Convergence Properties of

Noise Reduction by

Gradient Descent

David Ridout, BSc (Hons)

This thesis is presented for the Degree of Master of Science in Mathematics at

The University of Western Australia.

Department of Mathematics

March, 2001

Abstract

In this thesis, the gradient descent algorithm for noise reduction is investigated, particularly with respect to its convergence properties as the length of the trajectory required to be "noise reduced" tends to infinity. This investigation begins by considering the results of applying the gradient descent algorithm to noisy trajectories from some simple dynamical maps. These experiments suggest the type of convergence result that might hold, as well as indicating the type of dynamical system for which noise reduction might be expected to work. The aim of the rest of the thesis is to theoretically justify the conclusions of these experiments. In particular, the main focus is to rigorously prove that for a certain class of dynamical systems, the gradient descent algorithm will converge (as the length of the noisy trajectory given tends to infinity) onto the correct "clean" trajectory everywhere, except near the initial and final points. That is, that the errors between the "noise reduced" trajectory points and the original "clean" trajectory points can be made arbitrarily small by taking the length of these trajectories long enough, except near the initial and final points of the trajectories. The proof is based on relating the non-linear gradient descent algorithm to a linearised version. The convergence result for the noise reduction is shown for the linearised gradient descent by deriving explicitly computable analytic bounds for the errors at each point. These bounds are then shown to generalise to the non-linear gradient descent, provided that a certain condition is met. The question of which dynamical systems satisfy this condition remains open, but it is conjectured that it is satisfied, in fact, for any non-linear dynamical system of the class considered.

Acknowledgements

I would like to thank my supervisors Kevin Judd, Gary Froyland and Alistair Mees for all the time, effort, encouragement, and patience that they have invested in me during the course of this degree. Thanks to Kevin for not only providing the motivation for this study, but also for taking the time to follow all the twists and turns as I stumbled my way through this last year, and for frequently putting up with incoherent explanations of misunderstood results that often turned out to be irrelevant or just plain wrong! I would like to thank Gary for the invaluable technical assistance he has provided, and likewise, for his ability to tolerate weekly occurences of confident claims being retracted shortly thereafter. A big thankyou to Alistair for just generally being a great guy and making this whole experience a lot less scary than it could have been.

I should also like to take this opportunity to thank all the people whose friendship I have neglected over the last few months. I hope that if any of them ever read these acknowledgements, then they will recall that my plan to shortly correct this neglect, did indeed reach fulfillment.

Lastly, I need to acknowledge my partner, Penny, who has put up with all sorts of irritability, insomnia, and incoherent ravings and complaints about things she neither knows nor cares about, and for all this tolerance whilst she is trying to write up as well. Thankyou!

Contents

Chapter 1

Introduction

1.1 Noise Reduction

In the course of determining a quantity from an experimental system, some error will be introduced. These errors are generally referred to as noise. When the size of the errors are significant when compared to the quantity that is required to be measured, it is obviously important to know if the amount of noise can be reduced, and if so, how this reduction can be achieved. If the measurement is from a system which varies *slowly* with time (slowly with respect to how often the measurements are made), then it seems likely that the error will be reduced by averaging the results of many determinations. This of course assumes that there is no *bias* to the errors — that is, that the errors should average out to zero in the limit of infinitely many measurements. If the system does not vary slowly with time, then it cannot be assumed that the measurement will give the same answer (up to noise) each time. Instead, it is necessary to make an assumption about how the measured quantity should vary with time.

Successive determinations of such a quantity are usually termed a *time series*. Time series analysis is a vast area of study, and different assumptions about how the measured quantity should change with time lead to different analysis techniques (see [29] for an introduction to non-linear time series analysis). How a measured quantity varies with time is really a reflection of the *dynamics* of the system that is being investigated. One of the most common assumptions then, is that the system behaves as a *linear dynamical system*. This is an obvious generalisation of the no-change situation, and as such, should be expected to be a reasonable assumption for a broader range of experimental systems.

Given a linear dynamical system and knowledge of how the quantity of interest relates to it (the *physics* of the system), there are many ways of attempting to reduce the noise on each measurement made. A traditional method is to fit a *noise-free* time series to the observed measurements using a *least-squares error* approach. That is, the sum of the squares of the differences between each observed measurement and the corresponding fitted series value is minimised. As with the no-change situation, this procedure might be expected to reduce the amount of noise present at each measurement, provided there is no bias to the errors, *or in the fitting*.

A bias in the fitting can be caused by using the wrong dynamical system to generate noise-free data. Generally, there is no reason to expect that the experimenter knows the exact dynamics of their system. In fact, it is possible that this is what the experimenter hopes to elucidate. To generate noise-free data to fit to the measurements, it is therefore necessary to have a model (generally a class of models) which is assumed to include the experimental system (or at least a satisfactory approximation of it). The problem of noise reduction therefore leads, inexorably, to the problem of *modelling* a dynamical system. A good model of a dynamical system should be able to provide noise-free data that approximates the observed measurements. This artificial data is hopefully close to the *true values* of the experimental quantity being investigated, and may therefore represent successful noise reduction.

Hence the problem of noise reduction is intimately related to the modelling problem. It should also be noted that the fitting of the artificial data to the observed measurements after a model has been chosen, reflects the *state estimation* problem: Given an observed time series (noisy), can the state of the dynamical system at some particular time (say the initial or final point of the series) be determined? The point here is that a good model needs a good initial condition (or end condition, or both) in order to yield good artificial data. Noise reduction therefore is also intimately related to the state estimation problem. Note however, that state estimation asks for the state of the underlying dynamical system, whereas noise reduction is only concerned with the dependent quantity of interest. State estimation is therefore (generally) more fundamental.

For linear dynamical systems, the problems of modelling, state estimation and noise reduction have been investigated thoroughly, and answers are usually discussed in the language of *Kalman Filter theory* ([3, 4]). However, for non-linear dynamical systems, the

theory pertaining to these problems is far less developed, and has only recently become commonly applied to practical problems. In this thesis, a gradient descent algorithm for noise reduction is presented and its performance when applied to non-linear dynamical systems investigated. In particular, it is shown that noise reduction is guaranteed *asymptotically* — that is, in the limit of infinitely many measurements (covering infinite time) — at all data points except those near the initial and final points (subject to several conditions and assumptions of course). This algorithm is introduced below.

1.2   Gradient Descent and Other Algorithms

Throughout this thesis, it will always be assumed that the dynamics of the system under investigation will be known. This assumption is made to simplify the analysis. Further, the dynamical system will be assumed to be *discrete*:

$$y_{i+1} = f(y_i), \qquad i \in \mathbb{Z},$$

and the dynamical map, f, will be assumed to be a diffeomorphism from (a subset of) $\mathbb{R}^d$ into itself[1]. Let $\{x_i\}_{i=1}^n$, $x_i \in \mathbb{R}^d$, be the set of experimental measurements made (equally spaced in time). It is convenient to regard this set as a vector in $\mathbb{R}^{nd}$: $x = (x_1, x_2, \ldots, x_n)$. The output of the noise reduction algorithm is therefore another vector in $\mathbb{R}^{nd}$ which will be denoted by $\hat{x}$. Because the dynamical map is assumed to be known, there is no loss of generality in assuming that the noise on the measurements is additive. If the true value at time i is y_i then, the noisy measurements must satisfy

$$x_i = y_i + \delta_i$$

where the δ_i are a realisation of some noise distribution (assumed independently and identically distributed). In the forthcoming analysis (Chapter 5), this noise distribution will be assumed to be bounded. Clearly this is not a severe limitation.

The gradient descent algorithm for noise reduction is as follows. A *determinism function* $L : \mathbb{R}^{nd} \to \mathbb{R}$ is defined by

$$L(x) = \frac{1}{2} \sum_{i=1}^{n-1} \|x_{i+1} - f(x_i)\|^2. \tag{1.1}$$

---

[1]Later, f will be restricted to act on a compact manifold M for technical reasons. However, the noise reduction algorithms are more conveniently discussed in Euclidean space, and it is clear that their action can always be transferred back onto the manifold using the appropriate charts.

The norm used in this definition is completely arbitrary — the standard Euclidean norm is convenient (it has nice analytic properties) and will be used in what follows. Note that $L(x) = 0$ precisely when the points $x_i \in \mathbb{R}^d$ form a deterministic trajectory of f. Generally, the noisy measurements do not form a deterministic trajectory, but the idea is that one can be generated by minimising L. This can be done via gradient descent, hence the name. The algorithm therefore amounts to solving the set of differential equations

$$\dot{x}(t) = -\nabla L(x(t)), \qquad x(0) = x. \qquad (1.2)$$

The noise reduced trajectory, $\hat{x}$, is then given by $\hat{x} = \lim_{t \to \infty} x(t)$. The "time" variable t used in the gradient descent will be referred to as the *descent time* to distinguish it from the discrete "time" implicit in the iteration $y_i \mapsto y_{i+1} = f(y_i)$.

This algorithm was introduced (in this context) by Davies ([12]), as an alternative to the earlier algorithms of Kostelich and Yorke ([32]), Hammel ([21]), Schreiber and Grassberger ([16]) and Farmer and Sidorowich ([15]). The method of Kostelich and Yorke involved replacing each small segment of the noisy trajectory by a nearby trajectory that better fitted the *linearisation* of the dynamics in that segment. Hammel improved this by using a Newton-Raphson type algorithm to compute a common zero of the functions

$$\varepsilon_i = x_{i+1} - f(x_i).$$

Each iteration of the Newton algorithm amounts to finding the unique solution of a set of *algebraic* equations (rather than differential ones), provided that an extra constraint is imposed upon the ε_i. The constraint chosen by Hammel (on the basis of previous work) was to set ε_1 to be zero *in the stable direction* and ε_n to be zero *in the unstable direction*. These are directions (relative to the noisy points) where f locally contracts and expands distances, respectively (see Chapter 3). The algorithm of Farmer and Sidorowich on the other hand, aims to find the closest deterministic trajectory to the noisy data. This corresponds to minimising

$$L'(w) = \sum_{i=1}^{n} \|x_i - w_i\|^2$$

(where the x_i are the noisy data points) subject to the constraints $w_{i+1} = f(w_i)$, $i = 1, \ldots, n-1$. Using Lagrange multipliers, the minimisation can be achieved using Newton's method (again).

4

The justification for these algorithms typically involves the Shadowing Lemma ([8, 17]). This states that if the data came from a *hyperbolic* dynamical system (defined in Chapter 3), then for sufficiently small bounded noise, a nearby deterministic trajectory can always be found. If the data sequence is *infinite* (in past and future) then the nearby deterministic trajectory is unique. It seems plausible then that for a finite (but sufficiently long) data sequence, the nearby deterministic trajectory should be almost unique — all the sufficiently close trajectories are identical (to an excellent approximation) except perhaps near their initial and final points (where the finite length is important). In [13], Davies argues that therefore there is no point in explicitly trying to find the *closest* deterministic trajectory to the noisy data — any deterministic trajectory (reasonably near by) looks like the closest except around the initial and final points. The gradient descent algorithm (which just finds a deterministic trajectory) should thus be as good at noise reduction as that of Farmer and Sidorowich.

It is not the aim of this thesis to compare the performances of noise reduction algorithms. As stated before, the aim is to *prove rigorously* that the gradient descent algorithm will achieve noise reduction (under certain conditions). In recent work addressing a similar problem ([35]), Lalley notes that

> Although various *ad hoc* "noise reduction" algorithms have been proposed (some seemingly quite effective when tested on computer-generated data from low-dimensional chaotic systems, ...), their theoretical properties are largely unknown.

He then introduces another noise reduction algorithm and *proves* a result concerning its theoretical performance. The setup is as follows. Suppose that $\{y_i\}_{i=1}^n$ is a trajectory of a twice-differentiable *Axiom A* system (see [49] for a definition), that x_i is the noisy measurement of y_i and that the noise comes from a distribution bounded by δ (with mean zero). The performance of the algorithm will be evaluated in the limit that the number of points, n, tends to infinity. Let κ_n be a slowly increasing sequence of integers[2]. For each j between $\kappa_n + 1$ and $n - \kappa_n$, consider the *sub-trajectory* of $\{x_i\}$ consisting of the $2\kappa_n + 1$ points centred around x_j. This will be called the κ_n-trajectory about x_j. Then for each j, find all the x_k for which the κ_n-trajectories about x_k have each point within 3δ

---

[2]For technical reasons, Lalley suggests taking $\kappa_n \sim \log n / \log \log n$.

(say) of the corresponding point in the κ_n-trajectory about x_j. The noise reduced point $\hat{x}_j$ is defined to be the average of these x_k. If $j \leqslant \kappa_n$ or $j \geqslant n - \kappa_n + 1$, $\hat{x}_j$ is defined to be x_j. What this means is that all the sub-trajectories (of a given length) which "look" sufficiently like the sub-trajectory centred on x_j are found, and the central points of each of these sub-trajectories are averaged to get the noise-reduced point $\hat{x}_j$. Lalley's result is the following:

Theorem 1.1 (Lalley) *For the "averaging" algorithm described above, if the noise bound δ is sufficiently small, then for every $\varepsilon > 0$, the probability that the proportion of points where the noise reduction fails exceeds ε, tends to zero as n tends to infinity. That is,*

$$\lim_{n \to \infty} P\left(\frac{1}{n} |\{i : \|y_i - \hat{x}_i\| > \varepsilon\}| > \varepsilon \right) = 0.$$

Furthermore, if the initial point of the clean trajectory $\{y_i\}$ was chosen at random from the SRB-measure, then with probability one, the errors between the noise reduced points and the clean points converge uniformly to zero, except around the initial and final points. Specifically,

$$\lim_{n \to \infty} \max_{\kappa_n + 1 \leqslant i \leqslant n - \kappa_n} \|y_i - \hat{x}_i\| = 0$$

with probability one.

This result relies heavily on the statistical theory of dynamical systems (ergodic theory). Information on this (and SRB-measures) may be found in [7, 14, 53]. Lalley's result guarantees (with probability one at least) that in the limit of infinite data, the averaging algorithm returns the original clean trajectory *except* near the initial and final points (provided the clean trajectory is "typical"). In Chapter 5, a similar (though slightly stronger) result is proved for the gradient descent algorithm.

1.3   Indistinguishable States

The motivation behind studying the theoretical properties of the gradient descent algorithm did not, however, come from Lalley's result, nor did it come from consideration of the noise reduction literature. This study was in fact initiated by the theory of *indistinguishable states* of Judd and Smith ([28]), and relates to the state estimation problem.

This theory and its relation to the gradient descent algorithm will be summarised here for completeness.

Suppose that a *semi-infinite* time series is given, $(\ldots, x_{-2}, x_{-1}, x_0)$, terminating at time 0. The aim is to determine the state of the underlying dynamical system at time 0 so that the values of $x_1, x_2, \ldots$ may be predicted. If the time series is noise-free (and the underlying dynamical system is smooth), then Takens' Theorem ([51]) implies that an equivalent dynamical system can be constructed for which the state at time 0 is known exactly, and which gives the time series as the actual underlying dynamical system. Furthermore, only a finite (but sufficiently long) time series is required for this construction. If the time-series is noisy however, it is clear that the exact state cannot be determined in this way from a finite time series. The question is whether the exact state can be determined if the entire semi-infinite time series is used.

The noise on the time series will be assumed independent and identically distributed with distribution ρ. Given a noisy measurement x_i then, there are many candidates for the exact point that gave rise to it. These candidates are said to be *indistinguishable* from one another, on the basis of the measurement x_i. Obviously, knowledge of the noise distribution would let one quantify which candidates were more likely to have given rise to x_i. Take any two points y_i and y_i'. The probability of y_i and y_i' being indistinguishable (written $y_i \sim y_i'$) on the basis of a noisy measurement, is given by

$$\mathrm{P}\left(y_i \sim y_i'\right) = \frac{\displaystyle\int \rho\left(x_i - y_i\right) \rho\left(x_i - y_i'\right) dx_i}{\displaystyle\int \left[\rho\left(x_i\right)\right]^2 dx_i} \tag{1.3}$$

(see Figure 1.1). Note that $\mathrm{P}\left(y_i \sim y_i\right) = 1$ as it should.

Suppose now that $y = (\ldots, y_{-2}, y_{-1}, y_0)$ and $y' = (\ldots, y_{-2}', y_{-1}', y_0')$ are two possible semi-infinite time series, which shall be assumed to be noise-free. As the errors are assumed independent of one another and are identically distributed from point to point, the joint probability of y being indistinguishable from y' is given by

$$\mathrm{P}\left(y \sim y'\right) = \prod_{i=-\infty}^{0} \mathrm{P}\left(y_i \sim y_i'\right).$$

The interpretation is that if $\mathrm{P}\left(y \sim y'\right) = 0$, then there is enough evidence to conclude (with probability one) that y and y' cannot give rise to the same noisy time series. $\mathrm{P}\left(y \sim y'\right)$ is the probability that y and y' cannot be distinguished on the basis of the observation of some noisy version of y (or y').
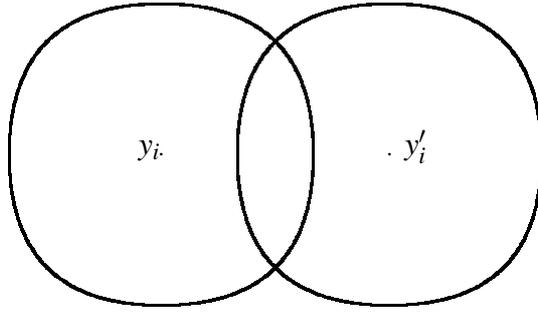
Figure 1.1: If the points y_i and y'_i are subjected to additive noise (which for clarity is assumed to be bounded and uniform) then the noisy points must lie within the indicated regions. If one such noisy measurement is observed, then it is clear that one can tell whether it came from y_i or y'_i unless the measurement lies in the intersection of the two regions. The *probability* that such a noisy measurement will distinguish between the points y_i and y'_i is therefore proportional to the area of overlap of the two regions. Equation 1.3 is a direct generalisation of this.

The question of whether a semi-infinite noisy trajectory allows the unique determination of the final state (and hence the state at all times) now reduces to the question of whether there are any other noise-free time series indistinguishable from the actual noise-free time series. If so, then the final state is not uniquely specified. To illustrate this, consider the simple case of a one-dimensional dynamical system with Gaussian noise. That is, suppose that

$$\rho(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-z^2/2\sigma^2}.$$

A quick calculation shows now that

$$
\begin{aligned}
\mathrm{P}\left(y_i \sim y'_i\right) &= e^{-\left(y_i - y'_i\right)^2/4\sigma^2} \\
\Rightarrow \quad \mathrm{P}\left(y \sim y'\right) &= \exp\left\{\frac{-1}{4\sigma^2} \sum_{i=-\infty}^{0} |y_i - y'_i|^2\right\}.
\end{aligned}
\tag{1.4}
$$

Therefore y and y' are indistinguishable if $\sum |y_i - y'_i|^2 < \infty$. Thus, $y - y'$ must converge to zero as i tends to $-\infty$. For *isotropic* Gaussian distributions in higher dimensions, this result generalises by replacing $|\cdot|$ by $\|\cdot\|$ (the Euclidean norm in higher dimensions). Similar results are true for more general distributions ([28]). It is convenient to introduce a definition here. The *unstable manifold* of the point y_0 (with respect to the dynamical

system) is the set $W_u(y_0)$ given by (see also Chapter 3):

$$W_u(y_0) = \left\{ y_0' : \lim_{i \to -\infty} \|y_i - y_i'\| = 0 \right\}.$$

It follows now that if y and y' are indistinguishable, then y_0' must lie on the unstable manifold of y_0. The convergence of $\|y_i - y_i'\|$ to zero in the definition of the unstable manifold is exponential (for nice systems anyway, see Chapter 3), so the converse is also true[3].

Hence the final state cannot be determined uniquely. However, all the possible final states must lie on the unstable manifold of the true final state. This suggests that rather than getting a "best estimate" of the final state and using it for prediction, one should use an *ensemble* sampled from the unstable manifold of the best estimate. This ensemble will contain the true final state (or at least a good approximation to it). How is the best estimate of the final state obtained? Judd and Smith use the gradient descent algorithm applied to the noisy data — the final point of the output becomes the best estimate. This may not be a superb estimate of the truth, but they claim (based on numerical evidence) that the ensemble constructed this way will contain the truth, at least if the noise is not too large. That is, that the gradient descent gives a best estimate on the unstable manifold of the true state. This claim is in fact equivalent to the claim that in the limit of infinite data, the gradient descent algorithm gives the true clean time series except near the initial and final points, as will be seen in Chapter 5.

1.4 Overview

The rest of this thesis consists of an attempt to give a rigorous proof of the claim that the gradient descent algorithm discussed above achieves noise reduction for a broad class of dynamical systems. More specifically, that in the limit as the number of points in the noisy trajectory given tends to infinity, the noise reduced trajectory given by gradient descent converges onto the original clean trajectory except near the initial and final points. This will be achieved by deriving analytic bounds on the errors at each point along the

---

[3]Actually, this holds only for unbounded noise distributions. If ρ is bounded, then y_0 and y_0' cannot be too far apart, otherwise they will be distinguishable on the basis of a single measurement. In this case then, the set of y_0' for which y and y' are indistinguishable form a bounded subset of the unstable manifold of y.

trajectory, and showing that these bounds converge to zero (as the length of the trajectory tends to infinity) except near the initial and final points.

In Chapter 2, the gradient descent algorithm is applied to artificial data from some two dimensional non-linear dynamical systems. These numerical experiments show the typical behaviour of noise reduction algorithms (as well as some atypical behaviour) and suggest various constraints that must be made upon the dynamical system in order to have any chance of establishing rigorous bounds on the errors. Of particular note here is the phenomenon of *tangencies*. Tangencies in non-linear dynamical systems are a bane of noise reduction algorithms. Here, a couple of examples are investigated numerically and some interesting behaviour noted. However, no attempt is made in what follows to treat the effect of tangencies in a rigorous manner as this would almost certainly require theory more developed that what is used in this thesis (Pesin theory for instance, see [42, 30]).

Chapter 3 introduces the parts of the standard theory of non-linear dynamical systems that is required in what follows. This includes summaries of contraction maps, stable manifold theory, centre manifold theory and other generalisations, as well as a couple of results concerning Lyapunov numbers. As a rule, proofs for all these results are referred to the literature. However, as the mandatory exception, there is one full proof. The result is a simple fact about the stretching and contracting in a dynamical system, but contains a couple of subtleties which are often ignored and occasionally misinterpreted. This fact (and its subtleties) are required for later proofs.

The technical work begins in Chapter 4. To get error bounds for the noise reduction of a non-linear system, the strategy employed is to relate it to the *linearised system* for which the analysis is somewhat easier. To this end then, the Hartman-Grobman Theorem is introduced and proven. This result is then extended to give *quantitative* information about the correspondence between the system and its linearisation (necessary for the analytic bounds to be set up later). A generalisation of the Hartman-Grobman Theorem due to Kurata is then presented and proven, and this too is extended to give quantitative information. The quantitative extensions given in this chapter are due to the author — this type of result is almost certainly known to experts but these cases do not seem to appear in the literature. Additionally, the proof of Kurata's generalisation has been modified by the author, both to clarify and explain the details of this elegant result, and to allow the required quantitative extension to be proven as simply as possible.

Finally then in Chapter 5, the proof that the gradient descent algorithm guarantees noise reduction begins. This result is first proven in the case of a symmetric linear dynamical system for simplicity. The surprisingly difficult generalisation for more general linear systems follows, and from here it is a small matter to pass to a linearised dynamical system. The transition to the full non-linear system is effected by constructing a commutative diagram between the non-linear gradient descent and its counterpart for the linearised system. It is here that the technical results of Chapter 4 are used, to ensure the components of the diagram have the correct properties. The quantitative nature of these results allow the bounds from the analysis of the linearised system to be carried across to the non-linear system, provided a particular condition is satisfied (Condition 5.12), and it is then shown that the non-linear gradient descent achieves noise reduction. This condition is argued (though not proven) to be satisfied by any dynamical system of the type suggested by the experiments of Chapter 2. All the results of this chapter are the work of the author.

Chapter 2

Numerical Experiments

In this chapter, some results of the gradient descent algorithm for noise reduction will be presented and discussed. As the algorithm consists of solving a set of differential equations (equations 1.2), it is very easy to implement. The following results were obtained using the `ode15s` function in MATLAB, by letting the descent-time variable increase until convergence appeared to have been established. The systems considered will all be artificial so that the original clean trajectory is known.

2.1 The Hénon Map

Perhaps the simplest example of a non-linear diffeomorphism, the Hénon map was introduced and investigated numerically (by Hénon) in 1976 ([22]). In fact, he investigated a family of maps, a version of which may be represented by the functions $f_{a,b} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by[1]

$$f_{a,b}(x,y) = \left(a - x^2 + by, x\right), \tag{2.1}$$

where a and b are given constants. Note that if $b \neq 0$, then this map is invertible. In what follows, a shall be set to $7/5$ and b to $3/10$.

Numerical iteration of any point sufficiently close to the origin by $f \equiv f_{7/5,3/10}$ gives a plot similar to that in Figure 2.1. All the points appear to lie on (or very near) a smooth curve. This curve folds back on itself so that its points remain bounded, and it seems reasonable to presume that the curve's length is in fact infinite. This curve is called the *Hénon attractor*, although f has not yet been rigorously proven to even possess such an

---

[1]This is not the form usually given for this map ([18]), but the qualitative features are the same.
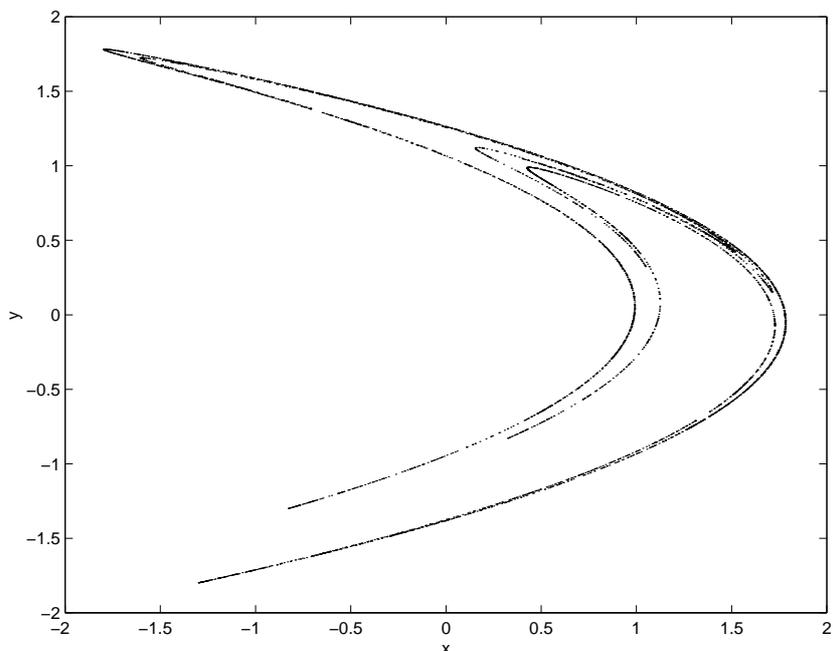
Figure 2.1: The Hénon attractor.

attractor in the technical sense of the word ([18], see also however, [6], and more recently, [52]). Transverse cross sections of the attractor have a strong resemblence to a Cantor set, and so the Hénon attractor is said to be *fractal* or *strange*.

The clean trajectory chosen for the gradient descent algorithm has ten points, with initial point $(-1.1709, 1.6318)$ and final point $(0.4552, -1.1989)$ (approximately). Each point of this trajectory is then perturbed by adding a random quantity to each coordinate. In this case, a normal random variable with mean zero and standard deviation $1/5$ was used. This "noised-up" trajectory was used as the initial condition for the gradient descent. The output from the differential equation solver used (`ode15s`) is shown in Figure 2.2 (left). From this it would appear that convergence was attained by $t = 50$. The ordinate axis here measures the deviation in each coordinate of each point of the noisy trajectory as the algorithm proceeds. That is, if $(x_1(t), y_1(t)), (x_2(t), y_2(t)), \ldots, (x_n(t), y_n(t))$ represents the coordinates of each point when the gradient descent has reached descent-time t, then the ordinate axis measures each $x_i(t) - x_i(0)$ and $y_i(t) - y_i(0)$ for $i = 1, 2, \ldots, n$.

The coordinates of the points after convergence has been achieved give the noise-reduced trajectory. In this case, the value of the determinism function L (see equation 1.1) was about 7×10^{-12}. The noise-reduced trajectory is then compared with the original

14
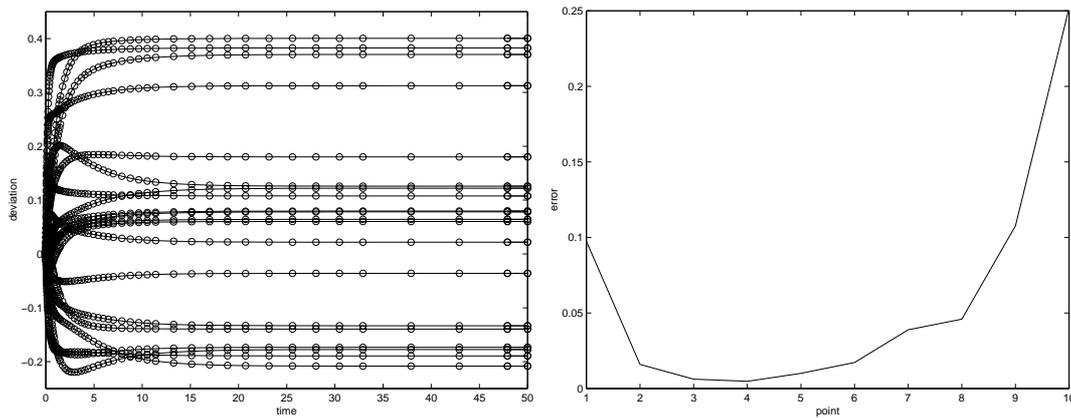
Figure 2.2: Convergence (left) and errors (right) from applying the gradient descent algorithm to a ten point trajectory of the Hénon map.

clean trajectory, and the errors (Euclidean) at each point plotted in Figure 2.2 (right). The result is also summarised pictorially in space in Figure 2.3, where the squares are points of the original clean trajectory, the circles are points of the noisy trajectory, and the lines leading from them represent the progress of the gradient descent algorithm.

The errors displayed in Figure 2.2 show a typical trend for noise reduction algorithms: They are quite small except near the initial and final points of the trajectory. The error values themselves depend on which particular noisy trajectory was generated from the clean one, but as may be seen in Figure 2.4, the "shape" of the error distribution remains the same, and its amplitude does not vary significantly in magnitude, even when a longer piece of trajectory is considered.

The clue to why this should be so is contained in Figure 2.3. There are three instances here where the gradient descent algorithm has converged to what is obviously the wrong point — the two points at the top left and one of the points at the bottom centre. It is easily checked (and should be obvious from the error graph in Figure 2.2) that these are the initial point, the final point and its predecessor. In fact, the final point and its predecessor are the points which appear to converge onto the attractor, whereas the initial point is the point which does not. Rather, the initial point appears to converge onto *the attractor*[2] *of* f^{-1}, the inverse Hénon map. It is somewhat difficult to get a picture of this inverse attractor,

---

[2]Actually this is not an attractor. Nearby trajectories do not come closer and closer to it, but rather tend further and further away. It is more correct to label this structure a *repellor*

15
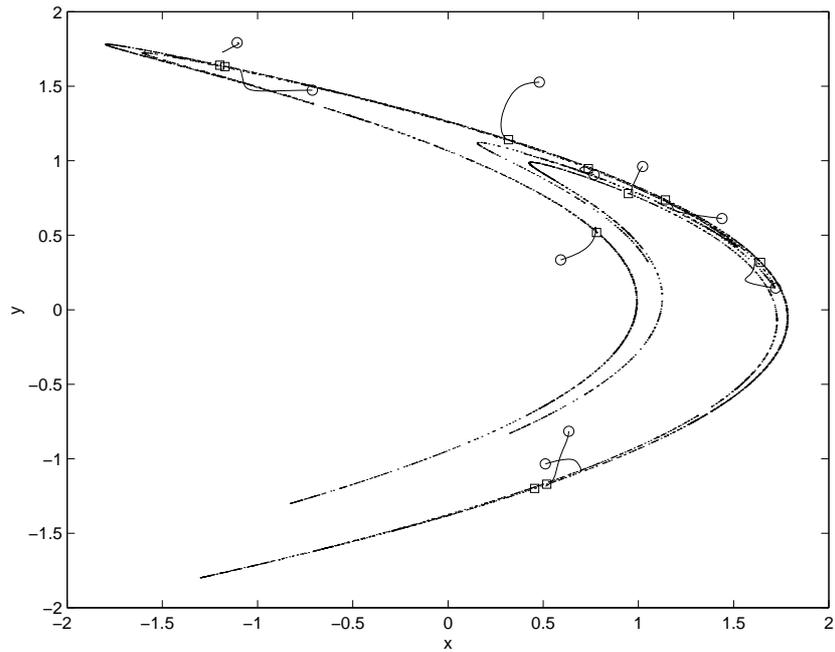
Figure 2.3: The Hénon attractor and the progress of the gradient descent algorithm. The squares label the points of the clean trajectory, the circles the points of the noisy trajectory, and the lines emanating from the circles represent the progress of the gradient descent algorithm.
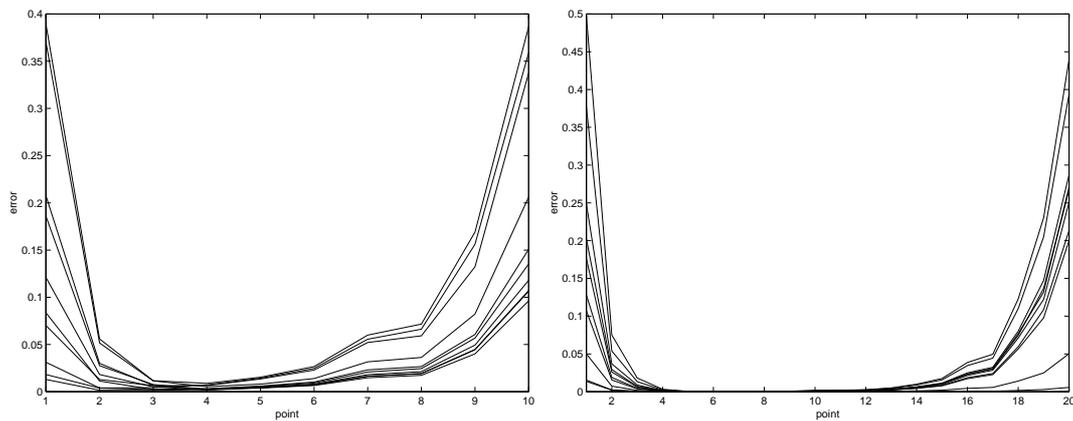
Figure 2.4: Errors from ten different noise realisations added to a ten-point trajectory (left) and a twenty-point trajectory (right) of the Hénon map. Note that the shape of the error distributions is largely unaffected by the particular noise realisation used, and the average size of the errors at the initial and final points seem to be unaffected by how many points are used.
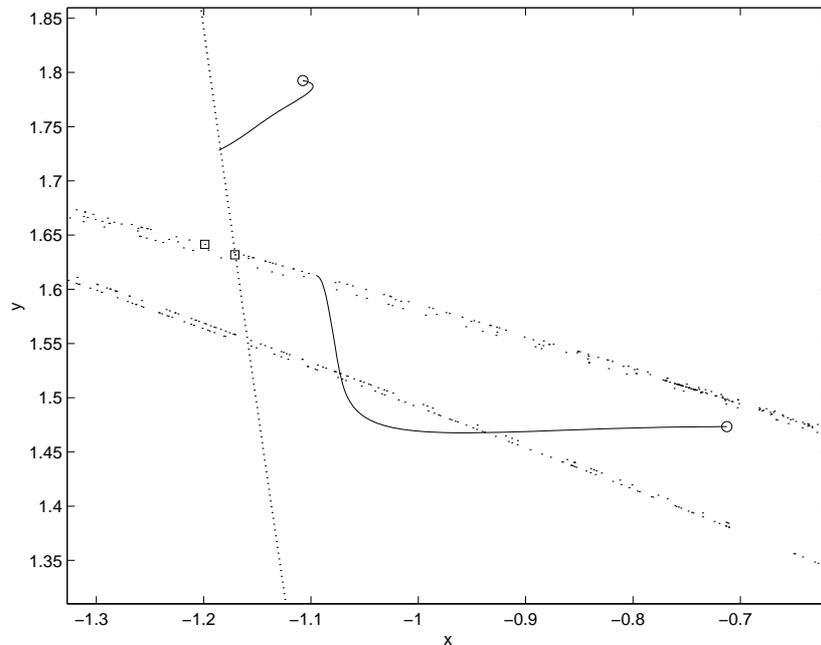
Figure 2.5: A magnification of the top-left of Figure 2.3 showing additionally, a piece of the inverse Hénon attractor. Note that one point appears to converge onto the attractor whereas the other appears to converge onto the inverse attractor

as it is unbounded so iterates under f^{-1} have a habit of ending up at infinity. However, the piece passing through the initial point of the clean trajectory is easily generated and is shown in Figure 2.5.

So, the final point of the noise-reduced trajectory lies on the piece of the attractor which passes through the final point of the original clean trajectory, and the initial point of the noise-reduced trajectory lies on the piece of the inverse attractor which passes through the initial point of the original clean trajectory. To make this statement a little less cumbersome, note that the relevant pieces of the attractor and inverse attractor appearing above are in fact, (local) unstable and stable manifolds[3], respectively. Therefore, after gradient descent, the computed final point should be on the unstable manifold of the correct final point, and the computed initial point should be on the stable manifold of the correct initial point.

---

[3]See chapter 3 for definitions.
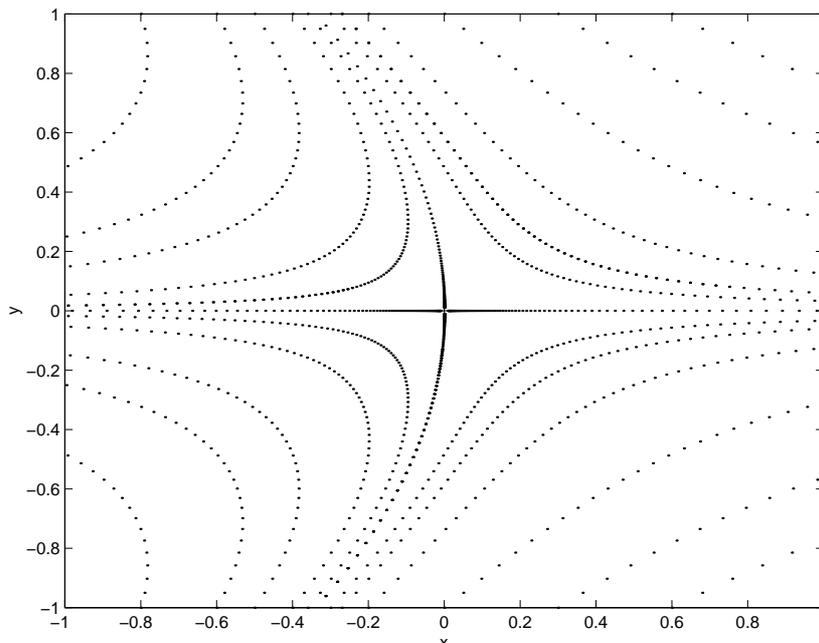
Figure 2.6: Phase portrait showing the dynamics of the map defined in equation 2.2. The fixed point at the origin is a saddle point with the vertical direction stable and the horizontal direction unstable.

2.2 Refinements

The conjecture of the previous paragraph is, unfortunately, false as it stands. As a counterexample, consider the system $g_\varepsilon : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$g_\varepsilon(x,y) = \left(\varepsilon y^2 + (1+\varepsilon)x, (1-\varepsilon)y\right). \tag{2.2}$$

This dynamical system has a fixed point at the origin. A phase diagram showing various trajectories is shown in Figure 2.6. A ten-point trajectory of $g_{1/20}$ was generated with initial point approximately $(0.0035, 0.2210)$. Gaussian noise with mean zero and standard deviation $1/50$ was added to each coordinate of each point of the trajectory, and the result was then noise reduced using the gradient descent algorithm. The results are shown in Figure 2.7. Note that the initial and final points of the noise-reduced trajectory *have not* converged onto the stable and unstable manifolds of the true initial and final points (respectively). Note also the shape of the error distribution — the errors near the middle of the trajectory are *not* small compared to those at the initial and final points.

The noise-reduction algorithm fails for this example because the algorithm has not
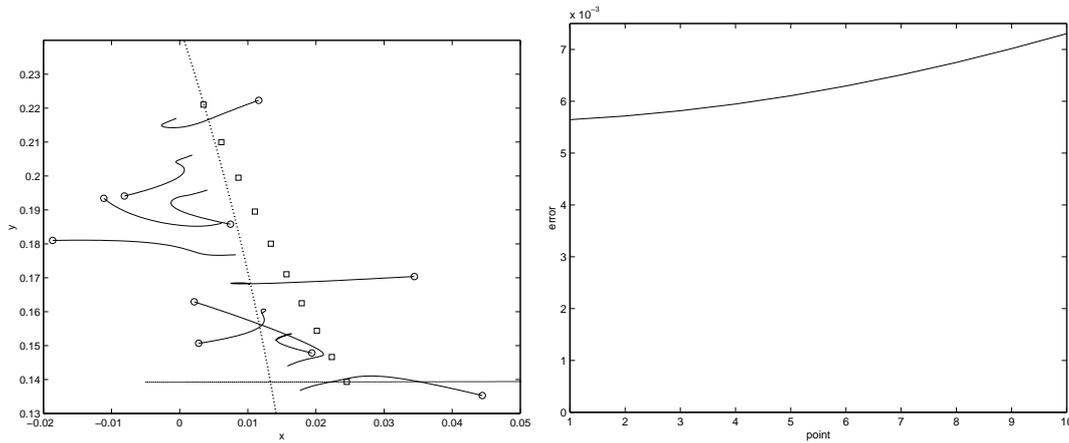
Figure 2.7: Results from applying the gradient descent algorithm to a ten-point trajectory from the map $g_{1/20}$. The symbols are used as they were in Figure 2.3. The stable manifold of the initial point of the clean trajectory and the unstable manifold of the final point are also shown (upper left and lower right respectively).

been given a long enough trajectory. If it is given forty points (with the same initial point say), then a typical result is shown in Figure 2.8. Note that now the initial and final points are on the stable and unstable manifolds (or at least much closer to them). Also, the error distribution now has the expected shape even if the errors in the middle of the distribution are still not small.

It would seem then that for a sufficiently long trajectory, the error distribution will be such that the errors decay at first, reach a minimum, and then increase towards the end. It is convenient to think of this distribution as having two components, one each corresponding to the stable and unstable manifolds of the points of the trajectory. The error along the stable manifold decreases (exponentially) and that along the unstable manifold increases (exponentially). If the rates of increase and decrease (called expansion and contraction rates) are close to unity, then a long trajectory will be required to make the errors in the middle small (see chapter 3, especially section 3.3.1 for more on these rates). This is the case with the map g_ε for ε small, whose expansion and contraction rates are $1 + \varepsilon$ and $1 - \varepsilon$ respectively (in fact, this map is a small perturbation of a linear map — indeed the linear map would have sufficed for the above discussion). For the Hénon map (equation 2.1), the expansion and contraction rates can be approximated numerically, and are about

19
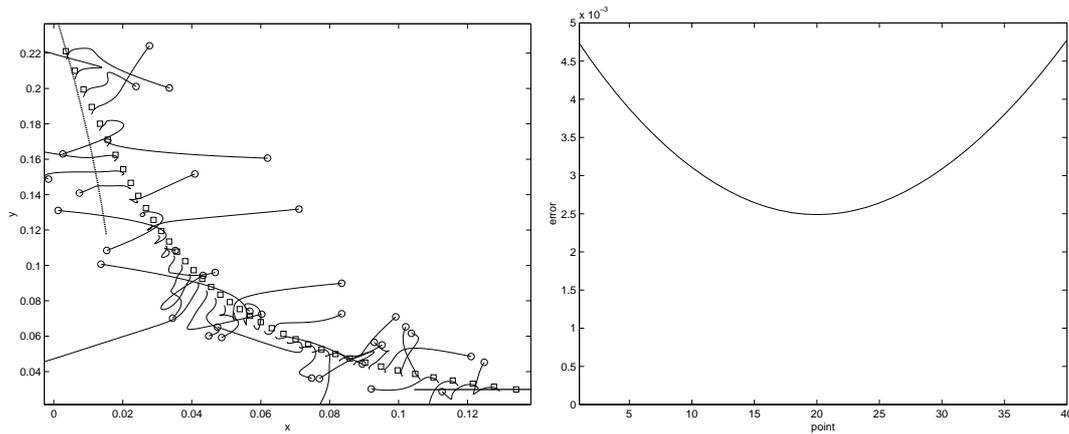
Figure 2.8: Results from applying the gradient descent algorithm to a forty-point trajectory from the map $g_{1/20}$. The initial point is at top-left, the final point at bottom-right.

1.52 and 0.198 ([2]). Therefore, only a few points are needed to make the errors in the middle of the trajectory small. Note that in Figure 2.4 the errors do indeed decrease much more quickly at the start, than they increase towards the end!

Of course, it is possible that the map used has an expansion or contraction rate equal to unity. That is, there is a direction where the map does not significantly expand or contract distances. On the basis of the previous paragraph then, the error distribution of a trajectory from such a map would not be expected to be small, no matter how long the trajectory was. A simple example confirming this statement is the identity map — the errors after noise reduction are necessarily constant from point to point. Such maps are referred to as being *non-hyperbolic*, to contrast with the maps whose rates of increase and decrease are not unity, called *hyperbolic* maps. Hyperbolicity will be discussed further in Chapters 3 and 4.

It should be mentioned here that the errors in the middle of the trajectory will decrease as the length of the trajectory increases *provided* that the errors at the initial and final points of the trajectory remain (approximately) constant. That is, these errors should be of the same order of magnitude as the amount of noise originally added, regardless of the length of the trajectory used. That this should be true is certainly a plausible statement to make, but it is by no means clear that it must hold[4]. If however, it is the case that

---

[4]One can try to force such a requirement by adding an extra term to the determinism function (equation 1.1) which penalises trajectories whose points are far away from the points of the original noisy trajectory.

all the errors are small, and that this holds for arbitrary long trajectories as well, then it would be expected that the errors would lie in the stable direction for the initial point and in the unstable direction for the final point. For otherwise, the component of the initial error in the unstable direction would be expected to grow exponentially, contradicting the smallness of the errors everywhere. Similarly with the final error. Therefore, the goal of the analysis of noise reduction algorithms should be to prove that the errors do remain small, regardless of the length of trajectory.

2.3   Tangencies

Consider now one last example, the Ikeda map $h_{a,b,c,d} : \mathbb{C} \to \mathbb{C}$ defined by

$$h_{a,b,c,d}(z) = a + be^{i\theta}z, \qquad \theta = c - \frac{d}{1 + |z|^2}, \tag{2.3}$$

introduced as a model for a cell in an optical computer ([25]). This can be expressed as a real function on $\mathbb{R}^2$, which shall also be denoted by $h_{a,b,c,d}$ and takes the form

$$h_{a,b,c,d}(x,y) = (a + b(x\cos\theta - y\sin\theta), b(x\sin\theta + y\cos\theta)), \tag{2.4}$$

where $\theta = c - d\left(1 + x^2 + y^2\right)^{-1}$. In what follows, $a = 1$, $b = 2/5$, $c = 9/10$ and $d = 6$. The attractor for $h \equiv h_{1,2/5,9/10,6}$ is shown in Figure 2.9.

A twenty-point trajectory was chosen with initial point $(0.9255, -1.0126)$ and final point $(1.1243, -2.1607)$ (approximately). Gaussian noise with mean zero and standard deviation $1/10$ was added to this trajectory before the gradient descent algorithm was applied. The resulting errors for *ten* different noise realisations are shown in Figure 2.10. The interesting feature here is the presence of "spikes" in the error distributions at points (times) 3, 8 and 16. Almost every noise distribution yields larger than expected errors at these points, and the points nearby also have larger errors than expected. It is as if at these points, the noise reduction algorithm fails, and the result is some sort of patching together of what might result if the algorithm were applied to the trajectories corresponding to points 1 to 3, 3 to 8, 8 to 16 and 16 to 20 *separately*.

The cause of this "spiking" phenomenon may be found by examining the stable and unstable manifolds of the offending points. These are shown in Figure 2.11. Notice that

---

The value of such a term has not been established however — see the discussion at the end of Chapter 5.
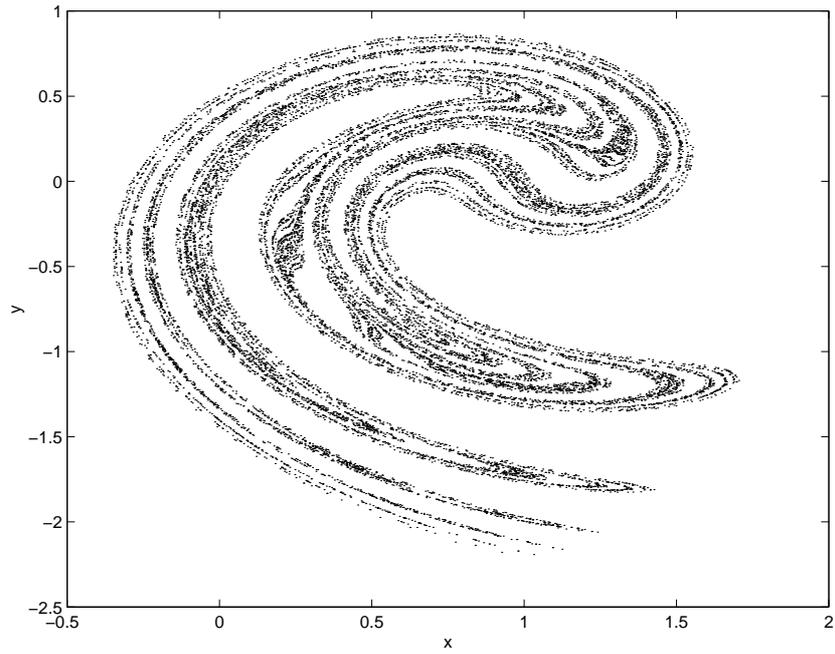
Figure 2.9: The Ikeda attractor.
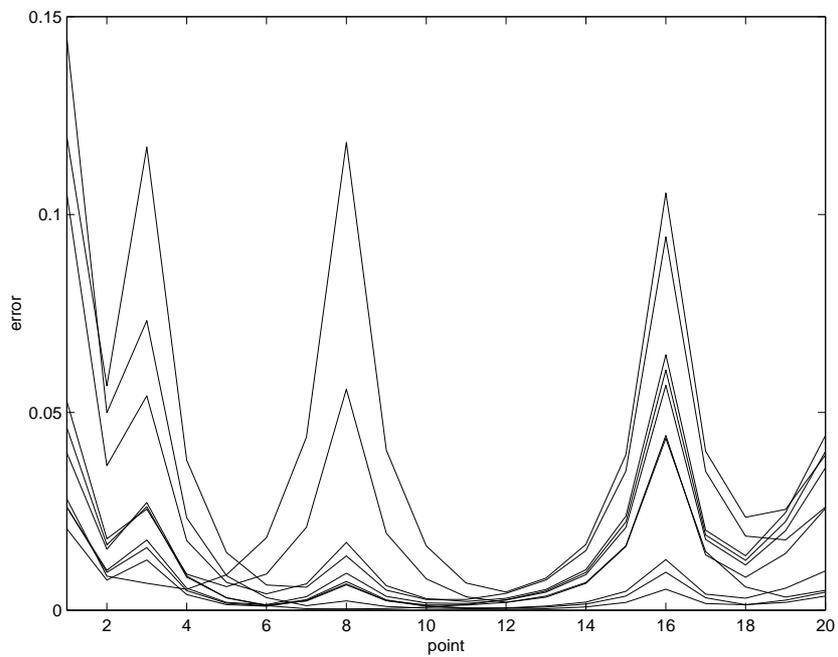
Figure 2.10: Errors from ten different noise realisations added to a trajectory from the Ikeda map.
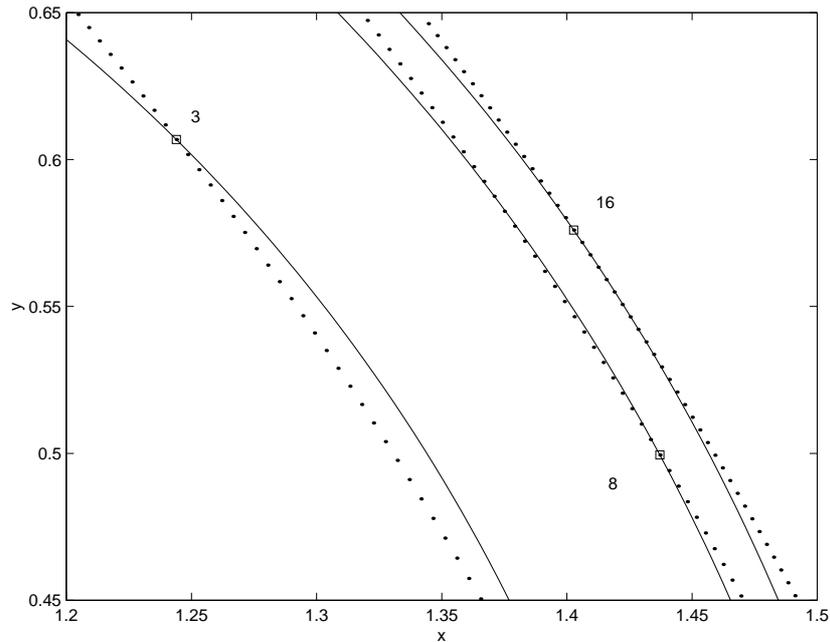
Figure 2.11: Points 3, 8 and 16 from the clean Ikeda trajectory considered, with their stable manifolds (dotted lines) and unstable manifolds (solid lines). Note that at each point, the angles between the stable and unstable manifolds are very small.

for points 8 and 16, the stable and unstable manifolds appear to be tangent to one another at the clean trajectory point, and for point 3, although the manifolds are not tangent there, the angle between them is quite slight. In fact, the angles[5] may be numerically approximated quite easily ([47]) and are (about) $8.2°$ for point 3, $2.8°$ for point 8, and $1.2°$ for point 16. Points where the angle between the stable and unstable manifolds is small are termed *near-tangency points* or often just *tangency points*. At these tangency points, the algorithm would seem to have difficulty distinguishing which direction is stable and which is unstable, and therein lies its failure.

There are certainly points on the Ikeda attractor where the angle between the stable and unstable manifold is exactly zero[6]. However, it might be expected that such points

---

[5]These are defined to be the angles between tangent lines for the stable and unstable manifolds at the point. See section 3.3.1.

[6]This is true of the Hénon attractor as well. It is the "folding" of the attractor back on itself (which is necessary for allowing expanding directions whilst keeping the attractor bounded) that forces tangencies to occur — hence they should be present in any system of this kind. Spiking occurs in noise reduction of Hénon trajectories too, though somewhat less frequently.
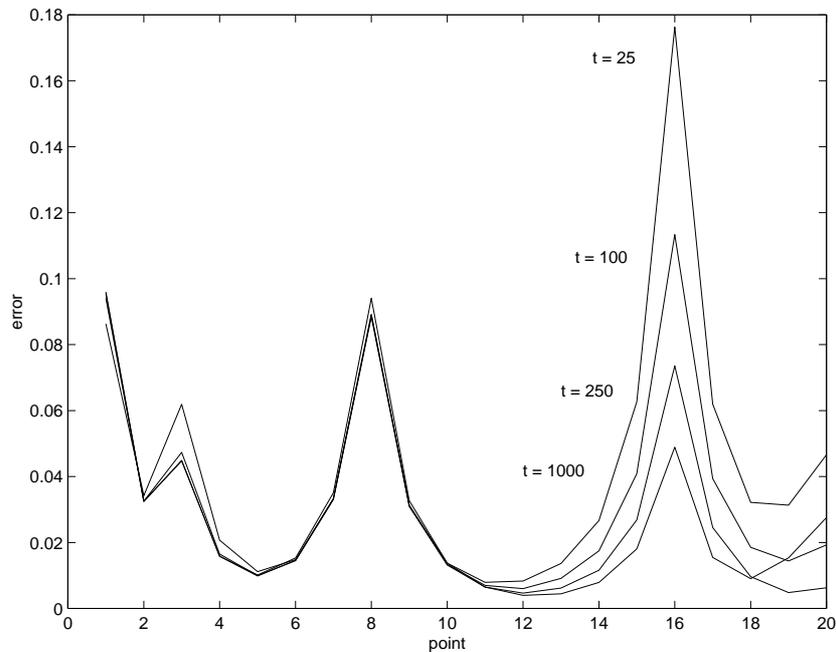
Figure 2.12: Errors after gradient descent from the Ikeda trajectory considered after the gradient descent algorithm has been to descent-times 25, 100, 250 and 1000. The distribution remained constant for descent-times greater than 1000.

are extremely rare (this expectation is justified in Theorem 3.5), and thus that they will never be encountered, practically. Therefore, for a given trajectory, the gradient descent algorithm should be able to sort out which direction is stable and which is unstable, *given enough time*. Perhaps the "spiking" effect seen in Figure 2.10 is an artefact of poor convergence.

Figure 2.12 shows the errors induced by the gradient descent algorithm for the Ikeda trajectory, at descent-times 25, 100, 250 and 1000. Recall that for the Hénon trajectory considered in section 2.1, convergence was complete by descent-time 50. It is apparent that convergence is obtained for the Ikeda trajectory by descent-time 100, except in the vicinity of the tangency at point 16. No further change in the error distribution was observed beyond descent-time 1000, however. Why points near one tangency are slow to converge compared with points near another, is not clear. In fact, with other noise realisations, it is sometimes observed that the tangency at point 8 is the one which takes a long time to converge, so it can only be concluded that the presence of tangency points in a trajectory *may* mean that the convergence of the gradient descent algorithm is very slow.
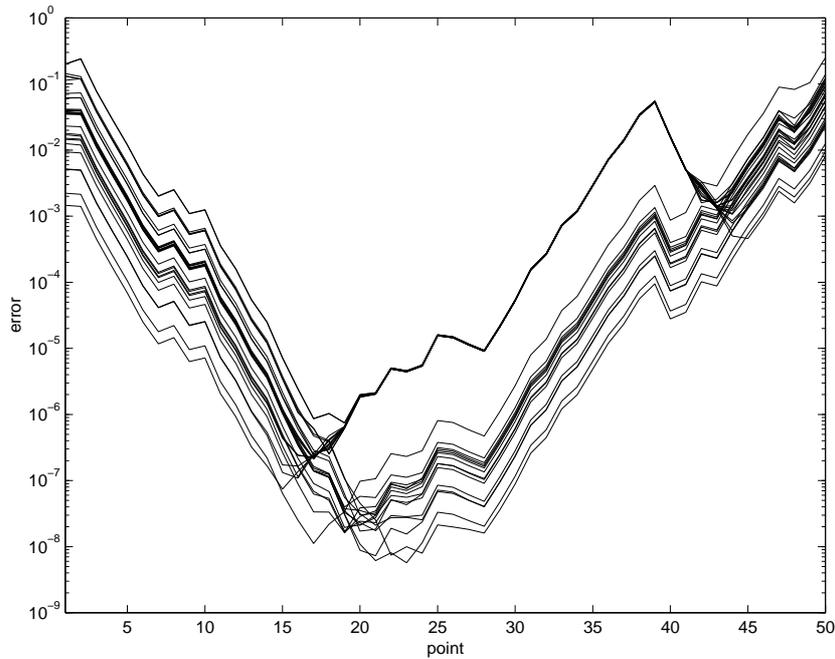
Figure 2.13: A logarithmic plot of the error distributions after noise reduction for thirty different noise realisations added to a fifty point trajectory. The noise distribution was Gaussian with standard deviation 1/10.

It is, however, clear that the errors at the tangency points remain larger than expected, even after the noise reduction algorithm has been run for very long times. Consider now Figure 2.13. This shows error distributions (in *logarithmic* scale) for thirty different noise realisations. The logarithmic scale shows the exponential decay and growth of the errors quite clearly. There is a tangency of approximately $3°$ at point 39 as well as lesser tangencies of between $10°$ and $20°$ at points 2, 10 and 25. These are visible in the error distributions. What is of greater interest is the observation that the error distributions around the tangency at point 39 form two quite distinct groups. The jump in the errors around point 39 is sometimes small and sometimes much larger[7].

These two groups are shown spatially (around the tangency point 39) in Figure 2.14. The large square marks point 39, the "+" signs mark the thirty noise reduced approximations of point 39, and the dotted and solid lines show the stable and unstable manifolds

---

[7]The errors corresponding to the larger jump are not *resolved* in this figure due to the logarithmic scale. In fact, approximately half the distributions show this larger jump. The reasons for this will become apparent shortly.

Figure 2.14: Spatial results for the 39th point of the thirty noise reduced trajectories of
Figure 2.13. The large square marks the correct point, the "+" signs mark the noise re-
duced points, and the dotted and solid lines are the stable and unstable manifolds through
the correct point.

through point 39 (respectively). *Note that the groups cluster about the points where the
stable and unstable manifolds intersect!* Points where the stable and unstable manifold
intersect are called *homoclinic intersection points*.

Why this clustering about homoclinic intersection points? Because this forces the
points to be close to the stable and unstable manifolds of the true point. Iterating forward
then means that the error must shrink (because the point is near the stable manifold).
The error along the unstable manifold must likewise grow, and this forces the unstable
manifold to bulge outwards (and the angle between the stable and unstable manifolds
to increase). Similarly, upon iterating backwards, the error along the unstable manifold
decreases and the error along the stable manifold grows, leading to a bulging of the stable
manifold (and a corresponding increase in the angle between the manifolds). This is

26

Figure 2.14: Spatial results for the 39th point of the thirty noise reduced trajectories of
Figure 2.13. The large square marks the correct point, the "+" signs mark the noise re-
duced points, and the dotted and solid lines are the stable and unstable manifolds through
the correct point.

through point 39 (respectively). *Note that the groups cluster about the points where the
stable and unstable manifolds intersect!* Points where the stable and unstable manifold
intersect are called *homoclinic intersection points*.

Why this clustering about homoclinic intersection points? Because this forces the
points to be close to the stable and unstable manifolds of the true point. Iterating forward
then means that the error must shrink (because the point is near the stable manifold).
The error along the unstable manifold must likewise grow, and this forces the unstable
manifold to bulge outwards (and the angle between the stable and unstable manifolds
to increase). Similarly, upon iterating backwards, the error along the unstable manifold
decreases and the error along the stable manifold grows, leading to a bulging of the stable
manifold (and a corresponding increase in the angle between the manifolds). This is

26

Figure 2.15: Spatial results for the 37th point (left) and the 41st point (right) of the thirty noise reduced trajectories of Figure 2.13, showing the stable (dotted) and unstable (solid) manifolds.

pictured in Figure 2.15. Thus the errors *decrease* in both directions (in fact, this argument also explains why the errors grow and decay *exponentially* around a tangency point). If the noise reduced points were not near a homoclinic intersection point, then by iterating forwards or backwards, the errors would have to eventually grow. Summarising, it can be said that the trajectories through the two homoclinic intersection points of Figure 2.14 (one of which is the true point) remain close together and so the noise reduction algorithm chooses one or the other depending on the particular noise realisation given.

Another way of saying this is that the trajectories through the homoclinic intersection points are difficult to distinguish on the basis of the given noise realisations. This difficulty can be quantified using the *indistinguishability theory* of section 1.3. The noise distribution used here was Gaussian with standard deviation $1/10$, so the probability that two trajectories y and y' will be indistinguishable given a random noise realisation is given by (equation 1.4):

$$\mathrm{P}\left(y \sim y'\right) = \exp\left\{-25\sum_i \left\|y_i - y'_i\right\|^2\right\}.$$

A plot of (an excellent approximation of) the indistinguishability of the correct trajectory

27
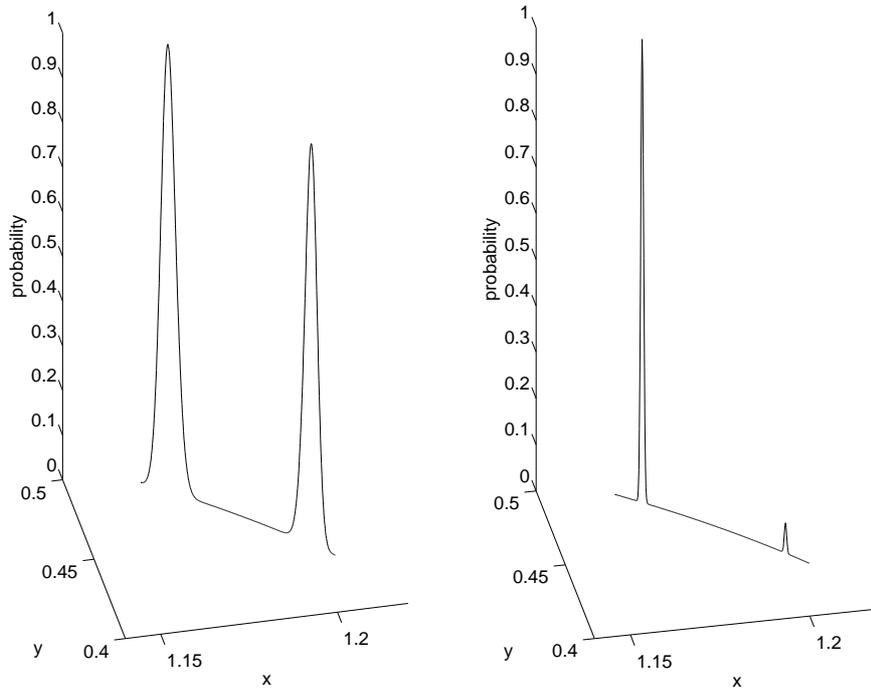
Figure 2.16: This shows the probability that the trajectory defined by the given point will be indistinguishable from the true trajectory. At left, the probability is computed assuming Gaussian noise with standard deviation $1/10$. At right, the standard deviation is $1/50$.

and the nearby trajectories is given in Figure 2.16 (left). The plot measures the probability of indistinguishability versus the point corresponding to point 39 of the correct trajectory. The two peaks correspond to the homoclinic intersection points (the peak with value 1 is obviously the correct point). The second peak has probability approximately 0.9. Therefore it is very likely that a given noise realisation will be unable to distinguish between the true trajectory and the trajectory through the other homoclinic intersection point. This explains why the numbers of noise reduced points clustered around each homoclinic intersection point are approximately equal — the two trajectories are usually indistinguishable so the noise reduction algorithm gives each with approximately equal probabilities.

The large peak around tangencies in the error distribution is therefore due to the algorithm choosing the wrong homoclinic intersection point. This is usually only observed when the angle between the stable and unstable manifold is quite small however. For small angles, the distance between the homoclinic intersection points is expected to be small *compared to the noise level* (and this forces the distances between the forward and

backward iterates of the homoclinic intersection points to decay exponentially). Therefore the algorithm is just as likely to converge onto the wrong homoclinic intersection point as the right one. In terms of indistinguishability, this is nicely pictured in Figure 2.16 (right) where the standard deviation of the noise has been dropped from 1/10 to 1/50. The probability that the trajectories through each of the homoclinic intersection points cannot be distinguished drops from 0.9 to about 0.06. At this noise level, the algorithm will only rarely choose the wrong homoclinic intersection point.

For *sufficiently small* noise then, it would seem that the noise reduced trajectory should match the true trajectory with significant errors only near the initial and final points, provided that the length of trajectory taken is sufficiently long. However, if the length of the trajectory is increased to achieve this goal, then it likely that a "bad" tangency will be introduced into the trajectory. A "bad" tangency here is one for which the distance between the homoclinic intersection points is not much larger than the noise level. To deal with this tangency then, the noise level must be reduced further. It is apparent then that in order to get any *convergence* results where the noise reduced trajectory converges onto the true trajectory (except near the initial and final points) as the length of the trajectory tends to infinity, it is necessary to exclude arbitrarily "bad" tangencies. A notion that does this will be introduced in Chapter 3.

2.4   Summary

It would seem then, on the basis of these numerical experiments, that the following claims may be made:

- The errors between the original clean trajectory and the noise reduced trajectory at the initial and final points, are of the same order of magnitude as the amount of noise added.

- If the map from which the trajectory was obtained is hyperbolic, and the trajectory sufficiently long, then the errors near the initial point decay exponentially, the errors near the final point grow exponentially, and the errors in the middle of the trajectory are negligible, *except around near-tangency points*.

- Around a near-tangency point, the error distribution may assume a "spike" shape

with the errors growing exponentially, reaching a maximum at the near-tangency point, and then decaying exponentially. The errors around a near-tangency point are often of the same order of magnitude as the noise level and are due to the presence of "close" homoclinic intersection points.

- For a hyperbolic map, as the length of the trajectory goes to infinity, the initial (final) point of the noise reduced trajectory converges onto a point from the stable (unstable) manifold of the initial (final) point of the original clean trajectory.

Since in practical applications, the clean trajectory is generally unknown, it is convenient to restate the last claim in the following form:

- For a hyperbolic map, as the length of the trajectory goes to infinity, the initial (final) point of each trajectory which could have produced the noisy trajectory (assuming a bounded noise distribution say) converges onto a point from the stable (unstable) manifold of the initial (final) point of the noise reduced trajectory.

These statements are equivalent because the notion of belonging to stable and unstable manifolds is *symmetric*. That is, if p belongs to the stable manifold of q, then q belongs to the stable manifold of p. The point is that the gradient descent algorithm does yield the correct clean trajectory except near the initial point, the final point and near tangency points. As mentioned above, if arbitrarily "bad" tangencies are not present in the dynamical system of interest, then for sufficiently long trajectories and sufficiently small noise, the errors in the middle will be negligible, so noise reduction will have been achieved everywhere except near the initial and final points.

Chapter 3

Stable Manifold Theory

This chapter briefly summarises stable manifold theory and some of its various gener-alisations. These results clarify the discussion of the behaviour of the gradient descent algorithm in Chapter 2. They will also be used repeatedly in the quantitative analysis of Chapter 5. To motivate the structures relevant to the generalisations, Oseledec's Multi-plicative Ergodic Theorem is introduced. This result also gives (as a corollary) a quan-tification of the stretching induced by the linearised dynamics and is also vital for the analysis of Chapter 5. With the exception of this corollary (Proposition 3.6), proofs of all the theorems in this chapter will be omitted.

3.1 Preliminaries

In this section, a couple of the most important results that underlie the theory of non-linear analysis are reviewed. These results are used to prove almost all of the major theorems in this area. For the purposes of dynamical systems theory, the most important is the Contraction Mapping Theorem of Banach. A map f between two metric spaces (X,d) and (X',d') is said to be *Lipschitz* if there is a constant $\kappa \geqslant 0$ for which

$$d'\left(f\left(x\right),f\left(x'\right)\right) \leqslant \kappa\, d\left(x,x'\right)$$

holds for every $x \in X$ and $x' \in X'$. The least such κ is called the *Lipschitz constant* of f and is denoted by $\operatorname{Lip} f$. If $\operatorname{Lip} f < 1$, then f is called a *contraction*. A proof of the Contraction Mapping Theorem may be found in most texts on analysis ([33] for instance).

Proposition 3.1 (Contraction Mapping Theorem) *Let $f : X \to X$ be a contraction on a complete metric space X. Then, there is a unique fixed point $x_0 \in X$ for f, and for any $x \in X$, $f^n(x) \to x_0$ as $n \to \infty$.*

In the theory of dynamical systems, X is usually a function space. Often it is required to prove the existence of a particular type of function satisfying a particular relation. The idea is to set up a map whose fixed point satisfies this relation, and then show that the map is a contraction on the space of functions of the given type (for instance the space of continuous functions). This guarantees the existence of a fixed point, and therefore of the function required. The Stable Manifold Theorem (Theorem 3.3 below) is proven in this way. The difficulty with this theorem and others like it, is to set the map up in such a way that it is a contraction. Note that since any point in X converges to the fixed point under iteration by a contraction map, it follows that the fixed point must belong to every *closed subset* of X which is preserved by the contraction.

The Inverse Function Theorem is another result used often in non-linear analysis. It is in fact, also proved using a contraction mapping argument, but it will be stated as a preliminary result here for convenience. The Inverse Function Theorem is usually stated as a local result (see [19] or [46]). However, the global forms ([27, 48]) are most suitable for the linearisation theory in the following chapter. The following result is a sharpening of the usual C^r-Inverse Function Theorem. $m(T)$ denotes the minimum dilation of an invertible linear map T:

$$m(T) = \inf_{\|x\|=1} \|Tx\| = \left\| T^{-1} \right\|^{-1}.$$

Proposition 3.2 (Lipschitz Inverse Function Theorem) *Let $T : E \to E$ be linear and invertible, where E is a Banach space, and let $\varphi : E \to E$ be Lipschitz with $\mathrm{Lip}\,\varphi < m(T)$, and $\varphi(0) = 0$. Then, $T + \varphi$ is invertible, and $(T + \varphi)^{-1}$ is Lipschitz.*

3.2   Stable, Centre and Unstable Manifolds

In this section the Stable Manifold Theorem and its generalisation, the Centre Manifold Theorem are presented and discussed. These relate to the non-linear dynamics near a hyperbolic and non-hyperbolic fixed point respectively. A differentiable map $f : M \to$

M on a compact manifold M is said to have a *hyperbolic fixed point* $p \in M$ if $f(p) = p$ and $df(p)$ has no eigenvalues on the unit circle. This restriction on the eigenvalues yields a significant conceptual simplification in that the derivative has only expanding and contracting directions.

3.2.1   Hyperbolic Fixed Points

Consider an invertible linear dynamical system A on a Banach space E with no eigenvalues on the unit circle. 0 is then a hyperbolic fixed point of A. There are eigenspaces E_s and E_u of A corresponding to those eigenvalues less than and greater than unity in modulus. These spaces have the following important characterisations:

$$E_s = \{x \in E : A^n x \to 0 \quad \text{as} \quad n \to \infty\}$$

$$\text{and} \quad E_u = \{x \in E : A^{-n} x \to 0 \quad \text{as} \quad n \to \infty\},$$

and are called the stable and unstable eigenspaces of A respectively. These characterisations clearly generalise (in an obvious manner) to non-linear dynamical systems. Given a diffeomorphism f on a manifold M then, and a hyperbolic fixed point p of f, define the *stable* and *unstable* manifolds to be the sets

$$W_s = \{x \in M : f^n(x) \to p \quad \text{as} \quad n \to \infty\}$$

$$\text{and} \quad W_u = \{x \in M : f^{-n}(x) \to p \quad \text{as} \quad n \to \infty\}.$$

Such sets are non-empty, for p belongs to both by definition. Clearly, they are invariant under f. The question of what forms these sets can take, and how they are oriented around p is the province of the Stable Manifold Theorem.

Theorem 3.3 (Stable Manifold Theorem) *Let p be a hyperbolic fixed point for a diffeomorphism $f : M \to M$. Then, there are* local *stable and unstable manifolds which are smooth submanifolds of M, diffeomorphic to a disc, and are tangent at p to the stable and unstable eigenspaces of $df(p) : T_p(M) \to T_p(M)$. That is, in the notation introduced above,*

$$T_p(W_s) = E_s \qquad and \qquad T_p(W_u) = E_u.$$

The local stable and unstable manifolds satisfy

$$W_s^{local} = \bigcap_{n=0}^{\infty} f^{-n}(B_r(p)) \qquad and \qquad W_u^{local} = \bigcap_{n=0}^{\infty} f^n(B_r(p)),$$
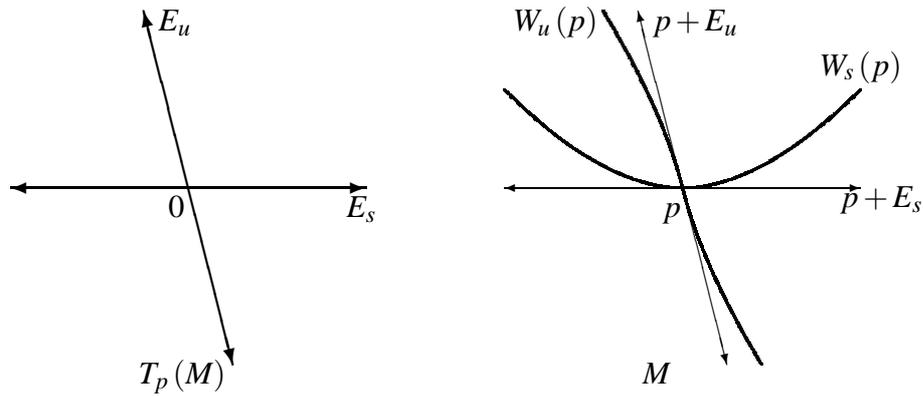
Figure 3.1: Stable and Unstable manifolds for a hyperbolic fixed point

where $B_r(p)$ is a ball of radius r about p.

This result is illustrated in Figure 3.1. It says that locally the stable manifold is a smooth manifold of the same dimension as the stable eigenspace. The global stable manifold is then constructed by taking the limit of the set of iterations of the local stable manifold under f^{-1}. Similarly, the global unstable manifold is the limit of the set of iterations of the local unstable manifold under f. These are often not genuine submanifolds of M however, as they can fold back infinitely close to themselves.

There are many proofs of the Stable Manifold Theorem. The most commonly encountered is the *graph transform* method of Hadamard. The idea here is that because the local stable manifold is tangent to the stable eigenspace at p, it can be represented (on a sufficiently small neighbourhood of p) as the graph of a function from E_s into E_u (refer again to Figure 3.1). Given any function $g : E_s \to E_u$ then, the graph transform Γ is defined in such a way that

$$\operatorname{graph}\Gamma(g) = f^{-1}(\operatorname{graph} g)$$

on the sufficiently small neighbourhood of p. It is expected that the inverse iterates will converge onto the local stable manifold, and therefore that iteration under Γ will yield a function whose graph is the local stable manifold. This is proved by showing that Γ is a contraction mapping on an appropriate function space. The smoothness of the fixed point of Γ then dictates the smoothness of the local stable manifold. Graph transform proofs of the Stable Manifold Theorem may be found in [30, 46, 48]. Other types of proof can be found in [41, 27, 40].

3.2.2 Non-hyperbolic Fixed Points

What happens when $df(p)$ has eigenvalues with unit modulus? There is an eigenspace corresponding to these eigenvalues called the *centre eigenspace*. It would seem plausible then that there must be a corresponding *centre manifold* tangent to the centre eigenspace, but it is not yet obvious how to characterise this manifold. The centre eigenspace may contain vectors that are fixed by $df(p)$ or are rotated by it. However, as it is an *algebraic* eigenspace, it can also contain vectors which are expanded by $df(p)$. For example,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} n \\ 0 \end{pmatrix}.$$

Note that the norms of the iterates of $(0,1)$ in this example increase in magnitude *polynomially* with n. In an unstable eigenspace, the norms must (eventually) increase exponentially. Similarly, in the stable eigenspace, the norms (eventually) decrease exponentially. A centre manifold is defined to be a smooth invariant manifold tangent at p to the centre eigenspace. The idea is that this set should consist of points of sub-exponential growth and decay (with respect to some metric on the manifold). However, because the manifold is assumed compact, the notion of an exponential growth rate is not well defined. Exponential decay rates *are* well defined but are not useful for characterising a centre manifold because only points on the stable and unstable manifolds are expected to decay exponentially with forward or backward iteration. It turns out anyway that centre manifolds may not be unique (see [18] for an example).

Theorem 3.4 (Centre Manifold Theorem) *Let p be a fixed point for a diffeomorphism $f : M \to M$. Then, there are local stable, centre and unstable manifolds, W_s, W_c and W_u respectively which are smooth submanifolds of M, diffeomorphic to a disc, and satisfying*

$$T_p(W_s) = E_s, \qquad T_p(W_c) = E_c \qquad and \qquad T_p(W_u) = E_u,$$

where E_s, E_c and E_u are the stable, centre and unstable eigenspaces of $df(p)$ respectively.

The local centre manifold can be extended to a global centre manifold by taking the union of its iterates under f and f^{-1}. Proofs of this theorem may be found in [1, 10, 48]. The idea here is that in the consideration of hyperbolicity, the number 1 is not all that special. For sufficiently small $\varepsilon > 0$, $df(p)$ will have no eigenvalues of modulus

$1 - \varepsilon$, and so there is a splitting into eigenspaces E_{ss} and E_{cu} whose eigenvalues have modulus less than or greater than $1 - \varepsilon$ respectively. If the unit ball is iterated forward under f, the resulting sets converge (in a manner that can be made precise) onto a set tangent to E_{cu}. This is because this eigenspace is associated with directions in which the dynamics do *not* contract, that is, the unstable and centre directions. This limiting set may not actually be a manifold (locally) but can always be decomposed into manifolds, W_{cu}, tangent at p to E_{cu}. These manifolds are called *centre-unstable manifolds* for obvious reasons. Similarly, iterating backwards gives a limiting set, W_{ss} tangent to E_{ss} called the *strong stable manifold*. For ε sufficiently small, this is the stable manifold and may be characterised by its exponential decay property. Repeating the process with $1 + \varepsilon$, gives *centre-stable manifolds*, W_{cs}, and a *strong-unstable manifold*, W_{uu}. Then, the transverse intersection of a centre-stable manifold with a centre-unstable manifold gives a centre manifold.

3.2.3 Flows

It should be noted at this point that the preceding theory (and in fact the rest of the chapter as well) applies equally well to flows as it does to diffeomorphisms, with the modification that hyperbolicity corresponds to the derivative having no eigenvalues on the imaginary axis. Although it is the stable and unstable manifolds of the diffeomorphism f which are of primary interest in the investigation of noise reduction by gradient descent, the gradient descent algorithm defines a flow, so it will be useful to keep in mind that these structures exist for flows too, especially in Chapter 5.

3.3 Global Hyperbolicity

3.3.1 The Multiplicative Ergodic Theorem

It is quite easy to extend Stable Manifold Theory to consider periodic orbits of a diffeomorphism f, essentially by considering a periodic point $p \in M$ as a fixed point of f^m (where m is the period of p). Hyperbolicity of the periodic orbit then corresponds to hyperbolicity of each p in the orbit, with respect to f^m. It is not so easy, however, to see how hyperbolicity can be defined on arbitrary trajectories. The correct notion to introduce is

that of an *average* rate of expansion or contraction. This is neatly expressed in Oseledec's Multiplicative Ergodic Theorem ([39]).

Before stating this result, the concept of ergodic theory must be introduced (briefly). Ergodic theory is a statistical theory of dynamical systems, dedicated to the study of probability measures left invariant by the action of the dynamical system (see [20, 36, 53]). A measure μ is said to be an *invariant measure* for $f : M \to M$ if

$$\mu(A) = \mu\left(f^{-1}(A)\right)$$

for every measurable set $A \subseteq M$. A measure is said to be *ergodic* if the manifold M cannot be decomposed into invariant measurable sets of positive measure. This guarantees that f does not act on (statistically significant) pieces of M in an independent manner.

Theorem 3.5 (Multiplicative Ergodic Theorem) *Let f be a diffeomorphism of a compact manifold M preserving an ergodic measure μ. Then, there is an f-invariant set $\Lambda \subseteq M$ of full measure such that for all $p \in \Lambda$, there is a splitting:*

$$T_p(M) = \bigoplus_{i=1}^{r} E_i(p),$$

varying measurably with p, invariant in that $df(p)(E_i(p)) = E_i(f(p))$, and non-negative numbers $\lambda_1 < \ldots < \lambda_r$ for which

$$\lim_{n \to \infty} \|df^n(p)x\|^{1/n} = \lambda_k \qquad \forall\, x \in \bigoplus_{i=1}^{k} E_i(p) \setminus \bigoplus_{i=1}^{k-1} E_i(p).$$

Proofs of the Multiplicative Ergodic Theorem may be found in [42], for the case of a two dimensional manifold, and [47] for the general case. The λ_i appearing in this theorem are called the *Lyapunov numbers* for f (with respect to μ). Each λ_i quantifies the average expansion and contraction for vectors in each family of spaces $\{E_i(p) : p \in \Lambda\}$. In analogy to the case of a fixed point, the spaces $E_i(p)$ will be referred to as eigenspaces. However, they are *not* generally eigenspaces of any of the derivatives $df(p)$. A more useful quantification of the expansion and contraction rates is the following.

Proposition 3.6 *Let f be a diffeomorphism on M. Then, for almost every p, and all $x \in T_p(M)$, there exists a $\lambda > 0$ such that given any $\varepsilon > 0$, there exists $C > 0$ (depending on ε but independent of x) for which*

$$\|df^n(p)x\| \leqslant C(\lambda + \varepsilon)^n \|x\|$$

for every $n \geqslant 0$.

Proof: Choose $p \in M$ so that the splitting of the Multiplicative Ergodic Theorem exists, choose $x \in T_p(M)$ non-zero, and let λ be the Lyapunov number for x. Suppose now, that the statement is false. Then, given any $C > 0$, there exist $n_1, n_2, n_3, \ldots$ such that $n_k \to \infty$ as $k \to \infty$, and $\|df^{n_k}(p)x\| \geqslant C(\lambda + \varepsilon)^{n_k} \|x\|$. Therefore,

$$\|df^{n_k}(p)x\|^{1/n_k} \geqslant C^{1/n_k}(\lambda + \varepsilon)\|x\|^{1/n_k}.$$

Since the n_k tend to ∞, it follows from Theorem 3.5 that

$$\lambda = \lim_{n\to\infty} \|df^n(p)x\|^{1/n} = \lim_{k\to\infty} \|df^{n_k}(p)x\|^{1/n_k} \geqslant \lim_{k\to\infty} C^{1/n_k}(\lambda + \varepsilon)\|x\|^{1/n_k} = \lambda + \varepsilon,$$

a contradiction for any $\varepsilon > 0$.

It remains to show that C may be chosen independent of x. Consider therefore the unit circle in each eigenspace, $S_i = E_i(p) \cap S^{d-1}$. The Lyapunov number is constant on this set, so the function $\|df^n(p)x\|(\lambda + \varepsilon)^{-n}$ is continuous on this set, and hence bounded by compactness. Let

$$K_i(\varepsilon) = \sup_{n \geqslant 0} \sup_{x \in S_i} \frac{\|df^n(p)x\|}{(\lambda + \varepsilon)^n}.$$

This is finite. It follows by linearity now, that

$$\|df^n(p)x\| \leqslant K_i(\varepsilon)(\lambda + \varepsilon)^n \|x\|$$

for all $x \in E_i$ and $n \geqslant 0$. Let P_i be the eigenprojection[1] onto the eigenspace $E_i(p)$ and λ_i be the Lyapunov number for vectors in $E_i(p)$. Then, for general x,

$$\begin{aligned} \|df^n(p)x\| &\leqslant \sum_i \|df^n(p)P_ix\| \leqslant \sum_i K_i(\varepsilon)(\lambda_i + \varepsilon)^n \|P_ix\| \\ &\leqslant \sum_i K_i(\varepsilon)(\lambda + \varepsilon)^n \|P_i\| \|x\| = \left[\sum_i K_i(\varepsilon)\|P_i\|\right](\lambda + \varepsilon)^n \|x\|, \end{aligned}$$

since λ_i, the Lyapunov number of P_ix, is less than or equal to λ, the Lyapunov number of x (whenever $P_ix \neq 0$). ∎

Using the same arguments as in the proof above, the bound

$$\|df^n(p)x\| \geqslant C'(\lambda - \varepsilon)^n \|x\| \tag{3.1}$$

may also be derived (for some $C' > 0$). Similarly, C' can be chosen independent of x *on each eigenspace*. However, this independence cannot be extended to general $x \neq 0$

---

[1]That is, the projection onto $E_i(p)$ parallel to the other $E_j(p)$, $j \neq i$.

without setting $C' = 0$. The reason for this is quite simple: The Lyapunov exponent of x with respect to f need not be related to the Lyapunov exponent of x with respect to f^{-1}. To clarify this, consider the inequality of Proposition 3.6 applied to f^{-1} in the tangent space $T_{f^n(p)}(M)$:

$$
\begin{aligned}
\left\| df^{-n}(f^n(p))x \right\| &\leqslant C \left(\lambda' + \varepsilon \right)^n \|x\| \\
\Rightarrow \qquad \|x\| &\leqslant C \left(\lambda' + \varepsilon \right)^n \| df^n(p)x \| \\
\Rightarrow \qquad \| df^n(p)x \| &\geqslant C^{-1} \left(\lambda' + \varepsilon \right)^{-n} \|x\| \\
&= C^{-1} \left((\lambda')^{-1} - \varepsilon' \right)^n \|x\|
\end{aligned}
$$

where in the second line, x was replaced by $df^n(p)x$, λ' is the Lyapunov number for x with respect to f^{-1}, and the equality in the last line holds for some $\varepsilon' > 0$. In this bound, C and hence C^{-1} can be chosen independent of x. Comparing with the bound in equation 3.1, it is clear that C' can be chosen independent of x if λ, the Lyapunov number for x with respect to f, is equal to $(\lambda')^{-1}$, the inverse of the Lyapunov number of x with respect to f^{-1}. This is true when x belongs to one of the eigenspaces, but is not true in general.

3.3.2 Hyperbolic Sets

The set Λ appearing in the statement of the Multiplicative Ergodic Theorem is said to be a *hyperbolic set* for f if none of the Lyapunov numbers are unity. The stable and unstable eigenspaces (in each tangent space $T_p(M)$) are defined to be

$$
E_s(p) = \bigoplus_{\lambda_i < 1} E_i(p) \qquad \text{and} \qquad E_u(p) = \bigoplus_{\lambda_i > 1} E_i(p).
$$

This gives a *hyperbolic splitting*, invariant under f. For such a hyperbolic splitting, the contraction-expansion estimates of Proposition 3.6 take the following form.

Proposition 3.7 *Suppose that Λ is an invariant hyperbolic set for f and that $\mu < 1 < \nu$ are chosen so that μ is greater than all the Lyapunov numbers for f less than unity, and ν is less than all the Lyapunov numbers for f greater than unity. Then, for each $p \in \Lambda$, there are $C_s \geqslant 1$ and $0 < C_u \leqslant 1$ such that*

$$
\| df^n(p)x_s \| \leqslant C_s \mu^n \|x_s\| \qquad \text{and} \qquad \| df^n(p)x_u \| \geqslant C_u \nu^n \|x_u\|,
$$

for all $x_s \in E_s(p)$ and all $x_u \in E_u(p)$.

This follows directly from Proposition 3.6 for the stable estimate, and from the discussion following it for the unstable estimate. Now, the C_s and C_u appearing in this result may be chosen independently of x_s and x_u respectively. Can they be chosen independent of p? If $\Lambda = M$, a compact manifold, then the answer is affirmative by a simple argument. However, the Multiplicative Ergodic Theorem asserts that Λ can exclude a set of measure zero from M, and hence be non-compact. It is possible then that the constants C_s and C_u may not be chosen independent of p.

It is necessary therefore to introduce compactness as an assumption. An invariant hyperbolic set Λ which is also compact is called an *invariant uniformly hyperbolic set*. For such a set, the constants C_s and C_u in Proposition 3.7 can be chosen independent of p. Another (more surprising) consequence of this compactness assumption is that the eigenspaces $E_s(p)$ and $E_u(p)$ vary *continuously* with p ([30]), rather than just measurably as claimed in Theorem 3.5. Therefore, the minimal angle between $E_s(p)$ and $E_u(p)$ is bounded away from zero, and so no exact tangencies can occur (recall the discussion of tangencies in section 2.3). The presence of tangencies therefore indicates a non-uniform structure (that is, Λ is not compact — the splitting fails at exact tangencies where the stable and unstable eigenspaces have non-trivial intersection). Obviously the tangencies can only occur at the points of M excluded from Λ — a set of measure zero. The numerical evidence of Chapter 2 therefore suggests that the Ikeda map and the Henon map are both examples of non-uniformly hyperbolic diffeomorphisms.

3.4   Generalised Stable, Centre and Unstable Manifolds

In this section the generalisation of the theory of section 3.2 to invariant uniformly hyperbolic sets is presented. This consists of taking the linear structure (guaranteed by the Multiplicative Ergodic Theorem) on the tangent spaces, and pulling it back onto the manifold, to get local stable and unstable manifolds through each point of the invariant uniformly hyperbolic set. An important question which doesn't arise in the fixed point case, is how all these stable and unstable manifolds fit together.

As in the fixed point case, vectors $x \in E_s(p)$ are eventually contracted under iteration by $df(p)$. The difference is that each iterate belongs to a different tangent space. That is, the distance between the iterates of x and the corresponding iterates of 0 (the zero of the

tangent space) contract. The corresponding notion on a manifold M (with a metric d) is therefore that the distance between iterates of x and iterates of p (under f now) contract (recall the discussion in section 3.2.2 indicating that decay is well defined on a manifold whereas growth is not). Thus, define

$$W_s(p) = \{x \in M : d(f^n(x), f^n(p)) \to 0 \quad \text{as} \quad n \to \infty\}$$
$$\text{and} \quad W_u(p) = \{x \in M : d(f^{-n}(x), f^{-n}(p)) \to 0 \quad \text{as} \quad n \to \infty\}$$

to be the *generalised stable* and *unstable manifolds of f through p* respectively. The properties of these sets are given in the following result ([23, 48]).

Theorem 3.8 (Generalised Stable Manifold Theorem) *Let $\Lambda \subseteq M$ be an invariant uniformly hyperbolic set for a diffeomorphism $f : M \to M$. Then, there are local generalised stable and unstable manifolds through each $p \in \Lambda$ which are smooth submanifolds of M, diffeomorphic to a disc, and tangent at p to the eigenspaces $E_s(p)$ and $E_u(p)$ determined in the Multiplicative Ergodic Theorem. Furthermore, the local generalised stable and unstable manifolds vary continuously with p.*

Global generalised stable and unstable manifolds can now be constructed from the local ones as

$$W_s(p) = \bigcup_{n \geqslant 0} f^{-n}\left(W_s^{\text{local}}(f^n(p))\right) \qquad \text{and} \qquad W_u(p) = \bigcup_{n \geqslant 0} f^n\left(W_u^{\text{local}}(f^{-n}(p))\right).$$

They form families invariant under f. Again, the global generalised stable and unstable manifolds are not usually submanifolds of M, but they clearly vary in a continuous manner with p also. In particular, it follows that the homoclinic intersection point closest to p (but different from p of course) also varies continuously, when it exists. Therefore, on an invariant uniformly hyperbolic set, the compactness forces the distance between p and its nearest homoclinic intersection point *to be bounded away from zero*. This is exactly the requirement suggested by the numerical experiments of section 2.3 for the noise reduction by gradient descent to work (for sufficiently small noise).

Finally, there is also a Generalised Centre Manifold Theorem ([23]) which will be used in a very special case in Chapter 5.

Theorem 3.9 (Generalised Centre Manifold Theorem) *Let* $\Lambda \subseteq M$ *be a compact invariant set for a diffeomorphism* $f : M \to M$ *such that at each* $p \in \Lambda$*, there is a continuous splitting into stable, centre and unstable directions:*

$$T_p(M) = E_s(p) \oplus E_c(p) \oplus E_u(p).$$

Then, there are local generalised stable, centre and unstable manifolds through each $p \in \Lambda$ *which are smooth submanifolds of M, diffeomorphic to a disc, and tangent at p to the stable, centre and unstable eigenspaces determined by the splitting. Furthermore, the local generalised stable and unstable manifolds vary continuously with p, and the local generalised centre manifolds, whilst not unique, may be chosen to vary continuously with p.*

Chapter 4

Linearisation Theory

4.1 Introduction

In this chapter, a generalisation, due to Kurata ([34]), of the Hartman-Grobman Theorem is presented. This result will be used in Chapter 5 to get quantitative bounds on the errors in the gradient descent algorithm for noise reduction. Kurata's proof is essentially an elegant application of the standard Hartman-Grobman Theorem on an infinite dimensional space. Therefore a detailed proof of the standard theorem is given below (Theorem 4.2). This theorem gives a qualitative correspondence between the behaviour of a non-linear system and its linearisation around a certain type of fixed point. Because quantitative information is needed for the analysis of the gradient descent algorithm, an extension of the Hartman-Grobman Theorem is also presented (Corollary 4.5). While the quantitative information provided by this extension (in the form of the *Hölder continuity* of the correspondence) is almost certainly well known to experts in the field, it does not seem to appear in the literature. Therefore this extension is the original work of the author. The same applies to Kurata's generalisation — the quantitative extension of the result (Proposition 4.17) is original work, and indeed, Kurata's proof has been modified so as to more easily accomodate this extension.

Before continuing, note that in all of the proofs in this chapter, the norms used will be *adapted* to the relevant dynamical system, or will be derived from an adapted norm. Given a hyperbolic invertible linear map T with stable and unstable eigenspaces E_s and

E_u, a norm, $\|\cdot\|$, is said to be adapted to T if

$$\|x\| = \max\left\{\|x_s\|_s, \|x_u\|_u\right\}$$

where $x = x_s + x_u$, with $x_s \in E_s$ and $x_u \in E_u$, and where the norms $\|\cdot\|_s$ on E_s and $\|\cdot\|_u$ on E_u satisfy

$$\left\|T|_{E_s}\right\|_s < 1 \quad \text{and} \quad \left\|T^{-1}|_{E_u}\right\|_u < 1$$

($g|_U$ denotes the restriction (projection) of a function g to a subset U of its domain). For proofs that such norms exist, see [46, 38]. It will be convenient to denote the restrictions of a function φ to E_s and E_u by φ_s and φ_u respectively. This convention will be used throughout this chapter. Note that for an adapted norm, the Lipschitz constant of φ is given by

$$\operatorname{Lip}\varphi = \max\left\{\operatorname{Lip}\varphi_s, \operatorname{Lip}\varphi_u\right\}.$$

It will become apparent in what follows, especially when considering Kurata's generalisation (section 4.3), that a proliferation of norms is required, most of which will be denoted by $\|\cdot\|$. It is hoped that the context is sufficient to clearly identify which norm is meant.

4.2 Fixed Points

4.2.1 The Hartman-Grobman Theorem

In this section the Hartman-Grobman Theorem for a diffeomorphism f is proven. This states that the non-linear dynamics around a hyperbolic fixed point is qualitatively similar (at least locally) to the dynamics of its linearisation about the fixed point. Specifically, it guarantees the existence of a homeomorphism h defined on a neighbourhood of the fixed point p, satisfying

$$f \circ h = h \circ df(p) \tag{4.1}$$

whenever this makes sense. A homeomorphism satisfying this relation is called a *topological conjugacy*. f and $df(p)$ are then said to be *locally topologically conjugate*. Note that h maps orbits of $df(p)$ onto orbits of f. In particular, the stable and unstable eigenspaces of $df(p)$ are mapped onto sets which must be local stable and unstable manifolds of f. In

one sense, the Hartman-Grobman Theorem gives more information than the Stable Manifold Theorem (Theorem 3.3) as it applies to a whole neighbourhood of the fixed point. However, the Stable Manifold Theorem asserts that the stable and unstable manifolds are as smooth as the diffeomorphism f, whereas the Hartman-Grobman Theorem can only assert continuity (since h is only a homeomorphism, see also section 4.2.2 below).

As with the Stable Manifold Theorem, the main part of the proof is an application of the contraction mapping theorem, and not a particularly difficult one at that. However, there is a conceptual difficulty in that the local conjugacy is not unique — essentially all that is required is to give a correspondence between *local* trajectories of the system and trajectories of its linearisation. There would seem to be an infinite number of ways that this can be done. Hence a local conjugacy cannot be exhibited as the fixed point of a contraction. But, there is a class of diffeomorphisms for which a unique *global* conjugacy can be specified. *Every* trajectory of these systems can be put into correspondence with the trajectories of their linearisations, and whilst this again might seem to be achievable in an infinite number of ways, it turns out that there is only one way that this correspondence may be constructed subject to a boundedness constraint. This conjugacy is then shown to apply to general diffeomorphisms, provided we restrict to a neighbourhood of the fixed point. The proof of the global theorem given below is due to Pugh ([43]). Since the main result (Theorem 4.2) is local, what follows can be simplified by working on a Banach space, rather than on the appropriate manifold. Note that since the non-linear behaviour is expected to be similar to the linearised behaviour near the fixed point, the conjugacy may be thought of as a perturbation of the identity. Thus the conjugacy will be written as $\mathrm{id} + h$ rather than h. Also recall from section 3.1, that $m(T)$ denotes the minimum dilation of a hyperbolic invertible linear map T.

Theorem 4.1 (Global Hartman-Grobman Theorem) *Let $T : E \to E$ be an invertible linear map on a Banach space E. Let $\varphi, \psi : E \to E$ be bounded with $\varphi(0) = 0$ and $\psi(0) = 0$, and Lipschitz with Lipschitz constant less than $\min\{1 - \|T_s\|, m(T_u) - 1, m(T)\}$. Denote the space of bounded continuous functions from E into itself (with the supremum norm) by $C_b^0(E)$. Then, there is a unique $h \in C_b^0(E)$ such that $\mathrm{id} + h$ conjugates $T + \phi$ and $T + \psi$.*

Proof: The aim is to solve the conjugacy equation

$$(T + \varphi) \circ (\mathrm{id} + h) \;=\; (\mathrm{id} + h) \circ (T + \psi) \tag{4.2}$$

$$\Rightarrow \qquad T \circ h + \varphi \circ (\mathrm{id} + h) \;=\; \psi + h \circ (T + \psi)$$

for h. The function h may be isolated in two ways from this equation:

$$h \;=\; (T \circ h - \psi + \varphi \circ (\mathrm{id} + h)) \circ (T + \psi)^{-1}$$

$$\text{or} \qquad h \;=\; T^{-1} \circ (\psi - \varphi \circ (\mathrm{id} + h) + h \circ (T + \psi)).$$

Note that $T + \psi$ is invertible by the Inverse Function Theorem (Proposition 3.2). Project the first of these equations onto E_s and the second onto E_u. This gives conspicuous factors T_s and T_u^{-1} which are both contractions. Define a map $\Theta : C_b^0(E) \to C_b^0(E)$ by

$$\Theta(h_s, h_u) \;=\; \Big((T_s \circ h_s - \psi_s + \varphi_s \circ (\mathrm{id} + h)) \circ (T + \psi)^{-1}$$

$$, T_u^{-1}(\psi_u - \varphi_u \circ (\mathrm{id} + h) + h_u \circ (T + \psi)) \Big). \tag{4.3}$$

Θ preserves $C_b^0(E)$ as it clearly maps continuous functions to continuous functions, and if h is bounded,

$$\|\Theta(h)\|_s \leqslant \|T_s\| \, \|h_s\| + \|\psi_s\| + \|\varphi_s\| \leqslant \|T_s\| \, \|h\| + \|\psi\| + \|\varphi\|$$

and similarly,

$$\|\Theta(h)\|_u \leqslant \|T_u^{-1}\| (\|\psi\| + \|\varphi\| + \|h\|),$$

so $\Theta(h)$ is bounded. Additionally, Θ is a contraction since for any $h, h' \in C_b^0(E)$,

$$\|\Theta(h) - \Theta(h')\|_s \;=\; \Big\| (T_s \circ h_s - \psi_s + \varphi_s \circ (\mathrm{id} + h)) \circ (T + \psi)^{-1}$$

$$- \big(T_s \circ h'_s - \psi_s + \varphi_s \circ (\mathrm{id} + h')\big) \circ (T + \psi)^{-1} \Big\|_s$$

$$\leqslant \Big\| T_s \circ (h_s - h'_s) \circ (T + \psi)^{-1} \Big\|_s$$

$$+ \Big\| \varphi_s \circ (\mathrm{id} + h) \circ (T + \psi)^{-1} - \varphi_s \circ (\mathrm{id} + h') \circ (T + \psi)^{-1} \Big\|_s$$

$$\leqslant \|T_s\| \, \|h_s - h'_s\|_s + \mathrm{Lip} \, \varphi_s \, \|h_s - h'_s\|_s$$

$$\leqslant (\|T_s\| + \mathrm{Lip} \, \varphi) \, \|h - h'\|$$

and similarly,

$$\|\Theta(h) - \Theta(h')\|_u \leqslant \|T_u^{-1}\| (1 + \mathrm{Lip} \, \varphi) \, \|h - h'\|.$$

Hence there is a unique $h \in C_b^0(E)$ fixed by Θ, and h therefore satisfies equation 4.2. Repeating the above construction with φ and ψ interchanged gives a unique $h' \in C_b^0(E)$ which satisfies

$$(T + \psi) \circ (\mathrm{id} + h') = (\mathrm{id} + h') \circ (T + \varphi). \qquad (4.4)$$

Using both equations 4.2 and 4.4 now gives

$$
\begin{aligned}
(T + \psi) \circ (\mathrm{id} + h') \circ (\mathrm{id} + h) &= (\mathrm{id} + h') \circ (T + \varphi) \circ (\mathrm{id} + h) \\
&= (\mathrm{id} + h') \circ (\mathrm{id} + h) \circ (T + \psi).
\end{aligned}
$$

That is, the function $(\mathrm{id} + h') \circ (\mathrm{id} + h)$ *commutes* with $T + \psi$. By what has just been proven (applied to the special case $\varphi = \psi$), there is a unique continuous function of the form $\mathrm{id} + h''$, h'' bounded, which commutes with $T + \psi$. Obviously, the identity function commutes ($h'' = 0$), and so it must follow that

$$(\mathrm{id} + h') \circ (\mathrm{id} + h) = \mathrm{id}.$$

A similar argument now shows that

$$(\mathrm{id} + h) \circ (\mathrm{id} + h') = \mathrm{id},$$

so that $\mathrm{id} + h$ is a homeomorphism. ∎

This is the global version of the theorem. In particular, it proves the existence of a unique conjugacy $\mathrm{id} + h$ with h bounded, defined everywhere, between T and $T + \varphi$, when φ is everywhere small in a Lipschitz sense. Now, any diffeomorphism f with a hyperbolic fixed point (translated to the origin for convenience) has the form $f = T + \varphi$, with $\varphi(0) = 0$ and $d\varphi(0) = 0$. Hence for any $\varepsilon > 0$, there is a neighbourhood of the origin in which $\mathrm{Lip}\,\varphi < \varepsilon$.

Theorem 4.2 (Hartman-Grobman Theorem) *Let $f : E \to E$ be a diffeomorphism on a Banach space E with a hyperbolic fixed point p. Then, there is a neighbourhood U of p and a neighbourhood V of the origin such that $f|_U$ is topologically conjugate to $df(p)|_V$.*

Proof: Without loss of generality, translate the fixed point to the origin. Given any $r > 0$, it is possible to choose an infinitely differentiable function $\gamma : \mathbb{R} \to \mathbb{R}$ such that $\gamma(x) = 1$

if $|x| < r/2$, $\gamma(x) = 0$ if $|x| > r$ and Lip γ is bounded. If $\varphi = f - df(0)$, then define a function $\tilde{\varphi} : E \to E$ by $\tilde{\varphi}(x) = \gamma(\|x\|)\varphi(x)$. It follows that $\tilde{\varphi}$ is Lipschitz since

$$
\begin{aligned}
\|\tilde{\varphi}(x) - \tilde{\varphi}(y)\| &= \|\gamma(\|x\|)\varphi(x) - \gamma(\|y\|)\varphi(y)\| \\
&\leqslant \|\gamma(\|x\|)(\varphi(x) - \varphi(y))\| + \|(\gamma(\|x\|) - \gamma(\|y\|))\varphi(y)\| \\
&\leqslant \left(\text{Lip } \varphi|_{B_r(0)} + \left\|\varphi|_{B_r(0)}\right\|\text{Lip }\gamma\right)\|x - y\|.
\end{aligned}
$$

Clearly by choosing r sufficiently small, it is possible to make

$$
\text{Lip }\tilde{\varphi} < \min\left\{1 - \|df(0)_s\|, m(df(0)_u) - 1, m(df(0))\right\}.
$$

Then, by Theorem 4.1, $df(0)$ and $df(0) + \tilde{\varphi}$ are globally topologically conjugate on E, with conjugacy h say. But, $df(0) + \tilde{\varphi} = df(0) + \varphi = f$ on $B_{r/2}(0)$, so there are neighbourhoods $U = B_{r/2}(0)$ and $V = h\left(B_{r/2}(0)\right)$ such that $df(0)|_V$ is topologically conjugate to $f|_U$ as required. ∎

Note that this proof holds for any Banach space, regardless of dimensionality. This is important as an infinite dimensional application of this theorem will be used to prove the generalisation required in section 4.3.2.

4.2.2 The Hölder Continuity of the Conjugacy

Recall that the conjugacy given by the Hartman-Grobman Theorem is only guaranteed to be continuous. This gives a qualitative correspondence between the dynamics of the system and its linearisation. To get quantitative information — analytic estimates for instance — a sharpening of the result is required. The major theorem concerning the smoothness of the conjugacy is that of Sternberg (see [37] for a detailed discussion) which says that the conjugacy will be smooth provided certain algebraic conditions on the eigenvalues of $df(p)$ are *not* satisfied. However, many important classes of dynamical systems automatically satisfy one or more of these conditions (Hamiltonian systems for instance). For general maps then, the conclusion is that the conjugacy may not be differentiable, and we must turn to weaker forms of quantitative information.

Lipschitz continuity of the conjugacy would be ideal. However, Irwin ([27]) gives a simple example of a hyperbolic system for which the conjugacy is not Lipschitz in any

neighbourhood of the fixed point. It seems that the best result we can get in this direction is that the conjugacy will be locally *Hölder continuous*. Recall that a function g is Hölder continuous if there are constants $\beta > 0$ and $0 < \alpha < 1$ such that for any x and y,

$$\|g(x) - g(y)\| \leqslant \beta \|x - y\|^{\alpha}.$$

That the conjugacy is Hölder continuous will be demonstrated by showing that the contraction map Θ of equation 4.3 preserves an appropriate space of Hölder functions. The set of all Hölder functions on a Banach space E form a linear subspace of $C_b^0(E)$. Unfortunately, this linear subspace is not usually closed, making it unsuitable for a contraction mapping argument. To see this, consider $E = [0, 1]$. If there is a function $\chi : E \rightarrow E$ which is continuous but not Hölder, then by the Weierstrass Approximation Theorem ([33]), χ is the uniform limit of polynomials, which are clearly Hölder (since E is bounded). It remains to exhibit such a function, and it is not difficult to show that $\chi : [0, 1] \rightarrow [0, 1]$ defined by

$$\chi(x) = \begin{cases} \frac{1}{1 - \ln x} & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases}$$

will do.

A subset of the Hölder functions which turns out to be useful is the set

$$\mathscr{H}_r^{\alpha,\beta}(E) = \left\{ g \in C_b^0(E) : \|g(x) - g(y)\| \leqslant \beta \|x - y\|^{\alpha} \quad \text{whenever} \quad \|x - y\| \leqslant r \right\}$$

where $0 < \alpha < 1$, $\beta > 0$ and $r > 0$ are fixed constants. This is not a linear space, but it is closed in $C_b^0(E)$.

Proposition 4.3 $\mathscr{H}_r^{\alpha,\beta}(E)$ *is a closed subset of* $C_b^0(E)$.

Proof: If $g_n \in \mathscr{H}_r^{\alpha,\beta}(E)$ for each n and $g_n \rightarrow g$ uniformly, then given any $\varepsilon > 0$, there exists n such that $\|g_n - g\| < \varepsilon$. Thus for $\|x - y\| \leqslant r$,

$$\begin{aligned} \|g(x) - g(y)\| &\leqslant \|g(x) - g_n(x)\| + \|g_n(x) - g_n(y)\| + \|g_n(y) - g(y)\| \\ &\leqslant \varepsilon + \beta \|x - y\|^{\alpha} + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $g \in \mathscr{H}_r^{\alpha,\beta}(E)$. ∎

Proposition 4.4 *Let* $\Theta : C_b^0(E) \to C_b^0(E)$ *be given by*

$$\Theta(h_s, h_u) = \left((T_s \circ h_s + \varphi_s \circ (\mathrm{id} + h)) \circ T^{-1}, T_u^{-1}(-\varphi_u \circ (\mathrm{id} + h) + h_u \circ T) \right).$$

Then, for any $\beta > 0$, Θ *preserves* $\mathcal{H}_r^{\alpha,\beta}(E)$ *for sufficiently small* α, $\mathrm{Lip}\,\varphi$ *and some* $r > 0$.

Proof: Θ is the contraction map of equation 4.3 with $\psi = 0$. It is convenient to choose $r = \beta^{1/1-\alpha}$ where α will be determined later. If $h \in \mathcal{H}_r^{\alpha,\beta}(E)$, then

$$
\begin{aligned}
\|\Theta(h)(x) - \Theta(h)(y)\|_s \;\leqslant\;& \left\| T_s h_s \left(T^{-1}x \right) - T_s h_s \left(T^{-1}y \right) \right\|_s \\
&+ \left\| \varphi_s \circ (\mathrm{id} + h) \left(T^{-1}x \right) - \varphi_s \circ (\mathrm{id} + h) \left(T^{-1}y \right) \right\|_s \\
\leqslant\;& \|T_s\| \left\| h_s \left(T^{-1}x \right) - h_s \left(T^{-1}y \right) \right\|_s \\
&+ \mathrm{Lip}\,\varphi_s \left[\left\| T^{-1}(x-y) \right\| + \left\| h \left(T^{-1}x \right) - h \left(T^{-1}x \right) \right\| \right] \\
\leqslant\;& \|T_s\| \beta \left\| T^{-1}(x-y) \right\|^\alpha \\
&+ \mathrm{Lip}\,\varphi_s \left[\left\| T^{-1} \right\| \|x-y\| + \beta \left\| T^{-1}(x-y) \right\|^\alpha \right] \\
\leqslant\;& \|T_s\| \left\| T^{-1} \right\|^\alpha \beta \|x-y\|^\alpha \\
&+ \mathrm{Lip}\,\varphi_s \left[\left\| T^{-1} \right\| \|x-y\| + \beta \left\| T^{-1} \right\|^\alpha \|x-y\|^\alpha \right].
\end{aligned}
$$

Now, r was chosen so that

$$\|x-y\| \leqslant r \quad \Rightarrow \quad \|x-y\|^{1-\alpha} \leqslant \beta \quad \Rightarrow \quad \|x-y\| \leqslant \beta \|x-y\|^\alpha.$$

It follows then that whenever $\|x-y\| \leqslant r$,

$$\|\Theta(h)(x) - \Theta(h)(y)\|_s \leqslant \left[\|T_s\| \left\| T^{-1} \right\|^\alpha + \mathrm{Lip}\,\varphi_s \left\| T^{-1} \right\| + \mathrm{Lip}\,\varphi_s \left\| T^{-1} \right\|^\alpha \right] \beta \|x-y\|^\alpha.$$

So, if α is sufficiently small and $\mathrm{Lip}\,\varphi$ is sufficiently small, the term in square brackets is less than unity. A similar calculation shows that whenever $\|x-y\| \leqslant r$,

$$\|\Theta(h)(x) - \Theta(h)(y)\|_u \leqslant \left[\left\| T_u^{-1} \right\| \left(\|T\|^\alpha + 2\mathrm{Lip}\,\varphi \right) \right] \beta \|x-y\|^\alpha.$$

Hence, $\Theta(h) \in \mathcal{H}_r^{\alpha,\beta}(E)$ for α and $\mathrm{Lip}\,\varphi$ sufficiently small. ∎

Corollary 4.5 *There are* local *topological conjugacies between a dynamical system and its linearisation about a hyperbolic fixed point that are Hölder continuous.*

Proof: Choose a neighbourhood of the fixed point for which $\operatorname{Lip} \varphi$ is sufficiently small, and $\alpha > 0$ small enough that the terms in square-brackets in the proof of Proposition 4.4 are less than unity. Set r to be the supremum of the distances between points in the chosen neighbourhood (which is finite since the neighbourhood can be taken to be bounded), and $\beta = r^{1-\alpha}$. Then, by Proposition 4.4 the contraction map of the Hartman-Grobman Theorem preserves the subset $\mathscr{H}_r^{\alpha,\beta}(E)$ which is closed in $C_b^0(E)$. Thus, the conjugacy obtained by adding the identity to the fixed point of the contraction is Hölder on the given neighbourhood (since the identity function is obviously Hölder on any bounded set, and the space of *all* Hölder functions is linear). ∎

4.3   Hyperbolic Sets

The Hartman-Grobman Theorem can be easily generalised from fixed points to hyperbolic periodic orbits, and such a treatment is found in most texts on the subject ([27, 46, 30]). The generalisation to invariant uniformly hyperbolic sets, however, appears to be quite obscure. Anosov mentions one such theorem in [5] and notes the following:

> Strangely enough this theorem has not achieved much publicity. Possibly the reason is that (by contrast with the Grobmann-Hartman theorem) it does not provide a good "model" for motions around (the hyperbolic set).

The theorem referred to by Anosov is due to Kurata ([34]), although the linearisation about hyperbolic sets had been investigated earlier in the work of Hirsch, Pugh and Shub (see [44] for instance). In the following, the theorem of Kurata is presented — it is more elementary than the results of Hirsch, Pugh and Shub, and as Irwin notes in his review ([26]), extremely elegant. This is then extended to investigate matters of Hölder continuity as was done for the fixed point case. To facilitate this, Kurata's proof will be modified somewhat. In particular, the contraction map employed will be analogous to the one used in section 4.2.1.

The proof of Kurata's theorem is phrased in the language of vector bundles, as are many proofs pertaining to hyperbolic sets — for instance, the Generalised Stable Manifold Theorem (Theorem 3.8). The appropriate concepts are reviewed now. A simple
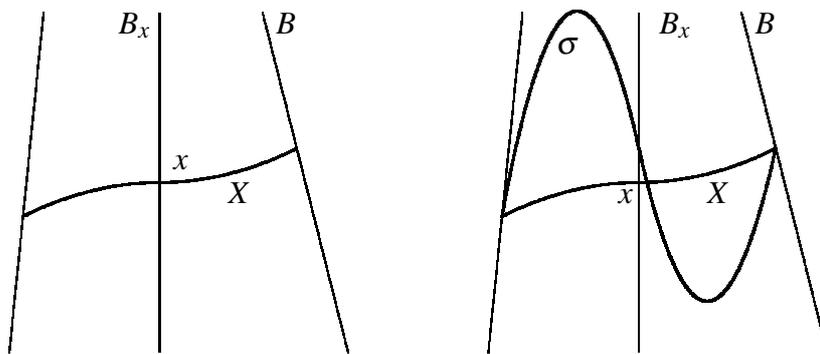
Figure 4.1: A vector bundle showing a fibre B_x and a (continuous) section σ

introduction to these concepts may be found in [9].

4.3.1 Vector Bundles

A *vector bundle* is a triple (B, π, X) where X is a topological manifold, $\pi : B \to X$ is surjective, and for each $x \in X$, $B_x = \pi^{-1}(x)$ is an n-dimensional vector space. Additionally, there is a third requirement, known as the axiom of local triviality. It states that for a sufficiently small neighbourhood U of any point $x \in X$, the restricted bundle $\pi^{-1}(U)$ can be mapped bijectively onto $U \times \mathbb{R}^d$ for some d (constant), and the bijections can be chosen to take each B_x *linearly* to $\{x\} \times \mathbb{R}^d$. The product topology on $U \times \mathbb{R}^d$ induces a topology on B by making these bijections *homeomorphisms*, and this gives the vector bundle a manifold structure (the homeomorphisms form the charts). See [9] for details. B is the *total space*, X the *base*, each B_x is a *fibre*, and π is the *projection* of the bundle. The archetypal vector bundle is of course the tangent bundle $T(X)$ of a manifold X. It is common to refer to the vector bundle as B (or sometimes π) instead of (B, π, X). Elements of the vector bundle will be denoted by u_x where the subscript indicates which fibre it belongs to. That is, $u_x \in B_x$. The projection π therefore has the action $\pi(u_x) = x$. A *section* of a vector bundle is a map $\sigma : X \to B$ taking each element of X to an element of the corresponding fibre: $\sigma(x) \in B_x$. Therefore, a section satisfies $\pi \circ \sigma = \mathrm{id}_X$. The *zero-section* is the section taking each x to 0_x. It is common to associate a section with its image just as with other functions. The zero-section is thus associated with X. A schematic picture of a vector bundle with a fibre and a section is shown in Figure 4.1.

Since every fibre of a vector bundle is a linear space, it can be given a norm. The

norm on each fibre B_x corresponding to $x \in X$ is usually denoted by $\|\cdot\|_x$. Under quite general conditions on X ([1]), these norms can be chosen so that they vary continuously with x (this gives a *Finsler* on B). An important class of maps on a vector bundle B are those which preserve the fibres, meaning that each fibre is mapped into another fibre. If a map $G : B \to B$ always takes the fibre B_x into the fibre $B_{g(x)}$ for a particular function $g : X \to X$, then G will be called a *bundle map over g*. This is, however, not quite standard terminology (see [9, 1]). Note that even though the set of maps from B to itself do not have a linear structure, the set of bundle maps over a function g do. A norm may be defined on this set by

$$\|G\| = \sup_{u_x \in B} \|G(u_x)\|_{g(x)}.$$

The set of *bounded* bundle maps over g is a Banach space.

The concept of the *exponential map* also needs to be introduced. This is a very useful map from the tangent bundle[1] $T(X)$ onto the manifold X. Its definition is rather complicated (see [9] for a discussion in terms of *sprays*), but it can be thought of as consisting of "projections", $\exp_x$ from each tangent space $T_x(X)$ onto X when X and the tangent space are embedded in Euclidean space (see Figure 4.2). The $\exp_x$ fit together smoothly to give the exponential map (denoted by exp) on the tangent bundle:

$$\exp(u_x) = \exp_x(u).$$

Each $\exp_x$ is a local diffeomorphism around 0 with $d\exp_x(0) = \mathrm{id}_{T_x(X)}$. It follows then that the map $(\pi, \exp) : T(X) \to X \times X$ is a diffeomorphism when restricted to a suitable neighbourhood of the zero-section. Since $(\pi, \exp)(0_x) = (x, x)$, the image of this neighbourhood is a neighbourhood of the diagonal in $X \times X$.

As $(\pi, \exp)$ acts on $T(X)$, it is necessary to introduce the tangent space of $T(X)$ in order to discuss its derivative. Along the zero-section, the tangent space of $T(X)$ can be decomposed into the "manifold" and "fibre" directions (Figure 4.3). Therefore,

$$T_{0_x}(T(X)) = T_x(X) \oplus T_{0_x}(T_x(X)) = T_x(X) \oplus T_x(X)$$

since $T_x(X)$ is a linear space. While these two factors are the same, it is important to remember that the first represents the direction along the zero-section (identified with X)

---

[1]Actually it is only defined on an open neighbourhood of the zero-section.
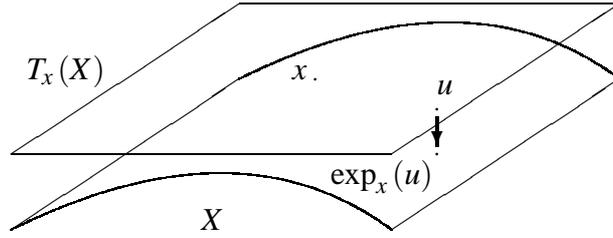
Figure 4.2: The action of an exponential map.
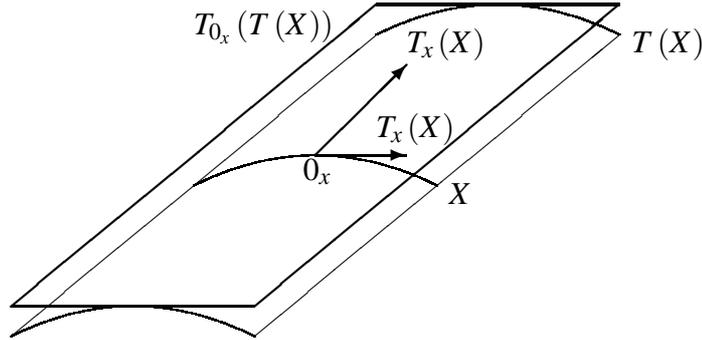
Figure 4.3: The double tangent space along the zero-section. It splits into two copies of $T_x(X)$, one tangent to the manifold X and the other tangent to (is) the fibre through x.

and the second, the direction along the fibre. This ordering will be maintained in what follows.

The derivative of $(\pi, \exp)$ on the zero-section may now be computed as follows. Since π and $\exp$ are both the identity on the zero-section ($0_x \mapsto x$), their derivatives at 0_x in the manifold direction are both the identity (on $T_x(X)$). Along the fibre, π is constant (by definition) so its derivative in the fibre direction is 0. $\exp$ however, reduces to $\exp_x$ on $T_x(X)$, and its derivative at 0 is the identity. Therefore (remembering the ordering),

$$d(\pi, \exp)(0_x) = \begin{pmatrix} \mathrm{id} & 0 \\ \mathrm{id} & \mathrm{id} \end{pmatrix} \qquad (4.5)$$

$$: T_x(X) \oplus T_x(X) = T_{0_x}(T(X)) \longrightarrow T_{(x,x)}(X \times X) = T_x(X) \oplus T_x(X).$$

4.3.2 The Generalised Hartman-Grobman Theorem

The goal of this section is to determine some kind of conjugacy between the non-linear map f and its derivatives $df(p)$ in the neighbourhood of a invariant uniformly hyperbolic

set, Λ. Specifically, a family of homeomorphisms $\{h_p : p \in \Lambda\}$ will be exhibited which vary continuously with p and satisfy the generalisation of equation 4.1:

$$\forall \, p \in \Lambda, \quad f \circ h_p = h_{f(p)} \circ df(p). \tag{4.6}$$

This will be achieved by abstracting this equation to a space where the standard fixed-point Hartman-Grobman Theorem may be applied. The bulk of the proof of the Generalised Hartman-Grobman Theorem therefore consists of building up this abstraction and then pulling it apart again to get the desired result. Note that in contrast to section 4.2, f must act on the manifold M rather than a model space E as the setting is no longer local.

The Jacobians $df(p)$ fit together to give the derivative $Tf : T(M) \to T(M)$. The derivative acts on the tangent bundle by

$$Tf(u_x) = (df(x)u)_{f(x)}.$$

That is, Tf takes a u from a tangent space $T_x(X)$, maps it to $df(u)$ and puts it in the tangent space $T_{f(x)}(X)$. Tf is therefore a bundle map over f. In order to make sense of a conjugacy between f and Tf then, it is necessary to "lift" the action of f onto the tangent bundle. Actually, the object of interest is the invariant uniformly hyperbolic set $\Lambda \subseteq M$, so the appropriate vector bundle to consider is $T_\Lambda(M)$, the bundle of tangent spaces for points in Λ.

The "lifting" of f to $T_\Lambda(M)$ is achieved using an exponential map as follows. Let U and V be neighbourhoods of the zero-section (in $T_\Lambda(M)$) and the diagonal of $\Lambda \times M$ for which $(\pi, \exp) : U \to V$ is a diffeomorphism. Define $F : U \to T_\Lambda(M)$ by

$$F = (\pi, \exp)^{-1} \circ (f \times f) \circ (\pi, \exp), \tag{4.7}$$

shrinking U if necessary to match up domains. Note that F must be defined in terms of two copies of f as the dimension of $T_\Lambda(M)$ is twice that of M. The first copy ensures that F acts along the zero-section like f (F is a bundle map over f), and the second that F acts in the fibre directions like f (distorted slightly by the exponential maps). The derivative

of F on the zero-section is therefore given by (using equation 4.5):

$$
\begin{aligned}
dF(0_x) &= \left[d(\pi, \exp)(0_{f(x)}) \right]^{-1} d(f \times f)(x, x) d(\pi, \exp)(0_x) \\
&= \begin{pmatrix} \mathrm{id} & 0 \\ \mathrm{id} & \mathrm{id} \end{pmatrix}^{-1} \begin{pmatrix} df(x) & 0 \\ 0 & df(x) \end{pmatrix} \begin{pmatrix} \mathrm{id} & 0 \\ \mathrm{id} & \mathrm{id} \end{pmatrix} \\
&= \begin{pmatrix} df(x) & 0 \\ 0 & df(x) \end{pmatrix} : T_x(M) \oplus T_x(M) \longrightarrow T_x(M) \oplus T_x(M). \quad (4.8)
\end{aligned}
$$

The idea now is to prove a bundle version of equation 4.6. That is, that there exists a homeomorphism $H : T_\Lambda(M) \to T_\Lambda(M)$ satisfying

$$
F \circ H = H \circ Tf, \qquad (4.9)
$$

where H is a bundle map formed by fitting together the h_p of equation 4.6. Since F and Tf are bundle maps over f, it is easy to see that H must be a bundle map over the identity.

Abstraction to the tangent bundle is, however, not sufficient to prove the Generalised Hartman-Grobman Theorem. Linearising about Λ corresponds to linearising about the zero-section of the tangent bundle. To abstract Λ to a fixed point (to which the standard Hartman-Grobman Theorem may be applied), the zero-section as a set is identified with the zero-section as a function — a single point in function-space. Therefore, the space $\Gamma(T_\Lambda(M))$ of bounded sections of $T_\Lambda(M)$ is introduced. This is a Banach space with the norm

$$
\|\sigma\| = \sup_{x \in \Lambda} \|\sigma(x)\|_x
$$

([1]). For simplicity, $\Gamma(T_\Lambda(M))$ will be denoted by Γ.

The link between Γ and the bundle $T_\Lambda(M)$ is provided by the observation that each $u_x \in T_\Lambda(M)$ corresponds to a bounded section $\delta_{u_x} \in \Gamma$ defined by

$$
\delta_{u_x}(x') = \begin{cases} u_x & \text{if } x = x' \\ 0_{x'} & \text{otherwise.} \end{cases}
$$

If G is a bundle map over a diffeomorphism g, then a map $G_* : \Gamma \to \Gamma$ may be defined by

$$
G_*(\sigma) = G \circ \sigma \circ g^{-1}.
$$

A quick calculation shows that if G preserves the zero-section (that is, $G(0_x) = 0_{g(x)}$ for all $x \in \Lambda$) then

$$
G(u_x) = v_{g(x)} \qquad \Rightarrow \qquad G_*(\delta_{u_x}) = \delta_{v_{g(x)}}.
$$

56

Therefore, bundle maps over a diffeomorphism which preserve the zero-section induce particularly well-behaved functions on Γ. To characterise these functions, it is necessary to know their action on arbitrary sections. This is given by

$$G_*\left(\sigma\right)\left(x\right) = G \circ \sigma \circ g^{-1}\left(x\right) = G \circ \delta_{\sigma \circ g^{-1}\left(x\right)} \circ g^{-1}\left(x\right) = G_*\left(\delta_{\sigma \circ g^{-1}\left(x\right)}\right)\left(x\right).$$

Note that if a map G_* satisfies this relation and takes any δ_{u_x} to some $\delta_{v_{g(x)}}$, then a zero-section preserving bundle map over G may be defined by

$$G\left(u_x\right) = G_*\left(\delta_{u_x}\right)\left(g\left(x\right)\right).$$

Summarising this gives the following result.

Proposition 4.6 *If $G : T_\Lambda\left(M\right) \to T_\Lambda\left(M\right)$ is a zero-section preserving bundle map over a diffeomorphism $g : \Lambda \to \Lambda$, then G induces a map $G_* : \Gamma \to \Gamma$ by $G_*\left(\sigma\right) = G \circ \sigma \circ g^{-1}$. G_* satisfies the relations:*

1. $G_*\left(\delta_{u_x}\right) = \delta_{v_{g(x)}}$

2. $G_*\left(\sigma\right)\left(x\right) = G_*\left(\delta_{\sigma \circ g^{-1}\left(x\right)}\right)\left(x\right)$.

Conversely, if $G_* : \Gamma \to \Gamma$ satisfies relations 1 and 2 for some diffeomorphism $g : \Lambda \to \Lambda$, then G_* induces a zero-section preserving bundle map over g, $G : T_\Lambda\left(M\right) \to T_\Lambda\left(M\right)$, by $G\left(u_x\right) = G_*\left(\delta_{u_x}\right)\left(g\left(x\right)\right)$.

Let $\mathscr{M}\left(\Gamma\right)$ denote the space of all maps from Γ into itself. The norm on Γ induces a norm on $\mathscr{M}\left(\Gamma\right)$ via

$$\|G_*\| = \sup_{\sigma \in \Gamma} \|G_*\left(\sigma\right)\|.$$

The set of bounded maps from Γ into itself, $\mathscr{B}\left(\Gamma\right)$, is a Banach space with this norm. Recall that the set of bounded bundle maps over a diffeomorphism g is a Banach space (section 4.3.1). The set of bounded bundle maps over g which preserve the zero-section clearly form a closed linear subspace, and hence a Banach space which will be denoted by $\mathscr{B}_g$. The set of functions in $\mathscr{M}\left(\Gamma\right)$ and $\mathscr{B}\left(\Gamma\right)$ which satisfy relations 1 and 2 (for a particular diffeomorphism g) will be denoted by $\mathscr{M}_g\left(\Gamma\right)$ and $\mathscr{B}_g\left(\Gamma\right)$ respectively. The following result completes the link between Γ and $T_\Lambda\left(M\right)$.

Proposition 4.7 *The map $* : G \mapsto G_*$ is an isometric linear isomorphism between $\mathscr{B}_g$ and $\mathscr{B}_g\left(\Gamma\right)$.*

Proof: It has already been shown that $*$ takes $\mathscr{B}_g$ into $\mathscr{M}_g(\Gamma)$, and $*$ is obviously linear. $*$ is an isometry because

$$\|G_*\| = \sup_{\sigma \in \Gamma} \sup_{x \in \Lambda} \|G_*(\sigma)(x)\| = \sup_{\sigma \in \Gamma} \sup_{x \in \Lambda} \left\|G\left(\sigma\left(g^{-1}(x)\right)\right)\right\| = \sup_{u_x \in T_\Lambda(M)} \|G(u_x)\| = \|G\|.$$

Linear isometries are always injective, so it remains to show that $*$ is surjective. Suppose then that $\mathscr{G} \in \mathscr{B}_g(\Gamma)$. Define G by $G(u_x) = \mathscr{G}(\delta_{u_x})(g(x))$. By Proposition 4.6, G is a zero-section preserving bundle map over g. G is bounded since

$$\|G(u_x)\|_{g(x)} = \|\mathscr{G}(\delta_{u_x})(g(x))\|_{g(x)} = \|\mathscr{G}(\delta_{u_x})\| \leqslant \|\mathscr{G}\|$$

for each $u_x \in T_\Lambda(M)$, so $G \in \mathscr{B}_g$. Finally,

$$G_*(\sigma)(x) = G\left(\sigma\left(g^{-1}(x)\right)\right) = \mathscr{G}\left(\delta_{\sigma(g^{-1}(x))}\right)(x) = \mathscr{G}(\sigma)(x)$$

for every $x \in \Lambda$ and $\sigma \in \Gamma$ (using relation 2 of Proposition 4.6). Thus, $G_* = \mathscr{G}$, and $*$ is invertible. ∎

As both F and Tf are (generally unbounded) zero-section preserving bundle maps over f, they induce maps F_* and Tf_* in $\mathscr{M}_f(\Gamma)$. It is easily verified that F_* and Tf_* leave the zero-section (in Γ) fixed. The standard Hartman-Grobman Theorem may therefore be applied to these functions (after some more technicalities have been dealt with) to guarantee the existence of a function H_* satisfying

$$F_* \circ H_* = H_* \circ Tf_*. \tag{4.10}$$

This H_* induces an H satisfying equation 4.9, which as remarked earlier, should be a bundle map over the identity. Therefore, H_* is expected to be an element of $\mathscr{B}_{\mathrm{id}}(\Gamma)$. However, to make H a *conjugacy*, continuity is required. Define then, for each diffeomorphism $g : \Lambda \to \Lambda$, the space $\mathscr{B}_g^0$ consisting of those bundle maps in $\mathscr{B}_g$ whose elements are continuous. Also introduce the spaces $\mathscr{M}_g^0(\Gamma)$ and $\mathscr{B}_g^0(\Gamma)$ consisting of the functions of $\mathscr{M}_g(\Gamma)$ and $\mathscr{B}_g(\Gamma)$ (respectively) which give continuous bundle maps under $*^{-1}$. H_* should therefore be an element of $\mathscr{B}_{\mathrm{id}}^0(\Gamma)$.

Lemma 4.8 $\mathscr{B}_g(\Gamma)$ *is a closed linear subspace of* $\mathscr{B}(\Gamma)$.

Proof: Linearity is obvious. Suppose $G_{n*} \in \mathscr{B}_g(\Gamma)$ and $G_{n*} \to G_*$, but $G_*(\delta_{u_x}) \neq \delta_{v_{g(x)}}$ for any $v \in T_{g(x)}(M)$. Then, $G_*(\delta_{u_x})(x') \neq 0_{x'}$ for some $x' \neq g(x)$. But,

$$\|G_{n*} - G_*\| \geqslant \|G_{n*}(\delta_{u_x})(x') - G_*(\delta_{u_x})(x')\| = \|G_*(\delta_{u_x})(x')\| \neq 0,$$

for every n. Hence $\|G_{n*} - G_*\|$ is bounded below by a positive number, a contradiction, so $G_*(\delta_{u_x}) = \delta_{v_{g(x)}}$ for some $v \in T_{g(x)}(M)$. Also,

$$G_*(\sigma)(x) = \lim_{n \to \infty} G_{n*}(\sigma)(x) = \lim_{n \to \infty} G_{n*}\left(\delta_{\sigma \circ g^{-1}(x)}\right)(x) = G_*\left(\delta_{\sigma \circ g^{-1}(x)}\right)(x),$$

so $\mathscr{B}_g(\Gamma)$ is closed. ∎

Lemma 4.9 $\mathscr{B}_g^0(\Gamma)$ *is a closed linear subspace of* $\mathscr{B}_g(\Gamma)$.

Proof: Since $\mathscr{B}_g^0(\Gamma) = *\left(\mathscr{B}_g^0\right)$ (by definition) and $*$ is an isometric linear isomorphism, $\mathscr{B}_g^0(\Gamma)$ is a closed linear subspace of $\mathscr{B}_g(\Gamma)$ if and only if $\mathscr{B}_g^0$ is a closed linear subspace of $\mathscr{B}_g$. This in turn is a consequence of the well-known result that the *uniform* limit of a sequence of continuous functions *between metric spaces* is also continuous ([45]). The details are as follows.

As M is a compact manifold, it admits a metric d — that is, there is a metric which gives the topology on M ([50]) — and Λ inherits this metric. By the axiom of local triviality (see section 4.3.1), there are open neighbourhoods U_x about each $x \in \Lambda$ and homeomorphisms $\chi_x : \pi^{-1}(U_x) \to U_x \times \mathbb{R}^d$. Each space $U_x \times \mathbb{R}^d$ is a metric space with metric D given by

$$D((x_1, y_1), (x_2, y_2)) = \max\{d(x_1, x_2), \|y_1 - y_2\|\}$$

(each $\mathbb{R}^d$ is given the norm $\|\cdot\|$). Suppose $G_n \to G$ where each $G_n \in \mathscr{B}_g^0$. Define $G'_n : U'_x \times \mathbb{R}^d \to U_{g(x)} \times \mathbb{R}^d$ by

$$G'_n = \chi_{g(x)} \circ G_n \circ \chi_x^{-1}\big|_{U'_x \times \mathbb{R}^d}$$

where $U'_x = U_x \cap g^{-1}\left(U_{g(x)}\right)$ is an open neighbourhood of x (g is continuous). Since χ_x and $\chi_{g(x)}$ are homeomorphisms, the G'_n are continuous, and converge *uniformly* (to G' say) as the convergence of the G_n is uniform. But, the G'_n act between metric spaces, so G' is continuous too. It follows that

$$\tilde{G} = \chi_{g(x)}^{-1} \circ G' \circ \chi_x\big|_{\pi^{-1}(U'_x)} : \pi^{-1}(U'_x) \to \pi^{-1}(U_{g(x)})$$

is also continuous. Finally then,

$$\tilde{G} = \lim_{n \to \infty} \chi_{g(x)}^{-1} \circ G'_n \circ \chi_x|_{\pi^{-1}(U'_x)} = \lim_{n \to \infty} G_n|_{\pi^{-1}(U'_x)} = G|_{\pi^{-1}(U'_x)}$$

so G is continuous on $\pi^{-1}(U'_x)$ for any $x \in \Lambda$. These sets cover $T_\Lambda(M)$ so $G \in \mathscr{B}_g^0$. ∎

To prove H_* exists, the contraction map of Theorem 4.1 is employed (equation 4.3). This will act on $\mathscr{B}(\Gamma)$, the space of bounded functions on Γ. However, it will also be shown to preserve the closed subspace $\mathscr{B}_{\mathrm{id}}^0(\Gamma)$, implying that the fixed point must belong to this subspace. As the contraction map will involve compositions of functions in $\mathscr{M}_g^0(\Gamma)$ and $\mathscr{B}_g^0(\Gamma)$ *for various g*, it is necessary to investigate how these spaces are related under function composition.

Lemma 4.10 *If $G_* \in \mathscr{M}_g(\Gamma)$ and $G'_* \in \mathscr{M}_{g'}(\Gamma)$, then $G_* \circ G'_* \in \mathscr{M}_{g \circ g'}(\Gamma)$. In fact,*

$$G_* \circ G'_* = (G \circ G')_*.$$

Additionally, if $G_* \in \mathscr{M}_g^0(\Gamma)$ and $G'_* \in \mathscr{M}_{g'}^0(\Gamma)$, then $G_* \circ G'_* \in \mathscr{M}_{g \circ g'}^0(\Gamma)$.

Proof: The first assertion is a simple matter of checking relations 1 and 2 of Proposition 4.6, and is easily verified. An equally simple calculation gives

$$G_* \circ G'_*(\sigma) = G_* \circ G' \circ \sigma \circ (g')^{-1} = G \circ G' \circ \sigma \circ (g')^{-1} \circ g^{-1} = (G \circ G')_*(\sigma),$$

so that $G \circ G'$ induces $G_* \circ G'_*$. Note that $G \circ G'$ is continuous if G and G' are. ∎

Lemma 4.11 *If $G_* \in \mathscr{M}_g(\Gamma)$ is invertible, then $(G_*)^{-1} \in \mathscr{M}_{g^{-1}}(\Gamma)$, G is invertible, and $(G_*)^{-1} = (G^{-1})_*$.*

Proof: Again, the first assertion is easily verified. It follows then that $(G_*)^{-1}$ induces a zero-section preserving bundle map over g^{-1}, G' say. By Lemma 4.10,

$$G_* \circ (G_*)^{-1} = \mathrm{id}_\Gamma \quad \Rightarrow \quad G \circ G' = \mathrm{id}_{T_\Lambda(M)},$$

and similarly, $G' \circ G = \mathrm{id}_{T_\Lambda(M)}$. Hence G is invertible and $G^{-1} = G'$. ∎

Clearly, both these assertions remain true when $\mathscr{M}$ is replaced by $\mathscr{B}$.

Lemma 4.12 *If $G_* \in \mathcal{M}_g^0(\Gamma)$ is invertible with a Lipschitz inverse, then $G_*^{-1} \in \mathcal{M}_{g^{-1}}^0(\Gamma)$.*
That is, G is a homeomorphism.

Proof: By Lemma 4.11, G is invertible, so it remains to demonstrate the continuity of G^{-1}. Let $B(r)$ be the ball of radius r around the zero section in $T_\Lambda(M)$:

$$B(r) = \{u_x \in T_\Lambda(M) : \|u_x\|_x \leqslant r\}.$$

Then, for $u_x \in B(r)$,

$$
\begin{aligned}
\left\|G^{-1}(u_x)\right\|_{g^{-1}(x)} &= \left\|G_*^{-1}(\delta_{u_x})(g^{-1}(x))\right\|_{g^{-1}(x)} = \left\|G_*^{-1}(\delta_{u_x})\right\| \\
&\leqslant \operatorname{Lip} G_*^{-1} \|\delta_{u_x}\| \leqslant r \operatorname{Lip} G_*^{-1}.
\end{aligned}
$$

Therefore, $G^{-1}(B(r)) \subseteq B\left(r \operatorname{Lip} G_*^{-1}\right)$, which is compact (since Λ is). As G is continuous by assumption, and invertible, it is easily seen that $G^{-1}(B(r))$ is closed, therefore compact. Consider therefore $G|_{G^{-1}(B(r))} : G^{-1}(B(r)) \to B(r)$. This is a continuous bijection from a compact set, hence a homeomorphism[2]. Therefore, G^{-1} is continuous on each $B(r)$ and since r was arbitrary, $G : T_\Lambda(M) \to T_\Lambda(M)$ is a homeomorphism. ∎

A global Hartman-Grobman Theorem can now be proved on Γ. The isomorphism $*$ can then be used to pull this back down to the tangent bundle. Finally, this is dissected fibre-wise to get the required result at each point of the hyperbolic set. First, however, a definition. Let $T : T_\Lambda(M) \to T_\Lambda(M)$ be an invertible linear bundle map over a diffeomorphism $g : M \to M$. Suppose there is a splitting of each fibre $T_x(M)$ into $T_x(M)_s \oplus T_x(M)_u$ which varies continuously with x, is invariant under T in that $T(T_x(M)_s) = T_{g(x)}(M)_s$ and $T(T_x(M)_u) = T_{g(x)}(M)_u$, and is such that

$$\|Tu_x\|_{g(x)} \leqslant \mu \|u_x\|_x \qquad \text{and} \qquad \|Tv_x\|_{g(x)} \geqslant \nu \|v_x\|_x$$

for some $\mu < 1 < \nu$ and all $u_x \in T_x(M)_s$ and $v_x \in T_x(M)_u$. Then, T is said to be hyperbolic, and $T_\Lambda(M)$ is said to have a hyperbolic splitting into $T_\Lambda(M)_s \oplus T_\Lambda(M)_u$. This splitting defines a notion of hyperbolicity on Γ. Define subspaces Γ_s and Γ_u to be the bounded

---

[2]Let $f : X \to Y$ be a continuous bijection, X compact, and Y Haussdorf. Then, if $A \subseteq X$ is open, $X \setminus A$ is closed hence compact, so $f(X \setminus A)$ is compact in Y (since f is continuous). Thus, $Y \setminus f(X \setminus A)$ is open, but this is $f(A)$ since f is a bijection. Therefore, f^{-1} is continuous.

sections of $T_x(M)$ that take values in $T_\Lambda(M)_s$ and $T_\Lambda(M)_u$ respectively. An invertible linear map $T_* : \Gamma \to \Gamma$ is then hyperbolic if

$$\|T_*\sigma_s\| \leqslant \mu \|\sigma_s\| \qquad \text{and} \qquad \|T_*\sigma_u\| \geqslant \nu \|\sigma_u\|$$

for some $\mu < 1 < \nu$ and all $\sigma_s \in \Gamma_s$ and $\sigma_u \in \Gamma_u$. It is easy to see that a hyperbolic linear bundle map T induces a hyperbolic linear section map T_*.

Theorem 4.13 *Suppose that $T_* \in \mathscr{M}_g^0(\Gamma)$ is linear, hyperbolic, and invertible. Then, there is an $\varepsilon > 0$ such that for any $\Psi_*, \Phi_* \in \mathscr{B}_g^0(\Gamma)$ with Lipschitz constants less than ε, there is a unique map $H_* \in \mathscr{B}_{\mathrm{id}}^0(\Gamma)$ satisfying*

$$(T_* + \Phi_*) \circ (\mathrm{id} + H_*) = (\mathrm{id} + H_*) \circ (T_* + \Psi_*).$$

$(\mathrm{id} + H_*)$ *is invertible, and* $(\mathrm{id} + H_*)^{-1} \in \mathscr{B}_{\mathrm{id}}^0(\Gamma)$.

Proof: This is the section map analogue of Theorem 4.1. In fact, the main contraction mapping argument is exactly the same, applied now to bounded functions on the infinite dimensional Banach space Γ. Comparing with equation 4.3, we define the map Θ by

$$\Theta(H_*) \;=\; \Big(\big(T_{*s} \circ H_{*s} - \Phi_{*s} \circ (\mathrm{id} + H_*)\big) \circ (T_* + \Psi_*)^{-1}$$
$$, T_{*u}^{-1} \big(\Psi_{*u} - \Phi_{*u} \circ (\mathrm{id} + H_*) + H_{*u} \circ (T_* + \Psi_*)\big), \Big) \quad (4.11)$$

where $T_* + \Psi_*$ is invertible for ε small enough by the Inverse Function Theorem (Theorem 3.2). The same calculations as in the proof of Theorem 4.1 show that Θ preserves $\mathscr{B}(\Gamma)$ and is a contraction on this space. Suppose now that $H_* \in \mathscr{B}_{\mathrm{id}}^0(\Gamma)$. By the *Lipschitz* Inverse Function Theorem (Theorem 3.2), $T_* + \Psi_* \in \mathscr{M}_g^0(\Gamma)$ has a Lipschitz inverse for sufficiently small ε. Thus, Lemma 4.12 implies that $(T_* + \Psi_*)^{-1} \in \mathscr{M}_{g^{-1}}^0(\Gamma)$. Similarly, $T_{*u}^{-1} \in \mathscr{M}_{g^{-1}}^0(\Gamma)$. Repeated use of Lemma 4.10 now shows that $\Theta(H_*)$ belongs to $\mathscr{M}_{\mathrm{id}}^0(\Gamma)$, hence $\mathscr{B}_{\mathrm{id}}^0(\Gamma)$ (since Θ preserves $\mathscr{B}(\Gamma)$). Since $\mathscr{B}_{\mathrm{id}}^0(\Gamma)$ is closed in $\mathscr{B}(\Gamma)$ (Lemmas 4.8 and 4.9), it follows that the fixed point of the contraction belongs to $\mathscr{B}_{\mathrm{id}}^0(\Gamma)$. Invertibility is proven as in Theorem 4.1. ∎

Theorem 4.14 (Global Generalised Hartman-Grobman Theorem) *Let $T : T_\Lambda(M) \to T_\Lambda(M)$ be a linear hyperbolic invertible bundle map over a diffeomorphism $g : \Lambda \to \Lambda$.*

Then, there is an $\varepsilon > 0$ such that for any $\Phi \in \mathscr{B}_g$ which is Lipschitz with $\mathrm{Lip}\,\Phi < \varepsilon$, there is a unique bounded zero-section preserving bundle map over the identity, H such that $\mathrm{id} + H$ conjugates T and $T + \Phi$.

Proof: T and Φ induce maps T_* and Φ_* on Γ as in Proposition 4.6. T_* is linear, hyperbolic and invertible, whilst Φ_* is bounded and it is easily verified that $\mathrm{Lip}\,\Phi_* = \mathrm{Lip}\,\Phi$. Thus the conditions of Theorem 4.13 are fulfilled, so there is a unique $H_* \in \mathscr{B}^0_{\mathrm{id}}(\Gamma)$ such that

$$T_* \circ (\mathrm{id} + H_*) = (\mathrm{id} + H_*) \circ (T_* + \Phi_*).$$

H_* therefore induces a continuous zero-section preserving bundle map over the identity, H, which satisfies

$$T \circ (\mathrm{id} + H) = (\mathrm{id} + H) \circ (T + \Phi),$$

by Lemma 4.10 (apply $*^{-1}$ to each side of the previous equation). Finally, $(\mathrm{id} + H_*)^{-1}$ induces $(\mathrm{id} + H)^{-1}$ by Lemma 4.11, and this is continuous by Theorem 4.13. Hence $\mathrm{id} + H$ is a homeomorphism. ∎

Theorem 4.15 (Generalised Hartman-Grobman Theorem) *Let $f : M \to M$ be a diffeomorphism and $\Lambda \subseteq M$ be an invariant uniformly hyperbolic set for f. Then, there is a neighbourhood U of the zero-section of $T_\Lambda(M)$ and a map h taking U homeomorphically onto a neighbourhood of the diagonal of $\Lambda \times M$, which satisfies*

$$(f \times f) \circ h = h \circ Tf$$

on $U \cap Tf(U)$.

Proof: f induces a map $F : U' \to T_\Lambda(M)$ as in equation 4.7, where U' is a neighbourhood of the zero-section. The derivative of F at 0_x is $df(x) \oplus df(x)$ by equation 4.8, so the restriction of TF to the zero-section is just Tf (in the fibre-directions). Define therefore, $\Phi = F - Tf|_{U'}$. By shrinking U' if necessary, it follows that Φ is bounded and Lipschitz with a suitably small Lipschitz constant. As in the proof of the Hartman-Grobman Theorem (Theorem 4.2), Φ is replaced by a globally defined function $\tilde{\Phi}$ which is bounded, has the same small Lipschitz constant, agrees with Φ on an open subset of U', U say, and is identically zero on the complement of U'. A global version of F is then given by

$\tilde{F} = Tf + \tilde{\Phi}$. By Theorem 4.14, there is a zero-section preserving bundle map over the identity, H, conjugating $\tilde{F}$ and Tf. Restricting this to U gives a local conjugacy between F and Tf. Finally, setting $h = (\pi, \exp) \circ H|_U$ gives

$$(f \times f) \circ h = h \circ Tf$$

as required.   ∎

Corollary 4.16 *Let $\Lambda \subseteq M$ be a invariant uniformly hyperbolic set for a diffeomorphism $f : M \to M$. Then, there is a family $\{h_p : p \in \Lambda\}$ of functions varying continuously with p, satisfying*

$$f \circ h_p = h_{f(p)} \circ df(p),$$

and such that for every $p \in \Lambda$, $h_p(p) = p$, and h_p maps some open neighbourhood of p homeomorphically onto another open neighbourhood of p.

Proof: Let π_M be the canonical projection from $\Lambda \times M$ onto M. Applying π_M to both sides of the conjugacy equation,

$$(f \times f) \circ (\pi, \exp) \circ H|_U = (\pi, \exp) \circ H|_U \circ Tf,$$

gives

$$f \circ \exp \circ H = \exp \circ H \circ Tf.$$

Restricting this to a fibre, $T_p(M)$, then gives

$$
\begin{aligned}
f \circ \exp_p \circ H_p &= \exp_{f(p)} \circ H_{f(p)} \circ df(p) \\
\Rightarrow \quad f \circ h_p &= h_{f(p)} \circ df(p)
\end{aligned}
$$

where $h_p = \exp_p \circ H_p$.   ∎

4.3.3 The Hölder Continuity of the h_p

It is now relatively simple to show that the h_p of equation 4.6 are Hölder continuous. Using exactly the same estimates as in Proposition 4.4, the fixed point, H_*, of the contraction

mapping of equation 4.11 is locally Hölder continuous: $H_* \in \mathscr{H}_r^{\alpha,\beta}(\Gamma)$ for sufficiently small exponent α. This induces the bundle map over the identity, H, and this satisfies

$$
\begin{aligned}
\left\| H(u_x) - H(u_x') \right\|_x &= \left\| H_*(\delta_{u_x})(x) - H_*(\delta_{u_x'})(x) \right\| \\
&= \left\| H_*(\delta_{u_x}) - H_*(\delta_{u_x'}) \right\| \\
&\leqslant \beta \left\| \delta_{u_x} - \delta_{u_x'} \right\|^\alpha \\
&= \beta \left\| u_x - u_x' \right\|_x^\alpha
\end{aligned}
$$

whenever $\left\| \delta_{u_x} - \delta_{u_x'} \right\| = \left\| u_x - u_x' \right\|_x \leqslant r$. Thus, the restriction of H to each fibre, $H|_{T_x(M)}$, is also locally Hölder continuous. Since each $\exp_p$ is a local diffeomorphism, it follows that $h_p = \exp_p \circ \left(\mathrm{id} + H|_{T_p(M)} \right)$ is Hölder continuous too.

Proposition 4.17 *Each h_p of Corollary 4.16 is Hölder continuous on a neighbourhood of p, and the Hölder constants and exponents may be chosen independent of p.*

Chapter 5

Analytical Results

Recall the gradient descent algorithm described in section 1.2. Given a noisy time series representing an approximation of some deterministic time series, this algorithm produces another deterministic time series which is *supposed* to be a much better approximation of the actual time series. This algorithm assumes that the dynamical system producing the time series is known. In this chapter, it is rigorously proven that the gradient descent algorithm satisfies this supposition, provided that certain conditions are met. More specifically, it is shown that *provided the time series comes from an invariant uniformly hyperbolic set, the noise is bounded, and the dynamical system induces a gradient descent satisfying Condition 5.12 below*, the approximation becomes arbitrarily good as the length of the time series tends to infinity, except near the initial and final points. This is accomplished by relating the gradient descent to its *linearisation*, for which the analysis is tractable. The transition back to the original non-linear gradient descent is then achieved using the linearisation results of Chapter 4. Note that the numerical results of section 2.3 suggest that the noise reduction will fail when the amount of noise added becomes comparable to the distance between homoclinic intersection points (at near-tangency points). As homoclinic intersections are impossible for linear systems, it seems reasonable that the noise levels which allow noise reduction correspond (approximately) to the neighbourhoods where the linearisation theory is valid! The forthcoming analysis needed for the linearised gradient descent is somewhat involved, so the case of a linear dynamical system is initially considered. First however, some general properties of gradient descent are addressed.

5.1 Gradient Descent Revisited

Suppose that $f : M \to M$ is a C^2-diffeomorphism defining a discrete dynamical system on a d-dimensional manifold M which will be assumed smooth and *compact*. Throughout this chapter, the manifold M will be locally identified with $\mathbb{R}^d$ using appropriate charts. The map induced by f on $\mathbb{R}^d$ will also be denoted by f. The determinism function $L : \mathbb{R}^{nd} \to \mathbb{R}$ may therefore be defined by equation 1.1:

$$L(x) = \frac{1}{2} \sum_{i=1}^{n-1} \|x_{i+1} - f(x_i)\|^2,$$

where $x = (x_1, \ldots, x_n)$, $x_i \in \mathbb{R}^d$, and (as before) the standard Euclidean norm is chosen on $\mathbb{R}^d$ for analytic convenience. The gradient descent algorithm then consists of solving equation 1.2:

$$\dot{x}(t) = -\nabla L(x(t)), \qquad x(0) = x,$$

where x represents the given noisy trajectory, and letting t tend to infinity. Now, $L(x) = 0$ if and only if the x_i form a deterministic trajectory for f, and clearly the deterministic trajectories are critical points of L. Conversely, by differentiating L:

$$\frac{\partial L}{\partial x_i} = \begin{cases} -df(x_1)^*(x_2 - f(x_1)) & \text{if } i = 1 \\ (x_i - f(x_{i-1})) - df(x_i)^*(x_{i+1} - f(x_i)) & \text{if } i = 2, \ldots, n-1 \\ (x_n - f(x_{n-1})) & \text{if } i = n \end{cases} \tag{5.1}$$

it is easily checked that these are the only critical points (here * denotes matrix transposition). If these critical points were *isolated*, then the gradient descent would have to converge to one of them, regardless of the initial point ([24]). However, the deterministic trajectories are not isolated — they vary continuously with the first coordinate (for instance). Therefore, more consideration is required before convergence to a deterministic trajectory can be claimed.

Choose a deterministic trajectory y. This is a fixed point of the gradient descent flow as are the other deterministic trajectories. With $q = (q_1, \ldots, q_n) \equiv \nabla L : \mathbb{R}^{nd} \to \mathbb{R}^{nd}$ defined by equation 5.1, the *linearisation* of the gradient descent flow about the fixed point y is given by

$$\dot{w}(t) = -dq(y)w(t), \qquad w(0) = x - y. \tag{5.2}$$

Since f is C^2 by assumption, $q_i(y + \delta)$, $i = 2, \ldots, n-1$, may be expanded (for δ small)

as

$$q_i(y+\delta) = (y+\delta)_i - f\left((y+\delta)_{i-1}\right) - df\left((y+\delta)_i\right)^*\left((y+\delta)_{i+1} - f\left((y+\delta)_i\right)\right)$$

$$= y_i + \delta_i - f(y_{i-1}) - df(y_{i-1})\delta_{i-1} - \ldots$$

$$-df(y_i+\delta_i)^*(y_{i+1}+\delta_{i+1} - f(y_i) - df(y_i)\delta_i - \ldots)$$

$$= \delta_i - df(y_{i-1})\delta_{i-1} - df(y_i)^*(\delta_{i+1} - df(y_i)\delta_i) + \ldots$$

where terms up to first order in δ have been retained. Similar computations when $i = 1$ and $i = n$ combine to give:

$$dq(y) = \begin{pmatrix} A_1^*A_1 & -A_1^* & & & & \\ -A_1 & I+A_2^*A_2 & -A_2^* & & & \\ & -A_2 & I+A_3^*A_3 & -A_3^* & & \\ & & \ddots & \ddots & \ddots & \\ & & & -A_{n-2} & I+A_{n-1}^*A_{n-1} & -A_{n-1}^* \\ & & & & -A_{n-1} & I \end{pmatrix} \tag{5.3}$$

where $A_i = df(y_i)$ (and all blank entries are zero). The deterministic trajectories for the linearised gradient descent are defined to be those satisfying

$$w_{i+1} = A_i w_i = df(y_i)w_i \quad \Rightarrow \quad w_{i+1} = df^i(y_1)w_1$$

for all $i = 1, \ldots, n-1$. The effect of the linearised gradient descent is now easy to determine.

Proposition 5.1 *The linearised gradient descent algorithm given by solving equation 5.2 and letting $t \to \infty$, is equivalent to projecting orthogonally onto the subspace of deterministic trajectories.*

Proof: Clearly the set of deterministic trajectories form a linear subspace. Consider $dq(y)$ defined by equation 5.3. This is a symmetric matrix (by inspection, but also because it is

a *Hessian* matrix for *L*). In fact, $dq(y)$ is positive semi-definite as

$$
\begin{aligned}
\langle w, dq(y)w \rangle &= \sum_{i,j=1}^{n} \left\langle w_i, (dq(y))_{ij} w_j \right\rangle \\
&= \langle w_1, A_1^* A_1 w_1 \rangle + \sum_{i=2}^{n-1} \langle w_i, (I + A_i^* A_i) w_i \rangle + \langle w_n, w_n \rangle \\
&\qquad - \sum_{i=1}^{n-1} \langle w_i, A_i^* w_{i+1} \rangle - \sum_{i=2}^{n} \langle w_i, A_{i-1} w_{i-1} \rangle \\
&= \sum_{i=1}^{n-1} \left[\|A_i w_i\|^2 - \langle A_i w_i, w_{i+1} \rangle \right] + \sum_{i=2}^{n} \left[\|w_i\|^2 - \langle w_i, A_{i-1} w_{i-1} \rangle \right] \\
&= \sum_{i=1}^{n-1} \left[\|A_i w_i\|^2 + \|w_{i+1}\|^2 - \langle A_i w_i, w_{i+1} \rangle - \langle w_{i+1}, A_i w_i \rangle \right] \\
&= \sum_{i=1}^{n-1} \|A_i w_i - w_{i+1}\|^2 \qquad\qquad (5.4) \\
&\geqslant 0
\end{aligned}
$$

where $\langle \cdot, \cdot \rangle$ refer to the standard inner-products[1] on $\mathbb{R}^d$ and $\mathbb{R}^{nd}$, and $A_i = df(y_i)$. The solution of the linearised gradient descent equations is

$$
w(t) = e^{-dq(y)t} w(0),
$$

and since $dq(y)$ is positive semi-definite, it follows that

$$
e^{-dq(y)t} \longrightarrow \mathscr{P}
$$

as $t \to \infty$, where $\mathscr{P}$ is the orthogonal projection onto $\ker dq(y)$. That is, $w(t) \to \hat{w} = \mathscr{P}w(0)$. It remains to show then, that $\ker dq(y)$ is the deterministic subspace. But, $\langle w, dq(y)w \rangle = 0 \iff dq(y)w = 0$ (because $dq(y)$ is positive semi-definite and therefore possesses a symmetric square root). Hence this is an easy consequence of equation 5.4. ∎

It follows from this result that for a linearised dynamical system, gradient descent finds the *closest* deterministic trajectory to the noisy one (in terms of the Euclidean norm on $\mathbb{R}^{nd}$). This is of interest in relation to the algorithm of Farmer and Sidorowich discussed in section 1.2, which aims to do exactly this, but for a non-linear system. It also follows that the linearised gradient descent flow (about a deterministic trajectory *y*) has a d-dimensional kernel, consisting of the deterministic trajectories for the linearised system,

---

[1]With the convention used in physics: $\langle x, y \rangle = x^* y$, rather than that standard in mathematics: $\langle x, y \rangle = y^* x$.

and that the dynamics orthogonal to the kernel is contracting. That is, the linearisation gives centre and stable eigenspaces. By the Centre Manifold Theorem then (Theorem 3.4), the non-linear gradient descent flow possesses centre and stable manifolds, tangent to these respective eigenspaces. The stable manifold is clearly the set of all initial conditions which give y after gradient descent. Consider the set of deterministic trajectories for the non-linear system. This is a smooth manifold as it can be represented as the graph of the smooth function

$$\left(f, f^2, \ldots, f^{n-1}\right) : \mathbb{R}^d \to \mathbb{R}^{(n-1)d}.$$

Since q is constant (zero) on this manifold, it follows that its tangent space at y is contained in the kernel of $dq(y)$, the centre eigenspace. As both these linear spaces have dimension d, they are in fact equal, and it follows that the smooth manifold of deterministic trajectories is tangent at y to the centre eigenspace. Being obviously invariant under q, the set of deterministic trajectories is therefore a centre manifold for the non-linear gradient descent! As y was an arbitrary deterministic trajectory, the set of deterministic trajectories is a centre manifold for *every* fixed point of the non-linear gradient descent, and will therefore be denoted by $\mathscr{W}_c$.

Now, note that $\mathscr{W}_c$ is closed (since L is continuous) hence compact (thinking of the set back on M). From its smoothness, it follows that the centre eigenspace at each point of $\mathscr{W}_c$ varies continuously with the point. Each stable eigenspace is the orthogonal complement of the corresponding centre eigenspace (by Proposition 5.1) so these also vary continuously with the point. Therefore, there is a continuous splitting along the compact invariant set $\mathscr{W}_c$ into stable and centre eigenspaces. By the Generalised Centre Manifold Theorem (Theorem 3.9), the generalised stable manifolds corresponding to each point in $\mathscr{W}_c$ vary continuously. These are of course just the stable manifolds for each fixed point. Because these vary continuously (on a local level at least), it follows that there is an open neighbourhood of $\mathscr{W}_c$ which is *laminated* by stable manifolds, meaning that the (disjoint) union of these stable manifolds contains the entire neighbourhood. Any point in this neighbourhood will therefore end up on $\mathscr{W}_c$ after the gradient descent algorithm has been completed. By extending this to global manifolds and making use of the compactness of M once more, it can be concluded that the global stable manifolds laminate all of M. Therefore, the following result is finally justified!

Proposition 5.2 *The gradient descent algorithm is guaranteed to converge onto a deterministic trajectory.*

As the linearised gradient descent algorithm will be the object of study in much of what follows, and as it is equivalent to projecting orthogonally (Proposition 5.1), it is convenient here to characterise the relevant projection *matrix* $\mathscr{P}$ (in terms of some basis of $\mathbb{R}^{nd}$). This projects onto the d-dimensional space of deterministic trajectories for the linearised gradient descent (about some $y \in \mathscr{W}_c$). Therefore the following representation is valid:

$$\mathscr{P} = \sum_{k=1}^{d} \frac{U_k U_k^*}{\|U_k\|^2} \tag{5.5}$$

where the

$$U_k = \begin{pmatrix} u_k \\ df(y_1) u_k \\ df^2(y_1) u_k \\ \vdots \\ df^{n-1}(y_1) u_k \end{pmatrix}, \qquad k = 1, \dots d,$$

form an orthogonal basis of the subspace of deterministic trajectories for the linearised system.

5.2   Gradient Descent for Linear Dynamical Systems

In this section the properties of the gradient descent algorithm are investigated in the special case where the dynamical map $f : \mathbb{R}^d \to \mathbb{R}^d$ is linear. For clarity, this linear map will be denoted by A rather than f. As A is a finite-dimensional linear operator, its spectrum (denoted by $\sigma(A)$) may be decomposed into three disjoint subsets: $\sigma(A) = \sigma_s(A) \cup \sigma_c(A) \cup \sigma_u(A)$ corresponding to eigenvalues of modulus less than, equal to, or greater than unity (respectively). This induces a splitting of $\mathbb{R}^d$ into stable, centre and unstable eigenspaces, denoted by E_s, E_c and E_u respectively. The eigenprojections corresponding to these eigenspaces will be denoted by P_s, P_c and P_u. These projections are not generally orthogonal.

As the derivative of a linear operator is that same linear operator, the gradient descent equations (equation 1.2, see also equation 5.1) are linear, and so by Proposition 5.1, the

gradient descent algorithm is equivalent to projecting orthogonally onto the deterministic subspace

$$\mathscr{E}_c = \left\{ y \in \mathbb{R}^{nd} : y_{i+1} = Ay_i, i = 1, \ldots, n-1 \right\}.$$

The investigation of gradient descent therefore consists of investigating the properties of this orthogonal projection, especially its relation to the non-orthogonal projections P_s, P_c and P_u. This is somewhat more difficult than it might appear, and so the analysis is initially simplified by assuming that the linear operator A is *symmetric*.

5.2.1 Symmetric Linear Systems

Suppose then that A is a symmetric linear operator on $\mathbb{R}^d$. Define $\pi_i : \mathbb{R}^{nd} \to \mathbb{R}^d$ to be the operator projecting out the ith point of a trajectory ($\pi_i x = x_i$). If $\mathscr{P}$ is the orthogonal projection effecting the gradient descent, then $\pi_i \mathscr{P}$ is the operator which gives the ith point of the noise reduced trajectory. It will be useful to consider the decomposition of these points in the stable, centre and unstable directions, and so the operators $P\pi_i \mathscr{P}$ (where P is one of the eigenprojections P_s, P_c or P_u) are the focus of what follows. Define a norm on $\mathbb{R}^{nd}$, $\|\cdot\|_\infty$, by

$$\|x\|_\infty = \max_{1 \leqslant i \leqslant n} \|x_i\|.$$

This norm reflects the noise level added to a trajectory. If x is a noisy version of a deterministic trajectory y where the noise distribution is bounded by ε say, then $\|x - y\|_\infty \leqslant \varepsilon$. This norm will be used on $\mathbb{R}^{nd}$ for the remainder of this thesis ($\|\cdot\|$ will still denote the Euclidean norm when applied to $\mathbb{R}^{nd}$ however).

Proposition 5.3 *Suppose that A is a symmetric linear operator from $\mathbb{R}^d$ into itself, with stable, centre and unstable eigenprojections P_s, P_c and P_u respectively, and $\mathscr{P}$ is the orthogonal projection in $\mathbb{R}^{nd}$ onto $\mathscr{E}_c$, the subspace of deterministic trajectories under A. Then,*

$$\|P_s \pi_i \mathscr{P}\| \leqslant \sum_{\lambda \in \sigma_s(A)} |\lambda|^{i-1} \frac{1 + |\lambda|}{1 + |\lambda|^n},$$

and the same inequality holds when P_s and $\sigma_s(A)$ are replaced by P_c and $\sigma_c(A)$ or P_u and $\sigma_u(A)$.

Proof: The norm on $P_s \pi_i \mathscr{P} : \mathbb{R}^{nd} \to \mathbb{R}^d$ is the operator norm:

$$\|P_s \pi_i \mathscr{P}\| = \sup_{x \neq 0} \frac{\|P_s \pi_i \mathscr{P} x\|}{\|x\|_\infty}.$$

Since A is symmetric, there is an orthonormal basis of $\mathbb{R}^d$ consisting of eigenvectors of A. Denote these by u_k, $k = 1, \ldots, d$, and let λ_k be the corresponding eigenvalues. The u_k define deterministic trajectories $U_k \in \mathcal{E}_c$ by

$$U_k = \begin{pmatrix} u_k \\ Au_k \\ A^2 u_k \\ \vdots \\ A^{n-1} u_k \end{pmatrix} = \begin{pmatrix} u_k \\ \lambda_k u_k \\ \lambda_k^2 u_k \\ \vdots \\ \lambda_k^{n-1} u_k \end{pmatrix}.$$

Since the $\{u_k\}$ are orthonormal,

$$\langle U_k, U_m \rangle = \sum_{i=1}^{n} \langle \lambda_k^{i-1} u_k, \lambda_m^{i-1} u_m \rangle = \begin{cases} \dfrac{|\lambda_k|^{2n} - 1}{|\lambda_k|^2 - 1} & \text{if } k = m \text{ and } |\lambda_k| \neq 1 \\ n & \text{if } k = m \text{ and } |\lambda_k| = 1 \\ 0 & \text{if } k \neq m. \end{cases} \tag{5.6}$$

So, the $\{U_k\}$ form an orthogonal basis for $\mathcal{E}_c$. Using equation 5.5 for $\mathcal{P}$, it follows that for all $x \in \mathbb{R}^{nd}$,

$$\begin{aligned} P_s \pi_i \mathcal{P} x &= P_s \pi_i \sum_{k=1}^{d} \langle U_k, x \rangle \|U_k\|^{-2} U_k \\ &= P_s \sum_{k=1}^{d} \sum_{j=1}^{n} \langle A^{j-1} u_k, x_j \rangle \frac{|\lambda_k|^2 - 1}{|\lambda_k|^{2n} - 1} A^{i-1} u_k \\ &= \sum_{u_k \in E_s} \sum_{j=1}^{n} \lambda^{j-1} \langle u_k, x_j \rangle \frac{|\lambda_k|^2 - 1}{|\lambda_k|^{2n} - 1} \lambda^{i-1} u_k \end{aligned}$$

using equation (5.6) and the fact that $P_s u_k = 0$ for all $u_k \notin E_s$. Therefore, using the Cauchy-Schwarz inequality,

$$\begin{aligned} \|P_s \pi_i \mathcal{P} x\| &\leqslant \sum_{u_k \in E_s} \sum_{j=1}^{n} \frac{1 - |\lambda_k|^2}{1 - |\lambda_k|^{2n}} |\lambda_k|^{i-1} |\lambda_k|^{j-1} \|x\|_\infty \\ &= \sum_{\lambda \in \sigma_s(A)} |\lambda|^{i-1} \frac{1 + |\lambda|}{1 + |\lambda|^n} \|x\|_\infty \end{aligned}$$

since the u_k were chosen to be normalised. The argument for the unstable case is identical, and the only difference with the centre case is that the $|\lambda_k| = 1$ part of equation 5.6 must be used. ∎

This estimate is used to prove the first result concerning the performance of the gradient descent algorithm.

Theorem 5.4 *Let A be a symmetric* hyperbolic *linear operator defining a discrete dynamical dynamical system on* $\mathbb{R}^d$, $x \in \mathbb{R}^{nd}$ *be a noisy trajectory, and* $\hat{x}$ *be the noise reduced trajectory given by the gradient descent algorithm. If the noise distribution is* bounded, *then the points of* any *deterministic trajectory that could be the true trajectory, differ from the points of* $\hat{x}$ *by an amount which tends to zero as n, the length of the trajectories, tends to infinity, except for points near the initial and final points. The errors at these points remain* bounded *as* $n \to \infty$.

Proof: Let $y \in \mathscr{E}_c$ be a candidate for the true trajectory. Then, if ε is the bound on the noise, $\|x - y\|_\infty \leqslant \varepsilon$. Define $\mu < 1 < \nu$ by

$$\mu = \max_{\lambda \in \sigma_s(A)} |\lambda| \qquad \text{and} \qquad \nu = \min_{\lambda \in \sigma_u(A)} |\lambda|.$$

Then, by Proposition 5.3, the error in the stable direction between each point of the noise reduced trajectory, $\hat{x}$, and y is given by

$$
\begin{aligned}
\|P_s \pi_i (\hat{x} - y)\| &= \|P_s \pi_i \mathscr{P} (x - y)\| \\
&\leqslant \sum_{\lambda \in \sigma_s(A)} |\lambda|^{i-1} \frac{1 + |\lambda|}{1 + |\lambda|^n} \|x - y\|_\infty \\
&\leqslant \dim E_s \mu^{i-1} (1 + \mu) \varepsilon.
\end{aligned}
$$

The unstable error is given by

$$
\begin{aligned}
\|P_u \pi_i (\hat{x} - y)\| &= \|P_u \pi_i \mathscr{P} (x - y)\| \\
&\leqslant \sum_{\lambda \in \sigma_u(A)} |\lambda|^{i-1} \frac{1 + |\lambda|}{1 + |\lambda|^n} \|x - y\|_\infty \\
&= \sum_{\lambda \in \sigma_u(A)} |\lambda|^{-(n-i)} \frac{|\lambda|^{n-1} + |\lambda|^n}{1 + |\lambda|^n} \|x - y\|_\infty \\
&= \sum_{\lambda \in \sigma_u(A)} |\lambda|^{-(n-i)} \left(1 + \frac{|\lambda|^{n-1} - 1}{|\lambda|^n + 1} \right) \|x - y\|_\infty \\
&\leqslant \dim E_u \nu^{-(n-i)} (1 + \nu^{-1}) \varepsilon.
\end{aligned}
$$

Now, ε, μ, ν, $\dim E_s$ and $\dim E_u$ are all constant, so the errors at point i (between $\hat{x}$ and y) decrease exponentially in the stable direction as i increases, and they are bounded above by $\dim E_s (1 + \mu) \varepsilon$ since $\mu < 1$. Similarly, the errors at point i increase exponentially in the unstable direction as i increases, but they are bounded above by $\dim E_u (1 + \nu^{-1}) \varepsilon$ since $\nu > 1$. As these bounds are independent of n, the length of the trajectory, it follows that the errors in both directions will be simultaneously arbitrarily small for a long

enough trajectory, except near the initial point where the stable error may be of the order of magnitude of the noise bound, and near the final point where the unstable error may be of the order of magnitude of the noise bound. Finally, since A is assumed hyperbolic (and symmetric), the total error at each point is the Pythagorean sum of the errors in the stable and unstable directions, so the total errors can be made arbitrarily small by taking a long enough trajectory, except near the initial and final points! $\blacksquare$

Note that from Proposition 5.3, the bound on $\|P_c \pi_i \mathscr{P}\|$ is independent of i (it reduces to $\dim E_c$). Therefore, when considering the generalisation of Theorem 5.4 to non-hyperbolic symmetric linear operators, one finds that the bounds in the stable and unstable directions behave as before, but the bound in the centre direction remains constant, and hence the bound on the total error does *not* become arbitrarily small anywhere as the length of the trajectory is increased. That is, the errors between that noise reduced trajectory $\hat{x}$ and the true trajectory y can be large, and the gradient descent algorithm fails. This is in accordance with the numerical results of Chapter 2. In what follows, attention will be restricted to hyperbolic systems.

It is easy to see how to generalise Theorem 5.4 to unbounded noise distributions. In this case, there is no strict bound on $\|x - y\|_\infty$ and the probability that a given noise realisation will exceed any bound imposed (as the length of the realisation goes to infinity) is 1. Instead, one considers some sort of average error at each point of the trajectory. This then relates to the corresponding average for the noise distribution. For instance, the *root-mean-square* errors at each point are bounded by the same expressions appearing in the proof of Theorem 5.4 with ε replaced by the *standard deviation* of the noise distribution. This obviously extends to confidence levels. The conclusion is then that with probability $p < 1$, the given bound holds with the appropriate ε. Clearly for an unbounded noise distribution, $\varepsilon \to \infty$ as $p \to 1^-$.

5.2.2 Hyperbolic Linear Systems

Before generalising the results of section 5.2.1, it is necessary to know how general linear operators expand and contract vectors. For symmetric linear operators, the existence of a basis of eigenvectors implies that these (and hence general vectors) are expanded or

contracted *exponentially* under iteration. However, this is not quite true for generalised eigenvectors — an example where a vector is expanded *polynomially* under iteration was given in section 3.2.2. A better estimate is provided by the results of section 3.3, especially Proposition 3.7 (adapted to the case where $f = A$ and hence $df^n(p) = A^n$). It is easy to see that the Lyapunov numbers for A are just the moduli of the eigenvalues of A. By Proposition 3.7 then, given $\mu < 1 < \nu$ such that

$$\mu > \max_{\lambda \in \sigma_s(A)} |\lambda| \qquad \text{and} \qquad \nu < \min_{\lambda \in \sigma_u(A)} |\lambda|,$$

there exist $C_s \geqslant 1$ and $0 < C_u \leqslant 1$ such that

$$\|A^n v_s\| \leqslant C_s \mu^n \|v_s\| \qquad \text{and} \qquad \|A^n v_u\| \geqslant C_u \nu^n \|v_u\|, \tag{5.7}$$

for all $v_s \in E_s$ and $v_u \in E_u$. μ and ν will be referred to as the *hyperbolicity bounds* for A.

In this section, the generalisation of Theorem 5.4 to a general hyperbolic linear operator, A, is proven. Because these operators admit non-orthogonal eigenspaces (and generalised eigenspaces), the analysis is more complicated. However, the idea of the proof remains the same: The orthogonal projection $\mathscr{P}$ onto $\mathscr{E}_c$ is decomposed into a sum of terms involving vectors (in $\mathscr{E}_c \subset \mathbb{R}^{nd}$) whose magnitudes can be related to the expansion and contraction rates of A. The following computation will also be required.

Lemma 5.5 *Suppose that a_j is a sequence of non-negative numbers satisfying*

$$a_j \leqslant C \kappa^{j-i} a_i$$

for all $j \geqslant i \geqslant 1$ where $a_1 > 0$, $0 \leqslant \kappa < 1$ and $C > 0$ are constants. Then,

$$\frac{\left(\sum_{j=1}^n a_j\right)^2}{\sum_{j=1}^n a_j^2} \leqslant \frac{1 + (2C - 1)\kappa}{1 - \kappa}.$$

Proof: Expand the numerator as

$$\left(\sum_{j=1}^n a_j\right)^2 = \sum_{j=1}^n a_j^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_i a_j.$$

By hypothesis,

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n a_i a_j \leqslant \sum_{i=1}^{n-1} \sum_{j=i+1}^n C \kappa^{j-i} a_i^2 = \frac{C\kappa}{1 - \kappa} \sum_{i=1}^{n-1} a_i^2 \left(1 - \kappa^{n-i}\right) \leqslant \frac{C\kappa}{1 - \kappa} \sum_{i=1}^n a_i^2,$$

77

which yields the estimate:

$$\frac{\left(\sum_{j=1}^n a_j\right)^2}{\sum_{j=1}^n a_j^2} = 1 + 2\frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^n a_i a_j}{\sum_{j=1}^n a_j^2} \leqslant \frac{1 + (2C - 1)\,\kappa}{1 - \kappa},$$

as required. ∎

Because of the presence of non-orthogonal eigenspaces, it will be convenient to consider the *minimal angle* between subspaces. For two subspaces E and E' of a Euclidean space, the minimal angle θ is defined to be the acute angle satisfying

$$\cos\theta = \sup\left\{\frac{\langle x, x'\rangle}{\|x\|\,\|x'\|} : x \in E \setminus \{0\} \text{ and } x' \in E' \setminus \{0\}\right\}.$$

Proposition 5.6 *Suppose that A is a hyperbolic linear operator from $\mathbb{R}^d$ into itself, with stable and unstable projections P_s and P_u respectively, and $\mathscr{P}$ is the orthogonal projection in $\mathbb{R}^{nd}$ onto $\mathscr{E}_c$, the subspace of deterministic trajectories for A. Then, the following bounds hold:*

$$\|P_s \pi_i \mathscr{P}\| \leqslant \dim E_s \frac{C_s \mu^{i-1}}{\sin\phi}\left(\frac{\sqrt{1 + (2C_s - 1)\,\mu}}{\sin\phi\,\sqrt{1-\mu}} + \frac{\sqrt{1 + (2C_u^{-1} - 1)\,v^{-1}}}{\tan\phi\,\sqrt{1-v^{-1}}}\right)$$

$$\|P_u \pi_i \mathscr{P}\| \leqslant \dim E_u \frac{C_u^{-1} v^{-(n-i)}}{\sin\phi}\left(\frac{\sqrt{1 + (2C_u^{-1} - 1)\,v^{-1}}}{\sin\phi\,\sqrt{1-v^{-1}}} + \frac{\sqrt{1 + (2C_s - 1)\,\mu}}{\tan\phi\,\sqrt{1-\mu}}\right)$$

where $\mu < 1 < v$ are hyperbolicity bounds for A, C_s and C_u are the associated constants (see equations 5.7), and ϕ is the minimal angle between E_s and E_u.

Proof: Let $\mathfrak{E}_s$ and $\mathfrak{E}_u$ be the deterministic trajectories whose points are in E_s and E_u respectively. That is, let

$$\mathfrak{E}_s = \left\{\begin{pmatrix} v \\ Av \\ A^2 v \\ \dots \\ A^{n-1}v \end{pmatrix} : v \in E_s\right\} \quad \text{and} \quad \mathfrak{E}_u = \left\{\begin{pmatrix} v \\ Av \\ A^2 v \\ \dots \\ A^{n-1}v \end{pmatrix} : v \in E_u\right\}.$$

Since A is hyperbolic, $\mathbb{R}^d = E_s \oplus E_u$, and this induces the decomposition $\mathscr{E}_c = \mathfrak{E}_s \oplus \mathfrak{E}_u$. If $\pi(E, E')$ denotes the projection onto the subspace E parallel to the subspace E'[2], then $\mathscr{P}$

---

[2] That is, $\pi(E, E')$ is the unique projection with image E and kernel E'.

may be decomposed as

$$\mathscr{P} = \pi\left(\mathfrak{E}_s, \mathfrak{E}_s^{\perp}\right) + \pi\left(\mathfrak{E}_s^{\perp}, \mathfrak{E}_s\right) \tag{5.8}$$

where $^{\perp}$ denotes orthogonal complementation. This obviously corresponds to a decomposition of $\mathscr{E}_c$ into $\mathfrak{E}_s \oplus \mathfrak{E}_s^{\perp}$, so the idea is to rewrite $\pi\left(\mathfrak{E}_s^{\perp}, \mathfrak{E}_s\right)$ so that it involves $\mathfrak{E}_u$. The constructions which achieve this are indicated schematically in Figure 5.1 for convenience.

Suppose then, that $\{W_k\}$ constitute an orthogonal basis for $\mathfrak{E}_s^{\perp}$. Since $\dim\mathfrak{E}_s^{\perp} = \dim\mathfrak{E}_u$, it follows that these linear spaces are isomorphic, and a convenient isomorphism is given by

$$\pi\left(\mathfrak{E}_u, \mathfrak{E}_s\right)\big|_{\mathfrak{E}_s^{\perp}},$$

(the projection onto $\mathfrak{E}_u$ parallel to $\mathfrak{E}_s$, *restricted* to $\mathfrak{E}_s^{\perp}$). This defines a (non-orthogonal) basis $\{U_k\}$ by $U_k = \pi\left(\mathfrak{E}_u, \mathfrak{E}_s\right) W_k$. However, it is clear that

$$\left[\pi\left(\mathfrak{E}_u, \mathfrak{E}_s\right)\big|_{\mathfrak{E}_s^{\perp}}\right]^{-1} = \pi\left(\mathfrak{E}_s^{\perp}, \mathfrak{E}_s\right)\Big|_{\mathfrak{E}_u},$$

the restriction of an *orthogonal* projection, so it follows that U_k has the orthogonal decomposition $U_k = W_k + V_k$, where $V_k = \pi\left(\mathfrak{E}_s, \mathfrak{E}_s^{\perp}\right) U_k \in \mathfrak{E}_s$. Note that the $\{V_k\}$ need not form a basis for $\mathfrak{E}_s$ — for instance, some (or even all) of the V_k may be zero. Now, if θ_k is the acute angle between U_k and V_k (defined to be $\pi/2$ if $V_k = 0$), then

$$\|W_k\| = \|U_k\|\sin\theta_k = \|V_k\|\tan\theta_k \tag{5.9}$$

(where the norm is the standard Euclidean norm on $\mathbb{R}^{nd}$).

As $\{W_k\}$ is an orthogonal basis for $\mathfrak{E}_s^{\perp}$ and $W_k = U_k - V_k$, the second term in the decomposition of $\mathscr{P}$ (equation 5.8) may be expanded as

$$\pi\left(\mathfrak{E}_s^{\perp}, \mathfrak{E}_s\right) = \sum_{k=1}^{d_u} \frac{W_k W_k^*}{\|W_k\|^2} = \sum_{k=1}^{d_u} \|W_k\|^{-2}\left(U_k U_k^* - U_k V_k^* - V_k U_k^* + V_k V_k^*\right),$$

where $d_u = \dim\mathfrak{E}_s^{\perp} = \dim\mathfrak{E}_u = \dim E_u$. Since $P_u \pi_i\left(\mathfrak{E}_s\right) = P_u\left(E_s\right) = \{0\}$, the first term in
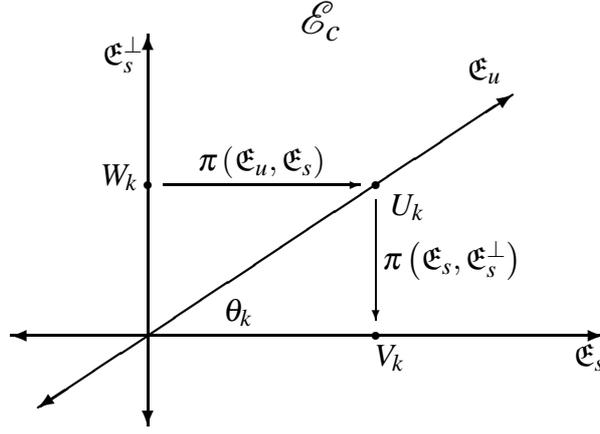
Figure 5.1: Construction of "unstable trajectories", $\{U_k\}$, and "stable trajectories", $\{V_k\}$, from the basis $\{W_k\}$ of $\mathfrak{E}_s^\perp$.

the decomposition of $\mathscr{P}$ (equation 5.8) is annihilated by $P_u \pi_i$. Therefore, for $x \in \mathbb{R}^{nd}$,

$$
\begin{aligned}
P_u \pi_i \mathscr{P} x &= P_u \pi_i \pi \left(\mathfrak{E}_s^\perp, \mathfrak{E}_s \right) x \\
&= P_u \pi_i \sum_{k=1}^{d_u} \|W_k\|^{-2} \left(U_k U_k^* - U_k V_k^* - V_k U_k^* + V_k V_k^* \right) x \\
&= P_u \sum_{k=1}^{d_u} \|W_k\|^{-2} \big(A^{i-1} u_k \langle U_k, x \rangle - A^{i-1} u_k \langle V_k, x \rangle \\
&\qquad\qquad - A^{i-1} v_k \langle U_k, x \rangle + A^{i-1} v_k \langle V_k, x \rangle \big) \\
&= \sum_{k=1}^{d_u} \|W_k\|^{-2} A^{i-1} u_k \sum_{j=1}^{n} \left(\langle A^{j-1} u_k, x_j \rangle - \langle A^{j-1} v_k, x_j \rangle \right),
\end{aligned}
$$

where $u_k = \pi_1 U_k \in E_u$ and $v_k = \pi_1 V_k \in E_s$. This expresses $\mathscr{P}$ in terms of vectors from the stable and unstable eigenspaces of A. Using equations 5.9 and the Cauchy-Schwarz inequality, this gives the bounds:

$$
\begin{aligned}
\|P_u \pi_i \mathscr{P} x\| &\leqslant \sum_{k=1}^{d_u} \frac{\|A^{i-1} u_k\|}{\|W_k\|^2} \sum_{j=1}^{n} \left(\|A^{j-1} u_k\| + \|A^{j-1} v_k\| \right) \|x_j\| \\
&\leqslant \sum_{k=1}^{d_u} \frac{\|A^{i-1} u_k\|}{\|W_k\|} \sum_{j=1}^{n} \left(\frac{\|A^{j-1} u_k\|}{\|W_k\|} + \frac{\|A^{j-1} v_k\|}{\|W_k\|} \right) \|x\|_\infty \\
&= \sum_{k=1}^{d_u} \frac{\|A^{i-1} u_k\|}{\|U_k\| \sin \theta_k} \sum_{j=1}^{n} \left(\frac{\|A^{j-1} u_k\|}{\|U_k\| \sin \theta_k} + \frac{\|A^{j-1} v_k\|}{\|V_k\| \tan \theta_k} \right) \|x\|_\infty . \quad (5.10)
\end{aligned}
$$

In the case where any of the V_k are zero, the corresponding v_k are zero, and so the second term in the parentheses above is zero.

80

Consider now the term

$$\sum_{j=1}^{n} \frac{\left\|A^{j-1}v_k\right\|}{\|V_k\|} = \frac{\sum_{j=1}^{n}\left\|A^{j-1}v_k\right\|}{\left[\sum_{j=1}^{n}\left\|A^{j-1}v_k\right\|^2\right]^{1/2}}.$$

If $a_j = \left\|A^{j-1}v_k\right\|$, then by equation 5.7,

$$a_j = \left\|A^{j-1}v_k\right\| \leqslant C_s\mu^{j-i}\left\|A^{i-1}v_k\right\| = C_s\mu^{j-i}a_i$$

where $\mu < 1$ is a (stable) hyperbolicity bound for A, and $C_s \geqslant 1$ is the associated constant. By Lemma 5.5 then,

$$\sum_{j=1}^{n} \frac{\left\|A^{j-1}v_k\right\|}{\|V_k\|} = \left[\frac{\left(\sum_{j=1}^{n}\left\|A^{j-1}v_k\right\|\right)^2}{\sum_{j=1}^{n}\left\|A^{j-1}v_k\right\|^2}\right]^{1/2} \leqslant \left[\frac{1+(2C_s-1)\mu}{1-\mu}\right]^{1/2} \tag{5.11}$$

Similarly, if $a_j = \left\|A^{n-j}u_k\right\|$, then $a_j \leqslant C_u^{-1}v^{-(j-i)}a_i$ where $v > 1$ is an (unstable) hyperbolicity bound for A and $C_u \leqslant 1$ is the associated constant. Therefore,

$$\sum_{j=1}^{n} \frac{\left\|A^{j-1}u_k\right\|}{\|U_k\|} \leqslant \left[\frac{1+\left(2C_u^{-1}-1\right)v^{-1}}{1-v^{-1}}\right]^{1/2} \tag{5.12}$$

One last application of equation 5.7 and Lemma 5.5 gives

$$\frac{\left\|A^{i-1}u_k\right\|}{\|U_k\|} \leqslant \frac{\left\|A^{i-1}u_k\right\|}{\left\|A^{n-1}u_k\right\|} \leqslant C_u^{-1}v^{-(n-i)},$$

so combining this with equations 5.11, 5.12 and 5.10, yields the estimate

$$\|P_u\pi_i\mathscr{P}\| \leqslant \sum_{k=1}^{d_u} \frac{C_u^{-1}v^{-(n-i)}}{\sin\theta_k}\left(\frac{\sqrt{1+\left(2C_u^{-1}-1\right)v^{-1}}}{\sin\theta_k\sqrt{1-v^{-1}}} + \frac{\sqrt{1+(2C_s-1)\mu}}{\tan\theta_k\sqrt{1-\mu}}\right).$$

This bound expresses the norm of $P_u\pi_i\mathscr{P}$ in terms of the constants μ, v, C_s, C_u and d_u — which depend on the hyperbolic linear operator A and not on the length of the trajectory n — and the angles θ_k. As the θ_k are angles between the *trajectories* U_k and V_k, they will generally vary with n. It remains then to show that they are bounded away from zero, so that $\sin\theta_k$ and $\tan\theta_k$ do not vanish as n tends to infinity. If ϕ is the minimal angle between the eigenspaces E_s and E_u (which only depends on A), then using the Cauchy-Schwarz

inequality for sums,

$$
\begin{aligned}
|\cos\theta_k| &= \frac{|\langle U_k, V_k\rangle|}{\|U_k\|\,\|V_k\|} \\
&= \frac{\left|\sum_{j=1}^{n}\langle A^{j-1}u_k, A^{j-1}v_k\rangle\right|}{\left[\sum_{j=1}^{n}\|A^{j-1}u_k\|^2\right]^{1/2}\left[\sum_{j=1}^{n}\|A^{j-1}v_k\|^2\right]^{1/2}} \\
&\leqslant \frac{\sum_{j=1}^{n}\|A^{j-1}u_k\|\,\|A^{j-1}v_k\|\cos\phi}{\left[\sum_{j=1}^{n}\|A^{j-1}u_k\|^2\right]^{1/2}\left[\sum_{j=1}^{n}\|A^{j-1}v_k\|^2\right]^{1/2}} \\
&\leqslant \frac{\left[\sum_{j=1}^{n}\|A^{j-1}u_k\|^2\right]^{1/2}\left[\sum_{j=1}^{n}\|A^{j-1}v_k\|^2\right]^{1/2}}{\left[\sum_{j=1}^{n}\|A^{j-1}u_k\|^2\right]^{1/2}\left[\sum_{j=1}^{n}\|A^{j-1}v_k\|^2\right]^{1/2}}\cos\phi \\
&= \cos\phi,
\end{aligned}
$$

since $A^{j-1}u_k \in E_u$ and $A^{j-1}v_k \in E_s$. Therefore, $\sin\theta_k \geqslant \sin\phi$ and $\tan\theta_k \geqslant \tan\phi$, so substitution gives the required unstable bound. The stable bound (for $P_s\pi_i\mathscr{P}$) is derived using the same technique, with s and u interchanged. ∎

Theorem 5.7 *Let A be a hyperbolic linear operator defining a discrete dynamical dynamical system on $\mathbb{R}^d$, $x \in \mathbb{R}^{nd}$ be a noisy trajectory, and $\hat{x}$ be the noise reduced trajectory given by the gradient descent algorithm. If the noise distribution is* bounded*, then the points of* any *deterministic trajectory that could be the true trajectory, differ from the points of $\hat{x}$ by an amount which tends to zero as n, the length of the trajectories, tends to infinity, except for points near the initial and final points. The errors at these points remain* bounded *as $n \to \infty$.*

Proof: If y is a candidate for the true trajectory and ε is the bound on the noise distribution, then $\|x - y\|_\infty \leqslant \varepsilon$. It follows now from Proposition 5.6, that

$$
\|P_s\pi_i(\hat{x}-y)\| = \|P_s\pi_i\mathscr{P}(x-y)\| \leqslant K_s\mu^{i-1}\|x-y\|_\infty \leqslant K_s\mu^{i-1}\varepsilon
$$

where K_s is a constant independent of i or the length of the trajectory n. Similarly,

$$
\|P_u\pi_i(\hat{x}-y)\| \leqslant K_u\nu^{-(n-i)}\varepsilon
$$

where K_u is also independent of i and n. The result now follows from the same argument as was used in the proof of Theorem 5.4. ∎

It is easy to see now that as the errors in the stable direction decay exponentially to zero as i increases, the difference between the final point of the noise reduced trajectory $\hat{x}$ and the final point of a candidate true trajectory y, *must converge onto the unstable eigenspace* as n tends to infinity. Another way of saying this is to consider the generalised unstable eigenspace through y (which is just the unstable eigenspace through 0 translated to y). Then, the final point of the noise reduced trajectory $\hat{x}_n$ can be made *arbitrarily close* to the generalised unstable eigenspace of the final point of the true trajectory y_n, by taking n sufficiently large. A similar statement is true for the first point and the generalised stable eigenspace. As this proves the conjecture of Judd and Smith mentioned at the end of section 1.3 (for hyperbolic linear dynamical systems and bounded noise), it is stated as a result.

Corollary 5.8 *If A is a hyperbolic linear operator defining a discrete dynamical dynamical system on $\mathbb{R}^d$, x a noisy trajectory of the system (derived from the true trajectory y and a bounded noise distribution), and $\hat{x}$ the noise reduced trajectory given by applying the gradient descent algorithm to x, then by taking n, the length of the trajectory, sufficiently large, $\hat{x}_1$ may be made arbitrarily close to the generalised stable eigenspace of y_1 and $\hat{x}_n$ may be made arbitrarily close to the generalised unstable eigenspace of y_n.*

Note that by the bounds of Proposition 5.6, the closer μ or ν are to unity, the larger n needs to be to make the errors near the middle of the trajectory small. The bounds also suggest that the errors at the initial and final points might be *larger* for systems with μ or ν close to unity, and for systems where the angle between the stable and unstable eigenspace is small! This should be contrasted with the numerical investigations of section 2.3, where tangencies and homoclinic intersection points were discussed. It would seem then, that even when there are no homoclinic intersection points, the small angle can magnify errors in the noise reduction procedure (by a factor up to the order of $1/\sin^2\theta$).

5.3 Gradient Descent for Non-linear Dynamical Systems

In this section, Theorem 5.7 is generalised to a non-linear dynamical system. This is done by extending the linear result to the linearisation of the system (see section 5.1) and then using the theory of Chapter 4 to recover the non-linear result. Let $f : M \to M$

be a C^2-diffeomorphism defining a discrete dynamical system on a smooth compact d-dimensional manifold M. The results of section 2.3 indicate that for noise reduction to work, the distance between homoclinic intersections for near-tangency points must be bounded away from zero. This uniform bound holds when f has an invariant uniformly hyperbolic set Λ (see section 3.4), and so the existence of such a Λ will be *assumed* in what follows.

5.3.1 Linearised Systems

Recall that the linearisation of the gradient descent flow for f, about a deterministic trajectory y is given by equation 5.2. Proposition 5.1 then asserts that the effect of the linearised gradient descent is to project orthogonally onto the subspace

$$
\mathscr{E}_c = \left\{ \begin{pmatrix} v \\ df(y_1)v \\ df^2(y_1)v \\ \vdots \\ df^{n-1}(y_1)v \end{pmatrix} : v \in \mathbb{R}^d \right\}.
$$

By the chain rule, these deterministic trajectories for the linearised system correspond to trajectories for a *linear* system, where the linear operator *changes* with each iteration. That is, at the first point, the operator is $df(y_1)$, at the second, $df(y_2)$, and so on. With this in mind, as well as the results of the Multiplicative Ergodic Theorem (Theorem 3.5), it is easy to generalise the results of section 5.2.2 to this case. Note that because Λ is hyperbolic, there exist $\mu < 1 < \nu$ such that μ is larger than any Lyapunov number for $f|_\Lambda$ less than unity, and ν is smaller than any Lyapunov number for $f|_\Lambda$ greater than unity. μ and ν will be called *hyperbolicity bounds* for $f|_\Lambda$.

Proposition 5.9 *Suppose that f is a C^2-diffeomorphism of a smooth compact d-dimensional manifold M possessing an invariant uniformly hyperbolic set Λ with splitting into stable and unstable eigenspaces $E_s(p)$ and $E_u(p)$, $p \in \Lambda$, and that y is a deterministic trajectory of length n for f. If $\mathscr{P}$ is the orthogonal projection (in $\mathbb{R}^{nd}$) onto $\mathscr{E}_c(y)$, the subspace of deterministic trajectories for the system linearised about y, and $P_s^{(i)}$ and $P_u^{(i)}$ are the stable and unstable projections onto $E_s(y_i)$ and $E_u(y_i)$ for $i = 1, \ldots, n$ (respectively), then the*

following bounds hold:

$$\left\| P_s^{(i)} \pi_i \mathscr{P} \right\| \leqslant d_s \frac{C_s \mu^{i-1}}{\sin \phi} \left(\frac{\sqrt{1 + (2C_s - 1)\,\mu}}{\sin \phi \sqrt{1 - \mu}} + \frac{\sqrt{1 + (2C_u^{-1} - 1)\,\nu^{-1}}}{\tan \phi \sqrt{1 - \nu^{-1}}} \right)$$

$$\left\| P_u^{(i)} \pi_i \mathscr{P} \right\| \leqslant d_u \frac{C_u^{-1} \nu^{-(n-i)}}{\sin \phi} \left(\frac{\sqrt{1 + (2C_u^{-1} - 1)\,\nu^{-1}}}{\sin \phi \sqrt{1 - \nu^{-1}}} + \frac{\sqrt{1 + (2C_s - 1)\,\mu}}{\tan \phi \sqrt{1 - \mu}} \right)$$

where $\mu < 1 < \nu$ are hyperbolicity bounds for $f|_\Lambda$, C_s and C_u are the associated constants (see Proposition 3.7), d_s and d_u are the common dimensions of the $E_s(p)$ and $E_u(p)$ (respectively), and $\phi \neq 0$ is the minimal angle between $E_s(p)$ and $E_u(p)$, $p \in \Lambda$ (non-zero by the discussion in section 3.3.2).

Proof: This proof is the same as that of Proposition 5.6 with a few modifications. In particular, A^m is replaced by $df^m(y_1)$ throughout. The subspaces $\mathfrak{E}_s$ and $\mathfrak{E}_u$ are then the trajectories in $\mathscr{E}_c(y)$ whose first point belongs to $E_s(y_1)$ and $E_u(y_1)$ respectively. The invariance of the $E_s(p)$ and the $E_u(p)$ given by the Multiplicative Ergodic Theorem (Theorem 3.5) and the fact that y was chosen to be a deterministic trajectory for f, show that $\mathfrak{E}_s$ and $\mathfrak{E}_u$ consist of trajectories whose points stay in stable and unstable eigenspaces (respectively). Hence, $\mathscr{E}_c(y) = \mathfrak{E}_s \oplus \mathfrak{E}_u$. Given an orthogonal basis of $\mathfrak{E}_s^\perp$ say, the construction of stable and unstable trajectories can proceed as in the proof of Proposition 5.6, and these can be used to derive the analogue of equation 5.10. The inequalities of Proposition 3.7 and Lemma 5.5 are then used to simplify this expression, noting that because Λ is compact, the constants C_s and C_u may be chosen independently of the points of the trajectory y (see the discussion after Proposition 3.7), and hence independent of n. The resulting expression still contains angles between stable and unstable trajectories — these are dealt with in exactly the same manner as in the proof of Proposition 5.6, noting that the angles between the $E_s(p)$ and the $E_u(p)$ are uniformly bounded away from zero. ∎

Obviously, analogues of Theorem 5.7 and Corollary 5.8 are also true. These are stated for completeness.

Theorem 5.10 *Let f be a C^2-diffeomorphism of a smooth compact d-dimensional manifold M possessing an invariant uniformly hyperbolic set Λ, $x \in \mathbb{R}^{nd}$ be a noisy trajectory of the linearised system, and $\hat{x}$ be the noise reduced trajectory given by the linearised*

gradient descent algorithm. If the noise distribution is bounded*, then the points of* any *deterministic trajectory (for the linearised system) that could be the true trajectory, differ from the points of* $\hat{x}$ *by an amount which tends to zero as n, the length of the trajectories, tends to infinity, except for points near the initial and final points. The errors at these points remain* bounded *as* $n \to \infty$.

Corollary 5.11 *If f is a C^2-diffeomorphism of a smooth compact d-dimensional manifold M possessing an invariant uniformly hyperbolic set Λ, x a noisy trajectory of the linearised system (derived from the true trajectory y of the linearised system and a bounded noise distribution), and $\hat{x}$ the noise reduced trajectory given by applying the linearised gradient descent algorithm to x, then by taking n, the length of the trajectory, sufficiently large, $\hat{x}_1$ may be made arbitrarily close to the generalised stable eigenspace of y_1 and $\hat{x}_n$ may be made arbitrarily close to the generalised unstable eigenspace of y_n.*

5.3.2 Non-linear Systems

Recall the discussion leading up to Proposition 5.2. There it was shown that for a non-linear system, the set of deterministic trajectories, $\mathscr{W}_c$, forms a centre manifold for *every* fixed point of the gradient descent flow, and there is a lamination of stable manifolds, $\{\mathscr{W}_s(y) : y \in \mathscr{W}_c\}$, orthogonal to this common centre manifold. The situation is exactly the same in the linearised case — here there is a *subspace* of deterministic trajectories which forms a centre eigenspace, and a lamination of stable eigenspaces given by the family of $(n-1)\,d$-dimensional hyperplanes parallel to $\ker \mathscr{P} = \mathscr{E}_c(y)$. These laminations are indicated in Figure 5.2. It would seem plausible then, that the non-linear gradient descent flow and its linearisation about some fixed point are qualitatively similar, that is, *topologically conjugate*, despite the presence of a centre manifold.

 This statement does in fact hold. In most discussions of centre manifold theory ([10, 11]), a statement to the effect that the *stability* of a fixed point with a stable and centre manifold is dictated by the dynamics on the centre manifold is quoted or proven. That is, if the dynamical system (continuous say) can be written in the form

$$
\begin{aligned}
\dot{u} &= Au + g(u, v) \\
\dot{v} &= Bv + h(u, v),
\end{aligned}
$$

86
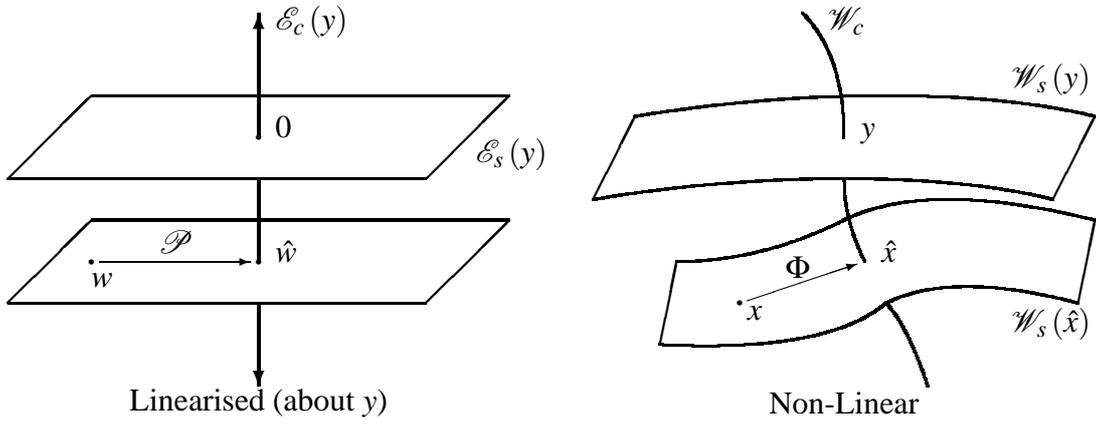
Figure 5.2: Stable manifold and eigenspace laminations in trajectory space $\mathbb{R}^{nd}$

where A has eigenvalues on the imaginary axis, B has eigenvalues with negative real part, $g(0,0) = h(0,0) = 0$ and $dg(0,0) = dh(0,0) = 0$, then the stability of the fixed point 0 is determined by the stability of the fixed point 0 of

$$\dot{u} = Au + g(u, \varphi_c(u))$$

where φ_c is a function from the stable eigenspace into the centre eigenspace whose graph gives a local centre manifold. What this suggests (but is more difficult to show) is that there is a local topological conjugacy between the full system and the system

$$\dot{u} = Au + g(u, \varphi_c(u))$$
$$\dot{v} = Bv.$$

Two proofs of this fact may be found in [31]. When the centre manifold consists entirely of fixed points, the dynamics on the centre manifold is nil. Hence $A = 0$ and $g \circ (\mathrm{id}, \varphi_c) = 0$ so the full system is locally topologically conjugate to its linearisation, verifying the claim of the previous paragraph.

Let the effect of the gradient descent algorithm be given by $\Phi : \mathbb{R}^{nd} \to \mathscr{W}_c$. That is, the gradient descent equations 1.2 give rise to a flow φ^t which converges, given any initial condition, as $t \to \infty$ (by Proposition 5.2). The pointwise limit of φ^t as $t \to \infty$ defines Φ. The idea in what follows is to make the diagram

$$
\begin{array}{ccccc}
\mathbb{R}^{nd} & \xrightarrow{\;\Phi\;} & \mathscr{W}_c & \xrightarrow{\;\pi_i\;} & \mathbb{R}^d \\
{\scriptstyle \mathscr{H}}\downarrow & & \uparrow{\scriptstyle H} & & \uparrow{\scriptstyle h_i} \\
\mathbb{R}^{nd} & \xrightarrow{\;\mathscr{P}\;} & \mathscr{E}_c(y) & \xrightarrow{\;\pi_i\;} & \mathbb{R}^d
\end{array}
\tag{5.13}
$$

87

commute for some functions $\mathcal{H}$, H and h_i, where $\mathcal{P}$ is the orthogonal projection effecting the linearised gradient descent about $\dot{y} \in \mathcal{W}_c$. Knowledge of the function $\mathcal{H}$ and the h_i then allows the study of the non-linear gradient descent to be reduced to the study of the linearised gradient descent. Of course, it is not enough to show that these functions exist. To generalise the results of section 5.3.1 to non-linear systems, it is necessary to demand that the h_i take points near the stable and unstable eigenspaces of the linearised system to points near the stable and unstable *manifolds* of the non-linear system, and that the *distortions* induced by using $\mathcal{H}$ and the h_i to switch between the non-linear and linearised spaces can be *bounded* as the length of the trajectories tend to infinity.

Consider $H : \mathcal{E}_c(y) \to \mathcal{W}_c$. To make the right square of diagram 5.13 commute, it follows that the function H must decompose as $H = (h_1, \ldots, h_n)$. As H maps deterministic trajectories for the linearised system onto deterministic trajectories for the non-linear system, its action may be written as

$$
H = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ \vdots \\ h_n \end{pmatrix} : \begin{pmatrix} u \\ df(y_1)u \\ df^2(y_1)u \\ \vdots \\ df^{n-1}(y_1) \end{pmatrix} \mapsto \begin{pmatrix} v \\ f(v) \\ f^2(v) \\ \vdots \\ f^{n-1}(v) \end{pmatrix}.
$$

Thus H is completely determined by how it takes u into v. If the function on $\mathbb{R}^d$ taking u into v is denoted by g then, it follows that

$$
h_i = f^{i-1} \circ g \circ \left[df^{i-1}(y_1) \right]^{-1} = f^{i-1} \circ g \circ df^{-(i-1)}(y_i).
$$

Note that the action of h_i on a neighbourhood of y_i will be to map the unstable eigenspace for y_i back onto the unstable eigenspace for y_1, distort it a little (the action of g), and then map them forward to a neighbourhood of y_i again. For i large enough then (and provided the action of g isn't too disruptive), the resulting set should be an excellent approximation (at least locally) of the generalised unstable manifold of y_i. In fact, there is a choice for g which makes h_i map the unstable eigenspace locally onto the local generalised unstable eigenspace *exactly*.

Consider the homeomorphisms $\tilde{h}_p$, $p \in \Lambda$, defined between neighbourhoods of p, satisfying equation 4.6:

$$
f \circ \tilde{h}_p = \tilde{h}_{f(p)} \circ df(p)
$$

(these exist by the Generalised Hartman-Grobman Theorem and specifically, Corollary 4.16). If g is defined to be $\tilde{h}_{y_1}$, then by equation 4.6,

$$
\begin{aligned}
h_1 &= g = \tilde{h}_{y_1}, \\
h_2 &= f \circ \tilde{h}_{y_1} \circ [df(y_1)]^{-1} = \tilde{h}_{y_2}, \\
h_3 &= f^2 \circ \tilde{h}_{y_1} \circ [df(y_1)]^{-1} \circ [df(y_2)]^{-1} = f \circ \tilde{h}_{y_2} \circ [df(y_2)]^{-1} = \tilde{h}_{y_3}, \\
&\ \ \vdots \\
h_n &= f^{n-1} \circ \tilde{h}_{y_1} \circ [df(y_1)]^{-1} \circ [df(y_2)]^{-1} \circ \cdots \circ [df(y_{n-1})]^{-1} \\
&= f \circ \tilde{h}_{y_{n-1}} \circ [df(y_{n-1})]^{-1} = \tilde{h}_{y_n}.
\end{aligned}
$$

As each $\tilde{h}_p$ maps the local stable and unstable eigenspaces of p onto the local generalised stable and unstable manifolds of p, the same is true of each h_i (for y_i). Furthermore, since each h_i is a homeomorphism, so is H, and the domain of H can be naturally extended to the product of the domains of the h_i, so H maps a neighbourhood of y homeomorphically onto another neighbourhood of y.

Consider now the left square of diagram 5.13. As Φ is the identity on $\mathscr{W}_c$ and $\mathscr{P}$ is the identity on $\mathscr{E}_c(y)$, it follows that

$$
\begin{aligned}
H|_{\mathscr{E}_c(y)} \circ \mathscr{H}|_{\mathscr{W}_c} &= \ \mathrm{id}|_{\mathscr{W}_c} \\
\Rightarrow \qquad \mathscr{H}|_{\mathscr{W}_c} &= \ H|_{\mathscr{E}_c(y)}^{-1} = H^{-1}|_{\mathscr{W}_c}.
\end{aligned}
$$

It would be very convenient if defining $\mathscr{H}$ to be H^{-1} made diagram 5.13 commute. However, there is no reason to expect this. Instead, note that

$$
\mathscr{P} \circ \mathscr{H} = H^{-1} \circ \Phi \qquad \Rightarrow \qquad \mathscr{H} = (I - \mathscr{P}) \circ \mathscr{H} + H^{-1} \circ \Phi,
$$

and that $(I - \mathscr{P}) \circ \mathscr{H}$ takes values in $\mathscr{E}_s(y)$ whereas $H^{-1} \circ \Phi$ takes values in $\mathscr{E}_c(y)$. In fact, it is clear that the commutativity requirement will still be fulfilled if the $\mathscr{H}$ appearing on the *right* of this equation is replaced by any function mapping $\mathscr{W}_c$ onto $\mathscr{E}_c(y)$. A convenient choice is the homeomorphism H^{-1}, as it is the only function satisfying this requirement whose properties are known. That is, define

$$
\mathscr{H} = (I - \mathscr{P}) \circ H^{-1} + H^{-1} \circ \Phi. \tag{5.14}
$$

$\mathscr{H}$ therefore maps a neighbourhood of y into another neighbourhood of y, and satisfies $\mathscr{P} \circ \mathscr{H} = \mathscr{H} \circ \Phi = H^{-1} \circ \Phi$ (whenever this makes sense). Geometrically, $\mathscr{H}$ takes the

centre manifold $\mathscr{W}_c$ onto the centre eigenspace $\mathscr{E}_c(y)$, *and* maps each stable manifold of the non-linear lamination onto some stable eigenspace of the linearised lamination (see Figure 5.2). The term $H^{-1} \circ \Phi$ specifies *which* stable eigenspace corresponds to a particular stable manifold, and the term $(I - \mathscr{P}) \circ H^{-1}$ specifies *where* on the stable eigenspace each point of the stable manifold is mapped[3].

It remains to consider the distortions induced by $\mathscr{H}$ and the h_i. That is, any stretching or contracting of distances caused by switching between the non-linear and linearised gradient descents. All these functions are *continuous* (Φ is continuous because the stable manifolds in the lamination vary continuously — see the discussion leading up to Proposition 5.2) on their respective domains, so this distortion can be made arbitrarily *small* by restricting their domains to be sufficiently small. However, the generalisation of Theorem 5.10 to non-linear systems must address the behaviour as the length of the relevant trajectories, n, tend to infinity. Therefore it is necessary to know how the distortion varies with n — specifically, it is necessary to show that the errors due to this distortion are *dominated* by the (expected) exponential decrease of the errors (away from the initial and final point of the trajectory) as n tends to infinity.

This can be done for the h_i using the quantitative information of Chapter 4. The h_i were chosen to be the homeomorphisms corresponding to y_i guaranteed by the Generalised Hartman-Grobman Theorem and Corollary 4.16. By Corollary 4.17, the h_i are even *Hölder continuous* on their respective domains, and the associated constants *may be chosen independently of the y_i and hence of n*. That is, there exist $\alpha, \beta > 0$ such that

$$\left\| h_i(u) - h_i(u') \right\| \leqslant \beta \left\| u - u' \right\|^{\alpha}$$

for all u, u' belonging to the domain of h_i, and where α and β may be chosen independently of i (and n). This bounds the distortion of the h_i nicely, and it is clear that the same

---

[3]Obviously, the "where" is unimportant in this application (as the linearised gradient descent projects along the stable eigenspace). However, if it were necessary to show that $\mathscr{H}$ conjugated the gradient descent flow and its linearisation for *finite descent times* rather than just in the limit $t \to \infty$, then this would be important. In fact, one can keep track of where points should be mapped by considering (in addition to the stable laminations) a lamination of (generalised) centre manifolds and eigenspaces (the eigenspaces are clearly the d-dimensional subspaces parallel to $\mathscr{E}_c(y)$). See [31] for details. In fact, the $\mathscr{H}$ constructed above has not been shown to be invertible, so it cannot even qualify as a conjugacy between Φ and $\mathscr{P}$. Because it makes the diagram 5.13 commute however, it is an example of a *semi-conjugacy*.

result holds for the inverses of the h_i (and α and β can be chosen to be Hölder exponents and constants for each h_i^{-1} too). Using this, the diagram 5.13, and Proposition 5.9, it follows that for $y \in \mathscr{W}_c$, $x \in \mathbb{R}^{nd}$ and $\hat{x} = \Phi(x)$,

$$
\begin{aligned}
\|\hat{x}_i - y_i\| &= \|\pi_i \Phi(x) - \pi_i \Phi(y)\| \\
&= \|h_i(\pi_i \mathscr{P} \mathscr{H}(x)) - h_i(\pi_i \mathscr{P} \mathscr{H}(y))\| \\
&\leqslant \beta \|\pi_i \mathscr{P}(\mathscr{H}(x) - \mathscr{H}(y))\|^\alpha \\
&\leqslant \beta \left(K_s \mu^{i-1} + K_u \nu^{-(n-i)}\right)^\alpha \|\mathscr{H}(x) - \mathscr{H}(y)\|_\infty^\alpha, \qquad (5.15)
\end{aligned}
$$

where K_s and K_u are constants (independent of n). It remains then to estimate the effect of (distortion due to) $\mathscr{H}$.

Note that since the h_i are Hölder continuous, it follows that H is Hölder continuous with respect to the norm $\|\cdot\|_\infty$:

$$
\begin{aligned}
\left\|H(x) - H(x')\right\|_\infty &= \sup_{1 \leqslant i \leqslant n} \left\|h_i(x_i) - h_i(x_i')\right\| \\
&\leqslant \sup_{1 \leqslant i \leqslant n} \beta \left\|x_i - x_i'\right\|^\alpha \\
&= \beta \left\|x - x'\right\|_\infty^\alpha,
\end{aligned}
$$

with constants α and β independent of n. As the inverses of the homeomorphisms h_i are also Hölder continuous, H^{-1} is Hölder continuous with respect to $\|\cdot\|_\infty$ too. It might seem plausible now that $\mathscr{H}$ is Hölder continuous as well (with respect to $\|\cdot\|_\infty$). Establishing this seems to be quite difficult however. Obviously the restriction of $\mathscr{H}$ to the centre manifold is Hölder as $\mathscr{H}|_{\mathscr{W}_c} = H^{-1}$, and it is quite easy to show that $\mathscr{H}$ restricted to each stable manifold is Hölder. This follows from the computation ($x, x' \in \mathscr{W}_s(y)$):

$$
\begin{aligned}
\mathscr{H}(x) - \mathscr{H}(x') &= (I - \mathscr{P})\left(H^{-1}(x) - H^{-1}(x')\right) + H^{-1}(\Phi(x)) - H^{-1}(\Phi(x')) \\
&= (I - \mathscr{P})\left(H^{-1}(x) - H^{-1}(x')\right) \\
\Rightarrow \quad \left\|\mathscr{H}(x) - \mathscr{H}(x')\right\|_\infty &\leqslant \|I - \mathscr{P}\|_\infty \beta \left\|x - x'\right\|_\infty^\alpha,
\end{aligned}
$$

where

$$
\|I - \mathscr{P}\|_\infty = \sup_{x \neq 0} \frac{\|(I - \mathscr{P})x\|_\infty}{\|x\|_\infty}.
$$

Note that

$$
\|I - \mathscr{P}\|_\infty \leqslant 1 + \|\mathscr{P}\|_\infty = 1 + \sup_i \|\pi_i \mathscr{P}\| \leqslant 1 + \sup_i \left(\|P_s \pi_i \mathscr{P}\| + \|P_u \pi_i \mathscr{P}\|\right),
$$

91

which is bounded uniformly in n by Proposition 5.9, so the Hölder constant and exponent of $\mathcal{H}|_{\mathcal{W}_s(y)}$ can be chosen independently of n (and y). It does not, however, follow from this that $\mathcal{H}$ itself (unrestricted) is Hölder with constant and exponent independent of n.

The problem seems to be that no quantitative information has been derived for the non-linear gradient descent flow Φ. If say a local Lipschitz condition was derived for Φ (with respect to $\|\cdot\|_\infty$) and the Lipschitz constant could be bounded independent of n, then the Hölder continuity of $\mathcal{H}$ would be established immediately from the definition (equation 5.14). However, getting any quantitative information about Φ appears to be hard. The following condition on f (and on the induced gradient descent Φ) suffices:

Condition 5.12 *Let* $y_1 \in \mathbb{R}^d$ *define deterministic trajectories* $y^{(n)} \in \mathcal{W}_c \subset \mathbb{R}^{nd}$ *(for each n) by* $y_{i+1}^{(n)} = f\left(y_i^{(n)}\right)$, $i = 1, \ldots, n-1$, *and let*

$$\mathscr{B}_\varepsilon\left(y^{(n)}\right) = \left\{x \in \mathbb{R}^{nd} : \left\|x - y^{(n)}\right\|_\infty \leqslant \varepsilon\right\}.$$

Then, for $\varepsilon > 0$ *(denoting the noise level) sufficiently small but fixed, the function*

$$\Omega_\varepsilon(n) = \sup_{x \in \mathscr{B}_\varepsilon\left(y^{(n)}\right)} \left\|\Phi(x) - y^{(n)}\right\|_\infty$$

increases sub-exponentially *with n.*

Since Φ is continuous, $\Omega_\varepsilon(n) \to 0$ as $\varepsilon \to 0$. However, the variance of this quantity with ε is not really important (although it might be nice to know). This is because of the following computation (with $\|x - y\|_\infty \leqslant \varepsilon$):

$$
\begin{aligned}
\|\mathcal{H}(x) - \mathcal{H}(y)\|_\infty &\leqslant \left\|(I - \mathscr{P})\left(H^{-1}(x) - H^{-1}(y)\right)\right\|_\infty + \left\|H^{-1} \circ \Phi(x) - H^{-1} \circ \Phi(y)\right\|_\infty \\
&\leqslant \|I - \mathscr{P}\|_\infty \beta \|x - y\|_\infty^\alpha + \beta \|\Phi(x) - y\|_\infty^\alpha \\
&\leqslant \|I - \mathscr{P}\|_\infty \beta \varepsilon^\alpha + \beta \Omega_\varepsilon(n)^\alpha.
\end{aligned}
$$

If Condition 5.12 is satisfied, then $\|\mathcal{H}(x) - \mathcal{H}(y)\|_\infty$ is bounded (for sufficiently small noise) by a quantity that increases only sub-exponentially with n. Therefore, by equation 5.15,

$$
\begin{aligned}
\|\hat{x}_i - y_i\| &\leqslant \beta \left(K_s \mu^{i-1} + K_u \nu^{-(n-i)}\right)^\alpha \|\mathcal{H}(x) - \mathcal{H}(y)\|_\infty^\alpha \\
&\leqslant \beta^{1+\alpha} \left[\left(K_s \mu^{i-1} + K_u \nu^{-(n-i)}\right)\left(\|I - \mathscr{P}\|_\infty \varepsilon^\alpha + \Omega_\varepsilon(n)^\alpha\right)\right]^\alpha,
\end{aligned}
$$

and as n increases, the exponential decay of the terms μ^{i-1} and $\nu^{-(n-i)}$ for $i \sim n/2$ dominates the sub-exponential increase of $\Omega_\varepsilon(n)$. That is, near the middle of the trajectory (hence away from the initial and final points), the errors converge to zero as the length of the trajectories, n, tends to infinity. Thus, when Condition 5.12 is satisfied, the generalisation of Theorem 5.10 to non-linear dynamical systems (with an invariant uniformly hyperbolic set) is proven.

Theorem 5.13 *Let f be a C^2-diffeomorphism of a smooth compact d-dimensional manifold M possessing an invariant uniformly hyperbolic set Λ and satisfying Condition 5.12, $x \in \mathbb{R}^{nd}$ be a noisy trajectory of the (non-linear) system, and $\hat{x}$ be the noise reduced trajectory given by the gradient descent algorithm. If the noise distribution is* bounded *by $\varepsilon > 0$ sufficiently small, then the points of* any *deterministic trajectory that could be the true trajectory, differ from the points of $\hat{x}$ by an amount which tends to zero as n, the length of the trajectories, tends to infinity, except for points near the initial and final points. The errors at these points are bounded if $\Omega_\varepsilon(n)$ is bounded (in n) and otherwise may increase subexponentially as $n \to \infty$.*

Additionally, if Condition 5.12 holds, then the generalisation of Corollary 5.11 to non-linear systems with an invariant uniformly hyperbolic set is also easy to prove. Assuming the hypotheses and notation of Theorem 5.13 (and Proposition 5.9), it is easy to see that the stable error after linearisation, $\left\| P_s^{(i)} \pi_i \mathscr{P}(\mathscr{H}(x) - \mathscr{H}(y)) \right\|$, will be negligible for all i sufficiently large (when n is sufficiently large). In particular, there will be an $i = n - m$ say, for which this stable error is negligible, *and* for which the corresponding unstable error after linearisation, $\left\| P_u^{(i)} \pi_i \mathscr{P}(\mathscr{H}(x) - \mathscr{H}(y)) \right\|$, is small enough that $\pi_{n-m} \mathscr{P}(\mathscr{H}(x) - \mathscr{H}(y))$ belongs to the neighbourhood of y_{n-m} where h_{n-m} conjugates the linear and non-linear systems. As the stable error is negligible, and h_{n-m} takes the local generalised unstable eigenspace onto the local generalised unstable manifold *quantitatively*, it follows that $\hat{x}_{n-m} = \pi_{n-m} \Phi(x) = h_{n-m}(\pi_{n-m} \mathscr{P}\mathscr{H}(x))$ will be extremely close to the *local* generalised unstable manifold of y_{n-m}. But, by Proposition 5.2, $\hat{x}$ is a deterministic trajectory for f, as is y, so $\hat{x}_{n-m}$ close to the generalised unstable manifold of y_{n-m} implies that $\hat{x}_n$ is even closer to the generalised unstable manifold of y_n (although the distance *along* the unstable manifold may be very large). The same argument gives the corresponding result for $\hat{x}_1$ and the generalised stable manifold of y_1.

Corollary 5.14 *If f is a C^2-diffeomorphism of a smooth compact d-dimensional manifold M possessing an invariant uniformly hyperbolic set Λ, and which satisfies Condition 5.12, x a noisy trajectory of the system (derived from the true trajectory y and a bounded noise distribution with sufficiently small bound), and $\hat{x}$ the noise reduced trajectory given by applying the gradient descent algorithm to x, then by taking n, the length of the trajectory, sufficiently large, $\hat{x}_1$ may be made arbitrarily close to the generalised stable eigenspace of y_1 and $\hat{x}_n$ may be made arbitrarily close to the generalised unstable eigenspace of y_n.*

Theorem 5.13 essentially states that the gradient descent algorithm is a good noise reduction algorithm for non-linear dynamical systems with an invariant uniformly hyperbolic set (that is, without arbitrarily bad tangencies), provided the noise level is sufficiently small. Corollary 5.14 then states that (as one might expect) the noise reduced trajectory begins on (or very near) the generalised stable manifold of the initial point of the true trajectory, and ends on (or very near) the generalised unstable manifold of the final point of the true trajectory. *Both results, however, rely on the Condition 5.12 being satisfied.* When does this condition hold? Perhaps a better question to ask would be: How could this condition possibly fail to hold? For a consequence of failure would be that the errors at the initial and final points could grow *exponentially* or even faster as the length of the trajectory increases. This is certainly at odds with the numerical experiments of Chapter 2, although these experiments are of course, not even remotely exhaustive. In fact, this limited set of experiments suggest that the errors are *bounded* as the length of trajectory tends to infinity, and this was proven for linear and linearised systems in Theorems 5.7 and 5.10. Furthermore, by Bowen's Shadowing Theorem ([8]), for sufficiently small noise, there is a *unique* deterministic trajectory that could produce any given noisy trajectory *of infinite length*. One would hope that a respectable noise reduction algorithm would converge (pointwise, not uniformly) onto this unique trajectory as the length of trajectory tends to infinity.

It seems reasonable therefore to *conjecture* that for any f possessing an invariant uniformly hyperbolic set (from which the time series was generated), Condition 5.12 is satisfied. Another reasonable conjecture to make is that Condition 5.12 is not only satisfied by such f, but $\Omega_\varepsilon(n)$ is in fact, bounded in n (for each ε small enough). The author believes strongly that the first conjecture holds. However, it is possible that the second

(strengthened) conjecture is not true in general. If this is the case, then it may happen that after gradient descent with small bounded noise, the initial and final errors might be quite large. This is obviously less than optimal for the state estimation problem (see sections 1.1 and 1.3). An alternate gradient descent algorithm ([28]) might therefore be of use in this case. The idea here is to try to limit how far the gradient descent can "move" each point of the trajectory being noise reduced, by replacing the determinism function L (defined in equation 1.1) by

$$\tilde{L}\left(x; x'\right) = L\left(x\right) + \eta \sum_{i=1}^{n} \left\| x_i - x_i' \right\|^2,$$

where x' is the original noisy trajectory and $\eta \geqslant 0$ is an arbitrary weighting. For $\eta \neq 0$, there is no reason why the trajectory obtained after gradient descent with $\tilde{L}$ should be deterministic, so to get a (nearly) deterministic trajectory, the algorithm is recursively applied to the point reached after gradient descent, using smaller values of η, until L is satisfactorily small[4]. It is expected that (with an appropriate choice of weightings η) this algorithm should give a deterministic trajectory that is "close" to the original noisy one. In a sense then, the question of whether $\Omega_\varepsilon\left(n\right)$ is in fact bounded in n or merely of subexponential growth, is equivalent to there being a theoretical justification for considering $\tilde{L}$ rather than L.

How then, might one establish the truth or falsity of the conjectures of the previous paragraph? The corresponding results for the linear and linearised systems was proven by deriving the analytic bounds of Propositions 5.6 and 5.9. This derivation was made possible because the linearised gradient descent is equivalent to the action of an orthogonal projection (by Proposition 5.1). This yields enough quantitative information to show that the errors are uniformly bounded. It seems likely therefore, that to get such a bound for the non-linear gradient descent, one must consider the stable manifolds (rather than the stable eigenspaces) with respect to which Φ behaves somewhat like a projection. It may be true that for small enough noise, the local stable manifolds are uniformly Lipschitz (as graphs of functions) with Lipschitz constants independent of n, but it is not clear, as

---

[4]Note that the extra term added to $\tilde{L}$ removes the *degeneracy* of the fixed point set. That is, the set of fixed points of the gradient descent with $\tilde{L}$ is zero-dimensional, and so they are isolated (by the compactness of M). Hence, the critical points of $\tilde{L}$ could be found using an algebraic solver, rather than a differential solver. With this in mind, this algorithm should be compared to the algorithm of Farmer and Sidorowich discussed in section 1.2

yet, how to relate this idea to the curvature of the centre manifold. The author would like to apologise for not pursuing this result further, pleading time constraints that force the matter to remain in question at this point.

Chapter 6

Conclusions

In this thesis, an algorithm for performing noise reduction was introduced, discussed, numerically tested, and an attempt was made to prove rigorously that the algorithm converges onto the clean data except near the initial and final points, in the limit that the number of data points used tends to infinity (and under additional mild assumptions). The algorithm studied was the gradient descent algorithm. Specifically, it was shown that the algorithm always reduces the noise level to zero, and that provided the data comes from an invariant uniformly hyperbolic set for a discrete dynamical system, the noise comes from a bounded distribution with sufficiently small bound, and Condition 5.12 is satisfied, then the possible noise-free data sets which could have given rise to the noisy data (assuming observational noise) only differ significantly near the initial and final points of the data set. It was argued that Condition 5.12 is in fact superfluous for uniformly hyperbolic systems, and an idea as to how this might be shown, mentioned.

The thesis begins by introducing noise reduction in general terms and its relation to the problems of modelling and state estimation. The gradient descent algorithm is introduced and compared with a couple of other simple "dynamical" noise reduction algorithms. The motivation behind investigating the theoretical properties of the gradient descent algorithm comes from the theory of indistinguishable states (in state estimation), so this was briefly introduced. This was followed by some numerical experiments consisting of applying the gradient descent algorithm to some simple non-linear maps (mostly the maps of Hénon and Ikeda). These experiments gave an indication of what sort of results should be expected for gradient descent, and also highlighted the importance of hyperbolicity and tangencies. In particular, it was shown that tangencies prevent the gradient descent

algorithm from achieving noise reduction, essentially because a tangency usually implies the presence of a (non-trivial) nearby homoclinic intersection point of generalised stable and unstable manifolds. It was further argued, on the basis of indistinguishability, that such tangencies will prevent noise reduction for any given algorithm, and that therefore, to show that noise reduction is always achieved, it is necessary to restrict attention to systems which do not exhibit tangencies.

The numerical experiments with tangencies also showed that in order to avoid homoclinic intersection points, it is necessary to consider small noise levels only. Specifically, the noise should be smaller than the distances between homoclinic intersection points. It was therefore argued that because the allowed noise levels correspond to neighbourhoods where one would expect the non-linear dynamics to be qualitatively similar to the linearised dynamics, it should be possible to prove the result about noise reduction for a non-linear system, by investigating what happens for the linearised system. However, the systems of interest do not generally consist of dynamics around a fixed point, for which the linearisation theory is well known (in the form of the Hartman-Grobman Theorem). Therefore, a more general linearisation theory had to be introduced. This took the form of a generalisation of the Hartman-Grobman Theorem to invariant uniformly hyperbolic sets.

The investigation of the gradient descent algorithm was then initiated. The algorithm itself amounts to solving a differential equation (with respect to time say) and computing the limit, as time tends to infinity, of the solution. It was shown that the linearised gradient descent algorithm (formed by linearising the differential equation about any fixed point) was equivalent to projecting orthogonally onto a subspace corresponding to trajectories of the linearised dynamical system. Using this result and general results from stable manifold theory, it was then shown that the non-linear gradient descent algorithm has to converge onto a noise-free trajectory. The gradient descent algorithm was then investigated for a hyperbolic linear dynamical system, and using the equivalence with an orthogonal projection, analytic bounds were derived at each point for the errors between the trajectory given by the algorithm, and the true trajectory. These were shown to imply that for a linear dynamical system, the errors converge to zero everywhere except near the initial and final points of the trajectory, as the length of the trajectory tends to infinity.

After proving this result, it was a simple matter to pass to the corresponding result for

the linearisation of a hyperbolic dynamical system (using the Multiplicative Ergodic Theorem). To generalise this to non-linear hyperbolic systems proved more difficult however. Mimicking the idea behind the proof of the Hartman-Grobman Theorem, a diagram was constructed around the non-linear and linearised gradient descent algorithms and functions sought to make this diagram commute. It was shown that there are many ways to do this, but that there were choices for the functions that were particularly convenient. In particular, corresponding to each point of the trajectory, it was shown to be convenient to choose the homeomorphism (guaranteed by the generalisation of the Hartman-Grobman Theorem) effecting the qualitative correspondence between the non-linear and linearised dynamics around that point. The homeomorphisms were actually shown earlier to provide a quantitative correspondence (in the form of Hölder continuity), and so the analytic bounds derived for the linearised system can be extended to deal with these homeomorphisms. It remained to choose a function which transformed the non-linear flow into a linearised flow. The functions that do this were shown to have an explicit form, and a convenient representative was chosen. Under the assumption that the non-linear gradient descent satisfies Condition 5.12, the commutative diagram was then shown to extend the analytic bounds derived for the linearised system to the full non-linear system.

Bibliography

[1] R Abraham and J Robbin. *Transversal Mappings and Flows*. W A Benjamin Inc, New York, 1967.

[2] K T Alligood, T D Sauer, and J A Yorke. *Chaos. An Introduction to Dynamical Systems*. Springer, New York, 1996.

[3] B D O Anderson and J B Moore. *Linear Optimal Control*. Prentice-Hall, New Jersey, 1971.

[4] B D O Anderson and J B Moore. *Optimal Filtering*. Prentice-Hall, New Jersey, 1979.

[5] D V Anosov. *Dynamical Systems IX*, volume 66 of *Encyclopaedia of Mathematical Sciences*. Springer, Berlin, 1995.

[6] M Benedicks and L Carleson. The Dynamics of the Hénon Map. *Annals of Mathematics*, 133:73–169, 1991.

[7] R Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1975.

[8] R Bowen. *On Axiom A Diffeomorphisms*, volume 35 of *Regional Conference Series in Mathematics*. American Mathematical Society, Providence, 1978.

[9] Th Bröcker and K Jänich. *Introduction to Differential Topology*. Cambridge University Press, Cambridge, 1982.

[10] J Carr. *Applications of Centre Manifold Theory*, volume 35 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1981.

[11] S N Chow and J K Hale. *Methods of Bifurcation Theory*, volume 251 of *A Series of Comprehensive Studies in Mathematics*. Springer-Verlag, New York, 1982.

[12] M Davies. Noise Reduction by Gradient Descent. *International Journal of Bifurcation and Chaos*, 3(1):113–118, 1992.

[13] M Davies. Noise Reduction Schemes for Chaotic Time Series. *Physica D*, 79:174–192, 1994.

[14] J-P Eckmann and D Ruelle. Ergodic Theory of Chaos and Strange Attractors. *Reviews of Modern Physics*, 57(3):617–656, July 1985.

[15] J D Farmer and J J Sidorowich. Optimal Shadowing and Noise Reduction. *Physica D*, 47(3):373–392, 1991.

[16] P Grassberger and T Schreiber. A Simple Noise-Reduction Method for Real Data. *Physics Letters A*, 160:411, 1991.

[17] C Grebogi, S M Hammel, T Sauer, and J A Yorke. Shadowing of Physical Trajectories in Chaotic Dynamics. *Physical Review Letters*, 65:1527, 1990.

[18] J Guckenheimer and P Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[19] V Guillemin and A Pollack. *Differential Topology*. Prentice-Hall, New Jersey, 1974.

[20] P R Halmos. *Lectures on Ergodic Theory*. Mathematical Society of Japan, Tokyo, 1956.

[21] S M Hammel. A Noise Reduction Method for Chaotic Systems. *Physics Letters A*, 148(8,9):421–428, September 1990.

[22] M Hénon. A Two-dimensional Mapping with a Strange Attractor. *Communications in Mathematical Physics*, 50:69–77, 1976.

[23] M W Hirsch, C C Pugh, and M Shub. *Invariant Manifolds*, volume 583 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1977.

[24] M W Hirsch and S Smale. *Differential Equations, Dynamical Systems and Linear Algebra*, volume 60 of *Pure and Applied Mathematics*. Academic Press, New York, 1974.

[25] K Ikeda. Multiple-valued Stationary State and its Instability of the Transmitted Light by a Ring Cavity System. *Optical Communications*, 30:257–, 1979.

[26] M C Irwin. Review for [34]. see `http://www.ams.org/mathscinet/`.

[27] M C Irwin. *Smooth Dynamical Systems*. Academic Press, London, 1980.

[28] K Judd and L Smith. Indistinguishable States I: Perfect Model Scenario. to appear in Physica D.

[29] H Kantz and T Schreiber. *Nonlinear Time Series Analysis*, volume 7 of *Cambridge Nonlinear Science Series*. Cambridge University Press, Cambridge, 1997.

[30] A Katok and B Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*, volume 54 of *Encyclopaedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1998.

[31] U Kirchgraber and K J Palmer. *Geometry in the Neighbourhood of Invariant Manifolds of Maps and Flows and Linearisation*, volume 233 of *Pitman Research Notes in Mathematics*. Longman Scientific and Technical, Essex, 1990.

[32] E J Kostelich and J A Yorke. Noise Reduction in Dynamical Systems. *Physical Review A*, 38(3):1649–1652, August 1988.

[33] E Kreyszig. *Introductory Functional Analysis with Applications*. Wiley, New York, 1978.

[34] M Kurata. Hartman's Theorem for Hyperbolic Sets. *Nagoya Mathematical Journal*, 67:41–52, 1977.

[35] S P Lalley. Beneath the Noise, Chaos. to appear in Annals of Statistics `http://www.stat.purdue.edu/~lalley/Papers/index.html`.

[36] R Mane. *Ergodic Theory and Differentiable Dynamics*. Springer-Verlag, Berlin, 1987.

[37] E Nelson. *Topics in Dynamics. I: Flows*. Princeton University Press, Princeton, 1969.

[38] Z Nitecki. *Differentiable Dynamics*. The MIT Press, Cambridge, Massachusetts, 1971.

[39] V I Oseledec. A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems. *Transactions of the Moscow Mathematical Society*, 19:197–221, 1968.

[40] J Palis Jr and W de Melo. *Geometric Theory of Dynamical Systems, An Introduction*. Springer-Verlag, New York, 1982.

[41] L Perko. *Differential Equations and Dynamical Systems*, volume 7 of *Texts in Applied Mathematics*. Springer, New York, 1996.

[42] M Pollicott. *Lectures on Ergodic Theory and Pesin Theory on Compact Manifolds*, volume 180 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1993.

[43] C C Pugh. On a Theorem of P Hartman. *American Journal of Mathematics*, 91:363–367, 1969.

[44] C C Pugh and M Shub. Linearisation of Normally Hyperbolic Diffeomorphisms and Flows. *Inventiones mathematicae*, 10:187–198, 1970.

[45] F Riesz and B Sz-Nagy. *Functional Analysis*. Dover, New York, 1955.

[46] C Robinson. *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. CRC Press, Boca Raton, 1995.

[47] D Ruelle. Ergodic Theory of Differentiable Dynamical Systems. *Publications Mathematiques. Institut des Hautes Etudes scientifiques*, 50:27–58, 1979.

[48] M Shub. *Global Stability of Dynamical Systems*. Springer-Verlag, New York, 1987.

[49] S Smale. Differentiable Dynamical Systems. In *The Mathematics of Time: Essays on Dynamical Systems, Economic Processes, and Related Topics*, pages 1–82, New York, 1980. Springer-Verlag.

[50] M Spivak. *A Comprehensive Introduction to Differential Geometry*, volume 1. Publish or Perish, Inc., Boston, 1970.

[51] F Takens. Detecting Strange Attractors in Turbulence. In D A Rand and L S Young, editors, *Dynamical Systems and Turbulence*, volume 898 of *Lecture Notes in Mathematics*, pages 365–381. Springer-Verlag, Berlin, 1981.

[52] M Viana and M Benedicks. Solution of the Basin Problem for Henon-like Attractors. `www.impa.br/~viana`, to appear in Inventiones Mathematicae.

[53] P Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, New York, 1982.