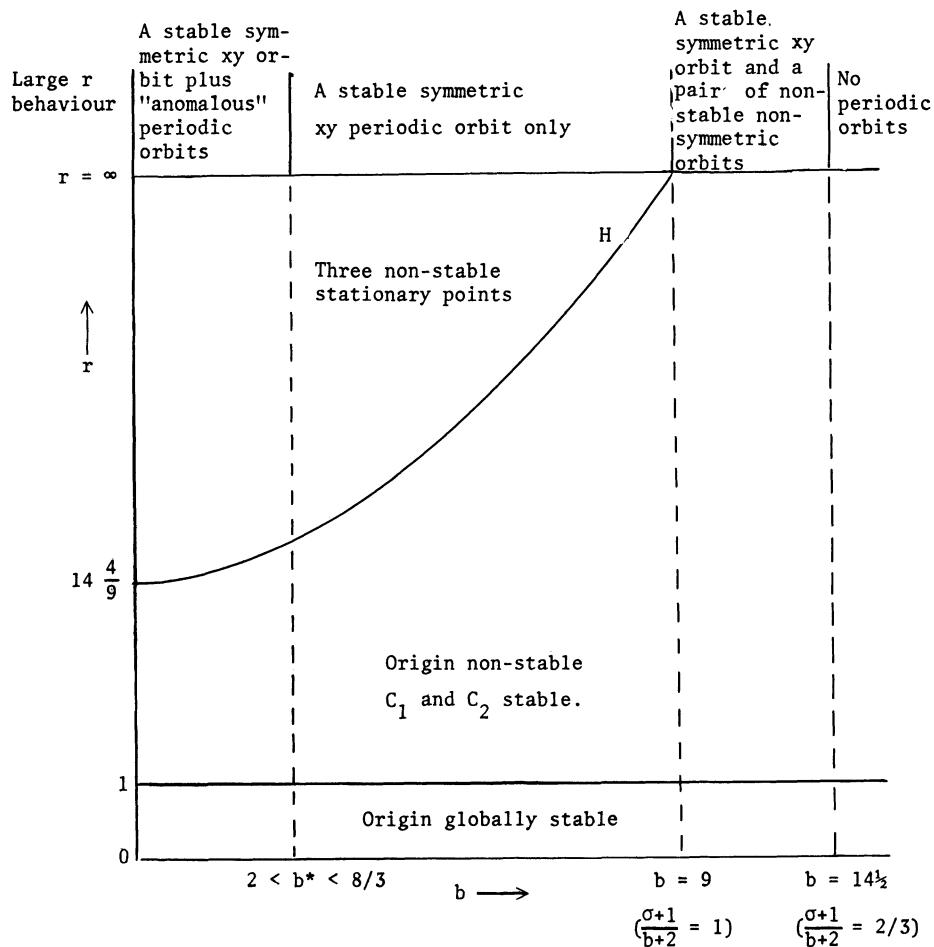
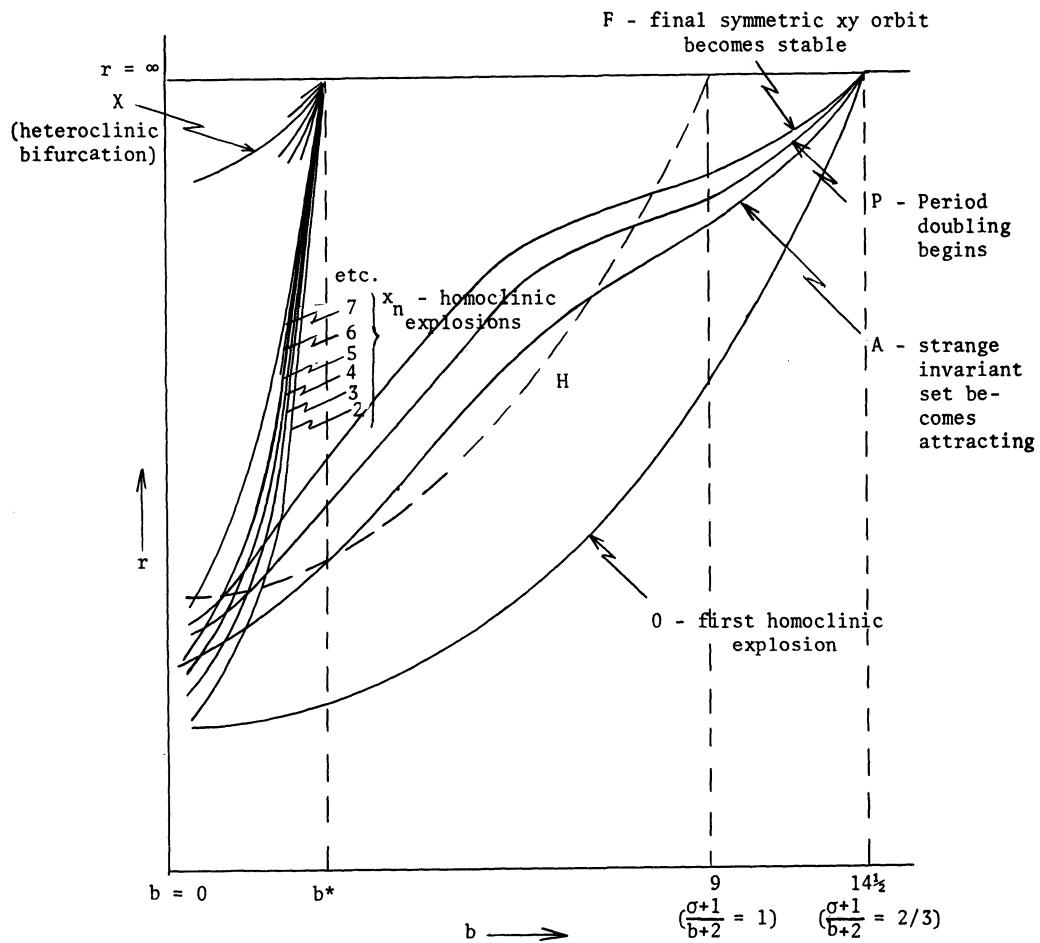


Figure 9.1. (a) Behaviour known analytically. $\sigma = 10$.

Figure 9.1. (b) Conjectured behaviour. $\sigma = 10$.

which the first homoclinic explosion occurred at one r -value but then reversed itself at a larger r -value. We can argue that we expect the curve O to lie always to the right of the Hopf curve H ; otherwise, as we increase r , we will see the Hopf bifurcation attempting to destroy some periodic orbits which do not yet exist. However, this argument is not rigorous. We can imagine that for some b -value there will be saddle-node bifurcations which produce, as r increases, two stable/non-stable pairs of periodic orbits which do not wind around the z -axis. The non-stable orbits from these pairs may be destroyed in the Hopf bifurcations at one r -value. At a larger r -value we may have the first homoclinic explosion, and at a yet larger r -value we may see the stable periodic orbits from the saddle-node bifurcations being annihilated in saddle-node bifurcations that also involve the two simplest non-stable periodic orbits born in the first homoclinic explosion.

It is not possible or useful to consider all the different alternatives for the bifurcation curves shown in Fig. 9.1. We have mentioned the two simple alternatives above to emphasize the conjectural nature of the curves displayed, and to show that alternative curves would not require us to make fundamental changes in our thinking about the Lorenz equations. In some sense, Fig. 9.1 shows the simplest possible curves; we should not be surprised if numerical experiments indicate that more complicated things can happen.

Curves A , P and F (Fig. 9.1(b)) show possible bifurcation curves for the events "original strange invariant set becomes attracting," "period doubling windows begin" (or, equivalently, "the first type (b) homoclinic explosion occurs"), and "final symmetric xy periodic orbit becomes stable". Between curves O and A we can expect preturbulence, between A and P we will have a strange attractor, and between P and F we will have period doubling behaviour (cf. Chapter 4). (We could, of course, draw many intermediate curves representing intermediate events. One example of such a curve might be " x^2y orbit becomes attracting".) However, we are not really sure which period doubling windows occur at which b -values. See the discussion at the end of Chapter 5.) Immediately to the left of curve F we will see the stable xy symmetric orbit. We already know from our study with $b = 0.25$ (Chapter 8) that this orbit is not stable in the whole of the region to the left of F ; there will be one or more regions to the left of F in which this orbit is non-stable and these regions will be bounded away from both F and $r = \infty$.

The curves 2,3,4,5, etc. (Fig. 9.1(b)) show possible curves for the extra twisted x_2, x_3, x_4, x_5 - homoclinic explosions (see Chapter 8) which occur in $b < b^*$. We know that there will be a countable infinity of these curves, that they will not cross, and that they will all terminate at $b = b^*, r = \infty$. For large enough r and $b > b^*$ we should be to the right of all these curves. This does not prevent some or all of these curves wandering temporarily into $b > b^*$ (cf. the argument about whether curve 0 wanders into $b > 14.5$ or not). The considerations of Chapter 8 indicate that we expect some curve X in $b < b^*$ which represents the heteroclinic bifurcation involving stationary points C_1 and C_2 .

Readers will notice that all the curves in Fig. 9.1 represent bifurcations connected either with homoclinic behaviour associated with the origin or with the large r behaviour. These are the two aspects of the behaviour which depend least on numerical experiments for their description, and of which we have the most "general" knowledge. They are also intimately related. Our understanding of the basic relationships between the curves shown on Fig. 9.1 and the large- r behaviour is the essence of our "global" understanding of the behaviour of the Lorenz equations or of Lorenz-like systems. It is worth restating (see Chapters 5 and 8) that this understanding may not allow us to predict accurately the numerically observed behaviour at many parameter values. There may be one or more attracting sets at some parameter values that are produced (as r increases, for example) in sequences of bifurcations unconnected with homoclinic or large r behaviour. All we can say about such sets is that we expect them to disappear (as we continue to increase r , for example) in sequences of bifurcations equally unrelated to homoclinic or large r behaviour; we can often expect that the sequences of bifurcations which produce and destroy these extra sets will be mirror images of one another (e.g., the sequence of bifurcations by which the final stable xy orbit loses and regains its stability when $b = 0.25$). In addition, Fig. 9.1 does not contain enough information for us to decide which, if any, parts of the homoclinically produced sections of the non-wandering set are stable at any particular b and r -value. This problem is particularly acute in $b < b^*$, though the considerations of Chapter 8 allow us to make some predictions.

If we now consider changes in the parameter σ , our bifurcation picture becomes three-dimensional. We can imagine that we have a continuum of figures such as Fig. 9.1 on top of one another, and events such as

"first homoclinic explosion" will now be represented by two-dimensional sheets which intersect each level plane ($\sigma = \text{constant}$) in a curve. When $\sigma > 10$ we have no reason to expect a picture qualitatively different from Fig. 9.1. We can conjecture that the range of behaviours will be very similar for any large enough σ . As σ decreases towards zero, we expect to lose the Hopf bifurcation when $\sigma = 1$ and all the interesting large r behaviour when $\sigma = 1/3$. This means that the curves H and O will move towards $b = 0$ as σ decreases, the former disappearing completely from the picture at $\sigma = 1$ and the latter disappearing at $\sigma = 1/3$. (We arrive at these σ -values by calculating conditions for it to be possible to have $\frac{\sigma+1}{b+2}$ greater than 1 and greater than $2/3$ respectively - see Chapter 7.) This tells us that we may expect behaviour similar to that seen with $\sigma = 10$ and $b = 8/3$ for small enough b -values whenever $\sigma > 1/3$, except that in $\sigma < 1$ we will have stable stationary points C_1 and C_2 as well. We do not know how the value b^* changes as σ changes except that we expect the relationship to be monotonic. It is possible that anomalous large r behaviour still occurs in $\sigma < 1$, though it is certain that there is some $\sigma^* \geq 1/3$ such that $\sigma < \sigma^*$ implies there is no anomalous large r behaviour for any $b > 0$.

Further investigation may add to Fig. 9.1. In particular, it would be interesting to know the behaviour in the limit $b \rightarrow 0$ for various parameters σ and r , since $b = 0$ is the line on which all our bifurcation curves terminate.

Fig. 9.1 also emphasizes the importance of the parameter value b^* . At this parameter value the large r behaviour changes, the homoclinic behaviour changes, and the behaviour of the stable manifolds of C_1 and C_2 changes.

9.2. OTHER APPROACHES

Even if we restrict our attention to the parameter range $\sigma = 10$, $b = 8/3$ and $0 < r < \infty$, we have by no means examined all the approaches used by other authors to study the Lorenz equations. This section is not intended to fill in all the gaps, but to point the way to some of the literature we have not mentioned so far.

Birman & Williams (1979) have examined the knots in the periodic orbits which exist in the geometric model of the Lorenz equations for parameter values near $b = 8/3$, $\sigma = 10$ and $r = 28$. The exact parameter values are not important, since the knot type of a periodic orbit will

not change with changing parameter. Their analysis applies to the periodic orbits born in the first homoclinic explosion in the Lorenz system and is valid whether or not the geometric model is appropriate, and could be extended to cover other orbits born in other homoclinic explosions.

Graham & Scholz (1980) sought to calculate analytically an invariant two-dimensional attracting manifold for each of the three stationary points in the Lorenz flow ($\sigma = 10$, $b = 8/3$, $r = 28$). They formulated a linear recursive scheme by which these manifolds could be determined in the form of power series expansions. For each of the stationary points C_1 and C_2 there was only one such manifold; this was the two-dimensional unstable manifold of the point. For the origin there was an infinite number of two-dimensional invariant manifolds, but only one which was locally a plane, locally attracting and which lay in $z \geq 0$. Graham & Scholz's paper does not prove that the power series which describe these manifolds converge. Nonetheless, by calculating coefficients up to tenth order they compute manifolds which together give a very good approximation of the Lorenz attractor. The three manifolds almost join up (plus or minus 1%) and the surface which results if you assume that they do join up lies incredibly close to the attractor obtained by direct numerical simulation of the equations.

Segur (1980) conjectures that there are precisely four sets of points in finite three-dimensional parameter space for which the Lorenz equations are completely integrable. These are

- i) $\sigma = 0$, for which the equations are linear,
- ii) $\sigma = 1/2$, $b = 1$, and $r = 0$,
- iii) $\sigma = 1$, $b = 2$, and $r = 1/9$, and
- iv) $\sigma = 1/3$, $b = 0$, and r arbitrary.

In addition, Segur reports that there are various regions (which include the ones listed above) for which the equations have a first integral. If we allow some of the parameters to become large, Segur reports that the equations can be solved if σ and b are finite and $r \rightarrow \infty$ (see Chapter 7), and that there is a first integral in either (r, b) finite with $\sigma \rightarrow \infty$, or in $(r/\sigma, b)$ finite and $r, \sigma \rightarrow \infty$.

Various authors have looked at the Lorenz equations using techniques developed in a more general context. Many of these techniques fall into the category of "attempts to say something interesting about chaotic systems with very incomplete information." One such technique is the analysis of spectra calculated from the behaviour of one variable over time. (See, for example, Crutchfield et al., 1980). This technique is interesting

mainly because it is a natural link between differential equations and real world problems; often it is not possible to measure three interesting variables in a real world experiment, even if it appears that the underlying dynamics are three-dimensional. (In addition, the study of spectra is one possible place to start incorporating the effect of noise into the study of chaotic systems.) However, the spectra for most chaotic systems of differential equations (including the Lorenz equations at chaotic parameter values) appear to consist of broad band noise plus some sharp peaks, and it is not yet possible to learn as much from a spectrum as can be learnt by more direct means. Another technique used is numerical calculation of the Liapunov characteristic exponent. This measures the degree of sensitive dependence on initial conditions by averaging (over non-transient sections of the flow) the rate at which nearby trajectories diverge. One obtains a positive value for the exponent if the system is chaotic, and a negative value if there is some stable periodic behaviour. The Liapunov characteristic exponent is very closely related to the concept of topological entropy, but, in the case of the Lorenz equations at least, we learn very little from the costly computations needed to determine it. (Numerical calculation of a positive characteristic exponent does not constitute "proof" that we have no stable periodic orbit.) See Shaw (1978), Shimada (1979) and Shimada & Nagashima (1977, 1979). See also Knobloch (1979) for a more statistical approach.

9.3. EXTENSIONS OF THE LORENZ SYSTEM

Various authors have studied systems of differential equations with dimension greater than three that are either extensions of the Lorenz equations or very nearly so. We mention a few here.

One way to build a Lorenz-like system is to use a different truncation of the infinite set of equations from which the Lorenz equations were originally derived. Rather than setting all but three modes to zero, one retains some greater number of modes. Curry (1978) shows how to define a "generalized Lorenz system" by considering different truncations of the infinite set of equations generated as a solution to the convection equations in the Boussinesq approximation. In the same paper, Curry studies a fourteen-dimensional truncation. As in the Lorenz equations, there is a stationary point at the origin and a parameter r . As r increases through one, two stable stationary points bifurcate away from the stationary point at the origin. Unlike the Lorenz system, these stationary

points lose their stability in a supercritical Hopf bifurcation, expelling two stable periodic orbits as r increases. At a larger r -value these periodic orbits surrender their stability to periodic orbits of twice the period (period doubling bifurcation), and at a yet larger r -value these doubled orbits appear to bifurcate to attracting invariant tori. (Invariant tori may occur in dissipative flows in greater than three dimensions. This is one difference which we might expect to see frequently when comparing three-dimensional and higher-dimensional dissipative flows.) Eventually the tori become non-stable, and, thereafter, Curry claims that they play the role of the stationary points C_1 and C_2 in the ordinary Lorenz system; i.e., trajectories wind away from one torus until they are spun over near to the other (symmetric image of the first) torus, etc. Trajectories do look quite Lorenz-like. In another paper, Curry (1979b) investigated the effect of a periodic forcing function on the toroidal regime (the parameter interval in which the tori are attracting), and in another (Curry, 1979c), he proves a boundedness theorem for trajectories of his generalized Lorenz systems.

Franceschini & Boldrighini (1979) and Franceschini & Tebaldi (1979) have studied a five-dimensional system also derived from the convection equations. The system has a four-fold symmetry and for some reasonably large r we find four stable stationary points which all simultaneously undergo supercritical Hopf bifurcations expelling four stable periodic orbits. For larger r -values the attracting set seems to be two separate Lorenz-like strange attractors, each near to two of the four (now non-stable) stationary points. Numerical simulations of these equations could easily be mistaken for simulations of the Lorenz equations. For even larger r -values there are (two) stable periodic orbits, both of which look very like the Lorenz equations' final stable symmetric xy orbit, but it seems that they appear in saddle-node bifurcations as r increases and that for large r both stable and non-stable periodic orbits exist. The papers mentioned above contain some details of the transitions between the different behaviours just mentioned, but the picture is still very sketchy and further investigation would be needed to clarify the relationship (or lack of it) between this system and the Lorenz equations. Franceschini & Tebaldi (1980) have now started work on a seven-dimensional system, and Franceschini (private communication) reports some results for a nine-dimensional system. In the seven-dimensional system there is an interesting sequence of "period doubling" bifurcations involving tori.

Da Costa et al. (1981), Knobloch et al. (1981) and Knobloch & Weiss (1981, 1982) have studied two slightly different five-dimensional systems of equations which reduce to the Lorenz equations when one of the parameters is set to zero. The equations possess a Lorenz-like symmetry. One set of equations is derived from a two-dimensional thermosolutal convection problem, the other from a problem of convection for a conducting fluid in the presence of a magnetic field. Between them these papers contain a detailed analysis of the stability of steady convection states, and a careful examination of some of the possible bifurcations (for varying parameters) as the "r" parameter increases. Of particular interest in these papers is the occurrence of chaotic behaviour for physically relevant parameter values (r near one). Much of this behaviour appears to be connected with heteroclinic orbits between two stationary points. It should also be noted that these papers give some thought to the relationship between the original problems and the equations used to model them; the higher-dimensional systems described earlier in this section are derived in a much more arbitrary fashion.

Gibbon (1981) and Gibbon & McGuinness (1981) have shown how to derive a fifth order system which may be viewed as a complex generalization of the Lorenz equations (the variables x and y are allowed to be complex, as are some parameters). These equations can describe behaviour in baroclinic instability and in nonlinear optics. Fowler, Gibbon & McGuinness (1981) show that in this system the origin becomes unstable by Hopf bifurcation (supercritical). The resulting stable periodic orbit (which has an exact analytic form) is the complex generalization of the stationary points C_1 and C_2 of the Lorenz equations, the complex equations having a rotational invariance in place of the symmetry of the Lorenz equations. This stable limit cycle appears to lose its stability by absorbing a non-stable torus as "r" increases (analogous to the sub-critical Hopf bifurcation in the Lorenz system) though thereafter it appears that the complex system behaves rather differently from the three-dimensional real system. Fowler & McGuinness (1982) have extended the large r results to cover the complex case, and make some interesting observations on the similarities and differences between the two systems.

9.4. AFTERWORD - A PERSONAL VIEW

We have examined the Lorenz equations in some detail - a detail which was warranted partly by the immense amount of interest that the equations have generated in the last eighteen years. It is time to ask whether we have learnt anything of general importance.

We should refer first to the physical relevance of the Lorenz equations. Various authors continue to maintain that the behaviour of some real world systems may be adequately modelled by the Lorenz equations or by some extension or generalization of the Lorenz equations. Unless or until this claim is shown to be false, an understanding of the Lorenz equations will be of physical importance. It should be noted, in particular, that turbulent behaviour can be observed in the Lorenz system at parameter values where the stationary points are still stable. Most other finite dimensional models of turbulence require the stationary points to lose stability before turbulent behaviour can occur. Thus, Lorenz type models may be particularly interesting when considering physical systems where the experimental transition to turbulence occurs before the theoretical loss of stability of the stationary solutions (e.g., plane Poiseuille flow and pipe Poiseuille flow). The understanding we have gained in these notes will be of greater importance if some of it can be carried over to Lorenz-like systems such as those described in the previous section. There are indications that this may be possible, which is most encouraging. At the moment, study of higher-than-three-dimensional systems is difficult and proceeds slowly; it is often not clear whether apparently related systems with different dimensionality actually behave quite differently, or whether it is just that the higher-dimensional systems present such a large number of different parameter values and regions of phase space for study that those parts of the behaviour which are similar to or derived from their lower-dimensional cousins have not yet been observed. This question of changing behaviour with changing dimensionality may be very important. If finite-dimensional systems are to be used to model infinite-dimensional systems, it seems important that similar behaviours should be observed in low-dimensional systems of differing dimension; there is no reason to think that any particular finite number of dimensions should be the "best" for modelling general real problems (though it is possible that there will be occasions when some particular finite number of dimensions is best for modelling some particular problem). The understanding we have gained of the Lorenz equations is not dimension dependent, and it will, if it is applicable, show us how and

where to look for Lorenz-like behaviour in higher dimensional systems. We may be able to locate Lorenz-like behaviours even where they are actually barely observable (as in Chapter 8), or, in some cases, where they cannot be observed directly at all. At the same time we may find some coherence in the behaviour which is easily observed.

The Chapter 8 study of $b = 0.25$, $\sigma = 10$, is an excellent illustration of how the earlier work can help in understanding a "new" and more complicated system. Initial, random numerical experiments for these parameter values were most confusing; in addition, they demonstrated that, with the extra twisting around the z-axis, our understanding from Chapter 5 was no longer adequate. However, an almost entirely non-rigorous extension to our arguments led to an understanding of a basic framework of bifurcations on which the rest of the observed behaviour could be hung. Adapting our understanding to higher dimensional systems may be more difficult, but if it yields any kind of global understanding at all it will be very worthwhile.

We should also consider how our study relates to general systems of chaotic differential equations. Clearly it is useful to have contributed information on some of the various types of chaotic behaviour that can occur. However, the Lorenz equations are rather special because of their symmetry; even if we extend our study to include slight perturbations of this symmetry - which we could do - we would still be studying a system which was "nearly" symmetrical and this alone would be sufficient to make the system special. Many systems of three-dimensional chaotic differential equations have just one stationary point which loses stability in a Hopf bifurcation by casting off a stable periodic orbit: this orbit undergoes successive period doublings and a chaotic regime is reached in which period doubling windows and chaotic behaviour alternate. It is not clear exactly how this behaviour is related to the behaviour observed in the Lorenz equations as r decreases (Chapter 4) though it is clear that there are both similarities and differences. For most systems there are no arguments to suggest the existence of a strange attractor in a whole range of parameter values; this state of affairs should be contrasted with the Lorenz situation where we believe that there is a strange attractor (albeit a constantly changing one) for every parameter value in some interval. These and other differences imply that much of our specific Lorenz knowledge will not be very helpful in understanding these more general systems. Nonetheless, it is becoming increasingly clear that sequences of homoclinic bifurcations may play an important role in many systems of

chaotic equations, though, regrettably, the homoclinic orbits are frequently of a saddle-focus type (with one real eigenvalue and a complex pair of eigenvalues at the stationary point) and these bifurcations are not easily analyzed. It is to be hoped, however, that the general approach used here will be useful for those who wish to know in some detail how a particular set of ordinary differential equations behaves with changing parameter.

Finally, we have learnt how fallible simple numerical experiments can be - the description (but not observation) of new behaviour expected for the Lorenz equations in much studied regions of period doubling behaviour (see Chapter 5, Section 6) suggests caution in the interpretation of any numerical results however carefully obtained. And I know that people who live in glass houses should not throw stones! Tinkle. Crash.

"An analysis terminates when the patient realizes
that it could go on forever." Hanns Sachs

Appendix A

Definitions

The Lorenz equations specify, for each value of the parameters, a vector field $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ which is independent of time. This vector field determines a flow, $\phi_t(\underline{x})$, where $\phi_t(\underline{x}), \underline{x} \in \mathbb{R}^3$ and $t \in \mathbb{R}$. The Lorenz equations are usually written in the form $\frac{d\phi}{dt} = f(\phi_t(\underline{x}))$.

The flow, ϕ , satisfies the usual conditions; $\phi_0(\underline{x}) = \underline{x}$, $\phi_s(\phi_t(\underline{x})) = \phi_{s+t}(\underline{x})$ and ϕ is continuous in both \underline{x} and t .

A stationary point, \underline{x}^* , of the flow is a point for which $\phi_t(\underline{x}^*) = \underline{x}^*$ for all $t \in \mathbb{R}$. Equivalently, $f(\underline{x}^*) = \underline{0}$.

The trajectory through a point \underline{x} is the set of points $\{\phi_t(\underline{x}): -\infty < t < \infty\}$. Sometimes, when we say, "the trajectory starting at \underline{x} ", we mean only that part of the trajectory given by $\{\phi_t(\underline{x}): 0 \leq t < \infty\}$.

A periodic orbit is a set of points $\{\phi_t(\underline{x}): 0 \leq t < p\}$ for some positive $p \neq 0$ such that $\phi_p(\underline{x}) = \underline{x}$ and $\phi_t(\underline{x}) \neq (\underline{x})$ for any $0 < t < p$. The period of the orbit is p . Occasionally we use period to describe the number of intersections that some periodic orbit makes with some suitable plane. The context should always make it clear which meaning we are using for "period".

An aperiodic trajectory is a trajectory which is not a periodic orbit.

We use the usual metric on \mathbb{R}^3 and define the distance $d(\underline{x}, \underline{y})$ between points \underline{x} and \underline{y} in the usual way. We also define the distance between a point \underline{x} and a set S in the usual way ($d(\underline{x}, S) = \inf_{\underline{y} \in S} d(\underline{x}, \underline{y})$). The trajectory through \underline{x} tends towards a set S if $\lim_{t \rightarrow \infty} d(\phi_t(\underline{x}), S) = 0$.

If $\lim_{t \rightarrow -\infty} d(\phi_t(\underline{x}), S) = 0$ we say the trajectory through \underline{x} tends towards S in reverse time.

A *homoclinic orbit* (of a stationary point \underline{x}^*) is a trajectory which tends to \underline{x}^* in both forwards and reverse time. A *heteroclinic orbit* (between stationary points \underline{x}^* and \underline{y}^*) is a trajectory which tends towards \underline{x}^* in reverse time and towards \underline{y}^* in forwards time.

An *invariant set*, S , is a set such that $\phi_t(S) = S$ for all t .

$$\text{def} \quad (\phi_t(S) = \bigcup_{x \in S} \phi_t(x))$$

In studying topological dynamics, or dynamical systems, it is usual to distinguish between a set which is stable and a set which is an attractor. The usual definitions are:

A closed invariant set S is (Liapunov) *stable* if for each $\epsilon > 0$ there is a $\delta \in (0, \epsilon]$ such that $d(\underline{x}, S) \leq \delta$ implies that for every $t > 0$, $d(\phi_t(\underline{x}), S) \leq \epsilon$; i.e., if you start close to S you remain close to S .

A closed invariant set S is an *attractor* if for all \underline{x} sufficiently close to S , the trajectory through \underline{x} tends to S ; i.e., if you start close to S you ultimately tend towards S . There are examples (e.g., Denjoy (1932)) of sets which are attractors but which are not stable. There is no reason to expect a similar pathology in the Lorenz system and it is likely that all the attractors we meet will be stable. Consequently, we fail to preserve this distinction in the body of the text. Normally we only call a set S stable (or an attractor) if no proper subset of S is stable (or an attractor).

A set S is *globally attracting* if all trajectories tend towards S .

A set S is *unstable* if it is stable when time is reversed (i.e., replace " $t > 0$ " with " $t < 0$ " in the definition of stable given above). A compact invariant set which is neither stable nor unstable is *non-stable*. In the Lorenz system there are no unstable sets (see Section 1.4 (iii)).

The *stable manifold* of a compact invariant set S is the set of points $\underline{x} \in \mathbb{R}^3$ such that the trajectories through \underline{x} tend towards S . The *unstable manifold* of S is the set of points \underline{x} in \mathbb{R}^3 such that the trajectories through \underline{x} tend towards S in reverse time.

A compact invariant set of special interest to us is the *non-wandering set*, Ω . A point \underline{x} *wanders* if we can find an open neighborhood U of \underline{x} and a $T > 0$ such that $\phi_t(U)$ does not contain \underline{x} for any $t > T$. The set of points which do not wander is the non-wandering set. Ω is closed, invariant, and contains all the recurrent behaviour of the flow.

Appendix B

Derivations of the Lorenz Equations From the Motion of a Laboratory Water Wheel

The derivation in this appendix follows Lorenz (1979) who reports that Professor W. Malkus of M.I.T. has constructed a laboratory water wheel whose equations of motion are the Lorenz equations. The advantage of this derivation over the original one (Lorenz, 1963) is that the ordinary differential equations are obtained directly, rather than as an approximation to a partial differential equation.

We take a wheel which is free to rotate around a horizontal axis. The circumference of the wheel is made up of leaky compartments. Water enters the compartments near the top of the wheel; this unbalances the wheel, which may start to move. Depending on the various physical constants associated with the experiment, the wheel may remain at rest, rotate steadily in one direction or the other, or reverse its direction of rotation at regular or irregular intervals.

We assume that the wheel has a radius a and that its whole mass consists of water lying exactly on the circumference. We assume that there is a continuous function $m(\theta, t)$ which describes the mass of water per unit arc around the circumference of the wheel. Here, t is the time and θ is the angular displacement measured anticlockwise. (θ refers to an angular displacement in space rather than to a fixed point on the wheel which may be rotating.) Water is added to the wheel at a constant rate. We assume that the various points on the circumference of the wheel gain water at a rate proportional to their instantaneous height above some reference level; this is either a rather crude approximation of a situation in which water is only added at the very top of the wheel, or a rather better approximation of a situation in which the wheel stands in a (vertical) rain. We assume that water leaks away at a rate propor-

tional to the function \bar{m} . The angular velocity of the wheel at time t is written $\Omega(t)$ and we assume that there is a frictional damping proportional to Ω .

If we use a bar, $\overline{\quad}$, to represent integration with respect to θ over the interval $[0, 2\pi]$ we have the two equations of motion,

$$\frac{d(a^2 \bar{m}\Omega)}{dt} = -g a \bar{m} \cos \theta - k a^2 \bar{m} \Omega \quad (1)$$

and

$$\frac{\partial \bar{m}}{\partial t} + \Omega \frac{\partial \bar{m}}{\partial \theta} = (A + 2B \sin \theta) - h \bar{m} \quad (2)$$

where A , B , h and k are positive constants. Equation (1) is simply the equation for angular momentum. Equation (2) is the mass equation; the right-hand side contains terms representing the rates of accretion and loss of mass at a particular point in space. The left-hand side contains a term that compensates for the fact that the wheel may be rotating.

The first thing to notice is that the total rate of accretion of mass is constant and that the total rate of mass loss is proportional to the total mass, \bar{m} . This implies that \bar{m} tends asymptotically towards an equilibrium value $2\pi A/h$. If we assume that \bar{m} has reached this equilibrium value then equation (1) becomes

$$\frac{d\Omega}{dt} = -k\Omega - \left(\frac{gh}{2\pi a A} \right) \bar{m} \cos \theta \quad (3)$$

and we can calculate expressions for the rates of change of the quantities $\bar{m} \sin \theta$ and $\bar{m} \cos \theta$ from equation (2). For example,

$$\begin{aligned} \frac{d(\bar{m} \sin \theta)}{dt} &= \int_0^{2\pi} \frac{d}{dt} (\bar{m} \sin \theta) d\theta \\ &= \int_0^{2\pi} \left(\frac{\partial \bar{m}}{\partial t} + \Omega \frac{\partial \bar{m}}{\partial \theta} \right) \sin \theta + \Omega \bar{m} \cos \theta d\theta \\ &= \int_0^{2\pi} (A \sin \theta + 2B \sin^2 \theta - h \bar{m} \sin \theta + \Omega \bar{m} \cos \theta) d\theta \\ &= \Omega \bar{m} \cos \theta - h \bar{m} \sin \theta + 2\pi B. \end{aligned} \quad (4)$$

Similarly,

$$\frac{d(\bar{m} \cos \theta)}{dt} = -\bar{m} \sin \theta - h \bar{m} \cos \theta. \quad (5)$$

A suitable linear change of coordinates converts equations (3), (4) and (5) into the Lorenz equations with the parameter $b = 1$.

Appendix C

Boundedness of the Lorenz Equations

1. EXISTENCE OF A BOUNDED ELLIPSOID WHICH ALL TRAJECTORIES ENTER

We wish to show that there is a bounded region E such that every trajectory eventually enters E and never thereafter leaves it.

Consider the Liapunov function

$$V = rx^2 + \sigma y^2 + \sigma(z - 2r)^2$$

which satisfies

$$\frac{dV}{dt} = -2\sigma(rx^2 + y^2 + bz^2 - 2brz).$$

This function is chosen to simplify the formulae, not to give the optimum result. Lorenz (1963, 1979) has used other Liapunov functions. Let D be the bounded region inside which $dV/dt \geq 0$ and let c be the maximum of V in D . Consider the bounded ellipsoid E in which $V \leq c + \epsilon$ for some small but positive ϵ . If a point \underline{x} lies outside E then it also lies outside D and so $\dot{V}(\underline{x}) \leq 0$. In fact, $\dot{V}(\underline{x}) \leq -\delta$ for some small but positive δ depending on ϵ . If we start a trajectory at this point \underline{x} outside E then the value of V associated with points on the trajectory must decrease as time goes by and within a finite time the trajectory will enter the ellipsoid E . Furthermore, all trajectories pass inwards through the boundary of E . A trajectory, once within E , will remain there forever.

(We can set both ϵ and δ to zero in the above argument. In this case all trajectories eventually enter E but they may take an infinite time to do so. For some choices of parameters and Liapunov function the stationary point at the origin may lie on the boundary of the region

$V \leq c$ and trajectories approaching the stationary point from outside the region $V \leq c$ will take an infinite time to reach it. This is the case for our choice of function V with some choices of parameters σ and b . See below.)

2. SIZE OF THE ELLIPSOID

The analysis of Chapter 7 requires that we have some idea of the size of E . It is obvious from the geometry that the maximum of V in D will occur somewhere on the boundary of D and hence at a point at which $2rx = 2rx\lambda$; $2\sigma y = 2y\lambda$; and $2\sigma(z-2r) = 2b(z-r)\lambda$ for some Lagrangian multiplier λ .

The possible solutions are

- (i) $\lambda = 1$, $y = 0$, $z = r(b-2\sigma)/(b-\sigma)$, and $x^2 = b^2 r(b-2\sigma)/(b-\sigma)^2$ provided $b \geq 2\sigma$, and here $V = b^2 r^2/(b-\sigma)$;
- (ii) $x = 0$, $\lambda = \sigma$, $z = r(b-2)/(b-1)$, $y^2 = b^2 r^2(b-2)/(b-1)^2$ provided $b \geq 2$, and here $V = b^2 r^2(b-2)/(b-1)$;
- (iii) $x = y = 0$, $z = 2r$ or $z = 0$, which give $V = 0$ and $V = 4\sigma r^2$ respectively.

Comparing the values of V at these various points it is easy to see that the maximum c of V in D is

$$c = \begin{cases} b^2 r^2 / (b-\sigma) & \text{if } b \geq 2\sigma, \sigma \leq 1 \\ b^2 r^2 / (b-1) & \text{if } b \geq 2, \text{ and } \sigma \geq 1 \\ 4\sigma r^2 & \text{otherwise.} \end{cases}$$

Therefore, for fixed σ and b , c is $O(r^2)$ and inside E the variable x is $O(r^{1/2})$, y is $O(r)$ and z is $O(r)$.

Notice that if we choose b and σ such that c takes the third value above, then we have proved that all trajectories eventually lie in $z \geq 0$ (because for these parameter values the ellipsoid E lies in $z \geq 0$). Different Liapunov functions allow us to prove that all trajectories eventually lie in $z \geq 0$ for other values of the parameters σ and b . For instance, the Liapunov function

$$V_1 = r^2 x^2 + \sigma y^2 + \sigma(z - r(r-1))^2$$

is adequate to prove the result in $b \leq \sigma+1$, which covers all the cases studied in these notes (in particular $b = 8/3$, $\sigma = 10$) as well as all

the cases already covered above. The disadvantages of this second Liapunov function, V_1 , are that the formulae are much uglier and the estimate for the size of the region which all trajectories eventually enter is much too large. It seems likely that Liapunov functions can be found to prove that all trajectories eventually enter and remain in the region $z \geq 0$ for all parameter values σ and b , but we shall proceed no further here.

3. EXISTENCE OF A BOUNDED SET OF ZERO VOLUME TOWARDS WHICH ALL TRAJECTORIES TEND

The divergence of the Lorenz flow is $-(\sigma+b+1)$. At times $1, 2, 3, \dots$ the surface of the closed ellipsoid E is taken by the flow into surfaces S_1, S_2, S_3, \dots which enclose regions E_1, E_2, E_3, \dots . If the volume of E is $\text{vol}(E)$ then each E_i has volume $e^{-(\sigma+b+1)i} \text{vol}(E)$ and the volumes of the E_i decrease exponentially to zero as i increases. Because all trajectories cross the boundary of E inwards we know that $E \supset E_1 \supset E_2 \supset E_3 \dots$ and hence that every trajectory is ultimately trapped in a region, E_∞ , of zero volume given by

$$E_\infty = \bigcap_{i \in \mathbb{Z}^+} E_i.$$

Because each E_i is closed and connected, E_∞ is closed and connected. Notice that the set E_∞ we have described above need not be the same as the non-wandering set, Ω , defined in Appendix A. It is easy to see that the non-wandering set, Ω , is a subset of the set E_∞ . That the two sets are not identical can be seen by considering the flow for r just larger than 1. The set E_∞ consists of the three stationary points plus the unstable manifold of the non-stable point (the origin). The two branches of this unstable manifold will join the non-stable point to the two stable points and the set E_∞ is connected. The non-wandering set, Ω , at the same r -value, contains only the three stationary points and is not connected. For most purposes the non-wandering set, Ω , is the most convenient set to consider. The set E_∞ contains some non-recurrent behaviour (in the example above, the two branches of the unstable manifold of the origin) in which we are not very interested.

Appendix D

Homoclinic Explosions

The results presented in this appendix apply to any system of three-dimensional differential equations with a Lorenz-type symmetry where the eigenvalues of the linearized flow near a stationary point on the axis of symmetry are real and satisfy $-\lambda_2 > \lambda_1 > -\lambda_3 > 0$ (where λ_3 is the eigenvalue associated with the eigenvector which lies on the axis of symmetry). The results will also apply to systems of higher dimension with appropriate conditions on the additional eigenvalues. I am indebted to Prof. Swinnerton-Dyer and to a paper by Kaplan & Yorke (1979a) for elements of this analysis. Our approach is similar to the approach used by Silnikov (1968, 1970) in papers on systems where the eigenvalues have different relative magnitudes, and Silnikov's papers contain the kind of estimates necessary to establish the rigour of the proofs used below. More recent work (e.g., Arnéodo, Coullet & Tresser, 1981a) is also of interest, though it does not have direct application to the Lorenz equations.

Suppose that for some $r = r'$ there is a homoclinic orbit to the origin. We shall investigate the behaviour for r near r' . First, it is convenient to change coordinates to (ξ, η, z) , the eigenvectors of the linearized flow at the origin. ξ corresponds to the positive eigenvalue λ_1 , η and z correspond to the negative eigenvalues λ_2 and λ_3 . We have $-\lambda_2 > \lambda_1 > -\lambda_3$ (see Chapter 1). The symmetry $(x, y, z) \rightarrow (-x, -y, z)$ carries over to a symmetry $(\xi, \eta, z) \rightarrow (-\xi, -\eta, z)$ in the new coordinates.

We consider a bounded domain \mathcal{D} consisting of a small box B around the origin with faces $|\xi| = c_1$, $|\eta| = c_2$, and $|z| = c_3$, together with two tubes T and S around the trajectories which leave the box B through the points $(\pm c_1, 0, 0)$. The tubes T and S only surround that part of these trajectories which lies between the points $(\pm c_1, 0, 0)$ and the points

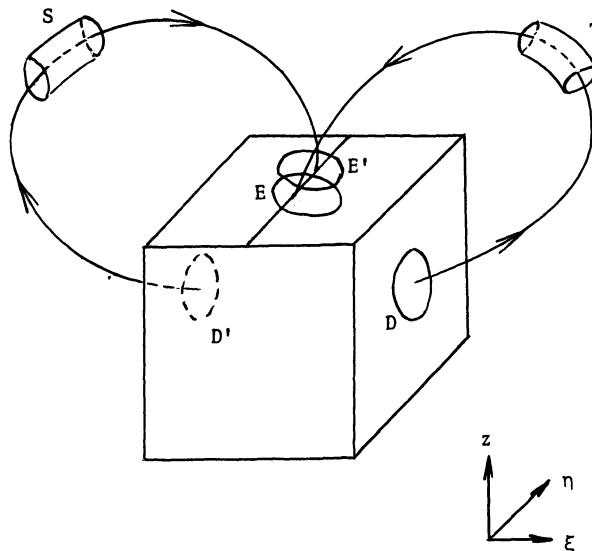


Figure D1. The box B . For the sake of clarity, only sections of the tubes T and S are shown.

where the trajectories re-enter the box B for the first time. Since these trajectories do re-enter the box B when $r = r'$, we can be sure that they will do so for r near r' . We know they re-enter the box through its top face ($z = c_3$) because the eigenvalue λ_2 is greater in absolute magnitude than the eigenvalue λ_3 .

Call the ends of the tube T , D and E , D being a disc on the face (c_1, n, z) of B , and E being a disc on the top face (ξ, n, c_3) of B . For $r = r'$ the disc E meets the line $(0, n, c_3)$ - which is the intersection of the stable manifold of the origin with the top face of B - and therefore it does so for r near r' . Call the ends of the tube S , D' and E' , where D' and E' are the images under the symmetry of D and E . See Figure D1.

Following trajectories through T induces a map

$$D \rightarrow E; (c_1, \eta, z) \rightarrow (f_r(\eta, z), g_r(\eta, z), c_3). \quad (1)$$

There is a similar map (given by the symmetry) induced by the tube S . Since our box B is small we can calculate a map $E \rightarrow D$ from the linear analysis. For $\xi > 0$,

$$(\xi, \eta, c_3) \rightarrow \left(c_1, \eta(\xi/c_1)^{-\lambda_2/\lambda_1}, c_3(\xi/c_1)^{-\lambda_3/\lambda_1} \right). \quad (2)$$

(This, and subsequent maps, should contain small error terms. They do not effect the subsequent analysis and so we ignore them.) Since ξ is small compared with c_1 (for thin enough tubes and r close to r'), the second coordinate in (2) above is much smaller than the third.

($-\lambda_2/\lambda_1 > 1$ and $-\lambda_3/\lambda_1 < 1$.) Combining (1) and (2) we get a map from $E \rightarrow E$ for $\xi > 0$ given by

$$(\xi, \eta, c_3) \rightarrow \left(f_r(\eta(\xi/c_1)^{-\lambda_2/\lambda_1}), c_3(\xi/c_1)^{-\lambda_3/\lambda_1}, g_r(\text{same arguments}), c_3 \right) \quad (3)$$

with an analogous formula for $\xi < 0$.

Now write ξ_r^* for $f_r(0, 0)$ and η_r^* for $g_r(0, 0)$. In other words, (ξ_r^*, η_r^*) is the point on the top of B where the right-hand branch of the unstable manifold of the origin first re-enters the box. Let $f_r^* = \frac{\partial f}{\partial z}(0, 0)$ and $g_r^* = \frac{\partial g}{\partial z}(0, 0)$ and expand (3) to first order. We ignore the terms in $\frac{\partial f}{\partial \eta}$ and $\frac{\partial g}{\partial \eta}$ because they are multiplied by the extremely small second coordinate of equation (2). We obtain a map $E \rightarrow E$

$$(\xi, \eta) \rightarrow (\xi_r^* + f_r^* c_3(\xi/c_1)^{-\lambda_3/\lambda_1}, \eta_r^* + g_r^* c_3(\xi/c_1)^{-\lambda_3/\lambda_1}). \quad (4)$$

Notice that we have now "decoupled" the first coordinate and can consider the one-dimensional first return map of this coordinate given by

$$\phi(\xi) = \xi_r^* + f_r^* c_3(\xi/c_1)^{-\lambda_3/\lambda_1}. \quad (5)$$

Providing that we assume that f_r^* does not change sign in a small r -interval near r' (which we do), we have only two cases to consider. These are (a) $f_r^* > 0$, and (b) $f_r^* < 0$. For each case we consider an $r_1 < r'$ and an $r_2 > r'$ and assume, without loss of generality, that $\xi_{r_1}^* > 0$ and $\xi_{r_2}^* < 0$. Fig. D2 shows the map ϕ (equation (5)) for r_1, r' and

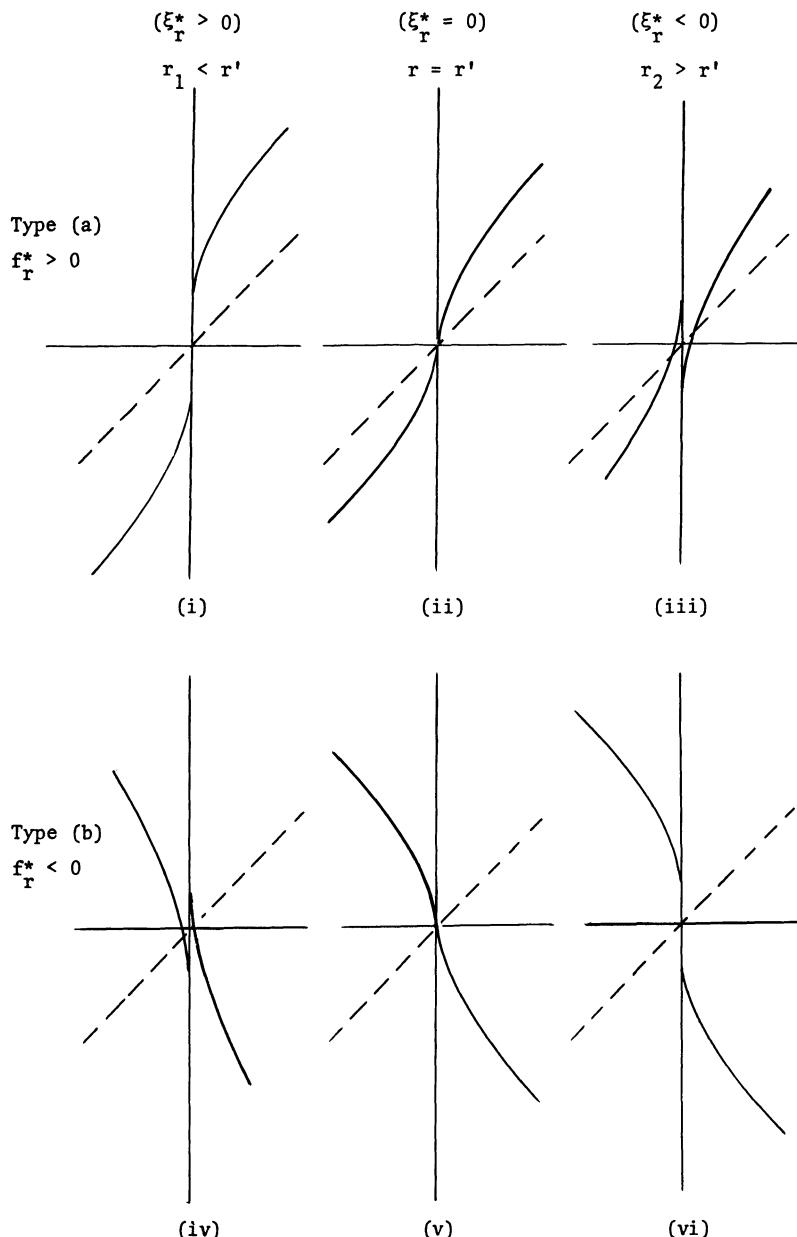


Figure D2. The one-dimensional return map $\phi(\xi) = \xi_r^* + f_r^* c_3(\xi/c_1)^{-\lambda_3/\lambda_1}$
 if $\xi > 0$, $\phi(\xi) = -\phi(-\xi)$ if $\xi < 0$.

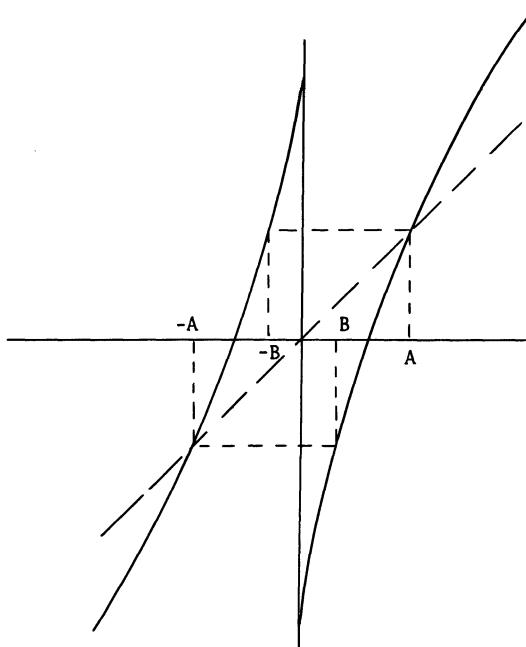


Figure D3. An expanded version of Fig. D2(iii).

r_2 in each of the two cases (a) and (b). Figs. D2 correspond in the obvious way to Figs. 2.6 of Chapter 2, as do the cases (or types) (a) and (b).

We wish to know how many trajectories, if any, will remain in our bounded domain \mathcal{D} forever. Trajectories will certainly not remain forever in this region if the map ϕ increases the absolute magnitude of the ξ coordinate on every return to the top face of the box B . Thus, Figs. D2(i), (ii), (iv) and (v) are easily interpreted; all trajectories other than the homoclinic orbits at $r = r'$ will have their ξ coordinate increased on each return to the top of the small box B and hence will

eventually leave the region of interest. Figs. D2(iii) and (iv) are harder to interpret. We shall look at (iii); similar arguments apply to (iv). Fig. D3 shows a more detailed version of Fig. D2(iii).

If $|\xi| > A$ (see Fig. D3) we have $|\phi(\xi)| > |\xi|$ and each return to the top face of B increases $|\xi|$. If $|\xi| < B$ then we have $|\phi(\xi)| > A$, with subsequent returns to the top face of B increasing $|\xi|$. Thus, trajectories started in $0 < |\xi| < B$ and in $|\xi| > A$, will eventually leave the region of interest, \mathcal{D} . However, ϕ maps each of the intervals $[-A, -B]$ and $[B, A]$ onto $[-A, A]$. This means that there are points in $[-A, -B]$ and $[B, A]$ that are mapped into either $[-A, -B]$ or $[B, A]$ by ϕ . We can continually extend this argument to cover more and more iterations of the map ϕ , eventually determining that there are some

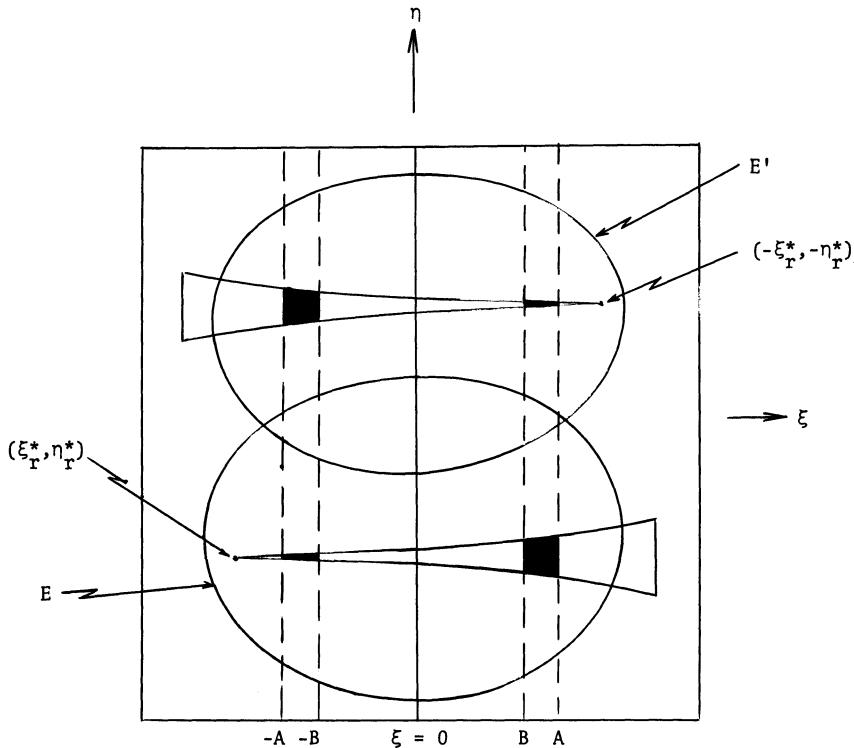


Figure D4. Return map on the top face of B. Each of the four shaded areas is mapped into a long thin area stretching all the way from $\xi = -A$ to $\xi = A$.

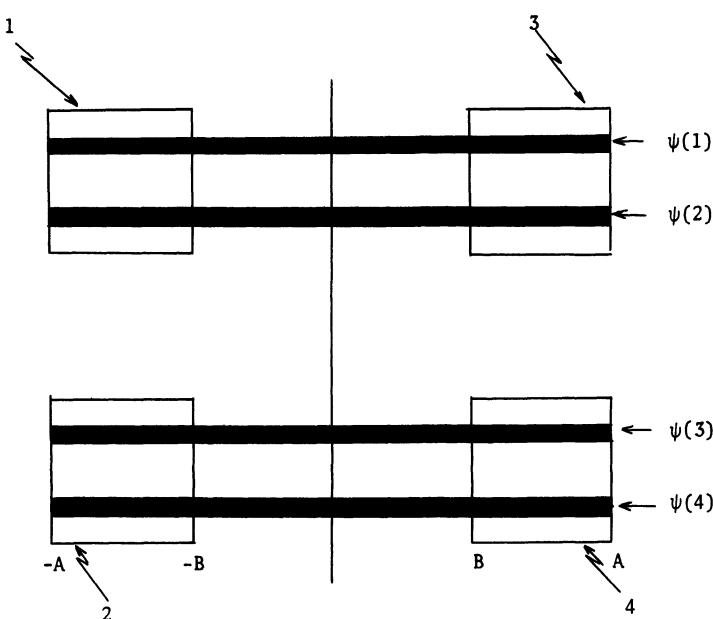
trajectories which remain in the union of $[-A, -B]$ and $[B, A]$ forever. First, we shall return to consideration of the full two-dimensional return map. Equation (4) indicates that there is a small interval of η -values corresponding to each of the two intervals $B < |\xi| < A$. Fig. D4 shows a schematic view of the return map on the top face of the box B (cf. Fig. 2.6(iii)).

Each of the four shaded areas in Fig. D4 will map, under one application of the return map, onto a very thin slice that stretches all the way from $\xi = -A$ to $\xi = A$. This slice is very thin because of the strong contraction in the η -direction. Fig. D5 shows the four areas and their first returns schematically.

Fig. D5(b) shows us the possible routes for trajectories to pass through the four shaded areas. For instance, if we want a trajectory that goes 1-3-4-2-... then we start in region 1; to get to region 3 we must be in the right-hand "third" of region 1; if we are going to go on from region 3 to region 4, we must be in the right-hand "third" of that "third"; and, to go on to region 2, in the left "third" of that "third", etc.. For any infinite sequence of symbols 1,2,3 and 4 allowed by Fig. D5(b) we can find a vertical line of points (arrived at by taking away "two-thirds" of an interval an infinite number of times) from which trajectories pass through the four shaded areas in the prescribed sequence. In addition, there will be vertical lines of points which generate finite sequences which represent trajectories which eventually strike the top face of B on the line $\xi = 0$ (and thereafter tend towards the origin and never strike the top face of B again). All these vertical lines, from which trajectories remain within \mathcal{D} for all future time, form a Cantor set, and trajectories which start between lines eventually fall into either of the intervals $0 < |\xi| < B$ and, hence, eventually leave our small region of interest.

If we now consider where trajectories came from, we can divide the four shaded areas into horizontal bands, each of which corresponds to some finite history. For example, if we are in region 1 (Fig. D5) and came from region 2, we know we are in the lower horizontal band which is the intersection of region 1 and $\psi(2)$. We can repeat this process indefinitely, arriving at a Cantor set of horizontal lines which represent the possible histories (allowed by Fig. D5(b)) of trajectories which remained within our small region of interest for all past time. The points which lie on the intersections of the horizontal and vertical Cantor sets will be points through which trajectories pass that do not leave our small

(a)



(b)

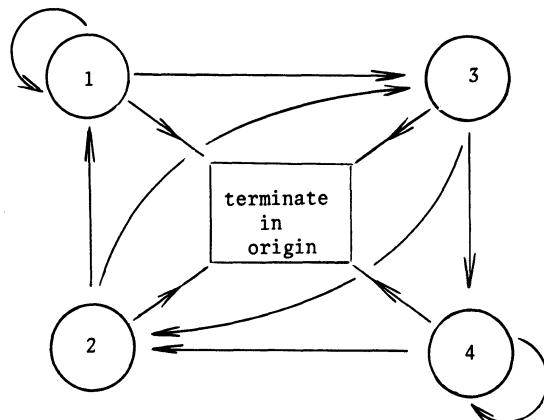


Figure D5. (a) Schematic version of the first returns of the shaded areas in Fig. D4; (b) Trajectories may go from one shaded area to another according to the arrows on this graph.

region of interest in either forwards or backwards time. Trajectories passing through any other points do leave the region of interest, or have come from outside it, or both.

Each point lying on an intersection will have a unique symbolic representation (which is infinite to the left, but which may terminate on the right) in terms of symbols 1,2,3, and 4. Conversely, each sequence of symbols, infinite on the left, will correspond to one point, providing that the sequence can be generated by following the arrows in Fig. D5(b). This last proviso actually tells us that we only need two symbols, T and S, to describe points. If we write T for 3 or 4, and S for 1 or 2, then any infinite sequence of T's and S's has one and only one representation in 1's, 2's, 3's and 4's which obeys the Fig. D5 rules. For example, STTTSST... must be 1344213... . Thus, the points which lie in our "invariant set" (the set of all points lying on intersections of the vertical and horizontal lines), each of which is taken by the return map to another such point, correspond 1-to-1 with symbolic sequences, infinite on the left, of T's and S's. The way we have chosen the symbols T and S also means, conveniently, that trajectories threading their way through the top face of the box B according to some sequence of T's and S's, make the same sequence of passes through the tubes, T and S.

We can tell how trajectories look by examining their corresponding sequence. Periodic trajectories (a countable infinity) will have repeating sequences, trajectories which terminate in the origin (an uncountable infinity) will have terminating sequences, trajectories that are asymptotically periodic (in either forwards or backwards time - an uncountable infinity) will have sequences which begin or end by repeating, and aperiodic trajectories (an uncountable infinity) will have aperiodic sequences. Notice that there will not be quite so many trajectories as points on the top face of B since trajectories will, in general, intersect this plane in a large (possibly infinite) number of points. When considering points, we have some central mark in the sequence to tell us where we are, and with trajectories we do not. For points, each application of the return map moves the central mark one place to the right. As an example, consider the "period 2" periodic orbit represented by the doubly infinite sequenceTSTSTSTS.... . There is only one such orbit. This orbit intersects the top face of the box B in two points, one of which lies to the left of the stable manifold of the origin and which is represented by the doubly infinite sequence (with mark)TSTSTSTS...., the other of which lies to the right of the stable manifold of the origin and which

is represented by the doubly infinite sequence (with mark)TSTSTSTS.... . Each application of the return map moves the mark one place to the right. Hence, we as follow the orbit around, we oscillate between one point and the other.

We call the collection of trajectories which remains within \mathcal{D} for ever a "strange invariant set". It is invariant because the flow takes it into (and onto) itself. The bifurcation which occurs as r passes through r' , and which creates the strange invariant set, we call a "homoclinic explosion". We can mention some interesting properties of the strange invariant set.

1) No single trajectory, nor any subset of the strange invariant set, is stable; we can find trajectories, as close as we like to any trajectory in the strange invariant set, which leave \mathcal{D} in both forwards and backwards time.

2) Even if we restrict our attention to those trajectories in our strange invariant set, we can see that almost all pairs of trajectories started at points close together on the top face of the box B will not remain close together as we follow them around. This is called "sensitive dependence on initial conditions", a phenomenon which persists even when the strange invariant set becomes attracting (see Chapter 3) and which gives the typical "chaotic", "turbulent" or "pseudo-random" behaviour which we associate with a "strange attractor". We can see why this phenomenon should occur. Two points on the top face of B are close together if the parts of their symbolic descriptions near the central mark are the same. As we follow trajectories starting at two close together points, we move to points with the same symbolic descriptions but with the central mark moved further and further to the right. Eventually we can expect the central mark (which tells us which region we are in at the moment) to have moved to parts of the two sequences which are quite different. This argument will only fail to apply to two close together points on the same vertical line ($\xi = \text{constant}$); such points will have converging futures (but distinct pasts). The vertical lines are an example of a contracting foliation; we discuss this in a more general context in Chapter 3. Notice that it is the existence of the contracting foliation which justified our decoupling of the η -coordinate in equation (4); the one-dimensional maps of Figs. D2 and D3 are derived by identifying together all the points on a vertical line (which have convergent futures). What is required to make the arguments in this appendix completely rigorous, is to prove that the ideal (vertical) foliation which we see in the schematic Fig. D5 persists under the small perturbations involved in moving from the "true" Fig. D4

to the "ideal" Fig. D5. Such proofs are available; see the references at the beginning of this appendix.

3) There is a very real sense in which the strange invariant set is a single object and not just the sum of its parts. It contains an uncountable infinity of dense trajectories; these are trajectories which pass as close as we like to all trajectories in the strange invariant set. (We can construct doubly infinite sequences of T's and S's which contain every possible finite sequence of T's and S's; the trajectories corresponding to these sequences pass as close as we like to any other trajectory in the invariant set. If it were not for the trajectories which terminate in the origin, we could construct aperiodic trajectories which pass as close as we liked to *every part of every trajectory* in the invariant set; this is the usual meaning of a dense trajectory. Kaplan & Yorke (1979a) appear to have ignored the terminating trajectories; hence their rather stronger statements.)

4) Periodic orbits are dense in the strange invariant set. Since periodic orbits can contain any finite number of symbols which repeat, it is clear that we can find a periodic orbit which passes arbitrarily close to any other trajectory in the strange invariant set.

The strange invariant set that we have been discussing is very similar to Smale's horseshoe example, introduced in his seminal (1967) paper. Our use of "symbolic dynamics" to describe the set also follows Smale.

We should consider the difference between type (a) and type (b) explosions. The strange invariant sets born in both types (i.e., from diagrams (iii) and (iv) of Fig. D2) are topologically identical, and the only differences between the two types that we have observed so far are that the strange invariant set exists when $\xi_r^* < 0$ in the type (a) explosion, and when $\xi_r^* > 0$ in the type (b) explosion, and that $f_r^* > 0$ in type (a) and $f_r^* < 0$ in type (b). In Chapter 2 we note that there is another difference (which can be deduced from the previous two) which turns out to be important when we consider whole sequences of homoclinic explosions later in these notes. The behaviour of the unstable manifold of the origin changes differently on either side of the different types of explosion. We can see this by considering the successive iterates of $0+$ in Figs. D2; these iterates will tell us the ξ -coordinate of the various successive points in which the right-hand branch of the unstable manifold of the origin intersects the top face of the box B. If we write a T for each iterate lying in $\xi > 0$, and an S for each iterate lying in $\xi < 0$, we obtain a sequence of T's and S's which describes the path taken

by the right-hand branch of the unstable manifold of the origin. This trajectory is not part of the strange invariant set, and leaves our small region of interest after a finite number of returns to the top face of the small box B. However, by choosing r-values closer and closer to r' , we can make this finite number as large as we like and, in this sense, we can write an infinite sequence of T's and S's describing the behaviour of the manifold in the limit as r approaches r' from above and from below. Examination of Figs. D2 shows us that we have: (i) TTTTT.....; (iii) TSSSS.....; (iv) TTSTSTST..., and (vi) TSTSTSTS.... . The differences between type (a) and type (b) explosions are summarized in Chapter 2.

Finally, we can ask what the effect would be of perturbing the symmetry of the system slightly. Our single homoclinic explosion would break up into an infinite number of homoclinic bifurcations, each of which produced some of the orbits and trajectories in our strange invariant set. The exact order in which these bifurcations occur would depend on the way in which we break the symmetry. We will not discuss this further here; it is not important for our study of the Lorenz equation and the techniques of Appendix G, appropriately modified, can be used to produce a non-symmetric theory.

Appendix E

Numerical Methods for Studying Return Maps and for Locating Periodic Orbits

The techniques described in this appendix are well known. Descriptions of the theory on which they are based can be found in many places. See, for example, Curry (1979a).

1. LOCAL PROPERTIES OF THE RETURN MAP

We write the Lorenz flow as $\phi_\tau(\underline{x}_0)$ where τ is the time and \underline{x}_0 is the initial position. We have $d\phi_\tau/dt = f(\phi_\tau(\underline{x}_0))$ where $f(\cdot)$ is the vector field given in the Lorenz equations. If we start a trajectory at a point $\underline{x}_0 = (x_0, y_0, r-1)$ on the plane $z = r-1$ it will strike the plane $z = r-1$ again at a point $\underline{x}_1 = (x_1, y_1, r-1)$ after a time τ^* . We can calculate x_1, y_1 , and τ^* by numerical integration of the equations. We have $\underline{x}_1 = \phi_{\tau^*}(\underline{x}_0)$.

If we start a trajectory on the plane $z = r-1$ and near to \underline{x}_0 we wish to know where it will strike the plane $z = r-1$ near to \underline{x}_1 .

If we write \underline{dx}_0 ($= (dx, dy, 0)$) for a small change in initial condition, $d\tau$ for the resulting small change in the return time, and $\underline{dx}_1 = (dx_1, dy_1, 0)$ for the resulting small change in the position where the trajectory next strikes the plane $z = r-1$, we wish to know \underline{dx}_1 in terms of \underline{dx}_0 .

Expanding to first order about \underline{x}_0 and τ^* we have

$$\phi_{\tau^*+d\tau}(\underline{x}_0 + \underline{dx}_0) = \phi_{\tau^*}(\underline{x}_0) + d\tau \frac{d}{dt} \phi_{\tau^*}(\underline{x}_0) + (D_x \phi_{\tau^*}) \underline{dx}_0$$

where $(D_x \phi_{\tau^*})$ is a matrix of partial derivatives which can be calculated by integrating the nine equations $\frac{d}{dt}(D_x \phi_\tau) = (Df)_{\phi_\tau} (D_x \phi_\tau)$ along with the original equations. The extra equations are, for instance,

$$\frac{d}{dt} \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right) = \sigma \left[\frac{\partial \mathbf{y}}{\partial \mathbf{x}_0} - \frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right]; \quad \frac{d}{dt} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}_0} \right) = r \frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} - \frac{\partial \mathbf{y}}{\partial \mathbf{x}_0} - x \frac{\partial z}{\partial \mathbf{x}_0} - z \frac{\partial \mathbf{x}}{\partial \mathbf{x}_0}$$

with initial conditions

$$\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0}(0) = 1; \quad \frac{\partial \mathbf{y}}{\partial \mathbf{x}_0}(0) = 0 \quad \text{etc. .}$$

We can rewrite equation (1) as

$$\begin{pmatrix} dx_1 \\ dy_1 \\ 0 \end{pmatrix} = dt \begin{pmatrix} f_1(\underline{x}_1) \\ f_2(\underline{x}_1) \\ f_3(\underline{x}_1) \end{pmatrix} + (D_{\mathbf{x}} \phi_{\tau^*}) \begin{pmatrix} dx_0 \\ dy_0 \\ 0 \end{pmatrix} \quad (2)$$

where the f_i are the components of the vector field \mathbf{f} . Equation (2) is a system of three linear equations in three unknowns (dx_1, dy_1 , and $d\tau$) which can therefore be calculated in terms of dx_0 and dy_0 . Notice that our choice of a return plane $z = \text{constant}$ means that we only have to compute six of the nine partial derivatives in the matrix $(D_{\mathbf{x}} \phi_{\tau^*})$.

This saves computation time and effort.

We are unlikely to be particularly interested in the quantity $d\tau$, so we can write the interesting part of equation (2) as

$$\begin{pmatrix} dx_1 \\ dy_1 \end{pmatrix} = A \begin{pmatrix} dx_0 \\ dy_0 \end{pmatrix} \quad (3)$$

where A is some real 2×2 matrix of known quantities. The matrix A will, of course, depend on \underline{x}_0 and we will have to recalculate it for each new \underline{x}_0 .

The preceding analysis can be extended to cope with other return planes; we may, however, need to calculate all nine partial derivatives if the plane is not of the form $z = \text{constant}$.

2. LOCATION OF PERIODIC ORBITS

We may wish to locate a periodic orbit. We do this by attempting to find a point on some suitable return plane (usually $z = r-1$) where the periodic orbit intersects the plane. What we wish for is a solution to the equation $\phi_{\tau^*}(\underline{x}^*) = \underline{x}^*$. If we have an initial guess \underline{x}_0, τ_0 for the position and period of the orbit we can calculate, by integrating the equations, the quantity $\underline{x}_1 = \phi_{\tau_0}(\underline{x}_0)$. In order to locate the periodic orbit we need to know quantities \underline{dx} and $d\tau$ such that $\underline{x}^* = \underline{x}_0 + \underline{dx}$

and $\tau^* = \tau_0 + d\tau$. Expanding to first order around x_0 and τ_0 we obtain, similar to equation (1) above,

$$x_0 + dx = x_1 + d\tau f(x_1) + (D_x \phi_{\tau_0}) dx . \quad (4)$$

In equation (4) we know the quantities x_1 , x_0 and $f(x_1)$, and can calculate the matrix of partial derivatives as explained in the last section. Thus equation (4) is a linear system of three equations in four unknowns (dx, dy, dz , and $d\tau$). We have one extra degree of freedom since it is possible to locate any point on the periodic orbit. We therefore specify that $dz = 0$ which ensures that we are looking for the point of intersection of the periodic orbit with the return plane. This choice allows us, once again, to get away with calculating only six of the nine partial derivatives in the matrix $(D_x \phi_{\tau})$. We can now rewrite equation (4) as

$$\begin{pmatrix} \frac{\partial x}{\partial x_0} - i & \frac{\partial x}{\partial y_0} & f_1(x_1) \\ \frac{\partial y}{\partial x_0} & \frac{\partial y}{\partial y_0} - i & f_2(x_1) \\ \frac{\partial z}{\partial x_0} & \frac{\partial z}{\partial y_0} & f_3(x_1) \end{pmatrix} \begin{pmatrix} dx \\ dy \\ d\tau \end{pmatrix} = \begin{pmatrix} ix_0 - x_1 \\ iy_0 - y_1 \\ z_0 - z_1 \end{pmatrix}$$

where $i = 1$ and where the f_j are the three components of the vector field.

From equation (5) we can calculate the three quantities dx , dy and $d\tau$ after integrating a set of (nine-dimensional) differential equations only once. If $x_0 + dx$ and $\tau_0 + d\tau$ are not sufficiently good estimates for the position and period of the orbit we can use them as new initial guesses and repeat the procedure. Before discussing convergence and practical use of this method, note the following modification of the method which may sometimes be useful.

Some of the periodic orbits in the Lorenz system are symmetric. If we wish to locate such an orbit we can either use the method described above, or we can look instead for a solution to the equation $\phi_{\tau^*}(x^*) = \bar{x}^*$ where \bar{x}^* is the symmetric image of x^* and τ^* is half the period of the orbit. The analysis of this method follows through exactly as before except that we must take $i = -1$ in equation (5). This method will locate only symmetric orbits and the value of the period which it returns is half the true period of the orbit. (Note: The orbit locating technique described in this section is essentially equivalent to the technique

for studying return maps described in 1) above. If we take equation (3) from 1) and consider it in conjunction with the equation $\underline{x}_1 + \frac{dx}{dr} = \underline{x}_0 + \frac{dx}{dr}$ then we obtain, unsurprisingly, the same equation for $\frac{dx}{dr}$ that we would obtain by using the appropriate return time as our guess for the period of the orbit in equation (5). Either technique can be used to locate periodic orbits, but, as explained in Section 4 below, it sometimes seems to be an advantage to be able to specify position and period guesses independently. The reasons for this are more practical than theoretical.)

The method described here works equally well on stable and non-stable orbits.

3. FOLLOWING ORBITS WITH CHANGING r

If we treat the parameter r as a variable, we can extend the methods described above so that they will locate the periodic orbit at a new value of the parameter r . In expanding to first order to obtain the equivalent of equation (4) we also expand to first order in r . The analysis carries through virtually unchanged except that the right-hand side of equation (5) becomes

$$\left[\begin{array}{l} x_0 - ix_1 - dr \frac{\partial x}{\partial r}(\tau_0) \\ y_0 - iy_1 - dr \frac{\partial y}{\partial r}(\tau_0) \\ z_0 - z_1 - dr \frac{\partial z}{\partial r}(\tau_0) \end{array} \right]$$

where dr is the desired change in r and the quantities $\partial x / \partial r$, etc., are calculated by integrating the three additional differential equations

$$\frac{d}{dt} \left(\frac{\partial x}{\partial r} \right) = \sigma \left(\frac{\partial y}{\partial r} - \frac{\partial x}{\partial r} \right)$$

$$\frac{d}{dt} \left(\frac{\partial y}{\partial r} \right) = r \frac{\partial x}{\partial r} + x - \frac{\partial y}{\partial r} - x \frac{\partial z}{\partial r} - z \frac{\partial x}{\partial r}$$

$$\frac{d}{dt} \left(\frac{\partial z}{\partial r} \right) = x \frac{\partial y}{\partial r} + y \frac{\partial x}{\partial r} - b \frac{\partial z}{\partial r}$$

along with the original equations. The initial conditions for these three variables are $\frac{\partial x}{\partial r} = \frac{\partial y}{\partial r} = \frac{\partial z}{\partial r} = 0$.

4. CONVERGENCE, PRACTICAL HINTS AND BIFURCATIONS

The convergence properties of the orbit locating method are described in Curry (1979a). In "normal" circumstances convergence is rapid (second order) as with Newton's method. Initial guesses for position and period of an orbit need only be moderately accurate. In practice, it seems that it is more important to have a good estimate for the period than it is to have a good estimate for the position. For parameter values where the system behaves chaotically, good guesses for the position of the orbit can be obtained by observing a chaotic solution until it performs a sequence of loops similar to the sequence of loops in the desired periodic orbit. A good guess for the period is then obtained by seeing how long it takes the trajectory to perform the required sequence of loops (starting and finishing on some suitable return plane). If the orbit locating method fails to converge onto the desired periodic orbit with these guesses then other nearby guesses can be tried. In practice it seems to work best if one tries different nearby guesses for the period. (See the remark at the end of Section 2.)

The orbit following technique works well if we restrict ourselves to small jumps in the parameter r . To avoid accumulating errors it is best to use the fixed- r orbit locating method to obtain an accurate position and period for the orbit at each new r -value before moving on to the next r -value. The orbit following technique does not work well for large jumps in the parameter r . However, having calculated the true position and period of an orbit at two close together r -values it is often possible to produce accurate estimates at a distant r -value by linear extrapolation. In addition, the orbit following program can be made to handle much larger r -jumps (and use much less computation time) if one records the differences between the true position and period of an orbit and the position and period first predicted by the program and then applies this correction (which can be constantly updated) to subsequent estimates at new r -values.

When following periodic orbits over wide ranges of parameter values it is important to remember that they will move in phase space and that they may become tangential to, and then cease to intersect, any one return plane. Thus, it may be wise to change the return plane on which the orbit is located as the parameter is changed. In general the techniques described here work best when working on a return plane that is intersected "perpendicularly" by the periodic orbits. If, for example, we are trying to locate and follow a periodic orbit like that shown schematically

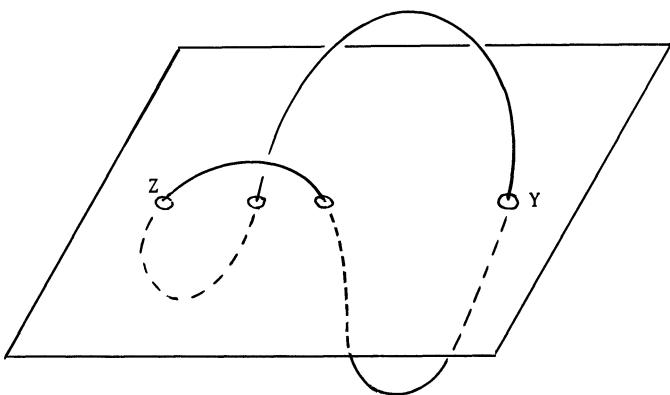


Figure E1. It is easier to locate periodic orbits where they intersect the return plane "perpendicularly" (Y).

in Fig. E1, we can expect better results if we use points near Y as our guesses for the position of the orbit than if we use points near Z .

There are three situations in which our techniques will not work.

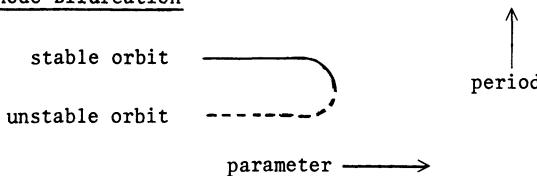
(i) When the period of an orbit is so long that appreciable numerical inaccuracies occur in integrating the equations over one period.

(ii) If the determinant of the matrix on the left-hand side of equation (5) becomes extremely large. This occurs when a trajectory passes near the origin and we get "extreme" sensitive dependence on initial conditions. Periodic orbits can only be followed a certain distance towards the homoclinic explosions which produce them; the figures in Chapter 5 that show almost homoclinic periodic orbits were obtained with a certain amount of difficulty. When attempting to locate periodic orbits in these almost homoclinic situations we can sometimes achieve better results by using quantities $x_0 + \frac{1}{2}dx$ and $\tau_0 + \frac{1}{2}d\tau$ as our new guesses for position and period of the orbit, rather than using the whole corrections dx and

$d\tau$ as calculated. In some circumstances the orbit locator will converge onto the orbit with these half corrections when it will not converge using the normal technique. (Situations (i) and (ii) above are different aspects of the same problem. Both involve numerical difficulties. (i) involves the general numerical problem of long-term integration in a chaotic system. Here, the orbit locator does not necessarily "know" that it is having difficulty, it just fails to converge onto the desired orbit. Situation (ii) involves the particular difficulties associated with integrating a trajectory which passes close to the origin - see Appendix F - and the orbit locator tells us it is having trouble when the determinant of the matrix on the left-hand side of equation (5) becomes large.)

(iii) The last situation in which we may have difficulty with our technique is when the determinant of the matrix on the left-hand side of equation (5) approaches zero. This occurs when there are two or more orbits of similar period near to one another and we are trying to locate one of them. This situation occurs if we are approaching a bifurcation of a non-homoclinic kind. Providing we proceed slowly near the bifurcation, it is usually possible to follow the periodic orbit we are interested in very close to the bifurcation value. Sometimes the orbit locator will jump onto one of the nearby periodic orbits; this may be an advantage if we wish to locate all the orbits involved in the bifurcation.

We shall now discuss the three kinds of bifurcation which we describe in Chapter 4. In what follows we will call the matrix on the left-hand side of equation (5) B_1 or B_{-1} , depending on whether we are using the technique to locate the whole of a periodic orbit ($i = 1$ in equation (5)) or the technique to locate half a symmetric orbit ($i = -1$ in equation (5)). The values of $\det(B)$ that we will be interested in are the values calculated with our best values for position and period of the orbit at each parameter value. Assuming we have located the periodic orbit accurately we can also consider the matrix A_1 or A_{-1} from equation (3), calculated for the appropriate return maps. Note the following properties of these various matrices: if we follow a symmetric orbit using the full-period locating method, then the return matrix A_1 so derived equals $(A_{-1})^2$ where A_{-1} is the matrix calculated when locating half the orbit; similarly, we may follow an orbit of period T as though it were an orbit of period $2T$ and obtain a return matrix A'_1 which equals $(A_1)^2$ where A_1 is the return matrix obtained normally; it is easy to show that $\det(B_1) = 0$ if and only if A_1 has an eigenvalue +1 and that $\det(B_{-1}) = 0$ if and only if A_{-1} has an eigenvalue -1.

a) Saddle-node Bifurcation

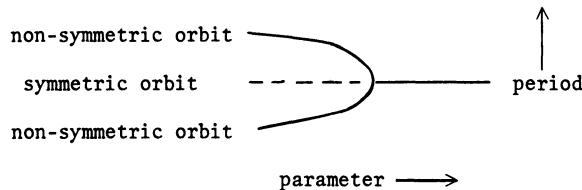
Suppose we are following a non-symmetric periodic orbit and $\det(B_1)$ approaches zero. Then the associated return matrix, A_1 , has an eigenvalue +1 and we are approaching a saddle-node bifurcation. This means that there is another orbit of similar period nearby. The value of r for which $\det(B_1) = 0$ is the bifurcation value, and attempts to locate the periodic orbit at r -values beyond the bifurcation value fail. If we are only a little beyond the bifurcation value the orbit locating program does converge onto the general area where the periodic orbit "used to exist" without ever converging to any particular value. (The explanation for this behaviour is essentially the same as the explanation for intermittent chaos - see Chapter 4.) We obtain a more accurate bifurcation value of the parameter by looking for $\det(B_1) = 0$ than we do by attempting to decide whether the orbit locating program is converging or not.

In the Lorenz system, one of the two periodic orbits involved in the saddle-node bifurcation is stable. Hence, if we are following the non-stable periodic orbit, the other orbit can be found easily. If the orbit we are following is the stable one, it is a little harder to locate the other (non-stable) orbit. We could develop more sophisticated, higher-order techniques which would allow us to locate both orbits. This would probably be necessary in higher-dimensional systems where both periodic orbits might be non-stable. In the Lorenz system it is easy enough to locate the other orbit by guessing its period and position. A good practical method to generate such guesses is to record the position and period of the orbit we are following as near as possible to the bifurcation value (period = T , position = x) and then to record the period and position of the same orbit at some nearby parameter value (period = $T + dT$, position = $x + dx$). The other orbit involved in this bifurcation will, at the same parameter value, have a period close to $T - dT$ and a position close to $x - dx$. If the orbit locating program does not converge onto the other orbit with these guesses, success can often be achieved by trying additional guesses $T - k.dT$ (where $k > 1$) for the period.

If we are following a symmetric orbit then both the available orbit locating methods indicate the presence of a saddle-node bifurcation. At

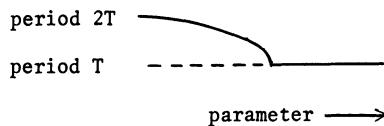
the bifurcation value the return matrices A_1 and A_{-1} have eigenvalues +1 and -1 respectively and both $\det(B_1)$ and $\det(B_{-1})$ equal zero.

b) Symmetric Saddle-Node Bifurcation



Since there is only one symmetric periodic orbit involved in this bifurcation, the bifurcation will not be detected by the orbit locating program if we are using it in the mode which locates half a symmetric orbit. This allows us, once we have located the symmetric orbit on the side of the bifurcation where it is stable, to follow it, without difficulty, through the bifurcation to the side where it is non-stable. If we wish to calculate the bifurcation value of the parameter, we can follow any of the orbits involved using the full-period orbit locating technique. We will have $\det(B_1) = 0$ at the bifurcation value, corresponding to the matrix A_1 having an eigenvalue +1. In the Lorenz system the bifurcation always appears to occur in such a way that the non-symmetric orbits involved are stable on the side of the bifurcation value where they exist; consequently, they are not difficult to locate. If the bifurcation were to occur with the non-symmetric orbits non-stable, we would need either higher-order techniques or good guesses in order to locate them.

Note the difference between this bifurcation and the saddle-node bifurcation involving two symmetric orbits. In the symmetric saddle-node bifurcation the return matrix A_{-1} obtained when we locate half the orbit has an eigenvalue +1 and $\det(B_{-1})$ does not equal zero. In the case of the saddle-node bifurcation between two symmetric orbits, A_{-1} has an eigenvalue -1 and $\det(B_{-1}) = 0$. In both cases the matrix $A_1 = (A_{-1})^2$ has an eigenvalue +1 and $\det(B_1) = 0$.

c) Period Doubling Bifurcation

Since there is only one periodic orbit with period T involved in this bifurcation, the bifurcation will not be detected if we are following this orbit. This allows us, once we have located this orbit on the side of the bifurcation where it is stable, to follow it through the bifurcation value into the region where it is non-stable. If we wish to determine the bifurcation value of the parameter we can follow any of the orbits involved using $2T$ as the period (a periodic orbit of period T is, as far as the orbit locating program is concerned, also a periodic orbit of period $2T$) and then we will obtain $\det(B_1) = 0$ at the bifurcation value. (If we look at the return map near the orbit of period T then the return matrix A_1 has eigenvalue -1 , which does not imply that $\det(B_1) = 0$. If we look at the return map generated when we follow twice around the orbit we obtain a return matrix $A'_1 = (A_1)^2$ and this has an eigenvalue $+1$ with $\det(B'_1) = 0$.) In the Lorenz system, the bifurcation always seems to occur in such a way that the orbit of period $2T$ is stable where it exists; consequently it can be easily located. If the bifurcation were to occur with this orbit non-stable we would need either higher-order techniques or good guesses to locate it.

Appendix F

Computational Difficulties Involved in Calculating Trajectories That Pass Close to the Origin

We wish to know, for trajectories which pass downwards through the plane $z = r-1$ (which all interesting trajectories seem to do), roughly how close to the stable manifold of the origin they must be on the plane $z = r-1$ if they are to pass within a certain distance of the origin. The calculation below is only approximate but should give an answer of the right order of magnitude.

Suppose a trajectory starts on the plane $z = r-1$ with x and y small. Then $\dot{z} \approx -bz$, and, providing the trajectory remains near the z -axis, the z -coordinate at time t is approximately $(r-1)e^{-bt}$.

At any particular r -value, if we assume z is fixed, the Lorenz equations are two-dimensional and linear with eigenvalues

$$\lambda_1(z), \lambda_2(z) = \frac{-(\sigma+1) \pm ((\sigma+1)^2 + 4\sigma(r-z-1))^{\frac{1}{2}}}{2}.$$

If we assume that the trajectory starts at a distance ε_0 from the stable manifold of the origin (on the plane $z = r-1$) and that it moves away from the stable manifold of the origin at a rate governed by the positive eigenvalue $\lambda_1(z)$, then we have $\frac{d\varepsilon}{dt} \approx \lambda_1(z)$ where $\lambda_1(z)$ is as above and $z = (r-1)e^{-bt}$. The solution to this differential equation is given by

$$\varepsilon(t) = \varepsilon_0 \left[\frac{p - (\sigma+1)}{p + (\sigma+1)} \right]^{p/2b} \left[\frac{p + \sqrt{p^2 - qe^{-bt}}}{p - \sqrt{p^2 - qe^{-bt}}} \right]^{p/2b} e^{\left(-\frac{(\sigma+1)}{2} t - \frac{1}{b} \sqrt{p^2 - qe^{-bt}} + \frac{(\sigma+1)}{b} \right)}$$

where $p^2 = (\sigma-1)^2 + 4\sigma r$ and $q = 4\sigma(r-1)$.

If we assume the trajectory strikes a plane $z = a$ near the origin it does so at time $t = \frac{1}{b} \ln(\frac{r-1}{a})$. The table below gives values $\ln(\frac{\epsilon}{\epsilon_0})$ for various values of b , r and a ($\sigma = 10$), calculated by substituting this value of t into the monstrous formula above.

r	b	a	$\ln(\epsilon/\epsilon_0)$
28	8/3	1.0	11.3
28	8/3	0.5	14.3
28	8/3	0.1	21.4
100	8/3	1.0	38.6
100	1/4	1.0	411
10000	8/3	100	464
10000	8/3	1.0	1000
10000	1/4	1000	2150

Examine, for example, the first line of this table. When $r = 28.0$ the intersection of the attractor with the plane $z = r-1$ stretches approximately nine units on either side of the stable manifold of the origin. If we assume that trajectories strike the plane $z = r-1$ near the attractor with a more or less uniform probability distribution, the probability that a trajectory will enter the box $|x| \leq 1$, $|y| \leq 1$, $|z| \leq 1$ after passing through the plane is approximately $\frac{1}{9} e^{-11.3}$. This is not a very large probability.

For larger r (and particularly with small b) we can see that it may well become impossible to compute a trajectory which starts on the plane $z = r-1$ and then passes close to the origin. With digital computation, trajectories can only start at a grid of points on $z = r-1$ and it may be that none of these points is close enough to the stable manifold of the origin for the trajectory to enter even a fairly large box around the origin. This is why we do not (in Chapter 8) try to perform numerical experiments at large r and small b , even though we believe there are some interesting trajectories which pass close to the origin at these parameter values.

Appendix G

Geometric Models of the Lorenz Equations

Several papers have been written on model flows which have attracting sets known as "Lorenz attractors". It is to be hoped that the Lorenz equations have a Lorenz attractor for some range of parameter values, but the global results needed to prove this are not available. In this appendix, we review some of the work done on model Lorenz attractors, ignoring questions of their relevance to the Lorenz equations.

The appendix is divided into two sections. In the first, we examine some results (including one new one) for one-dimensional maps similar to the map of Fig. 3.7 in Chapter 3. In the second, we see how these maps are related to model Lorenz flows. We shall give very few proofs, preferring to outline the major results while referring the reader to the original papers for details. All of this appendix relies heavily on a series of papers by Williams (W) and/or Guckenheimer (G) (W-1977, 1979, 1980, G-1976, G & W-1980).

1. ONE-DIMENSIONAL LORENZ MAPS

The maps we study satisfy the following conditions:

1. $f: I \rightarrow I$, $I = [0,1]$
2. f is continuous and differentiable except at c , $0 < c < 1$.
3. f is monotonic and strictly increasing on $[0,c)$ and $(c,1]$.
4. $\lim_{x \rightarrow c^-} f(x) = 1$, $\lim_{x \rightarrow c^+} f(x) = 0$, $f(c) = c$.
5. $f'(x) \rightarrow \infty$ as $x \rightarrow c$ from either side.
6. f is locally eventually onto (l.e.o.) (this condition is explained below).

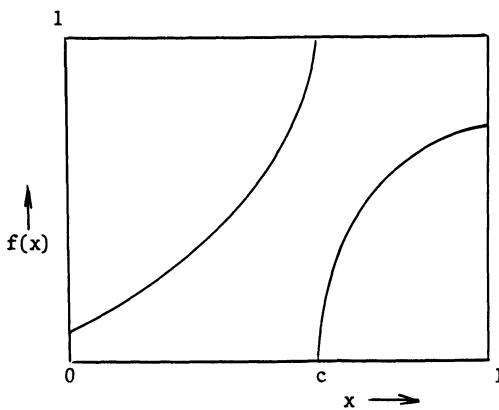


Figure G1. A map f satisfying conditions 1 - 6.

Fig. G1 shows a typical map f satisfying these conditions. Notice that we have not required that f has any symmetry properties. This extra generality produces few extra difficulties and is worthwhile.

The condition "locally eventually onto" (l.e.o.) means that for every interval $J \subset I$ there is some n such that $f^n(J) = I$. It will become clear why this condition is important as we proceed. Williams (1980) shows that f is l.e.o. if the slope of f is everywhere greater than $\sqrt{2}$. This fact is interesting but not important, and f may be l.e.o. even if the slope is not greater than $\sqrt{2}$ everywhere.

The first step (see also Rand, 1978) is to define a kneading sequence, $k(x)$, for every point in I . This is a sequence of symbols, 0 and 1, defined as follows:

$$\begin{aligned} k(x) &= s_0 s_1 s_2 s_3 s_4 \dots \dots \quad \text{where each } s_i = 0 \text{ or } 1 \text{ with} \\ s_i &= 0 \text{ if } f^i(x) < c \\ s_i &= 1 \text{ if } f^i(x) > c \\ \text{the sequence terminates if } f^i(x) &= c. \end{aligned}$$

We can see that $k(x)$ may be either finite or infinite depending on whether some iterate $f^n(x)$ of x equals c or not.

It is easy to see that a symbol sequence k corresponds to at most one point $x \in [0,1]$ since the l.e.o. condition implies that there is some n for which $f^n(x_1)$ and $f^n(x_2)$ are on different sides of c if

$x_1 \neq x_2$. The ordering of the real numbers in I carries over to an ordering on the kneading sequences. This ordering is essentially the binary ordering with the additional condition that the empty symbol is intermediate between 0 and 1. For example, $000\dots < 00 < 001\dots$. With this ordering, $x < y$ implies that $k(x) < k(y)$. This means that for all $x \in I$, $k(0) \leq k(x) \leq k(1)$.

We do not yet know which sequences of 0's and 1's correspond to some $k(x)$. The condition just given is necessary but not sufficient. Now notice that $k(x)$ and $k(f(x))$ are related in a very simple way. If $k(x) = s_0 s_1 s_2 s_3 \dots$, and $k(f(x)) = t_0 t_1 t_2 t_3 \dots$, then $t_i = s_{i+1}$ since $f^n(f(x))$ equals $f^{n+1}(x)$ (see definition of $k(x)$ above). This motivates our definition of the map s (the shift map) on sequences defined by:

$$\text{If } k = s_0 s_1 s_2 s_3 \dots, \text{ then } s(k) = s_1 s_2 s_3 s_4 \dots.$$

Thus $k(f(x)) = s(k(x))$ and by induction, $k(f^i(x)) = s^i(k(x))$. Now, if some point x lies in I , $f^i(x)$ lies in I also. Consequently, for a symbol sequence k to be a kneading sequence $k(x)$ for some x , we must have $k(0) \leq s^i(k) \leq k(1)$ for all $i = 0, 1, 2, 3, \dots$ (*). It can be shown that this condition (*) is sufficient as well as necessary and so any sequence k satisfying (*) is a kneading sequence for some point $x \in I$. This means that the two kneading sequences $k(0)$ and $k(1)$ completely determine which kneading sequences $k(x)$ can occur.

The kneading sequences, $k(x)$, tell us the kinds of behaviour we can expect from f . Periodic points of f will have infinitely repeating kneading sequences, points eventually mapped onto c will have finite kneading sequences, and points whose successive iterates under f wander around I aperiodically will have aperiodic sequences. Hence, $k(0)$ and $k(1)$, by determining the sequences $k(x)$ that can occur, also determine all the different types of behaviour possible under f . The converse of this remark is also true, though slightly harder to prove (see G & W, 1980 or Rand, 1978). This means that we can state, after defining $k'(x)$ to be the sequence obtained from $k(x)$ by changing all the symbols (replacing 0's by 1's and vice versa):

Theorem 1. Let f_1 and f_2 be two maps satisfying the conditions given at the beginning of this section. Then f_1 and f_2 are topologically equivalent if and only if $k_{f_1}(0) = k_{f_2}(0)$ and $k_{f_1}(1) = k_{f_2}(1)$, or if $k'_{f_1}(0) = k'_{f_2}(1)$ and $k'_{f_1}(1) = k'_{f_2}(0)$. (G & W, 1980.)

The last part of this theorem is necessary because two maps f_1 and f_2 will be topologically equivalent if one can be obtained from the other by rotating Fig. G1 through 180° . Notice that if f is symmetrical ($c = 1/2$, $f(x) = 1 - f(1-x)$), then $k(0) = k'(1)$ and we only need one kneading sequence to determine the topological conjugacy class of f . We call the pair of kneading sequences, $k_f(0)$ and $k_f(1)$, the kneading invariant for the map f .

It is interesting to note that there is one particularly simple representative of each topological class. This representative does not satisfy the condition that there be infinite slope on either side of the discontinuity, but that condition has not been used in any of the preceding analysis and was only specified in order that we can later build a Lorenz model flow around the maps f . Parry (1976) showed that each f satisfying our given conditions is topologically conjugate to a piecewise linear map, g , with constant slope (m) and one discontinuity (at c'). See Fig. G2. This gives a natural continuous mapping of the different conjugacy classes into a two-dimensional disc with "Parry"-coordinates (m, c') . In another paper, Parry (1977) discusses the construction of invariant probability measures for these piecewise linear maps and proves that they are weak-Bernoulli with respect to a unique probability measure absolutely continuous with respect to Lebesgue measure.

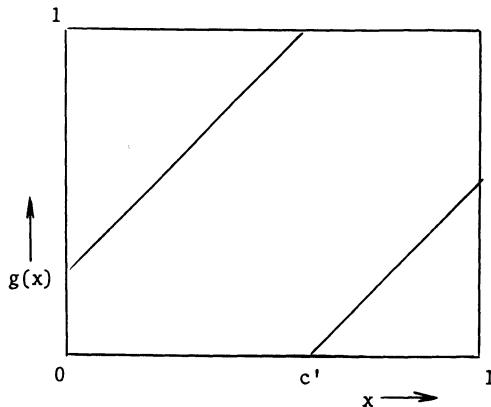


Figure G2. If a map F satisfies conditions 1 - 6 it is topologically conjugate to a piecewise linear map with constant slope.

We can ask how many different conjugacy classes there are. It is fairly clear that there is an infinite number, but we can obtain a better result:

Theorem 2. In any C^0 -open neighborhood of a map f satisfying our conditions, there is an uncountable number of maps (also satisfying our conditions) that have mutually distinct kneading invariants. (W 1980, Rand 1978. Note that Williams' 1980 paper was actually in wide circulation as a preprint as early as 1976. It is referred to in papers such as Rand, 1978. Rand's paper is a little easier to understand.)

This theorem is not really surprising. It seems reasonable, for example, that any small change in m or c' for the Parry piecewise linear maps will produce a map with different kneading invariants, and that every Parry map ($m > \sqrt{2}$) is topologically equivalent to a map f satisfying our conditions. However, the theorem does not tell us which pairs of symbol sequences will be possible kneading invariants for some map f . Not every pair of symbol sequences will be "realizable" as a kneading invariant, as we see below. This explains why the kneading sequences themselves (with some suitable topology) are not a very natural coordinate system for mapping the different topological conjugacy classes onto a two-dimensional disc. This next section is a little more difficult, but it is included because some previous work suggests that things are simpler than they actually are.

Given two symbol sequences, k^0 and k^1 , we want to know if there is some Lorenz map f which has the pair (k^0, k^1) as a kneading invariant. Clearly we must have:

$$\underline{\text{Axiom 1. }} k^0 < k^1.$$

In addition, the sequences k^0 and k^1 must themselves satisfy condition (*) from our previous discussion. Thus:

$$\underline{\text{Axiom 2. }} k^0 \leq s^i(k^0) \leq k^1 \text{ and } k^0 \leq s^i(k^1) \leq k^1 \text{ for all } i = 0, 1, 2, 3, \dots.$$

Williams (1979) gives these two axioms as necessary and sufficient for the two sequences to be realizable as a kneading invariant for some Parry map, and hence for them to be realizable as some Lorenz map. We can see that this is not quite right. First, for the Parry map to be equivalent to a Lorenz map we will need the slope, m , of the Parry map to be greater than $\sqrt{2}$. This will require some condition which may be quite complicated but which will be roughly, " k^0 starts with enough 0's

and k^1 starts with enough 1's". Assume that we have such a condition. We then have the following problem.

The argument in Chapter 3 that only certain homoclinic orbits can occur will generate an argument that only certain kneading invariants can occur. Some of those which are disallowed will satisfy axioms (1) and (2). Loosely speaking, these will be the pairs which correspond to periodic orbits and trajectories that we expect to be "dealt with" in some other homoclinic situation which does occur. Suppose, for example, that the pair (k^0, k^1) is a finite, realizable kneading invariant. Then other pairs of sequences, $(k^{0'}, k^{1'})$ given by

$$k^{0'} = k^0(0 \ k^1)^{n_1}(1 \ k^0)^{n_2}(0 \ k^1)^{n_3}(1 \ k^0)^{n_4} \dots$$

$$k^{1'} = k^1(1 \ k^0)^{p_1}(0 \ k^1)^{p_2}(1 \ k^0)^{p_3}(0 \ k^1)^{p_4} \dots$$

will not be realizable. (We need to add the extra symbols "0" in front of k^1 and "1" in front of k^0 because we have chosen to use $k(0)$ and $k(1)$ as our kneading invariant rather than using $k(c^+)$ and $k(c^-)$ which correspond in a more obvious way with the behaviour of the two branches of the unstable manifold of the origin.) We can give a simple example. (000, 111) is a realizable kneading invariant. Consider the possible kneading invariant (0000111, 1111000) which satisfies axioms (1) and (2) but which is of the form given above ($n_1 = p_1 = 1$, $n_2 = p_2 = 0$, etc.). We can construct a map h which has this pair as its kneading invariant. However, though h may look like Fig. G1, it will not be locally eventually onto. The easiest way to see this is by drawing a flow on a branched manifold which realizes the map h . See Fig. G3. We can see that the interval marked [A,B] on Fig. G3 is mapped eventually onto itself and does not expand to fill the whole of I. Axioms (1) and (2) are the conditions necessary to be able to draw a figure like Fig. G3. We require another axiom to ensure that the map generated by this figure is l.e.o..

The appropriate form of the extra axiom is:

Axiom 3. Let w_1 and w_2 be two sequences of the symbols 0 and 1. Both are at least two symbols long and one or both are finite. w_1 begins with a 0. w_2 begins with a 1. Then, k^0 and k^1 are not of the form:

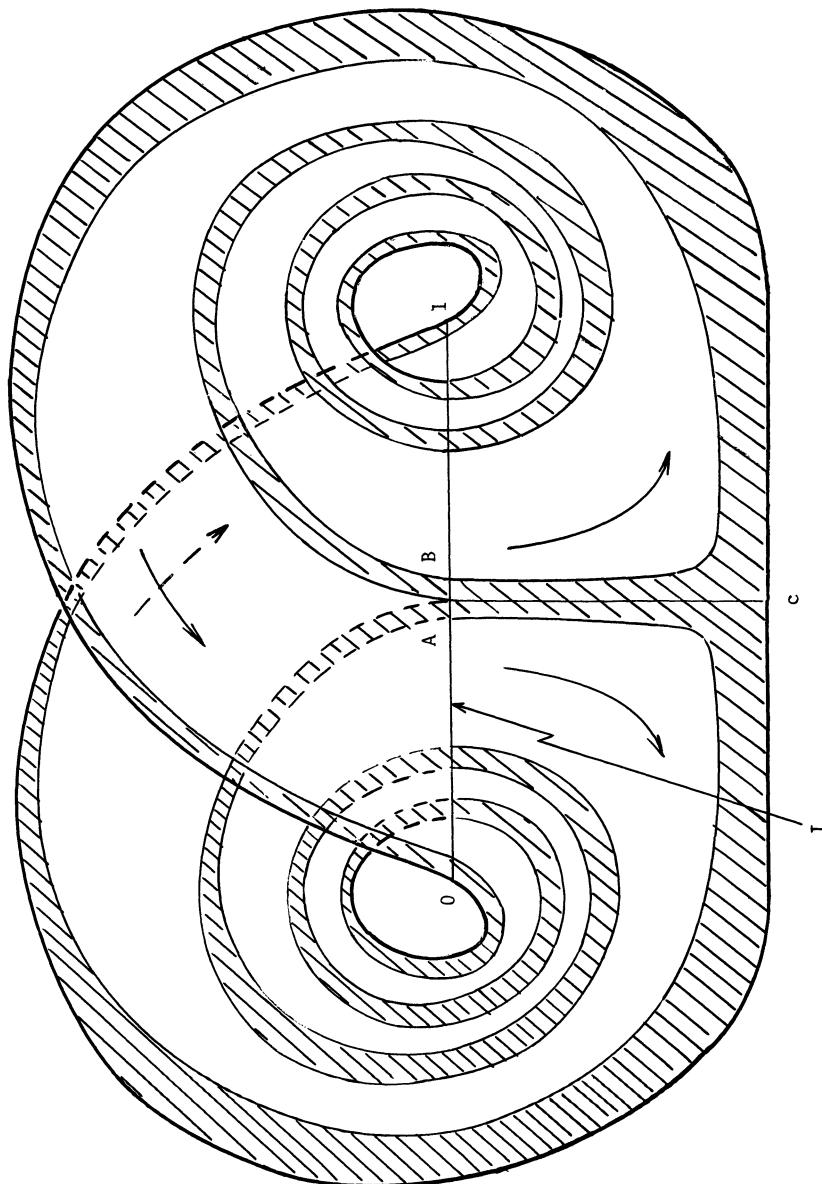


Figure G3. A semi-flow on a two-dimensional branched manifold. Trajectories started to the right of c , below the line I , travel round on the right-hand branch which lies in front of the left-hand branch. The two branches join on the line I . The map h induced on the line I (a one-dimensional map of I to itself) has kneading invariants $(000011, 1111000)$ as can be seen by following the trajectories starting at 0 and 1 . The map h is not l.e.o. because the interval $[A, B]$ is not eventually mapped onto the whole of I .

$$k^0 = w_1(0 \ w_2)^{n_1}(1 \ w_1)^{n_2}(0 \ w_2)^{n_3}(1 \ w_1)^{n_4} \dots \dots$$

$$k^1 = w_2(1 \ w_1)^{p_1}(0 \ w_2)^{p_2}(1 \ w_1)^{p_3}(0 \ w_2)^{p_4} \dots \dots$$

where, when w_1 and w_2 are finite, the sequences k^0 and k^1 can be finite or infinite (but $n_i, p_i > 0$), and where k^0 and k^1 are both infinite (with either n_1 or $p_1 = 1$ and all the rest of the $n_i, p_i = 0$) when one of w_1 and w_2 is infinite.

It remains to be proved that axioms (1), (2) and (3) are sufficient to characterize kneading invariants that are realizable (in some appropriate part of kneading invariant space), though it is fairly easy to see that they are necessary. Williams (private communication) has outlined a possible proof.

2. MODEL FLOWS

We now sketch the main steps needed to relate the maps of the last section to some model flows. We then give the two main results for those flows.

Under the one-dimensional maps f , points $x \in I$ have unique futures but non-unique pasts (the maps f are not invertible). If we want to construct a flow, we must give a unique past to each point x . We can construct a map \hat{f} on a space \hat{I} which is called the pinched inverse limit of f . Every point $\hat{x} \in \hat{I}$ consists of a point $x \in I$ together with a choice of a history for the point x (i.e., a sequence of points $\dots, x_{-3}, x_{-2}, x_{-1}$ such that $f(x_{-i}) = x_{-i+1}$). Since f is non-invertible, there will be several points in \hat{I} for each point in I . There will be some difficulties introduced by those points \hat{x} for which we choose some point $x_{-i} = 0$ or 1, since there is no point $x \in I$ for which $f(x) = 0$ or 1. For these points we allow the history to terminate at the point x_{-i} which equals 0 or 1.

Next, it is possible to construct a model return map, F , of the square into itself which has all the nice properties we would like the return maps of the real Lorenz system to have. The map is 1-1 everywhere except on a discontinuity $c \times I$, it has a contracting foliation, and the one-dimensional map which we obtain by shrinking each contracting arc to a point is one of our Lorenz maps f . The map looks like Fig. G4.

Since F is contracting there will be some attracting set, Λ , contained within the square. As we explained for the real Lorenz map (rather

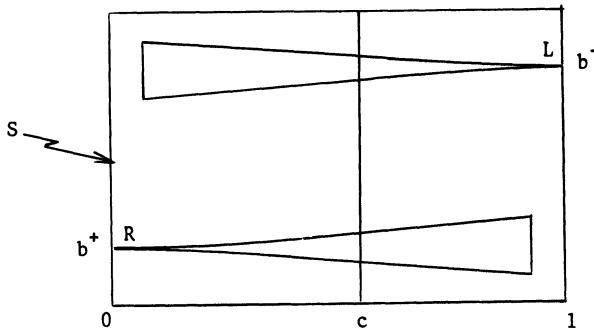


Figure G4. A two-dimensional map F of a square (S) to itself. F satisfies:

- (a) $F(x, y) = (f(x), g(x, y))$
- (b) F is 1-1 on the complement of $c \times I$.
- (c) $\frac{\partial g}{\partial y} < \frac{1}{2}$ (contraction in y direction)
- (d) As $x \rightarrow c^\pm$ the functions $g|_{x \times I}$ tend uniformly to constant functions b^\pm .

hopefully), this will consist of a Cantor set of arcs, and most points (x, y) on the attractor will lie on an arc with a Cantor set of other arcs nearby. However, notice that we have forced F to pinch the lines (c^+, I) and (c^-, I) to points. This means that in the neighborhood of points such as R and L (Fig. G4), the attractor will look like a Cantor fan of arcs. We then show that the map F restricted to A is topologically equivalent to the pinched inverse limit of f . Those special points in \hat{I} with finite histories will correspond to the special points on the attractor which have a Cantor fan neighborhood. Recalling our interpretation of points in \hat{I} as being points $x \in I$ plus a history of the point x , it is the history which determines which arc we lie on, and the point x which determines where on the arc we lie.

Finally, it is possible to show that we can construct a model flow, Φ , which has F as a return map. Once again, this is done in the way suggested by the real Lorenz equations. See Fig. G5 and its caption. The stationary point which lies below the square in Fig. G5 deals with all the troublesome pieces of the return map; the two-dimensional stable manifold attracts the line of discontinuity, $c \times I$, and the one-dimensional unstable manifold of the origin intersects the square in those points which have only finite histories.

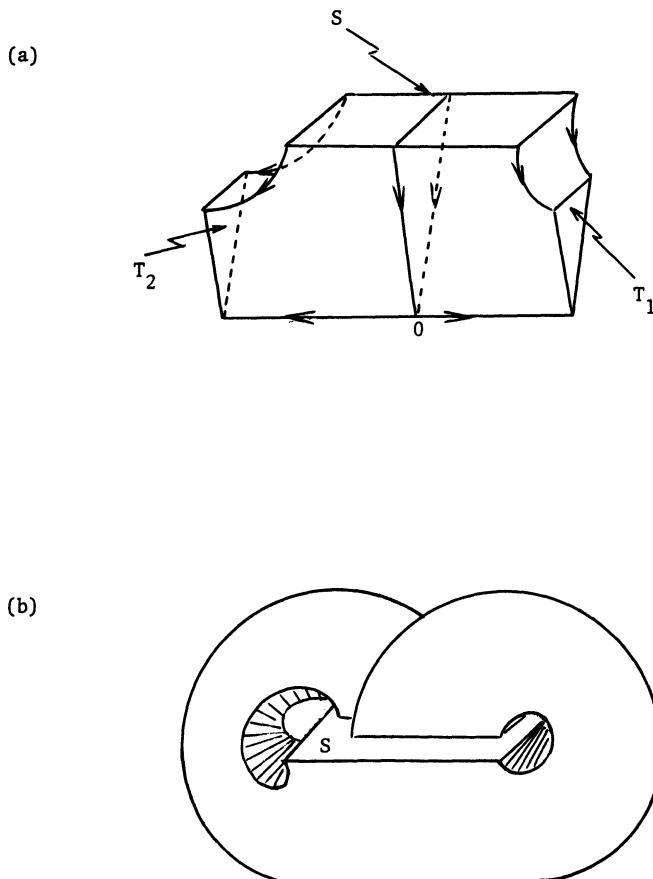


Figure G5. To build a flow from the return map of Fig. G4 we put the square S over a non-stable stationary point that has a two-dimensional stable manifold and a one-dimensional unstable manifold (a). The flow in the cell shown in (a) is linear. We then imbed this cell in a smooth vector field which takes the pieces T_1 and T_2 back to S in the way determined by Fig. G4. See (b).

We now have one "Lorenz attractor" in a model flow. It is possible to show that any small perturbation of this flow produces a perturbed flow, Φ' , which has a return map F' on some return surface S' such that F' has all the nice properties we need to extract a one-dimensional map f' of the interval to itself. Thus Φ' also has a Lorenz attractor. After

demonstrating that the two Lorenz attractors are topologically conjugate if and only if the two maps f and f' are topologically equivalent, we can state:

Theorem 3 (Williams & Guckenheimer, 1980). There is an open set V in the space of all vector fields in \mathbb{R}^3 , and a continuous mapping k of V into a two-dimensional disk, such that

- (A) Each $X \in V$ has a two-dimensional ("Lorenz") attractor which contains a stationary point;
- (b) X and $Y \in V$ are topologically conjugate by a homeomorphism close to the identity if and only if they have the same image under k .

The two-dimensional disc of this theorem is the disc generated by the kneading invariant of the maps f obtained from the flows (with some appropriate coordinate system). If we restrict our attention to symmetrical systems then we require only one kneading sequence to specify the kneading invariant and we can replace "two-dimensional" by "one-dimensional" in the theorem above. The phrase "homeomorphism close to the identity" is required to eliminate the problem associated with the possibility of producing conjugate attractors by using two maps f and f' that can be obtained from one another by rotating Fig. G1 through 180° .

Theorem 3 tells us how "good" things are. For example, in the symmetrical case it tells us that we can see all the different types of Lorenz attractor which exist in some neighborhood of a particular flow by changing only one parameter. This is called co-dimension one structural stability. There is another theorem due to Williams (1980), improving on a theorem of Guckenheimer (1976), which tells us how "bad" things are:

Theorem 4 (Williams, 1980). There are uncountably many mutually distinct "Lorenz" attractors in each open set of vector fields having "Lorenz" attractors.

Theorem 4 is clearly connected with Theorem 2 of the previous section. It is true whether we restrict our attention to symmetrical systems or not. This theorem implies that there will be homoclinic explosions at a dense set of r -values if the Lorenz system has a "Lorenz" attractor.

Appendix H

One-Dimensional Maps from Successive Local Maxima in z

1. $\sigma = 10$, $b = 8/3$, $r < 28$. (See Chapter 3)

The one-dimensional maps of Appendix G are difficult to compute numerically. However, an alternative one-dimensional map can be extracted very easily from the Lorenz flow by recording the successive local maximum values of z attained by trajectories, and plotting them one against the next. Providing the trajectory we consider lies near the non-wandering set, all the points obtained in this way appear to lie on a curve like Fig. H1. Several authors, notably Lorenz himself (1963, 1979, 1980a,b), have studied the Lorenz equations by looking at these maps.

Fig. H1 is the appropriate map for r -values such as 28.0 ($\sigma = 10$, $b = 8/3$). The diagonal dotted line is $z_{n+1} = z_n$ and we call the return map h . The cusp at z_M is a true cusp; we know that h is bounded because the Lorenz equations are, and the discontinuity in the derivative at $z = z_M$ corresponds to the discontinuity in two-dimensional return maps that is caused by the stable manifold of the origin. If an itinerary on Fig. H1 falls onto the point of the cusp at $z = z_M$, this corresponds to a trajectory in the real system which terminates in the origin. When we run a trajectory for a long time we obtain very few points near the cusp because the trajectory passes near the origin very seldom. However, by judiciously choosing a number of different initial conditions, we can calculate more points on the cusp if we wish.

The section of Fig. H1 within the dotted box is attracting; all itineraries move towards the dotted box and then remain there forever. If the absolute magnitude of the slope of h is everywhere greater than one (which seems to be the case for $r = 28.0$) we can immediately conclude that there are no stable periodic orbits. A more complicated analy-

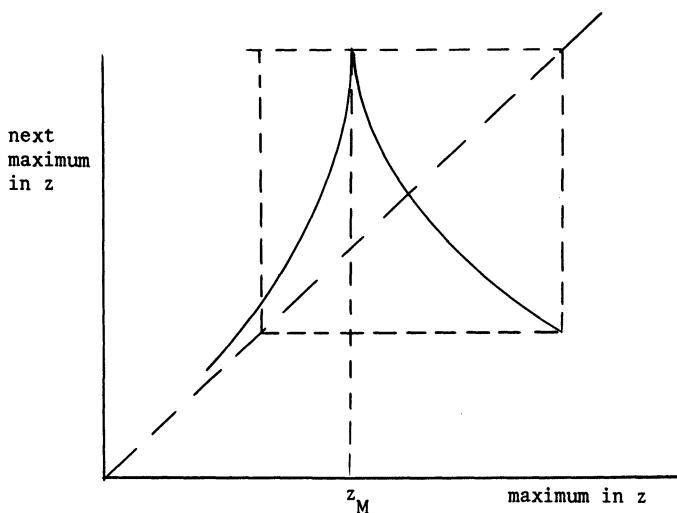


Figure H1. Plotting local maxima in z against one another produces approximately one-dimensional maps like that shown here (r near 28, $\sigma = 10$, $b = 8/3$).

sis will allow us to deduce something of the structure of the strange attractor.

Notice that the map in Fig. H1 is not symmetrical. By considering only the z -values attained by trajectories, we are identifying together all those points which are images under the natural symmetry. Each symmetric orbit of the Lorenz equations will be represented on Fig. H1 by an itinerary which has half as many points as the orbit has local maxima in z . Similarly, symmetric pairs of non-symmetric orbits will both be represented by the same itinerary on Fig. H1. For example, the fixed point of h represents the symmetric xy orbit of the Lorenz equations, and were there to be another fixed point of h to the left of z_M , this would represent both the simple x and y orbits (see below).

Yorke & Yorke (1979) studied the phenomenon of preturbulence using these maps. In the parameter range $r < 24.06$, the appropriate map is shown in Fig. H2.

The point C (Fig. H2) represents both of the stationary points C_1 and C_2 of the Lorenz flow, and is stable. The point A represents both of the simple x and y "Hopf" periodic orbits. Within the dotted

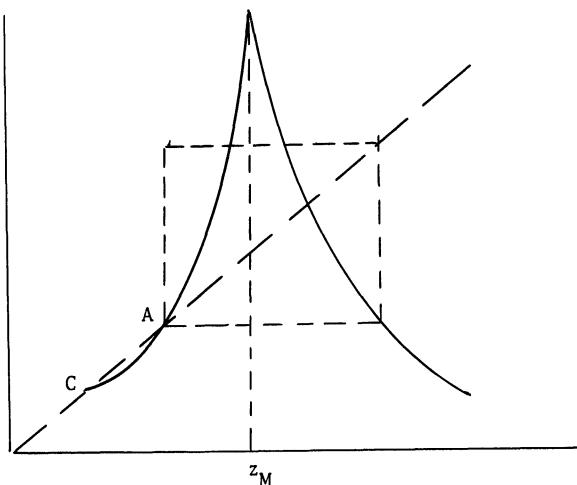


Figure H2. The map for $r < 24.06$.

box is a non-stable strange invariant set, but most itineraries started within the box eventually fall on those sections of the curve near to z_M which lie outside the box, after which they tend to C. Yorke & Yorke point out that near the cusp, $h(z) \approx h(z_M) - k|z_M - z|^\alpha$, where k and α are positive constants, $\alpha < 1$. The values of α which they calculate from numerical experiments are very close to the ratio in absolute magnitude of the two "strongest" eigenvalues of the linearized flow near the origin. As Yorke & Yorke remark, this is to be expected. The equation for the cusp allows us to calculate the width of the cusp at the critical height (above which trajectories are mapped in two steps to the left of A and thence to C - i.e., the top of the dotted box). The ratio of this width to the width of the whole dotted box will give some idea of the probability of a preturbulent trajectory escaping from the box on any one step. Yorke & Yorke give values of this probability for several values of r (as well as showing how it will change with changing r). When $r = 22.0$, the probability of escape on any one step is approximately 0.017. When $r = 23.0$, it is 0.0032. As r approaches 24.06 this probability tends to zero very rapidly and for a large part of $23 < r < 24.06$ it is negligible. Hence, numerically calculated trajectories (of the Lorenz equations) in this parameter range are hardly ever observed to escape from their preturbulent behaviour.

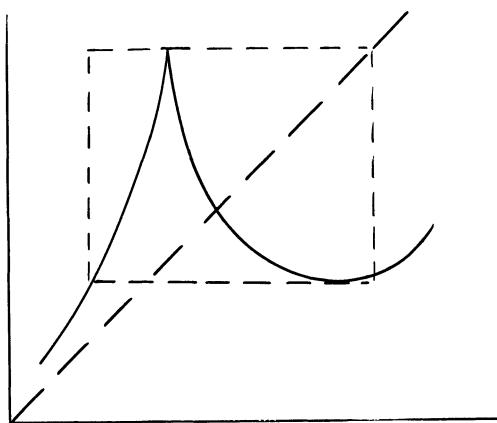


Figure H3. The map for $r > 30.2$.

2. $\sigma = 10$, $b = 8/3$, $r > 30.2$. (See Chapter 5.)

When $r > 30.2$, the appropriate map is shown in Fig. H3. As r increases from 28.0 (Fig. H1) the "tail" of the curve slopes less and less, and for $r > 30.2$ it slopes upwards again. The slope within the dotted box is no longer everywhere greater than one in absolute value and stable periodic orbits are possible. This corresponds (see Chapter 5) with the appearance of hooked return maps and the start of period doubling windows in the real Lorenz system.

3. b/σ SMALL. (See Chapter 8.)

Lorenz (1979) noticed that when b is small (r = Hopf bifurcation value, σ chosen to minimize r) the one-dimensional map extracted from the flow has two cusps. See Fig. H4. Once again, periodic orbits are possible. The curve now has three sections; points to the left of z_{M_1} correspond to sections of real trajectories which have two successive maxima in z in the same half-space $x < 0$ or $x > 0$; points between z_{M_1} and z_{M_2} correspond to sections of real trajectories which have one

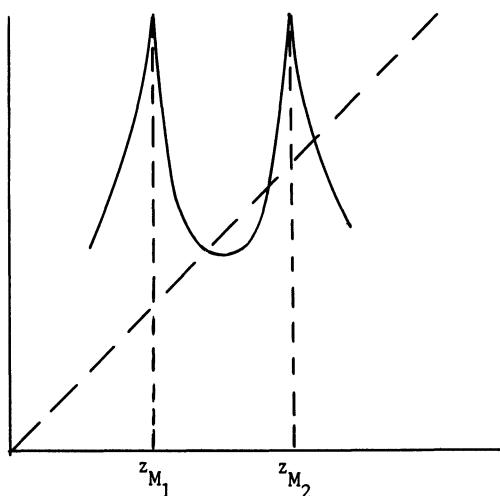


Figure H4. Multiply-cusped maps occur when b/σ is small.

maximum in one half-space and the next maximum in the other half-space; the new section of the curve to the right of z_{M_2} corresponds to sections of trajectories which have two successive maxima in the same half-space but which twist once around the z -axis in between. This behaviour is not observed for any r when $\sigma = 10$ and $b = 8/3$. See Chapter 8.

Fowler & McGuinness (1981a,b) have done analytical calculations, based on the ideas of matched asymptotic expansions, which predict multiply-cusped maps in a limit $\sigma \sim r \gg 1$, $b \sim 1$. Their work is discussed in Chapter 8.

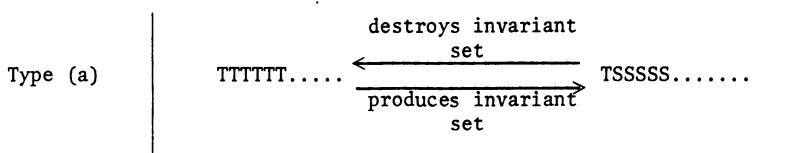
Appendix I

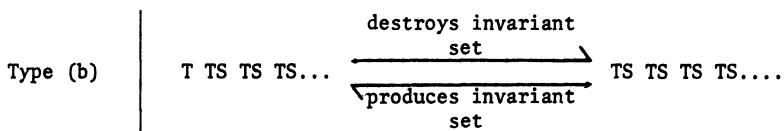
Numerically Computed Values of $k(r)$ for $\sigma = 10$ and $b = 8/3$

This appendix contains a list of numerically computed values of the symbolic sequence $k(r)$ which describes the behaviour of the right-hand branch of the unstable manifold of the origin. The list was calculated as follows.

A trajectory was started near the origin and displaced from it in the direction associated with the positive eigenvalue of the linearized flow. This trajectory was followed until it had reached seven local maxima in the variable z . Each time a maximum fell in $x > 0$ the program recorded an "x", and each time a maximum fell in $x < 0$ the program recorded a "y". (This method of assigning a symbol sequence to the trajectory is equivalent to other methods used in these notes - see Chapter 6.) The value of $k(r)$ - up to the seventh symbol - was sampled for increasing values of r . It was assumed that $k(r)$ only changed between two values of r if there was a homoclinic explosion for some intermediate r -value. It was also assumed that each homoclinic explosion was either type (a) or type (b). In Chapter 2 we showed that we know how the sequence $k(r)$ changes on either side of a homoclinic explosion. The rules were:

If we have a T -homoclinic explosion (where T is some finite sequence of x's and y's), and if we write S for the symmetric image of T (obtained by replacing all x's with y's and vice versa), then $k(r)$ changes as follows.





For example, if we have an xy -homoclinic orbit at some r -value, then $k(r)$ must change (with either increasing or decreasing r) from

$xy\ xy\ xy\ x$ to $xy\ yx\ yx\ y$

if the explosion is type (a), and from

$xy\ xy\ yx\ x$ to $xy\ yx\ xy\ y$

if the explosion is type (b). This result is obtained by substituting $T = xy$ and $S = yx$ into the rules above, and restricting our attention to the first seven symbols for the purposes of this appendix.

If, when we computed $k(r)$ for two different r -values, the two values obtained were neither the same, nor different in a way that could be explained by a single homoclinic explosion, then the r -interval was halved repeatedly until a series of values of $k(r)$ was computed, each of which was either the same as the previous one or different in a way that could be explained by a single homoclinic explosion. (This was always possible, as we would expect.) Thus, each change in $k(r)$ in the list can be explained by a single homoclinic explosion. It is possible that some changes in $k(r)$ have been missed. (If the value of $k(r)$ was the same at two r -values, it was assumed that there was no homoclinic explosion between these r -values; it is always possible that there was actually a sequence of explosions in this interval which changed $k(r)$ and then returned it to its original value. However, there are arguments in Appendix J that suggest that we have not missed many - if any - of the changes which can be detected by following $k(r)$ for seven symbols. There is, of course, an infinite number of changes in $k(r)$, but most of these are associated with homoclinic orbits which require seven or more symbols x and y for their description. We will not detect these changes if we only compute the first seven symbols in $k(r)$.)

The list of values of $k(r)$ is in two parts. The first part (Table II) covers the parameter range $r < 60.8$. In this list we have indicated the homoclinic explosion necessary to explain each observed change in $k(r)$. We have also indicated the type of each explosion. (The type of the explosions can be determined directly from the list in only two cases. One is the xy^2 -explosion for r between 46.5 and 47.5; the other is the xy -explosion for r near 54.646. In all the other cases - where the

explosion needs four or more symbols to describe it - the type (a) and type (b) changes are indistinguishable if we only know $k(r)$ to seven symbols. The types of these other explosions were determined by other means. See Chapter 2.) Notice that every explosion is proceeding in the direction which produces a strange invariant set. This is to be expected (see Appendix J). Fifteen explosions have been numbered. These are the explosions which produce periodic orbits needed for period doubling windows which involve a periodic orbit that existed in the original strange invariant set born in the first homoclinic explosion at $r \approx 13.926\dots$. These are also the only explosions which change $k(r)$ in a "permanent" way. All other explosions in the list occur in pairs which have no net effect on $k(r)$. These pairs are sometimes nested as marked on Table I1. For some r -value near 60.8, it seems that the unstable manifold of the origin is attracted to the non-stable symmetric xy orbit which exists in the non-wandering set. At this r -value $k(r) = xyxyxy\dots$ which is also the value of $k(\infty)$. Near the xy -explosion (number 13), $k(r)$ changes for very small changes in r . This is not unexpected, since at these r -values we are following the trajectory for three passes close to the origin. We have not indicated the exact r -values for which $k(r)$ takes its various values near this bifurcation since these are extremely sensitive to alterations in integrating routine, initial conditions, etc. However, the order in which these changes occurs is not sensitive to these alterations; any sufficiently accurate application of the technique used to compute this list appears to produce a series of close together r -values for which $k(r)$ has the different values given.

The second part of the list (Table I2) records the successive values of $k(r)$ in $r > 60.8$. The total effect of all these changes is nothing (in terms of the value of $k(r)$), and no changes can be detected in $r > 500$. The various explosions needed to explain the changes in $k(r)$ have not been marked, but can be deduced by comparing successive values of $k(r)$. Once again, each explosion appears to proceed in the direction which produces a strange invariant set, though not all have been checked. The nested "pairs" of changes which return $k(r)$ to its original value have been marked. Notice that these "pairs" are not strictly pairs; in the case of the xyx -explosions at $r \approx 64.569$ and $r \approx 124$ we actually have four related changes which return $k(r)$ to its original value. A similar phenomenon would be observed for all the "pairs" if we followed $k(r)$ for a greater number of symbols. This will be discussed further in Appendix J.

TABLE II

r	$k(r)$	Explosion	Type	
28.0	$\overline{xy\cdots\cdots} \ xy^5$	(b)	1	
39.2	$\overline{xy\cdots\cdots} \ xy^4$	(b)	2	
41.0	$\overline{xy\cdots\cdots} \ xy^4x$	(b)	3	
42.5	$\overline{xy\cdots\cdots} \ xy^3$	(b)	4	
43.5	$\overline{xy\cdots\cdots} \ xy^3xy$	(b)	5	
44.5	$\overline{xy\cdots\cdots} \ xy^3x$	(b)	6	
45.5	$\overline{xy\cdots\cdots} \ xy^3x^2$	(b)	7	
46.5	$\overline{xy\cdots\cdots} \ xy^2$	(b)	8	
47.5	$\overline{xy\cdots\cdots} \ xy^2xy^2$	(b)	9	
48.25	$\overline{xy\cdots\cdots} \ xy^2xy$	(b)	10	
48.75	$\overline{xy\cdots\cdots} \ xy^2xyx$	(b)	11	
49.5	$\overline{xy\cdots\cdots} \ xy^2x$	(b)	12	
50.5	$\overline{xy\cdots\cdots} \ xy^2x^2y$	(b)		
51.5	$\overline{xy\cdots\cdots} \ xy^2x^2$	(b)		
52.5	$\overline{xy\cdots\cdots} \ xy^2x^3$	(b)		
53.5	$\overline{xy\cdots\cdots} \ xy^2x^3$	(a)		
54.63	$\overline{xy\cdots\cdots} \ xy^2x^2$	(a)		
54.644	$\overline{xy\cdots\cdots} \ xy^2x^2y$	(a)		
54.646	$\overline{xy\cdots\cdots} \ xy$	(b)	13	
54.646	$\overline{xy\cdots\cdots} \ xyxy^2x$	(a)	14	
54.646	$\overline{xy\cdots\cdots} \ xyxy^2$	(a)		
54.646	$\overline{xy\cdots\cdots} \ xyxy^3$	(a)		
54.66	$\overline{xy\cdots\cdots} \ xyxy^3$	(b)		
56.0	$\overline{xy\cdots\cdots} \ xyxy^2$	(b)		
57.0	$\overline{xy\cdots\cdots} \ xyxy^2x$	(b)		
58.0	$\overline{xy\cdots\cdots} \ xyxy^2x$	(a)		
59.267	$\overline{xy\cdots\cdots} \ xyxy$	(b)	15	
59.267	$\overline{xy\cdots\cdots} \ xyxyx$	(a)		
59.267	$\overline{xy\cdots\cdots} \ xyxyy$	(b)		
60.8	$\overline{xy\cdots\cdots} \ xyxyx$	(b)		

TABLE I2

Appendix J

Sequences of Homoclinic Explosions

In this appendix we consider the arguments of Section 5.6 in more detail. There are few proofs, but the conjectures we give improve our understanding of the changes in behaviour which we may expect as σ and b vary.

1. SEQUENCES OF EXPLOSIONS AS A COMBINATORIAL PROBLEM

We can consider some aspects of the sequence of homoclinic explosions which occurs in the Lorenz system as a purely combinatorial problem, without reference to the dynamics of the equations. In order to make precise statements it is necessary to consider sequences k_i of symbols x and y that are finite of length n , though n can be arbitrarily large. Suppose that we are given two such sequences, $k_1 = \dots x x \dots$ and $k_{xy} = x y x y x y x y \dots x y$, and two rules which tell us how we may change a sequence. The two rules are:

- (a) Let T be a finite sequence of symbols x and y (of length less than n), and let S be the sequence obtained from T by changing all x 's to y 's and all y 's to x 's. Then if we have a sequence k_i which is the same as $T T T T T T T \dots$ up to the n^{th} symbol, we can change k_i to $T S S S S \dots$. It is also possible to make the change in reverse.
- (b) If T and S are defined as above, and k_i is the same as $T T S T S T S T \dots$ up to the n^{th} symbol, then we can change it to $T S T S T S T \dots$. Again, we can also make the reverse change.

These two rules are clearly motivated by the way in which the sequence $k(r)$ (which describes the unstable manifold of the origin) changes as we pass through a type (a) or a type (b) homoclinic explosion.

We will define a *Lorenz list* to be a list of values of k_i which starts with k_1 and ends with k_{xy} and for which each sequence in the list can be obtained from the previous one by one of the two rules. A trivial example of a Lorenz list with $n = 3$ is:

$$\begin{aligned} k_1 &= \text{xxx} \\ k_2 &= \text{xyy} \\ k_{xy} &= \text{xyx} \end{aligned}$$

where we get from k_1 to k_2 by substituting $T = x$ and $S = y$ in rule (a), and from k_2 to k_{xy} by substituting $T = xy$ and $S = yx$ in either rule (a) or rule (b). A trivial example of a list which is not a Lorenz list is:

$$\begin{aligned} k_1 &= \text{xxx} \\ k_{xy} &= \text{xyx}. \end{aligned}$$

Here there is no value of T that will get us from k_1 to k_{xy} in one step if we are to obey either rule (a) or rule (b). Our explanation of the way in which we computed the list in Appendix I (which is a Lorenz list) could be summed up by saying that we continued to look for intermediate values of r for which the value of $k(r)$ was different from the values we already had until the list was a Lorenz list.

We will define a *minimal Lorenz list* to be one which never takes the same value of k twice (or more times). We will define a *basic Lorenz list* to be a minimal list which uses only rule (a) to explain the changes between successive values of k in the list. It is fairly easy to see that there is only one basic Lorenz list for each n (the length of the sequences k that we are considering), and that it starts with $k_1 = \text{xxxx.....x}$ and $k_2 = \text{yyyy...y}$. The example we gave of a Lorenz list (above) is the basic Lorenz list for $n = 3$ (each change can be explained by rule (a) and the list does not repeat).

(The observation that basic Lorenz lists are unique means that if the Lorenz system has properties ensuring that only type (a) explosions occur - as it appears to have in $r < 30.1$ - then there is only one possible sequence of homoclinic explosions which can occur providing the development of the system is minimal, i.e., "monotonic". This conclusion is true even if the additional conditions for the study of the geometric

model to be appropriate do not hold. The statement - see Appendix G - that the symmetrical system has a one-dimensional space of kneading sequences, is true whenever we have only type (a) explosions.)

Now, for a general Lorenz list, we must create some kind of register which contains finite symbol sequences (\approx periodic orbits) which we add to or subtract from as we proceed down the list from k_1 to k_{xy} . We add and subtract from the register in the obvious way; i.e., between each pair of successive values, k_i and k_{i+1} , we say that we have either a type (a) or type (b) T-explosion (though we don't need to define this in any dynamical way) according to whether the change from k_i to k_{i+1} was made using rule (a) or rule (b) and where T is the sequence used in the application of the relevant rule. The finite words added to or subtracted from the register correspond in the obvious way to the periodic orbits born or destroyed in the equivalent dynamical T-homoclinic explosion, though we only consider periodic orbits of length $n-1$ or less. (We need $n-1$ rather than n since, with k only length n , we do not detect explosions of length n and hence will miss some changes which add or subtract periodic orbits of length n .) Thus, we add all the periodic orbits of length $n-1$ or less that are contained in a T-generated strange invariant set if we have either rule (a) going in the direction TTTT... to TSSS..., or rule (b) going in the direction TSTS.... to TTST..., and subtract the same orbits otherwise (cf. the rules for homoclinic explosions given in Chapter 2 and repeated in Appendix I). All this can be defined without reference to any dynamics (though it seems unnecessary to spell it all out), as can a property "symmetry" which we do or do not assign to words in the register in the obvious way. In what follows we use dynamical terms freely without continually reminding readers that they can be given non-dynamical meanings.

We now define *non-destructive Lorenz lists*. If we state that the register of periodic orbits contains nothing at the top of a Lorenz list (equivalent to assuming that there are no periodic orbits in the Lorenz system until after the first homoclinic explosion), then a non-destructive Lorenz list is one which, in general, never attempts to remove a periodic orbit from the register which it did not previously place there. The one exception which we allow is that we permit an xy explosion to remove periodic orbits from the register if the only two orbits which are missing are two non-symmetric xy orbits. We henceforth restrict our attention to non-destructive Lorenz lists since we claim that destructive Lorenz lists are not dynamically interesting. (Dynamically speaking, the restric-

tion to non-destructive lists does not seem reasonable if we only consider the problem for one particular value of n . Each change in the list will try to remove at most a finite number of periodic orbits of length $n-1$ or less, and it is easy to conceive of a series of simple non-homoclinic bifurcations in the Lorenz system which would provide these orbits. However, it is fairly easy to see - combinatorially - that a homoclinic explosion which removes a finite number of orbits not created in previous explosions, also removes a countable infinity of other orbits - of length greater than n - that were also not produced in previous explosions. It is hard to conceive of a sequence of non-homoclinic bifurcations which could produce all these orbits and put them together into a strange invariant set suitable for removal in a homoclinic explosion. We allow the exception because all the orbits for an xy -explosion - except two non-symmetric xy orbits - can be produced in a previous explosion; that explosion would be the original x -explosion. We discuss this point further in a moment.)

First, observe that the basic Lorenz lists are non-destructive. The first change in the basic lists is a type (a) x -explosion which adds the original x -generated strange invariant set to the register. All subsequent changes (also type (a) by definition) remove periodic orbits from the register (cf. the development of the strange attractor in Chapter 3). The final change in the basic lists (from $xy\ yx\ yx\ yx\ \dots$ to $k_{xy} = xyxyxy\dots$) is a type (a) xy -explosion for which we have all the orbits except two non-symmetric xy orbits. What we have left at the end of a basic list of changes is one x orbit, one y orbit, and one symmetric xy orbit. All the foregoing statements are easy to prove (see, for instance, Section 3 of this appendix). We can ask whether it is plausible that the Lorenz system realizes the basic list of changes for some parameter values σ and b . The answer is that it is plausible (though not necessarily the case); all we need to do is to assume that the behaviour develops in such a way that type (a) explosions continue to remove periodic orbits from the non-wandering set as r increases (and the hooks never appear in the return map), with the non-stable symmetric xy orbit gaining stability in a symmetric saddle-node bifurcation - see Chapter 4 - in which it throws off the two non-stable non-symmetric xy orbits needed for the final xy -explosion. This sequence of bifurcations has the happy advantage of leaving us with a stable symmetric xy orbit which we expect to exist at large r -values, and there is no reason to suppose that the Lorenz system does not behave in this way for some σ and b values.

Returning to the consideration of general non-destructive Lorenz lists we can conjecture:

Conjecture A. Every non-destructive Lorenz list passes, in the correct order, through the values of the basic Lorenz list.

As an example, Appendix I contains a non-destructive Lorenz list with $n = 7$. If we start with $k_1 = \text{xxxxxx}$, proceed through the values of $k(r)$ listed for r -values from 28.0 to 45.5, jump to the values given for 48.25 and 48.75, and end with $k_{xy} = \text{xyxyxyx}$, these eleven sequences, taken in the order given, form the unique basic Lorenz list for $n = 7$. (Each change in this list of eleven sequences agrees with rule (a) for suitable T's up to the seventh symbol, regardless of what the type of the explosions in Appendix I actually are.)

The point of this conjecture is to allow us to break any non-destructive Lorenz list into "blocks", each block corresponding to a single rule (a) change in the basic list. We can then consider the effect of each block separately. We can conjecture that each block of changes in a non-destructive Lorenz list is equivalent to the single rule (a) change from the basic list to which it corresponds, in the following sense:

Conjecture B. The first block (which changes k from xxx....x to xyy....y) adds an original x -generated strange invariant set to the register. Subsequent blocks (except the last one) either remove the same orbits from the register that the equivalent rule (a) change would have done or add exactly the right number and type of periodic orbits to the register to dispose of the same orbits in period doubling windows which terminate in saddle-node bifurcations. The last block either removes the same periodic orbits from the register that the equivalent rule (a) xy -change would have done or adds exactly the right number and type of periodic orbits to the register to dispose of the same orbits in period doubling windows which terminate in saddle-node bifurcations, plus exactly the right number and type of periodic orbits for a period doubling window which leaves us with a stable symmetric xy orbit. Every block may also add "extra" orbits to the register, but the number and type of "extra" orbits produced in each block will be exactly right for disposal in period doubling windows terminating in saddle-node bifurcations.

(For the purposes of this conjecture, we could give a non-dynamical definition of "orbits required for a period doubling window". We refrain from doing so.) Once again we can use Appendix I as an example. Consider the explosions numbered 7, 8 and 9. These, together, change $k(r)$ from

$xyy\ yxx\ y$ to $xyy\ xyy\ x$. This change is the change effected by the single rule (a) xy^2 -change in the basic list. Such a change would remove orbits xy^2 , yx^2 , and x^3y^3 (symmetric) from the register (considering only orbits up to length 6). Instead, we have three explosions, all of which create orbits. Explosion 7 gives us xy^3x^2 and yx^3y^2 (= two non-symmetric x^3y^3) orbits, explosion 8 gives us an xy^2 , a yx^2 and a symmetric x^3y^3 orbit, and explosion 9 gives us xy^2xy^2 and yx^2yx^2 orbits. Together with the orbits we already have (which would have been removed in the type (a) change), these orbits are exactly what is needed (considering only orbits up to length 6) for two period doubling windows. One is a symmetric x^3y^3 window (like the x^2y^2 window of Chapter 4), the other is the xy^2 window (studied in Chapter 4). Again considering Appendix I, everything after explosion 10 forms the last block of changes. Explosions 11, 12, 13 and 15 are the ones which deal with orbits already in the register at the beginning of the block. Explosion 14 is there to deal with extra orbits produced in explosion 13 that are not needed for disposal of orbits in the register at the beginning of the block. As we increase n we will see more and more of such detail in any non-destructive Lorenz list which contains some rule (b) changes. We will find that conjectures A and B, with appropriate modifications, will apply to such details, and to sub-blocks within sub-blocks within blocks etc., rather in the spirit of renormalization group techniques now often used to study some maps of an interval to itself. Similar arguments will show that the cancelling groups of "extra" changes (of which we see many in Appendix I, particularly for $r > 60.8$) either do nothing or produce period-doubling-windows-worths of "extra" orbits. Conjectures A and B, together with what we know about the effect of the basic Lorenz lists, will prove Conjecture 1 of Section 5.6. Conjectures A and B may be fairly easy to prove, though I have not been able to do so in the time available. It is almost certain that they are true, just because of the way in which they will apply (appropriately modified) to ever finer details of the different non-destructive Lorenz lists that we can construct as n increases; it is not really possible for the conjectures to be false in some minor detail - such a detail would spread its effect through the whole list, and it is likely that we would see some sign of it.

2. EXPLOSION EXPLOSIONS

Conjectures A and B indicate how we may expect some aspects of the behaviour of the Lorenz equations to change as we alter parameters σ and b . Suppose that we fix σ and allow b to vary (without loss of generality). This will almost certainly have the effect of changing the point in the Lorenz list of changes in $k(r)$ at which type (b) explosions start to occur. We may suppose that, at one b -value, a certain collection of periodic orbits is destroyed in a single type (a) explosion as r increases. At another b -value, these same orbits will be removed in period doubling windows with the help of periodic orbits born in an infinite number of homoclinic explosions, all belonging to the block of explosions which corresponds to the single type (a) explosion in the basic list. We can expect a "homoclinic explosion explosion", where one explosion becomes an infinite number of explosions, as b passes through some critical value b^* . Since the dynamics of the Lorenz equations will ensure that the period doubling windows occur almost instantly near the critical b -value, (since when the hooks first appear, the points R' and B are very close together - see Sections 5.1 through 5.3) we can expect an extremely complicated bifurcation at b^* and the appropriate r -value. This complicated bifurcation (in two-dimensional parameter space) is associated with the problem of explosions which are neither type (a) nor type (b), and explains our reluctance to consider that case elsewhere in these notes. We do consider the problem a little in Section 3 of this appendix. We may summarize the change in behaviour as b passes through b^* as follows:

Single type (a) explosion \Rightarrow	<div style="display: inline-block; vertical-align: top;"> <p>A countable infinity of explosions directly connected with the orbits previously destroyed in the type (a) explosion. These explosions are needed for us to have a combinatorially possible non-destructive Lorenz list.</p> <p>A countable infinity of period doubling windows involving the orbits from the explosions above.</p> <p>A countable infinity of "extra" explosions, such as those seen in Appendix I, which must happen between the explosion mentioned above for dynamical reasons - see later in this appendix.</p> <p>A countable infinity of period doubling windows involving orbits from the "extra" explosions. One or more of these windows may occur simultaneously with those mentioned above - see the arguments in Section 5.6.</p> </div>
--	--

The distinction between "extra" explosions and the others is a real one, dynamically as well as combinatorially. If we assume that b is increasing through b^* in the diagram above, each "extra" explosion only

appears for some b greater than b^* , though there is a countable infinity of such explosions for any b greater than b^* . On the other hand, each of the explosions needed combinatorially (that produce orbits with the same symbolic descriptions as the orbits that would have been destroyed in the type (a) explosion), exist for all b greater than b^* .

In addition, we should remember that there may be other type (a) explosions occurring arbitrarily close to the one we have considered, which may "explode" at b -values arbitrarily close to b^* .

3. ONE-DIMENSIONAL MAPS AND LORENZ LISTS

We may think of our combinatorial treatment of Lorenz lists as a generalization of a partial kneading sequence theory for certain families of one-dimensional maps of an interval to itself. A particularly simple family of such maps is shown schematically in Fig. J1.

A family of maps like that shown in Fig. J1 is an obvious choice if we wish to try and model the behaviour of the Lorenz equations in the parameter range $\sigma = 10$, $b = 8/3$ and $r > 24.06$. (Compare with the maps in Appendix G.) The discontinuity models the origin (and we have infinite slope near this discontinuity), and the behaviour of the successive iterates of 0_+ will model the behaviour of the right-hand branch of the unstable manifold of the origin. We can define a symbol sequence k^λ according to whether these iterates lie in $t > 0$ or $t < 0$, and k^λ will model $k(r)$. If some iterate of 0_+ lies on zero we have a homoclinic situation, and this will be type (a) if an even number of iterates lie on parts of the map with negative slope, type (b) if an odd number of iterates lie on parts of the map with negative slope, and neither type (a) nor type (b) if some iterate of 0_+ lies on a turning point of the map. Ignoring the last case for the moment, k^λ will change according to the (a) and (b) rules as we pass through type (a) and type (b) homoclinic situations, and the production and destruction of periodic orbits will occur in the usual way in each of these situations.

Numerical experiments indicate that families of maps like those shown in Fig. J1 model most aspects of the behaviour of the Lorenz equations very well. In $0 < \lambda < \lambda_1$ (Fig. J1), all homoclinic explosions are type (a), modelling the parameter range $24.06 < r < 30.1$. At $\lambda = \lambda_1$ the turning points appear and the map is no longer monotonic. In $\lambda > \lambda_1$ some homoclinic explosions will be type (b). At $\lambda = \lambda_2$ we have the xy-homoclinic explosion ($r \approx 54.6$), and at $\lambda = \lambda_4$ we see the turning points

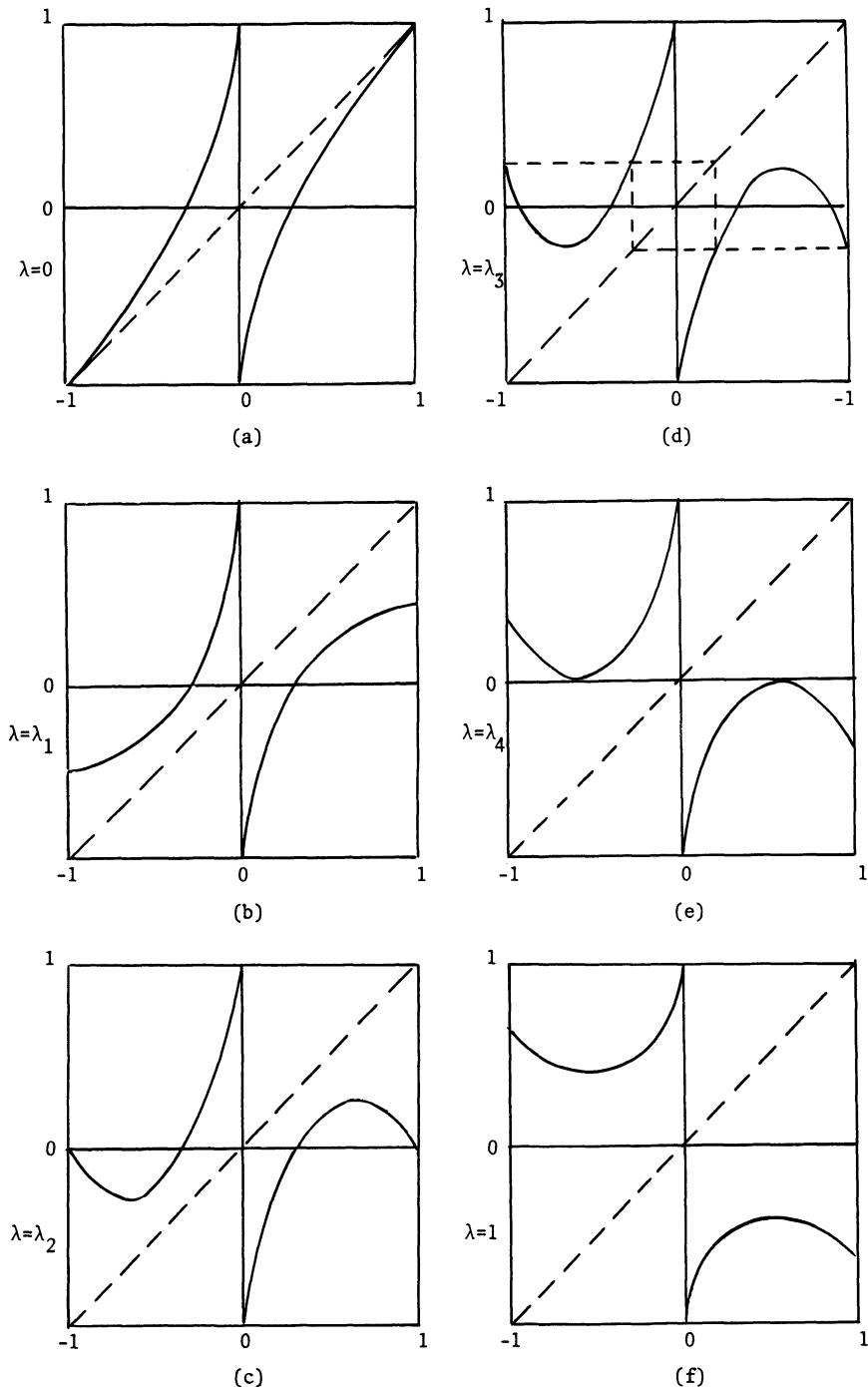


Figure J1. A family of maps, $f_\lambda: [-1,1] \rightarrow [-1,1]$. The maps have the symmetry $f(t) = -f(-t)$ and have infinite slope as you approach zero.

cross zero ($r \approx 197.6$). There is some λ_3 a little larger than λ_2 for which $f^2(0+) = -f^3(0+)$ (see Fig. J1), so that iterates of $0+$ lie alternatively on the two points $f^2(0+)$ and $f^3(0+)$ and $k^{\lambda_3} = xyxyxyx\ldots$. This parameter value models the r -value for which the unstable manifold of the origin is included in the stable manifold of the symmetric xy orbit. (See Conjecture 2 of Section 5.6. Arguments on these one-dimensional maps are equivalent to arguments on simplified return maps like those shown in Fig. 5.8 of Chapter 5, though they are slightly easier to explain.) It is clear that maps like Fig. J1 do not model all the behaviour of the Lorenz equations; for example, in $\lambda > \lambda_4$ all trajectories oscillate $xyxyxy\ldots$ forever, whereas in $r > 197.6$ (the parameter value at which all easily observed trajectories oscillate $xyxyxy\ldots$ forever) we have continuing changes in the sequence $k(r)$. Nonetheless, a great number of the details of the changes in $k(r)$ can be predicted using these maps. We will give two examples. Fig. J2 shows the successive iterates of $0+$ for some homoclinic λ -values near λ_3 . The development shown in Fig. J2 is a reasonable monotonic development in the behaviour of the map, and the sequence of homoclinic situations shown is identical to the sequence of explosions on either side of $r = 60.8$ in Appendix I. Notice that it is Fig. J2 which permits us to decide that $r = 60.8$ is the right choice for the r -value nearest to the special r -value at which we have $k(r) = xyxyxy\ldots$ forever. There are several r -values in Appendix I for which $k(r) = xyxyxy$ up to the seventh symbol, but arguments on figures like Fig. J2 quickly convince us that at these other points in the list, $k(r)$ cannot continue to oscillate $xyxyxy\ldots$ forever. Fig. J3 shows two situations for $\lambda > \lambda_3$. These predict the xyx -explosions at r near 64.569 and r near 124 (see Appendix I). More detailed arguments allow us to predict exactly the complete sequence of homoclinic explosions that lies between these two explosions (see Appendix I), though it is necessary that we make some assumptions about the relative movement of the point $f_\lambda(-1)$ and the turning point of the map as λ increases. (Both can increase with λ , but, if the turning point moves upwards too quickly, we do not see all the explosions observed for the actual Lorenz system.) Notice that all the homoclinic explosions we have just discussed are proceeding in the direction which produces a strange invariant set (Conjecture 3, and hence Conjecture 4, of Chapter 5).

Different families of maps, like those in Fig. J1, will produce different non-destructive Lorenz lists of explosions. It is not clear which

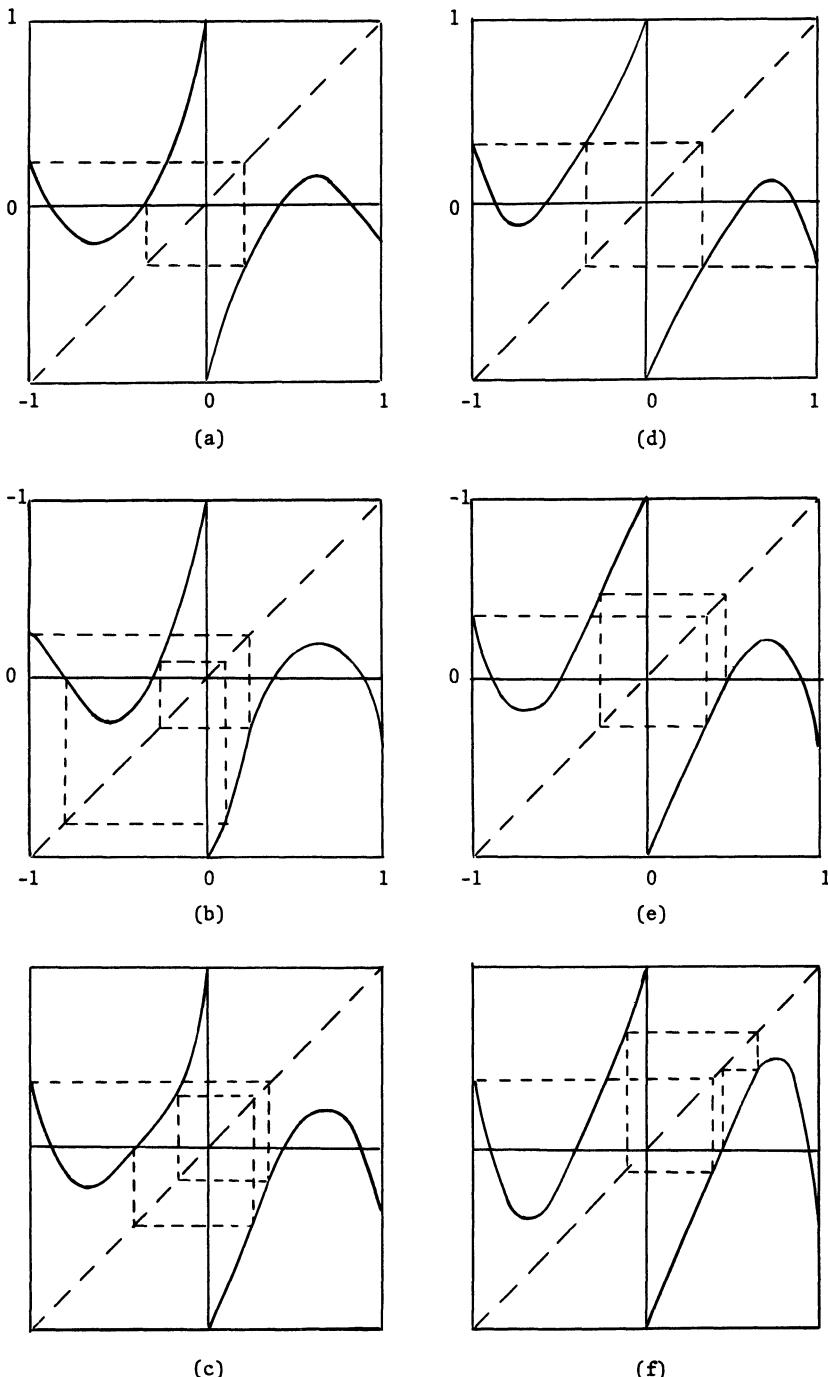


Figure J2. Sketches of a sequence of homoclinic situations as λ increases near λ_3 . The homoclinic situations are type (a) if an even number of iterates of $0+$ lie on sections of the map with negative slope and type (b) otherwise. (a) xyxy-homoclinic; type (b); (b) λ yxyxy-homoclinic; type (a); (c) yxyxy-homoclinic; type (b); (d) $\lambda = \lambda_3$; $k^3 = xyxyxy\dots\dots$; (e) yxyxy-homoclinic; type (b); (f) xyxyxx-homoclinic; type (b).

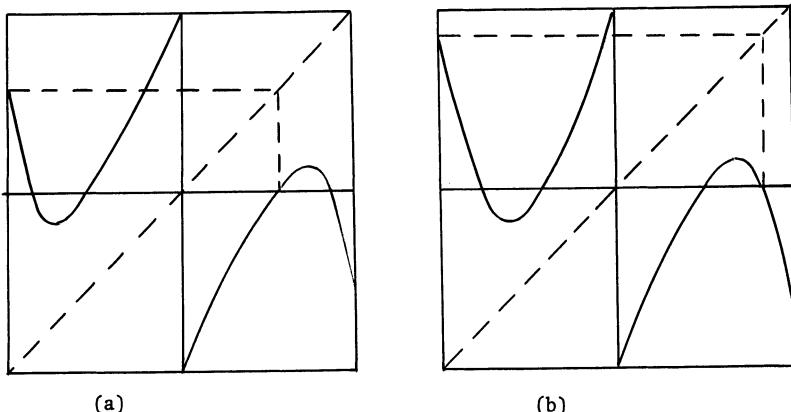


Figure J3. Sketches of xyx -homoclinic situations as λ increases in $\lambda_3 < \lambda < \lambda_4$. (a) type (b); (b) type (a).

combinatorially possible non-destructive Lorenz lists will be realizable by one-dimensional maps, though this will clearly depend to some extent on the precise restrictions which one chooses to put on the maps. Notice, though, that we expect one-dimensional maps to realize Lorenz lists which include "extra" homoclinic explosions (as above) and so we know that the "extra" explosions in Appendix I are not completely a product of the three-dimensional nature of the Lorenz flow. A complete kneading sequence theory for maps like Fig. J1 would be quite complicated and would require study of an additional symbolic sequence describing the behaviour of successive iterates of one of the turning points of the map (thereby keeping track of the period doubling windows as well as the homoclinic behaviour). It is not clear that a complete kneading sequence theory for these maps would tell us anything about the Lorenz equations that we do not already know; most of the results that depended on knowledge of the behaviour of successive iterates of the turning points would not carry over in a rigorous way to the Lorenz equations (see Section 5.5).

Before leaving one-dimensional maps, it is worth mentioning two special cases.

- 1) We could prove the remarks we made about the basic Lorenz lists in Section 1 of this appendix. If we studied a family of one-dimensional maps which never develops a turning point, but for which $f_\lambda(-1) > 0$ for large enough λ , all homoclinic situations will be type (a) and the Lorenz list that the maps realize will be the basic Lorenz list. See also Parry

(1977) and Appendix G. By drawing the maps f^2 for this family, it is easy to see that the symmetric xy orbit must become stable by casting off two non-symmetric non-stable xy orbits at some λ value before $f_\lambda(-1)$ crosses zero. When $f_\lambda(-1) = 0$, we have a type (a) xy -homoclinic explosion which removes these two orbits (along with the rest of an xy -generated strange invariant set) and, with appropriate conditions on the maps, we are left with a stable symmetric xy orbit only.

2) We could use one-dimensional maps to study the special homoclinic situation which is neither type (a) nor type (b). Though we could not be sure that our results would necessarily apply directly to the full three-dimensional flow, they would give us some idea of what was going on. There are several different possibilities in the one-dimensional map, depending on the shape of the turning point and on the behaviour near the discontinuity. The most relevant combination of conditions is probably a quadratic turning point (non-zero quadratic coefficient in the Taylor expansion about the turning point), and behaviour near the discontinuity of the form $f(\epsilon) = -1 + \epsilon^\alpha$ where $0 < \alpha < \frac{1}{2}$. The analysis of this case is non-trivial; it is clear, however, that the results of the analysis will not invalidate any of the combinatorial arguments which rested on the assumption that all homoclinic situations were type (a) or type (b). Since the analysis involves examination of terms of higher order than those needed for the analysis of type (a) and (b) explosions, when applying any results to the Lorenz system it will be necessary to consider the possible contribution of terms in the other variables (which we ignored in Appendix D because they were small compared with the first order terms of the most important variable) to the behaviour at the homoclinic parameter values.

Appendix K

Large r ; The Formulae

I am indebted to Prof. Swinnerton-Dyer for his most valuable assistance in producing this analysis. A justification for the method of averaging used in this appendix can be found in his paper, Swinnerton-Dyer (1980). I am also grateful to Prof. Fowler for some comments on an earlier draft of Chapter 7 and this appendix. Fowler & McGuinness (1982) have recently extended the analysis here to cover a complex version of the Lorenz equations, at the same time making interesting connections with some of their previous work on the real Lorenz equations in a limit as σ and r both become large (Fowler & McGuinness, 1981a,b). This earlier work, and the connection with our present analysis, is discussed in Chapter 8.

We have (Chapter 7) the Lorenz equations in the form

$$\begin{aligned}\xi' &= \eta - \epsilon\sigma\xi \\ \eta' &= -\xi z - \epsilon\eta \\ z' &= \xi\eta - \epsilon b(z+\sigma)\end{aligned}\tag{1}$$

where ϵ is small and ξ , η and z are $O(1)$. In the limit, as $\epsilon \rightarrow 0$, we have the approximate equations

$$\xi' = \eta; \quad \eta' = -\xi z; \quad z' = \xi\eta.\tag{2}$$

Equations (2) have the obvious integrals

$$\xi^2 - 2z = 2A, \quad \eta^2 + z^2 = B^2.\tag{3}$$

We choose $B > 0$ and consider regions $A > B > 0$ and $|A| < B$. To complete the integration of (2) we use (3) to obtain

$$z' = \{2(B-z)(z+B)(z+A)\}^{1/2}.\tag{4}$$

This allows us to express z as an elliptic integral.

We use the standard notations of elliptic function theory; the letters k , E and K will have their canonical meanings. The formulae we need are in Byrd & Friedman (1954).

In $A > B > 0$ we write (Byrd & Friedman, §236)

$$z = B(1 - 2\operatorname{sn}^2(u)) \quad \text{where } k^2 = \frac{2B}{(A+B)} < 1. \quad (5)$$

From (3) we now have

$$\xi = \pm\{2(A+B)\}^{\frac{1}{2}}\operatorname{cn}(u), \quad \eta = \mp 2B\operatorname{sn}(u)\operatorname{cn}(u) \quad (6)$$

where the upper signs are for $\xi > 0$ and the lower signs are for $\xi < 0$. Substituting in (2) gives

$$u' = \{ \frac{1}{2}(A+B) \}^{\frac{1}{2}}. \quad (7)$$

Going once around an orbit corresponds to increasing u by $2K$.

In $|A| < B$ we have, using a different transformation,

$$z = B - (A+B) \operatorname{sn}^2(u) \quad \text{where } k^2 = \frac{(A+B)}{2B} < 1. \quad (8)$$

From (3) we obtain, in this case,

$$\xi = \{2(A+B)\}^{\frac{1}{2}}\operatorname{cn}(u), \quad \eta = -2kB \operatorname{dn}(u)\operatorname{cn}(u) \quad (9)$$

where the choice of sign can be made because we can change the sign of u and increase it by $2K$. Substituting in (2) we have

$$u' = B^{\frac{1}{2}} \quad (10)$$

and going once around an orbit corresponds to increasing u by $4K$.

Now, for the Lorenz equations, (1), define new variables A and B from (3). We obtain

$$\begin{aligned} A' &= \epsilon(-\sigma\xi^2 + bz + b\sigma) \\ BB' &= -\epsilon(\eta^2 + bz^2 + b\sigma z). \end{aligned}$$

A and B are slowly varying. The change in A , for example, in the time interval $(0, T)$ is

$$\Delta A = \int_{u(0)}^{u(T)} \frac{A'}{u'} du$$

where u is defined by (5) or (8) depending on the region of the A, B plane which interests us. For instance, if we are interested in $A > B > 0$,

we define u from (5), substitute for u' from (7) (which holds with error $O(\epsilon)$) and for ξ , η and z from (6) (which remains exactly true). We then integrate over the interval $(0, 4K)$ to obtain the "averaged" equation; this gives the average rate of change of A . Performing this operation for both A and B and in both regions (in $|A| < B$ we substitute from (8), (9), and (10)) we obtain:

(i) In $A > B > 0$

$$\begin{aligned} KA' &= \epsilon\{b\sigma K - 4B\sigma k^{-2}E - bBk^{-2}((2 - k^2)K - 2E)\} \\ 3Kk^4B' &= -\epsilon\{4B[(2 - k^2)E - 2(1 - k^2)K] + bB[4(k^2 - 2)E + (3k^4 - 8k^2 + 8)K] \\ &\quad + 3b\sigma k^2[2E - K(2 - k^2)]\} \end{aligned}$$

(ii) In $|A| < B$

$$\begin{aligned} KA' &= \epsilon\{Kb\sigma + bB(2E - K) - 4B\sigma(E - (1 - k^2)K)\} \\ 3KB' &= -\epsilon\{4B[K(1 - k^2) + E(2k^2 - 1)] + 3b\sigma(2E - K) \\ &\quad + bB[K(4k^2 - 1) + 4E(1 - 2k^2)]\}. \end{aligned}$$

Figs. 7.5 were calculated by integrating these equations numerically. k is determined from the values of A and B (and depends on which region you are in) and K and E were determined in various different ways (Gauss transformation, series expansion, or differential equations satisfied by K and E) depending on the values of A and B . There are always problems on the boundaries of the two regions in the A, B plane (where $k = 0$ or 1) and the Figs. 7.5 are only crude approximations to the truth.

We look for stationary points of the averaged equations. In $A > B > 0$ notice that

$$(2 - k^2)K - 2E = \frac{1}{4} \int_0^{4K} \text{sn}^2(u) \text{cd}^2(u) du > 0.$$

Hence, the equation $A' = 0$ gives a unique positive B for each value of k . Substituting this into the equation $B' = 0$ we obtain an equation for k which we can write as

$$2\lambda - 1 = \frac{K\{(2 - k^2)E - 2(1 - k^2)K\}}{3E\{(2 - k^2)K - 2E\}} \quad (11)$$

where $\lambda = (\sigma+1)/(b+2)$. The right-hand side decreases monotonically from 1 to $\frac{1}{3}$ as k increases from 0 to 1. Thus there is a unique k satisfying (11) if and only if $\frac{2}{3} < \lambda < 1$.

In $|A| < B$, notice that the two terms in square brackets on the right-hand side of the equation for B' are strictly positive. Thus, there will only be a positive B satisfying $B' = 0$ if

$$K > 2E. \quad (12)$$

Elimination of B between the equations $A' = 0$ and $B' = 0$ gives

$$2\lambda - 1 = \frac{K\{K(1 - k^2) + E(2k^2 - 1)\}}{3(K - 2E)\{E - (1 - k^2)K\}}. \quad (13)$$

As k increases from the solution of (12) (≈ 0.82) to 1, the right-hand side of (13) decreases monotonically from ∞ to $\frac{1}{3}$. Hence, there is a singular point if $\lambda > \frac{2}{3}$ but not otherwise.

On the line $B = 0$ we have $k = 0$, and Byrd & Friedman (§900) give expressions for K and E near $k = 0$. Retaining only the dominant terms we have

$$A' \sim \varepsilon\sigma(b - 2A)$$

$$B' \sim \frac{1}{4}\varepsilon B(b\sigma^{-1} - 2b - 2)$$

near the line $B = 0$, and hence a singular point at $A = \frac{1}{2}b$, $B = 0$. Trajectories moving near the line $B = 0$ will move towards $B = 0$ if $A > \frac{b\sigma}{2(b+1)}$. Hence the stationary point at $(\frac{1}{2}b, 0)$ will be stable if $\lambda < 1$ and non-stable if $\lambda > 1$.

On the line $A = -B$ we have $k = 0$ again (though this is the boundary of the other region so the equations are different). Here, the dominant terms do not give us sufficient information and we must also form a differential equation for k^2 . We obtain

$$B' \sim -\varepsilon b(B + \sigma)$$

$$(k^2)' \sim \varepsilon k^2(\frac{3}{2}b - \sigma).$$

Hence, the line $A = -B$ is a trajectory with B decreasing. Trajectories moving near the line $A = -B$ will move towards it (k decreases) providing $\sigma > \frac{3}{2}b$.

On the line $A = B$ (the boundary between the two regions) we have $k = 1$. Considering the line to be the boundary of $A > B > 0$ (similar results are obtained if we consider it to be the boundary of $|A| < B$), we have $A' \sim B' \sim \varepsilon b(\sigma - B)$. This tells us that we have a stationary point at (σ, σ) towards which trajectories on $A = B$ tend, but does not tell us how trajectories move near $A = B$. Forming the differential equation for k^2 again, we obtain

$$(k^2) \sim \varepsilon K^{-1} (2\sigma - \frac{2}{3} - \frac{b}{3} - b\sigma B^{-1}).$$

Equivalently, (Fowler, private communication) we can write $\phi = A - B > 0$ and obtain

$$(-\ln(\phi))\phi' \sim C = 4\varepsilon(b\sigma + \frac{1}{3}B[b + 2 - 6\sigma])$$

or

$$[\ln \frac{1}{\phi} + 1]\phi \sim C\tau + c_0 \text{ where } c_0 \text{ is a constant.}$$

Hence, trajectories reach the line $A = B$ in finite (positive or negative) time. Trajectories move towards $A = B$ (from $A > B > 0$) if $6\sigma > b+2$ and $B > \frac{3b\sigma}{(6\sigma - b - 2)}$.

Finally, we consider the flow in $|A| < B$ in order to determine the stability of the stationary point there. Referring to the original equations (1), we have the divergence formula

$$\frac{d}{d\tau} \log \iiint d\xi d\eta dz = -\varepsilon(b + \sigma + 1)$$

where elements of volume are to be regarded as being carried with the flow. Making the transformations (8) and (9) - which are the appropriate ones for the region $|A| < B$ - we find the Jacobian

$$J = \frac{\partial(\xi, \eta, z)}{(A, B, u)} = B^{\frac{1}{2}}$$

and hence

$$\frac{d}{d\tau} \log \iiint B^{\frac{1}{2}} dA dB du = -\varepsilon(b + \sigma + 1).$$

Here, the process of averaging over u and replacing the true equations for A and B by the averaged equations can only produce errors of $O(\varepsilon^2)$ which are not important. We deduce

$$\frac{d}{d\tau} \log \iint B^{\frac{1}{2}} K dA dB = -\varepsilon(b + \sigma + 1)$$

which now refers to the flow of the averaged equations and implies that this flow is contracting. Thus, there are no periodic orbits contained entirely within $|A| < B$ and any stationary points there are stable.

These are the results we used to draw Figs. 7.2.

Bibliography

- Alligood, K. G., Mallet-Paret, J. and Yorke, J. A. (1981). "Families of periodic orbits: local continuability does not imply global continuability". *J. Diff. Geom.*, in press.
- Arnéodo, A., Coullet, P. and Tresser, C. (1981a). "A possible mechanism for the onset of turbulence". *Phys. L.*, 81A (4), 197-201.
- Arnéodo, A., Coullet, P. and Tresser, C. (1981b). "Possible new strange attractors with spiral structure". *Commun. Math. Phys.*, 79, 573-579. Also other papers by the same authors.
- Birman, J. S. and Williams, R. F. (1979). "Knotted periodic orbits in dynamical systems I: Lorenz's equations". Preprint, Northwestern University.
- Brindley, J. and Moroz, I. M. (1980). "Lorenz attractor behaviour in a continuously stratified baroclinic fluid". *Phys. L.*, 77A (6), 441-444.
- Byrd, P. F. and Friedman, M. D. (1954). "Handbook of elliptic integrals for engineers and physicists". Springer-Verlag.
- Collet, P. and Eckmann, J.-P. (1980). "Iterated maps on the interval as dynamical systems". Birkhäuser, Boston.
- Collet, P., Eckmann, J.-P., and Koch, H. (1980). "Period doubling bifurcations for families of maps on \mathbb{R}^n ". Preprint, Université de Genève.
- Crutchfield, J., Farmer, D., Packard, N., Shaw, R., Jones, G. and Donnelly, R. J. (1980). "Power spectral analysis of a dynamical system". *Phys. L.*, 76A, 1-4.
- Curry, J. H. (1978). "A generalised Lorenz system". *Commun. Math. Phys.*, 60, 193-204.
- Curry, J. H. (1979a). "An algorithm for finding closed orbits". In "Global Theory of Dynamical Systems". Edited by Z. Nitecki and C. Robinson. *Lect. Notes in Math.*, 819, 111-120, Springer-Verlag.

- Curry, J. H. (1979b). "Chaotic response to periodic modulation of model in a vonvecting fluid". *Phys. Rev. L.*, 43, 1013.
- Curry, J. H. (1979c). "Bounded solutions of finite dimensional approximations to the Boussinesq equations". *SIAM J. Math.*, 10, 71-77.
- Da Costa, L. N., Knobloch, E. and Weiss, N. O. (1981). "Oscillations in double-diffusive convection". *J. Fluid Mech.*, 109, 25-43.
- Denjoy, A. (1932). "Sur les courbes définies par les équations différentielles à la surface du tore". *J. de. Math.*, Ser. 9, 11, 333-375.
- Feigenbaum, M. J. (1978). "Quantitative universality for a class of nonlinear transformations". *J. Stat. Phys.*, 19, 25-52.
- Fowler, A. C., Gibbon J. D. and McGuinness, M. J. (1981). "The complex Lorenz equations". *Physics. D*, in press.
- Fowler, A. C. and McGuinness, J. M. (1981a). "A description of the Lorenz attractor at high Prandtl number". *Physica D*, to appear.
- Fowler, A. C. and McGuinness, M. J. (1981b). "Hysteresis, period doubling, and intermittence at high Prandtl number in the Lorenz equations". Preprint, Department of Mathematics, M.I.T.
- Fowler, A. C. and McGuinness, M. J. (1982). "On the nature of the torus in the complex Lorenz equations". Preprint, Department of Mathematics, M.I.T.
- Franceschini, V. and Boldrighini, C. (1979). "A five-dimensional truncation of the plane incompressible Navier-Stokes equations". *Commun. Math. Phys.*, 64, 159-170.
- Franceschini, V. and Tebaldi, C. (1979). "Sequences of infinite bifurcations and turbulence in a 5-mode truncation of the Navier-Stokes equations". *J. Stat. Phys.*, 21, 707-726.
- Franceschini, V. and Tebaldi, C. (1980). "A seven mode truncation of the Navier-Stokes equations". Preprint, University di Modena, Italy.
- Franceschini. V. (1980). "A Feigenbaum sequence of bifurcations in the Lorenz model". *J. Stat. Phys.*, 22 (3), 397-406.
- Gibbon, J. D. (1981). In "Synergetics", Springer-Verlag. Edited by H. Haken.
- Gibbon, J. D. and McGuinness, M. J. (1980). "A derivation of the Lorenz equations for some unstable dispersive physical systems". *Phys. L.*, 77A, 295-299.
- Gibbon, J. D. and McGuinness, M. J. (1981). "The real and complex Lorenz equations in rotating fluids and lasers". *Physica D*, in press.
- Graham, R. and Scholz, H. J. (1980). "Analytic approximation of the Lorenz attractor by invariant manifolds". *Phys. Rev.*, 22A, 1198-1204.
- Guckenheimer, J. (1976). "A strange, strange attractor". In Marsden and McCracken, 1976.

- Guckenheimer, J. (1979a). "The bifurcations of quadratic functions". In *Bifurcation Theory and its Applications in Scientific Disciplines*. Edited by O. Gurel and O. E. Rössler. *Annal. N. Y. Acad. Sci.*, 316, 78-85.
- Guckenheimer, J. (1979b). "Sensitive dependence on initial conditions for one-dimensional maps". *Commun. Math. Phys.*, 70, 133-160.
- Guckenheimer, J. (1980). "One-dimensional dynamics". In *Nonlinear Dynamics*. Edited by R. H. G. Helleman. *Annal. N. Y. Acad. Sci.*, 357, 343-347.
- Guckenheimer, G., Moser, J. and Newhouse, S. E. (1980). "Dynamical Systems". *Progress in Mathematics* 8. *Birkhäuser*, Boston.
- Guckheimer, J. and Williams, R. F. (1980). "Structural stability of the Lorenz attractor". *Publications Mathématiques*, I.H.E.S., 50, 73-100.
- Haken, H. (1975). "Analogy between higher instabilities in fluids and lasers". *Phys. L.*, 53A, 77-78.
- Hassard, B. D. (1980). "Computation of invariant manifolds". In *New Approaches to Non-Linear Problems in Dynamics*. Edited by P. J. Holmes. *SIAM*, Philadelphia.
- Hassard, B. D., Kazarinoff, N. D. and Wan, Y.-H. (1981). "Theory and application of Hopf bifurcation". *Lon. Math. Soc. Lecture Notes* 41.
- Helleman, R. H. G. (1980). "Self-generated chaotic behaviour in nonlinear mechanics". In *Fundamental Problems in Statistical Mechanics*, Vol. 5, 165-233. Edited by E. G. D. Cohen. *North Holland Publ.*, Amsterdam and New York.
- Hénon, M. and Pomeau, Y. (1976). "Two strange attractors with a simple structure". In *Turbulence and Navier-Stokes Equations*. Edited by R. Temam. *Lect. Notes in Math.*, 565, 29-68, Springer-Verlag.
- Ibañez, J. L. and Pomeau, Y. (1978). "A simple case of non-periodic (strange) attractor". *J. Non. Equib. Thermodyn.*, 3, 135-152.
- Kaplan, J. L. and Yorke, J. A. (1979a). "Preturbulence: A regime observed in a fluid flow model of Lorenz". *Commun. Math. Phys.*, 67, 93-108.
- Kaplan, J. L. and Yorke, J. A. (1979b). "The onset of chaos in a fluid flow model of Lorenz". In *Bifurcation Theory and Applications in Scientific Disciplines*. Edited by O. Gurel and O. E. Rössler. *Annal. N. Y. Acad. Sci.*, 316, 400-407.
- Knobloch, E. (1979). "On the statistical dynamics of the Lorenz model". *J. Stat. Phys.*, 20, 695-709.
- Knobloch, E. (1981). "Chaos in a segmented disc dynamo". *Phys. L.*, 82A (1), 439-440.
- Knobloch, E., Weiss, N. O. and Da Costa, L. N. (1981). "Oscillatory and steady convection in a magnetic field". *J. Fluid Mech.*, 113, 153.
- Knobloch, E. and Weiss, N. O. (1981). "Bifurcations in a model of double-diffusive convection". *Phys. L.*, 85A, 127-130.

- Knobloch, E. and Weiss, N. O. (1982). "Bifurcations in a model of magnetoconvection". Preprint, Physics Department, U. C. Berkeley.
- Lanford, O. E. (III) (1976). "Qualitative and statistical theory of dissipative systems". Preprint, University of California at Berkeley.
- Lichtenberg, A. and Lieberman, M. (1982). "Regular and Stochastic Motion". Appl. Math. Sci., 38, Springer-Verlag.
- Lorenz, E. N. (1963). "Deterministic non-periodic flows". J. Atmos. Sci., 20, 130-141.
- Lorenz, E. N. (1979). "On the prevalence of aperiodicity in simple systems". In Global Analysis. Edited by M. Grmela and J. E. Marsden. Lect. Notes in Math., 755, 53-75. Springer-Verlag.
- Lorenz, E. N. (1980a). "Noisy periodicity and reverse bifurcation". In Nonlinear Dynamics. Edited by R. H. G. Helleman. Annal. N. Y. Acad. Sci., 357, 282-291.
- Lorenz, E. N. (1980b). "Attractor sets and quasi-geostrophic equilibrium". J. Atmos. Sci., 36, 1685-1699.
- Malkus, W. V. R. (1972). "Non-periodic convection at high and low Prandtl number". Mémoires Société Royale des Sciences de Liège, Series, 6, Vol. IV, 125-128.
- Manneville, P. and Pomeau, Y. (1980). "Different ways to turbulence in dissipative dynamical systems". Physica, 1 D, 219-226.
- Manneville, P. and Pomeau, Y. (1979). "Intermittency and the Lorenz model". Phys. L., 75A, 1.
- Marsden, J. (1977). "Attempts to relate the Navier-Stokes equations to turbulence". In Turbulence Seminar. Edited by P. Bernard. Lect. Notes in Math., 615, 1-23. Springer-Verlag.
- Marsden, J. and McCracken, M. (1976). "The Hopf bifurcation and its applications". Appl. Math. Sci., 19, Springer-Verlag.
- May, R. M. (1976). "Simple Mathematical models with very complicated dynamics". In Bifurcation Theory and Its Applications in Scientific Disciplines. Edited by O. Gurel and O. E. Rössler. Annal. N. Y. Acad. Sci., 316, 517-529.
- McLaughlin, J. B. and Martin, P. C. (1975). "Transitions to turbulence in a statistically stressed fluid system". Phys. Rev., A12, 186-203.
- Parry, W. (1976). "Symbolic dynamics and transformations of the unit interval". Trans. Amer. Math. Soc., 222, 368-378.
- Parry, W. (1977). "The Lorenz attractor and a related population model". In Ergodic Theory Conf. Oberwolfach. Edited by M. Denker & K. Jacobs. Lect. Notes in Math., 729, 169-187. Springer-Verlag (1979).
- Pedlosky, J. (1972). "Limit cycles and unstable baroclinic waves". J. Atmos. Sci., 29, 53.
- Pedlosky, J. and Frenzen, C. (1980). "Chaotic and periodic behaviour of finite amplitude baroclinic waves". J. Atmos. Sci., 37, 1177-1196.

- Rand, D. (1978). "The topological classification of Lorenz attractors". *Math. Proc. Camb. Phil. Soc.*, 83, 451-460.
- Robbins, K. (1979). "Periodic solutions and bifurcation structure at high r in the Lorenz system". *SIAM J. Appl. Math.*, 36, 457-472.
- Ruelle, D. and Takens, F. (1971). "On the nature of turbulence". *Commun. Math. Physics*, 20, 167-192.
- Ruelle, D. (1976). "The Lorenz attractor and the problem of turbulence". In *Turbulence and Navier-Stokes equations*. Edited by R. Temam. *Lect. Notes in Math.*, 565, 146-155. Springer-Verlag.
- Ruelle, D. (1977). "Dynamical systems with turbulent behaviour". *Proc. Int. Math. Physics Conf.*, Rome 1977. Edited by G. Dell'Antonio, S. Doplicher and G. Jona-Lasinio. *Lect. Notes in Physics*, 80, 341-360. Springer-Verlag.
- Ruelle, D. (1979). "Sensitive dependence on initial conditions and turbulent behaviour in dynamical systems". In *Bifurcation Theory and its Applications in Scientific Disciplines*. Edited by O. Gurel and O. E. Rossler. *Annal. N. Y. Acad. Sci.*, 316, 408-416.
- Segur, H. (1980). "Solitons and inverse scattering transform". Preprint. Aeronautical Research Associates of Princeton Inc.
- Shaw, R. (1978). "Strange attractors, chaotic behaviour and information flow". Ph.D. Thesis. Preprint. University of California at Santa Cruz.
- Shimada, I. and Nagashima, T. (1977). "On the C-like property of the Lorenz system". *Prog. Theor. Phys.*, 58, 1318-1319.
- Shimada, I. and Nagashima, T. (1978). "The iterative transition phenomenon between periodic and turbulent states in a dissipative dynamical system". *Prog. Theor. Phys.*, 59, 1033.
- Shimada, I. and Nagashima, T. (1979). "A numerical approach to ergodic problem of dissipative dynamical systems". *Prog. Theor. Phys.*, 61, 1605-1616.
- Shimada, I. (1979). "Gibbsian distribution on the Lorenz attractor". *Prog. Theor. Phys.*, 62, 61-69.
- Shimizu, T. and Morioka, N. (1978a). "Chaos and limit cycles in the Lorenz equations". *Phys. L.*, 66A, 182.
- Shimizu, T. and Morioka, N. (1978b). "Transient behaviour in periodic regions of the Lorenz model". *Phys. L.*, 66A, 447.
- Shimizu, T. and Morioka, N. (1978c). "Transition between turbulent and periodic states in the Lorenz model". *Phys. L.*, 69A, 148.
- Shimizu, T. (1979). "Analytic form of the simplest limit cycle in the Lorenz model". *Physica*, 97A, 383-398.
- Sil'nikov, L. P. (1965). "A case of the existence of a denumerable set of periodic motions". *Soviet. Math. Dokl.*, 6, 163-166.

- Sil'nikov, L. P. (1968). "On the generation of a periodic motion from trajectories doubly asymptotic to an equilibrium state of saddle type". *Math. USSR. Sbornik*, 6, 428-438.
- Sil'nikov, L. P. (1970). "A contribution to the problem of the structure of an extended neighborhood of a rough equilibrium state of saddle-focus type". *Math. USSR, Sbornik*, 10, 91-102.
- Smale, S. (1967). "Differentiable dynamical systems". *Bull. Amer. Math. Soc.*, 73, 747-817.
- Sparrow, C. and Swinnerton-Dyer, H. P. F. (1982). "The Lorenz equations for large r'. To appear.
- Swinnerton-Dyer, H. P. F. (1980). "The method of averaging for some almost conservative differential equations". *J. Lon. Math. Soc.*, 22 (3), 534-542.
- Van Gib, S. (1979). "Hopf bifurcation and attractivity". *Stichting Mathematisch Centrum, Amsterdam, TN93*.
- Williams, R. F. (1977). "The structure of Lorenz attractors". In *Turbulence Seminar*. Edited by P. Bernard. *Lect. Notes in Math.*, 615, 93-113. Springer-Verlag.
- Williams, R. F. (1979). "The bifurcation space of the Lorenz attractor". In *Bifurcation Theory and Its Applications in Scientific Disciplines*. Edited by O. Gurel and O. E. Rössler. *Annal. N. Y. Acad. Sci.*, 316, 393-399.
- Williams, R. F. (1980). "Structure of Lorenz attractors". *Publications Mathématiques, I.H.E.S.*, 50, 59-72.
- Yorke, E. D. and Yorke, J. A. (1979). "Metastable chaos: transition to sustained chaotic behaviour in the Lorenz model". *J. Stat. Phys.*, 21 (3), 263-277.

Index

- almost periodic behaviour, 63, 64
- attractor, 5, 9, 66, 74, 103, 185, 234
- averaging, 132-150, 259-261
- baroclinic instability, 188
- basin of attraction, 5, 48
- bifurcation, 6, 7, 9, 16, 20, 22, 23, 39, 40, 42, 50-75, 91-93, 99, 106, 107, 115, 118, 139, 149, 154, 161-172, 179, 183-191, 208, 210, 218-220
- bounded system, 5
- branched manifold, 44-46, 229, 230
- Cantor set, 21, 36, 37, 205, 231
- chaotic attractor, 5, 174, 176
- chaotic behaviour, 4-7, 31, 32, 36, 58, 62-75, 83, 84, 110, 110, 155, 159, 160, 168, 174, 177, 178, 190, 208, 215, 217
- characteristic exponent, 186
- combinatorics, 39, 50, 104, 171-174, 244
- contracting foliation, 43, 44, 49, 102, 208, 230
- contraction, 9, 18, 20, 35, 36, 43, 49, 52, 77-81, 102, 198, 205, 220, 231, 261
- cyclic permutation, 23, 39
- deterministic systems, 3
- dissipative systems, 4, 5, 9, 100, 167, 196-198
- divergence 9, 198
- fluid convection, 1-3
- free group, 125
- Hamiltonian systems, 5
- Hénon map, 70, 75
- heteroclinic bifurcation, 167-172, 176, 183
- heteroclinic orbit, 127, 164-172, 188
- homoclinic explosion, 12, 16, 22-24, 26, 28, 32, 37-42, 44, 50, 86, 90-94, 100, 103-116, 121, 131, 139, 140, 149, 153-166, 170-172, 176, 179, 182-185, 190, 191, 199-210, 233, 239-256
- homoclinic explosion explosion, 114, 250, 251
- homoclinic orbit, 12, 15, 16, 20, 21, 23, 37, 39, 40, 42, 50, 90, 91, 115, 126, 133, 134, 155, 156, 170, 172, 191, 199-210, 228
- homotopy group, 125
- hooked return maps, 76-92, 100-102, 109, 110, 114, 236, 247
- Hopf bifurcation, 11, 12, 25-28, 32, 38, 42, 49, 107, 113, 137-140, 150, 153, 158, 172, 179, 186, 188, 190, 235, 236
- horseshoe, 109, 111, 209
- intermittent chaos, 62-69, 74, 174
- kneading invariants, 226-229
- kneading sequences, 101, 224-256
- laminar behaviour, 62-66
- Liapunov functions, 9, 196-198
- linearization, 9, 11, 12, 16, 17, 20, 35, 44, 47, 119, 151, 153, 199, 200, 211, 212, 221, 236, 239
- locally eventually onto, definition, 224

Lorenz, E. N., 1, 3, 8, 9, 32, 66, 70, 194
Lorenz list
definition, 245
basic, 245
minimal, 245
non-destructive, 246

maxima-in-z, 45, 86, 101, 115-119, 130, 131, 173, 175, 234-238
meta-stable chaos, 31
meteorology, 1
models, 1, 4, 188, 189, 194, 195
monstrous formula, 221, 222

noise, 3, 186
noisy periodicity, 69-73
non-destructive sequence,
definition, 106
non-symmetric definition, 8
non-stable, definition, 193
non-wandering set, 5, 6, 21, 24, 49, 64, 66, 81-83, 88, 91, 103, 104, 110, 113, 127, 130, 132, 177, 193, 198, 234

one-dimensional interval maps, 56, 58, 64, 65, 69, 70, 75, 76, 100-102, 173-176, 201-204, 208, 223-238, 251-256
optics, 188
original, definition, 24

periodic orbits, 3, 6, 8, 12, 21-28, 36-44, 47, 50-79, 82-100, 103, 106-118, 126-190, 207, 209, 212-220, 234-256, 261
periodic orbits without extra twisting, definition, 155
Poiseuille flow, 189
Prandtl number, 2
preturbulence, 31, 32, 35, 48, 235, 236

Rayleigh number, 2
recurrent behaviour, 5
return map, definition, 17
rotation, direction of, 8, 154

saddle-node bifurcation, 52-54, 58, 59, 64-66, 74, 91, 92, 100, 103, 149, 164, 173, 174, 182, 187, 218, 219, 247, 248
saddles, 9
semi-flow, 44-46, 229, 230, 233
semi-periodicity, definition, 70
sensitive dependence on initial conditions, 22, 36, 208, 216, 241
sequences of symbols, definition, 37
sinks, 9
sources, 9
stable, definition, 193
stable manifold, definitions, 193
stationary points, 9
stiff equations, 5
strange attractor, 5, 24, 31-49, 70, 74, 76, 103, 113, 190, 208, 209, 223, 231-233
strange invariant set, 21-42, 48, 50, 90, 92, 103, 108, 113, 153, 156, 162, 208-210, 236, 239-241, 247, 248
structural stability, 233
subcritical Hopf bifurcation, 11, 32
supercritical Hopf bifurcation, 11, 186, 188
symmetric definition, 8
symmetry, 8, 16, 20-23, 32, 42, 52-74, 100, 107, 109, 119, 133, 174, 187, 188, 190, 201, 210, 235, 252

T-homoclinic explosion, definition, 23
T-homoclinic orbit, definition, 22
T-generated, definition, 23
thermosolutal convection, 188
torus, invariant, 9, 55, 187, 188
transience, 3
truncation, 3, 186
turbulence, 3, 4, 189, 208
type (a), definition, 20
type (b), definition, 20

unstable, definition, 193
unstable manifold, definition, 193

weather, 1